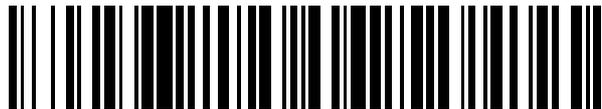


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 656 464**

51 Int. Cl.:

G06F 11/07 (2006.01)

G06F 11/30 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

- 86 Fecha de presentación y número de la solicitud internacional: **11.09.2013 PCT/CN2013/083325**
- 87 Fecha y número de publicación internacional: **19.03.2015 WO15035574**
- 96 Fecha de presentación y número de la solicitud europea: **11.09.2013 E 13882632 (6)**
- 97 Fecha y número de publicación de la concesión europea: **06.12.2017 EP 2869201**

54 Título: **Procedimiento, sistema informático y aparato de procesamiento de fallo**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
27.02.2018

73 Titular/es:
**HUAWEI TECHNOLOGIES CO., LTD. (100.0%)
Huawei Administration Building, Bantian
Longgang District
Shenzhen, Guangdong 518129, CN**

72 Inventor/es:

**LIN, MUHUI;
WANG, JUNJIE y
WANG, RUILING**

74 Agente/Representante:

LEHMANN NOVO, María Isabel

ES 2 656 464 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

Procedimiento, sistema informático y aparato de procesamiento de fallo

SECTOR TÉCNICO

5 La presente invención se refiere a tecnologías informáticas, y en particular, a un procedimiento, un sistema informático y un aparato de aislamiento de fallo.

ANTECEDENTES

10 La tecnología de buses de interconexión de componentes periféricos exprés (Peripheral Component Interconnect Express, PCIe) es una tecnología de buses de alto rendimiento utilizada para interconectar una CPU con un dispositivo periférico. La PCIe, como estándar de interfaz y bus de nueva generación, lleva a cabo transmisión de datos a modo de interconexión en serie y de tipo punto a punto, lo que aumenta enormemente la velocidad de transmisión y crea asimismo las condiciones para aumentar adicionalmente la frecuencia. La PCIe se aplica ampliamente a servidores industriales, dispositivos PC, computación/comunicación empotrada, estaciones de trabajo y similares, y sustituye gradualmente los buses tales como PCI y AGP. Actualmente, los fallos de un dispositivo PCIe constituyen la mayor parte de todos los fallos de un sistema. Monitorizar el sistema en tiempo real, identificar la aparición de un error, y detectar y procesar los fallos de sistema puede evitar de manera efectiva la interrupción total del funcionamiento del sistema, y es una característica RAS (fiabilidad, disponibilidad y mantenimiento (Reliability, Availability, and Serviceability)) que garantiza una disponibilidad continua del sistema.

15 En la técnica anterior, cuando un dispositivo PCIe está en fallo, se puede generar un paquete de error, y el paquete de error es encaminado desde el dispositivo en fallo hasta un complejo raíz; después de obtener el paquete de error, el complejo raíz genera una interrupción de sistema y notifica el paquete de error a un sistema operativo; el sistema operativo lleva a cabo un tratamiento de errores en función del paquete de error. En la técnica anterior, existe una ventana de tiempo desde el momento en que el dispositivo en fallo genera un paquete de error hasta el momento en que el sistema operativo procesa el paquete de error; dentro de la ventana de tiempo, una CPU u otro dispositivo de punto extremo PCIe y el dispositivo en fallo pueden seguir accediendo uno al otro, de tal modo que el dispositivo en fallo no puede ser aislado de manera efectiva, lo que puede provocar la propagación del fallo y afectar a la fiabilidad del sistema.

20 El documento US 2011/320860 A1 da a conocer la habilitación, detección, notificación y/o procesamiento de eventos, tales como errores asociados con adaptadores. Un hardware detecta un evento, pone uno o varios adaptadores en estado de error para impedir el acceso hacia/desde los adaptadores y notifica el evento al sistema operativo.

RESUMEN

30 La presente invención da a conocer un procedimiento de aislamiento de fallo según la reivindicación 1, un sistema según la reivindicación 13 y un aparato según la reivindicación 10, que pueden aislar un dispositivo en fallo, impedir la propagación de un fallo y mejorar la fiabilidad del sistema.

35 De acuerdo con un primer aspecto, una realización de la presente invención da a conocer un procedimiento de aislamiento de fallo, utilizado en un sistema informático interconectado con PCIe, donde el sistema informático incluye un dominio principal y un dominio extendido, el dominio principal incluye un complejo raíz, un primer dispositivo de punto extremo y un dispositivo de punto extremo del complejo raíz, y el dominio extendido incluye el dispositivo de punto extremo del complejo raíz y un segundo dispositivo de punto extremo; y el procedimiento incluye:

40 monitorizar el estado del segundo dispositivo de punto extremo en el dominio extendido;

establecer un registro del estado del dispositivo de acuerdo con el estado del segundo dispositivo de punto extremo, donde el registro del estado del dispositivo incluye una correspondencia entre una información de identificador del segundo dispositivo de punto extremo y el estado del segundo dispositivo de punto extremo;

45 recibir una solicitud de acceso, donde la solicitud de acceso incluye una solicitud de acceso desde el segundo dispositivo de punto extremo al dominio principal o una solicitud de acceso desde el dominio principal al segundo dispositivo de punto extremo;

50 consultar el registro del estado del dispositivo de acuerdo con la información de identificador que es del segundo dispositivo de punto extremo y está en la solicitud de acceso, y determinar el estado del segundo dispositivo de punto extremo; y

desechar la solicitud de acceso para impedir la comunicación entre el segundo dispositivo de punto extremo del dominio extendido y el dominio principal cuando el estado del segundo dispositivo de punto extremo es un estado de fallo;

donde el procedimiento comprende además:

- 5 configurar una primera dirección de memoria y una segunda dirección de memoria para el segundo dispositivo de punto extremo, la primera dirección de memoria se utiliza para representar el segundo dispositivo de punto extremo en el dominio principal, la segunda dirección de memoria se utiliza para representar el segundo dispositivo de punto extremo en el dominio extendido;
- configurar un identificador de bus/dispositivo/función BDF para el segundo dispositivo de punto extremo;
- almacenar una relación de mapeo entre el identificador BDF del segundo dispositivo de punto extremo y la segunda dirección de memoria del segundo dispositivo de punto extremo; y
- 10 almacenar una relación de mapeo entre la primera dirección de memoria del segundo dispositivo de punto extremo y la segunda dirección de memoria del segundo dispositivo de punto extremo.
- Haciendo referencia al primer aspecto, en un primer posible modo de implementación,
- la monitorización del estado del segundo dispositivo de punto extremo en el dominio extendido incluye: recibir un mensaje de error enviado por el segundo dispositivo de punto extremo, o recibir un mensaje de respuesta de detección del dispositivo, utilizado para indicar si el segundo dispositivo de punto extremo existe; determinar el estado del segundo dispositivo de punto extremo de acuerdo con el mensaje de error o el mensaje de respuesta de detección del dispositivo; obtener el identificador BDF del segundo dispositivo de punto extremo y transportado en el mensaje de error o el mensaje de respuesta de detección del dispositivo; y registrar, en el registro del estado del dispositivo, una correspondencia entre el identificador BDF del segundo dispositivo de punto extremo y el estado del segundo dispositivo de punto extremo.
- 15
- 20 Haciendo referencia al primer posible modo de implementación del primer aspecto, en un segundo posible modo de implementación, obtener la primera dirección de memoria del segundo dispositivo de punto extremo de acuerdo con el identificador BDF, y registrar, en el registro del estado del dispositivo, una correspondencia entre la primera dirección de memoria del segundo dispositivo de punto extremo y el estado del segundo dispositivo de punto extremo
- 25 Haciendo referencia al primer modo de implementación posible del primer aspecto, en un segundo posible modo de implementación, la obtención de la primera dirección de memoria del segundo dispositivo de punto extremo de acuerdo con el identificador BDF o la segunda dirección de memoria incluye: obtener la segunda dirección de memoria del segundo dispositivo de punto extremo de acuerdo con una relación de mapeo almacenada entre el identificador BDF del segundo dispositivo de punto extremo y la segunda dirección de memoria del segundo dispositivo de punto extremo, y transformar la segunda dirección de memoria en la primera dirección de memoria del segundo dispositivo de punto extremo de acuerdo con una relación de mapeo almacenada entre la segunda dirección de memoria del segundo dispositivo de punto extremo y la primera dirección de memoria del segundo dispositivo de punto extremo.
- 30
- 35 Haciendo referencia al segundo posible modo de implementación del primer aspecto, en un tercer posible modo de implementación, la primera dirección de memoria del segundo dispositivo de punto extremo incluye una primera dirección de memoria de acceso al espacio de configuración, una primera dirección de memoria de acceso a interrupciones señalizadas por mensaje, una primera dirección de memoria de acceso de entrada/salida con mapeo a memoria, y una primera dirección de memoria de acceso directo a memoria; y el registro, en el registro del estado del dispositivo, de una correspondencia entre la primera dirección de memoria del segundo dispositivo de punto extremo y el estado del segundo dispositivo de punto extremo incluye: registrar, en el registro del estado del dispositivo, la correspondencia entre cada primera dirección de memoria del segundo dispositivo de punto extremo y el estado del segundo dispositivo de punto extremo.
- 40
- 45 De acuerdo con un segundo aspecto, una realización de la presente invención da a conocer un aparato de aislamiento de fallo, utilizado en un sistema informático interconectado con PCIe, donde el sistema informático incluye un dominio principal y un dominio extendido, el dominio principal incluye un complejo raíz, un primer dispositivo de punto extremo y un dispositivo de punto extremo del complejo raíz, y el dominio extendido incluye el dispositivo de punto extremo del complejo raíz y un segundo dispositivo de punto extremo; y el aparato incluye:
- una unidad de monitorización, configurada para monitorizar un estado del segundo dispositivo de punto extremo en el dominio extendido;
- 50 una unidad de registro, configurada para establecer un registro del estado del dispositivo de acuerdo con el estado del segundo dispositivo de punto extremo, donde el registro del estado del dispositivo incluye una correspondencia entre una información de identificador del segundo dispositivo de punto extremo y el estado del segundo dispositivo de punto extremo;
- 55 una unidad de recepción, configurada para recibir una solicitud de acceso, donde la solicitud de acceso incluye una solicitud de acceso del segundo dispositivo de punto extremo al dominio principal o una solicitud de acceso del dominio principal al segundo dispositivo de punto extremo;

una unidad de determinación, configurada para consultar el registro del estado del dispositivo de acuerdo con la información de identificador que es del segundo dispositivo de punto extremo y está en la solicitud de acceso, y determinar el estado del segundo dispositivo de punto extremo; y

5 una unidad de procesamiento, configurada para: cuando el estado del segundo dispositivo de punto extremo es un estado de fallo, desechar la solicitud de acceso para impedir la comunicación entre el segundo dispositivo de punto extremo y el dominio principal;

10 donde la unidad de determinación está configurada para configurar una primera dirección de memoria y una segunda dirección de memoria para el segundo dispositivo de punto extremo, la primera dirección de memoria se utiliza para representar el segundo dispositivo de punto extremo en el dominio principal, la segunda dirección de memoria se utiliza para representar el segundo dispositivo de punto extremo en el dominio extendido, configurar un identificador BDF de bus/dispositivo/función para el segundo dispositivo de punto extremo, almacenar una relación de mapeo entre el identificador BDF del segundo dispositivo de punto extremo y la segunda dirección de memoria del segundo dispositivo de punto extremo, y almacenar una relación de mapeo entre la primera dirección de memoria del segundo dispositivo de punto extremo y la segunda dirección de memoria del segundo dispositivo de punto extremo.

15 Haciendo referencia al segundo aspecto, en un primer posible modo de implementación, la unidad de monitorización está configurada específicamente para: recibir un mensaje de error enviado por el segundo dispositivo de punto extremo, o recibir un mensaje de respuesta de detección del dispositivo, utilizado para indicar si el segundo dispositivo de punto extremo existe; y determinar el estado del segundo dispositivo de punto extremo de acuerdo con el mensaje de error o el mensaje de respuesta de detección del dispositivo;

20 la unidad de registro está configurada específicamente para: obtener el identificador BDF del segundo dispositivo de punto extremo y transportado en el mensaje de error o el mensaje de respuesta de detección del dispositivo, y registrar, en el registro del estado del dispositivo, una correspondencia entre el identificador BDF del segundo dispositivo de punto extremo y el estado del segundo dispositivo de punto extremo.

25 Haciendo referencia al primer posible modo de implementación del segundo aspecto, en un segundo posible modo de implementación, la unidad de registro está configurada específicamente para: obtener la primera dirección de memoria del segundo dispositivo de punto extremo de acuerdo con el identificador BDF; y

registrar, en el registro del estado del dispositivo, una correspondencia entre la primera dirección de memoria del segundo dispositivo de punto extremo y el estado del segundo dispositivo de punto extremo.

30 En las realizaciones de la presente invención, se puede monitorizar el estado de un segundo dispositivo de punto extremo en un dominio extendido, y se ajusta un registro del estado del dispositivo de acuerdo con el estado del segundo dispositivo de punto extremo; después de la recepción de una solicitud de acceso entre el segundo dispositivo de punto extremo y un dominio principal, se consulta el registro del estado del dispositivo de acuerdo con la información de identificador que es del segundo dispositivo de punto extremo y está en la solicitud de acceso, y se determina el estado del segundo dispositivo de punto extremo; si el estado del segundo dispositivo de punto extremo está en un estado de fallo, la solicitud de acceso se desecha para impedir la comunicación entre el segundo dispositivo de punto extremo en fallo y el dominio principal e impedir la propagación de un fallo al dominio principal, asegurando de ese modo la fiabilidad del sistema.

BREVE DESCRIPCIÓN DE LOS DIBUJOS

40 Para describir más claramente las soluciones técnicas en las realizaciones de la presente invención, a continuación se introducen brevemente los dibujos adjuntos necesarios para describir la técnica anterior o las realizaciones. Evidentemente, los dibujos adjuntos en la siguiente descripción muestran solamente algunas realizaciones de la presente invención, y un experto en la materia obtendrá sin esfuerzos creativos otros dibujos a partir de estos dibujos adjuntos.

45 La figura 1 es un diagrama de bloques de sistema, de un sistema informático que incluye un dominio principal PCIe y un dominio extendido PCIe, de acuerdo con una realización de la presente invención;

la figura 2 (a) es un diagrama de asignación de direcciones de memoria de dispositivos de punto extremo en el dominio principal y el dominio extendido en el sistema informático mostrado en la figura 1;

50 la figura 2 (b) es una relación de mapeo entre direcciones de memoria en el dominio principal en el sistema informático mostrado en la figura 1 y direcciones de memoria en el dominio extendido en el sistema informático mostrado en la figura 1;

la figura 3 es un diagrama de flujo de un procedimiento de aislamiento de fallo, según la realización 1 de la presente invención;

55 la figura 4 es un diagrama de flujo de un procedimiento de aislamiento de fallo, según la realización 2 de la presente invención;

la figura 5 es un diagrama de flujo de un procedimiento de aislamiento de fallo, según la realización 3 de la presente invención;

la figura 6 es un diagrama de flujo de un procedimiento de aislamiento de fallo, según la realización 4 de la presente invención;

5 la figura 7 es un diagrama estructural de un aparato de aislamiento de fallo, según una realización de la presente invención;

la figura 8 es un diagrama estructural de un aparato de aislamiento de fallo, según una realización de la presente invención;

10 la figura 9 es un diagrama estructural de un sistema de aislamiento de fallo, según una realización de la presente invención; y

la figura 10 es un diagrama estructural de un aparato de aislamiento de fallo, según una realización de la presente invención.

DESCRIPCIÓN DE REALIZACIONES

15 Las realizaciones de la presente invención dan a conocer un procedimiento, un sistema informático y un aparato de aislamiento de fallo, que se aplican a un sistema informático que incluye un dominio principal PCIe y un dominio extendido PCIe. Un dispositivo de punto extremo del complejo raíz en el dominio extendido es un dispositivo de punto extremo en el dominio principal. Cuando algún dispositivo de punto extremo en el dominio extendido está en fallo, en las realizaciones de la presente invención, se puede impedir el acceso mutuo entre el dominio principal y el dispositivo de punto extremo en el dominio extendido, lo que impide la propagación de un fallo y asegura la disponibilidad del sistema.

20 La figura 1 es un diagrama de bloques de sistema de un sistema informático que incluye un dominio principal PCIe y un dominio extendido PCIe. El sistema informático incluye un dominio principal 100 y un dominio extendido 118, y el sistema informático extiende el dominio PCIe utilizando un dispositivo de punto extremo de complejo raíz (punto extremo de complejo raíz, RCEP) 106; el dominio principal 100 incluye un complejo raíz (complejo raíz, RC) 102, un conmutador (conmutador) 104 y por lo menos un dispositivo de punto extremo PCIe 107; el complejo raíz 102 está conectado a un puerto de arriba 104A del conmutador 104 utilizando un puerto raíz 103; un puerto de abajo 104B del conmutador 104 está conectado al dispositivo de punto extremo PCIe 107, de tal modo que el complejo raíz 102 está conectado al dispositivo de punto extremo PCIe 107 utilizando el conmutador 104; el complejo raíz 102 puede estar integrado en una CPU principal 101. La figura anterior toma un ejemplo en el que el dominio principal 100 incluye un conmutador. En otra realización, el dominio principal 100 puede incluir una serie de conmutadores, y cada conmutador puede estar conectado a uno o varios dispositivos de punto extremo PCIe.

35 El complejo raíz 102 está configurado para procesar y enviar una solicitud entre la CPU principal 101 y el dispositivo de punto extremo PCIe 107; el conmutador 104 está configurado para encaminar una solicitud hacia abajo al dispositivo de punto extremo PCIe conectado al puerto de abajo 104B, y encaminar una solicitud hacia arriba desde cada puerto de abajo independiente a un solo complejo raíz, y puede estar configurado además para encaminar una solicitud desde un puerto de abajo a otro puerto de abajo. El dispositivo de punto extremo PCIe 107 tiene una función de iniciar una solicitud y completar un proceso de transacción de PCIe, y el dispositivo de punto extremo PCIe 107 puede ser un dispositivo de almacenamiento, un adaptador de red, una tarjeta de audio o similar.

40 El dispositivo de punto extremo PCIe 107 en el dominio principal 100 incluye el RCEP 106. El RCEP 106 puede no sólo iniciar una solicitud y completar el proceso de transacción de PCIe, sino asimismo, dado que el RCEP 106 está dotado de un módulo de hardware, que tiene la misma función que un complejo raíz, y de un controlador de dispositivo, implementar la conexión entre el dominio extendido 118 y el dominio principal 100, y gestionar y enviar una solicitud entre el dominio extendido 118 y el dominio principal 100. Tal como se muestra en la figura 1, el dominio extendido 118 incluye el RCEP 106, que sirve como el complejo raíz en el dominio extendido, un conmutador 112, y segundos dispositivos de punto extremo 114 y 116 (pueden existir múltiples segundos dispositivos de punto extremo, y puede asimismo haber uno, y esta realización de la presente invención utiliza dos dispositivos de punto extremo como ejemplo). Los segundos dispositivos de punto extremo 114 y 116 están conectados independientemente al RCEP 116 utilizando el conmutador 112, y los segundos dispositivos de punto extremo 114 y 116 pueden ser dispositivos de almacenamiento, adaptadores de red, tarjetas de audio y similares. El dominio extendido 118 puede incluir además múltiples puertos raíz y múltiples conmutadores, y cada conmutador puede estar conectado a múltiples dispositivos de punto extremo.

