



19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA

11 Número de publicación: **2 267 850**

51 Int. Cl.:
H04L 12/56 (2006.01)
H04L 29/14 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Número de solicitud europea: **01992133 .7**
86 Fecha de presentación : **06.12.2001**
87 Número de publicación de la solicitud: **1342340**
87 Fecha de publicación de la solicitud: **10.09.2003**

54 Título: **Encaminador y redundancia de protocolos de encaminamiento.**

30 Prioridad: **07.12.2000 US 733284**

45 Fecha de publicación de la mención BOPI:
16.03.2007

45 Fecha de la publicación del folleto de la patente:
16.03.2007

73 Titular/es: **Nokia Inc.**
313 Fairchild Drive, Building A
Mountain View, California 94043, US

72 Inventor/es: **Ho, Chi, Fai;**
Gupta, Amar;
Grandhi, Madhu y
Bachmutsky, Alex

74 Agente: **Curell Suñol, Marcelino**

ES 2 267 850 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

Encaminador y redundancia de protocolos de encaminamiento.

5 Campo de la invención

La presente invención pertenece al campo de la interconexión de redes y dispositivos de interconexión de redes. Más particularmente, la presente invención se refiere a encaminadores y a protocolos de encaminamiento de red. Específicamente, la presente invención se refiere a un encaminador y a la redundancia de protocolo de encaminamiento.

10 Antecedentes de la invención

Una red es un conjunto de dispositivos interconectados que permite a los usuarios acceder a recursos y datos. Los tipos habituales de dispositivos de red incluyen servidores, encaminadores, puentes, interruptores, puertas de enlace y concentradores (“hub”). Una red muy conocida es Internet. Internet es un sistema mundial de redes interconectadas que ejecuta el Protocolo de Internet (IP, “Internet Protocol”) para transferir datos (por ejemplo, paquetes). Dado que un paquete puede llegar a su destino atravesando una serie de límites entre redes en Internet, el IP incluye un servicio de capa “3” que proporciona funciones de encaminamiento y reenvío de manera que el paquete puede llegar a su destino empleando una ruta óptima.

Un dispositivo de red habitual que proporciona el servicio de capa 3 del IP es un encaminador. Un encaminador encamina paquetes determinando una ruta óptima basándose en su visión actual de la red y envía el paquete a través de los límites entre redes hacia un destino usando la ruta óptima. Basándose en su visión de la red, un encaminador genera y mantiene una tabla de encaminamiento de rutas disponibles conocidas por el encaminador. El encaminador emplea la tabla de encaminamiento para crear una tabla de información de reenvío (FIB).

La FIB es una tabla de rutas que emplea el encaminador para reenviar paquetes a su destino.

Un encaminador emplea un protocolo de encaminamiento para intercambiar información con otros encaminadores para mantener una visión consistente de la red (es decir, una FIB consistente). Para reenviar paquetes de manera adecuada, cada encaminador debe presentar una FIB consistente con otros encaminadores en la red. Es decir, los encaminadores que tienen tablas de información de reenvío (FIB) inconsistentes no harán pasar paquetes a través de la red de un modo predecible. Como tales, pueden producirse bucles de encaminamiento o un encaminamiento inapropiado de paquetes.

Por lo tanto, un problema crítico que puede aparecer en la red es un fallo de un encaminador. Un encaminador puede fallar por diversos motivos tales como una mala configuración, ataques de piratas informáticos (“hackers”), fallos de hardware y de software. Estos fallos son impredecibles. Desafortunadamente, un fallo de encaminador puede provocar el cambio en la topología de la red. En particular, la topología puede cambiar debido a que ciertos enlaces o rutas desaparecen. Además, la información del protocolo de encaminamiento puede perderse debido a que ciertos nodos no pueden alcanzarse o cierta información no puede propagarse por la red. Además, los paquetes pueden ser incapaces de llegar a un destino porque determinadas direcciones son inalcanzables.

Un fallo de encaminador puede por tanto provocar una serie de problemas tales como un corte del servicio, un empeoramiento del servicio (encaminamiento no óptimo), y un corte del servicio debido a un tiempo de convergencia largo de las tablas de encaminamiento. Un encaminador defectuoso puede hacer que otros encaminadores reenvíen paquetes usando rutas no óptimas provocando un empeoramiento del servicio ya que los paquetes pueden tardar más en llegar a su destino. Un encaminador defectuoso también hará que sus pares y otros encaminadores en la red que atraviesa estos pares actualicen sus tablas de encaminamiento (“convergencia”) provocando un corte o empeoramiento del servicio para realizar esta convergencia.

Por ejemplo, si un encaminador falla y los protocolos de encaminamiento de nodos pares o encaminadores vecinos observan el fallo, los protocolos de encaminamiento propagarán la noticia del encaminador defectuoso por toda la red para que las tablas de encaminamiento se actualicen de forma correspondiente. En consecuencia, antes de que la red pueda reanudar el servicio completo, se produce un corte o empeoramiento del servicio para actualizar las tablas de encaminamiento en los encaminadores en funcionamiento para que puedan generar FIB consistentes entre sí. Esta reconfiguración de la red puede tardar varios segundos, minutos u horas antes de que se reestablezca toda la red. Para servicios de finalidad crítica, este comportamiento es inaceptable.

Un método para tratar un fallo de encaminador es presentar una redundancia de hardware para aumentar la disponibilidad del sistema. Este tipo de redundancia se designa habitualmente como redundancia de capa 2. Un sistema de redundancia de capa 2 puede incluir tarjetas de línea, puertos o tarjetas de control redundantes. Si una tarjeta de línea, un puerto o una tarjeta de control falla, la tarjeta de línea, el puerto o la tarjeta de control redundante puede reestablecer el funcionamiento. Sin embargo, una desventaja de la redundancia de capa 2 es que no proporciona una redundancia de protocolo de encaminamiento en tiempo real. Por ejemplo, los numerosos estados de software que se generan por los protocolos de encaminamiento en tiempo real no se mantienen en el hardware redundante haciendo que las sesiones de protocolo se caigan. Por lo tanto, en un sistema de redundancia de capa 2, las sesiones de protocolo se caen provocando un cambio en la topología de la red y por tanto un corte del servicio o empeoramiento del servicio.

Otro método para tratar un fallo de encaminador es presentar un encaminador de seguridad. Un esquema de este tipo se designa habitualmente como Protocolo de redundancia de encaminador virtual (VRRP). En un esquema VRRP, si un encaminador par reconoce que un encaminador principal ha fallado iniciará la comunicación con un encaminador de seguridad. Una desventaja del VRRP es que puede tardar mucho (“tiempo de técnico”) en conectarse al encaminador de seguridad. Otra desventaja del VRRP es que las sesiones entre pares del encaminador defectuoso se caen o desconectan y no pueden retomarse por el encaminador de seguridad, provocando por tanto un fallo del servicio.

Otra desventaja del VRRP es que, o bien se desconectan todas las sesiones de encaminamiento, o el encaminador de seguridad tiene sesiones entre pares separadas con los mismos vecinos que el encaminador principal, provocando una sobrecarga significativa para procesar el encaminamiento. En cualquier caso, existe un tiempo de convergencia cuando el encaminador principal falla porque las sesiones entre pares del encaminador principal se caerán.

Otro ejemplo de redundancia se introduce en el documento US nº 5.473.593.

15 Sumario de la invención

Según un primer aspecto de la invención, se proporciona un procedimiento según se reivindica en la reivindicación 1 en la presente memoria.

Según un segundo aspecto de la invención, se proporciona un dispositivo de red según se reivindica en la reivindicación 9 en la presente memoria.

Según un tercer aspecto de la invención, se proporciona un sistema según se reivindica en la reivindicación 18 en la presente memoria.

Según un cuarto aspecto de la invención, se proporciona un programa informático según se reivindica en la reivindicación 19 en la presente memoria.

Las características preferidas de la invención se definen en las reivindicaciones dependientes.

En lo sucesivo, el término “realización” no debería equipararse con realización de la invención.

Breve descripción de los dibujos

La presente invención se ilustra a modo de ejemplo y no pretende estar limitada por las figuras de los dibujos adjuntos en los que las referencias iguales indican elementos similares y en los que:

la figura 1 ilustra una red a modo de ejemplo en la que puede ponerse en práctica la presente invención;

la figura 2 ilustra un modelo de arquitectura por capas que puede usarse por los nodos mostrados en la figura 1 según una realización;

la figura 3 ilustra un diagrama de flujo de un funcionamiento básico de un nodo redundante según una realización;

la figura 4A ilustra un diagrama para mostrar la replicación de la información del protocolo de encaminamiento desde una tarjeta activa hasta una tarjeta de reserva;

la figura 4B ilustra un diagrama de flujo de una operación para replicar información de cambio de estado del protocolo de encaminamiento según una realización;

la figura 5 ilustra un diagrama de flujo de una operación detallada para reestablecer el funcionamiento en la figura 3 con un cambio de conexión rápido al sistema de control de reserva según una realización;

la figura 6A ilustra un diagrama para mostrar la simulación de un nodo par para evitar que se observe un fallo y un cambio de conexión según una realización;

la figura 6B ilustra un diagrama para mostrar la simulación de un nodo par para evitar que observe un fallo y un cambio de conexión según otra realización;

la figura 7 ilustra un diagrama de flujo de una operación de envío de una orden a un nodo par según una realización;

la figura 8 ilustra un dispositivo de red que presenta un sistema de control activo y un sistema de control de reserva según una realización;

la figura 9 ilustra una plataforma de redundancia para un dispositivo de red según una realización;

la figura 10A ilustra un diagrama de flujo de una operación de una secuencia de arranque según una realización;

ES 2 267 850 T3

la figura 10B ilustra un diagrama de flujo de una operación de una secuencia de arranque según otra realización;

las figuras 11A y 11B ilustran un diagrama de flujo de una operación para que una tarjeta activa realice un cambio de conexión de cortesía a una tarjeta de reserva según una realización;

5

las figuras 12A y 12B ilustran un diagrama de flujo de una operación para que una tarjeta de reserva realice el cambio de conexión de cortesía según una realización;

10

la figura 13 ilustra un diagrama de flujo de una operación para que una tarjeta de reserva realice un cambio de conexión no de cortesía según una realización;

la figura 14 ilustra un diagrama de flujo de una operación para actualizar la información persistente en una tarjeta de reserva según una realización;

15

la figura 15 ilustra un diagrama de flujo de una operación para actualizar la información no persistente en una tarjeta de reserva según una realización;

la figura 16 ilustra un diagrama de flujo de una operación para realizar un tratamiento de errores según una realización;

20

la figura 17 ilustra un diagrama para mostrar la interacción del protocolo de encaminamiento dentro de un dispositivo de red según una realización;

la figura 18 ilustra un diagrama para mostrar la interacción del protocolo de encaminamiento entre una tarjeta activa y una tarjeta de reserva según una realización;

25

la figura 19 ilustra una arquitectura para la redundancia del protocolo de encaminamiento según una realización;

la figura 20 ilustra un diagrama a modo de ejemplo para mostrar una interacción entre BGP, TCP e IP;

30

la figura 21A ilustra un diagrama de flujo de una operación para replicar los cambios de estado BGP recibidos o generados según una realización;

la figura 21B ilustra un diagrama de flujo de una operación para replicar los cambios de estado TCP recibidos o generados según una realización;

35

la figura 22 ilustra una realización de un diálogo entre un TCP activo que funciona en una tarjeta activa y un TCP de reserva que funciona en una tarjeta de reserva para mostrar un requisito de etapa de bloqueo para un mensaje BGP enviado a un nodo par;

40

la figura 23 ilustra una realización de un diálogo entre un TCP activo que funciona en una tarjeta activa y un TCP de reserva que funciona en una tarjeta de reserva para mostrar un requisito de etapa de bloqueo para un mensaje BGP recibido desde un nodo par;

la figura 24 ilustra una arquitectura BGP para mostrar una actualización delta para información BGP individual enviada a un nodo par según una realización;

45

la figura 25 ilustra una arquitectura BGP para mostrar una actualización delta para información BGP individual enviada a un nodo par según una realización;

50

la figura 26 ilustra una arquitectura BGP para mostrar una actualización delta para información BGP individual recibida de un nodo par según una realización;

la figura 27 ilustra un diagrama de flujo para llevar a cabo un mensaje BGP según una realización;

55

la figura 28 ilustra un diagrama de flujo de una operación para realizar una actualización masiva de la redundancia del protocolo BGP según una realización;

la figura 29 ilustra un diagrama de flujo de una operación para realizar una actualización masiva de la redundancia del protocolo IS-IS según una realización;

60

la figura 30 ilustra un diagrama de flujo de una operación para realizar una actualización delta de los mensajes IS-IS recibidos o enviados según una realización;

la figura 31 ilustra un diagrama de flujo de una operación para realizar una actualización masiva del OSPF según una realización; y

65

la figura 32 ilustra un diagrama de flujo de una operación para realizar una actualización delta de los mensajes OSPF recibidos o enviados según una realización.

Descripción detallada

5 Se describen un encaminador y una redundancia de protocolo de encaminamiento para reducir los cortes o deterioros del servicio para un dispositivo de red y así aumentar la disponibilidad de servicio en una red debido a fallos de software o de hardware del dispositivo de red. Para una realización, un dispositivo de red tal como un encaminador incluye una plataforma de redundancia que presenta un sistema de control activo y un sistema de control de reserva.
10 Un cambio de estado de protocolo de encaminamiento se recibe o genera en el sistema de control activo. El cambio de estado del protocolo de encaminamiento recibido o generado se replica al sistema de control de reserva. Mediante la replicación del cambio de estado del protocolo de encaminamiento, el sistema de control de reserva puede mantener las sesiones del protocolo de encaminamiento para el dispositivo de red en caso de producirse un fallo en el sistema de control activo. Además, los estados del protocolo de encaminamiento se mantienen en tiempo real para abordar los
15 cambios dinámicos creados por los protocolos de encaminamiento.

Las técnicas de redundancia descritas en la presente memoria permiten que un dispositivo de red defectuoso vuelva a estar en servicio en un breve periodo de tiempo para evitar cortes del servicio. Las técnicas de redundancia también permiten que un sistema de control de seguridad o de reserva vuelva a poner en servicio el dispositivo de red que ha
20 fallado en el estado operativo del sistema de control activo anterior al fallo. Las técnicas de redundancia también impiden que los nodos pares de un dispositivo de red observen el fallo del dispositivo de red. Las técnicas de redundancia también impiden que las sesiones del protocolo de encaminamiento se caigan en el caso de un cambio de conexión desde un sistema de control activo a un sistema de control de reserva manteniendo las sesiones del protocolo en tiempo real. Las técnicas de redundancia también mantienen una visión consistente de la red en un sistema de control de
25 reserva.

En la siguiente descripción, se describen técnicas de redundancia con respecto a encaminadores de red y protocolos de encaminamiento. Sin embargo, las técnicas de encaminamiento descritas en la presente memoria no pretenden
30 quedar limitadas a ningún tipo particular de dispositivo de red y pueden implementarse en otros tipos de dispositivos de red, que pueden presentar fallos de hardware o de software o realizar funciones de protocolos de encaminamiento tales como, por ejemplo, interruptores de red, interruptores ópticos de red, puentes, concentradores o puertas de enlace.

Además, en la siguiente descripción, la redundancia de encaminador se refiere a un encaminador que presenta un sistema de control de seguridad (es decir, sistema de control de reserva) para un sistema de control activo. El sistema de
35 control de reserva puede reanudar el funcionamiento del sistema de control activo si el sistema de control activo falla. Además, la redundancia de protocolo de encaminamiento se refiere al mantenimiento de las sesiones del protocolo que se están efectuando en el sistema de control activo en el sistema de control de reserva y al mantenimiento de un encaminamiento y reenvío de información en el sistema de control de reserva consistente con el sistema de control activo.

40 *Visión general del encaminador y la redundancia de protocolo de encaminamiento*

Red con redundancia de encaminamiento ilustrativa

La figura 1 ilustra una red 100 ilustrativa en la que puede ponerse en práctica la presente invención. En referencia a
45 la figura 1, un sistema de red ilustrativo incluye un nodo 104 que presenta una plataforma 900 de redundancia (“nodo redundante 104”) interconectado con una pluralidad de nodos pares 102A y 102B. Para la finalidad de la ilustración se muestran tres nodos, sin embargo, la red 100 puede incluir cualquier número de nodos. Los nodos pares 102A y 102B son nodos que tienen una “sesión” o “conexión lógica” con el nodo redundante 104.

50 Para una realización, los nodos 102A, 102B y el nodo redundante 104 representan dispositivos de red tales como, por ejemplo, encaminadores de red que realizan servicios de capa 3 del IP. Alternativamente, los nodos 102A, 102B y el nodo redundante 104 pueden ser otro tipo de dispositivos de red tales como, por ejemplo, interruptores, puentes, concentradores o puertas de enlace que pueden llevar a cabo servicios de capa 3 del IP o incluso servicios de mayor nivel, incluso servicios de aplicación. En otras realizaciones, los nodos 102A, 102B y el nodo redundante 104 pueden
55 realizar servicios de conmutación de etiquetas multiprotocolo (MPLS).

Los nodos 102A, 102B y el nodo redundante 104 pueden representar encaminadores de red que se usan para reenviar información (es decir, paquetes) a través de un grupo particular de redes bajo la misma entidad de administración y control, que normalmente se denomina como Sistema Autónomo (AS). Como tal, los nodos 102A, 102B y el nodo
60 redundante 104 pueden representar “encaminadores internos” que emplean un Protocolo de puerta de enlace interna (IGP) para el intercambio de información con el AS.

En caso de actuar como IGP, los nodos 102A, 102B y el nodo redundante 104 pueden emplear protocolos de encaminamiento tales como un protocolo de sistema intermedio a sistema intermedio (IS-IS), un protocolo de abrir primero la ruta más corta (OSPF) y un protocolo de información de encaminamiento (RIP). El protocolo IS-IS y el
65 protocolo OSPF son protocolos de estado de enlace. Un protocolo de estado de enlace emplea paquetes de estado de enlace para mantener una visión consistente de la red. El protocolo RIP es un protocolo simple basado en vectores de distancia que emplean un cálculo de la ruta más corta.

ES 2 267 850 T3

Alternativamente, los nodos 102A, 102B y 104 pueden representar encaminadores de red que se emplean para reenviar información entre AS, en cuyo caso los encaminadores se denominan “encaminadores externos” y utilizan protocolos de puerta de enlace externa (EGP). Si actúan como EGP, el nodo 102A, 102B y el nodo redundante 104 pueden utilizar un protocolo de encaminamiento tal como un protocolo de puerta de enlace de frontera (BGP). El protocolo BGP intercambia información de conectividad a través de un protocolo de transporte fiable tal como el protocolo de control de transporte (TCP, “Transport Control Protocol”) y no tiene capacidades de control de errores. Los nodos 102A, 102B y el nodo redundante 104, sin embargo pueden representar cualquier combinación de encaminadores internos o encaminadores externos y puede representarse cualquier número de encaminadores en la red 100.

Por lo tanto, los nodos 102A, 102B y el nodo redundante 104 pueden mantener un encaminamiento e información de estado del protocolo de encaminamiento consistentes. Si se actualiza una ruta, la ruta debe actualizarse en los nodos pares para mantener una visión consistente de la red. Para una realización, los nodos 102A, 102B y el nodo redundante 104 pueden determinar vecinos enviando un paquete de “saludo”. Si un nodo par establecido no responde al paquete de “saludo” en un determinado periodo de tiempo, el nodo par se considerará inoperativo o que ha “dado un fallo”.

Como tal, el nodo redundante 104 es un tipo especial de nodo que presenta una plataforma de redundancia con un sistema de control activo (tarjeta activa 910) y un sistema de control de reserva (tarjeta de reserva 950) que puede evitar que los fallos en el nodo se observen por los nodos pares 102A y 102B. Adicionalmente, si el sistema de control activo falla, el sistema de control de reserva puede reanudar las sesiones de protocolo con los nodos pares de tal forma que los nodos pares no observan que el sistema de control activo ha fallado. Por ejemplo, si el nodo 102A envía un paquete de “saludo” al nodo redundante 104, que cualquier motivo se pierde, y se produce un cambio de conexión y el nodo 102A vuelve a enviar el paquete de saludo, el sistema de control de reserva puede reanudar el funcionamiento del nodo redundante 104 y acusar recibo del paquete de saludo reenviado antes del tiempo máximo. Así, el nodo 102A no observa el cambio de conexión al sistema de reserva.

Para una realización, el nodo 104 representa un encaminador que presenta una plataforma 900 de redundancia tal como se muestra en la figura 9. La plataforma 900 de redundancia, incluye una tarjeta activa 910 y una tarjeta de reserva 950 para reanudar el funcionamiento si se produce un fallo en la tarjeta activa 910. La tarjeta activa 910 y la tarjeta de reserva 950 incluyen módulos de hardware y de software funcionando en su interior. Para una realización, tanto la tarjeta activa 910 como la tarjeta de reserva 950 pueden emplear potencialmente la misma o diferentes versiones de software. Tal como se explicará en más detalle más adelante, la plataforma 900 de redundancia proporciona el soporte para presentar el encaminador y la redundancia de protocolo de encaminamiento para el nodo 104, que evita que los nodos pares observen fallos y mantiene las sesiones de protocolo de encaminamiento para el nodo redundante 104 con sus nodos pares.

Modelo de arquitectura por capas ilustrativo

La figura 2 ilustra un modelo 200 de arquitectura por capas que puede usarse por los nodos mostrados en la figura 1 según una realización. Para una realización, la tarjeta activa 910 y la tarjeta de reserva 950 funcionan usando el modelo 200 de arquitectura por capas. El modelo 200 de arquitectura por capas puede basarse en un modelo de referencia de 7 capas estándar para comunicaciones de red. Para la finalidad de la explicación, el modelo 200 de arquitectura por capas representa una realización de las diferentes capas en la que puede funcionar un encaminador IP.

En referencia a la figura 2, el modelo 200 de arquitectura por capas incluye una capa física 202, una capa de enlace 204, una capa IP 206, una capa de protocolo de control de la transmisión (TCP) 208, una capa de protocolo de datagrama de usuario (UDP) 208, una capa de protocolo de mensaje de Internet (ICMP) 218, una capa de protocolos de encaminamiento 220, que incluye un protocolo de puerta de enlace de frontera (BGP) 226, una capa de protocolo de Internet de encaminamiento (RIP) 222, una capa de protocolo de abrir primero la ruta más corta (OSPF) 224, una capa de protocolo de sistema intermedio a sistema intermedio (IS-IS) 214, capas de conexión 210, una capa de aplicación 212. Los protocolos de encaminamiento 220 pueden mantener una tabla de encaminamiento para generar una tabla de información de reenvío (FIB) 216. La FIB 216 se usa por la capa IP 209 y la capa de enlace 216. Las capas anteriores proporcionan servicios para los nodos de una red.

La capa física 202 proporciona el servicio de mover datos entre nodos en un enlace físico. La capa de enlace 204 proporciona el servicio de tratar los datos que están siendo transferidos en el enlace físico. La capa IP 209 (“capa 3 del IP”) proporciona servicios de encaminamiento y reenvío a través de las capas física y de enlace. La capa TCP 208 proporciona el servicio de garantizar la transferencia de datos completa realizando una comprobación de errores y asegurando que todos los datos han llegado. La capa TCP 208 funciona sobre la capa 3 del IP. Por lo tanto, los nodos en una red pueden transmitir datos usando un servicio TCP sobre un servicio de capa 3 del IP.

La capa ICMP 218 funciona encima, y es una parte integrante de la capa IP 206. Es decir, el servicio de capa 3 es inherentemente poco fiable y los paquetes de datos pueden caerse. Por tanto, la capa ICMP 218 proporciona un control de los mensajes e informes de errores para el servicio de capa 3 del IP. La capa UDP 209 proporciona un servicio alternativo al servicio que proporciona la capa TCP 208. En particular, la capa UDP 209 funciona encima de la capa de servicio de capa 3 del IP para proporcionar un protocolo de transmisión sin conexión para transmitir datagramas. Es decir, la capa UDP 209 no proporciona detección de extremo a extremo. La capa de conexión 210 proporciona un punto de extremo de una comunicación de doble sentido entre la capa de aplicaciones 212 o los protocolos de encaminamiento

ES 2 267 850 T3

220 de un nodo que funciona sobre una red. La capa de aplicaciones 212 incluye aplicaciones que funcionan sobre un nodo. La capa de aplicaciones 212 puede usar las capas inferiores para comunicarse con aplicaciones de otros nodos.

5 Los protocolos de encaminamiento 220 proporcionan el servicio de determinar rutas óptimas, reenviar paquetes, y garantizar que las actualizaciones de las rutas son consistentes en toda la red. Al analizar las actualizaciones de rutas de todos los encaminadores, un encaminador puede formar una visión detallada de la red. En los protocolos de encaminamiento 220, pueden utilizarse varios protocolos de encaminamiento. Por ejemplo, el protocolo BGP 226, el protocolo RIP 222, el protocolo OSPF 224, y el protocolo IS-IS 220 pueden funcionar todos dentro de la capa de protocolos de encaminamiento 220. Los protocolos de encaminamiento 220 pueden usar FIB 216 para la transmisión de datos (por ejemplo, paquetes) en el servicio de capa 3 proporcionado por la capa IP 206.

15 El protocolo BGP 226 no es un protocolo de encaminamiento fiable. Por tanto, el BGP 226 funciona encima del TCP 208 para la transferencia fiable de mensajes o paquetes. En particular, el BGP 226 no reenvía mensajes o paquetes sino que se basa en el TCP 208 para tratar los mensajes o paquetes perdidos. El RIP 222 utiliza UDP 209 para la transferencia de mensajes o paquetes. El OSPF 224 y el IS-IS 214 tienen mecanismos de transferencia de datos fiables en sus respectivos protocolos de encaminamiento. El OSPF 224 funciona sobre la capa IP 206 y el IS-IS 214 funciona directamente sobre la capa de enlace 204.

20 Tal como se explicará más detalladamente más adelante, los protocolos de encaminamiento que funcionan en el nodo redundante 104 pueden funcionar junto con la plataforma 900 de redundancia para obtener el encaminador y la redundancia de protocolos de encaminamiento. La plataforma 900 de redundancia proporciona el soporte necesario para presentar redundancia de protocolos de encaminamiento en tiempo real. Es decir, los protocolos de encaminamiento son dinámicos por lo que las actualizaciones de las rutas se realizan a intervalos regulares o irregulares dependiendo de los protocolos de encaminamiento. Para obtener una redundancia total, estas actualizaciones deben mantenerse en tiempo real.

30 Por ejemplo, los estados de las sesiones de los protocolos de encaminamiento para los RIP 222, OSPF 224, BGP 226 y IS-IS 214 que pueden estar ejecutándose en la tarjeta activa 910 para el nodo redundante 104 pueden mantenerse en tiempo real en la tarjeta de reserva 950 empleando la plataforma 900 de redundancia. La tarjeta de reserva 950 puede reanudar los mismos estados de las sesiones de los protocolos de encaminamiento si la tarjeta activa 910 falla. Además, la FIB 216 también se mantiene en la tarjeta de reserva 950 de tal manera que la tarjeta de reserva 950 tendrá la visión más actualizada de la red en caso de asumir el control del nodo redundante 104.

Operación de redundancia básica

35 La figura 3 ilustra un diagrama de flujo de una operación 300 básica para el nodo redundante 104 según una realización. La siguiente operación 300 puede implementarse por el nodo redundante 104 que presenta una tarjeta activa 910 y una tarjeta de reserva 950 tal como se muestra en la figura 1. Para la finalidad de la explicación, el nodo redundante 104 es un encaminador de red 104 que puede realizar servicios de capa 3 del IP o MPLS y la operación 300 comienza en la operación 302.

45 En la operación 302, el nodo redundante 104 mantiene el estado actual de la tarjeta activa 910 en la tarjeta de reserva 950. En particular, el nodo redundante 104 utiliza la plataforma 900 de redundancia para replicar o copiar la información de configuración actual, la información global, la información de la tabla de encaminamiento, la información de la tabla de reenvío, la información de la sesión de protocolo, o la información de bases de datos en la tarjeta activa 910 a la tarjeta de reserva 950.

50 En la operación 304, la tarjeta activa 910 detecta un fallo. Por ejemplo, la tarjeta activa 910 puede detectar un fallo de hardware o un fallo de software en el nodo redundante 104 que provocará que el nodo redundante 104 cambie la conexión del funcionamiento de la tarjeta activa 910 a la tarjeta de reserva 950.

55 En la operación 306, la tarjeta de reserva 950 reanudará el funcionamiento del estado actual de la tarjeta activa 910 antes del fallo. La tarjeta de reserva 950 reanudará el funcionamiento de tal manera que el fallo no se observa en los nodos pares 102A o 102B. Tal como se describirá más detalladamente en las figuras 4A a 7, el nodo redundante 104 puede evitar que los fallos se observen en los nodos pares al mantener en tiempo real los estados de las sesiones de protocolo de encaminamiento de la tarjeta activa 910 en la tarjeta de reserva 950 y obteniendo un cambio de conexión rápido a la tarjeta de reserva 950 de tal forma que no se caiga una sesión de protocolo. Así, un cambio en la topología de la red no tiene que propagarse por toda la red y el tiempo de convergencia se reduce.

Mantenimiento de las sesiones de protocolo y la información de encaminamiento

60 La figura 4A ilustra un diagrama 400 para mostrar la replicación de la información de protocolo de encaminamiento desde la tarjeta activa 910 hasta la tarjeta de reserva 950. La plataforma 900 de redundancia, tal como se muestra en la figura 9, proporciona el soporte para el mantenimiento del protocolo de información de protocolo de encaminamiento en tiempo real en la tarjeta de reserva 950. En referencia a la figura 4A, el diagrama muestra el nodo par 102A en comunicación con el nodo redundante 104. El nodo redundante 104 incluye una plataforma 900 de redundancia que presenta una tarjeta activa 910 que se comunica con el nodo par 102A. El nodo redundante 104 también incluye una tarjeta de reserva 950 para reanudar el funcionamiento si se produce un fallo en la tarjeta activa 910.