55 La figura 2 es un diagrama de asignación de direcciones de memoria de dispositivos de punto extremo en el dominio principal y el dominio extendido en el sistema informático mostrado en la figura 1. Una dirección física de 64 bits 202 (específicamente, la dirección puede ser una dirección de entrada/salida con mapeo a memoria entrada/salida con mapeo a memoria, dirección MMIO) de la CPU principal 101 se puede dividir en una dirección de memoria 203 del dominio principal y una dirección de memoria 204 del dominio extendido.

En el sistema informático mostrado en la figura 1, cuando el sistema carga controladores de dispositivos de punto extremo en el dominio principal, el sistema asigna una dirección de memoria cada dispositivo de punto extremo en el dominio principal, por ejemplo, el sistema asigna direcciones de memoria al RCEP 106 y a un primer dispositivo de punto extremo 108. Tal como se muestra en la figura 2(a), una parte de la dirección MMIO 202, es decir, una dirección de memoria 205 del RCEP 106, se asigna al RCEP 106, y otra parte de la dirección MMIO 202, es decir, una dirección de memoria 210 del primer dispositivo de punto extremo 108, se asigna al primer dispositivo de punto extremo 108. Dado que el RCEP 106 y el primer dispositivo de punto extremo 108 son ambos dispositivos de punto extremo en el dominio principal, la dirección de memoria 205 y la dirección de memoria 210 forman conjuntamente la dirección de memoria 203 del dominio principal.

Cuando el sistema carga el controlador del RCEP 106, el sistema detecta controladores de ambos segundos dispositivos de punto extremo 114 y 116 en el dominio extendido, desencadena el escaneo sobre todos los segundos dispositivos de punto extremo en todo el dominio extendido, y asigna una dirección de memoria a cada segundo dispositivo de punto extremo en el dominio extendido. Específicamente, una parte de la dirección MMIO 202 se asigna a los segundos dispositivos de punto extremo en el dominio extendido, es decir, unas segundas direcciones de memoria (las segundas direcciones de memoria son direcciones de memoria, en el dominio extendido, de los segundos dispositivos de punto extremo, y se utilizan para representar, en el dominio extendido, los segundos dispositivos de punto extremo), por ejemplo, las segundas direcciones de memoria 206 y 207 de los segundos dispositivos de punto extremo 114 y 116 mostrados en la figura 2(a), se asignan a los segundos dispositivos de punto extremo. La cantidad de segundos dispositivos de punto extremo es igual que la cantidad de segundas direcciones de memoria de los segundos dispositivos de punto extremo. Esta realización utiliza dos segundos dispositivos de punto extremo como ejemplo, y por lo tanto se asignan las segundas direcciones de memoria 206 y 207 de los segundos dispositivos de punto extremo. Además, junto a las segundas direcciones de memoria que se asignan a los segundos dispositivos de punto extremo, es necesario asimismo asignar unas primeras direcciones de memoria (las primeras direcciones de memoria son direcciones de memoria, en el dominio principal, de los segundos dispositivos de punto extremo, y se utilizan para representar, en el dominio principal, los segundos dispositivos de punto extremo). Específicamente, las primeras direcciones de memoria de los segundos dispositivos de punto extremo son principalmente la totalidad o una parte de la dirección de memoria del RCEP, y la dirección de memoria 205 del RCEP se divide en función de la cantidad de segundos dispositivos de punto extremo en el dominio extendido. Si la cantidad de segundos dispositivos de punto extremo es n , la dirección de memoria 205 del RCEP se divide en n partes, y cada parte de la dirección memoria corresponde a un segundo dispositivo de punto extremo. La dirección de memoria 205 del RCEP mostrada en la figura 2(a) está dividida en dos partes: 208 y 209, donde 208 corresponde a la dirección de memoria 206 del segundo dispositivo de punto extremo 114 y es la primera dirección de memoria del segundo dispositivo de punto extremo 114; 209 corresponde a la dirección de memoria 207 del segundo dispositivo de punto extremo 116 y es la primera dirección de memoria del segundo dispositivo de punto extremo 116.

Cuando los segundos dispositivos de punto extremo 114 y 116 en el dominio extendido llevan a cabo un intercambio de mensajes con el dominio principal, se pueden utilizar cuatro modos de acceso, que son específicamente: acceso al espacio de configuración, acceso a interrupciones señalizadas por mensaje (interrupciones señalizadas por mensaje, MSI), acceso de entrada/salida con mapeo a memoria MMIO, y acceso directo a memoria (acceso directo a memoria, DMA), de tal modo que cada una de la segunda dirección de memoria 206 y la primera dirección de memoria 208 que el sistema asigna al segundo dispositivo de punto extremo pueden ser de hecho cuatro tipos de direcciones de memoria. En una relación de mapeo entre primeras direcciones de memoria en el dominio principal y segundas direcciones de memoria en el dominio extendido que se muestra en la figura 2 (b), la segunda dirección de memoria 206 del segundo dispositivo de punto extremo 114 se puede dividir en 206a, 206b, 206c y 206d, que se utilizan para llevar a cabo acceso al espacio de configuración, acceso MSI, acceso MMIO y acceso DMA en el segundo dispositivo de punto extremo 114, respectivamente. Por consiguiente, cuando el sistema asigna los cuatro tipos de direcciones de memoria al segundo dispositivo de punto extremo 114, el sistema divide la primera dirección de memoria 208 que está en la dirección de memoria del RCEP y corresponde al segundo dispositivo de punto extremo 114 en cuatro partes, que son específicamente una dirección del espacio de configuración 208a, una dirección MSI 208b, una dirección MMIO 208c y una dirección DMA 208d. Existe una relación de mapeo entre las primeras direcciones de memoria 208a, 208b, 208c y 208d del segundo dispositivo de punto extremo 114 y las segundas direcciones de memoria 206a, 206b, 206c y 206d del segundo dispositivo de punto extremo 114. Específicamente, la relación de mapeo se puede reflejar utilizando una relación de desplazamiento de direcciones, por ejemplo, existe una primera relación de desplazamiento de direcciones entre 208a y 206a, existe una segunda relación de desplazamiento de direcciones entre 208b y 206b, existe una tercera relación de desplazamiento de direcciones entre 208c y 206c y existe una cuarta relación de desplazamiento de direcciones entre 208d y 206d. La relación de mapeo entre las primeras direcciones de memoria del segundo dispositivo de punto extremo y las segundas direcciones de memoria del segundo dispositivo de punto extremo se puede almacenar en el RCEP 106, por ejemplo, almacenar en un módulo de transformación de direcciones del RCEP 106, donde el módulo de transformación de direcciones almacena las relaciones de desplazamiento de direcciones, y el módulo de transformación de direcciones puede llevar a cabo la transformación de direcciones en función de la relación de mapeo almacenada entre las segundas direcciones de memoria y las primeras direcciones de memoria.

En un proceso en que el sistema escanea los segundos dispositivos de punto extremo en todo el dominio extendido, además de asignar una segunda dirección de memoria a cada segundo dispositivo de punto extremo en el dominio extendido, el sistema asigna un identificador de bus/dispositivo/función (bus/dispositivo/función, BDF) a cada segundo dispositivo de punto extremo en el dominio extendido. Existe una relación de mapeo entre el identificador BDF del segundo dispositivo de punto extremo y la segunda dirección de memoria del segundo dispositivo de punto extremo, y el RCEP puede almacenar la relación de mapeo. Por ejemplo, el módulo de transformación de direcciones del RCEP 106 almacena una relación de mapeo entre el identificador BDF del segundo dispositivo de punto extremo 114 y la segunda dirección de memoria del segundo dispositivo de punto extremo 114, de tal modo que el RCEP 106 lleva a cabo una transformación mutua entre el identificador BDF del segundo dispositivo de punto extremo 114 y la segunda dirección de memoria del segundo dispositivo de punto extremo 114 de acuerdo con la relación de mapeo almacenada entre el identificador BDF del segundo dispositivo de punto extremo 114 y la segunda dirección de memoria del segundo dispositivo de punto extremo 114.

Si el segundo dispositivo de punto extremo 114 está en fallo, dado que existe una ventana de tiempo desde el momento en que el segundo dispositivo de punto extremo 114 genera un mensaje de interrupción hasta el momento en que un sistema operativo procesa el mensaje de interrupción, dentro de la ventana de tiempo, el segundo dispositivo de punto extremo en fallo 114 en el dominio extendido y otro dispositivo de punto extremo pueden seguir accediendo uno a otro, por ejemplo, el segundo dispositivo de punto extremo 114 comunica con un dispositivo de punto extremo en el dominio principal o comunica, utilizando una CPU en el dominio principal, con otro dispositivo, o la CPU u otro dispositivo de punto extremo en el dominio principal puede asimismo acceder al segundo dispositivo de punto extremo 114. Sin embargo, un acceso o comunicación que involucre el segundo dispositivo de punto extremo en fallo 114 pueden provocar que otro dispositivo falle, por ejemplo, provocar que el primer dispositivo de punto extremo 108 falle o provocar que la CPU lleve a cabo un procesamiento innecesario y repetido de un mensaje de error, afectando de ese modo al rendimiento del sistema y afectando severamente a la fiabilidad del sistema.

Procedimiento de aislamiento de fallo

Realización 1 de la presente invención

Esta realización de la presente invención da a conocer un procedimiento de aislamiento de fallo, que se utiliza para impedir el acceso mutuo entre un dominio principal y un dispositivo de punto extremo en un dominio extendido cuando el dispositivo de punto extremo en el dominio extendido está en fallo, impidiendo de ese modo la propagación de un fallo al dominio principal.

La figura 3 es un diagrama de flujo del procedimiento de aislamiento de fallo según esta realización de la presente invención, utilizado en un sistema informático interconectado por PCIe, donde el sistema informático incluye un dominio principal y un dominio extendido, el dominio principal incluye un complejo raíz, un primer dispositivo de punto extremo y un RCEP, y el dominio extendido incluye el RCEP y un segundo dispositivo de punto extremo. El procedimiento incluye:

101. Monitorizar el estado del segundo dispositivo de punto extremo en el dominio extendido.

El estado del segundo dispositivo de punto extremo puede incluir un estado de fallo y un estado sin fallo, donde el estado de fallo indica que el segundo dispositivo de punto extremo está en fallo y no puede funcionar normalmente, y el estado sin fallo indica que el segundo dispositivo de punto extremo en el dominio extendido puede funcionar normalmente. Que el RCEP monitorice el estado del segundo dispositivo de punto extremo en el dominio extendido puede consistir en: recibir un mensaje de error enviado por el segundo dispositivo de punto extremo, o recibir un mensaje de respuesta de detección del dispositivo, utilizado para indicar si existe el segundo dispositivo de punto extremo, y determinar el estado del segundo dispositivo de punto extremo de acuerdo con el mensaje de error o el mensaje de respuesta de detección del dispositivo.

102. Establecer un registro del estado del dispositivo de acuerdo con el estado del segundo dispositivo de punto extremo, donde el registro del estado del dispositivo incluye una correspondencia entre una información de identificador del segundo dispositivo de punto extremo y el estado del segundo dispositivo de punto extremo;

El RCEP puede establecer el registro del estado del dispositivo según el estado del segundo dispositivo de punto extremo, donde el registro del estado del dispositivo incluye la correspondencia entre la información de identificador del segundo dispositivo de punto extremo y el estado del segundo dispositivo de punto extremo, de tal modo que el RCEP puede determinar el estado del segundo dispositivo de punto extremo en función de la información de identificador del segundo dispositivo de punto extremo.

103. Recibir una solicitud de acceso, donde la solicitud de acceso incluye una solicitud de acceso desde el segundo dispositivo de punto extremo al dominio principal o una solicitud de acceso desde el dominio principal al segundo dispositivo de punto extremo.

Cuando el segundo dispositivo de punto extremo en el dominio extendido accede al dominio principal utilizando la solicitud de acceso, o el dominio principal accede al segundo dispositivo de punto extremo utilizando la solicitud de acceso, la solicitud de acceso es encaminada al RCEP, y el RCEP recibe la solicitud de acceso.

104. Consultar el registro del estado del dispositivo de acuerdo con la información de identificador que es del segundo dispositivo de punto extremo y está en la solicitud de acceso, y determinar el estado del segundo dispositivo de punto extremo.

5 La solicitud de acceso lleva la información de identificador del segundo dispositivo de punto extremo, y el RCEP puede consultar, en el registro del estado del dispositivo, la correspondencia entre la información de identificador del segundo dispositivo de punto extremo y el estado del segundo dispositivo de punto extremo, y determinar el estado del segundo dispositivo de punto extremo.

105. Si el estado del segundo dispositivo de punto extremo es un estado de fallo, desechar la solicitud de acceso para impedir una comunicación entre el segundo dispositivo de punto extremo y el dominio principal.

10 Cuando se determina que el estado del segundo dispositivo de punto extremo es el estado de fallo, la solicitud de acceso se desecha para impedir que el RCEP envíe la solicitud de acceso, impidiendo de ese modo la comunicación entre el segundo dispositivo de punto extremo y el dominio principal.

15 En esta realización de la presente invención, se monitoriza el estado del segundo dispositivo de punto extremo en el dominio extendido, y se establece el registro del estado del dispositivo de acuerdo con el estado del segundo dispositivo de punto extremo, donde el registro del estado del dispositivo incluye la correspondencia entre la información de identificador del segundo dispositivo de punto extremo y el estado del segundo dispositivo de punto extremo; después de la recepción de la solicitud de acceso entre el segundo dispositivo de punto extremo y el dominio principal, se consulta el registro del estado del dispositivo según la información de identificador que es del segundo dispositivo de punto extremo y está en la solicitud de acceso, y se determina el estado del segundo dispositivo de punto extremo; si el estado del segundo dispositivo de punto extremo es el estado de fallo, se desecha la solicitud de acceso para impedir la comunicación entre el segundo dispositivo de punto extremo en fallo y el dominio principal e impedir la propagación de un fallo al dominio principal, asegurando de ese modo la fiabilidad del sistema.

Realización 2 de la presente invención

25 La figura 4 es un diagrama de flujo de un procedimiento de aislamiento de fallo según esta realización de la presente invención, utilizado en un sistema informático interconectado por PCIe, donde el sistema informático incluye un dominio principal y un dominio extendido, el dominio principal incluye un complejo raíz, un primer dispositivo de punto extremo y un RCEP, y el dominio extendido incluye el RCEP y un segundo dispositivo de punto extremo; y el segundo dispositivo de punto extremo lleva a cabo una interacción de comunicación con el complejo raíz o el primer dispositivo de punto extremo en el dominio principal utilizando el RCEP. El procedimiento puede incluir:

201. Monitorizar el estado del segundo dispositivo de punto extremo en el dominio extendido.

35 El estado del segundo dispositivo de punto extremo incluye un estado de fallo y un estado sin fallo, donde el estado de fallo indica que el segundo dispositivo de punto extremo está en fallo y no puede funcionar normalmente, y el estado sin fallo indica que el segundo dispositivo de punto extremo en el dominio extendido puede funcionar normalmente. Que el RCEP monitorice el estado del segundo dispositivo de punto extremo en el dominio extendido incluye: recibir un mensaje de error enviado por el segundo dispositivo de punto extremo, o recibir un mensaje de respuesta de detección del dispositivo, utilizado para indicar si existe el segundo dispositivo de punto extremo, y determinar el estado del segundo dispositivo de punto extremo de acuerdo con el mensaje de error o el mensaje de respuesta de detección del dispositivo.

40 Específicamente, el RCEP puede enviar un mensaje de detección de dispositivo a un registro del espacio de configuración del segundo dispositivo de punto extremo, y obtener un mensaje de respuesta de detección del dispositivo devuelto por el segundo dispositivo de punto extremo; si el mensaje de respuesta de detección del dispositivo indica que el segundo dispositivo de punto extremo no existe, que indica que debido a un fallo el segundo dispositivo de punto extremo no puede ser detectado, el RCEP determina que el estado del segundo dispositivo de punto extremo es el estado de fallo; de lo contrario, el RCEP determina que el estado del segundo dispositivo de punto extremo es el estado sin fallo; o

45 el RCEP recibe un mensaje de error del segundo dispositivo de punto extremo, y determina un tipo de mensaje de error en función del mensaje de error; si el tipo de mensaje de error es un tipo de error no corregible (error no corregible), el RCEP determina que el estado del segundo dispositivo de punto extremo es el estado de fallo; de lo contrario, el RCEP determina que el estado del segundo dispositivo de punto extremo es el estado sin fallo.

50 Preferentemente, después de recibir el mensaje de error desde el segundo dispositivo de punto extremo y determinar que el estado del segundo dispositivo de punto extremo es el estado de fallo, el RCEP puede determinar además si el mensaje de error es un mensaje de error enviado repetidamente; si el mensaje de error es un mensaje de error enviado repetidamente, que indica que el segundo dispositivo de punto extremo ha enviado ya el mensaje de error al dominio principal para la realización del correspondiente tratamiento de errores, el RCEP desecha el mensaje de error para impedir que una CPU lleve a cabo de manera innecesaria y repetida el procesamiento del mensaje de error y para asegurar la fiabilidad del sistema; si el mensaje de error no es un mensaje de error enviado

repetidamente, que indica que el mensaje de error es un mensaje de error enviado por el segundo dispositivo de punto extremo por primera vez, el RCEP envía el mensaje de error a la CPU, de tal modo que la CPU lleva a cabo un tratamiento de errores sobre el segundo dispositivo de punto extremo.

Determinar si el mensaje de error es un mensaje de error enviado repetidamente, incluye específicamente:

- 5 obtener un identificador BDF o segundas direcciones de memoria que son del segundo dispositivo de punto extremo y transportadas en el mensaje de error;

consultar una correspondencia que existe entre las segundas direcciones de memoria del segundo dispositivo de punto extremo y el estado del segundo dispositivo de punto extremo y está registrada en el registro del estado del dispositivo, y determinar el estado del segundo dispositivo de punto extremo; o consultar una correspondencia que existe entre el identificador BDF del segundo dispositivo de punto extremo y el estado del segundo dispositivo de punto extremo y está registrada en el registro del estado del dispositivo, y determinar el estado del segundo dispositivo de punto extremo; o determinar primeras direcciones de memoria del segundo dispositivo de punto extremo de acuerdo con las segundas direcciones de memoria o el identificador BDF que es del segundo dispositivo de punto extremo y está en la solicitud de acceso, y determinar el estado del segundo dispositivo de punto extremo consultando una correspondencia que existe entre las primeras direcciones de memoria del segundo dispositivo de punto extremo y el estado del segundo dispositivo de punto extremo y está registrada en el registro del estado del dispositivo; y

si se determina que el estado del segundo dispositivo de punto extremo es el estado de fallo, determinar que el mensaje de error es un mensaje de error enviado repetidamente; y si se determina que el estado del segundo dispositivo de punto extremo es el estado sin fallo, determinar que el mensaje de error no es un mensaje de error enviado repetidamente.

Las primeras direcciones de memoria son direcciones de memoria, en el dominio principal, del segundo dispositivo de punto extremo, y se utilizan para representar, en el dominio principal, el segundo dispositivo de punto extremo; las segundas direcciones de memoria son direcciones de memoria, en el dominio extendido, del segundo dispositivo de punto extremo, y se utilizan para representar, en el dominio extendido, el segundo dispositivo de punto extremo.