ES 2 267 850 T3

Para una realización, el nodo par 102A incluye información de protocolo 415 que incluye datos persistentes 411, estados de las sesiones 412 y la tabla de encaminamiento 413, que se usa para generar la FIB 432. La información de protocolo 415 debe ser consistente con la información de protocolo 405A en la tarjeta activa 910. Es decir, la información de los datos persistentes 411, los estados de las sesiones 412 y la tabla de encaminamiento 413 del nodo par 102A debe ser consistente con los datos persistentes 401A, los estados de las sesiones 402A y la tabla de encaminamiento 403A de modo que la tarjeta de reserva 950 pueda replicarse con la misma información para garantizar la redundancia. Además, si las tablas de encaminamiento 413 y 403A no son consistentes, la FIB 432 en el nodo par 102A no será consistente con la FIB 422A en la tarjeta activa 910. Como tal, el nodo par 102A puede considerar que el nodo redundante 104 presenta una visión diferente de la red que éste tiene y el nodo par 102A puede derribar las sesiones de protocolo de encaminamiento con el nodo redundante 104.

Para conseguir una redundancia completa, los cambios recibidos o realizados por la tarjeta activa 910 en la información de protocolo 405A deben replicarse a la información de protocolo 405B en la tarjeta de reserva 950. Específicamente, los cambios recibidos o realizados por la tarjeta activa 910 en los datos persistentes 401A, los estados de las sesiones 402A y la tabla de encaminamiento 403A se replican en los datos persistentes 401B, los estados de las sesiones 402B y la tabla de encaminamiento 403B en la tarjeta de reserva 950. Si los cambios no se mantienen, la redundancia se rompe.

La información de protocolo puede referirse a los protocolos de encaminamiento tales como, por ejemplo, los protocolos de encaminamiento BGP, RIP, OSPF e IS-IS. Los datos persistentes pueden incluir información de configuración para cada protocolo de encaminamiento que sea de naturaleza más permanente. La información del estado de las sesiones incluye cambios en el estado del protocolo de encaminamiento para cada protocolo de encaminamiento que funciona en un nodo. La información sobre el estado de las sesiones es de naturaleza dinámica y puede cambiar a intervalos regulares o irregulares. Por ejemplo, los datos de los cambios en el estado del protocolo de encaminamiento pueden incluir información relativa a las normas de comunicación entre nodos, el estado de cada ruta recibida desde un par, el estado de cada ruta enviada a un par, parámetros de tiempo máximo, el historial de las rutas eliminadas por cada par, etc., para cada protocolo de encaminamiento. La información de la tabla de encaminamiento incluye las rutas conocidas por un nodo para cada protocolo de encaminamiento. La información de la tabla de encaminamiento se usa para generar la tabla FIB, que se usa para reenviar paquetes.

Dado que la tarjeta de reserva 950 está poblada con la información pertinente requerida para cada protocolo de encaminamiento ejecutado en la tarjeta activa 910, si se produce un fallo en la tarjeta activa 910, la tarjeta de reserva 950 puede reanudar las sesiones de protocolo de encaminamiento de la tarjeta activa 910 usando una FIB consistente con la tarjeta activa 910. Por tanto, la tarjeta de reserva 950 puede reanudar las mismas sesiones de protocolo usando los mismos estados de la tarjeta activa 910. En tal caso, el nodo par 102A se comunicará con la tarjeta de reserva 950 creyendo que sigue comunicándose con la tarjeta activa 910 evitando de este modo un corte del servicio.

La figura 4B ilustra un diagrama de flujo de una operación 450 para replicar la información del cambio de estado del protocolo de encaminamiento según una realización. La siguiente operación 450 puede implementarse por el nodo redundante 104 que presenta una tarjeta activa 910 y una tarjeta de reserva 950 tal como se muestra en la figura 1. Para la finalidad de la explicación, el nodo redundante 104 es un encaminador de red 104 que realiza servicios de capa 3 del IP o MPLS y la operación 450 comienza en la operación 452.

En la operación 452, la información del cambio de estado del protocolo de encaminamiento se recibe o genera en la tarjeta activa 910. Por ejemplo, la tarjeta activa 910 puede generar un cambio en los datos persistentes 401A, en los estados de las sesiones 402A y en la tabla de encaminamiento 403A. Alternativamente, la tarjeta activa 910 puede recibir un cambio desde un nodo par 102A en los datos persistentes 401A, en los estados de las sesiones 402A y en la tabla de encaminamiento 403A.

En la operación 454, el cambio de estado del protocolo de encaminamiento recibido o generado en la tarjeta activa 910 se replica a la tarjeta de reserva 950. Por ejemplo, la plataforma 900 de redundancia tal como se muestra en la figura 9, proporciona el soporte para realizar la replicación de los datos persistentes 401A, los estados de las sesiones 402A y la tabla de encaminamiento 403A de la tarjeta activa 910 en los datos persistentes 401B, los estados de las sesiones 402B y la tabla de encaminamiento 403B de su par en la tarjeta de reserva 950. Dicha operación de replicación se realiza en tiempo real. Por tanto, si la tarjeta activa 910 falla, la tarjeta de reserva 950 puede reanudar el funcionamiento usando la misma información de la tarjeta activa 910.

Cambio de conexión rápido/Simulación del fallo

Las figuras 5 a 7 ilustran cómo un nodo con una plataforma 900 de redundancia puede evitar que los nodos pares observen los fallos y cambios de conexión. Al mantener la información de los datos persistentes, los estados de sesión y la tabla de encaminamiento en un sistema de control de reserva (tarjeta de reserva 950) consistente con el resto de la red, puede realizarse un cambio de conexión sin interrupciones, uniforme y rápido. El cambio de conexión es suficientemente rápido (por ejemplo, de unos pocos milisegundos) para que los nodos pares no observen que el nodo redundante 104 puede haberse retrasado ligeramente para realizar el cambio de conexión.

ES 2 267 850 T3

La figura 5 ilustra un diagrama de flujo de una operación 306 detallada de la figura 3 para reanudar el funcionamiento mediante un sistema de control de reserva que presenta un cambio de conexión rápido según una realización. Para la finalidad de la explicación, la operación 306 comienza en la operación 502.

5 En referencia a la figura 5, en la operación 502, el nodo redundante 104 realiza un cambio de conexión desde la tarjeta activa 910 hasta la tarjeta de reserva 950. El cambio de conexión puede realizarse en unos pocos milisegundos. La pequeña cantidad de tiempo para realizar el cambio de conexión (es decir, el tiempo de “fallo técnico”) es demasiado pequeño para que un nodo par 102A no observe que se ha producido un cambio de conexión. En la operación 10 504, la tarjeta de reserva 350 reanuda el funcionamiento con el nodo par 102A sin derribar una sesión de protocolo. Es decir, el tiempo de fallo técnico es tan pequeño que el nodo par 102A sabe que ha habido un tiempo de “fallo técnico” en el nodo redundante 104.

A continuación se da una descripción detallada para explicar el rápido cambio de conexión. Dado que el IP es inherentemente poco fiable, los paquetes pueden caerse. Si un paquete se recibe en la tarjeta activa 910 desde el nodo par 15 102A y se produce un fallo tal que la tarjeta activa 910 no acusa recibo del paquete, el nodo par 102A puede volver a enviar el paquete. En esta situación, incluso si hay un cambio de conexión, la tarjeta de reserva 950 todavía puede recibir el paquete reenviado desde el nodo par 102A y acusar recibo de ese paquete. Así, el nodo par 102A observará un suceso normal (es decir, el reenvío de un paquete y la recepción de un acuse de recibo del paquete reenviado) y no creará que se ha producido un fallo o cambio de conexión en el nodo redundante 104 incluso aunque el paquete se haya caído. 20

Puesto que el cambio de conexión se realiza de forma rápida, la tarjeta de reserva 950 puede reanudar el funcionamiento para el nodo redundante 104 sin que se rompa una sesión del protocolo de encaminamiento. Por ejemplo, cada protocolo de encaminamiento incluye un periodo de tiempo máximo en el que romperá una sesión si no se acusa el recibo de un cierto número de paquetes en un determinado periodo de tiempo. En consecuencia, incluso si algunos 25 paquetes se caen durante el cambio de conexión, el tiempo de fallo técnico es corto para el nodo redundante 104 que presenta la plataforma 900 de redundancia. Es decir, la tarjeta de reserva 950 puede reanudar el funcionamiento de la tarjeta activa 910 sin que se rompan sesiones o fallen los servicios del nodo redundante 104.

La figura 6A ilustra un diagrama 600 para mostrar la simulación de un nodo par para evitar que se observe un 30 fallo y un cambio de conexión según una realización. Dado que hay un aspecto de tiempo real en los protocolos de encaminamiento, un nodo par 102A del nodo redundante 104 requiere que una transacción que se está llevando a cabo por la tarjeta activa 910 se lleve a cabo también por la tarjeta de reserva 950. Por ejemplo, si la tarjeta activa 910 lleva a cabo una actualización del encaminamiento, la actualización del encaminamiento debe realizarse también en la tarjeta de reserva 950. Si la transacción llevada a cabo por la tarjeta activa 910 no se lleva a cabo por la tarjeta de reserva 35 950 y se produce un cambio de conexión, el nodo par 102A romperá su sesión con el nodo redundante 104 porque la tarjeta de reserva 950 no tendrá una visión de la red consistente con el nodo par 102A. Específicamente, la tarjeta de reserva 950 no llevó a cabo la actualización.

Por lo tanto, para evitar que el nodo par 102A rompa la sesión con el nodo redundante 104, el cambio de conexión 40 a la tarjeta de reserva 950 debe hacerse rápidamente y las transacciones llevadas a cabo por la tarjeta activa 910 deben llevarse a cabo por la tarjeta de reserva 950. El nodo redundante 104 que presenta una plataforma 900 de redundancia puede realizar dicho cambio de conexión rápido y llevar a cabo transacciones en la tarjeta de reserva 950 evitando que un nodo par 102A rompa las sesiones de protocolo.

En referencia a la figura 6, el diagrama ilustra una tarjeta activa 910 que recibe un mensaje (MSG A) desde un nodo 45 par. El mensaje puede ser, por ejemplo, para informar al nodo redundante 104 de que el nodo 102B ha fallado y que su ruta hacia 102B debe actualizarse de forma correspondiente. El nodo 102A necesita obtener una confirmación de que el nodo 104 ha recibido la actualización y está realizando los cambios necesarios (es decir, llevar a cabo el mensaje). Si el nodo 104 no lleva a cabo el mensaje, el nodo redundante 104 presentará una visión de la red inconsistente con el 50 nodo par 102A. Así, si el nodo par 102A cree que el nodo redundante 104 presenta una visión de la red inconsistente, el nodo redundante 102A romperá su sesión con el nodo redundante 104 provocando un corte del servicio.

Ahora bien, si la tarjeta activa 910 está procesando el mensaje y falla en los puntos de fallo 1 ó 2 (el nodo redundante 104 no llevó a cabo el mensaje en esos puntos) y se produce un cambio de conexión en la tarjeta de reserva 950, el 55 nodo par 102A no romperá su sesión con el nodo 104 porque puede volver a enviar el mensaje nuevamente (hasta un cierto número de reintentos) y la tarjeta de reserva 950 recibirá el mensaje reenviado y responderá en consecuencia llevando a cabo la actualización de la ruta de que el nodo par 102B ha fallado. Siempre que el cambio de conexión se produzca rápidamente y la tarjeta de reserva 950 reanude el funcionamiento antes del número máximo de reintentos, la tarjeta de reserva 950 puede reanudar el funcionamiento del nodo redundante 104 sin que el fallo y el cambio de 60 conexión se observen por el nodo par 102A.

Sin embargo, si se produce un fallo en el punto de confirmación por la tarjeta activa 910, la tarjeta de reserva 950 también debe llevar a cabo el mensaje. Es decir, si la tarjeta activa 910 llevó a cabo el mensaje y la tarjeta de reserva 950 no llevó a cabo el mensaje, la redundancia se rompe y habrá una inconsistencia de la información en la tarjeta 65 activa 910 y la tarjeta de reserva 950 que puede provocar un fallo del servicio. Por ejemplo, la transacción llevada a cabo se refería a una actualización de ruta y, si la tarjeta de reserva 950 no lleva a cabo esa actualización, presentará una visión de la red inconsistente con el nodo par 102A. Por tanto, la tarjeta de reserva 950 debe llevar a cabo el mensaje llevado a cabo por la tarjeta activa 910.

ES 2 267 850 T3

En el ejemplo de la figura 6A, para garantizar que un mensaje llevado a cabo por la tarjeta activa 910 se lleva a cabo por la tarjeta de reserva 950, la tarjeta activa 910 no llevará a cabo un mensaje a menos que la tarjeta de reserva 950 haya llevado a cabo el mensaje. Tal como se muestra en la figura 6A, tras recibir el MSG A, una tarjeta activa 910 envía el MSG A a la tarjeta de reserva 950. La tarjeta de reserva 950 envía un acuse de recibo del MSG A a la tarjeta activa 910 (por ejemplo, “MSG A recibido y llevado a cabo el MSG A”). Una vez que la tarjeta activa 910 recibe el acuse de recibo desde la tarjeta de reserva 950, la tarjeta activa 910 llevará a cabo el mensaje y enviará la confirmación (“acuse de recibo de que el MSG A se ha llevado a cabo”) al par remoto. La tarjeta activa 910 enviará entonces el MSG A a través de las capas superiores.

La figura 6B ilustra un diagrama 650 para mostrar la simulación de un nodo par para evitar que se observe un fallo y un cambio de conexión según otra realización. En el ejemplo de la figura 6B, la tarjeta activa 910 puede enviar el MSG A directamente a través de las capas superiores, pero no llevará a cabo el MSG A hasta haber recibido un acuse de recibo de que la tarjeta de reserva 950 ha llevado a cabo el MSG A. En ese punto, la tarjeta activa 910 enviará una “confirmación” al par remoto. Tal como se muestra en la figura 6B, el MSG A pasará más rápidamente a través de las capas superiores de la tarjeta activa 910 que el MSG A en la figura 6A.

La plataforma 900 de redundancia proporciona el soporte para la actualización de la tarjeta de reserva 950 con información relativa a transacciones llevadas a cabo en la tarjeta activa 910. Tal como se explicará más adelante, los mensajes o transacciones llevados a cabo pueden requerir pequeñas actualizaciones o grandes actualizaciones. Para transacciones individuales, se requiere una actualización pequeña o “delta” de la tarjeta de reserva 950. Para reproducir un largo historial de transacciones, se requiere una actualización grande o “masiva” de la tarjeta de reserva 950. Así, todas las transacciones llevadas a cabo por la tarjeta activa 910 pueden mantenerse en la tarjeta de reserva 950.

La figura 7 ilustra un diagrama de flujo de una operación 700 para enviar una orden a un nodo par según una realización. En relación a la figura 7, en la operación 702, se recibe un mensaje desde el nodo par. Por ejemplo, la tarjeta activa 910 puede recibir un mensaje desde un nodo par 102A de que el estado de una ruta ha cambiado y debe realizarse una actualización.

En la operación 704, la información relativa al mensaje se envía a la tarjeta de reserva 950. En la operación 706, la tarjeta de reserva 950 procesa el mensaje y lleva a cabo el mensaje cambiando el estado de la ruta. Al llevar a cabo el mensaje, la tarjeta de reserva 950 envía un acuse de recibo a la tarjeta activa 910. La tarjeta activa 910 recibe así el acuse de recibo desde la tarjeta de reserva 950.

En la operación 708, tras recibir el acuse de recibo de la tarjeta de reserva 950, la tarjeta activa 910 lleva a cabo el mensaje. En este punto, la tarjeta de reserva 950 también cambiará el estado de la ruta.

En la operación 710, tras llevar a cabo el mensaje, la tarjeta activa 910 puede enviar una “orden” al nodo par informando al nodo par de que se ha realizado la actualización de la ruta, manteniendo así una visión de la red consistente no sólo en la tarjeta activa 910 sino también en la tarjeta de reserva 950.

Redundancia del encaminador

Hardware de redundancia del encaminador

La figura 8 ilustra un dispositivo 104 de red que presenta un sistema de control activo y un sistema de control de reserva redundante según una realización. En referencia a la figura 8, el dispositivo 104 de red incluye una pluralidad de puertos 814. Los puertos 814 pueden soportar señales eléctricas o señales ópticas con tasas de transferencia de datos variables. Una vista 810 en despiece ordenado ilustra los componentes internos básicos del dispositivo 104 de red, que incluye una tarjeta de línea 812A, una tarjeta activa 910, una tarjeta de reserva 950, una tarjeta de línea 812, y una placa de conexiones 814 que acopla las tarjetas entre sí. Otros tipos de componentes también pueden incluirse, tales como una tarjeta de control de sistema. Para una realización, el dispositivo 104 de red es un encaminador de red para proporcionar servicios de capa 3 del IP. En otras realizaciones, el dispositivo 104 de red puede proporcionar servicios de nivel de capa superior, hasta servicios de memoria de aplicación. El dispositivo 104 de red también puede proporcionar servicios de conmutación de etiquetas multiprotocolo (MPLS).

Para que el encaminador 104 sea redundante, la tarjeta activa 910 y la tarjeta de reserva 950 presentan componentes y módulos de hardware y de software idénticos. La tarjeta activa 910 y la tarjeta de reserva 950 pueden incluir procesadores de red de gran velocidad, controlador de memoria estándar para controlar los dispositivos de memoria tales como dispositivos de memoria de acceso aleatorio estática (SRAM), dispositivos de memoria de acceso aleatorio dinámica (DRAM), o dispositivos de memoria similares. Tales dispositivos de memoria pueden usarse para almacenar información de protocolo, información global, o información de configuración sobre la tarjeta. Los dispositivos de memoria también pueden almacenar instrucciones, módulos de software y sistemas operativos para controlar las tarjetas.

Para una realización, la placa de conexiones 814 es pasiva y permite la comunicación entre la tarjeta de línea 812A, la tarjeta activa 910, la tarjeta de reserva 950 y la tarjeta de línea 812B. En otras realizaciones, la placa de conexiones 814 puede soportar una redundancia de tarjetas de > 2 de líneas, de modo que la tarjeta activa 910 y la tarjeta de reserva 950 pueden controlar más de una tarjeta de línea. La tarjeta activa 910 incluye hardware y/o software para

ES 2 267 850 T3

detectar fallos en la tarjeta activa 910 o en la tarjeta de línea 812A y transferir el funcionamiento a la tarjeta de reserva 950. La tarjeta de reserva 950 también incluye hardware y/o software para reanudar el funcionamiento de la tarjeta activa 910 si ésta falla.

5 Tal como se explicará con más detalle con relación a la plataforma 900 de redundancia, la tarjeta activa 910 y la tarjeta de reserva 950 incluyen subsistemas de reenvío y recuperación de datos para mantener la información consistente en la tarjeta de reserva 950. La tarjeta activa 910 puede comunicarse con la tarjeta de reserva 950 por un enlace de comunicaciones a través de la placa de conexiones 814. Por ejemplo, el enlace de comunicaciones puede ser un enlace de interfaz de control periférica (PCI) o un enlace Ethernet.

10 El encaminador 104 puede proporcionar los siguientes tipos de redundancia de hardware o de encaminador: (a) redundancia de conjunto de tarjetas, (b) redundancia de tarjeta de sistema, (c) redundancia de puerto, o (d) redundancia de tarjeta de línea. La redundancia de conjunto de tarjetas se refiere a la tarjeta de reserva 950 y a la tarjeta de línea 812B que actúa como un par de redundancia respecto a la tarjeta activa 910 y a la tarjeta de línea 812A. Para una
15 realización, la tarjeta de línea 812A y la tarjeta activa 910 pueden insertarse en las ranuras 0 y 1 y la tarjeta de reserva 950 y la tarjeta de línea 812B puede insertarse en las ranuras 2 y 3 para el encaminador 104. Por tanto, un fallo en la tarjeta de línea 812A o en la tarjeta activa 910 provocará un cambio de conexión a la tarjeta de línea 812B y la tarjeta de reserva 950.

20 La redundancia de tarjeta de sistema se refiere a la tarjeta de reserva 950 actuando como una tarjeta de sistema redundante con respecto a la tarjeta activa 910. Para una realización, la redundancia de tarjeta de sistema es la configuración por defecto para el encaminador 104 y es independiente de la redundancia de puerto y puede habilitarse con o sin redundancia de puerto. La redundancia de puerto se refiere a presentar puertos redundantes 814. Por ejemplo, un
25 a redundancia de cable tipo “Y” puede implementarse para los puertos 814. Para una realización, la redundancia de puerto se aplica sólo a las tarjetas de línea individuales. La redundancia de tarjeta de línea se refiere a presentar una tarjeta de línea redundante para una tarjeta de línea activa. Por ejemplo, las tarjetas de línea 812A pueden presentar una tarjeta de línea redundante y la tarjeta de línea 812B puede también presentar una tarjeta de línea redundante.

Estados de tarjeta

30 La tarjeta activa 910 y la tarjeta de reserva 950 deben ser conscientes de dos estados importantes, que son un “estado activo” y un “estado de reserva”. Dependiendo de en qué estado está funcionando la tarjeta, cada tarjeta realizará diferentes tipos de operaciones. Por ejemplo, una tarjeta que funciona en el estado activo actualizará la información de configuración, estado y aprendida en una tarjeta de reserva que funciona con un estado de reserva. La
35 tarjeta que funciona en el estado de reserva recibirá información desde la tarjeta activa y actualizará sus subsistemas de almacenamiento de forma correspondiente. Tal como se explicará más detalladamente más adelante, hay dos clases de actualizaciones de la tarjeta de reserva 950: una actualización grande (“masiva”) y una actualización pequeña o incremental (“delta”).

40 La tarjeta activa 910 se considera que está en un “estado activo” si todos sus diagnósticos y auto-test son capaces de recibir y enviar tráfico de datos desde y hacia los nodos pares y tiene maestría. Una tarjeta activa es por tanto accesible para fines de gestión y aprovisionamiento. Para una realización, una determinación de si una tarjeta está activa puede hacerse mediante un indicador global. Además, un usuario puede determinar qué tarjeta está activa basándose en un
45 indicador de diodo emisor de luz (LED) (por ejemplo, un indicador verde) en el encaminador 104. La tarjeta de reserva 950 se considera que está en “estado de reserva” si todos sus diagnósticos y auto-test están conformes y puede convertirse en una tarjeta activa y no tiene maestría. Para una realización, una tarjeta de reserva es accesible para la gestión pero no para fines de aprovisionamiento. En una realización, una determinación de si una tarjeta está en reserva puede hacerse también mediante un indicador global y un usuario puede determinar qué tarjeta está en reserva basándose en un indicador LED (por ejemplo, un indicador amarillo) en el encaminador 104.

50 Si la tarjeta activa 910 o la tarjeta de reserva 950 no pueden estar operativas, la tarjeta entra en un “estado de fallo”. El estado de fallo también puede determinarse por un indicador LED (por ejemplo un indicador rojo). Una tarjeta se define como “redundante” si la configuración del estado activo se refleja en una tarjeta redundante. La comunicación entre la tarjeta activa y la tarjeta de reserva debería existir en todo momento para mantener la redundancia. En
55 particular, la tarjeta redundante 950 debería ser capaz de activarse si la tarjeta activa 910 falla.

Plataforma de redundancia

Requisitos básicos

60 La figura 9 ilustra una realización de una plataforma 900 de redundancia para el nodo 104. Para que el nodo 104 evite que los nodos pares observen un fallo que se produzca en el nodo, se requiere que la plataforma 900 de redundancia (a) mantenga las conexiones del nodo 104 con los nodos pares (es decir, sesiones con los nodos pares 102A y 102B) sin que se caigan y (b) mantenga la información consistente en la tarjeta activa 910 y en la tarjeta de
65 reserva 950. Como tal, si la tarjeta de reserva 950 reanuda el funcionamiento de una tarjeta activa 910 que ha fallado, la tarjeta de reserva 950 funciona justo como si fuera la tarjeta activa 910.

ES 2 267 850 T3

Arquitectura ilustrativa

En referencia a la figura 9, la plataforma 900 de redundancia muestra una arquitectura de plataforma ilustrativa que presenta una combinación de componentes o módulos de hardware y de software. La plataforma de redundancia incluye una tarjeta activa 910 y una tarjeta de reserva 950 que tiene subsistemas de memoria redundantes y módulos de software. Por ejemplo, los subsistemas de memoria incluyen una memoria de acceso aleatorio (RAM) para almacenar estructuras de datos y datos no persistentes y un disco de memoria flash para almacenar datos persistentes. Además, los módulos de software incluyen gestores de redundancia de software, tareas de aplicación, gestores de control de redundancia y almacenes de datos.

La tarjeta activa 910 incluye un gestor de redundancia de software (SRM, “software redundancy manager”) 918 que se comunica con una tarea de aplicación 916. La tarea de aplicación 916 puede enviar cambios de información a estructuras de datos RAM 912 y datos no persistentes 914. Los datos no persistentes 914 almacenan información que cambia tras una actualización, tal como información de encaminamiento y estados. La tarea de aplicación 916 puede enviar actualizaciones al almacén de datos 922 en el que los cambios de las estructuras de datos RAM 912 y los datos no persistentes 914 pueden almacenarse de manera permanente y redundante en el disco de memoria flash 924. Un gestor de control de redundancia (RCM) 920 puede comunicarse con la tarea de aplicación 920 y el almacén de datos 922. El RCM 920 envía información de actualización a su RCM par 960 en la tarjeta de reserva 950. La tarjeta de reserva 950 incluye una tarea de aplicación par 956, que es reflejar el funcionamiento de la tarea de aplicación 916. La tarea de aplicación par 956 puede comunicarse con el SRM 958 y enviar los cambios a las estructuras de datos RAM 952 y datos no persistentes 954 que se hicieron en la tarjeta activa 910. El RCM 960 también puede enviar los cambios al almacén de datos 962 para actualizar el disco de memoria flash 964 con el fin de mantener la consistencia con el disco de memoria flash 924.

Las estructuras de datos RAM 912 pueden almacenar estados de las sesiones del protocolo de encaminamiento de la tarjeta activa 910 con los nodos pares. Las estructuras de datos RAM 952 en la tarjeta de reserva 950 son para mantener información consistente con las estructuras de datos RAM 912 en la tarjeta activa 910. Los datos no persistentes 914 representan la información almacenada en la RAM. En particular, los datos no persistentes pueden ser paquetes almacenados temporalmente en una memoria intermedia (“buffer”). La información de la FIB puede incluir datos no persistentes, que se actualizan a intervalos regulares o irregulares. Los datos persistentes son datos que se almacenan de manera permanente en un disco de memoria flash empleando el almacén de datos. Por ejemplo, los datos persistentes pueden ser datos de configuración del protocolo de encaminamiento, que no cambian con frecuencia.

El SRM 918 en la tarjeta activa 910 es responsable de detectar fallos de software, notificar a su SRM par 958 cualquier fallo y cambiar la conexión del funcionamiento a la tarjeta de reserva 950. Específicamente, el SRM 916 en la tarjeta activa 910 determina si la tarjeta activa 910 se ha deteriorado o ha fallado. Si el SRM 916 detecta un deterioro o fallo de este tipo en la tarjeta activa 910, el SRM 916 facilita un cambio de conexión a la tarjeta de reserva 950 y puede coordinarse con otras tareas ejecutadas en la tarjeta de reserva 950.

Para que el nodo 104 funcione correctamente, los demás módulos deben estar en un estado “listo”. En particular, el RCM 960, la tarea de aplicación 956 y el almacén de datos 962 deben estar en un estado “listo”. Un estado listo es el estado en el que puede tener lugar un cambio de conexión sin interrupciones. Es decir, cuando la información en las estructuras de datos RAM 912, los datos persistentes 914, el disco de memoria flash 924 en la tarjeta activa 910 son consistentes con la misma en las estructuras de datos RAM 952, los datos persistentes 954 y el disco de memoria flash 964 en la tarjeta de reserva 950.

El RCM 920 en la tarjeta activa 910 se comunica con el RCM 960 en la tarjeta de reserva 950 para “reflejar” la información de la tarjeta activa 912 en la tarjeta de reserva 950 en cada instante de tiempo y sincronizar los procesos de la tarjeta activa 910 con la tarjeta de reserva 950. El RCM 920 es el responsable del movimiento selectivo de datos a la tarjeta de reserva 950. En particular, el RCM 920 es responsable de la actualización de transacción individual pequeña, que se denomina actualización “delta”, y de la actualización de transacción grande, que se denomina actualización “masiva”.

Comunicación entre tarjetas

Una capa física soporta la comunicación entre tarjetas entre la tarjeta activa 910 y la tarjeta de reserva 950. Por ejemplo, un enlace Ethernet puede usarse para soportar la comunicación entre la tarjeta activa 910 y la tarjeta de reserva 950. El RCM 920 en la tarjeta activa 910 puede comunicarse con el RCM 960 en la tarjeta de reserva 950 empleando dicho enlace.

Los RCM 920 y 960 son módulos de software que se ejecutan sobre las comunicaciones entre tarjetas. Los gestores de control de redundancia (RCM) determinan el papel y la maestría en sus tarjetas respectivas. Los RCM también pueden comunicarse con lógica de software para determinar el papel y la maestría. Los RCM soportan la transferencia de la actualización de la tarjeta de reserva 950 con información consistente en la tarjeta activa 910. Por ejemplo, los RCM controlan las actualizaciones grandes “masivas” y las actualizaciones pequeñas incrementales “delta” entre la tarjeta activa 910 y la tarjeta de reserva 950. Las actualizaciones masivas se realizan normalmente si se ha insertado una nueva tarjeta de redundancia. Las actualizaciones delta se consideran actualizaciones de reserva y se realizan cuando

ES 2 267 850 T3

se hacen cambios individuales en la tarjeta activa 910. Para una realización, el RCM 960 en la tarjeta de reserva 950 puede realizar algunas transacciones más atrás para facilitar un proceso de actualización más eficaz.

5 La plataforma 900 subyacente soporta la comunicación entre tarjetas por lo que la replicación de información entre la tarjeta activa 910 y la tarjeta de reserva 950, es decir las actualizaciones masivas y las actualizaciones delta, pueden tener lugar. La comunicación entre tarjetas puede facilitar el acuse de recibo de mensajes entre la tarjeta activa 910 y la tarjeta de reserva 950 y mantener la consistencia de la información en la tarjeta activa 910 y la tarjeta de reserva 950.