202. Establecer un registro del estado del dispositivo de acuerdo con el estado del segundo dispositivo de punto extremo, donde el registro del estado del dispositivo incluye una correspondencia entre primeras direcciones de memoria del segundo dispositivo de punto extremo y el estado del segundo dispositivo de punto extremo.

El RCEP establece el registro del estado del dispositivo de acuerdo con el estado del segundo dispositivo de punto extremo, donde el registro del estado del dispositivo incluye la correspondencia entre las primeras direcciones de memoria del segundo dispositivo de punto extremo y el estado del segundo dispositivo de punto extremo, de tal modo que el RCEP puede determinar el estado del segundo dispositivo de punto extremo de acuerdo con las primeras direcciones de memoria del segundo dispositivo de punto extremo.

El establecimiento de un registro del estado del dispositivo según el estado del segundo dispositivo de punto extremo puede incluir específicamente:

obtener el identificador BDF o las segundas direcciones de memoria que son del segundo dispositivo de punto extremo y transportadas en el mensaje de error o el mensaje de respuesta de detección del dispositivo.

Obtener las primeras direcciones de memoria del segundo dispositivo de punto extremo de acuerdo con el identificador BDF o las segundas direcciones de memoria puede consistir en:

40 transformar, mediante el RCEP, las segundas direcciones de memoria en las primeras direcciones de memoria del segundo dispositivo de punto extremo de acuerdo con una relación de mapeo almacenada entre las segundas direcciones de memoria del segundo dispositivo de punto extremo y las primeras direcciones de memoria del segundo dispositivo de punto extremo; u obtener en primer lugar, mediante el RCEP, las segundas direcciones de memoria del segundo dispositivo de punto extremo de acuerdo con una relación de mapeo almacenada entre el identificador BDF del segundo dispositivo de punto extremo y las segundas direcciones de memoria del segundo dispositivo de punto extremo, y transformar a continuación las segundas direcciones de memoria en las primeras direcciones de memoria del segundo dispositivo de punto extremo de acuerdo con una relación de mapeo almacenada entre las segundas direcciones de memoria del segundo dispositivo de punto extremo y las primeras direcciones de memoria del segundo dispositivo de punto extremo.

50 La correspondencia entre las primeras direcciones de memoria del segundo dispositivo de punto extremo y el estado del segundo dispositivo de punto extremo está registrada en el registro del estado del dispositivo, de tal modo que el RCEP puede determinar el estado del segundo dispositivo de punto extremo de acuerdo con las primeras direcciones de memoria del segundo dispositivo de punto extremo.

55 Además, la correspondencia entre las segundas direcciones de memoria del segundo dispositivo de punto extremo y el estado del segundo dispositivo de punto extremo, o la correspondencia entre el identificador BDF del segundo dispositivo de punto extremo y el estado del segundo dispositivo de punto extremo se puede registrar asimismo en el

registro del estado del dispositivo, de tal modo que el RCEP puede determinar además el estado del segundo dispositivo de punto extremo de acuerdo con las primeras direcciones de memoria o el identificador BDF del segundo dispositivo de punto extremo.

5 Específicamente, dado que las primeras direcciones de memoria del segundo dispositivo de punto extremo incluyen una primera dirección de memoria de acceso al espacio de configuración, una primera dirección de memoria de acceso MSI, una primera dirección de memoria de acceso MMIO y una primera dirección de memoria de acceso DMA, y las segundas direcciones de memoria del segundo dispositivo de punto extremo incluyen una segunda dirección de memoria del acceso al espacio de configuración, una segunda dirección de memoria del acceso MSI, una segunda dirección de memoria del acceso MMIO y una segunda dirección de memoria del acceso DMA, el
10 RCEP puede obtener la primera dirección de memoria del acceso al espacio de configuración, la primera dirección de memoria del acceso MSI, la primera dirección de memoria del acceso MMIO y la primera dirección de memoria del acceso DMA del segundo dispositivo de punto extremo de acuerdo con una relación de mapeo almacenada entre cada segunda dirección de memoria del segundo dispositivo de punto extremo y cada primera dirección de memoria del segundo dispositivo de punto extremo;

15 alternativamente, el RCEP puede obtener la segunda dirección de memoria del acceso al espacio de configuración, la segunda dirección de memoria del acceso MSI, la segunda dirección de memoria del acceso MMIO y la segunda dirección de memoria del DMA del segundo dispositivo de punto extremo de acuerdo con la relación de mapeo entre el identificador BDF del segundo dispositivo de punto extremo y cada segunda dirección de memoria del segundo dispositivo de punto extremo, y obtener a continuación la primera dirección de memoria del acceso al espacio de
20 configuración, la primera dirección de memoria del acceso MSI, la primera dirección de memoria del acceso MMIO y la primera dirección de memoria del acceso DMA del segundo dispositivo de punto extremo de acuerdo con una relación de mapeo almacenada entre cada segunda dirección de memoria del segundo dispositivo de punto extremo y cada primera dirección de memoria del segundo dispositivo de punto extremo;

25 entonces, que la correspondencia entre las primeras direcciones de memoria del segundo dispositivo de punto extremo y el estado del segundo dispositivo de punto extremo se registre en el registro del estado del dispositivo se refiere específicamente a: registrar una correspondencia entre cada primera dirección de memoria del segundo dispositivo de punto extremo y el estado del segundo dispositivo de punto extremo último; que la correspondencia entre las segundas direcciones de memoria del segundo dispositivo de punto extremo y el estado del segundo dispositivo de punto extremo se registre en el registro del estado del dispositivo se refiere específicamente a:
30 registrar una correspondencia entre cada segunda dirección de memoria del segundo dispositivo de punto extremo y el estado del segundo dispositivo de punto extremo.

203. Recibir una solicitud de acceso, donde la solicitud de acceso incluye una solicitud de acceso desde el segundo dispositivo de punto extremo al dominio principal o una solicitud de acceso desde el dominio principal al segundo dispositivo de punto extremo.

35 Cuando el primer dispositivo de punto extremo y el dominio principal llevan a cabo intercambio de mensajes entre sí utilizando la solicitud de acceso, la solicitud de acceso puede ser una solicitud de acceso desde el segundo dispositivo de punto extremo en el dominio extendido y para acceder al dominio principal, y puede ser asimismo una solicitud de acceso desde el complejo raíz en el dominio principal o el primer dispositivo de punto extremo en el dominio principal y para acceder al segundo dispositivo de punto extremo; cuando la solicitud de acceso es desde el
40 dominio principal, la solicitud de acceso lleva las primeras direcciones de memoria del segundo dispositivo de punto extremo; cuando la solicitud de acceso es desde el dominio extendido, la solicitud de acceso lleva las segundas direcciones de memoria del segundo dispositivo de punto extremo o el identificador BDF del segundo dispositivo de punto extremo.

45 204. Consultar el registro del estado del dispositivo de acuerdo con la información de identificador que es del segundo dispositivo de punto extremo y está en la solicitud de acceso, y determinar el estado del segundo dispositivo de punto extremo.

La información de identificador del segundo dispositivo de punto extremo incluye un elemento de la siguiente información o una combinación de la misma: las primeras direcciones de memoria del segundo dispositivo de punto extremo y las segundas direcciones de memoria del segundo dispositivo de punto extremo.

50 Cuando la solicitud de acceso es desde el dominio principal, la correspondencia que existe entre las primeras direcciones de memoria del segundo dispositivo de punto extremo y el estado del segundo dispositivo de punto extremo y está registrada en el registro del estado del dispositivo es consultada de acuerdo con las primeras direcciones de memoria que son del segundo dispositivo de punto extremo y está en la solicitud de acceso, y se determina el estado del segundo dispositivo de punto extremo. Por ejemplo, cuando se utiliza un acceso MMIO, la
55 solicitud de acceso lleva la primera dirección de memoria del acceso MMIO del segundo dispositivo de punto extremo, y el registro del estado del dispositivo registra la correspondencia entre cada primera dirección de memoria del segundo dispositivo de punto extremo y el estado del segundo dispositivo de punto extremo. El RCEP puede consultar el registro del estado del dispositivo utilizando la primera dirección de memoria del acceso MMIO que es del segundo dispositivo de punto extremo y está en la solicitud de acceso, y determinar el estado del segundo
60 dispositivo de punto extremo.

5 Cuando la solicitud de acceso es desde el dominio extendido, el RCEP consulta el registro del estado del dispositivo de acuerdo con las segundas direcciones de memoria o el identificador BDF que es del segundo dispositivo de punto extremo y está en la solicitud de acceso; si el registro del estado del dispositivo no registra la correspondencia entre las segundas direcciones de memoria o el identificador BDF del segundo dispositivo de punto extremo y el estado del segundo dispositivo de punto extremo, el RCEP obtiene las primeras direcciones de memoria del segundo dispositivo de punto extremo de acuerdo con las segundas direcciones de memoria o el identificador BDF, y determina el estado del segundo dispositivo de punto extremo consultando la correspondencia que existe entre las primeras direcciones de memoria del segundo dispositivo de punto extremo y el estado del segundo dispositivo de punto extremo y está registrada en el registro del estado del dispositivo; si el registro del estado del dispositivo registra la correspondencia entre las segundas direcciones de memoria o el identificador BDF del segundo dispositivo de punto extremo y el estado del segundo dispositivo de punto extremo, el RCEP consulta directamente la correspondencia que existe entre las segundas direcciones de memoria o el identificador BDF del segundo dispositivo de punto extremo y el estado del segundo dispositivo de punto extremo y está registrada en el registro del estado del dispositivo, y determina el estado del segundo dispositivo de punto extremo, lo que evita transformar las segundas direcciones de memoria o el identificador BDF del segundo dispositivo de punto extremo en las primeras direcciones de memoria del segundo dispositivo de punto extremo, acelerando de ese modo el proceso de determinación del estado del dispositivo.

205. Si el estado del segundo dispositivo de punto extremo es un estado de fallo, desechar la solicitud de acceso para impedir la comunicación entre el segundo dispositivo de punto extremo y el dominio principal.

20 Cuando se determina que el estado del segundo dispositivo de punto extremo es el estado de fallo, la solicitud de acceso se desecha para impedir que el RCEP envíe la solicitud de acceso, impidiendo de ese modo la comunicación entre el segundo dispositivo de punto extremo y el dominio principal.

Además, el procedimiento puede incluir también:

206. Si el estado del segundo dispositivo de punto extremo es el estado de fallo, enviar un mensaje de aislamiento de fallo a una CPU, donde el mensaje de aislamiento de fallo se utiliza para ordenar a la CPU en el dominio principal que interrumpa el acceso al segundo dispositivo de punto extremo en el dominio extendido, y el mensaje de aislamiento de fallo lleva las primeras direcciones de memoria del segundo dispositivo de punto extremo.

Si el estado del segundo dispositivo de punto extremo es el estado de fallo, el RCEP puede enviar el mensaje de aislamiento de fallo a la CPU en el dominio principal, de tal modo que la CPU en el dominio principal interrumpe el acceso al segundo dispositivo de punto extremo en el dominio extendido, por ejemplo, la CPU puede desinstalar un controlador del segundo dispositivo de punto extremo en fallo o aislar un trayecto de E/S para acceder al segundo dispositivo de punto extremo en fallo.

Además, si la solicitud de acceso es una solicitud de acceso que es desde el dominio principal y para acceder al segundo dispositivo de punto extremo, el procedimiento puede incluir además:

206'. Si la solicitud de acceso es una solicitud de acceso enviada por el dominio principal, devolver al dominio principal un paquete de respuesta analógica en respuesta a la solicitud de acceso.

Específicamente, cuando la solicitud de acceso procedente del dominio principal y para acceder al segundo dispositivo de punto extremo es una solicitud de acceso de tipo con confirmación, es necesario devolver un mensaje de respuesta en respuesta a la solicitud de acceso; de lo contrario, el dominio principal puede generar un error de tiempo de espera de devolución de paquete, que provoca un reinicio del sistema informático. Sin embargo, cuando el segundo dispositivo de punto extremo está en fallo, la solicitud de acceso puede no llegar al segundo dispositivo de punto extremo, o aunque la solicitud de acceso llegue al segundo dispositivo de punto extremo, el segundo dispositivo de punto extremo no puede generar un mensaje de respuesta normal debido a un fallo. El RCEP puede generar el paquete de respuesta analógica en respuesta a la solicitud de acceso, y devolver el paquete de respuesta analógica al dominio principal, con el fin de impedir generar un error de tiempo de espera de devolución de paquete y provocar el reinicio del sistema informático, donde el paquete de respuesta analógica puede ser un paquete de solicitud no soportada (UR, Unsupported Request) o un paquete de abortar finalización (CA, Completion Abort).

La etapa 206 y la etapa 206' son dos etapas opcionales, y no son etapas que haya que llevar a cabo simultáneamente.

50 En esta realización de la presente invención, se monitoriza el estado del segundo dispositivo de punto extremo en el dominio extendido, y se establece el registro del estado del dispositivo de acuerdo con el estado del segundo dispositivo de punto extremo, donde el registro del estado del dispositivo incluye la correspondencia entre las primeras direcciones de memoria del segundo dispositivo de punto extremo y el estado del segundo dispositivo de punto extremo; después de que se recibe la solicitud de acceso entre el segundo dispositivo de punto extremo y el dominio principal, se obtienen las primeras direcciones de memoria que son del segundo dispositivo de punto extremo y están en la solicitud de acceso, o se obtienen las primeras direcciones de memoria del segundo dispositivo de punto extremo de acuerdo con el identificador BDF o las segundas direcciones de memoria que son del segundo dispositivo de punto extremo y están en la solicitud de acceso, y se consulta la correspondencia que

5 existe entre las primeras direcciones de memoria del segundo dispositivo de punto extremo y el estado del segundo dispositivo de punto extremo y en el registro del estado del dispositivo, y se determina el estado del segundo dispositivo de punto extremo; si el estado del segundo dispositivo de punto extremo es el estado de fallo, se desecha la solicitud de acceso para impedir la comunicación entre el segundo dispositivo de punto extremo en fallo y el dominio principal. Además, el mensaje de aislamiento de fallo se puede enviar asimismo a la CPU para ordenar a la CPU que interrumpa el acceso al segundo dispositivo de punto extremo en el dominio extendido, evitando de ese modo la propagación de un fallo al dominio principal.

10 Además, en esta realización de la presente invención, el registro del estado del dispositivo puede registrar además la correspondencia entre el identificador BDF o las segundas direcciones de memoria del segundo dispositivo de punto extremo y el estado del segundo dispositivo de punto extremo, de tal modo que el estado del segundo dispositivo de punto extremo se puede determinar consultando directamente el registro del estado del dispositivo de acuerdo con el identificador BDF, las segundas direcciones de memoria o las primeras direcciones de memoria que son del segundo dispositivo de punto extremo y están en la solicitud de acceso, lo que evita transformar las segundas direcciones de memoria o el identificador BDF del segundo dispositivo de punto extremo en las primeras direcciones de memoria del segundo dispositivo de punto extremo, acelerando de ese modo el proceso de determinación del estado del segundo dispositivo de punto extremo.

15 Además, en esta realización de la presente invención, cuando se monitoriza el estado del segundo dispositivo de punto extremo, después de que se recibe el mensaje de error enviado por el segundo dispositivo de punto extremo y se determina que el estado del segundo dispositivo de punto extremo es el estado de fallo en función del tipo de mensaje de error, se puede determinar además si el mensaje de error enviado por el segundo dispositivo de punto extremo es un mensaje de error enviado repetidamente; si el mensaje de error es un mensaje de error enviado repetidamente, se desecha el mensaje de error para impedir que el mensaje de error llegue al dominio principal, impedir la propagación de un error e impedir que la CPU realice de manera innecesaria y repetida el procesamiento del mensaje de error, asegurando de este modo la fiabilidad del sistema.

20 Realización 3 de la presente invención

25 Haciendo referencia al sistema informático mostrado en la figura 1, esta realización de la presente invención da a conocer un procedimiento de aislamiento de fallo mostrado en la figura 5. El segundo dispositivo de punto extremo 116 en el dominio extendido es un dispositivo en fallo, y el primer dispositivo de punto extremo 108 en el dominio principal envía una solicitud de acceso de tipo con confirmación para acceder al segundo dispositivo de punto extremo en fallo 116 en un modo de acceso DMA, donde la solicitud de acceso se encamina en primer lugar al RCEP 106. Cuando el segundo dispositivo de punto extremo está en fallo, la solicitud de acceso puede haber cruzado ya un límite del RCEP 106, es decir, puede ya haber sido enviada por el RCEP 106, o puede no haber cruzado el límite del RCEP 106, es decir, no ha sido enviada por el RCEP 106; por lo tanto, el procedimiento puede incluir específicamente:

30 301. El RCEP 106 monitoriza estados de todos los segundos dispositivos de punto extremo en el dominio extendido.

35 Los estados de los dispositivos incluyen un estado de fallo y un estado sin fallo. Que el RCEP 106 monitorice estados del segundo dispositivo de punto extremo 114 y el segundo dispositivo de punto extremo 116 en el dominio extendido, incluye específicamente: recibir un mensaje de error enviado por el segundo dispositivo de punto extremo 114 ó 116, o recibir un mensaje de respuesta de detección del dispositivo, utilizado para indicar si el segundo dispositivo de punto extremo 114 ó 116 existe; y determinar el estado del segundo dispositivo de punto extremo 114 ó 116 en función del mensaje de error o del mensaje de respuesta de detección del dispositivo.

40 302. Establecer un registro del estado del dispositivo de acuerdo con los estados de todos los segundos dispositivos de punto extremo en el dominio extendido, donde el registro del estado del dispositivo incluye una correspondencia entre segundas direcciones de memoria, en el dominio extendido, de los segundos dispositivos de punto extremo y los estados de los dispositivos.

45 Por ejemplo, el mensaje de error enviado por el segundo dispositivo de punto extremo 116 incluye un identificador BDF del segundo dispositivo de punto extremo 116; el RCEP 106 obtiene el identificador BDF del segundo dispositivo de punto extremo 116, obtiene una segunda dirección de memoria del acceso al espacio de configuración, una segunda dirección de memoria del acceso MSI, una segunda dirección de memoria del acceso MMIO y una segunda dirección de memoria del acceso DMA del segundo dispositivo de punto extremo 116 de acuerdo con una relación de mapeo entre el identificador BDF del segundo dispositivo de punto extremo 116 y cada segunda dirección de memoria del segundo dispositivo de punto extremo 116, obtiene una primera dirección de memoria del acceso al espacio de configuración, una primera dirección de memoria del acceso MSI, una primera dirección de memoria del acceso MMIO y una primera dirección de memoria del acceso DMA del segundo dispositivo de punto extremo 116 de acuerdo con una relación de mapeo almacenada entre cada segunda dirección de memoria del segundo dispositivo de punto extremo 116 y cada primera dirección de memoria del segundo dispositivo de punto extremo 116, y registra, en el registro del estado del dispositivo, una correspondencia entre cada primera dirección de memoria del segundo dispositivo de punto extremo 116 y el estado del segundo dispositivo de punto extremo 116, por ejemplo, marca cada primera dirección de memoria del segundo dispositivo de punto extremo de 116 en fallo.

Análogamente, si el segundo dispositivo de punto extremo 114 está en fallo, se registra una correspondencia entre cada primera dirección de memoria del segundo dispositivo de punto extremo 114 y el estado del segundo dispositivo de punto extremo 114 en el registro del estado del dispositivo, por ejemplo, cada primera dirección de memoria del segundo dispositivo de punto extremo 114 se marca en fallo.