10 *Maestría*

Una determinación de maestría se produce si hay un sistema redundante (es decir, una tarjeta de reserva insertada para el nodo o encaminador 104). En una configuración de una única tarjeta, por ejemplo con sólo una tarjeta activa 910, la tarjeta activa 910 obtiene automáticamente la maestría del encaminador 104. Para una realización, la tarjeta activa 910 determina automáticamente si está presente una tarjeta de reserva 950 en un encaminador 104. Si no está presente ninguna tarjeta redundante, la tarjeta activa 910 adopta la maestría. Sin embargo, si la tarjeta activa 910 determina que hay una tarjeta redundante, se realiza una determinación para decidir la maestría entre la tarjeta activa 910 y la tarjeta de reserva 950. Esta determinación puede realizarse en software y/o hardware. Por ejemplo, la lógica de arbitraje o la lógica de maestría puede residir tanto en la tarjeta activa 910 como en la tarjeta de reserva 950 empleando una serie de técnicas para determinar la maestría. Por ejemplo, la lógica de arbitraje puede determinar la maestría basándose en la ID de ranura de las tarjetas. Por ejemplo, la tarjeta insertada en la ID de ranura 1 puede determinarse como la “tarjeta activa” y la tarjeta insertada en la ID de ranura 2 puede determinarse como la “tarjeta redundante”. Alternativamente, la lógica de arbitraje puede comprobar un número de ID en la tarjeta para ver si coincide con un número de ID del encaminador. Si coincide, esa tarjeta se convertirá en la “tarjeta activa”. Como resulta evidente, pueden usarse varias técnicas para determinar la maestría.

Secuencia de arranque (Boot)

En una configuración de una única tarjeta de control, la maestría pertenece a la tarjeta de control. Sin embargo, en un sistema redundante, se realiza una determinación en el momento de arranque para determinar qué tarjeta será la “activa”. La figura 10A ilustra un diagrama de flujo de una operación 1000 de una secuencia de arranque según una realización.

En referencia a la figura 10A, en la operación 1002, se realiza una determinación para ver cual de las tarjetas en el sistema redundante va a funcionar como activa o de reserva. Para una implementación, si una de las tarjetas era activa antes de la secuencia de arranque, se le da la maestría y se convierte en activa. Para otra implementación, las ID de ranura se usan para determinar qué tarjeta es activa o de reserva.

En la operación 1004, la tarjeta determinada como activa pide la maestría. El mismo proceso puede tener lugar en la otra tarjeta, es decir, que ambas tarjetas pueden ejecutarse al mismo tiempo y cada tarjeta puede ejecutar un proceso para obtener la maestría. Puede usarse cualquier número de algoritmos o técnicas de maestría o selección (por ejemplo, una técnica de número aleatorio) para romper el empate en la situación en la que ambas tarjetas eran activas antes y están preparadas para ser activas.

En la operación 1004, la lógica de arbitraje concede la maestría a una de las tarjetas 910 ó 950. Para la finalidad de la explicación, la lógica de arbitraje en la tarjeta activa 910 concede la maestría a la tarjeta activa 910 y la lógica de arbitraje en la tarjeta de reserva 950 le confiere un estado de reserva.

Suponiendo que la tarjeta activa 910 funciona como activa y la tarjeta de reserva 950 funciona como reserva, la tarjeta de reserva 950 necesita actualizarse para reflejar la tarjeta activa 910 durante la secuencia de arranque. Es decir, la tarjeta activa 910 se inicia y el SRM 918 lee la información de configuración y de estado en la tarjeta activa 910, que empieza a funcionar de manera correspondiente en el estado activo. El SRM 958 en la tarjeta de reserva 950 también leerá la información de configuración y de estado y funcionará en un estado de reserva e informará al SRM 918 de que está preparada y en un estado de reserva. La figura 10B ilustra un diagrama de flujo de una operación 1010 para una secuencia de arranque según otra realización.

En referencia a la figura 10B, en la operación 1012, la tarjeta activa 910 realiza una copia masiva de todos sus contenidos almacenados en las estructuras de datos RAM 912, disco de memoria flash 924 y datos no persistentes 914 a su par en la tarjeta de reserva 950 empleando la tarea de aplicación 916, el SRM 918 y el RCM 920. Por ejemplo, las aplicaciones de protocolo de encaminamiento que se ejecutan en la tarjeta activa 910 empezarán a copiar toda la información pertinente en las estructuras de datos RAM 912 a su par 952 en la tarjeta de reserva 950. Para una realización, sólo deben copiarse los datos esenciales tales como datos de estado, datos de conexión, bases de datos de rutas privadas, etc. Además, los datos no esenciales tales como información de contabilización no tienen que copiarse.

En la operación 1014, la tarjeta activa 910 realiza una copia incremental de cualquier nueva información o datos generados a partir de nuevos mensajes de encaminamiento que llegan desde los nodos pares. La nueva información debe copiarse a la tarjeta de reserva 950, independientemente de si se ha completado la operación de copia masiva. Para una implementación, puede usarse una técnica de marcado y barrido para determinar qué información se acaba

ES 2 267 850 T3

de generar. Para otra implementación, no se permite el copiado incremental y masivo concurrente. En tal caso, la plataforma 900 de redundancia puede incluir cualquier número de colas para realizar la actualización incremental después de una operación de copia masiva. Para garantizar una información de protocolo de encaminamiento válida y consistente, el almacén de datos 922 realiza una copia redundante del disco de memoria flash 924 en el almacén de datos 962 y el disco de memoria flash 964 en la tarjeta de reserva 950 antes de que pueda realizarse ningún cambio en el disco de memoria flash 924.

En la operación 1016, tanto la tarjeta activa 910 como la tarjeta de reserva 950 implementan una verificación de la consistencia. Por ejemplo, cada protocolo de encaminamiento necesita implementar una verificación de que se ha copiado o replicado la información consistente desde la tarjeta activa 910 a la tarjeta de reserva 950. Una vez verificada la consistencia, los protocolos de encaminamiento puede declararse a sí mismos redundantes en la tarjeta de reserva 950 y si los protocolos de encaminamiento se declaran a sí mismos redundantes, el encaminador 104 se declara redundante.

15 *Cambio de conexión de cortesía*

Un cambio de conexión de cortesía se refiere a un cambio de conexión que se inicia por un usuario o un software de una manera controlada. Por ejemplo, un usuario puede iniciar una orden en una línea de órdenes para conmutar el funcionamiento desde una tarjeta activa 910 hasta una tarjeta de reserva 950 para el nodo 104. En un cambio de conexión de cortesía, el SRM 918 es consciente de un cambio de maestría a la tarjeta de reserva 950. Por tanto, el SRM 918 prepara el cambio de conexión de una manera más controlada y la tarjeta activa 910 puede ceder de manera suave el control del encaminador 104 a la tarjeta de reserva 950. El SRM 918 mantiene un mapa de bits de todas las funciones críticas que son necesarias para que la tarjeta activa 910 funcione en un “estado activo”. Las tareas a través de la tarea de aplicación 916 envían sus estados al SRM 918 en la tarjeta de reserva 950 para un cambio de conexión sin interrupciones y rápido.

Las figuras 11A y 11B ilustran un diagrama de flujo de una operación 1100 para que una tarjeta activa 910 realice un cambio de conexión de cortesía a una tarjeta de reserva 950 según una realización. En la siguiente operación, el SRM 918 y el RCM 920 de la tarjeta activa 910 pueden usarse para facilitar el cambio de conexión rápido y suave.

En referencia a las figuras 11A y 11B, en la operación 1102, la tarjeta activa 910 verifica que un cambio de conexión no está bloqueado comprobando si la tarjeta de reserva 950 está insertada en el encaminador 104 o comprobando si la tarjeta de reserva 950 no está fuera de línea. La tarjeta de reserva 950 puede ponerse en modo “fuera de línea” con fines de actualización de versión o por otros motivos, lo que elimina efectivamente la funcionalidad de redundancia. El mecanismo de comunicación entre tarjetas puede usarse para verificar si la tarjeta de reserva 950 está en línea o fuera de línea. Por ejemplo, el SRM 918 puede pedir el estado de la tarjeta de reserva 950. Los estados pueden estar relacionados con un estado de tarjeta, tipo de tarjeta, hardware, software y sumas de control de base de datos. Así, si el estado devuelto es “fuera de línea”, la tarjeta activa 910 rechazará un cambio de conexión.

En la operación 1104, la tarjeta activa 910 verifica que la tarjeta de reserva 950 no está fuera de línea.

En la operación 1106, el SRM 918 informa al panel de control o a las capas de aplicación del encaminador 104 que se está realizando una conmutación.

En la operación 1108, la tarjeta activa 910 bloquea cualquier nueva actualización al disco de memoria flash 924. Por ejemplo, el SRM 918 y el almacén de datos 922 pueden bloquear cualquier actualización a las estructuras de datos RAM 912, los datos no persistentes 914 y el disco de memoria flash 924. Para una realización, el SRM 918 puede enviar un mensaje a la tarea de aplicación 916 de que se impide a todas las tareas realizar una actualización tal como, por ejemplo, un cambio en una tabla de encaminamiento en el disco de memoria flash 924. El bloqueo bloqueará también cualquier nueva actualización de la tarjeta de reserva 950, lo que es obligatorio para un cambio de conexión. Para otra realización, la tarea de aplicación 916 puede determinar si ciertos cambios de datos en las estructuras de datos RAM 912 y datos no persistentes 914 que no son fatales para un cambio de conexión no deberían bloquearse.

En la operación 1110, el SRM 918 y el RMC 920 replican información en las estructuras de datos RAM 912, datos persistentes 914 y disco de memoria flash 924 a su par en la tarjeta de reserva 950 para completar las actualizaciones de bases de datos.

En la operación 1112, el SRM 918 verifica que la información de configuración en la tarjeta de reserva 950 sea idéntica a la tarjeta activa 910. Por ejemplo, el SRM 958 puede intercambiar sumas de control en su información de la base de datos en la tarjeta de reserva 950 con la tarjeta activa 910. Para una realización, si las sumas de control no coinciden, el SRM 958 replicará bases de datos en la tarjeta activa 918 de nuevo y realizará nuevamente el proceso de verificación. Si las sumas de control no coinciden la segunda vez, el cambio de conexión no tendrá lugar.

En la operación 1114, el SRM 918 informa a la tarjeta de reserva 950 que esté preparada para activarse después de que la replicación se haya completado. La tarjeta activa 910 informa a la tarjeta de reserva 950 de que se prepare para “activarse”. Para una implementación, la tarjeta activa 910 envía un mensaje a la tarjeta de reserva 950 para que se prepare para activarse. La tarjeta activa 910 puede así esperar a un acuse de recibo (es decir, “preparada” o “no preparada”). Si la tarjeta activa 910 responde con un “no preparada”, el cambio de conexión se cancela.

ES 2 267 850 T3

En la operación 1116, el SRM 918 informa selectivamente a la tarea de aplicación 916 de que ciertas tareas que se están ejecutando van a pasar a la reserva. Por ejemplo, el SRM 918 enviará un mensaje a un grupo de selección de tareas y es necesario un acuse de recibo de este mensaje. El mensaje es para informar las tareas que está parando la tarjeta activa 910 hasta un estado de reserva.

5

En la operación 1118, la tarjeta activa 910 cede su maestría del encaminador 104. Por ejemplo, el SRM 918 puede llamar a un controlador (“driver”) I/O para destituir un estado “maestro” para ceder la maestría. El hardware de la tarjeta activa 910 cede entonces el control inmediato o maestría a la tarjeta de reserva 950. Como tal, esta activación reenviará todos los datos procedentes del encaminador 104 a la tarjeta de reserva 950.

10

En la operación 1120, el SRM 918 informa a la tarea de aplicación 916 de que el resto de tareas pasen a reserva. Es decir, las funciones de algunas tareas cambian cuando la tarjeta cambia de estado. Dicho cambio puede propagarse a todas las tareas.

15

En la operación 1122, el SRM 918 pregunta a las tareas por sus estados y espera hasta que las tareas pasan a reserva. Esta operación se requiere principalmente para tareas que están en estado activo podrían ser funciones de ejecución que sólo debería poder hacer la tarjeta activa 910, por ejemplo responder a la estación de gestión o transmitir datos en los puertos de enlace ascendente/acceso, responder a alarmas de línea, etc. Mediante este cuestionario/saludo, se garantiza que todas las tareas críticas pasen a estado de reserva.

20

En la operación 1124, la tarjeta activa 910 establece comunicación con la tarjeta (de reserva 950) activa para la sincronización de las bases de datos tanto para la información persistente como la no persistente. Una vez sincronizadas las bases de datos, el estado de la tarjeta activa 910 está listo para el estado de reserva.

25

En la operación 1126, el SRM 918 elimina el bloqueo puesto a la tarea de aplicación 916 y al almacén de datos 922.

En la operación 1128, la tarjeta activa 910 se pasa al estado de reserva.

30

La operación anterior se refiere a eventos de la tarjeta activa 910. Los eventos de la tarjeta de reserva 950 se describen con respecto a las figuras 12A y 12B.

35

Las figuras 12A y 12B ilustran un diagrama de flujo de una operación 1200 para que una tarjeta de reserva 950 realice el cambio de conexión de cortesía según una realización. En la siguiente operación, el SRM 958 y el RMC 960 de la tarjeta de reserva 950 pueden usarse para facilitar un cambio de conexión rápido y suave. En referencia a las figuras 12A y 12B, en la operación 1202, el SRM 958 proporciona información sobre el estado de la tarjeta e información de auto-test a la tarjeta activa 910 a través del SRM 918.

40

En la operación 1204, el RCM 960 actualiza las bases de datos pares en la tarjeta de reserva 950 desde la tarjeta activa 910.

En la operación 1206, el SRM 958 proporciona sumas de control de las bases de datos para la tarjeta activa 910 mediante el SRM 958.

45

En la operación 1208, el SRM 958 informa a la tarea de aplicación par 956 que ciertas tareas deben prepararse para el estado activo.

50

En la operación 1210, la tarjeta de reserva 950 obtiene la maestría del encaminador 104 en el momento en que la tarjeta activa 910 cede la maestría.

55

En la operación 1212, el SRM 958 informa a la tarea de aplicación par 956 de que el resto de tareas deben presentar un estado activo. El SRM 958 también actualiza la información de estado en la tarjeta de reserva 950 que tiene la maestría.

60

En la operación 1214, la tarea de aplicación par 956 pregunta a las tareas por su estado y espera hasta que están en el estado activo. Para una implementación, algunas tareas pueden reiniciarse como activas en caso necesario.

En la operación 1216, el SRM 958 bloquea nuevas actualizaciones de red a las estructuras de datos RAM 952, datos no persistentes 954 y disco de memoria flash 964 en la tarjeta de reserva 950.

65

En la operación 1218, la tarjeta de reserva 950 cambia su estado a activo.

En la operación 1220, el SRM 958 informa al panel de control o a la capa de aplicaciones ejecutada en la tarjeta de reserva 950 del cambio de conexión.

65

En la operación 1222, el SRM 958 espera hasta que la otra tarjeta (la tarjeta activa 910) esté en estado de reserva.

ES 2 267 850 T3

En la operación 1224, el SRM 958 y/o el RCM 960 verifican que los datos de las bases de datos de la tarjeta de reserva 950 son consistentes con las bases de datos de la tarjeta activa 910.

5 En la operación 1226, el SRM 958 elimina la condición de bloqueo a las bases de datos en la tarjeta de reserva 950.

Cambio de conexión de no cortesía

10 Un cambio de conexión de no cortesía se refiere a un cambio de conexión que se inicia por un fallo en la tarjeta activa 910 sin aviso. Por ejemplo, una tarjeta activa 910 puede fallar por cualquier razón de hardware o de software tal como se explicará más detalladamente más adelante. El cambio de conexión de no cortesía es muy similar al cambio de conexión de cortesía salvo porque no ha preparación para el cambio de conexión. Es decir, el cambio de conexión podría producirse en cualquier momento temporal para un sistema redundante y las actualizaciones de las bases de datos podrían estar pendientes o las bases de datos podrían estar, por ejemplo, a mitad de una actualización de una tabla de encaminamiento o FIB. Además puede perderse parte de información. Para una realización, puede implementarse un mecanismo de recuperación para recuperar información perdida.

20 La figura 13 ilustra un diagrama de flujo de una operación 1300 para que una tarjeta de reserva 950 realice un cambio de conexión de no cortesía según una realización. La siguiente operación 1300 está relacionada con la tarjeta de reserva 950. Para una realización, la tarjeta activa 910 realiza la operación 1100 en la figura 11 para el cambio de conexión de no cortesía.

25 En relación a la figura 13, en la operación 1302, el SRM 958 determina que la otra tarjeta (la tarjeta activa 910) no posee la maestría del encaminador 104.

30 En la operación 1304, el SRM 958 informa a la tarea de aplicación par 956 de que todas las tareas deben “activarse” y de que el estado de la tarjeta de reserva 950 debe actualizarse como “activo”.

35 En la operación 1306, el SRM 958 pregunta a las tareas por sus estados, y espera hasta que se vuelven “activos”.

En la operación 1308, el estado de la tarjeta activa 910 cambia a “no activo” o “en reserva”.

40 En la operación 1310, la tarjeta de reserva 950 informa al panel de control o a la capa de aplicaciones de que la tarjeta de reserva 950 posee la maestría del encaminador 104 y de que se ha producido un cambio de conexión.

Sincronización de la redundancia de la tarjeta de reserva

45 Hay dos componentes principales que necesitan sincronizarse para que la redundancia funcione con la tarjeta activa 910 y la tarjeta de reserva 950. En primer lugar, debe haber sincronización de la “información persistente”, que se almacena en el disco de memoria flash 924. La información persistente puede incluir, por ejemplo, información de configuración y actualizaciones relativas a archivos asociados, registros, informes, etc. En segundo lugar, debe haber una sincronización de “información no persistente”, que se almacena en RAM (por ejemplo, estructuras de datos Ram 912 y datos no persistentes 914). La información no persistente incluye por ejemplo, tablas de encaminamiento, conexiones de sesión, etc. Cualquier tarea que se ejecute en el encaminador 104 tiene redundancia como parte de su diseño. Es decir, cualquier tarea se centra en la “duplicación de información persistente” y en la “duplicación de información no persistente”. Los módulos de software de almacén de datos en la tarjeta activa 910 y en la tarjeta de reserva 950 colaboran en la duplicación o actualización persistente y no persistente.

Actualización persistente

50 La figura 14 ilustra un diagrama de flujo de una operación 1400 para actualizar información persistente en una tarjeta de reserva 950. La operación 1400 se refiere a cualquier tarea que invoque el almacén de datos 922 en la tarjeta activa 910 para guardar la información de configuración en el disco de memoria flash 924. El almacén de datos 922, después de actualizar el disco de memoria flash local 924, envía un mensaje a través de su almacén de datos par 962 en la tarjeta de reserva 950 para copiar la misma información en el disco de memoria flash 964. Además, el mensaje al almacén de datos 962 puede enviarse a la tarea de aplicación par 956 para actualizar las estructuras de datos RAM 952 y los datos no persistentes 954 con la nueva información de configuración. Para una realización, puede haber acuses de recibo positivos o negativos para todas las transacciones.

60 En referencia a la figura 14, en la operación 1402, la tarea de aplicación 916 envía un mensaje al almacén de datos 922 para guardar la información de configuración que se ha hecho en el disco de memoria flash 924. Por ejemplo, el mensaje puede contener una identificación de registro para la actualización. Alternativamente, puede haber múltiples actualizaciones de registro para una única transacción.

65 En la operación 1404, el almacén de datos 924 actualiza el disco de memoria flash 924 con la información de configuración. Después de actualizar el disco de memoria flash local 924, el almacén de datos 924 envía un acuse de recibo a la tarea de aplicación 916 de que el disco de memoria flash 924 se ha actualizado. El almacén de datos 924 envía entonces el mismo mensaje a su almacén de datos par 962 en la tarjeta de reserva 950 a través del RCM 920 y 960.

ES 2 267 850 T3

En la operación 1406, el almacén de datos 962 en la tarjeta de reserva 950 actualiza el disco de memoria flash 964 con la misma actualización que el disco de memoria flash 924 en la tarjeta activa 910. El almacén de datos 962 envía el mismo conjunto de mensajes a la tarea de aplicación par 956 que se envió a la tarea de aplicación 916 por el almacén de datos 922 en la tarjeta activa 910.

5

En la operación 1408, la tarea de aplicación par 956 interpreta los mensajes y actualiza las estructuras de datos RAM 952 y datos no persistentes 954 en consecuencia.

Actualización no persistente

10

La figura 15 ilustra un diagrama de flujo de una operación 1500 para actualizar la información no persistente en la tarjeta de reserva 950. La operación 1500 se refiere a una tarea de aplicación 916 que transmite un cambio de configuración a sus estructuras de datos RAM 912 y datos no persistentes 914 en la tarjeta activa 910 y un mensaje a su tarea de aplicación par 956 en la tarjeta de reserva 950 para actualizar sus estructuras de datos RAM 952 y datos no persistentes 954 respectivos. Para una implementación, pueden usarse acuses de recibo positivos o negativos en todas las transacciones.

15

En referencia a la figura 15, en la operación 1502, la tarea de aplicación 916 actualiza sus estructuras de datos RAM 952 o datos no persistentes 954 y envía un mensaje al RCM 920 para transmitir el mensaje a la tarjeta de reserva 950.

20

En la operación 1504, el RCM 920 transmite el mensaje a la tarea de aplicación par 956 a través del RCM 960. El RCM 920 transmite el mismo conjunto de mensajes a la tarea de aplicación par 956 que el que transmitió la tarea de aplicación 916 en la tarjeta activa.

25

En la operación 1506, la tarea de aplicación par 956 actualiza sus estructuras de datos RAM 952 y datos no persistentes 954 en consecuencia. La tarea de aplicación par 956 puede enviar un acuse de recibo a la tarea de aplicación 916 en la tarjeta activa 910 de que ha realizado el cambio.

30

En la operación 1508, el SRM 918 en la tarjeta activa 910, al recibir el acuse de recibo, desbloquea el proceso de bloqueo que puede haberse implementado para impedir cambios en las estructuras de datos RAM 912, datos no persistentes 914 y disco de memoria flash 924.

Las operaciones de actualización anteriores puede realizarse para dos tipos de actualización en la tarjeta de reserva 950, que se denominan “actualizaciones masivas” y “actualizaciones delta (pequeñas)”.

35

Actualizaciones masivas

Una actualización masiva se refiere a una actualización de la tarjeta de reserva 950 cuando la tarjeta de reserva 950 se inserta por primera vez en el encaminador 104 y la tarjeta activa 910 ya estaba funcionando en el modo activo. La tarjeta recién insertada puede ser una tarjeta nueva o una tarjeta libre o una tarjeta de otro encaminador. Los SRM que funcionan en la tarjeta activa 910 y en la tarjeta de reserva 950 determinan la validez y el estado de sí mismos determinando el estado activo o el estado de reserva. Para la finalidad de la explicación, la tarjeta de reserva 950 que se inserta obtiene un estado de reserva. Si la tarjeta insertada es funcionalmente capaz de pasar a ser de reserva, la tarjeta de reserva 950 debe sincronizarse con la tarjeta activa 910.

40

En este punto tiene lugar una gran “actualización masiva” entre la tarjeta activa 910 y la tarjeta de reserva 950 puede presentar una gran tabla de encaminamiento funcionando en su interior. En particular, la actualización se denomina “masiva” porque debe copiarse toda la información en la tarjeta activa 910 a la tarjeta de reserva 950, que puede incluir millones de entradas de encaminamiento.

50

Al inicio de la actualización masiva, el SRM 918 bloquea cualquier cambio en la configuración de la tarjeta activa 910 tal como, por ejemplo, cambios en la línea de órdenes o cambios en el establecimiento de sesiones. Para una realización, las actualizaciones de ruta de red pueden quedar en cola porque los cambios en la topología de la red deben conocerse siempre, incluso durante una actualización masiva.

55

El almacén de datos 922 en la tarjeta activa 910 garantizará que las bases de datos que almacenan los datos persistentes (es decir, la información almacenada en el disco de memoria flash 924) se refleje en una memoria o disco de memoria flash 964 de la tarjeta de reserva 950. Por ejemplo, una tabla de encaminamiento, información de conexión, etc. se reflejan en la tarjeta activa 910 y en la tarjeta de reserva 950. El RCM 920 informa a los módulos en la tarjeta activa 910 de la existencia de la tarjeta de reserva 950. Para una realización, la actualización puede ser diferente para cada aplicación que pueda usarse para diferentes tipos de protocolos de encaminamiento ejecutados en la tarjeta activa 910.

60

Actualizaciones delta

En la tarjeta activa 910 pueden producirse dos tipos de cambios. En primer lugar, un cambio que es necesario replicarse o actualizarse en la tarjeta de reserva 950. Por ejemplo, es necesario actualizar los cambios de configuración,

65

ES 2 267 850 T3

una actualización de la tabla de encaminamiento, cambios en el nombre del nodo, etc. en la tarjeta de reserva 950. En segundo lugar, un cambio que no es necesario actualizar en la tarjeta de reserva 950. Por ejemplo, no es necesario actualizar las actualizaciones de contabilización o alarmas no críticas en la tarjeta de reserva 950. Sin embargo, tales cambios pueden replicarse a la tarjeta de reserva 950.

La actualización delta puede realizarse empleando las operaciones descritas anteriormente en relación a la actualización persistente y no persistente. Para una realización, incluso si una actualización delta falla y la tarjeta de reserva 950 reanuda el funcionamiento, la falta de la actualización delta no hará necesariamente que la tarjeta de reserva 950 falle porque un nodo par volverá a enviar el mensaje. Mientras que el mensaje no se haya llevado a cabo, la tarjeta de reserva 950 no requerirá necesariamente que la actualización delta reanude la operación si la tarjeta activa 910 falla. Es decir, si un mensaje o cambio se lleva a cabo por la tarjeta activa 910, debe realizarse en la tarjeta de reserva 950 en una actualización delta para mantener la consistencia con los nodos pares.

Tratamiento de errores

La tarjeta activa 910 incluye capacidades de tratamiento de errores de software y hardware. Por ejemplo, el SRM 918 puede tratar errores de software y la lógica de errores puede tratar errores de hardware para la tarjeta activa 910. La figura 16 ilustra un diagrama de flujo de una operación 1600 para realizar el tratamiento de errores según una realización.

En referencia a la figura 16, en la operación 1602, la tarjeta activa 910 detecta un error, tal como un error de software o de hardware. La tarjeta activa 910 detecta si el error requiere un cambio de conexión. Si el error requiere un cambio de conexión, la tarjeta activa 910 y la tarjeta de reserva 950 pueden llevar a cabo un cambio de conexión no de cortesía tal como se ha descrito anteriormente.

En la operación 1604, tiene lugar un cambio de conexión de hardware a la tarjeta de reserva 950. Se requiere un cierto periodo de tiempo para realizar el cambio de conexión físico, que se encuentra en el orden de unos pocos milisegundos.

En la operación 1606, la tarjeta de reserva 950 reanuda el funcionamiento del encaminador 104. La tarjeta de reserva 950 debe reanudar el funcionamiento rápidamente porque puede expirar el tiempo máximo de una sesión de protocolo con el encaminador 104. Gracias a la actualización delta de información no persistente y persistente para cambios de información relevantes de la tarjeta activa 910, la tarjeta de reserva 950 puede reanudar el funcionamiento sin interrupciones y rápidamente.

Fallos de software/hardware

Un fallo de software es el tipo de fallo más crítico. Es decir, los errores de software están relacionados con un número de estados y variables de software que requieren consistencia en el sistema de redundancia. Además, los errores de software pueden ser difíciles de detectar. Los tipos más comunes de fallos de software incluyen defectos de segmentación, daños de la memoria, agotamiento de la memoria, cambio de conexión forzado por la aplicación, y bucles infinitos.

Un defecto de segmentación se produce si hay un acceso no válido a la memoria. Si hay un acceso erróneo a la memoria, el hardware o el software puede detectar el acceso erróneo y generar un error hacia un SRM para provocar un cambio de conexión. En particular, un acceso erróneo a la memoria puede provocar el almacenamiento de información incorrecta, lo que puede crear el almacenamiento de información inconsistente en la tabla de encaminamiento. Un error de agotamiento de la memoria se produce si se está usando demasiado espacio en la memoria. Para una implementación, puede facilitarse un aviso si los espacios de memoria usados alcanzan un nivel de aviso, y puede producirse un cambio de conexión si el espacio usado de la memoria sobrepasa un umbral determinado.

Un cambio de conexión forzado por la aplicación puede producirse en el software cuando un usuario fuerza el cambio de conexión mediante una instrucción en la línea de órdenes. Por ejemplo, insertando una nueva tarjeta que debe presentar el estado activo. Un bucle infinito también puede provocar que un procesador procese otras instrucciones. Para una realización, pueden usarse temporizadores de vigilancia para determinar si una instrucción es un bucle infinito que provoca un error de software. Alternativamente, una tarea de poca prioridad puede usarse para determinar si el procesador está atascado procesando de manera infinita otra tarea. Es decir, si la tarea de poca prioridad nunca obtiene un tiempo de procesamiento, puede determinarse un bucle infinito.

Un fallo de hardware es menos grave que un fallo de software gracias al hardware redundante en la tarjeta de reserva 950. Tipos comunes de fallos de hardware son los fallos de diagnóstico ASIC, los fallos de bus, los fallos de memoria, o un fallo de tarjeta durante una secuencia de activación o de arranque. Este tipo de fallos de hardware también provocará que una tarjeta activa ceda la maestría y provoque un cambio de conexión a la tarjeta de reserva 950.

ES 2 267 850 T3

Redundancia de protocolo de encaminamiento

Requisitos básicos de la redundancia de protocolos de encaminamiento

5 Para presentar una redundancia en el nivel del protocolo de encaminamiento, la tarjeta de reserva 950 debe poblarse con toda la información pertinente requerida para cada uno de los protocolos de encaminamiento ejecutados en la tarjeta activa 910. Cada módulo de protocolo de encaminamiento ejecutado en la tarjeta activa 910 y en la tarjeta de reserva 950 es responsable de mantener una copia idéntica de su información de protocolo tanto en la tarjeta activa 910 como en la tarjeta de reserva 950. Así, si la tarjeta activa 910 falla, la tarjeta de reserva 950 puede reanudar todas las sesiones de los protocolos de encaminamiento de la tarjeta activa 910. La tarjeta de reserva 950 reanuda el funcionamiento antes de que cualquiera de los estados de sesión de protocolo de encaminamiento supere su tiempo máximo, evitando así que el fallo se observe en los nodos pares.

Interacción de protocolo de encaminamiento ilustrativa

15 La figura 17 ilustra un diagrama 1700 para mostrar la interacción de protocolo de encaminamiento en un nodo según una realización. En referencia a la figura 17, cada protocolo de encaminamiento BGP 1726, OSPF 1724 y IS-IS 1714 está asociado a su propia base de datos 1731, 1732 y 1733, respectivamente. Estas bases de datos pueden incluir información o rutas de protocolo de encaminamiento específicas. Además, las bases de datos 1731, 1732 y 1733 para cada protocolo de encaminamiento pueden almacenar estructuras de datos para dispositivos de estado y estadísticas que funcionan sin protocolos de encaminamiento BGP 1726, OSPF 1724 y IS-IS 1714.

20 La suma de todas las rutas agregadas se almacena en una tabla de encaminamiento IP 1702. La tabla de reenvío (FIB) 1702 puede generarse basándose en las rutas de la tabla de encaminamiento IP 1702. La FIB 1716 puede incluir información de reenvío para reenviar paquetes para el encaminador 104. Para una realización, una tercera memoria de contenido direccionable (TCAM) 1706, puede almacenar las rutas en la FIB 1716. En otras realizaciones, cualquier combinación de un procesador y un sistema de memoria puede usarse para almacenar y mantener la FIB 1716.

25 Las interacciones de protocolo de encaminamiento anteriores deben funcionar de la misma manera en la tarjeta de reserva 950 en caso de que la tarjeta activa 910 falle. Así, la información en las bases de datos 1731, 1732, 1733 para BGP 1726, OSPF 1724 y IS-IS para la tarjeta activa 910 se replican a las mismas en la tarjeta de reserva 950. Además, la tabla de encaminamiento IP 1702 y FIB 1716 se replican a la misma en la tarjeta de reserva 950. Por tanto, la TCAM 1706 para la tarjeta de reserva 950 conmutará y reenviará paquetes empleando información de reenvío consistente para el encaminador 104.

30 La figura 18 ilustra un diagrama 1800 para mostrar la interacción de protocolo de encaminamiento entre un punto de control activo y un punto de control de reserva según otra realización. En referencia a la figura 18, se mantiene una copia de la información de encaminamiento en las bases de datos 1731A, 1732A y 1733A para protocolos de encaminamiento BGP 1726A, OSPF 1724A y IS-IS 1714A, respectivamente, en sus bases de datos pares en el punto de control de reserva.

35 Para una realización, para evitar sobrecargar la conexión entre la tarjeta activa 910 y la tarjeta de reserva 950, la tarjeta activa 910 sólo replicará o copiará rutas estáticas 1820A a sus rutas estáticas pares 1820B en el punto de control de reserva. Las rutas estáticas son rutas nativas. La tarjeta de reserva 950 se ejecuta de la misma manera que la tarjeta activa 910 salvo porque no tiene acceso a los puertos usados para la comunicación con la red 100 para el encaminador 104. En particular, la vía de código regular en la tarjeta de reserva 950 sólo redistribuirá y poblará su tabla de encaminamiento IP 1716.

40 En el caso de un fallo del punto de control activo, el encaminador 104 cambiará la conexión del funcionamiento al punto de control de reserva. El encaminador 104 continuará encaminando el tráfico sin interrupciones porque el punto de control de reserva ha generado una tabla de reenvío válida a partir del almacén de datos privado de cada protocolo de encaminamiento. Como tal, un nodo par que se comunica con el encaminador 104 puede mantener la sesión de protocolo de encaminamiento con el encaminador 104. Además, el encaminador 104 puede evitar que los fallos se observen por el nodo par y evitar que los cambios de conexión se observen por el nodo par.

Arquitectura ilustrativa para la redundancia de protocolo de encaminamiento

45 La figura 19 ilustra una arquitectura 1900 ilustrativa para la redundancia de protocolo de encaminamiento. La arquitectura 1900 ilustrativa incluye bases de datos y módulos de protocolo de encaminamiento para la tarjeta activa 910 y una tarjeta de reserva 950 para soportar la redundancia de protocolo de encaminamiento. Cada uno de los módulos para la tarjeta activa 910 y la tarjeta de reserva 950 presenta dos tipos de bases de datos para la redundancia, que son las bases de datos redundantes (RDB, "redundant databases") y las bases de datos persistentes (PDB).

50 La tarjeta activa 910 incluye el módulo de protocolos de puerta de enlace interior (IGP) 1992A de acceso a las RDB IGP 1942A y RDB IGP 1924A. El módulo IGP incluye módulos para los protocolos OSPF, RIP e IS-IS. La tarjeta activa 910 también incluye el módulo BGP 1726A de acceso a RDB BGP 1927A y PDB BGP 1731A, el módulo TCP 1932A de acceso a RDB TCP 1933A y PDB TCP 1926A, y el módulo IP 1930A de acceso a RDB IP 1931A y PDB IP 1928A.

ES 2 267 850 T3

La tarjeta de reserva 950 incluye módulos pares de la tarjeta activa 910 de acceso a las PDB y RDB pares. En particular, la tarjeta de reserva 950 incluye el módulo IGP 1992B de acceso a RDB IGP 1942B y RDB IGP 1924B. La tarjeta de reserva 950 también incluye el módulo BGP 1726B de acceso a RDB BGP 1927B y PDB BGP 1731B, el módulo TCP 1932B de acceso a RDB TCP 1933B y PDB TCP 1926B, y un módulo IP 1930B de acceso a RDB IP 1931B y PDB IP 1928B.

La redundancia de datos persistentes (PDB) y datos no persistentes (RDB) se trata de manera diferente. La redundancia de datos persistentes se trata mediante un módulo de almacén de datos interno de la tarjeta activa 910 y la tarjeta de reserva 950. Si el módulo de almacén de datos de la tarjeta activa 910 almacena datos en un disco de memoria flash, los mismos datos se pasan de manera transparente a la tarjeta de reserva 950 y a su disco de memoria flash. Además, el módulo de almacén de datos par en la tarjeta de reserva 950 recibe notificación del cambio y los datos que se están cambiando pasan también como parte de la notificación. Cada uno de los módulos es también responsable de la redundancia de sus, y sólo sus datos no persistentes propios. Por ejemplo, si OSPF recibe alguna actualización desde BGP, OSPF no la pasará a la tarjeta de reserva 950. En esta situación, BGP enviará una actualización a su BGP par en la tarjeta de reserva 950.

La tarjeta activa 910 y la tarjeta de reserva 950 realizan la redistribución del encaminamiento de forma idéntica. Cada uno de los módulos de protocolo de encaminamiento en ambas tarjeta activa 910 y tarjeta de reserva 950 es responsable de enviar sus mejores rutas a los gestores de la tabla de encaminamiento (RTM) 1940A y 1940B, respectivamente. Si el RTM 1940A en la tarjeta activa 910 está configurado para redistribuir las rutas a otros protocolos, el RTM 1940A lo hará igual en la tarjeta de reserva 950 también. Para una realización, los ajustes de redistribución del RTM se consideran información de configuración y se hacen redundantes como datos persistentes.

Las tablas de información de reenvío (FIB) 1716A y 1716B se construyen de forma idéntica en la tarjeta activa 910 y en la tarjeta de reserva 950, respectivamente. Las FIB 1716A y 1716B se basan en las mejores rutas y en la distancia administrativa configurada para cada protocolo. Para una realización, la información de ajuste de distancia de protocolo para los RTM se considera información de configuración y se hace redundante como datos persistentes. Las terceras memorias de contenido direccionable (TCAM, “tertiary content addressable memories”) 1706A y 1706B funcionan de manera idéntica en la tarjeta activa 910 y en la tarjeta de reserva 950, respectivamente. Las TCAM 1706A y 1706B pueden programarse basándose en datos persistentes de diferentes módulos y datos FIB dinámicos de los RTM 1940A y 1940B. La tarjeta de reserva 950 (si actúa en modo de reserva) no presenta línea física conectada al encaminador 104. Como tal, la tarjeta de reserva 950 es responsable de ser un gestor de interfaz para tratar los estados de interfaz que deben estar en sincronización con la tarjeta activa 910. Es decir, la tarjeta de reserva 950 proporciona información a sus módulos y módulos pares en la tarjeta activa 910.

Redundancia de protocolo de puerta de enlace de frontera (BGP)

Requisitos básicos de la redundancia BGP

El BGP es el protocolo de encaminamiento de uso más extendido en Internet. El BGP es un protocolo de puerta de enlace externa usado por encaminadores de diferentes sistemas autónomos (SA). Un encaminador BGP encamina paquetes entre fronteras de red. Por tanto, una tabla de reenvío o encaminamiento BGP puede ser muy grande, capaz de almacenar millones de rutas. El BGP, sin embargo, ofrece un desafío diferente al de otros protocolos de encaminamiento. El BGP utiliza TCP para la conectividad y la transferencia fiable de datos. En consecuencia, si el BGP pierde su conexión TCP con un par, el par reacciona haciendo caer inmediatamente todas las rutas que ha aprendido desde ese par vecino. Por este motivo, para presentar una redundancia del protocolo de encaminamiento BGP, TCP también debe hacerse redundante para evitar que las rutas aprendidas por el BGP se vuelvan inaccesibles.

En las siguientes realizaciones, la plataforma 900 de redundancia mostrada en la figura 9 proporciona el soporte para presentar redundancia de protocolo BGP y TCP. La plataforma 900 de redundancia permite que una tarjeta de reserva 950 vuelva a aprender la información de reenvío y encaminamiento BGP en un corto periodo de tiempo suficiente para no constituir un corte del servicio. Además, la plataforma 900 de redundancia proporciona el soporte para realizar las siguientes operaciones para obtener la redundancia de protocolo BGP y TCP.

Requisitos de redundancia de nivel TCP

El nivel TCP de redundancia es un nivel adicional de redundancia para obtener la redundancia del protocolo de encaminamiento BGP. La siguiente realización ilustra la interacción a modo de ejemplo entre BGP y TCP para el nodo redundante 104.

La figura 20 ilustra un diagrama 2000 a modo de ejemplo para mostrar la interacción entre BGP, TCP e IP. Para la finalidad de la explicación, el diagrama 2000 se refiere a la tarjeta activa 910 en “modo activo”. En referencia a la figura 20, el diagrama 2000 muestra el BGP 1726A enviando tres mensajes al TCP 1932A. El mensaje 1 tiene una longitud de 19 bytes. El mensaje 2 tiene 70 bytes de longitud, y el mensaje 3 tiene 26 bytes de longitud con un número de total de 115 bytes.

El TCP 1932A es un protocolo de flujo continuo de bytes. El TCP 1932A considera los tres mensajes del BGP 1726A como un flujo continuo de bytes. Por ejemplo, el TCP 1932A puede enviar los 115 bytes del BGP 1726A como

ES 2 267 850 T3

dos mensajes al IP 1930A. El mensaje 1 con 85 bytes y el mensaje 2 con 30 bytes. Así, el IP 1930A puede recibir los dos mensajes del TCP 1932A. El IP 1932A puede almacenar temporalmente los dos mensajes en una memoria intermedia de transmisión 2002.

5 Dado que el TCP 1932A considera los mensajes como un flujo continuo de bytes, el TCP 1932A puede almacenar números secuenciales en el mensaje para indicar dónde está situado el mensaje en el flujo de bytes. Por ejemplo, el TCP 1932A puede almacenar un número enviado siguiente (NS) y un número recibido siguiente (NR) para determinar el orden de los mensajes. El número NS es un identificador que identifica el mensaje o paquete. El número NR es un identificador que identifica el siguiente mensaje o paquete en el flujo de bytes recibidos desde el par remoto. En
10 referencia a la figura 20, el mensaje 1 al IP 1930A puede presentar el número NS = 1000 generado de manera aleatoria. Si un nodo par recibe el mensaje 1 con un número NS = 1000, el nodo par sabe que el siguiente mensaje (es decir, el mensaje 2) debería presentar un NS = 1085. Si no es así, el nodo par sabrá que el mensaje está desordenado respecto al mensaje 1 y determinará que algo va mal.

15 Otro parámetro que puede usar el TCP 1932A es el tamaño de ventana. El tamaño de ventana es el número máximo de bytes que TCP 1932A puede enviar antes de recibir un acuse de recibo desde el nodo par. El parámetro del tamaño de ventana puede negociarse entre los pares. Por ejemplo, el tamaño de ventana puede ser de 8 K o 16 K. Para una realización, los datos que pasan desde y a través del BGP 1726A, TCP 1932A, IP 1930A, y la memoria intermedia de
20 transmisión 2002 se replican o copian en la tarjeta de reserva 950 hasta que se haya acusado recibo de los mensajes por el nodo par receptor.

La figura 21A ilustra un diagrama de flujo de una operación 2100 para replicar los cambios de estado BGP recibidos o generados según una realización. La siguiente operación 2100 puede implementarse por el nodo redundante 104 que presenta una tarjeta activa 910 y una tarjeta de reserva 950, tal como se indicó anteriormente. Para la finalidad de la
25 explicación, la operación 2100 empieza en la operación 2102. En la operación 2102, un cambio de estado BGP se recibe o genera por la tarjeta activa 910. Por ejemplo, un nodo par puede enviar al nodo redundante 104 que una ruta BGP ya no está disponible como un cambio de estado.

En la operación 2104, el cambio de estado BGP recibido o generado se replica desde la tarjeta activa 910 hasta
30 la tarjeta de reserva 950. Por ejemplo, la plataforma 900 de redundancia puede realizar una “actualización delta” del cambio de estado BGP en la tarjeta de reserva 950, tal como se explicó anteriormente.

La figura 21B ilustra un diagrama de flujo de una operación 2150 para replicar cambios de estado TCP recibidos o generados según una realización. Para la finalidad de la explicación, la operación 2150 empieza en la operación 2152. En la operación 2152, un cambio de estado TCP se recibe o genera por la tarjeta activa 910. Por ejemplo, un nodo par
35 puede enviar a un nodo redundante 104 que una conexión TCP se ha caído y ya no está disponible como un cambio de estado.

En la operación 2104, el cambio de estado TCP recibido o generado se replica desde la tarjeta activa 910 hasta
40 la tarjeta de reserva 950. Por ejemplo, la plataforma 900 de redundancia puede realizar una “actualización delta” del cambio de estado TCP en la tarjeta de reserva 950, tal como se explicó anteriormente.

Las operaciones 2100 y 2150 anteriores permiten la redundancia BGP y TCP para cambios de estado BGP y TCP. En otras realizaciones, las operaciones 2100 y 2150 anteriores pueden implementarse para proporcionar redundancia
45 selectiva para mensajes BGP y TCP. Es decir, algunos o todos los mensajes BGP y TCP pueden hacerse redundantes en la tarjeta de reserva 950.

Etapa de bloqueo TCP para mensajes BGP se están enviando

50 En una realización, un requisito para el nivel TCP de redundancia es un requisito de “etapa de bloqueo” BGP y TCP. El requisito de la etapa de bloqueo requiere que de cada mensaje que se envía o se recibe por una tarjeta activa 910 debe haber un acuse de recibo de que la tarjeta de reserva 950 ha almacenado el mensaje enviado o recibido antes de que la tarjeta activa 910 pueda enviar o recibir otro mensaje. Si se produce un cambio de conexión y la tarjeta de reserva 950 no guardó el mensaje, la redundancia se rompe.

55 La figura 22 ilustra un diálogo ilustrativo entre un TCP activo que funciona sobre una tarjeta activa 910 y un TCP de reserva que funciona sobre la tarjeta de reserva 950 para mostrar el requisito de la etapa de bloqueo para un mensaje BGP que está siendo enviado a un nodo par. En referencia a la figura 22, el diálogo 2200 muestra un BGP activo 1726A que envía un mensaje BGP (mensaje 1) al TCP activo 1932A. La tarjeta activa 910 se está preparando para enviar el
60 mensaje 1 a otro encaminador o nodo par. El TCP activo 1932A envía el mensaje 1 al TCP de reserva 1932B de modo que el mensaje 1 puede replicarse en la tarjeta de reserva 950. La tarjeta de reserva 950 envía un acuse de recibo al TCP activo 1932 de que ha replicado el mensaje 1.

65 Para una realización, el nodo redundante 104 que presenta una plataforma 900 de redundancia no enviará el mensaje 1 al nodo par a través de IP 1932A hasta que haya recibido un acuse de recibo de que el mensaje se ha replicado. Además, el nodo redundante 104 no enviará otro mensaje (es decir, el mensaje 2) hasta que haya recibido el acuse de recibo desde el TCP de reserva 1932B de que el mensaje 1 se ha guardado. Como se indicó anteriormente, si el mensaje 1 no se ha guardado en la tarjeta de reserva 950, la redundancia se romperá y si se produce un cambio

ES 2 267 850 T3

de conexión, la tarjeta de reserva 950 no podrá reanudar el funcionamiento en el estado actual de la tarjeta activa 910.

Por lo tanto, después de que el TCP activo 1932A recibe un acuse de recibo del mensaje 1, enviará el “ack” al BGP 1726A y entonces el BGP 1726A puede enviar el segundo mensaje 2. Asimismo, la tarjeta activa 910 no enviará el mensaje al nodo par a través de IP 1932A hasta que haya recibido un acuse de recibo de que el TCP de reserva 1932B ha almacenado el mensaje 2 en la tarjeta de reserva 950. Al mantener esta etapa de bloqueo se garantiza que la tarjeta de reserva 950 tiene los mismos mensajes BGP que se están preparando para enviar a un nodo par en la tarjeta activa 910.

Etapas de bloqueo TCP para mensajes BGP que se están recibiendo

La figura 23 ilustra un diálogo a modo de ejemplo entre un TCP activo que funciona sobre la tarjeta activa 910 y un TCP de reserva que funciona sobre la tarjeta de reserva 950 para mostrar el requisito de la etapa de bloqueo para un mensaje BGP que se está recibiendo desde un nodo par. En referencia a la figura 23, el diálogo 2300 muestra un TCP activo 1932A en la tarjeta activa 910 que recibe un mensaje BGP (mensaje A) desde el nodo remoto 102A. Antes de que el TCP 1932A pueda enviar un acuse de recibo al nodo remoto 102A, el TCP 1932A debe garantizar que el mensaje A se ha replicado en la tarjeta de reserva 950. Si el mensaje A no se ha replicado en la tarjeta de reserva 950, la redundancia se rompe.

Como tal, el TCP 1932A envía el mensaje A a su TCP de reserva 1932B. El TCP de reserva 1932B envía el mensaje A al BGP de reserva 1726B en la tarjeta de reserva 950. El TCP de reserva 1932B envía entonces un acuse de recibo al TCP activo 1932A de que el mensaje A se ha replicado. Tras recibir el acuse de recibo desde dicho TCP de reserva 1932B, el TCP activo 1932A envía el mensaje A al BGP activo 1726A y puede entonces enviar un acuse de recibo del mensaje A al nodo remoto 102A.

Si un segundo mensaje BGP (mensaje B) se recibe en el TCP activo 1932A, también realizará la misma operación que con el mensaje A para replicar el mensaje B en la tarjeta de reserva 950 y esperará al acuse de recibo de que el mensaje B se ha replicado. Tras recibir el acuse de recibo desde el TCP de reserva 1932B, el TCP activo 1932A enviará el mensaje B al BGP activo 1726A y puede entonces enviar un acuse de recibo del mensaje B al nodo remoto 102A. El acuse de recibo del mensaje B no se producirá hasta que el mensaje A se haya replicado. Así, al mantener esta etapa de bloqueo se garantiza que la tarjeta de reserva 950 presenta los mismos mensajes BGP que se han recibido en la tarjeta activa 910.

Actualización incremental (delta) para la redundancia de protocolo BGP

Las figuras 24 y 25 muestran realizaciones variadas de la actualización delta de mensajes BGP individuales que se están enviando a un nodo par implementando el requisito de etapa de bloqueo TCP tal como se ilustra en la figura 22 anteriormente. La figura 26 muestra una realización de la actualización delta para mensajes BGP que se están recibiendo desde un nodo par que implementa el requisito de la etapa de bloqueo TCP tal como se ilustra en la figura 23 anterior.

La figura 24 ilustra una arquitectura BGP 2400 para mostrar la actualización delta para mensajes BGP individuales que se están enviando a un nodo par según una realización. La arquitectura BGP 2400 funciona sobre la plataforma 900 de redundancia tal como se muestra en la figura 9. En la siguiente arquitectura 2400, cada mensaje BGP generado por el protocolo BGP ejecutado en la tarjeta activa 910 se actualiza o replica en una tarjeta de reserva 950. El mensaje BGP pasa a través de un número de etapas de memorización intermedia, que requieren redundancia en cada etapa.

En referencia a la figura 24, se explicará el funcionamiento de la arquitectura 2400 con relación a los puntos de referencia 1 a 34. En el punto de referencia 1, el BGP 1926A envía un mensaje que se almacena en la memoria intermedia BGP 1927A. En el punto de referencia 2, el mensaje debe reflejarse o replicarse en la tarjeta de reserva 950. Así, el mensaje se envía al gestor de redundancia 920, y en el punto de referencia 3, el gestor de redundancia 920 envía el mensaje a su gestor de redundancia par 960 en la tarjeta de reserva 950 a través de un enlace entre tarjetas (por ejemplo, un “cable”). El gestor de redundancia 960 actualiza ahora la memoria intermedia BGP 1927B con el mensaje obteniendo de este modo una copia idéntica del mensaje almacenado en la memoria intermedia BGP 1927A.

En el punto de referencia 5, se envía un acuse de recibo al gestor de redundancia 960 de que el mensaje se ha actualizado. En el punto de referencia 6, el gestor de redundancia 960 en la tarjeta de reserva 950 envía el acuse de recibo al gestor de redundancia 920 en la tarjeta activa 910. El gestor de redundancia 920 almacena por tanto el acuse de recibo en la memoria intermedia BGP 1927. Posteriormente, el mensaje pasa a la cola de conexión 2403A.

Aquí, las operaciones anteriores se realizan para los puntos de referencia 9 a 28 para propagar el mensaje a través de la cola de conexión 2403A, la memoria intermedia de conexión 2404A y la memoria intermedia TCP 1933A en la tarjeta activa 910. Así, el mensaje que se está propagando en la cola de conexión 2403A, la memoria intermedia de conexión 2404A y la memoria intermedia TCP 1933A se refleja o replica en su cola de conexión par 2403B, memoria intermedia de conexión 2404B y memoria intermedia TCP 1933B en la tarjeta de reserva 950. Después de almacenar el mensaje en la memoria intermedia TCP 1933A, para los puntos de referencia 29 a 34, el mensaje se pasa por IP 1930A, la cola del gestor de cadena de protocolo (PCM) 2405A, el PCM 2006A, la cola del controlador (“driver”)

ES 2 267 850 T3

2407A, y el controlador 2408A y sale en un cable hacia el nodo par. Las colas de conexión almacenan información de punto final para el protocolo BGP, estando la información de punto final relacionada con el protocolo BGP ejecutado en otro nodo. El PCM gestiona mensajes designados para cada tipo de protocolo de encaminamiento. Las colas del PCM almacenan mensajes para protocolos de encaminamiento individuales.

5

La figura 25 ilustra una arquitectura BGP 2500 a modo de ejemplo para mostrar la actualización delta para que los mensajes BGP individuales sean para la redundancia de transmisión de datos según otra realización. La arquitectura BGP 2500 ilustrativa reduce el número de etapas de memorización intermedia entre BGP y TCP/IP. En particular, el BGP puede enviar mensajes (“paquetes”) directamente a la memoria intermedia usada por TCP/IP. Por ejemplo, la memoria intermedia puede ser una “memoria intermedia de anillo”. El BGP puede controlar un indicador de “escritura” en la memoria intermedia de anillo y el TCP/IP puede controlar un indicador de “lectura” en la memoria intermedia de anillo.

10

En referencia a la figura 25, se explicará ahora el funcionamiento de la arquitectura 2500 con respecto a los puntos de referencia 1 a 13. En el punto de referencia 1, el BGP 1926A envía un mensaje “paquete” que se almacena en el anillo de datos de transmisión 2508A. En el punto de referencia 2, el mensaje debe reflejarse o replicarse en la tarjeta de reserva 950. Por tanto, el mensaje se envía al gestor de redundancia 920 y, en el punto de referencia 3, el gestor de redundancia 920 envía el mensaje a su gestor de redundancia par 960 en la tarjeta de reserva 950 a través de un enlace entre tarjetas (por ejemplo un “cable”). El gestor de redundancia 960 actualiza ahora el anillo de datos de transmisión 2108B con el mensaje, de manera que se obtiene una copia idéntica del mensaje almacenado en el anillo de datos de transmisión 2508A.

15

En el punto de referencia 5, se envía un acuse de recibo al gestor de redundancia 960 de que el mensaje se ha actualizado. En el punto de referencia 6, el gestor de redundancia 960 en la tarjeta de reserva 950 envía el acuse de recibo al gestor de redundancia 920 en la tarjeta activa 910. El gestor de redundancia 920 almacena por tanto el acuse de recibo en el anillo de datos de transmisión 2508A. Posteriormente, el mensaje pasa por TCP/IP 1930A, la cola PCM 2405A, el PCM 2406A, la cola del controlador 2407A y el controlador 2408A y sale en un cable hacia el nodo par.

25

La figura 26 ilustra una arquitectura BGP 2600 ilustrativa para datos recibidos según una realización. La arquitectura BGP 2600 ilustrativa es similar a la arquitectura BGP 2500 con una memoria intermedia de anillo adicional (es decir, la memoria intermedia de anillo de datos de recepción 2660A y 2660B). En particular, el BGP puede recibir mensajes BGP desde la memoria intermedia de anillo de datos.

30

En referencia a la figura 26, se explicará ahora el funcionamiento de la arquitectura 2600 con respecto a los puntos de referencia 1 a 14. En el punto de referencia 1, la tarjeta de reserva 950 recibe un mensaje BGP desde una conexión por cable por el controlador 2408A. En el punto de referencia 2, el controlador 2408A envía el mensaje a la cola PCM 2405A. En el punto de referencia 3, la cola PCM 2405A envía el mensaje al PCM 2406A. En el punto de referencia 4, el PCM 2406A envía el mensaje a la cola TCP/IP 2631A. En el punto de referencia 5, la cola TCP/IP 2631A envía el mensaje al TCP/IP 1930A. En el punto de referencia 6, el TCP/IP 1930A envía el mensaje a la memoria intermedia de anillo de recepción de datos 2660A.

35

40

En el punto de referencia 7, la memoria intermedia de anillo de recepción de datos 2660A envía el mensaje al gestor de redundancia 920. En el punto de referencia 8, el gestor de redundancia 920 envía el mensaje al gestor de redundancia 960 en la tarjeta de reserva 950 para replicarlo. En el punto de referencia 9, el gestor de redundancia 960 envía el mensaje a la memoria intermedia de anillo de recepción de datos 2660B. En el punto de referencia 10, un acuse de recibo del mensaje se almacena en la memoria intermedia de anillo de transmisión de datos 2550B. En el punto de referencia 11 y 11a, el mensaje se envía al BGP 1726B y el acuse de recibo se envía al gestor de redundancia 960. En el punto de referencia 12, el acuse de recibo se envía al gestor de redundancia 920.

50

En el punto de referencia 13, el gestor de redundancia 920 envía el acuse de recibo a la memoria intermedia de anillo de recepción de datos 2660A. En el punto de referencia 14, el acuse de recibo desde el gestor de redundancia 960 se almacena en la memoria intermedia de anillo de transmisión 2550A. En el punto de referencia 15, el mensaje se envía al BGP 1726A. En la operación anterior, el mensaje recibido por la tarjeta activa 910 no se enviará al BGP 1726A hasta que se haya recibido por el BGP 1726B. Además, las operaciones anteriores, ilustran la actualización incremental de un mensaje BGP recibido desde un nodo par o vecino.

55

La figura 27 ilustra un diagrama de flujo de una operación 2700 para llevar a cabo el mensaje BGP según una realización. La siguiente operación 2700 puede implementarse por el nodo redundante 104 que presenta una tarjeta activa 910 y una tarjeta de reserva 950 tal como se mostró anteriormente. Para la finalidad de la invención, la operación 2700 se refiere a la figura 19 y se inicia en la operación 2702.

60

En la operación 2702, la tarjeta activa 910 recibe un mensaje BGP. La tarjeta activa 910 envía el mensaje a través de las capas superiores hacia una aplicación (BGP), es decir, BGP 1726A. La tarjeta activa 910 también envía el mensaje al TCP 1932B en la tarjeta de reserva 950.

65

En la operación 2704, la tarjeta activa 910 lleva a cabo el mensaje.

ES 2 267 850 T3

En la operación 2706, la tarjeta de reserva 950 recibe el mensaje y envía el mensaje del TCP 1932B a la aplicación (BGP) de reserva, es decir, BGP 1726B.

5 En la operación 2708, la tarjeta de reserva 950 lleva a cabo el mensaje y envía la confirmación al TCP 1932A en la tarjeta activa 910.

En la operación 2710, la tarjeta de reserva 950 recibe la confirmación desde la tarjeta de reserva 950 y convierte la confirmación en una confirmación de sistema.

10 En la operación 2712, la tarjeta activa 910 envía la confirmación de sistema al par remoto.

La operación anterior 2700 utiliza el gestor de redundancia 920 y 960 para facilitar la transferencia de mensajes entre la tarjeta activa 910 y la tarjeta de reserva 950. En otras realizaciones, las operaciones 2702 a 2710 pueden repetirse para otros mensajes, sin embargo, la confirmación de sistema para un particular no se enviará a un nodo par hasta que la tarjeta de reserva 950 la haya confirmado. Las operaciones anteriores permiten que un mensaje BGP pase rápidamente a través de las capas superiores tanto en la tarjeta activa 910 como en la tarjeta de reserva 950.

Secuencia de arranque/actualización masiva para la redundancia de protocolo BGP

20 La figura 28 ilustra un diagrama de flujo de una operación 2800 para realizar la actualización masiva para la redundancia de protocolo BGP según una realización. Para la finalidad de la explicación, la operación 2800 se refiere a una secuencia de arranque en la que una tarjeta activa 910 y una tarjeta de reserva 950 se ejecutan en un encaminador 104.

25 En referencia a la figura 28, en la operación 2802, las bases de datos BGP y TCP se replican desde la tarjeta activa 910 hasta la tarjeta de reserva 950. Por ejemplo, tal como se muestra en la figura 19, las RDB BGP 1927A y PDB BGP 1731A en la tarjeta activa 910 se replican en RDB BGP 1927B y PDB BGP 1731B en la tarjeta de reserva 950. La operación de actualización masiva tal como se explica en la plataforma 900 de redundancia puede usarse para replicar las bases de datos BGP. Además, las RDB TCP 1933A y PDB TCP 1926A en la tarjeta activa 910 se replican en RDB TCP 1933B y PDB TCP 1926B en la tarjeta de reserva 950 usando la misma operación que para las bases de datos BGP. Alternativamente, la memoria no volátil así como otras bases de datos pueden también replicarse.