- 5 303. Recibir una solicitud de acceso que procede del primer dispositivo de punto extremo 108 y es para acceder al segundo dispositivo de punto extremo 116, donde la solicitud de acceso lleva la primera dirección de memoria de acceso DMA del segundo dispositivo de punto extremo 116.

10 Cuando el primer dispositivo de punto extremo 108 realiza el acceso DMA sobre el segundo dispositivo de punto extremo 116, la solicitud de acceso se envía al RCEP 106 por medio de encaminamiento de direcciones, y el RCEP obtiene la primera dirección de memoria del acceso DMA que es del segundo dispositivo de punto extremo 116 y es transportada en la solicitud de acceso

304. Determinar el estado del segundo dispositivo de punto extremo 116 de acuerdo con la primera dirección de memoria del acceso DMA que es del segundo dispositivo de punto extremo 116 y está en la solicitud de acceso.

15 La correspondencia que existe entre las primeras direcciones de memoria de los segundos dispositivos de punto extremo y el estado del segundo dispositivo de punto extremo y está registrada en el registro del estado del dispositivo se consulta en función de la primera dirección de memoria del acceso DMA que es del segundo dispositivo de punto extremo 116 y está en la solicitud de acceso, y se determina el estado del segundo dispositivo de punto extremo 116.

20 Específicamente, cuando el segundo dispositivo de punto extremo 116 está en fallo, si la solicitud de acceso no ha cruzado el límite del RCEP, el estado del segundo dispositivo de punto extremo registrado en el registro del estado del dispositivo es el estado de fallo; se determina que el estado del segundo dispositivo de punto extremo 116 es el estado de fallo consultando la correspondencia que existe entre la primera dirección de memoria del acceso DMA del segundo dispositivo de punto extremo 116 y el estado del segundo dispositivo de punto extremo 116 y está registrada en el registro del estado del dispositivo; cuando el segundo dispositivo de punto extremo 116 está en fallo, si la solicitud de acceso ha cruzado ya el límite del RCEP, el estado del segundo dispositivo de punto extremo registrado en el registro del estado del dispositivo es el estado sin fallo, se consulta la correspondencia que existe entre la primera dirección de memoria del acceso DMA del segundo dispositivo de punto extremo 116 y el estado del segundo dispositivo de punto extremo 116 y está registrada en el registro del estado del dispositivo, y se determina que el estado del segundo dispositivo de punto extremo 116 es el estado sin fallo.

- 30 305. Si el estado del segundo dispositivo de punto extremo 116 es un estado de fallo, desechar la solicitud de acceso para impedir que el primer dispositivo de punto extremo 108 acceda al segundo dispositivo de punto extremo 116, y llevar a cabo a continuación la etapa 306.

35 Cuando el segundo dispositivo de punto extremo 116 está en fallo, si la solicitud de acceso no ha cruzado el límite del RCEP 106, el RCEP determina que el estado del segundo dispositivo de punto extremo 116 a acceder es el estado de fallo después de recibir la solicitud de acceso, y desecha la solicitud de acceso para impedir que el primer dispositivo de punto extremo 108 acceda al segundo dispositivo de punto extremo 116, evitando de ese modo propagar un fallo al dominio principal.

306. Devolver un paquete de respuesta analógica en respuesta a la solicitud de acceso al primer dispositivo de punto extremo 108.

40 Dado que la solicitud de acceso es una solicitud de acceso de tipo con confirmación, se genera el paquete de respuesta analógica en respuesta a la solicitud de acceso, y el paquete de respuesta analógica generado se devuelve al primer dispositivo de punto extremo 108, lo que impide que la CPU en el dominio principal genere un error de tiempo de espera de devolución de paquete y evita un reinicio del sistema informático.

45 En esta realización de la presente invención, el RCEP 106 monitoriza los estados de todos los segundos dispositivos de punto extremo en el dominio extendido, y establece el registro del estado del dispositivo de acuerdo con los estados de todos los segundos dispositivos de punto extremo en el dominio extendido; cuando el primer dispositivo de punto extremo 108 en el dominio principal envía una solicitud de acceso para acceder al segundo dispositivo de punto extremo 116 en el dominio extendido, el RCEP 106 recibe la solicitud de acceso, consulta el registro del estado del dispositivo de acuerdo con la primera dirección de memoria del acceso DMA en la solicitud de acceso, y determina el estado del segundo dispositivo de punto extremo. Si la solicitud de acceso no ha cruzado el límite del RCEP 106 cuando el segundo dispositivo de punto extremo 116 está en fallo, se determina en la etapa 304 que el estado del segundo dispositivo de punto extremo 116 es el estado de fallo. En este caso, el RCEP106 desecha la solicitud de acceso para impedir que el primer dispositivo de punto extremo 108 acceda al segundo dispositivo de punto extremo 116, evitando de ese modo la propagación del fallo al dominio principal. Además, el RCEP puede devolver asimismo el paquete de respuesta analógica en respuesta a la solicitud de acceso al primer dispositivo de punto extremo 108, lo que impide que la CPU en el dominio principal genere el error de tiempo de espera de devolución de paquete y evita el reinicio del sistema informático.

Además, si la solicitud de acceso ha cruzado ya el límite del RCEP 106 cuando el segundo dispositivo de punto extremo 116 está en fallo, se determina en la etapa 304 que el estado del segundo dispositivo de punto extremo 116 es el estado sin fallo. El RCEP 106 envía la solicitud de acceso al segundo dispositivo de punto extremo 116 según un proceso de funcionamiento normal; después de recibir la solicitud de acceso, el segundo dispositivo de punto extremo en fallo 116 puede ser activado por la solicitud de acceso para enviar un mensaje de error al RCEP 106, y el segundo dispositivo de punto extremo en fallo 116 puede enviar activamente un mensaje de error al RCEP 106 para notificar un fallo; el RCEP106 recibe del segundo dispositivo de punto extremo el mensaje de error y monitoriza el estado del segundo dispositivo de punto extremo; si el tipo del mensaje de error es un tipo de mensaje de error no corregible (error no corregible), el RCEP determina que el estado del segundo dispositivo de punto extremo está en el estado de fallo. El RCEP puede consultar además el registro del estado del dispositivo, y determinar si el mensaje de error es un mensaje de error enviado repetidamente; si el mensaje de error es un mensaje de error enviado repetidamente, el RCEP desecha el mensaje de error enviado repetidamente, lo que impide que la CPU lleve a cabo de manera innecesaria y repetida el procesamiento del mensaje de error, garantizando de ese modo la fiabilidad del sistema.

15 Realización 4 de la presente invención

Haciendo referencia al sistema informático mostrado en la figura 1, esta realización de la presente invención da a conocer un procedimiento de aislamiento de fallo mostrado en la figura 6. El segundo dispositivo de punto extremo 116 en el dominio extendido es un dispositivo en fallo, y el segundo dispositivo de punto extremo en fallo 116 envía una solicitud de acceso de tipo con confirmación para acceder a la CPU 101 en el dominio principal en un modo de acceso MMIO, donde la solicitud de acceso se encamina primero al RCEP 106. Cuando el segundo dispositivo de punto extremo está en fallo, la solicitud de acceso puede haber cruzado ya un límite del RCEP 106, es decir, puede ya haber sido enviada por el RCEP 106, o puede no haber cruzado el límite del RCEP 106, es decir, no ha sido enviada por el RCEP 106; por lo tanto, el procedimiento puede incluir específicamente:

401. El RCEP 106 monitoriza estados de todos los segundos dispositivos de punto extremo en el dominio extendido.

Los estados de los dispositivos incluyen un estado de fallo y un estado sin fallo. Que el RCEP 106 monitorice estados del segundo dispositivo de punto extremo 114 y el segundo dispositivo de punto extremo 116 en el dominio extendido, incluye específicamente: recibir un mensaje de error enviado por el segundo dispositivo de punto extremo 114 ó 116, o recibir un mensaje de respuesta de detección del dispositivo, utilizado para indicar si el segundo dispositivo de punto extremo 114 ó 116 existe; y determinar el estado del segundo dispositivo de punto extremo 114 ó 116 en función del mensaje de error o del mensaje de respuesta de detección del dispositivo.

402. Establecer un registro del estado del dispositivo de acuerdo con los estados de todos los segundos dispositivos de punto extremo en el dominio extendido, donde el registro del estado del dispositivo incluye una correspondencia entre segundas direcciones de memoria, en el dominio extendido, de los segundos dispositivos de punto extremo y los estados de los dispositivos.

Por ejemplo, el mensaje de error enviado por el segundo dispositivo de punto extremo 116 incluye un identificador BDF del segundo dispositivo de punto extremo 116; el RCEP 106 obtiene el identificador BDF del segundo dispositivo de punto extremo 116, obtiene una segunda dirección de memoria del acceso al espacio de configuración, una segunda dirección de memoria del acceso MSI, una segunda dirección de memoria del acceso MMIO y una segunda dirección de memoria del acceso DMA del segundo dispositivo de punto extremo 116 de acuerdo con una relación de mapeo entre el identificador BDF del segundo dispositivo de punto extremo 116 y las segundas direcciones de memoria del segundo dispositivo de punto extremo 116, obtiene una primera dirección de memoria del acceso al espacio de configuración, una primera dirección de memoria del acceso MSI, una primera dirección de memoria del acceso MMIO y una primera dirección de memoria del acceso DMA del segundo dispositivo de punto extremo 116 de acuerdo con una relación de mapeo almacenada entre las segundas direcciones de memoria del segundo dispositivo de punto extremo 116 y las primeras direcciones de memoria del segundo dispositivo de punto extremo 116, y registra, en el registro del estado del dispositivo, una correspondencia entre cada primera dirección de memoria del segundo dispositivo de punto extremo 116 y el estado del segundo dispositivo de punto extremo 116, por ejemplo, marca cada primera dirección de memoria del segundo dispositivo de punto extremo de 116 en fallo.

Análogamente, si el segundo dispositivo de punto extremo 114 está en fallo, se registra en el registro del estado del dispositivo una correspondencia entre cada primera dirección de memoria del segundo dispositivo de punto extremo 114 y el estado del segundo dispositivo de punto extremo 114 y una correspondencia entre un identificador BDF del segundo dispositivo de punto extremo 114 y el estado del segundo dispositivo de punto extremo 114.

403. Recibir una solicitud de acceso que procede del segundo dispositivo de punto extremo 116 y es para acceder a la CPU principal 101, donde la solicitud de acceso lleva la segunda dirección de memoria del acceso MMIO del segundo dispositivo de punto extremo 116.

Cuando el segundo dispositivo de punto extremo 116 lleva a cabo el acceso MMIO sobre la CPU principal 101, la solicitud de acceso se envía al RCEP 106 por medio de encaminamiento de direcciones, y el RCEP106 obtiene la

segunda dirección de memoria del acceso MMIO que es del segundo dispositivo de punto extremo 116 y es transportada en la solicitud de acceso.

404. Determinar el estado del segundo dispositivo de punto extremo 116 de acuerdo con la segunda dirección de memoria del acceso MMIO que es del segundo dispositivo de punto extremo 116 y está en la solicitud de acceso.

5 La primera dirección de memoria del acceso MMIO del segundo dispositivo de punto extremo 116 se obtiene de acuerdo con la segunda dirección de memoria del acceso MMIO que es del segundo dispositivo de punto extremo 116 y está en la solicitud de acceso, y utilizando una relación de mapeo almacenada entre cada primera dirección de memoria del segundo dispositivo de punto extremo y cada segunda dirección de memoria del segundo dispositivo de punto extremo; se consulta una correspondencia que existe entre las primeras direcciones de memoria del segundo dispositivo de punto extremo y el estado del segundo dispositivo de punto extremo y está registrada en el registro del estado del dispositivo, y se determina el estado del segundo dispositivo de punto extremo 116.

15 Específicamente, cuando el segundo dispositivo de punto extremo 116 está en fallo, si la solicitud de acceso no ha cruzado el límite del RCEP, se determina que el estado del segundo dispositivo de punto extremo 116 es el estado de fallo consultando una relación que existe entre la primera dirección de memoria del acceso DMA del segundo dispositivo de punto extremo 116 y el estado del segundo dispositivo de punto extremo 116 y está registrada en el registro del estado del dispositivo; cuando el segundo dispositivo de punto extremo 116 está en fallo, si la solicitud de acceso ha cruzado ya el límite del RCEP, se consulta la relación que existe entre la primera dirección de memoria del acceso DMA del segundo dispositivo de punto extremo 116 y el estado del segundo dispositivo de punto extremo 116 y está registrada en el registro del estado del dispositivo, y se determina que el estado del segundo dispositivo de punto extremo 116 es el estado sin fallo.

20 405. Si el estado del segundo dispositivo de punto extremo 116 es en un estado de fallo, desechar la solicitud de acceso para impedir que el segundo dispositivo de punto extremo 116 acceda a la CPU principal 101.

25 Cuando el segundo dispositivo de punto extremo 116 está en fallo, si la solicitud de acceso no ha cruzado el límite del RCEP 106, el RCEP106 determina que el estado del segundo dispositivo de punto extremo 116 accedido a acceder es el estado de fallo después de recibir la solicitud de acceso, y desecha la solicitud de acceso para impedir que el segundo dispositivo de punto extremo 116 acceda a la CPU principal 101, evitando de ese modo propagar un fallo al dominio principal.

30 En esta realización de la presente invención, el RCEP 106 monitoriza los estados de todos los segundos dispositivos de punto extremo en el dominio extendido, y establece el registro del estado del dispositivo de acuerdo con los estados de todos los segundos dispositivos de punto extremo en el dominio extendido; cuando el segundo dispositivo de punto extremo 116 en el dominio extendido envía una solicitud de acceso para acceder a la CPU principal 101 en el dominio principal, el RCEP106 recibe la solicitud de acceso, obtiene la primera dirección de memoria del acceso MMIO del segundo dispositivo de punto extremo de acuerdo con la segunda dirección de memoria del acceso MMIO que está en la solicitud de acceso, consulta el registro del estado del dispositivo y determina el estado del segundo dispositivo de punto extremo; si el segundo dispositivo de punto extremo 116 está en fallo y la solicitud de acceso no ha cruzado el límite del RCEP 106, se determina en la etapa 404 que el estado del segundo dispositivo de punto extremo 116 es el estado de fallo. En este caso, el RCEP desecha la solicitud de acceso para impedir que el segundo dispositivo de punto extremo 116 acceda a la CPU principal 101, evitando de ese modo propagar el fallo al dominio principal.

40 Además, si la solicitud de acceso ha cruzado ya el límite del RCEP 106 cuando el segundo dispositivo de punto extremo 116 está en fallo, se determina en la etapa 404 que el estado del segundo dispositivo de punto extremo 116 es el estado sin fallo. El RCEP 106 envía la solicitud de acceso a la CPU principal 101 según un proceso de funcionamiento normal; después de recibir la solicitud de acceso, la CPU principal 101 devuelve un paquete de respuesta, en respuesta a la solicitud de acceso, donde el paquete de respuesta devuelto alcanza en primer lugar el RCEP. Dado que el segundo dispositivo de punto extremo se ha puesto ya en fallo, enviar el paquete de respuesta devuelto al segundo dispositivo de punto extremo en fallo 116 carece de sentido, y puede desencadenar que el segundo dispositivo de punto extremo en fallo 116 envíe repetidamente un mensaje de error. Por lo tanto, el RCEP puede desechar el paquete de respuesta devuelto.

50 Además, el segundo dispositivo de punto extremo en fallo 116 puede enviar activamente un mensaje de error al RCEP para notificar el fallo; después de recibir el mensaje de error desde el segundo dispositivo de punto extremo, el RCEP monitoriza el estado del segundo dispositivo de punto extremo; si el tipo de mensaje de error es un tipo de mensaje de error no corregible (error no corregible), se determina que el estado del segundo dispositivo de punto extremo es el estado de fallo, y el RCEP puede consultar además el registro del estado del dispositivo, y determinar si el mensaje de error es un mensaje de error enviado repetidamente; si el mensaje de error es un mensaje de error enviado repetidamente, el RCEP desecha el mensaje de error enviado repetidamente para evitar la propagación del fallo.

Aparatos acordes con realizaciones de la presente invención

Las realizaciones de la presente invención dan a conocer aparatos de aislamiento de fallo, que se utilizan para impedir el acceso mutuo entre un dominio principal y un dispositivo de punto extremo en un dominio extendido cuando el dispositivo de punto extremo en el dominio extendido está en fallo, evitando de ese modo propagar un fallo al dominio principal.

La figura 7 es un diagrama estructural del aparato de aislamiento de fallo según una realización de la presente invención, utilizado en un sistema informático interconectado por PCIe, donde el sistema informático incluye un dominio principal y un dominio extendido, el dominio principal incluye un complejo raíz, un primer dispositivo de punto extremo y un RCEP, y el dominio extendido incluye el RCEP y un segundo dispositivo de punto extremo. El aparato incluye:

una unidad de monitorización 701, configurada para monitorizar un estado del segundo dispositivo de punto extremo en el dominio extendido;

una unidad de registro 702, configurada para establecer un registro del estado del dispositivo de acuerdo con el estado del segundo dispositivo de punto extremo, donde el registro del estado del dispositivo incluye una correspondencia entre información de identificador del segundo dispositivo de punto extremo y el estado del segundo dispositivo de punto extremo;

una unidad de recepción 703, configurada para recibir una solicitud de acceso, donde la solicitud de acceso incluye una solicitud de acceso del segundo dispositivo de punto extremo al dominio principal o una solicitud de acceso del dominio principal al segundo dispositivo de punto extremo;

una unidad de determinación 704, configurada para consultar el registro del estado del dispositivo de acuerdo con la información de identificador que es del segundo dispositivo de punto extremo y está en la solicitud de acceso, y determinar el estado del segundo dispositivo de punto extremo; y

una unidad de procesamiento 705, configurada para: cuando el estado del segundo dispositivo de punto extremo es un estado de fallo, desechar la solicitud de acceso para impedir la comunicación entre el segundo dispositivo de punto extremo y el dominio principal.

Específicamente, el estado del segundo dispositivo de punto extremo incluye un estado de fallo y un estado sin fallo; la unidad de monitorización 701 puede recibir un mensaje de error enviado por el segundo dispositivo de punto extremo o recibir un mensaje de respuesta de detección del dispositivo, utilizado para indicar si el segundo dispositivo de punto extremo existe, y determinar el estado del segundo dispositivo de punto extremo en función del mensaje de error o del mensaje de respuesta de detección del dispositivo; la unidad de registro 702 establece un registro del estado del dispositivo de acuerdo con el estado del segundo dispositivo de punto extremo en la unidad de monitorización 701, donde el registro del estado del dispositivo incluye la correspondencia entre la información de identificador del segundo dispositivo de punto extremo y el estado del segundo dispositivo de punto extremo; después de que la unidad de recepción 703 reciba la solicitud de acceso entre el segundo dispositivo de punto extremo y el dominio principal, la unidad de determinación 704 consulta el registro del estado del dispositivo según la información de identificador que es del segundo dispositivo de punto extremo y está en la solicitud de acceso, y determina el estado del segundo dispositivo de punto extremo; la unidad de procesamiento 705 está configurada para: cuando la unidad de determinación 704 determina que el estado del segundo dispositivo de punto extremo es el estado de fallo, desechar la solicitud de acceso para impedir la comunicación entre el segundo dispositivo de punto extremo en fallo y el dominio principal e impedir la propagación de un fallo al dominio principal, asegurando de ese modo la fiabilidad del sistema.

La figura 8 es un diagrama estructural de un aparato de aislamiento de fallo según una realización de la presente invención, utilizado en un sistema informático interconectado por PCIe, donde el sistema informático incluye un dominio principal y un dominio extendido, el dominio principal incluye un complejo raíz, un primer dispositivo de punto extremo y un RCEP, y el dominio extendido incluye el RCEP y un segundo dispositivo de punto extremo; y el segundo dispositivo de punto extremo lleva a cabo una interacción de comunicación con el complejo raíz o el primer dispositivo de punto extremo en el dominio principal utilizando el RCEP. El aparato puede incluir una unidad de monitorización 801, una unidad de registro 802, una unidad de recepción 803, una unidad de determinación 804 y una unidad de procesamiento 805, y el aparato de aislamiento de fallo puede ser el RCEP.

La unidad de monitorización 801 está configurada para monitorizar un estado del segundo dispositivo de punto extremo en el dominio extendido, donde el estado del segundo dispositivo de punto extremo incluye un estado de fallo y un estado sin fallo, donde el estado de fallo indica que el segundo dispositivo de punto extremo está en fallo y no puede funcionar normalmente, y el estado sin fallo indica que el segundo dispositivo de punto extremo en el dominio extendido puede funcionar normalmente. Que la unidad de monitorización 801 monitorice el estado del segundo dispositivo de punto extremo en el dominio extendido incluye: recibir un mensaje de error enviado por el segundo dispositivo de punto extremo, o recibir un mensaje de respuesta de detección del dispositivo, utilizado para indicar si existe el segundo dispositivo de punto extremo, y determinar el estado del segundo dispositivo de punto extremo de acuerdo con el mensaje de error o el mensaje de respuesta de detección del dispositivo.

Específicamente, la unidad de monitorización 801 puede enviar un mensaje de detección de dispositivo a un registro del espacio de configuración del segundo dispositivo de punto extremo, y obtener un mensaje de respuesta de detección del dispositivo devuelto por el segundo dispositivo de punto extremo; si el mensaje de respuesta de detección del dispositivo indica que el segundo dispositivo de punto extremo no existe, lo que indica que debido a un fallo el segundo dispositivo de punto extremo no puede ser detectado, se determina que el estado del segundo dispositivo de punto extremo es el estado de fallo; de lo contrario, se determina que el estado del segundo dispositivo de punto extremo es el estado sin fallo; o la unidad de monitorización 801 recibe un mensaje de error desde el segundo dispositivo de punto extremo, y determina un tipo de mensaje de error en función del mensaje de error; si el tipo del mensaje de error es un tipo de mensaje de error no corregible (error no corregible), se determina que el estado del segundo dispositivo de punto extremo es el estado de fallo; de lo contrario, se determina que el estado del segundo dispositivo de punto extremo es el estado sin fallo.

La unidad de registro 802 está configurada para establecer un registro del estado del dispositivo de acuerdo con el estado del segundo dispositivo de punto extremo, donde el registro del estado del dispositivo incluye una correspondencia entre primeras direcciones de memoria del segundo dispositivo de punto extremo y el estado del segundo dispositivo de punto extremo, donde las primeras direcciones de memoria son direcciones de memoria, en el dominio principal, del segundo dispositivo de punto extremo, y se utilizan para representar, en el dominio principal, el segundo dispositivo de punto extremo.

La unidad de registro 802 incluye específicamente: una subunidad del módulo de transformación de direcciones 802a y una subunidad de registro de estado 802b, donde la subunidad del módulo de transformación de direcciones 802a está configurada para obtener un identificador BDF o segundas direcciones de memoria que son del segundo dispositivo de punto extremo y se transportan en el mensaje de error o el mensaje de respuesta de detección del dispositivo, obtener las primeras direcciones de memoria del segundo dispositivo de punto extremo de acuerdo con el identificador BDF o las segundas direcciones de memoria; la subunidad del módulo de transformación de direcciones 802a almacena una relación de mapeo entre las segundas direcciones de memoria del segundo dispositivo de punto extremo y las primeras direcciones de memoria del segundo dispositivo de punto extremo y una relación de mapeo entre el identificador BDF del segundo dispositivo de punto extremo y las segundas direcciones de memoria del segundo dispositivo de punto extremo, y transforma las segundas direcciones de memoria en las primeras direcciones de memoria del segundo dispositivo de punto extremo de acuerdo con la relación de mapeo almacenada entre las segundas direcciones de memoria del segundo dispositivo de punto extremo y las primeras direcciones de memoria del segundo dispositivo de punto extremo, u obtiene las segundas direcciones de memoria del segundo dispositivo de punto extremo en función de la relación de mapeo almacenada entre el identificador BDF del segundo dispositivo de punto extremo y las segundas direcciones de memoria del segundo dispositivo de punto extremo, y transforma las segundas direcciones de memoria en las primeras direcciones de memoria del segundo dispositivo de punto extremo de acuerdo con la relación de mapeo almacenada entre las segundas direcciones de memoria del segundo dispositivo de punto extremo y las primeras direcciones de memoria del segundo dispositivo de punto extremo; la subunidad de registro de estado 802b está configurada para registrar, en el registro del estado del dispositivo, una correspondencia entre las primeras direcciones de memoria del segundo dispositivo de punto extremo y el estado del segundo dispositivo de punto extremo, de tal modo que el RCEP puede determinar el estado del segundo dispositivo de punto extremo de acuerdo con las primeras direcciones de memoria del segundo dispositivo de punto extremo, donde las segundas direcciones de memoria del segundo dispositivo de punto extremo son direcciones de memoria, en el dominio extendido, del segundo dispositivo de punto extremo, y se utilizan para representar, en el dominio extendido, el segundo dispositivo de punto extremo.

Además, la subunidad de registro de estado 802b puede estar configurada adicionalmente para registrar, en el registro del estado del dispositivo, la correspondencia entre las primeras direcciones de memoria del segundo dispositivo de punto extremo y el estado del segundo dispositivo de punto extremo, o la correspondencia entre el identificador BDF del segundo dispositivo de punto extremo y el estado del segundo dispositivo de punto extremo, de tal modo que la unidad de determinación 803 puede determinar además el estado del segundo dispositivo de punto extremo de acuerdo con las primeras direcciones de memoria o el identificador BDF del segundo dispositivo de punto extremo.

Específicamente, debido a que las primeras direcciones de memoria del segundo dispositivo de punto extremo incluyen una primera dirección de memoria del acceso al espacio de configuración, una primera dirección de memoria de acceso MSI, una primera dirección de memoria de acceso MMIO y una primera dirección de memoria de acceso DMA, y las segundas direcciones de memoria del segundo dispositivo de punto extremo incluyen una segunda dirección de memoria del acceso al espacio de configuración, una segunda dirección de memoria del acceso MSI, una segunda dirección de memoria del acceso MMIO y una segunda dirección de memoria del acceso DMA, la subunidad del módulo de transformación de direcciones 802a está configurada específicamente para: almacenar una relación de mapeo entre cada segunda dirección de memoria del segundo dispositivo de punto extremo y cada primera dirección de memoria del segundo dispositivo de punto extremo y una relación de mapeo entre el identificador BDF del segundo dispositivo de punto extremo y cada segunda dirección de memoria del segundo dispositivo de punto extremo; y obtener la primera dirección de memoria del acceso al espacio de configuración, la primera dirección de memoria del acceso MSI, la primera dirección de memoria del acceso MMIO y la primera dirección de memoria del acceso DMA del segundo dispositivo de punto extremo de acuerdo con la relación de mapeo almacenada entre cada segunda dirección de memoria del segundo dispositivo de punto extremo

y cada primera dirección de memoria del segundo dispositivo de punto extremo, u obtener primero la segunda dirección de memoria del acceso al espacio de configuración, la segunda dirección de memoria del acceso MSI, la segunda dirección de memoria del acceso MMIO y la segunda dirección de memoria del acceso DMA del segundo dispositivo de punto extremo de acuerdo con la relación de mapeo entre el identificador BDF del segundo dispositivo de punto extremo y cada segunda dirección de memoria del segundo dispositivo de punto extremo, y obtener a continuación la primera dirección de memoria del acceso al espacio de configuración, la primera dirección de memoria del acceso MSI, la primera dirección de memoria del acceso MMIO y la primera dirección de memoria del acceso DMA del segundo dispositivo de punto extremo de acuerdo con la relación de mapeo almacenada entre cada segunda dirección de memoria del segundo dispositivo de punto extremo y cada primera dirección de memoria del segundo dispositivo de punto extremo; la subunidad de registro 802b está configurada específicamente para registrar una correspondencia entre cada primera dirección de memoria del segundo dispositivo de punto extremo y el estado del segundo dispositivo de punto extremo, o está configurada además para registrar una correspondencia entre cada segunda dirección de memoria del segundo dispositivo de punto extremo y el estado del segundo dispositivo de punto extremo, o está configurada además para registrar la correspondencia entre el identificador BDF del segundo dispositivo de punto extremo y el estado del segundo dispositivo de punto extremo.

La unidad de recepción 803 está configurada para recibir una solicitud de acceso, donde la solicitud de acceso incluye una solicitud de acceso desde el segundo dispositivo de punto extremo al dominio principal o una solicitud de acceso desde el dominio principal al segundo dispositivo de punto extremo; cuando la solicitud de acceso es desde el dominio principal, la solicitud de acceso lleva las primeras direcciones de memoria del segundo dispositivo de punto extremo; cuando la solicitud de acceso es desde el dominio extendido, la solicitud de acceso lleva las segundas direcciones de memoria del segundo dispositivo de punto extremo o el identificador BDF del segundo dispositivo de punto extremo.

La unidad de determinación 804 está configurada para consultar el registro del estado del dispositivo de acuerdo con la información de identificador que es del segundo dispositivo de punto extremo y está en la solicitud de acceso, y determinar el estado del segundo dispositivo de punto extremo, donde la información de identificador del segundo dispositivo de punto extremo incluye un elemento de la siguiente información o una combinación de los mismos: las primeras direcciones de memoria del segundo dispositivo de punto extremo, las segundas direcciones de memoria del segundo dispositivo de punto extremo y el identificador BDF del segundo dispositivo de punto extremo. Específicamente, cuando la solicitud de acceso es desde el dominio principal, la unidad de determinación 804 consulta, de acuerdo con las primeras direcciones de memoria que son del segundo dispositivo de punto extremo y están en la solicitud de acceso, la correspondencia que existe entre las primeras direcciones de memoria del segundo dispositivo de punto extremo y el estado del segundo dispositivo de punto extremo y está registrada en el registro del estado del dispositivo, y determina el estado del segundo dispositivo de punto extremo. Por ejemplo, cuando se utiliza un acceso MMIO, la solicitud de acceso lleva la primera dirección de memoria del acceso MMIO del segundo dispositivo de punto extremo, y el registro del estado del dispositivo registra la correspondencia entre cada primera dirección de memoria del segundo dispositivo de punto extremo y el estado del segundo dispositivo de punto extremo. La unidad de determinación 804 puede consultar el registro del estado del dispositivo utilizando la primera dirección de memoria del acceso MMIO que es del segundo dispositivo de punto extremo y está en la solicitud de acceso, y determinar el estado del segundo dispositivo de punto extremo. Cuando la solicitud de acceso es desde el dominio extendido, la unidad de determinación 804 consulta el registro del estado del dispositivo de acuerdo con las segundas direcciones de memoria o el identificador BDF que es del segundo dispositivo de punto extremo y está en la solicitud de acceso; si el registro del estado del dispositivo no registra la correspondencia entre las segundas direcciones de memoria o el identificador BDF del segundo dispositivo de punto extremo y el estado del segundo dispositivo de punto extremo, la unidad de determinación 804 obtiene las primeras direcciones de memoria del segundo dispositivo de punto extremo de acuerdo con las segundas direcciones de memoria o el identificador BDF, y determina el estado del segundo dispositivo de punto extremo consultando la correspondencia que existe entre las primeras direcciones de memoria del segundo dispositivo de punto extremo y el estado del segundo dispositivo de punto extremo y está registrada en el registro del estado del dispositivo; si el registro del estado del dispositivo registra la correspondencia entre las segundas direcciones de memoria o el identificador BDF del segundo dispositivo de punto extremo y el estado del segundo dispositivo de punto extremo, la unidad de determinación 804 determina el estado del segundo dispositivo de punto extremo consultando directamente la correspondencia que existe entre las segundas direcciones de memoria o el identificador BDF del segundo dispositivo de punto extremo y el estado del segundo dispositivo de punto extremo y está registrada en el registro del estado del dispositivo, lo que evita transformar las segundas direcciones de memoria o el identificador BDF del segundo dispositivo de punto extremo en las primeras direcciones de memoria del segundo dispositivo de punto extremo, acelerando de ese modo el proceso de determinación del estado del dispositivo.

La unidad de procesamiento 805 está configurada para: cuando la unidad de determinación 804 determina que el estado del segundo dispositivo de punto extremo es el estado de fallo, desechar la solicitud de acceso para impedir la comunicación entre el segundo dispositivo de punto extremo y el dominio principal.

La unidad de procesamiento 805 está configurada además para: cuando la unidad de determinación 804 determina que el estado del segundo dispositivo de punto extremo es el estado de fallo, enviar un mensaje de aislamiento de fallo a una CPU, de tal modo que la CPU en el dominio principal interrumpe el acceso al segundo dispositivo de punto extremo en el dominio extendido, por ejemplo, la CPU puede desinstalar un controlador del segundo

dispositivo de punto extremo en fallo o aislar un trayecto de E/S que es para acceder al segundo dispositivo de punto extremo en fallo, donde el mensaje de aislamiento de fallo lleva las primeras direcciones de memoria del segundo dispositivo de punto extremo.

5 La unidad de procesamiento 805 está configurada además para: cuando la solicitud de acceso es una solicitud de acceso enviada por el dominio principal, devolver al dominio principal un paquete de respuesta analógica en respuesta a la solicitud de acceso. Específicamente, cuando la solicitud de acceso procedente del dominio principal y para acceder al segundo dispositivo de punto extremo es una solicitud de acceso de tipo con confirmación, es necesario devolver un mensaje de respuesta en respuesta a la solicitud de acceso; de lo contrario, el dominio principal puede generar un error de tiempo de espera de devolución de paquete, que provoca un reinicio del sistema informático. Sin embargo, cuando el segundo dispositivo de punto extremo está en fallo, la solicitud de acceso puede no llegar al segundo dispositivo de punto extremo, o aunque la solicitud de acceso llegue al segundo dispositivo de punto extremo, el segundo dispositivo de punto extremo no puede generar un mensaje de respuesta normal debido a un fallo. La unidad de procesamiento 805 puede generar el paquete de respuesta analógica en respuesta a la solicitud de acceso, y devolver el paquete de respuesta analógica al dominio principal, con el fin de impedir generar un error de tiempo de espera de devolución de paquete y provocar el reinicio del sistema informático, donde el paquete de respuesta analógica puede ser un paquete de solicitud no soportada (UR, Unsupported Request) o un paquete de abortar finalización (CA, Completion Abort).

20 Preferentemente, después de determinar que el estado del segundo dispositivo de punto extremo es el estado de fallo de acuerdo con el mensaje de error recibido, la unidad de monitorización 801 está configurada además para: determinar si el mensaje de error es un mensaje de error enviado repetidamente; si el mensaje de error es un mensaje de error enviado repetidamente, que indica que el segundo dispositivo de punto extremo ha enviado ya el mensaje de error al dominio principal para la realización del correspondiente tratamiento de errores, desechar el mensaje de error para impedir que la CPU lleve a cabo un procesamiento innecesario y repetido sobre el mensaje de error y para asegurar la fiabilidad del sistema; y si el mensaje de error no es un mensaje de error enviado repetidamente, que indica que el mensaje de error es un mensaje de error enviado por el segundo dispositivo de punto extremo por primera vez, el RCEP envía el mensaje de error a la CPU, de tal modo que la CPU lleva a cabo tratamiento de errores sobre el segundo dispositivo de punto extremo. Determinar si el mensaje de error es un mensaje de error enviado repetidamente, incluye específicamente:

30 obtener el identificador BDF o las segundas direcciones de memoria que son del segundo dispositivo de punto extremo y se transportan en el mensaje de error; consultar una correspondencia que existe entre las segundas direcciones de memoria del segundo punto extremo y el estado del segundo dispositivo de punto extremo y está registrada en el registro del estado del dispositivo, y determinar el estado del segundo dispositivo de punto extremo; o consultar la correspondencia que existe entre el identificador BDF del segundo dispositivo de punto extremo y el estado del segundo dispositivo de punto extremo y está registrada en el registro del estado del dispositivo, y determinar el estado del segundo dispositivo de punto extremo; o determinar las primeras direcciones de memoria del segundo dispositivo de punto extremo de acuerdo con las segundas direcciones de memoria o el identificador BDF que es del segundo dispositivo de punto extremo y está en la solicitud de acceso, y determinar el estado del segundo dispositivo de punto extremo consultando la correspondencia que existe entre las primeras direcciones de memoria del segundo dispositivo de punto extremo y el estado del segundo dispositivo de punto extremo y está registrada en el registro del estado del dispositivo; si se determina que el estado del segundo dispositivo de punto extremo es el estado de fallo, determinar que el mensaje de error es un mensaje de error enviado repetidamente, y si se determina que el estado del segundo dispositivo de punto extremo es el estado sin fallo, determinar que el mensaje de error no es un mensaje de error enviado repetidamente.

45 En esta realización de la presente invención, la unidad de monitorización 801 monitoriza el estado del segundo dispositivo de punto extremo en el dominio extendido; la unidad de registro 802 establece el registro del estado del dispositivo de acuerdo con el estado que es del segundo dispositivo de punto extremo y está determinado por la unidad de monitorización 801, donde el registro del estado del dispositivo incluye la correspondencia entre las primeras direcciones de memoria del segundo dispositivo de punto extremo y el estado del segundo dispositivo de punto extremo; después de que la unidad de recepción 803 recibe la solicitud de acceso entre el segundo dispositivo de punto extremo y el dominio principal, la unidad de determinación 804 obtiene las primeras direcciones de memoria que son del segundo dispositivo de punto extremo y están en la solicitud de acceso, u obtiene las primeras direcciones de memoria del segundo dispositivo de punto extremo de acuerdo con el identificador BDF o las segundas direcciones de memoria que son del segundo dispositivo de punto extremo y están en la solicitud de acceso, consulta la correspondencia que existe entre las primeras direcciones de memoria del segundo dispositivo de punto extremo y el estado del segundo dispositivo de punto extremo y está registrada en el registro del estado del dispositivo, y determina el estado del segundo dispositivo de punto extremo; si la unidad de determinación 804 determina que el estado del segundo dispositivo de punto extremo es el estado de fallo, la unidad de procesamiento 805 desecha la solicitud de acceso para impedir la comunicación entre el segundo dispositivo de punto extremo en fallo y el dominio principal. Además, la unidad de procesamiento 805 puede enviar asimismo el mensaje de aislamiento de fallo a la CPU para ordenar a la CPU interrumpir el acceso al segundo dispositivo de punto extremo en el dominio extendido, evitando de ese modo la propagación de un fallo al dominio principal.