30 En la operación 2804, cualquier mensaje BGP recibido o enviado se pone en cola de manera que pueda realizarse en la tarjeta de reserva 950 después del proceso de actualización masiva. En la operación 2806, cualquier mensaje BGP recibido o enviado se actualiza en delta en la tarjeta de reserva 950 usando las operaciones de actualización delta tal como se indicaron en las figuras 24 a 26.

Redundancia de protocolo sistema intermedio a sistema intermedio (IS-IS)

40 *Requisitos básicos para la redundancia de protocolo IS-IS*

El protocolo IS-IS es un protocolo de estado de enlace. Un encaminador en una zona/dominio que genera un paquete de protocolo IS-IS inunda todos los encaminadores en la zona/dominio con el paquete. Es decir, el paquete generado por un encaminador IS-IS se almacena en todos los encaminadores IS-IS en la zona o dominio. Por tanto, cada encaminador IS-IS tiene una visión completa y consistente de la red de los demás encaminadores IS-IS. Estos paquetes se denominan paquetes de estado de enlace (LSP, "link state packets"). Un paquete LSP incluye información sobre el encaminador IS-IS que ha generado el paquete. Como tal, cada encaminador que ejecuta un protocolo IS-IS incluye una base de datos de LSP o una base de datos IS-IS que almacena paquetes LSP.

50 Para obtener la redundancia del protocolo IS-IS, el sistema de control de reserva debe mantener o conocer la información de configuración/global, la información de circuito, la información de adyacencia, y la información del paquete de estado de enlace (LSP) en el sistema de control activo. La información de configuración incluye información global tal como el estado global de una tarjeta activa 910, es decir si está activa o en reserva. La información de circuito incluye los estados de los circuitos que se están ejecutando. Por ejemplo, si los circuitos están habilitados o deshabilitados. La información de adyacencia incluye información sobre la adyacencia de la activa, es decir quienes son sus vecinos. La información de estado de enlace incluye información sobre paquetes LSP. La plataforma 900 de redundancia proporciona el soporte para mantener la redundancia del protocolo IS-IS.

Secuencia de arranque/actualización masiva para la redundancia de protocolo IS-IS

60 La figura 29 ilustra un diagrama de flujo de una operación 2900 para realizar una actualización masiva para la redundancia del protocolo IS-IS según una realización. Para la finalidad de la explicación, la operación 2900 se refiere a una secuencia de arranque en la que una tarjeta activa 910 y una tarjeta de reserva 950 se ejecutan en un encaminador 104.

65 En referencia a la figura 29, en la operación 2902, las bases de datos IS-IS se replican desde la tarjeta activa 910 a la tarjeta de reserva 950. Por ejemplo, tal como se muestra en la figura 19, las RDB IGP 1942A y PDB IGP 1924A y FIB 1716A en la tarjeta activa 910 se replican en las RDB IGP 1942B y PDB IGP 1924B y FIB 1726B en la tarjeta de

ES 2 267 850 T3

reserva 950. Las bases de datos IGP pueden copiarse selectivamente sólo para la información de las bases de datos IS-IS. La operación de actualización masiva, tal como se explicó en la plataforma 900 de redundancia puede usarse para replicar las bases de datos IBP IS-IS.

5 En la operación 2904, cualquier mensaje IS-IS recibido o enviado se pone en cola de manera que pueda realizarse en la tarjeta de reserva 950 después de l proceso de actualización masiva. En la operación 2906, cualquier mensaje IS-IS recibido o enviado se actualiza en delta en la tarjeta de reserva 950 utilizando las operaciones de actualización delta tal como se mostró en las figuras 24 a 26.

10 *Actualización incremental (delta) para la redundancia del protocolo IS-IS*

La figura 30 ilustra un diagrama de flujo de una operación 3000 para realizar la actualización incremental (delta) para mensajes IS-IS individuales recibidos o enviados según una realización. Para la finalidad de la explicación, la operación 3000 se refiere a un encaminador que presenta una tarjeta activa 910 y una tarjeta de reserva 950.

15 En referencia a la figura 30, en la operación 3002, se recibe o envía un mensaje IS-IS. Por ejemplo, el protocolo IS-IS puede generar un LSP o recibir un LSP desde un nodo par. En la operación 3004, la tarjeta activa 910 envía el paquete LSP a la tarjeta de reserva 950, que se trata como un vecino. Por tanto, la tarjeta de reserva 950 replica el paquete LSP.

20 *Redundancia del protocolo de abrir primero la ruta más corta (OSPF)*

Requisitos básicos para la redundancia del protocolo OSPF

25 El protocolo OSPF es un protocolo de encaminamiento de estado de enlace intra-dominio y se basa en el protocolo IP para transmitir y recibir paquetes. El OSPF no utiliza TCP o UDP para la transferencia fiable de paquetes. El protocolo OSPF construye adyacencias con los nodos pares vecinos intercambiando información de red con los nodos pares. El OSPF se actualiza en la FIB y otros protocolos se realizan por el gestor de la tabla de encaminamiento (RTM). Los requisitos básicos para la redundancia del protocolo OSPF es mantener los servicios del protocolo OSPF sin perturbación para el RTM en la tarjeta de reserva 950 con el RTM en los nodos pares. En consecuencia, para la redundancia del protocolo OSPF, toda la información de estado del protocolo, la información de la base de datos OSPF, la información de configuración deben mantenerse en la tarjeta de reserva 950.

Secuencia de arranque/actualización masiva para la redundancia del protocolo OSPF

35 La figura 31 ilustra un diagrama de flujo de una operación 3100 para realizar la actualización masiva para la redundancia del protocolo OSPF según una realización. Para la finalidad de la explicación, la operación 3100 se refiere a una secuencia de arranque en la que una tarjeta activa 910 y una tarjeta de reserva 950 funcionan en un encaminador 104.

40 En referencia a la figura 31, en la operación 3102, las bases de datos OSPF se replican desde la tarjeta activa 910 a la tarjeta de reserva 950. Por ejemplo, tal como se muestra en la figura 19, las RDB IGP 1942A y PDB IGP 1924A y FIB 1716A en la tarjeta activa se replican en las RDB IGP 1942B y PDB IGP 1924B y FIB 1716B en la tarjeta de reserva 950. Las bases de datos IGP puede copiarse selectivamente sólo para la información de las bases de datos OSPF. La operación de actualización masiva tal como se explicó en la plataforma 900 de redundancia puede usarse para replicar las bases de datos IGP OSPF.

45 En la operación 3104, cualquier mensaje OSPF recibido o generado se pone en cola de manera que pueda realizarse en la tarjeta de reserva 950 después del proceso de actualización masiva. En la operación 3106, cualquier mensaje OSPF recibido o enviado se actualiza en delta en la tarjeta de reserva 950 usando las operaciones de actualización delta tal como se mostró en las figuras 24 a 26.

Actualización incremental (delta) para la actualización del protocolo OSPF

55 La figura 32 ilustra un diagrama de flujo de una operación 3200 para realizar la actualización incremental (delta) para mensajes OSPF individuales recibidos o enviados según una realización. Para la finalidad de la explicación, la operación 3200 se refiere a un encaminador que tiene una tarjeta activa 910 y una tarjeta de reserva 950. En referencia a la figura 32, en la operación 3202, se recibe o genera un mensaje OSPF. En la operación 3004, la tarjeta activa 910 envía el mensaje OSPF recibido o generado a la tarjeta de reserva 950 usando la actualización delta tal como se explica en la plataforma 900 de redundancia. Las técnicas y las operaciones de redundancia de protocolos de encaminamiento descritas anteriormente son de naturaleza ilustrativa y pueden aplicarse a otros tipos de protocolos de encaminamiento tales como, por ejemplo, el protocolo de Internet de encaminamiento (RIP). Por ejemplo, la plataforma 900 de redundancia puede usarse para realizar actualizaciones masivas, delta, de datos no persistentes y de datos persistentes para la información RIP tal como se ha descrito anteriormente.

65 El encaminador y las operaciones de redundancia de protocolo de encaminamiento anteriores pueden implementarse como rutinas de software ejecutadas por un procesador. Para un procesador dado, las rutinas de software pueden almacenarse en un dispositivo de almacenamiento, tal como una memoria permanente. Alternativamente, las rutinas

ES 2 267 850 T3

de software pueden ser instrucciones ejecutables por ordenador almacenadas en cualquier medio de almacenamiento legible por ordenador, tal como un disquete, un CD-ROM, una cinta magnética, disco de video digital o versátil (DVD), disco láser, ROM, memoria flash, u otro tipo de dispositivos de memoria similares. La serie de instrucciones no tienen que almacenarse localmente, y podrían recibirse desde un dispositivo de almacenamiento remoto, tal como un servidor en una red, un dispositivo de CD ROM, un disco flexible, etc. Las instrucciones pueden copiarse desde el dispositivo de almacenamiento a una memoria temporal y después accederse a ellas y ejecutarse por un procesador. Para una implementación, estas rutinas de software pueden escribirse en el lenguaje de programación C. Debe apreciarse, sin embargo, que estas rutinas pueden implementarse en cualquiera de una gran variedad de lenguajes de programación.

Para realizaciones alternativas, el encaminador y las operaciones de redundancia de protocolos de encaminamiento pueden implementarse en hardware o firmware discretos. Por ejemplo, uno o más circuitos integrados específicos de una aplicación (ASIC) podrían programarse para realizar las operaciones de redundancia descritas anteriormente. En otro ejemplo, las operaciones de redundancia pueden implementarse en uno o más ASIC sobre placas de circuito impreso adicionales y las placas de circuito impreso podrían insertarse en el encaminador o nodo con redundancia tal como se de ha descrito anteriormente. En otro ejemplo, pueden usarse matrices de puertas programables de campo (FPGA) o matrices de puertas programables estáticas (SPGA) para implementar las operaciones de redundancia descritas en la presente memoria. En otro ejemplo más, podría usarse una combinación de hardware y software para implementar las operaciones de redundancia descritas en la presente memoria.

Por lo tanto, se ha descrito un encaminador y una redundancia de protocolos de encaminamiento. En la memoria anterior, la invención se ha descrito con referencia a realizaciones ilustrativas específicas de la misma. La memoria y los dibujos deben considerarse en un sentido ilustrativo y no en un sentido restrictivo.

REIVINDICACIONES

5 1. Procedimiento de funcionamiento de un dispositivo de red que presenta una plataforma de redundancia que incluye un sistema de control activo y un sistema de control de reserva, comprendiendo dicho procedimiento:

recibir un cambio de estado de protocolo de encaminamiento desde un nodo par o generar el cambio de estado de protocolo de encaminamiento mediante el sistema de control activo;

10 enviar el cambio de estado de protocolo de encaminamiento al sistema de control de reserva;

caracterizado porque

15 se recibe una confirmación del cambio de estado de protocolo de encaminamiento por el sistema de control activo desde el sistema de control de reserva;

se lleva a cabo el cambio de estado de protocolo de encaminamiento en el sistema de control activo; y

20 se envía la confirmación al nodo par por el sistema de control activo.

2. Procedimiento según la reivindicación 1, que comprende además:

25 enviar el cambio de estado de protocolo de encaminamiento para actualizar un protocolo de encaminamiento después de recibir la confirmación desde el sistema de control de reserva.

3. Procedimiento según la reivindicación 1, que comprende además:

30 enviar el cambio de estado de protocolo de encaminamiento a un protocolo de encaminamiento después de recibir el cambio de estado de protocolo de encaminamiento desde el nodo par.

4. Procedimiento según la reivindicación 1, en el que el cambio de estado de protocolo de encaminamiento incluye un cambio de estado de protocolo de puerta de enlace de frontera, BGP, un cambio de estado de protocolo de Internet de encaminamiento, RIP, un cambio de estado de protocolo de abrir primero la ruta más corta, OSPF, o un cambio de estado de protocolo de sistema intermedio a sistema intermedio, IS-IS.

35 5. Procedimiento según la reivindicación 1, que comprende además:

recibir o generar un mensaje de protocolo de encaminamiento por el sistema de control activo; y

40 replicar de manera selectiva el mensaje de protocolo de encaminamiento recibido o generado en el sistema de control de reserva.

6. Procedimiento según la reivindicación 5, en el que el mensaje de protocolo de encaminamiento incluye un mensaje BGP, un mensaje RIP, un mensaje OSPF o un mensaje IS-IS.

45 7. Procedimiento según la reivindicación 1, que comprende además:

detectar un fallo en el sistema de control activo; y

50 mantener en el sistema de control de reserva los mismos cambios de estado de protocolo de encaminamiento en el sistema de control activo antes de fallo.

8. Procedimiento según la reivindicación 1, que comprende además:

55 realizar el servicio de capa 3 del protocolo de Internet, IP, o un servicio de conmutación de etiquetas multiprotocolo, MPLS.

9. Dispositivo de red, que comprende:

60 un controlador de reserva; y

un controlador activo adaptado para recibir un cambio de estado de protocolo de encaminamiento desde un nodo par o para generar el cambio de estado de protocolo de encaminamiento, adaptado para enviar el cambio de estado de protocolo de encaminamiento al controlador de reserva,

65 **caracterizado** porque

ES 2 267 850 T3

dicho controlador activo está adaptado para recibir una confirmación con el cambio de estado de protocolo de encaminamiento desde el sistema de control de reserva, adaptado para llevar a cabo el cambio de estado de protocolo de encaminamiento después de recibir la confirmación desde el controlador de reserva, y adaptado para enviar la confirmación al nodo par.

5

10. Dispositivo de red según la reivindicación 9, en el que el controlador activo está adaptado para enviar el cambio de estado de protocolo de encaminamiento a un protocolo de encaminamiento después de recibir la confirmación desde el controlador de reserva.

10

11. Dispositivo de red según la reivindicación 9, en el que el controlador activo está adaptado para enviar el cambio de estado de protocolo de encaminamiento a un protocolo de encaminamiento después de recibir el cambio de estado de protocolo de encaminamiento desde el nodo par.

15

12. Dispositivo de red según la reivindicación 1, en el que el cambio de estado de protocolo de encaminamiento incluye un cambio de estado de protocolo de puerta de enlace de frontera, BGP, un cambio de estado de protocolo de Internet de encaminamiento, RIP, un cambio de estado de protocolo de abrir primero la ruta más corta, o un cambio de estado de protocolo de sistema intermedio a sistema intermedio, IS-IS.

20

13. Dispositivo de red según la reivindicación 1, en el que el sistema de control activo está adaptado para recibir o generar un mensaje de protocolo de encaminamiento y replicar de manera selectiva el mensaje de protocolo de encaminamiento recibido o generado al sistema de control de reserva.

25

14. Dispositivo de red según la reivindicación 13, en el que el mensaje de protocolo de encaminamiento incluye un mensaje BGP, un mensaje RIP, un mensaje OSPF o un mensaje IS-IS.

30

15. Dispositivo de red según la reivindicación 1, en el que el sistema de control activo está adaptado para detectar un fallo y el sistema de control de reserva está adaptado para mantener los mismos cambios de estado de protocolo de encaminamiento en el sistema de control activo antes del fallo.

30

16. Dispositivo de red según la reivindicación 9, en el que el sistema de control activo o el sistema de control de reserva está adaptado para realizar un servicio de capa 3 del protocolo de Internet, IP, o un servicio de conmutación de etiquetas multiprotocolo, MPLS.

35

17. Dispositivo de red según la reivindicación 9, en el que el dispositivo de red incluye un encaminador de red, un conmutador, un conmutador óptico, un puente, un concentrador o una puerta de enlace.

40

18. Sistema que presenta una plataforma de redundancia que incluye un sistema de control activo y un sistema de control de reserva, comprendiendo el sistema:

primeros medios de recepción adaptados para recibir un cambio de estado de protocolo de encaminamiento desde un nodo par por el sistema de control activo;

primeros medios de envío adaptados para enviar el cambio de estado de protocolo de encaminamiento al sistema de control de reserva;

45

caracterizado por

segundos medios de recepción adaptados para recibir una confirmación con el cambio de estado de protocolo de encaminamiento por el sistema de control activo desde el sistema de control de reserva;

50

medios de confirmación adaptados para llevar a cabo el cambio de estado de protocolo de encaminamiento en el sistema de control activo; y

55

segundos medios de envío adaptados para enviar la confirmación al nodo par por el sistema de control activo.

19. Programa informático que comprende medios de codificación de programa informático adaptados para realizar todas las etapas según cualquiera de las reivindicaciones 1 a 8.

60

20. Programa informático según la reivindicación 19 cuando se realiza en un medio legible por ordenador.

65

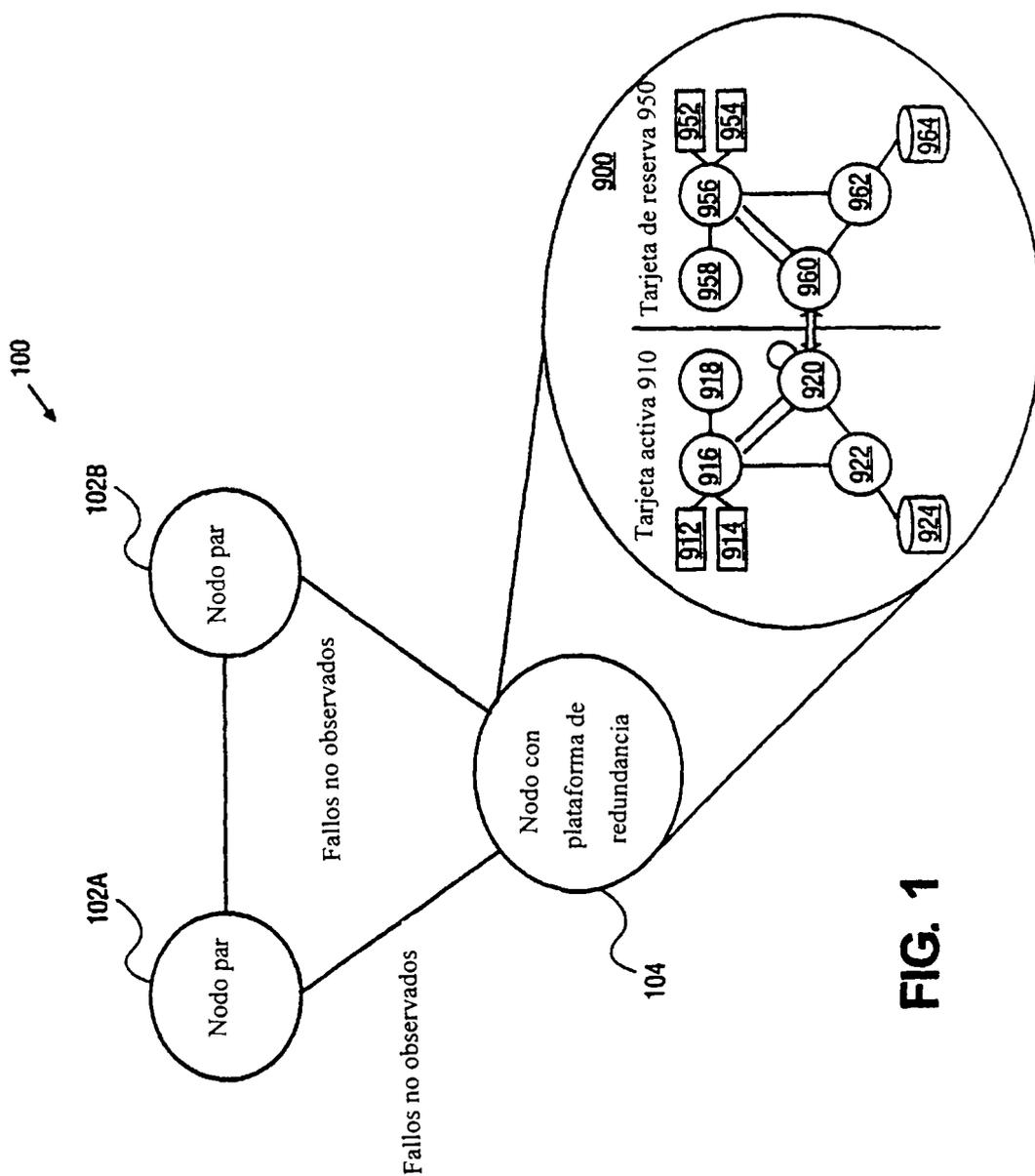


FIG. 1

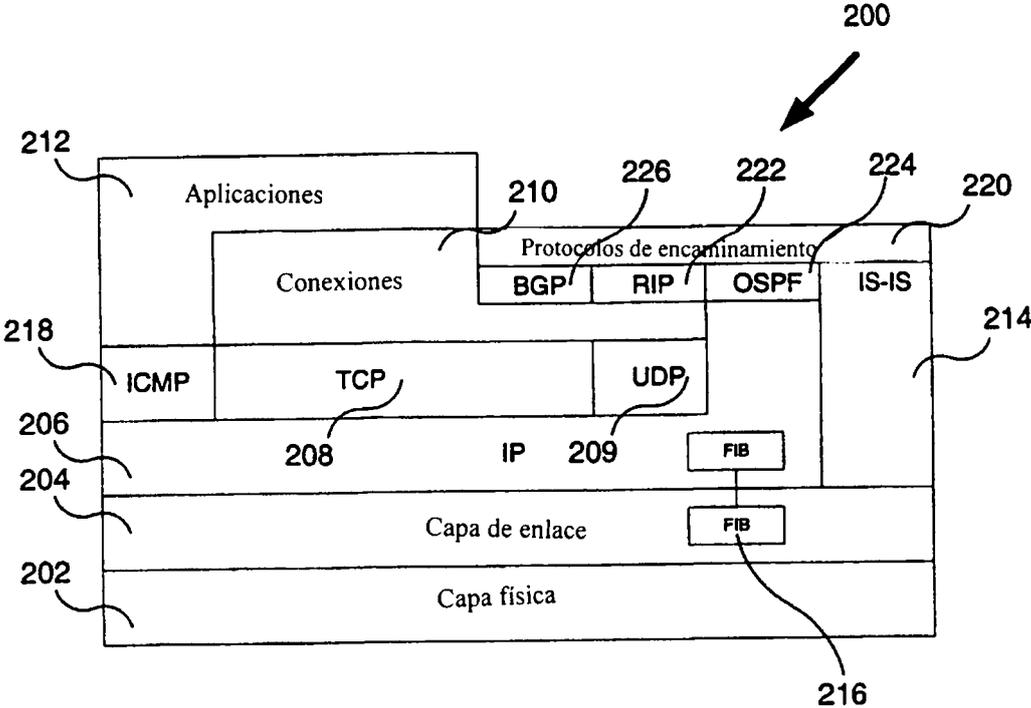


FIG. 2

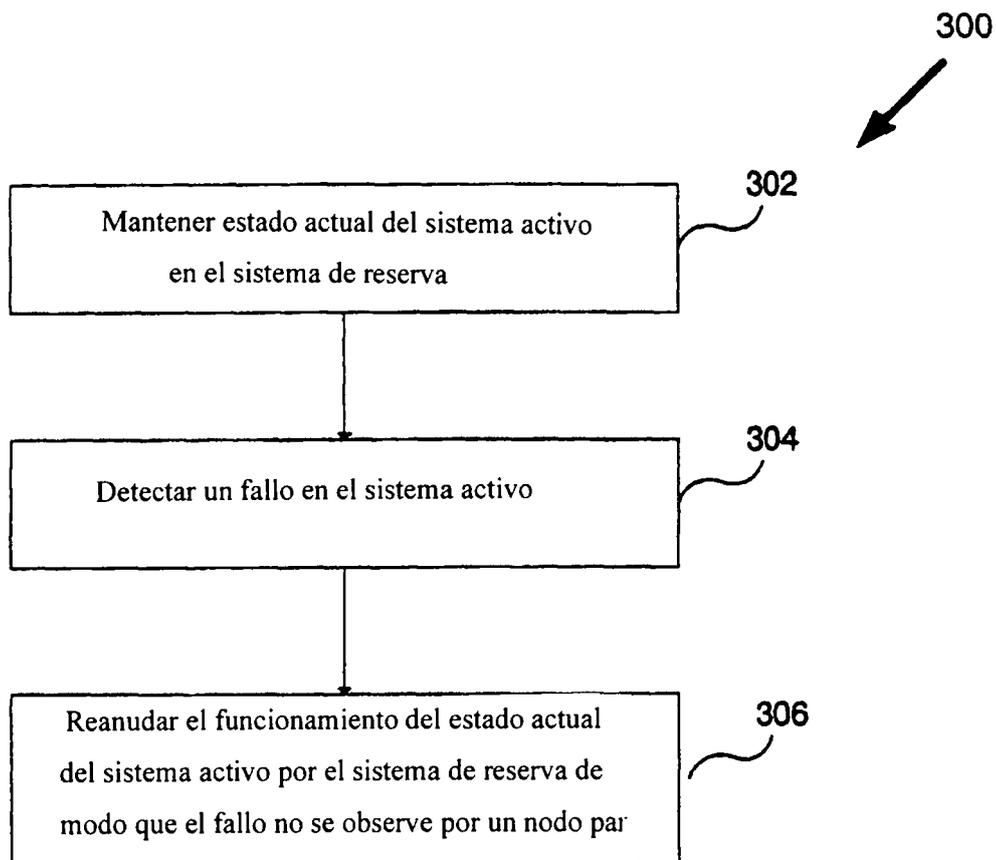


FIG. 3

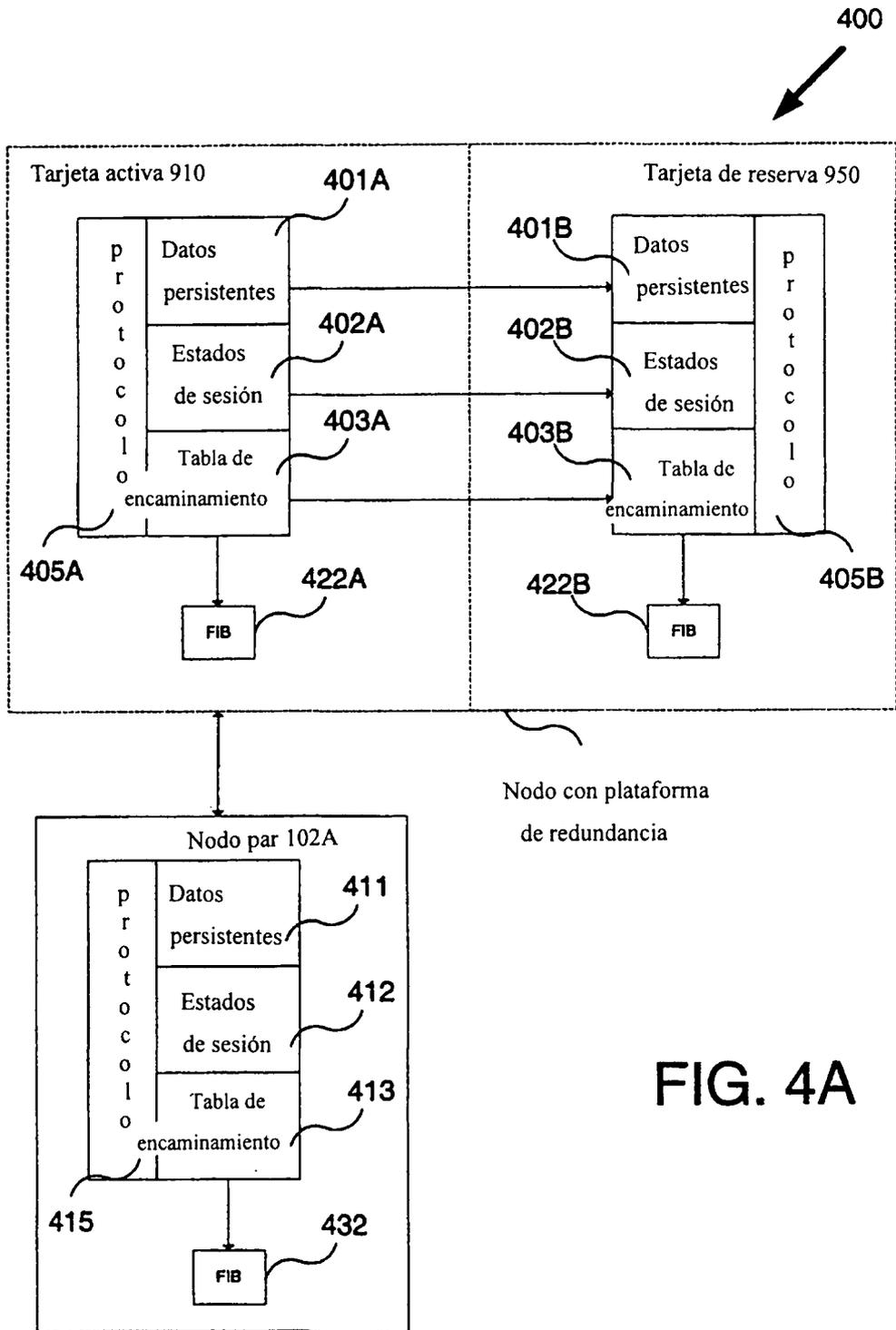


FIG. 4A

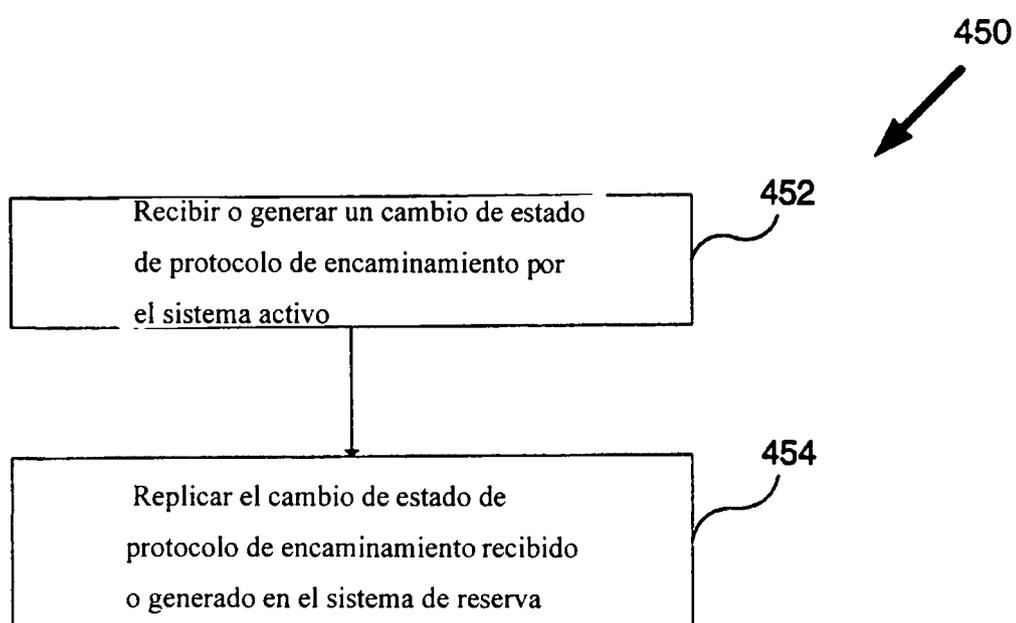


FIG. 4B

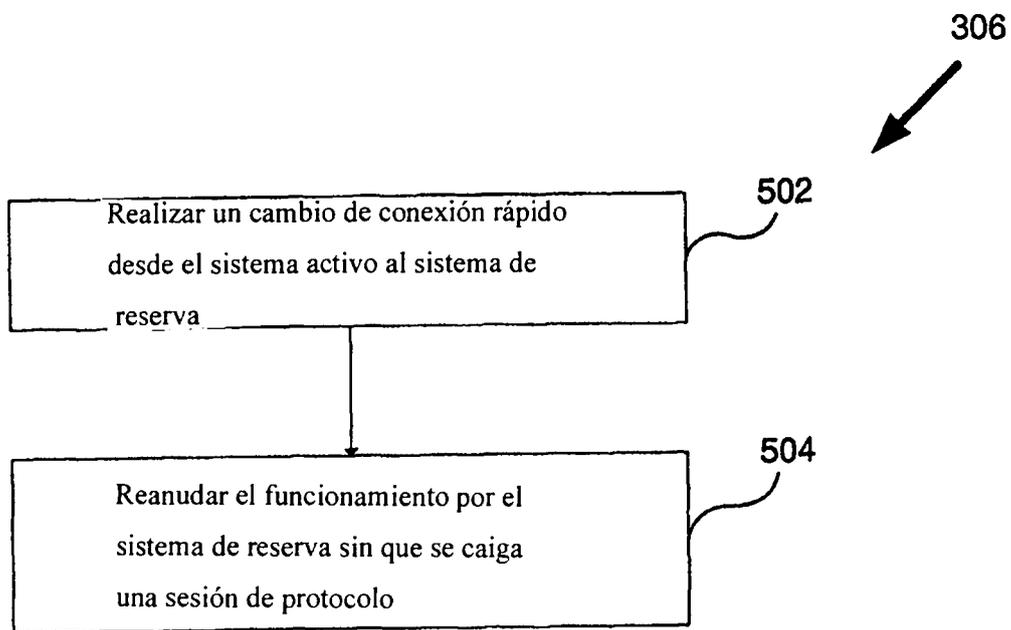


FIG. 5

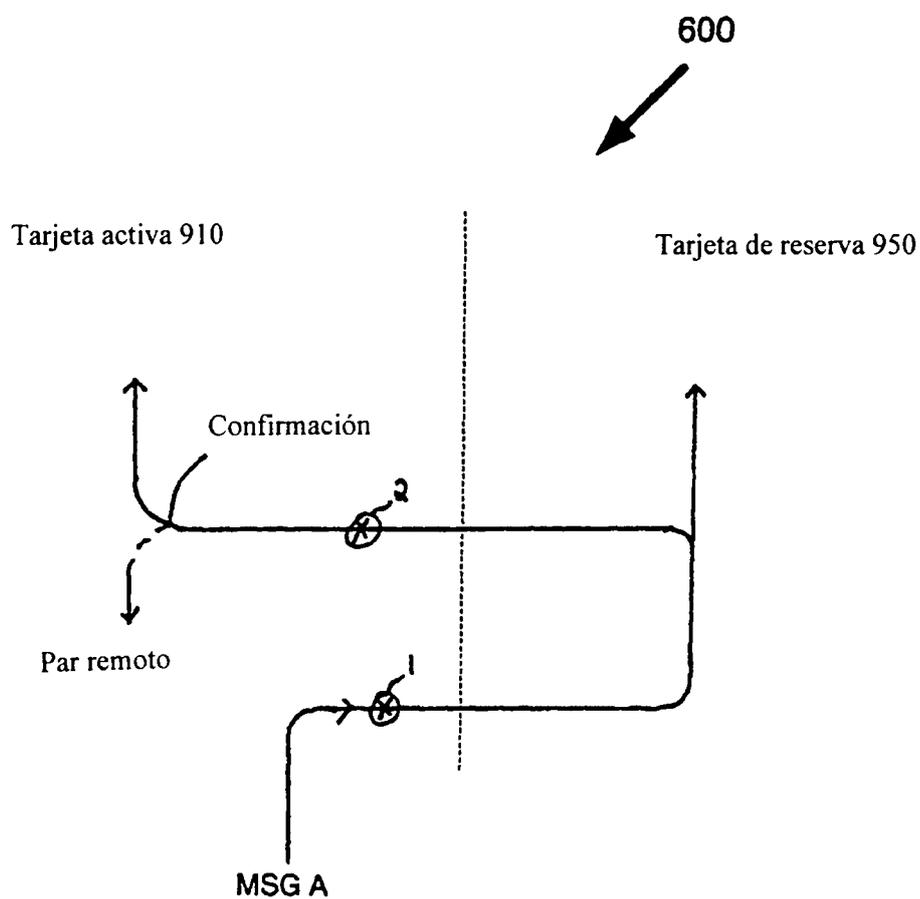


FIG. 6A

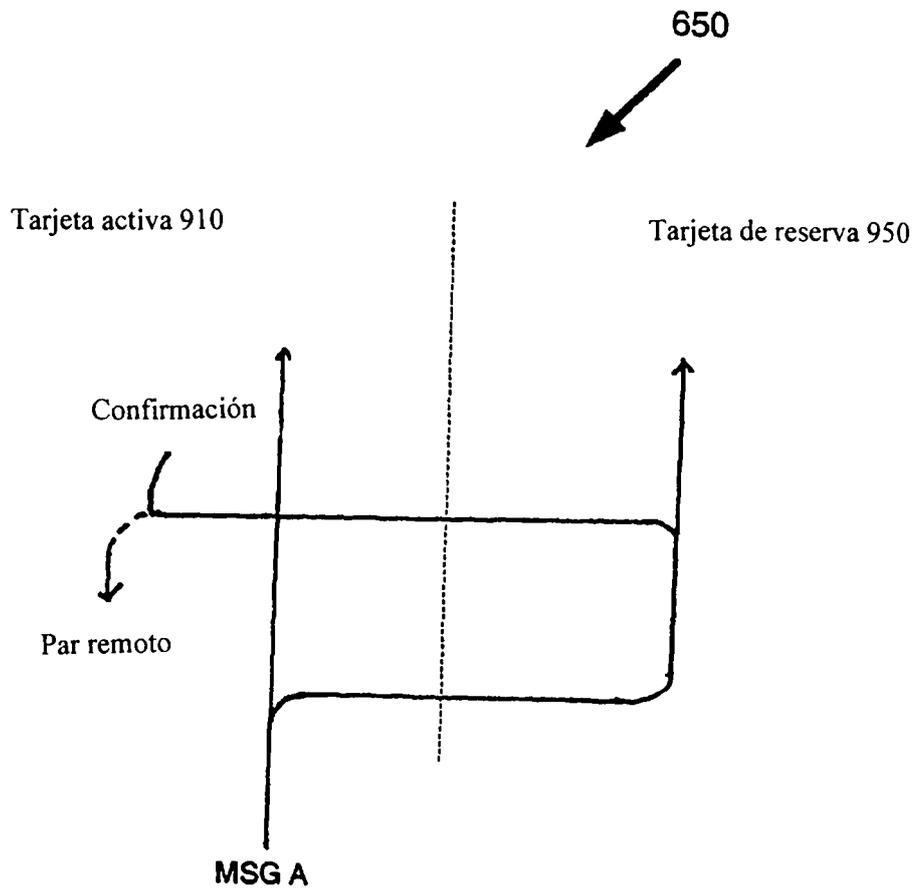


FIG. 6B

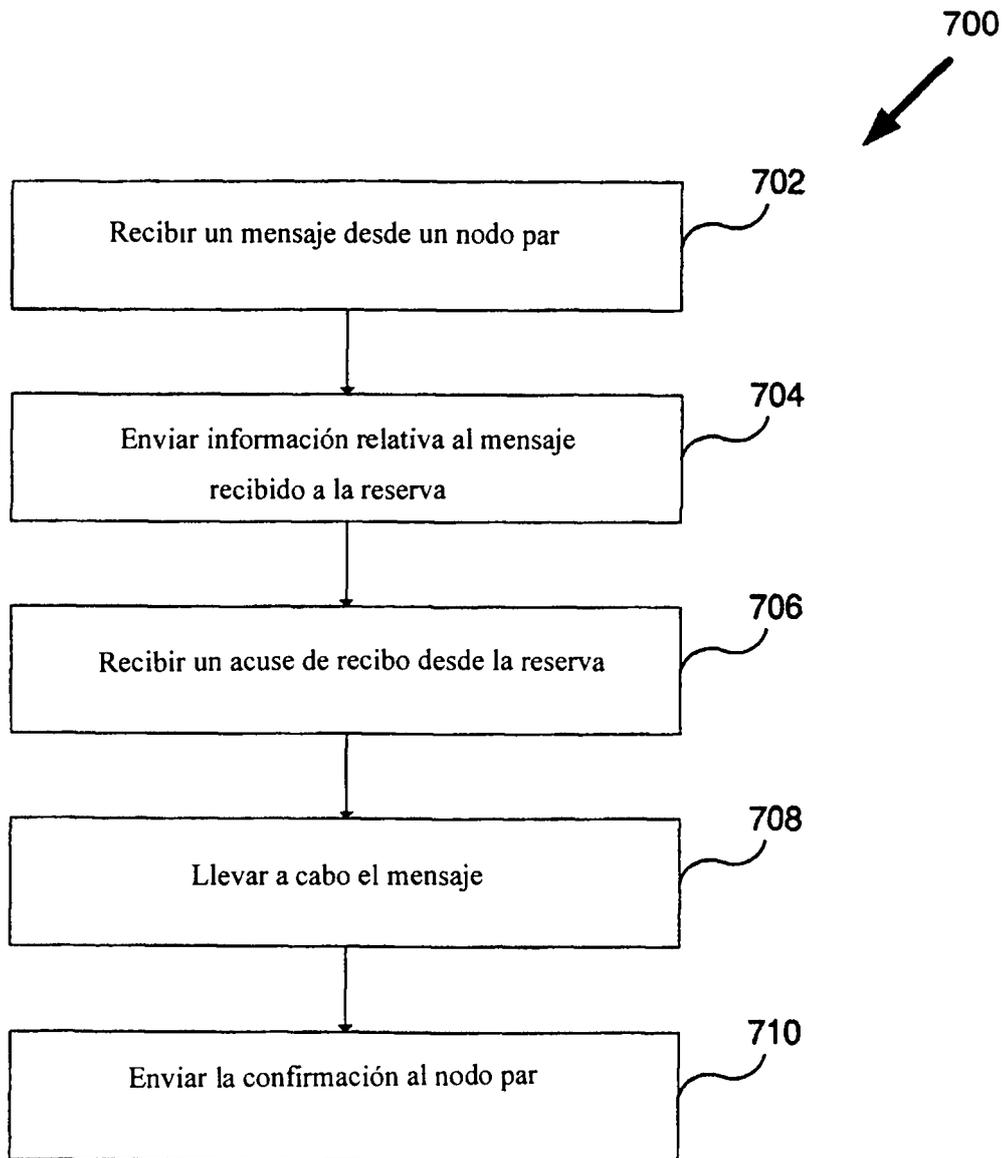


FIG. 7

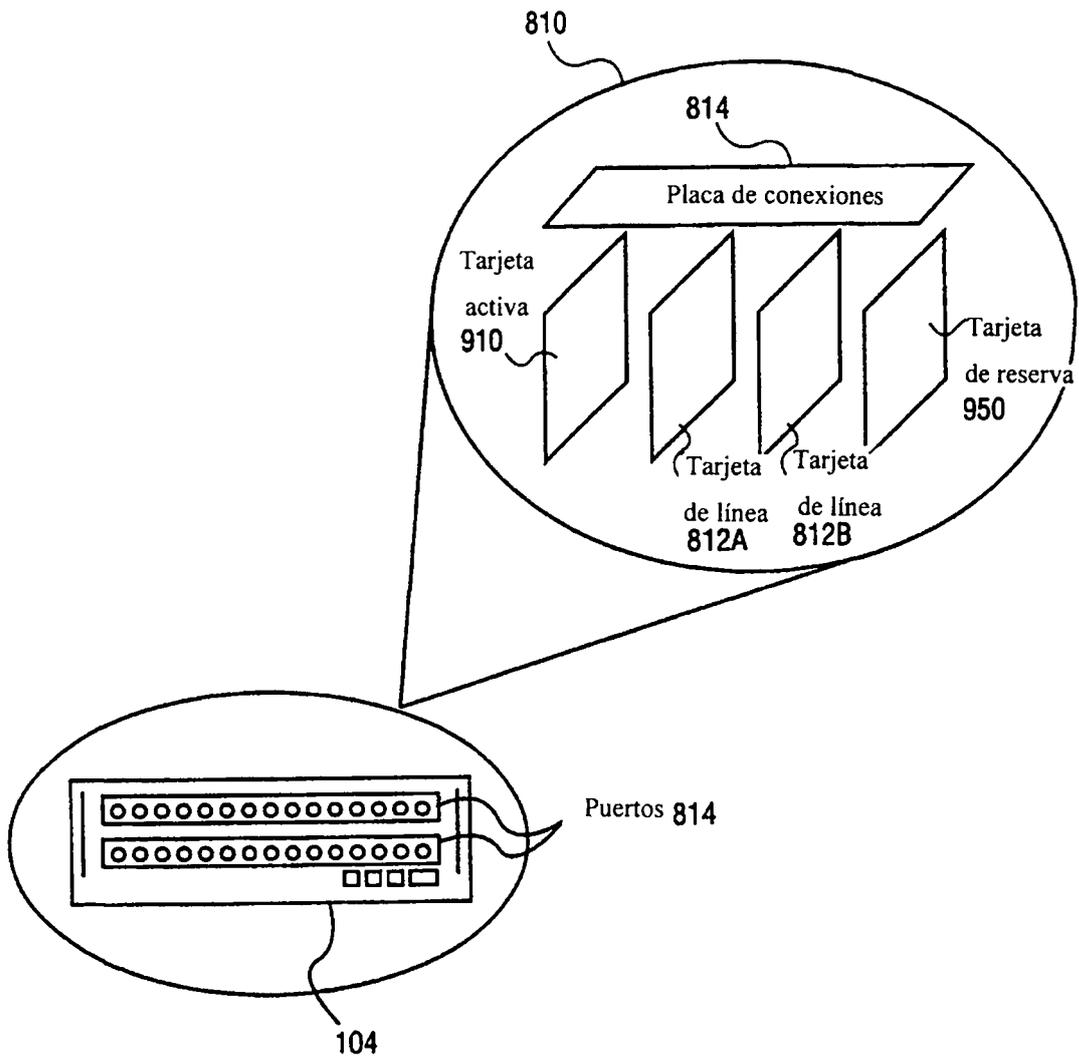