Además, en esta realización de la presente invención, la unidad de registro 802 puede registrar además, en el registro del estado del dispositivo, la correspondencia entre el identificador BDF o las segundas direcciones de memoria del segundo dispositivo de punto extremo y el estado del segundo dispositivo de punto extremo, de tal modo que la unidad de determinación 804 puede determinar el estado del segundo dispositivo de punto extremo consultando directamente el registro del estado del dispositivo de acuerdo con el identificador BDF, las segundas direcciones de memoria o las primeras direcciones de memoria que son del segundo dispositivo de punto extremo y están en la solicitud de acceso, lo que evita transformar las segundas direcciones de memoria o el identificador BDF del segundo dispositivo de punto extremo en las primeras direcciones de memoria del segundo dispositivo de punto extremo, acelerando de ese modo el proceso de determinación del estado del segundo dispositivo de punto extremo.

Además, en esta realización de la presente invención, cuando la unidad de monitorización 801 monitoriza el estado del segundo dispositivo de punto extremo, después de recibir el mensaje de error enviado por el segundo dispositivo de punto extremo y determinar que el estado del segundo dispositivo de punto extremo es el estado de fallo de acuerdo con el tipo de mensaje de error, la unidad de monitorización 801 puede determinar además si el mensaje de error enviado por el segundo dispositivo de punto extremo es un mensaje de error enviado repetidamente; si el mensaje de error es un mensaje de error enviado repetidamente, el mensaje de error se desecha para impedir que el mensaje de error llegue al dominio principal y evitar que la CPU lleve a cabo de manera innecesaria y repetida el procesamiento del mensaje de error, asegurando de ese modo la fiabilidad del sistema.

La figura 9 es un sistema de aislamiento de fallo 900 de acuerdo con una realización de la presente invención, donde el sistema 900 incluye un dominio principal PCIe 910 y un dominio extendido PCIe 920, donde el dominio principal incluye un complejo raíz 911, un primer dispositivo de punto extremo 912 y un dispositivo de punto extremo del complejo raíz 913, y el dominio extendido 920 incluye el dispositivo de punto extremo del complejo raíz 913 y un segundo dispositivo de punto extremo 921. El dispositivo de punto extremo del complejo raíz 913 está configurado para: monitorizar un estado del segundo dispositivo de punto extremo 921 en el dominio extendido; establecer un registro del estado del dispositivo de acuerdo con el estado del segundo dispositivo de punto extremo 921, donde el registro del estado del dispositivo incluye una correspondencia entre información de identificador del segundo dispositivo de punto extremo 921 y el estado del segundo dispositivo de punto extremo 921; recibir una solicitud de acceso, donde la solicitud de acceso incluye una solicitud de acceso desde el segundo dispositivo de punto extremo 921 al dominio principal 910 o una solicitud de acceso desde el dominio principal 910 al segundo dispositivo de punto extremo 921; consultar el registro del estado del dispositivo de acuerdo con la información de identificador que es del segundo dispositivo de punto extremo 921 y está en la solicitud de acceso, y determinar el estado del segundo dispositivo de punto extremo 921; y si el estado del segundo dispositivo de punto extremo 921 es un estado de fallo, desechar la solicitud de acceso para impedir la comunicación entre el segundo dispositivo de punto extremo 921 y el dominio principal 910.

En esta realización de la presente invención, se monitoriza el estado del segundo dispositivo de punto extremo en el dominio extendido, y se establece el registro del estado del dispositivo de acuerdo con el estado del segundo dispositivo de punto extremo; después de que se recibe la solicitud de acceso entre el segundo dispositivo de punto extremo y el dominio principal, se consulta el registro del estado del dispositivo de acuerdo con la información de identificador que es del segundo dispositivo de punto extremo y está en la solicitud de acceso, y se determina el estado del segundo dispositivo de punto extremo; si el estado del segundo dispositivo de punto extremo es el estado de fallo, se desecha la solicitud de acceso para impedir la comunicación entre el segundo dispositivo de punto extremo en fallo y el dominio principal, y evitar la propagación de un fallo al dominio principal, asegurando de ese modo la fiabilidad del sistema.

La figura 10 es un diagrama estructural esquemático de un aparato de aislamiento de fallo de acuerdo con una realización de la presente invención. El aparato de aislamiento de fallo de acuerdo con esta realización de la presente invención se utiliza en un sistema informático interconectado por PCIe, donde el sistema informático incluye un dominio principal y un dominio extendido, el dominio principal incluye un complejo raíz, un primer dispositivo de punto extremo y un dispositivo de punto extremo del complejo raíz, y el dominio extendido incluye el dispositivo de punto extremo del complejo raíz y un segundo dispositivo de punto extremo.

El aparato de aislamiento de fallo puede incluir:

un procesador 1001, una memoria 1002, un bus del sistema 1004 y una interfaz de comunicación el 1005, donde el procesador 1001, la memoria 1002 y la interfaz de comunicaciones 1005 están conectados mediante el bus del sistema 1004 e implementan comunicación mutua utilizándolo.

El procesador 1001 puede ser una unidad central de procesamiento de un solo núcleo o de múltiples núcleos, o un circuito integrado de aplicación específica, o puede estar configurado como uno o varios circuitos integrados que implementan esta realización de la presente invención.

La memoria 1002 puede ser una memoria RAM de alta velocidad o una memoria no volátil (memoria no volátil), por ejemplo, por lo menos una memoria de disco magnético.

La memoria 1002 se utiliza para una instrucción 1003 ejecutada por ordenador. Específicamente, la instrucción 1003 ejecutada por ordenador puede incluir código de programa.

Cuando el aparato de aislamiento de fallo está funcionando, el procesador 1001 ejecuta la instrucción 1003 ejecutada por ordenador, y puede llevar a cabo el siguiente procedimiento:

monitorizar el estado del segundo dispositivo de punto extremo en el dominio extendido;

5 establecer un registro del estado del dispositivo de acuerdo con el estado del segundo dispositivo de punto extremo, donde el registro del estado del dispositivo incluye una correspondencia entre una información de identificador del segundo dispositivo de punto extremo y el estado del segundo dispositivo de punto extremo;

recibir una solicitud de acceso, donde la solicitud de acceso incluye una solicitud de acceso desde el segundo dispositivo de punto extremo al dominio principal o una solicitud de acceso desde el dominio principal al segundo dispositivo de punto extremo;

10 consultar el registro del estado del dispositivo de acuerdo con la información de identificador que es del segundo dispositivo de punto extremo y está en la solicitud de acceso, y determinar el estado del segundo dispositivo de punto extremo; y

si el estado del segundo dispositivo de punto extremo es un estado de fallo, desechar la solicitud de acceso para impedir la comunicación entre el segundo dispositivo de punto extremo y el dominio principal.

15 El procedimiento puede además incluir específicamente:

monitorizar el estado del segundo dispositivo de punto extremo en el dominio extendido;

establecer un registro del estado del dispositivo de acuerdo con el estado del segundo dispositivo de punto extremo, donde el registro del estado del dispositivo incluye una correspondencia entre primeras direcciones de memoria del segundo dispositivo de punto extremo y el estado del segundo dispositivo de punto extremo;

20 recibir una solicitud de acceso, donde la solicitud de acceso incluye una solicitud de acceso desde el segundo dispositivo de punto extremo al dominio principal o una solicitud de acceso desde el dominio principal al segundo dispositivo de punto extremo;

25 consultar el registro del estado del dispositivo de acuerdo con la información de identificador que es del segundo dispositivo de punto extremo y está en la solicitud de acceso, y determinar el estado del segundo dispositivo de punto extremo; y

si el estado del segundo dispositivo de punto extremo es el estado de fallo, desechar la solicitud de acceso para impedir la comunicación entre el segundo dispositivo de punto extremo y el dominio principal;

30 si el estado del segundo dispositivo de punto extremo es el estado de fallo, enviar un mensaje de aislamiento de fallo a una CPU, donde el mensaje de aislamiento de fallo se utiliza para ordenar a la CPU en el dominio principal que interrumpa el acceso al segundo dispositivo de punto extremo en el dominio extendido, y el mensaje de aislamiento de fallo lleva las primeras direcciones de memoria del segundo dispositivo de punto extremo; o

si la solicitud de acceso es una solicitud de acceso enviada por el dominio principal, devolver al dominio principal un paquete de respuesta analógica en respuesta a la solicitud de acceso.

35 Un experto en la materia deberá comprender que cada aspecto de la presente invención o un posible modo de implementación de cada aspecto se puede implementar específicamente como un sistema, un procedimiento o un producto de programa informático. Por lo tanto, cada aspecto de la presente invención o el posible modo de implementación de cada aspecto puede utilizar una forma de una realización íntegra en hardware, una realización íntegra en software (incluyendo software inalterable, software residente y similares) o una realización que combina software y hardware, que se denomina de manera uniforme "circuito", "módulo" o "sistema". Además, cada aspecto
40 de la presente invención o el posible modo de implementación de cada aspecto puede utilizar una forma de producto de programa informático, donde el producto de programa informático se refiere a un código de programa legible por ordenador almacenado en un medio legible por ordenador.

45 El medio legible por ordenador puede ser un medio de señal legible por ordenador o un medio de almacenamiento legible por ordenador. El medio de almacenamiento legible por ordenador incluye, de forma no limitativa, un sistema, dispositivo o aparato electrónico, magnético, óptico, electromagnético, de infrarrojos o semiconductor, o cualquier combinación adecuada de los mismos, tal como una memoria de acceso aleatorio (RAM), una memoria de sólo lectura (ROM), una memoria de sólo lectura programable regrabable (EPROM o memoria flash), una fibra óptica, una memoria de sólo lectura portátil (CD-ROM).

50 Un procesador en un ordenador lee código de programa legible por ordenador almacenado en un medio legible por ordenador, de tal modo que el procesador puede ejecutar una función y una acción especificadas en cada etapa o en una combinación de etapas en un diagrama de flujo, y se genera un aparato para implementar una función y una acción especificada en cada bloque o en una combinación de bloques en un diagrama de bloques.

5 El código de programa legible por ordenador puede ser ejecutado completamente en un ordenador de un usuario, ejecutado parcialmente en el ordenador del usuario, utilizado como un paquete de software independiente, ejecutado parcialmente en el ordenador del usuario y ejecutado parcialmente en un ordenador remoto, o ejecutado completamente en el ordenador remoto o en un servidor. Se debe observar asimismo que en algunos esquemas de implementación alternativos, las funciones indicadas en cada etapa de los diagramas de flujo o en cada bloque de los diagramas de bloque pueden no producirse en una secuencia mostrada en los diagramas. Por ejemplo, dependiendo de la función involucrada, dos etapas o dos bloques marcados sucesivamente pueden llevarse a cabo de manera aproximadamente simultánea, o estos bloques pueden en ocasiones realizarse en orden inverso.

10 Un experto en la materia puede ser consciente de que, en combinación con los ejemplos descritos en las realizaciones dadas a conocer en esta memoria descriptiva, las unidades y etapas de algoritmo se pueden implementar mediante hardware electrónico o una combinación de software informático y hardware electrónico. Que las funciones se lleven a cabo mediante hardware o software depende de las aplicaciones particulares y de las condiciones de los límites de diseño de las soluciones técnicas. Un experto en la materia puede utilizar diferentes procedimientos para implementar las funciones descritas para cada aplicación particular, pero no se deberá
15 considerar que esta implementación rebasa el alcance de la presente invención.

Las descripciones anteriores son tan solo modos de implementación específicos de la presente invención, pero no están destinadas a limitar el alcance de protección de la presente invención. Cualquier variación o sustitución que se le ocurra fácilmente a un experto en la materia dentro del alcance técnico dado a conocer en la presente invención quedará dentro del alcance de protección de la presente invención. Por lo tanto, el alcance de protección de la
20 presente invención estará sujeto al alcance de protección de las reivindicaciones.

REIVINDICACIONES

1. Un procedimiento de aislamiento de fallo, utilizado en un sistema informático interconectado por interconexión de componentes periféricos exprés, PCIe, donde el sistema informático comprende un dominio principal y un dominio extendido, el dominio principal comprende un complejo raíz, un primer dispositivo de punto extremo y un dispositivo de punto extremo del complejo raíz, y el dominio extendido comprende el dispositivo de punto extremo del complejo raíz y un segundo dispositivo de punto extremo, donde el dispositivo de punto extremo del complejo raíz sirve como el complejo raíz en el dominio extendido; y el procedimiento comprende:
- 5 monitorizar un estado del segundo dispositivo de punto extremo en el dominio extendido;
- establecer un registro del estado del dispositivo de acuerdo con el estado del segundo dispositivo de punto extremo,
- 10 donde el registro del estado del dispositivo comprende una correspondencia entre información de identificador del segundo dispositivo de punto extremo y el estado del segundo dispositivo de punto extremo;
- recibir una solicitud de acceso, donde la solicitud de acceso comprende una solicitud de acceso desde el segundo dispositivo de punto extremo del dominio extendido hasta el dominio principal o una solicitud de acceso desde el dominio principal hasta el segundo dispositivo de punto extremo del dominio extendido;
- 15 consultar el registro del estado del dispositivo de acuerdo con la información de identificador que es del segundo dispositivo de punto extremo y está en la solicitud de acceso, y determinar el estado del segundo dispositivo de punto extremo; y
- desechar la solicitud de acceso para impedir la comunicación entre el segundo dispositivo de punto extremo del dominio extendido y el dominio principal cuando el estado del segundo dispositivo de punto extremo es un estado de fallo;
- 20 donde el procedimiento comprende además:
- configurar una primera dirección de memoria y una segunda dirección de memoria para el segundo dispositivo de punto extremo, la primera dirección de memoria se utiliza para representar el segundo dispositivo de punto extremo en el dominio principal, la segunda dirección de memoria se utiliza para representar el segundo dispositivo de punto extremo en el dominio extendido;
- 25 configurar un identificador de bus/dispositivo/función BDF para el segundo dispositivo de punto extremo;
- almacenar una relación de mapeo entre el identificador BDF del segundo dispositivo de punto extremo y la segunda dirección de memoria del segundo dispositivo de punto extremo; y
- 30 almacenar una relación de mapeo entre la primera dirección de memoria del segundo dispositivo de punto extremo y la segunda dirección de memoria del segundo dispositivo de punto extremo.
2. El procedimiento según la reivindicación 1, en el que la monitorización de un estado del segundo dispositivo de punto extremo en el dominio extendido y el establecimiento de un registro del estado del dispositivo de acuerdo con el estado del segundo dispositivo de punto extremo, comprende:
- 35 recibir un mensaje de error enviado por el segundo dispositivo de punto extremo, o recibir un mensaje de respuesta de detección del dispositivo, utilizado para indicar si existe el segundo dispositivo de punto extremo;
- determinar el estado del segundo dispositivo de punto extremo de acuerdo con el mensaje de error o el mensaje de respuesta de detección del dispositivo;
- obtener el identificador BDF del segundo dispositivo de punto extremo y transportado en el mensaje de error o el mensaje de respuesta de detección del dispositivo; y
- 40 registrar, en el registro del estado del dispositivo, una correspondencia entre el identificador BDF del segundo dispositivo de punto extremo y el estado del segundo dispositivo de punto extremo.
3. El procedimiento según la reivindicación 2, comprendiendo además el establecimiento:
- obtener la primera dirección de memoria del segundo dispositivo de punto extremo de acuerdo con el identificador BDF; y
- 45 registrar, en el registro del estado del dispositivo, una correspondencia entre la primera dirección de memoria del segundo dispositivo de punto extremo y el estado del segundo dispositivo de punto extremo.
4. El procedimiento según la reivindicación 3, que comprende además: registrar, en el registro del estado del dispositivo, una correspondencia entre la segunda dirección de memoria del segundo dispositivo de punto extremo y el estado del segundo dispositivo de punto extremo.

5. El procedimiento según la reivindicación 3, en el que obtener la primera dirección de memoria del segundo dispositivo de punto extremo de acuerdo con el identificador BDF comprende:

5 obtener la segunda dirección de memoria del segundo dispositivo de punto extremo de acuerdo con la relación de mapeo entre el identificador BDF del segundo dispositivo de punto extremo y la segunda dirección de memoria del segundo dispositivo de punto extremo, y transformar la segunda dirección de memoria en la primera dirección de memoria del segundo dispositivo de punto extremo de acuerdo con la relación de mapeo entre la segunda dirección de memoria del segundo dispositivo de punto extremo y la primera dirección de memoria del segundo dispositivo de punto extremo.

10 6. El procedimiento según las reivindicaciones 2 ó 5, en el que la primera dirección de memoria del segundo dispositivo de punto extremo comprende una primera dirección de memoria de acceso al espacio de configuración, una primera dirección de memoria de acceso a interrupciones señalizadas por mensaje, una primera dirección de memoria de acceso de entrada/salida con mapeo a memoria y una primera dirección de memoria de acceso DMA; y registrar, en el registro del estado del dispositivo, una correspondencia entre la primera dirección de memoria del segundo dispositivo de punto extremo y el estado del segundo dispositivo de punto extremo, comprende:

15 registrar, en el registro del estado del dispositivo, una correspondencia entre cada primera dirección de memoria del segundo dispositivo de punto extremo y el estado del segundo dispositivo de punto extremo.

7. El procedimiento según cualquiera de las reivindicaciones 2 a 6, en el que el procedimiento comprende además:

20 enviar un mensaje de aislamiento de fallo a una CPU, donde el mensaje de aislamiento de fallo se utiliza para ordenar a la CPU en el dominio principal que interrumpa el acceso al segundo dispositivo de punto extremo en el dominio extendido, y el mensaje de aislamiento de fallo lleva la primera dirección de memoria del segundo dispositivo de punto extremo.

8. El procedimiento según cualquiera de las reivindicaciones 2 a 6, en el que si la solicitud de acceso es una solicitud de acceso desde el dominio principal al segundo dispositivo de punto extremo, el procedimiento comprende además:

devolver al dominio principal un paquete de respuesta analógica en respuesta a la solicitud de acceso.

25 9. El procedimiento según la reivindicación 2, en el que después de recibir un mensaje de error enviado por el segundo dispositivo de punto extremo, el procedimiento comprende además:

determinar si el mensaje de error es un mensaje de error enviado repetidamente; y si el mensaje de error es un mensaje de error enviado repetidamente, desechar el mensaje de error.