FIG. 8

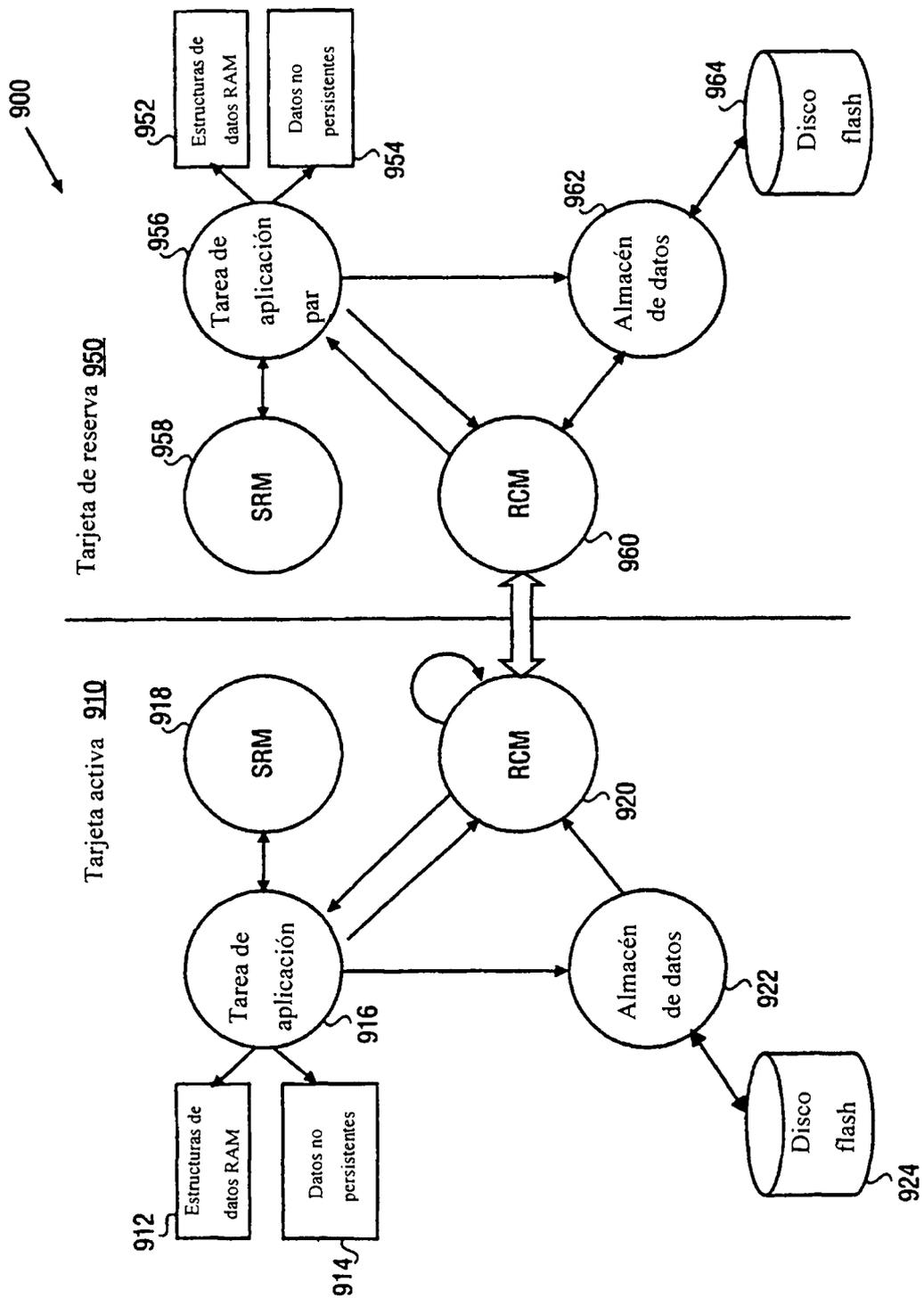


FIG. 9

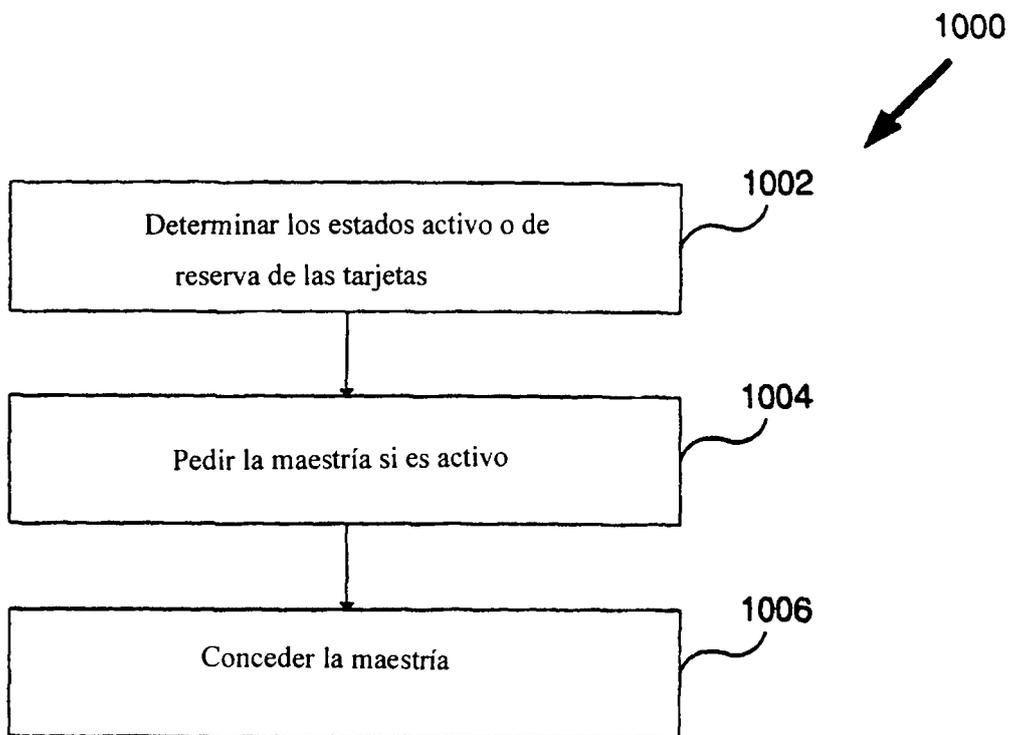


FIG. 10A

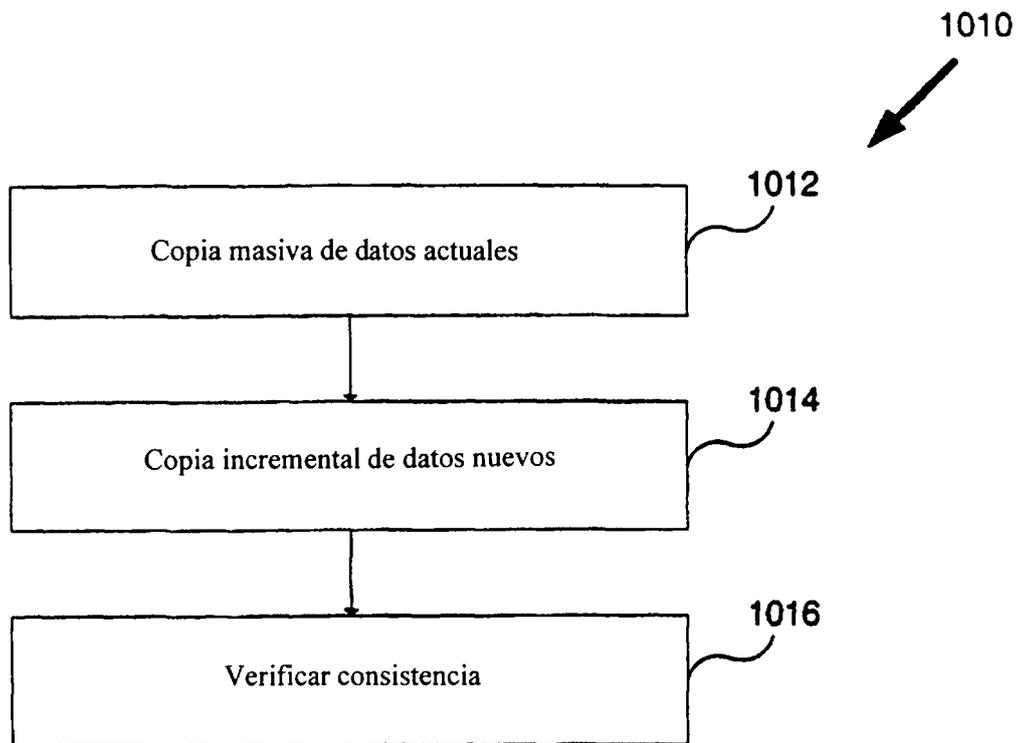


FIG. 10B

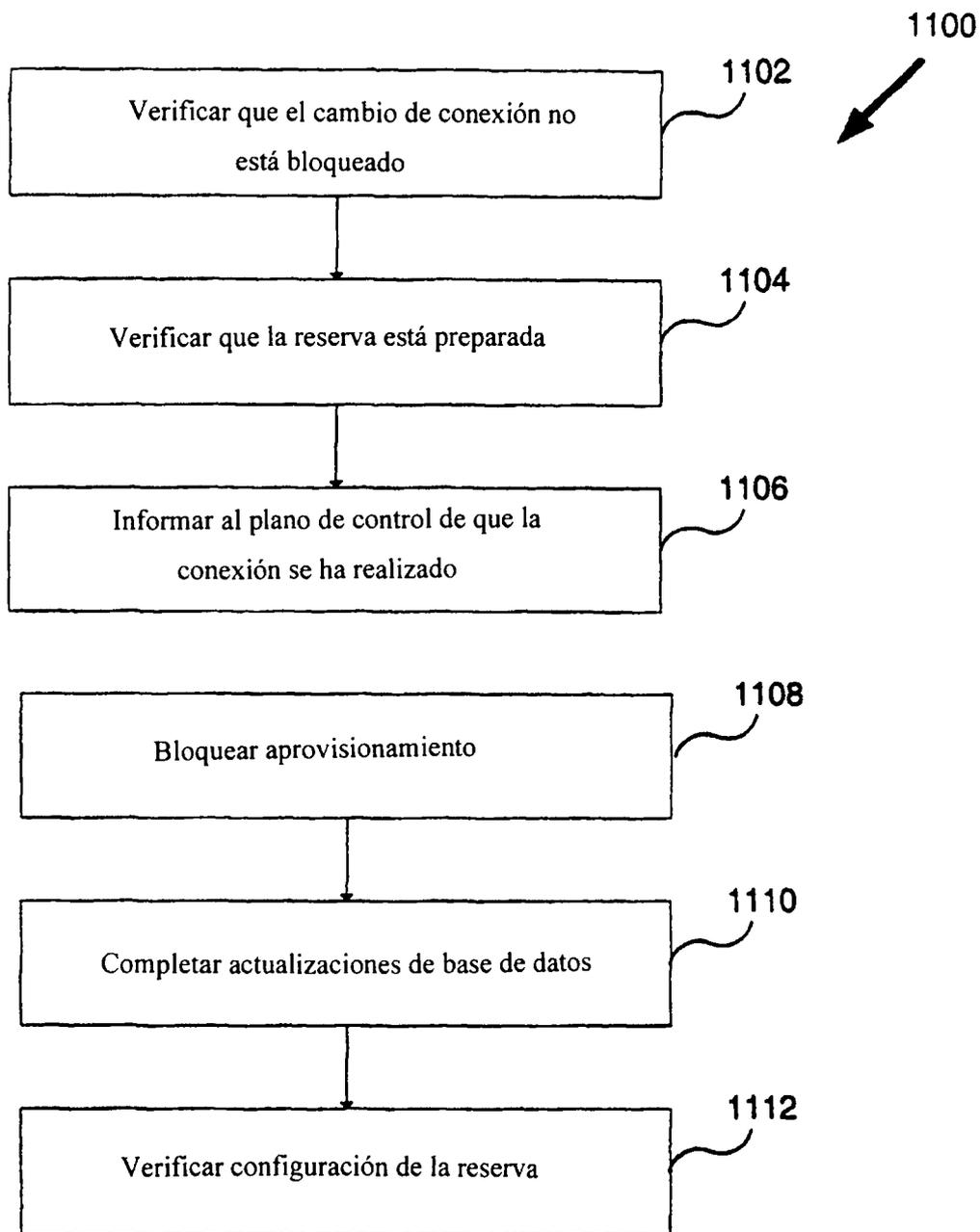


FIG. 11A

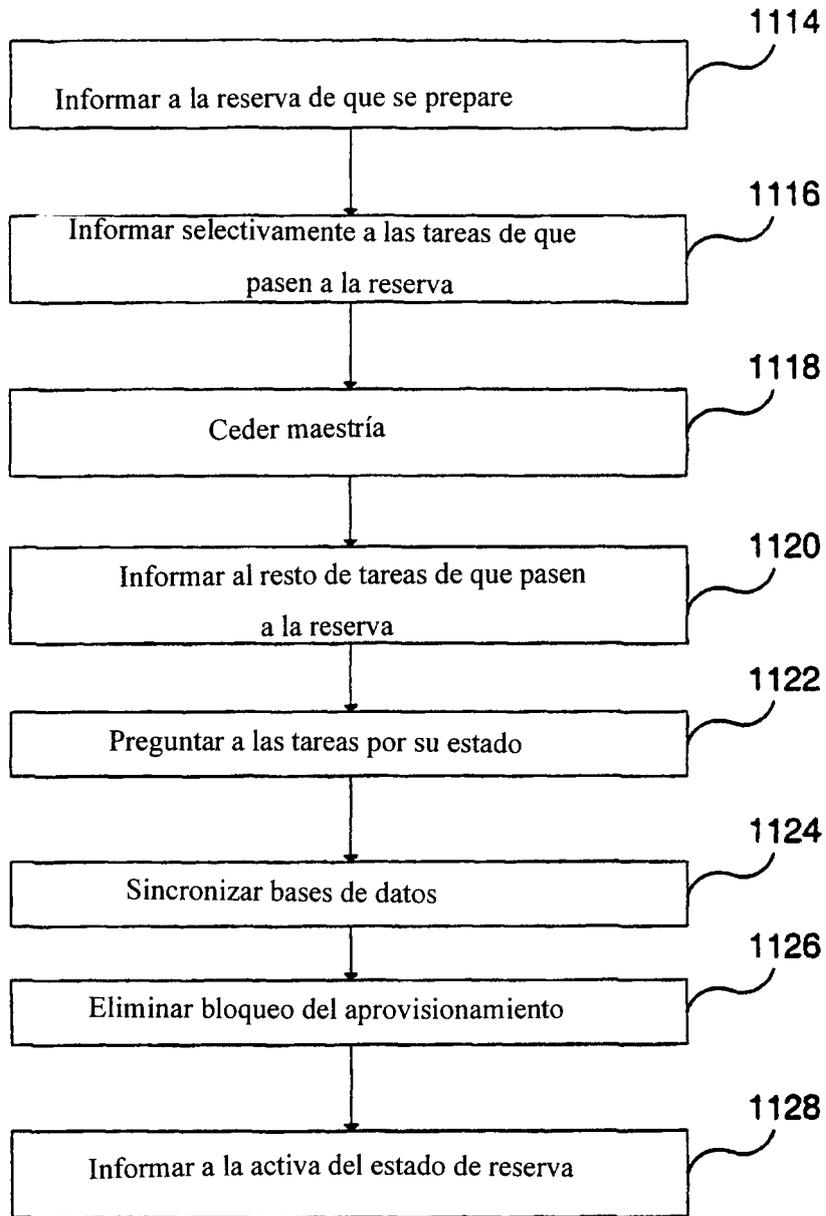


FIG. 11B

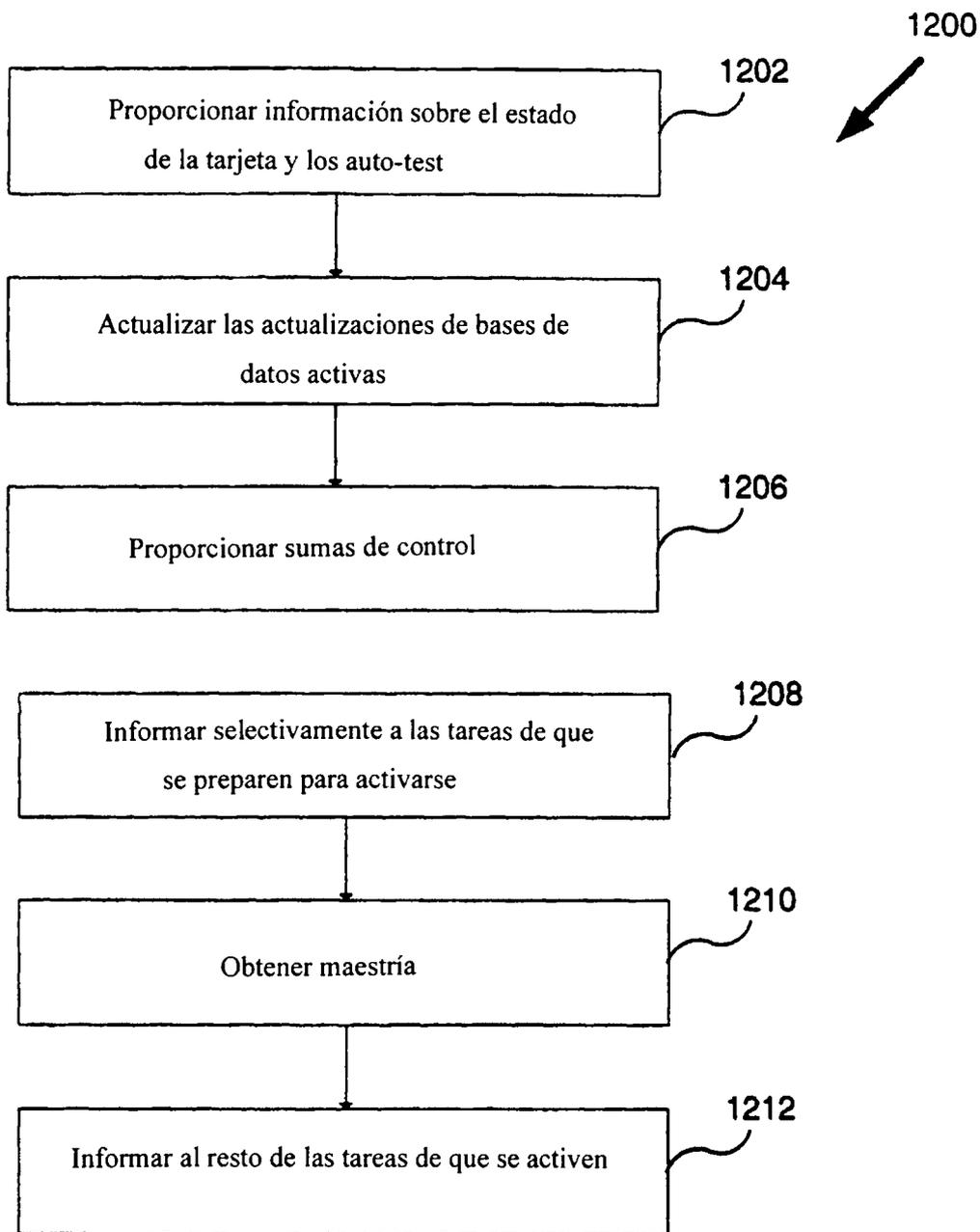


FIG. 12A

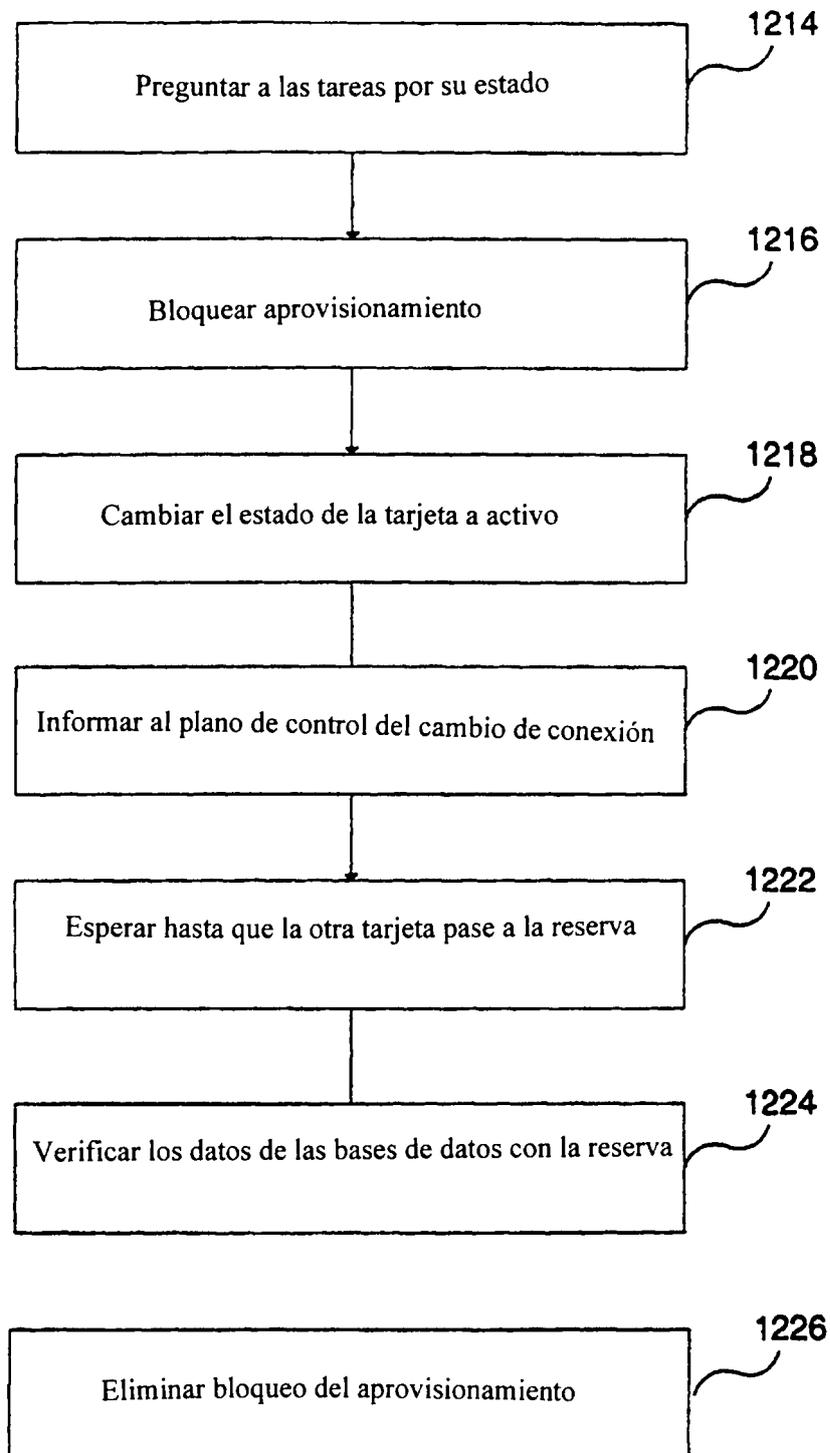


FIG. 12B

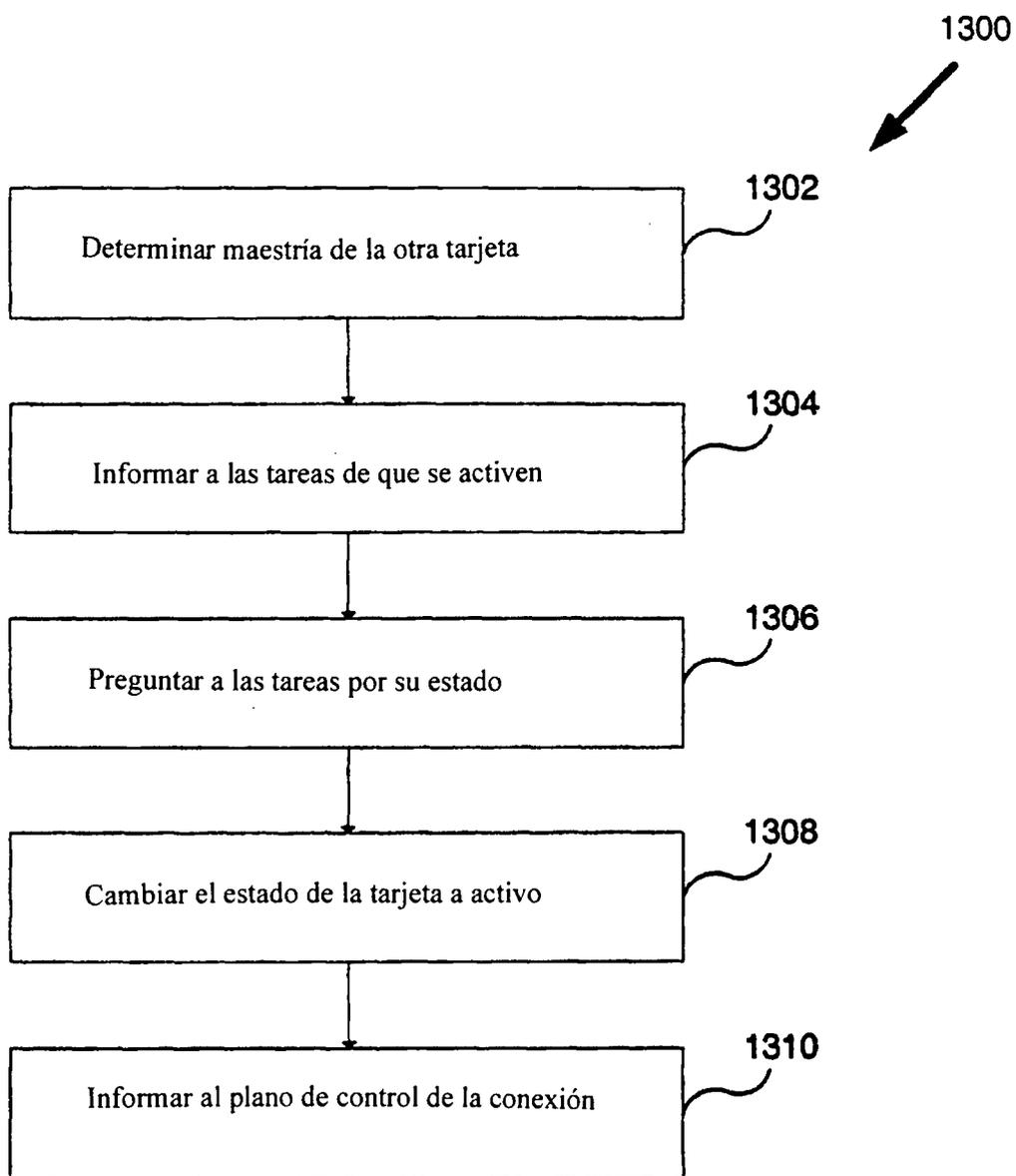


FIG. 13

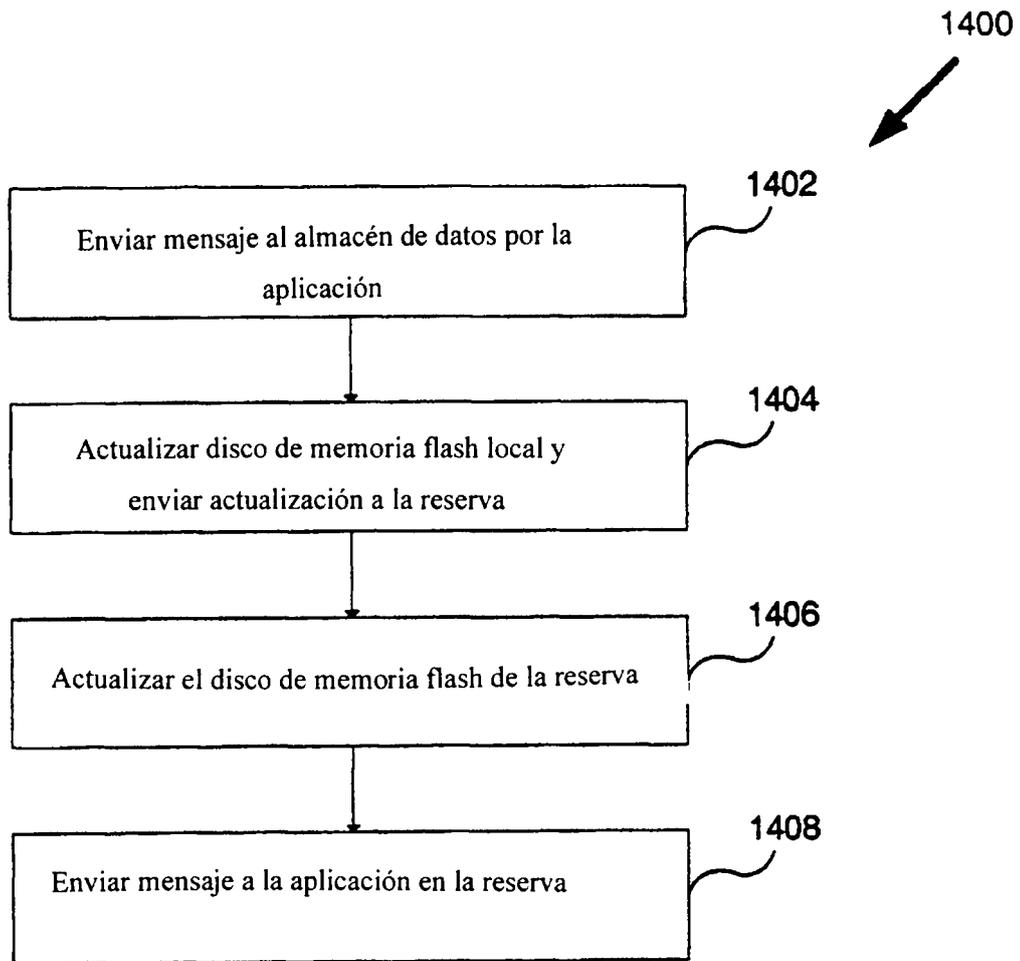


FIG. 14

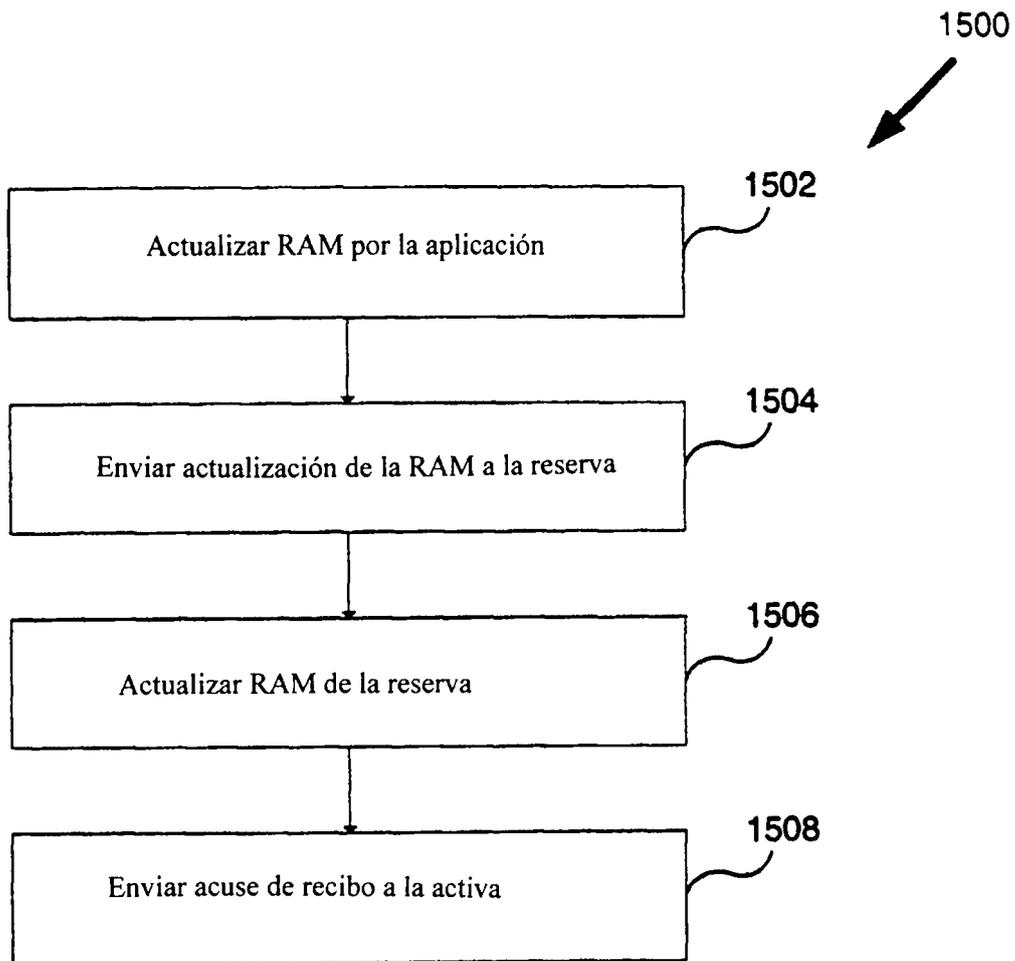


FIG. 15

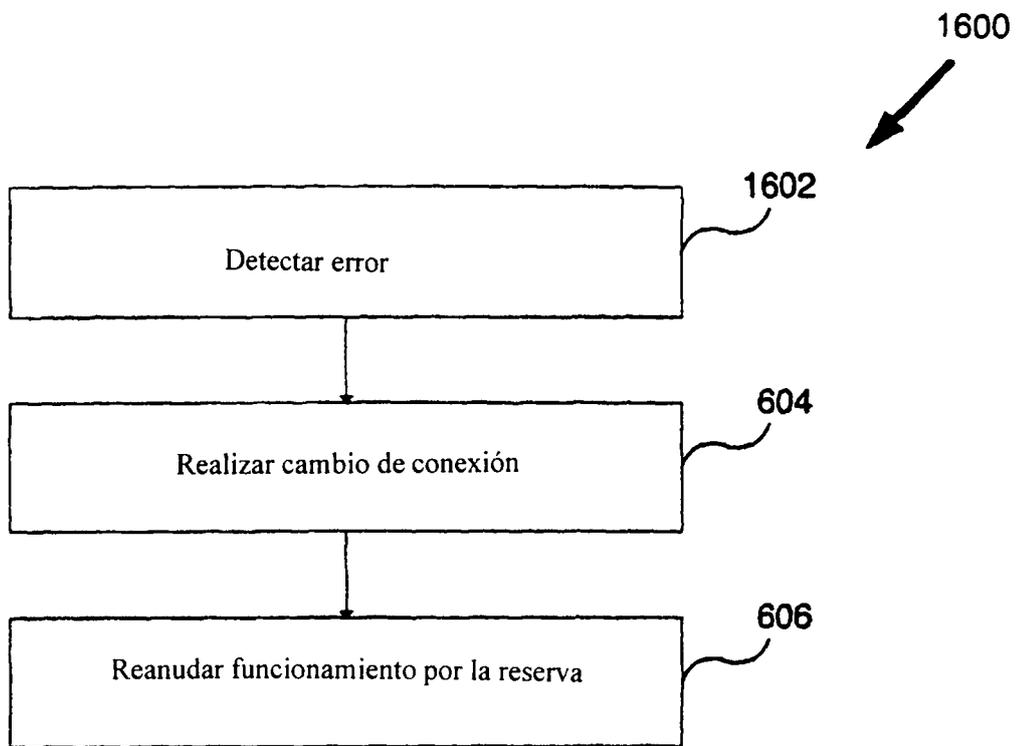


FIG. 16

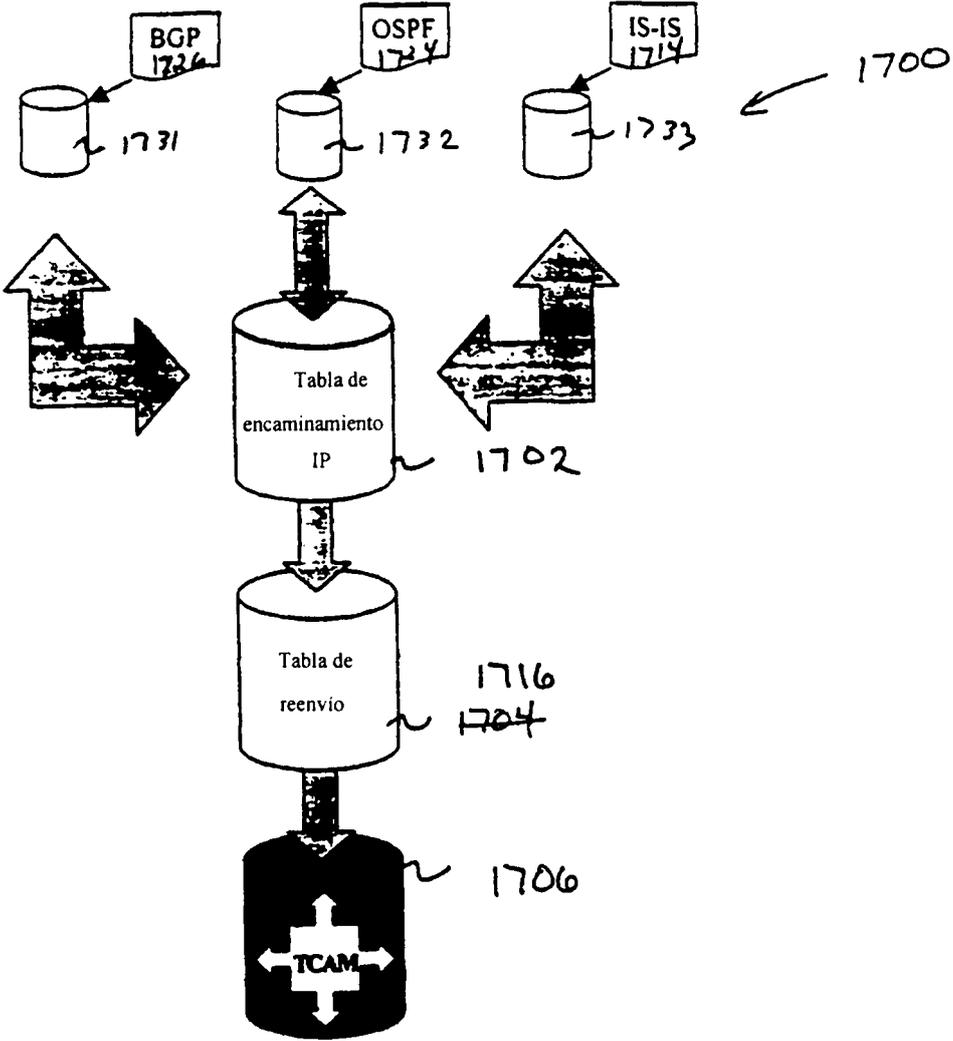