30 10. Un aparato de aislamiento de fallo, utilizado en un sistema informático interconectado por interconexión de componentes periféricos exprés PCIe, donde el sistema informático comprende un dominio principal y un dominio extendido, el dominio principal comprende un complejo raíz, un primer dispositivo de punto extremo y un dispositivo de punto extremo del complejo raíz, y el dominio extendido comprende el dispositivo de punto extremo del complejo raíz y un segundo dispositivo de punto extremo, donde el dispositivo de punto extremo del complejo raíz sirve como el complejo raíz en el dominio extendido; y el aparato comprende:

35 una unidad de monitorización, configurada para monitorizar un estado del segundo dispositivo de punto extremo en el dominio extendido;

40 una unidad de registro, configurada para establecer un registro del estado del dispositivo de acuerdo con el estado del segundo dispositivo de punto extremo, donde el registro del estado del dispositivo comprende una correspondencia entre una información de identificador del segundo dispositivo de punto extremo y el estado del segundo dispositivo de punto extremo;

una unidad de recepción, configurada para recibir una solicitud de acceso, donde la solicitud de acceso comprende una solicitud de acceso desde el segundo dispositivo de punto extremo al dominio principal o una solicitud de acceso desde el dominio principal al segundo dispositivo de punto extremo;

45 una unidad de determinación, configurada para consultar el registro del estado del dispositivo de acuerdo con la información de identificador que es del segundo dispositivo de punto extremo y está en la solicitud de acceso, y determinar el estado del segundo dispositivo de punto extremo; y

una unidad de procesamiento, configurada para: cuando el estado del segundo dispositivo de punto extremo es un estado de fallo, desechar la solicitud de acceso para impedir la comunicación entre el segundo dispositivo de punto extremo y el dominio principal;

50 donde la unidad de determinación está configurada para configurar una primera dirección de memoria y una segunda dirección de memoria para el segundo dispositivo de punto extremo, la primera dirección de memoria se utiliza para representar el segundo dispositivo de punto extremo en el dominio principal, la segunda dirección de memoria se utiliza para representar el segundo dispositivo de punto extremo en el dominio extendido, configurar un

- 5 identificador de bus/dispositivo/función BDF para el segundo dispositivo de punto extremo, almacenar una relación de mapeo entre el identificador BDF del segundo dispositivo de punto extremo y la segunda dirección de memoria del segundo dispositivo de punto extremo, y almacenar una relación de mapeo entre la primera dirección de memoria del segundo dispositivo de punto extremo y la segunda dirección de memoria del segundo dispositivo de punto extremo.
- 10 11. El aparato según la reivindicación 10, en el que la unidad de monitorización está configurada específicamente para:
- 15 recibir un mensaje de error enviado por el segundo dispositivo de punto extremo, o recibir un mensaje de respuesta de detección del dispositivo utilizado para indicar si existe el segundo dispositivo de punto extremo, y determinar el estado del segundo dispositivo de punto extremo de acuerdo con el mensaje de error o el mensaje de respuesta de detección del dispositivo;
- 20 la unidad de registro está configurada específicamente para: obtener el identificador BDF del segundo dispositivo de punto extremo y transportado en el mensaje de error o el mensaje de respuesta de detección del dispositivo, y registrar, en el registro del estado del dispositivo, una correspondencia entre el identificador BDF del segundo dispositivo de punto extremo y el estado del segundo dispositivo de punto extremo.
- 25 12. El aparato según la reivindicación 11, en el que la unidad de registro está configurada además para:
- 30 obtener la primera dirección de memoria del segundo dispositivo de punto extremo de acuerdo con el identificador BDF; y
- 35 registrar, en el registro del estado del dispositivo, una correspondencia entre la primera dirección de memoria del segundo dispositivo de punto extremo y el estado del segundo dispositivo de punto extremo.
- 40 13. Un sistema de aislamiento de fallo, en el que el sistema comprende un dominio principal de interconexión de componentes periféricos exprés, PCIe, y un dominio extendido de interconexión de componentes periféricos exprés, PCIe, el dominio principal comprende un complejo raíz, un primer dispositivo de punto extremo y un dispositivo de punto extremo del complejo raíz, y el dominio extendido comprende el dispositivo de punto extremo del complejo raíz y un segundo dispositivo de punto extremo; y el dispositivo de punto extremo del complejo raíz está configurado para llevar a cabo el procedimiento según la reivindicación 1.

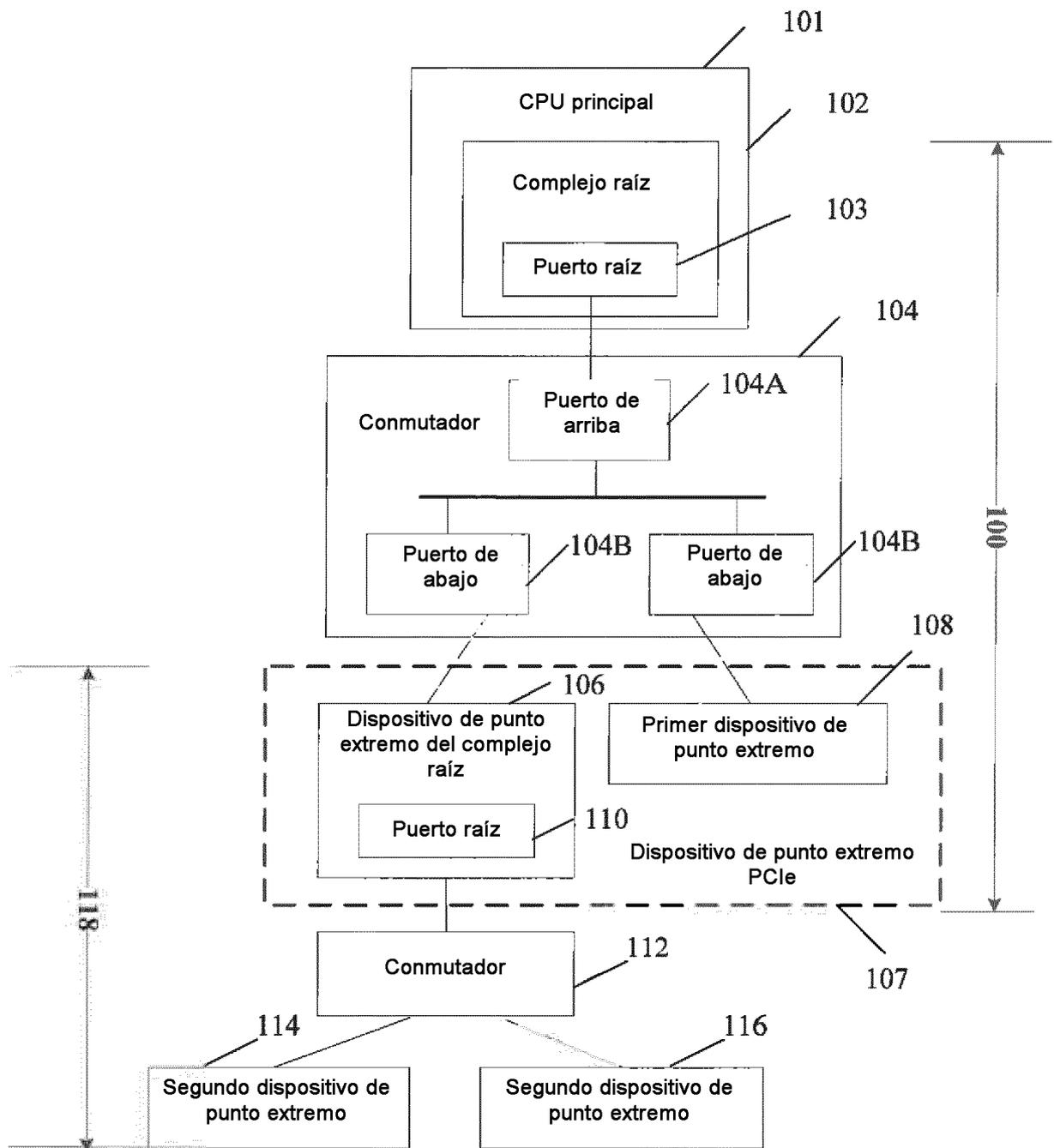


FIG. 1

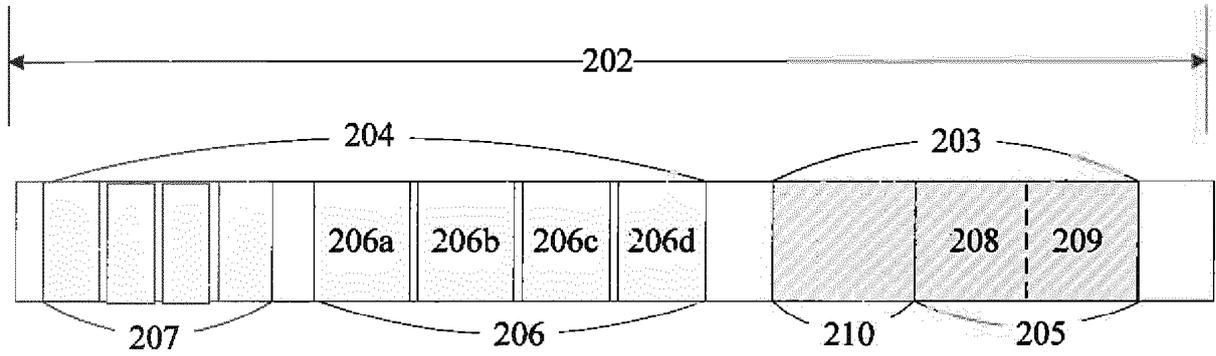


FIG. 2 (a)

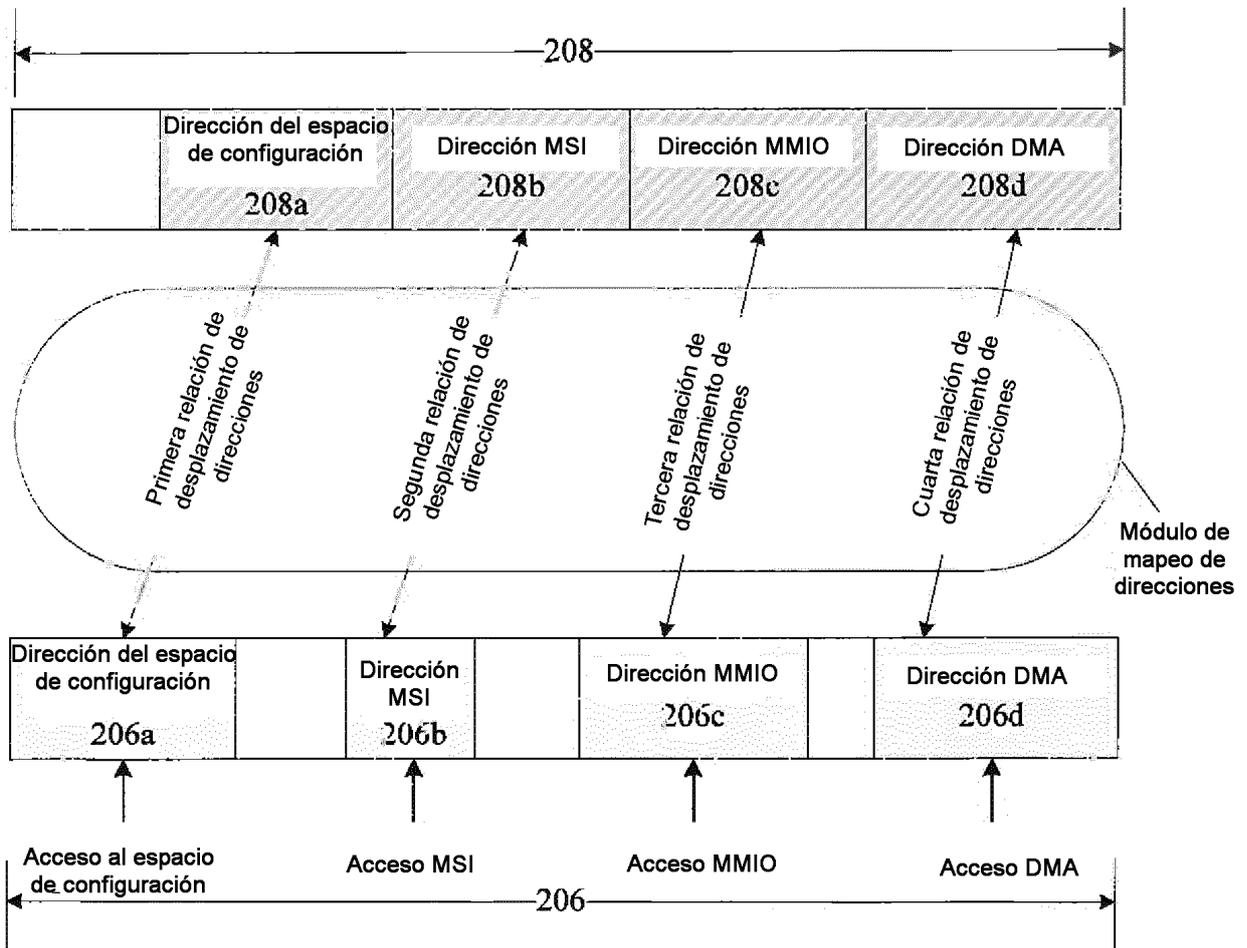


FIG. 2 (b)

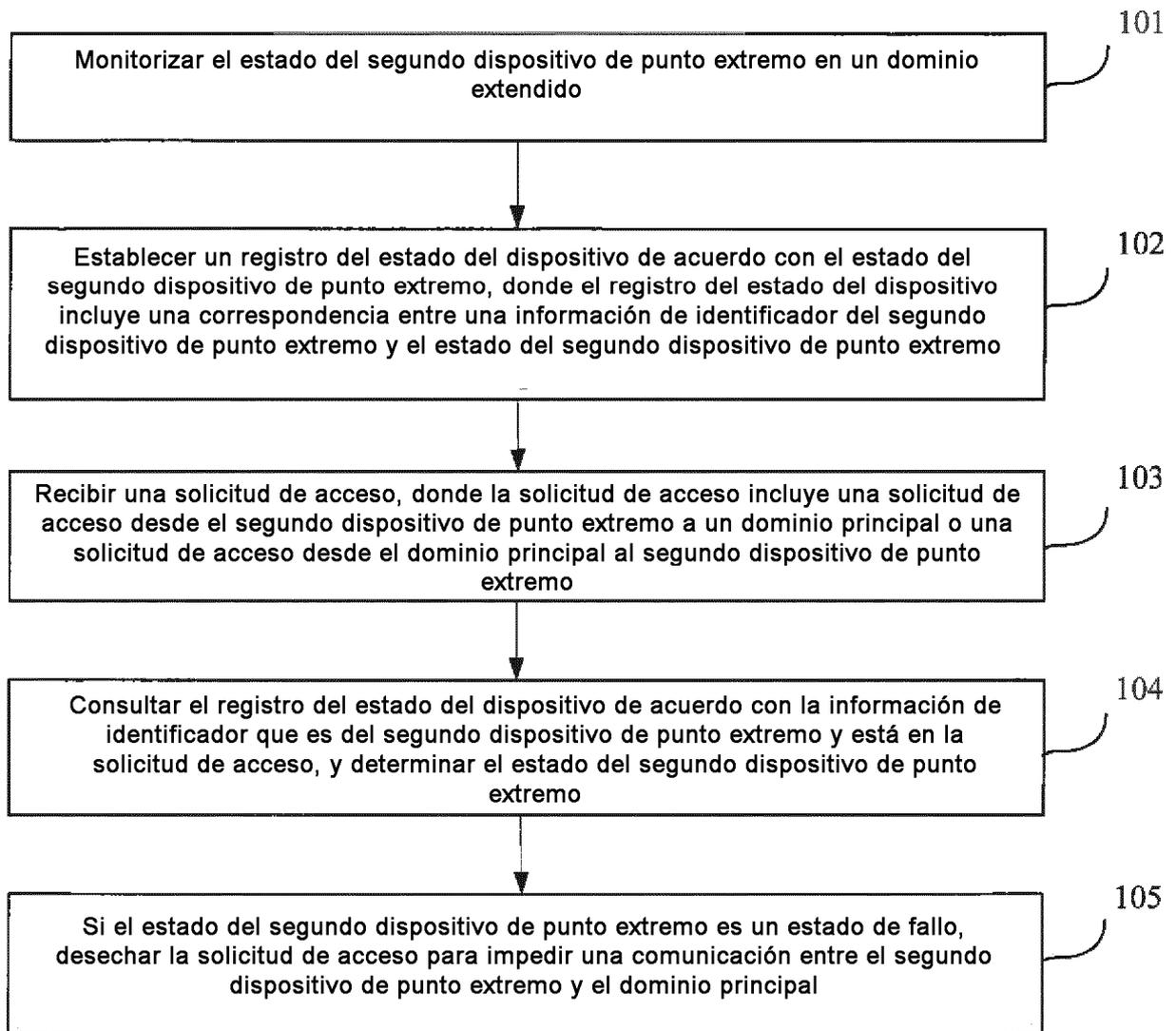


FIG. 3

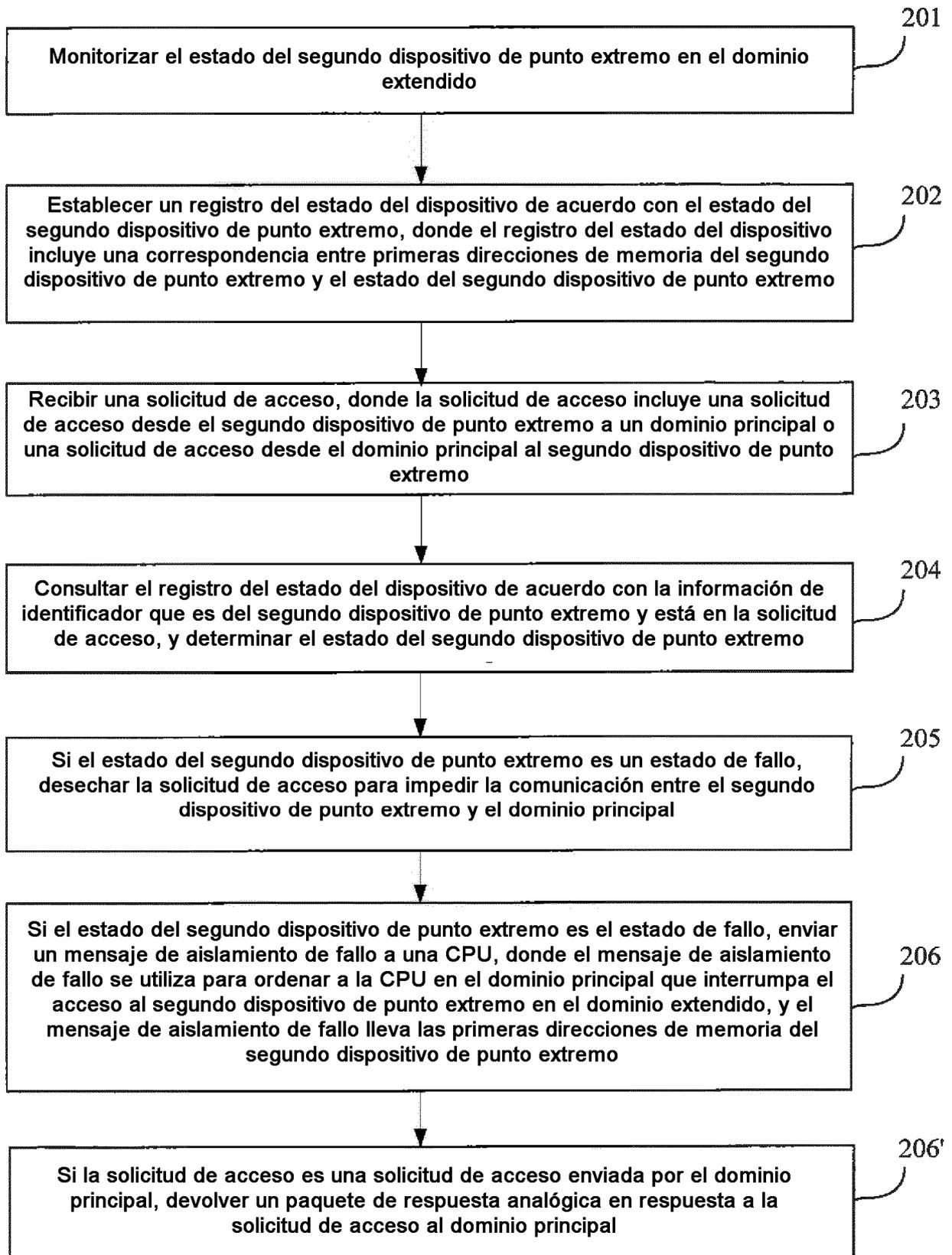


FIG. 4

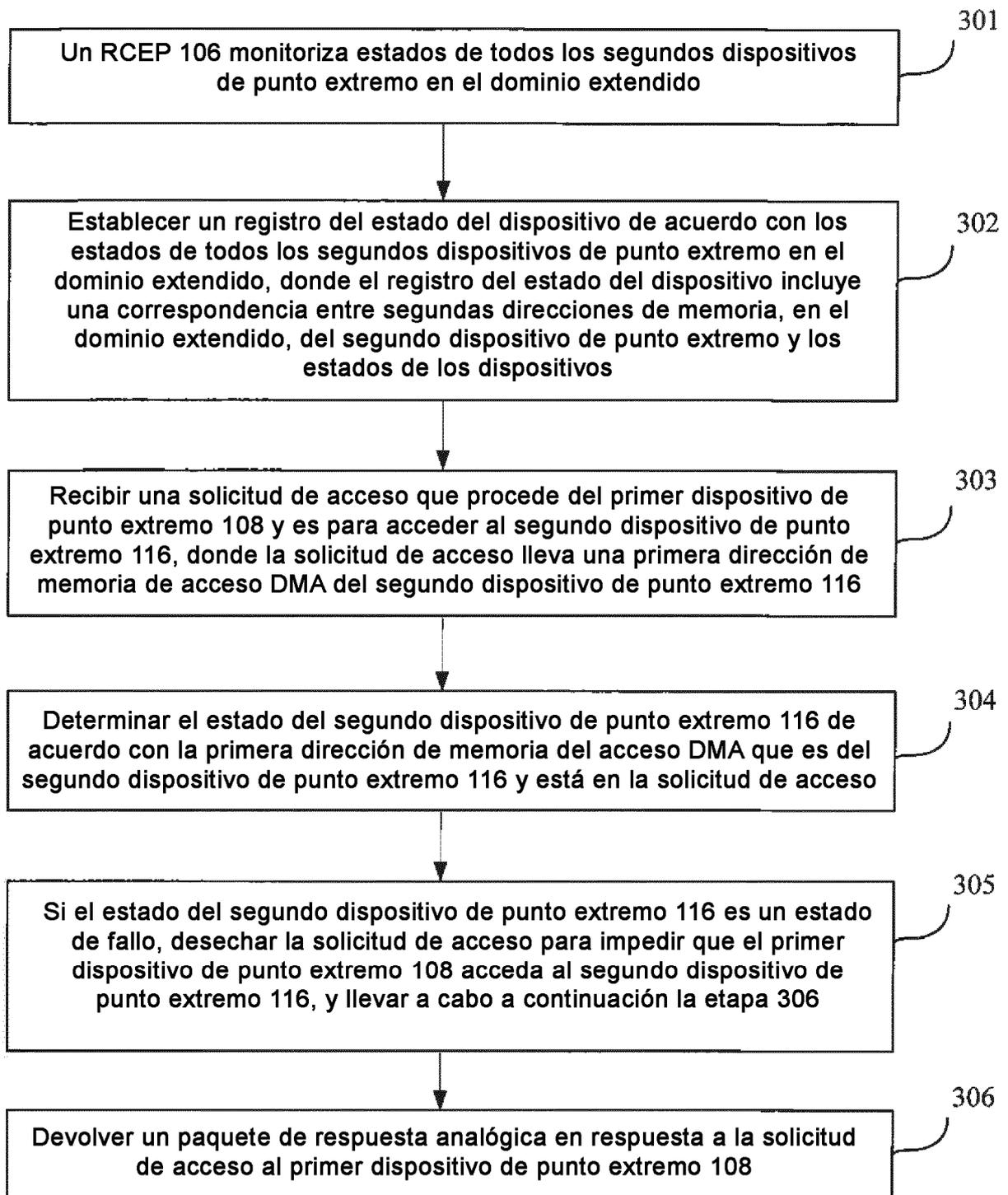


FIG. 5

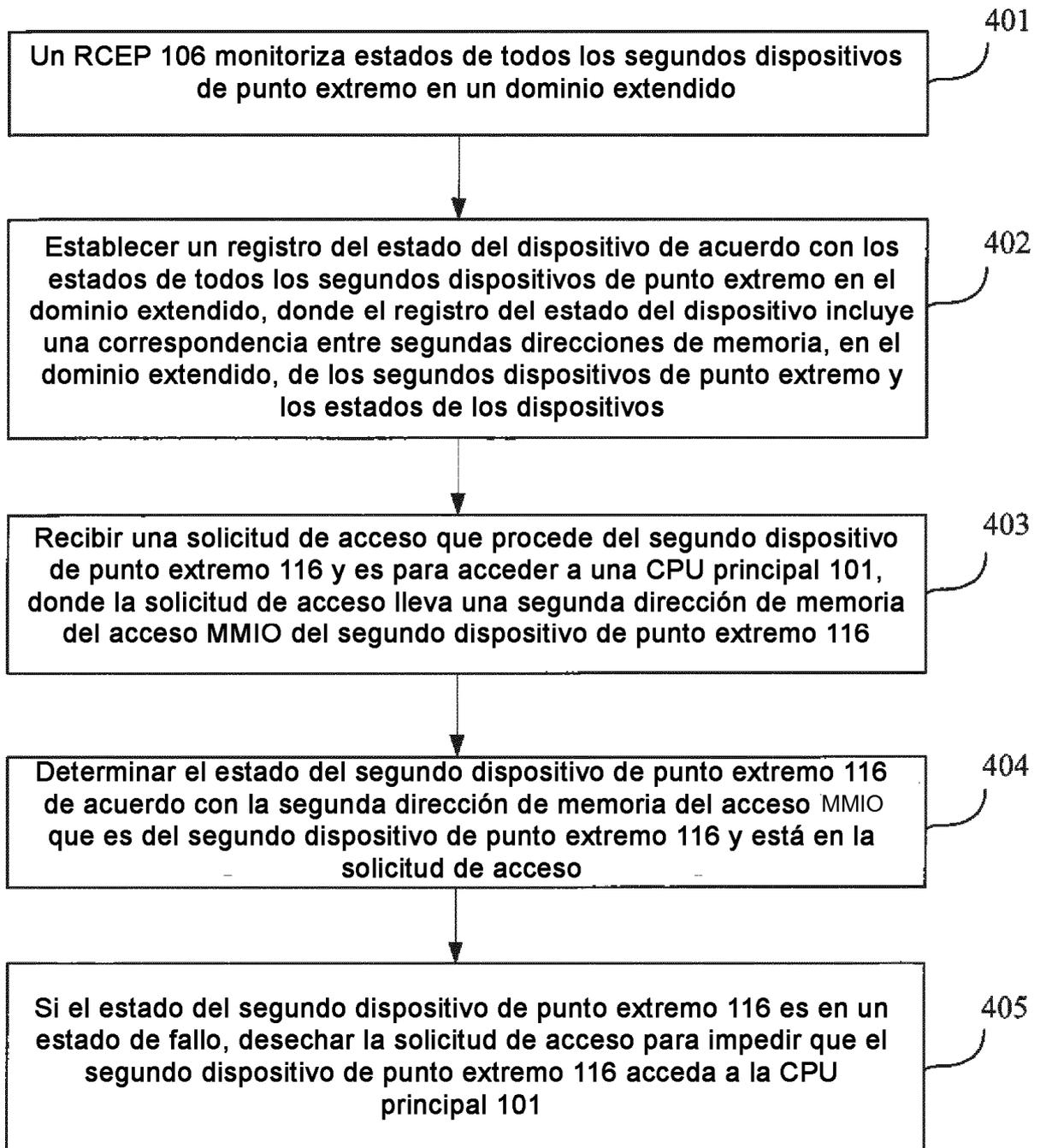


FIG. 6

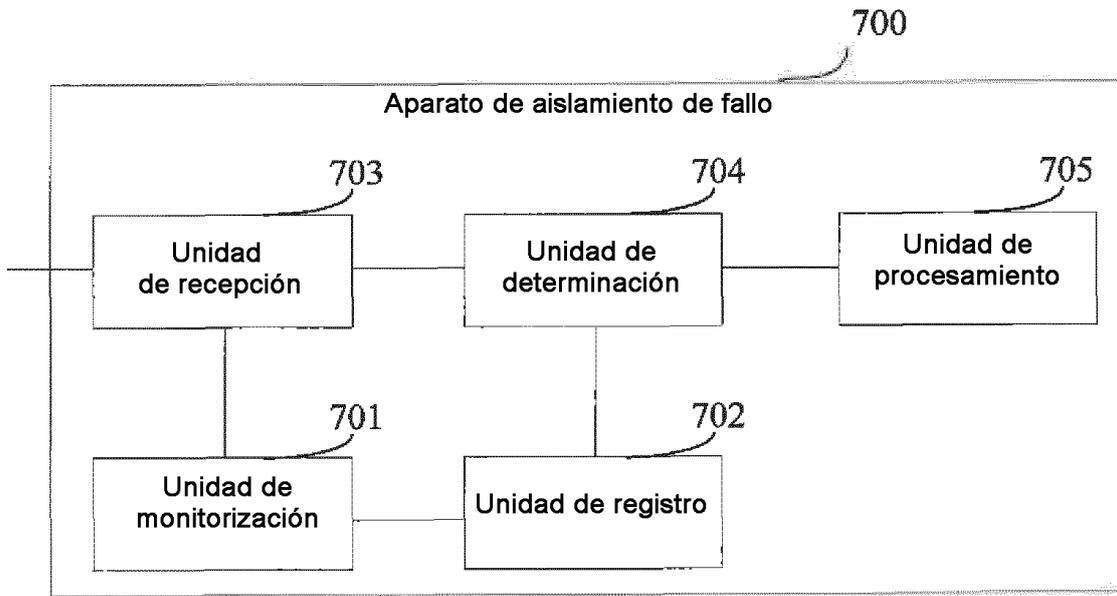


FIG. 7

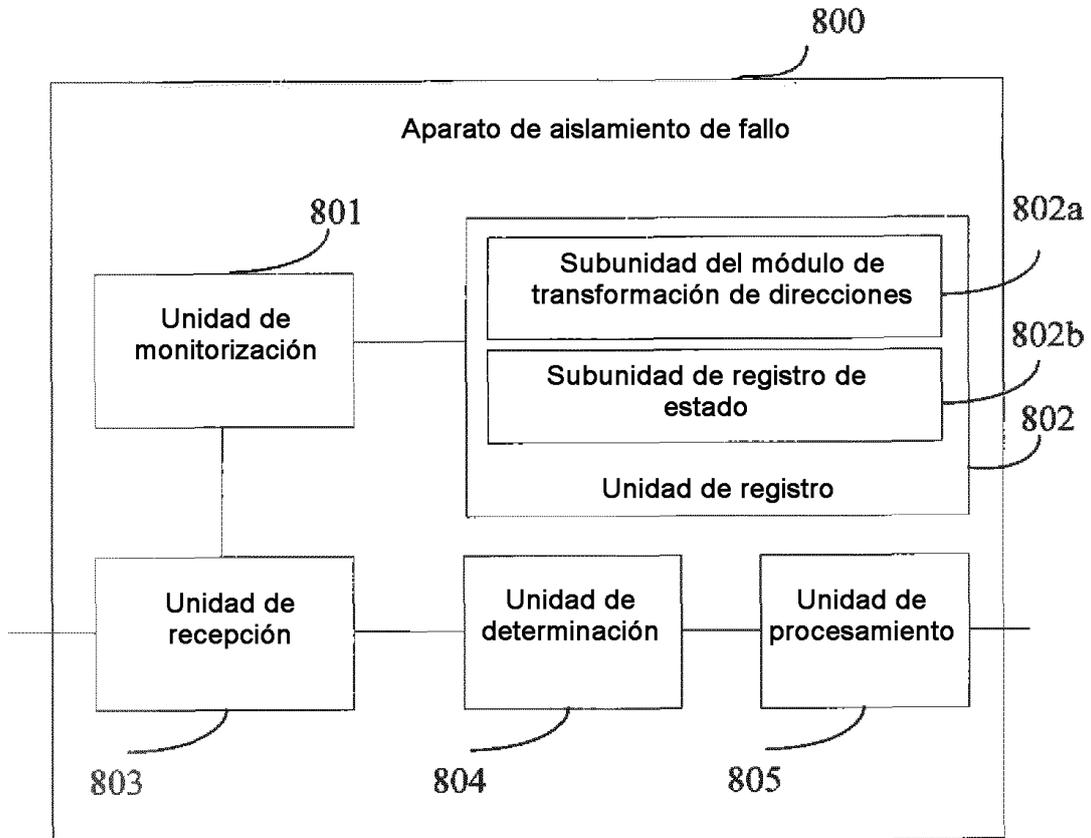


FIG. 8

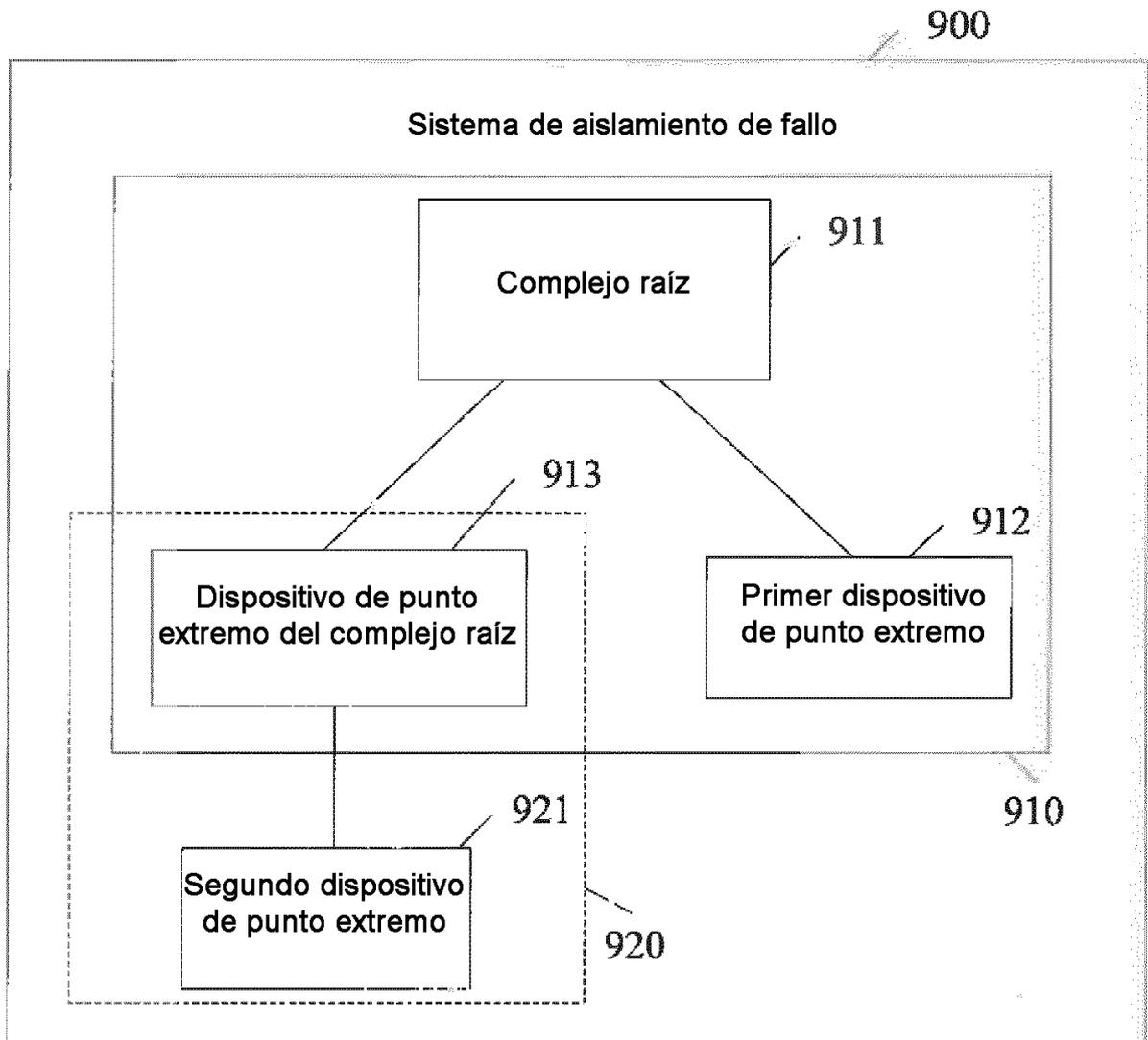


FIG. 9

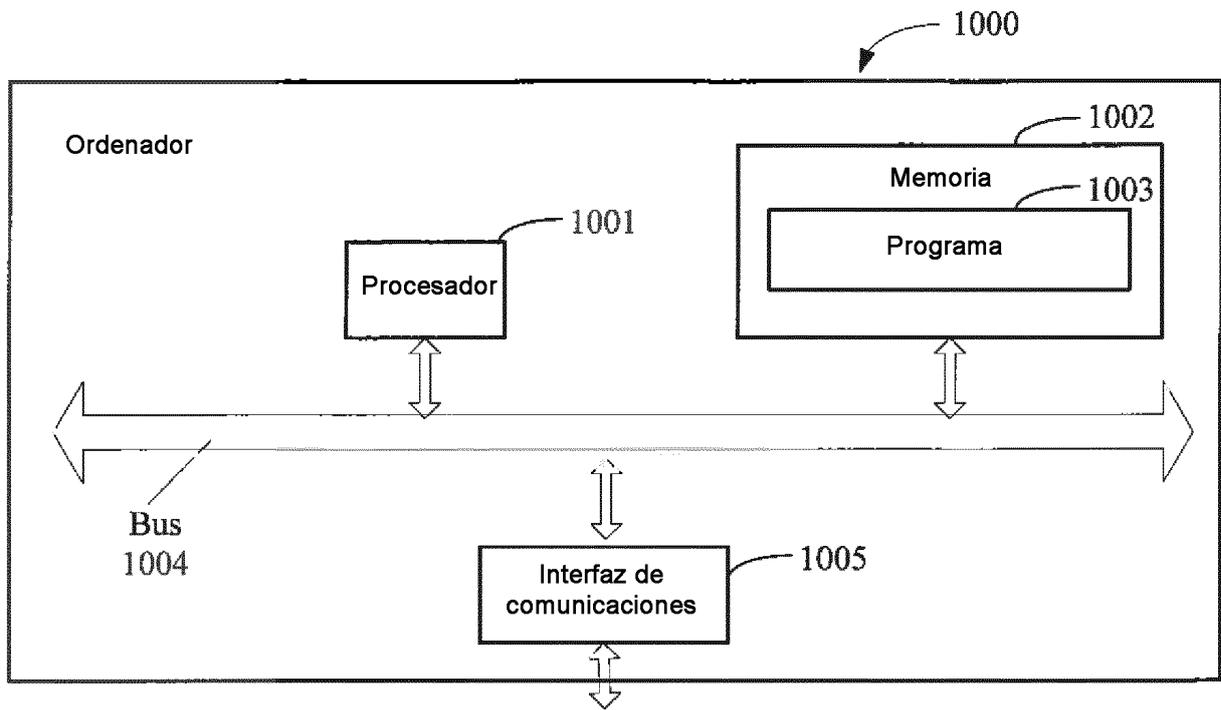


FIG. 10