Figura 17

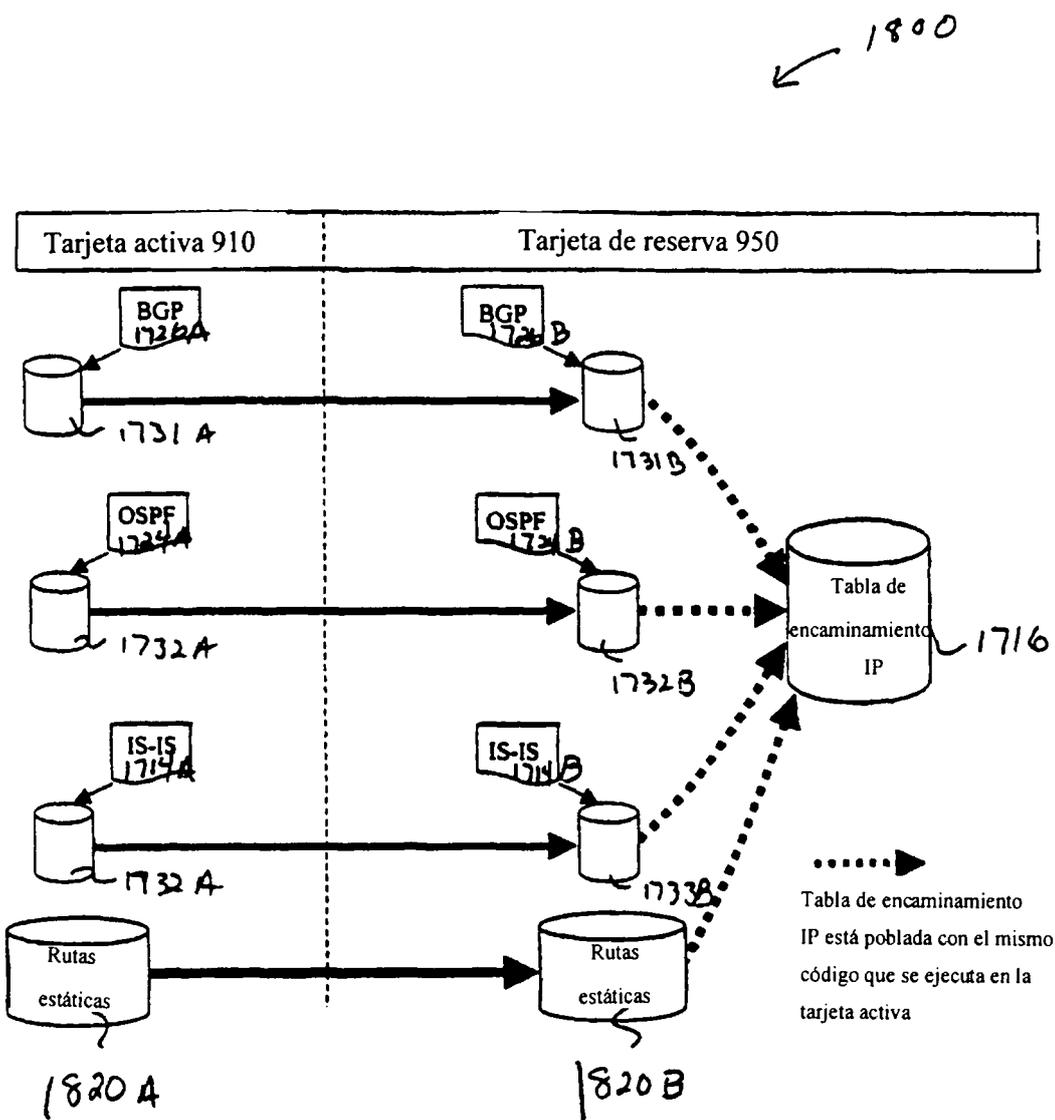


Figura 18

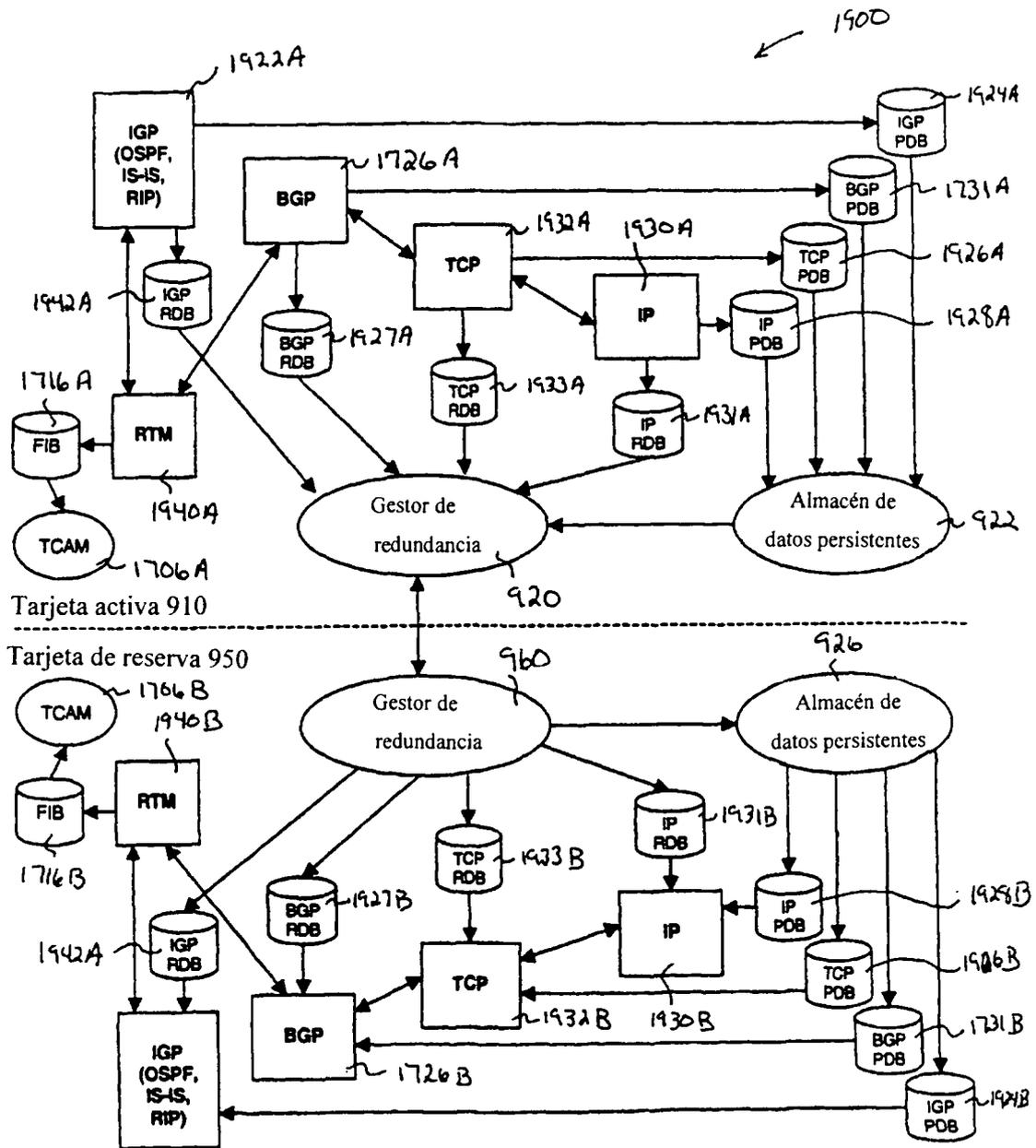


Figura 19

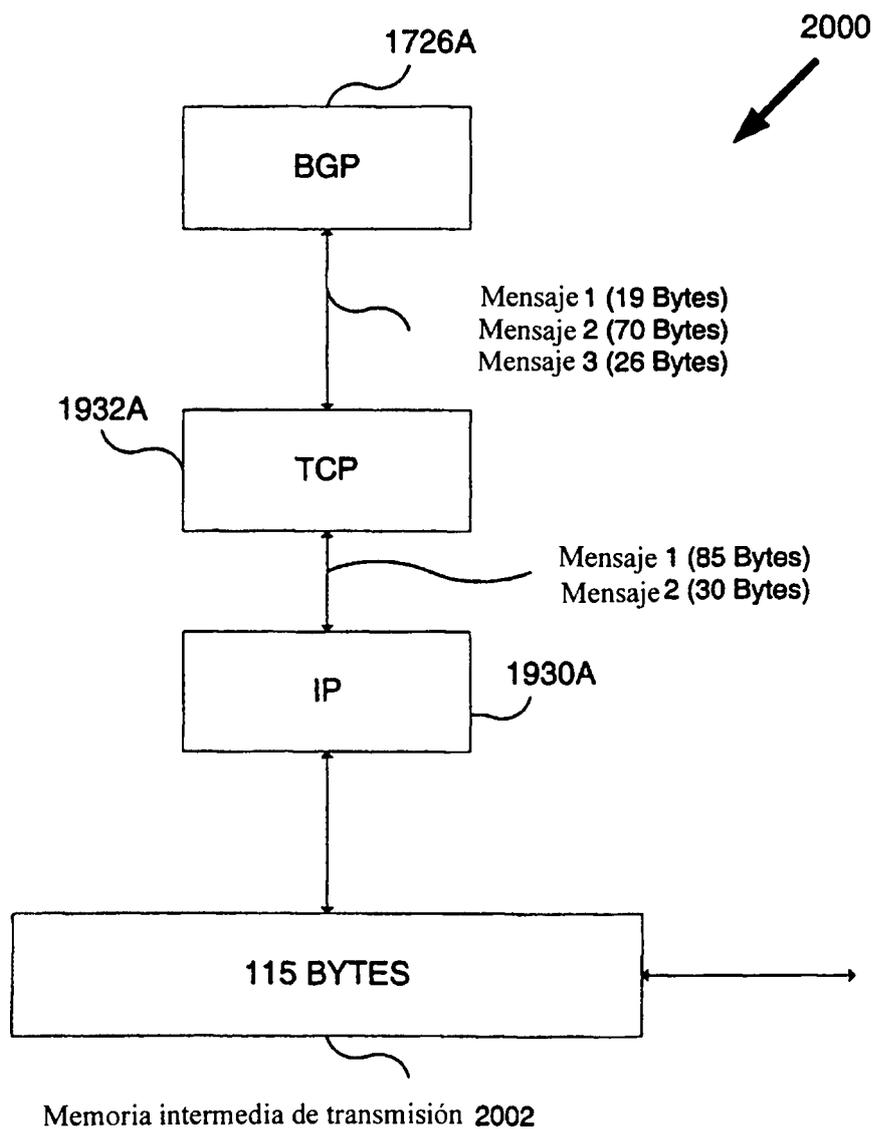


FIG. 20

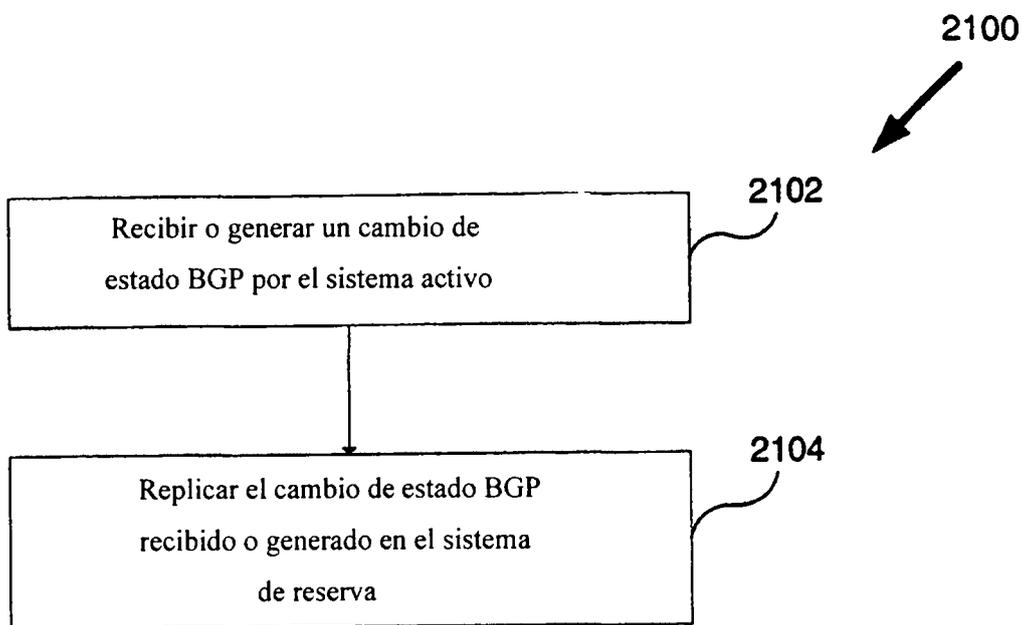


FIG. 21A

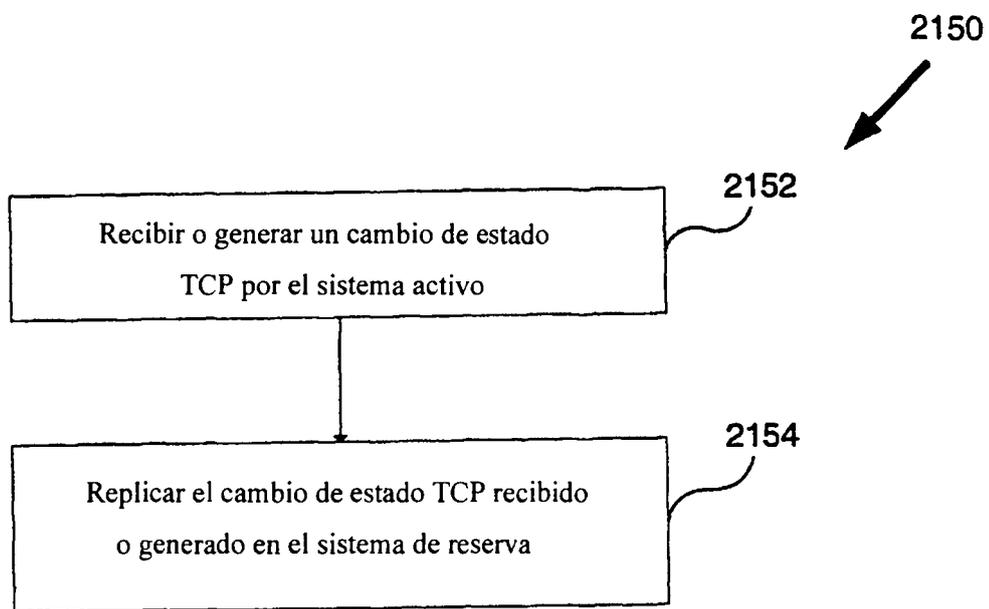


FIG. 21B

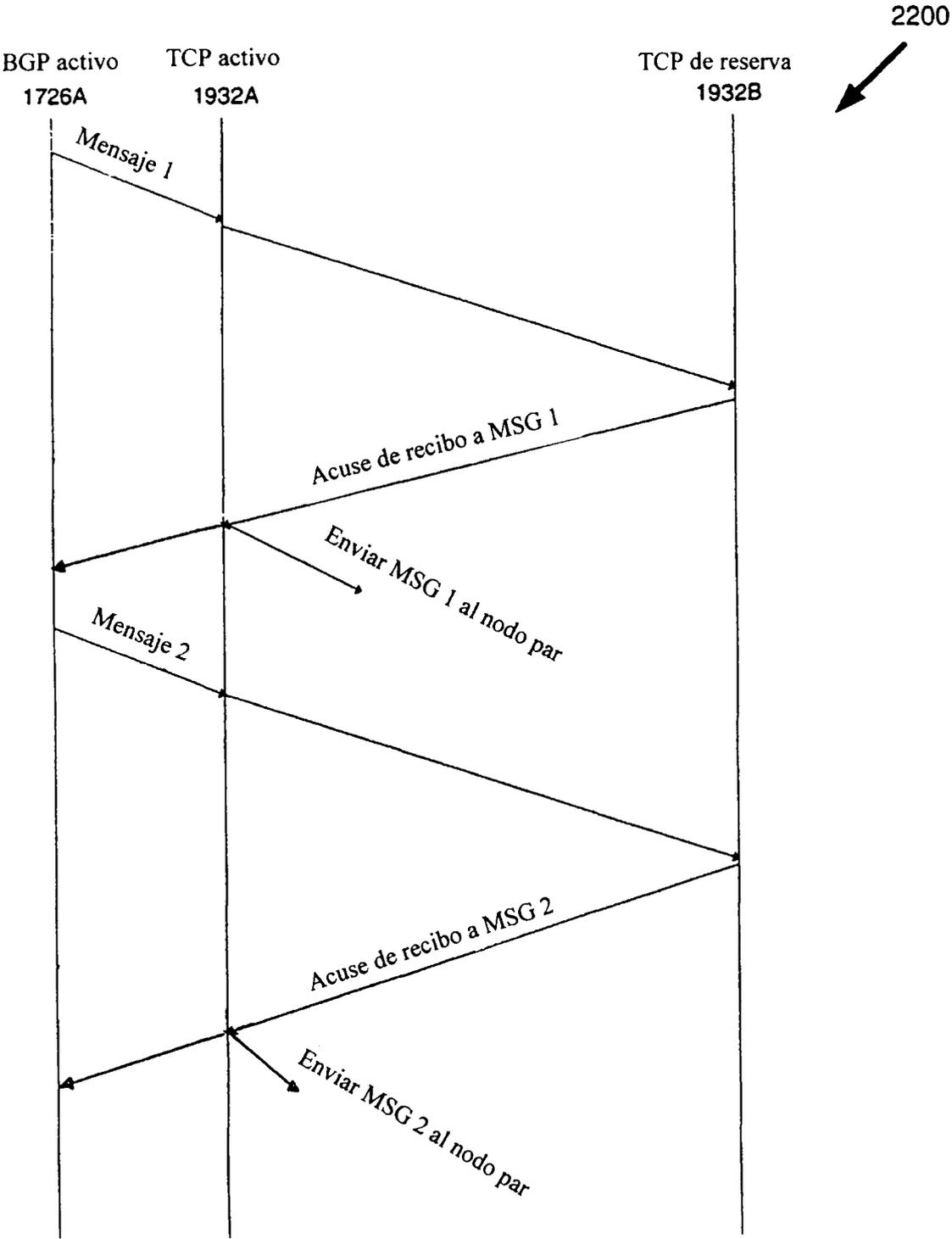


FIG. 22

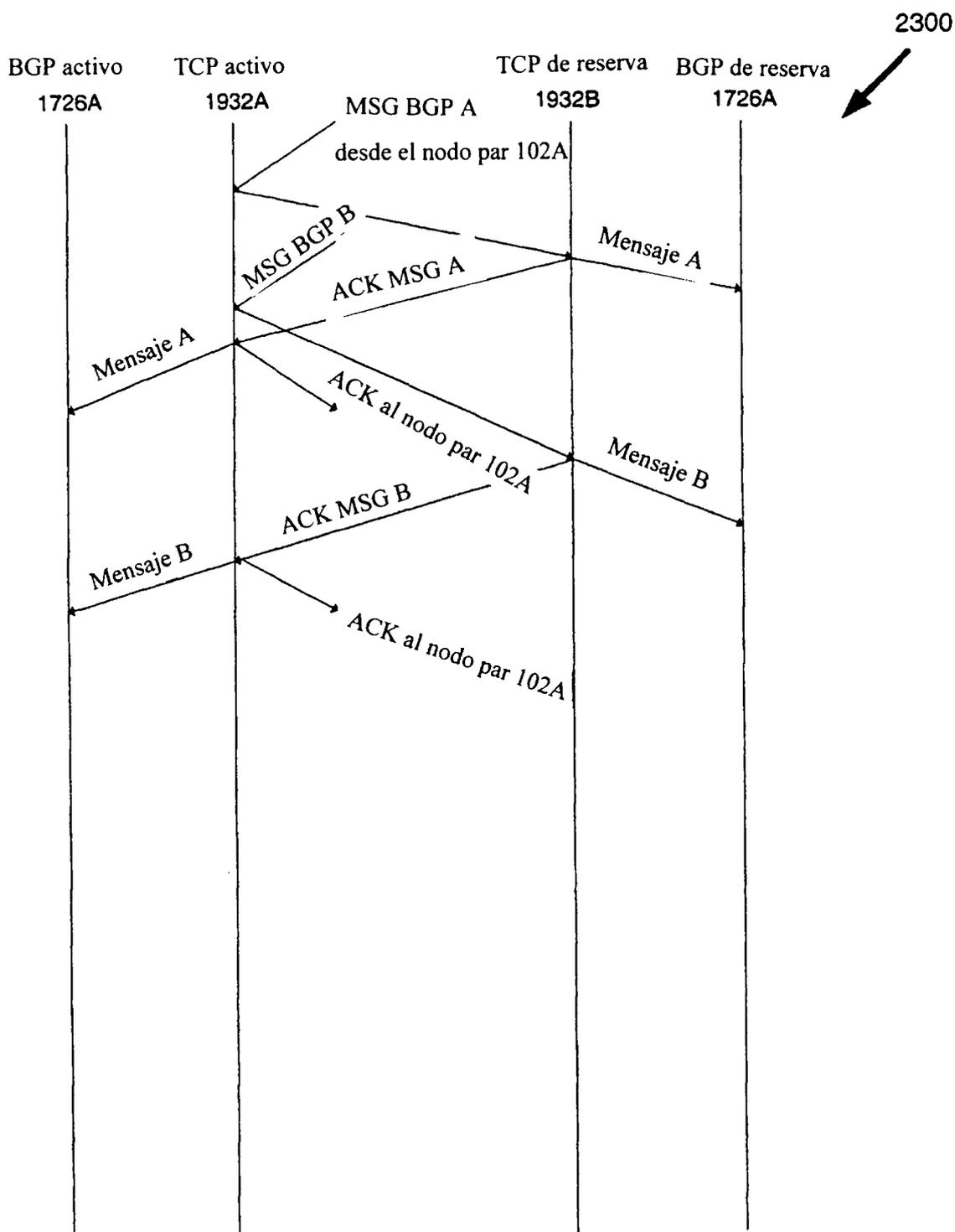


FIG. 23

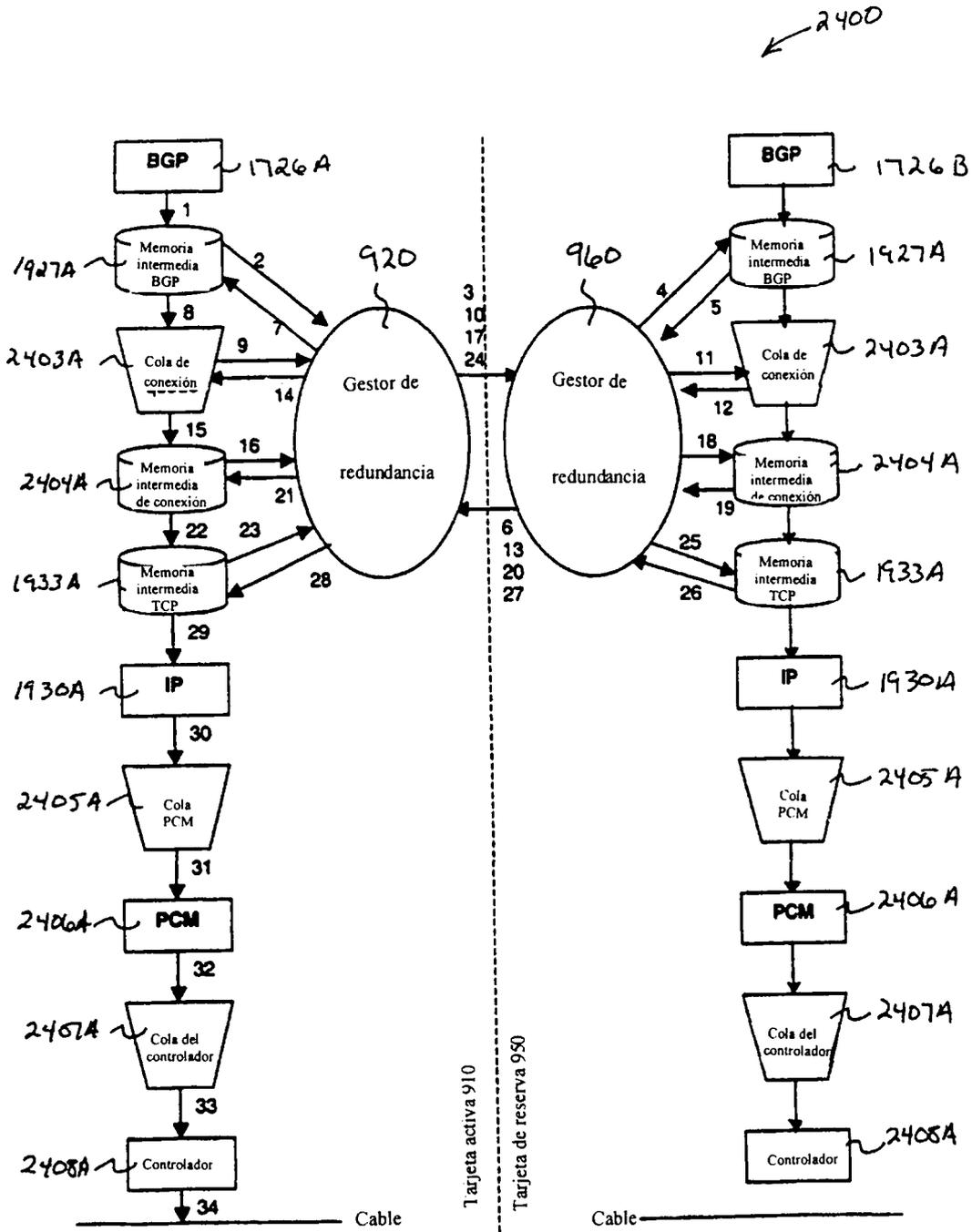


Figura 24

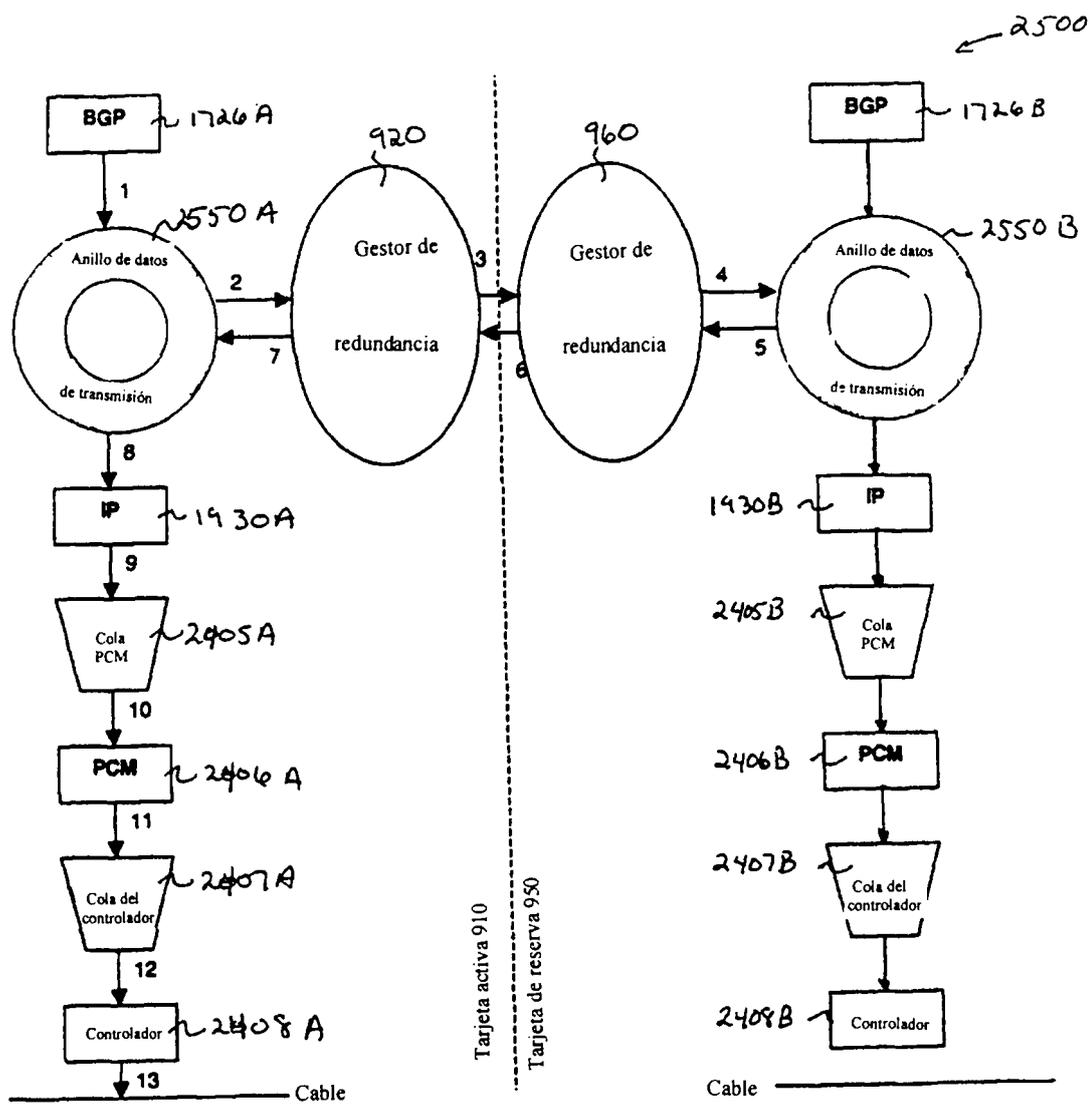


Figura 25

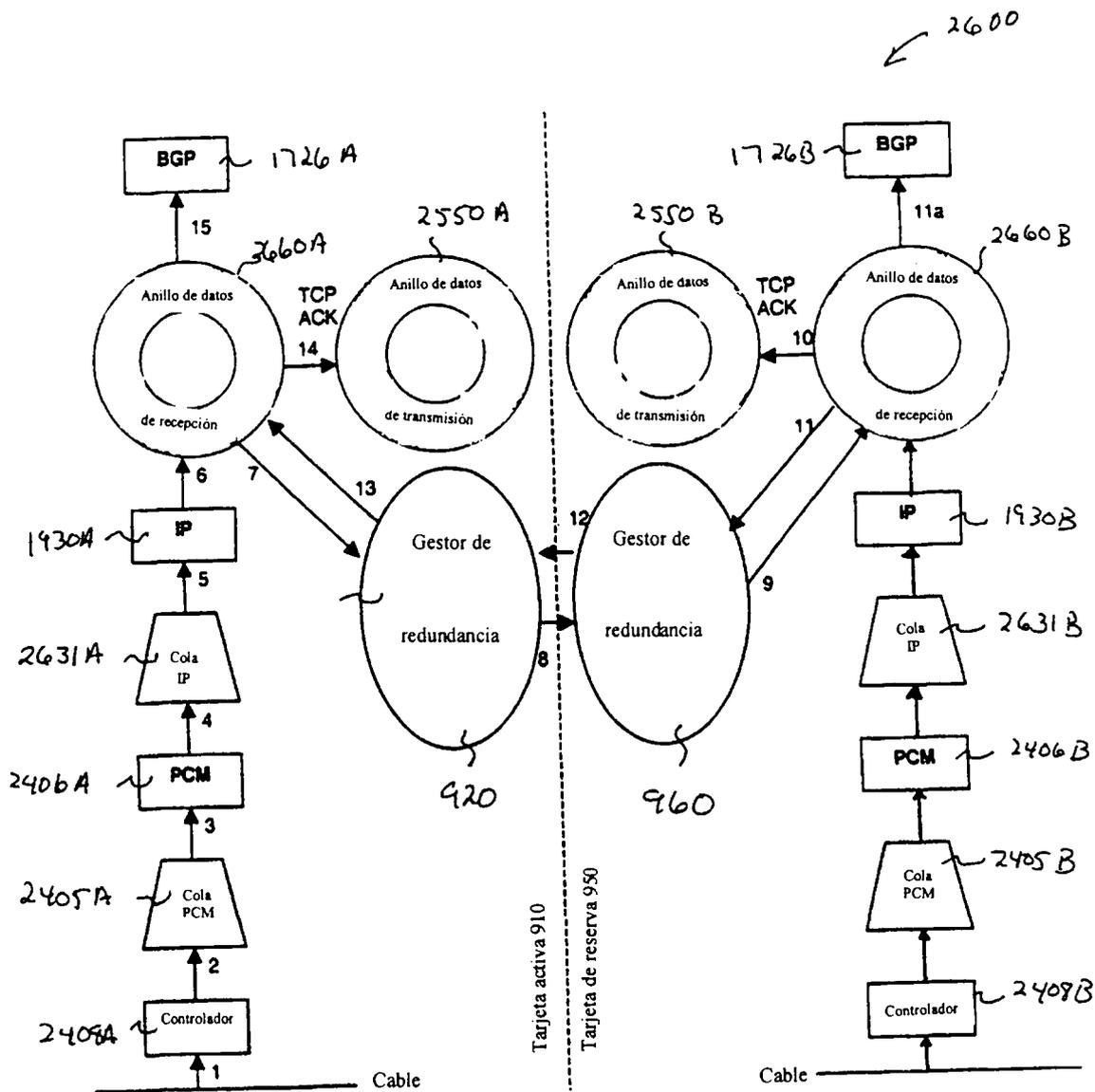


Figura 26

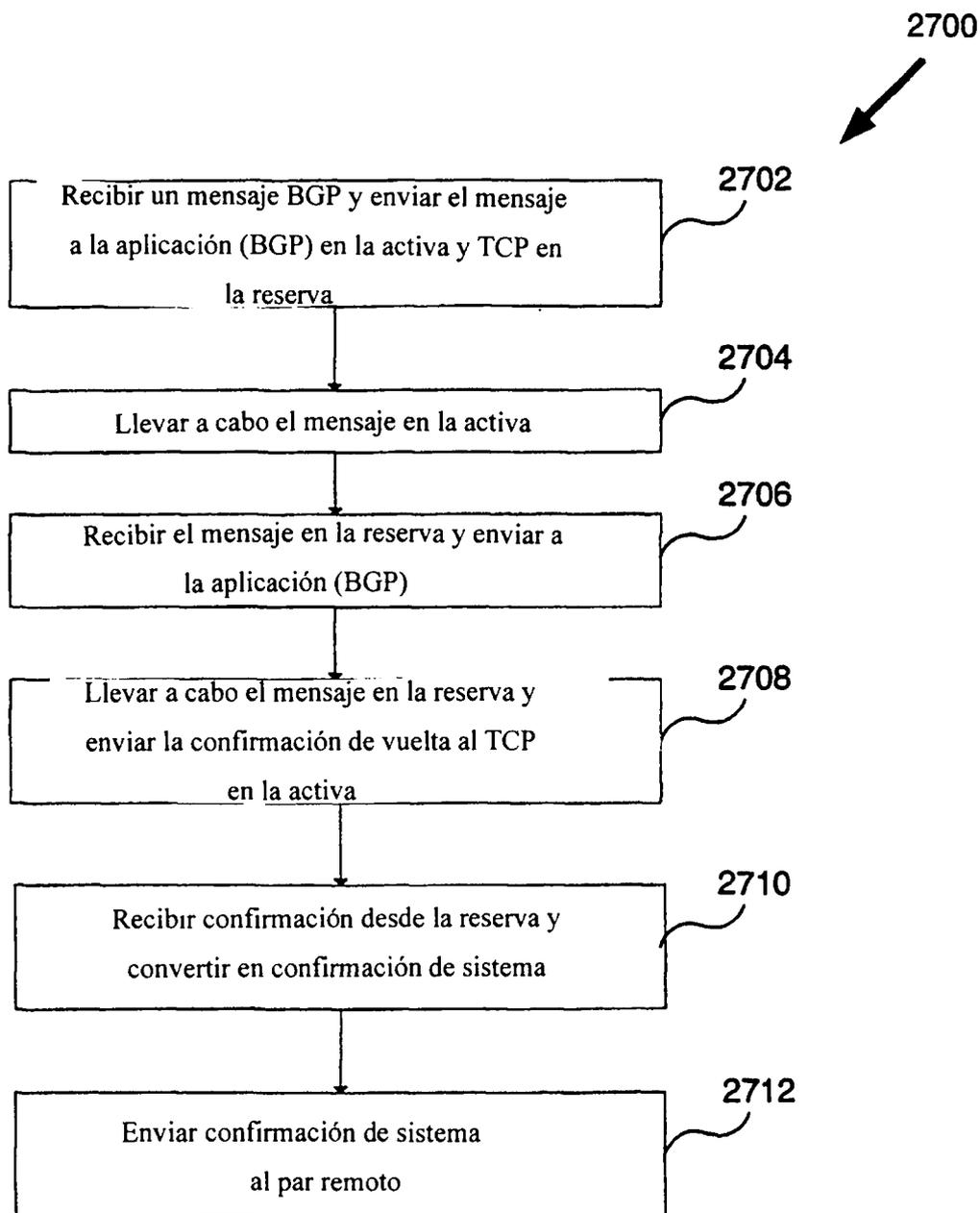


FIG. 27

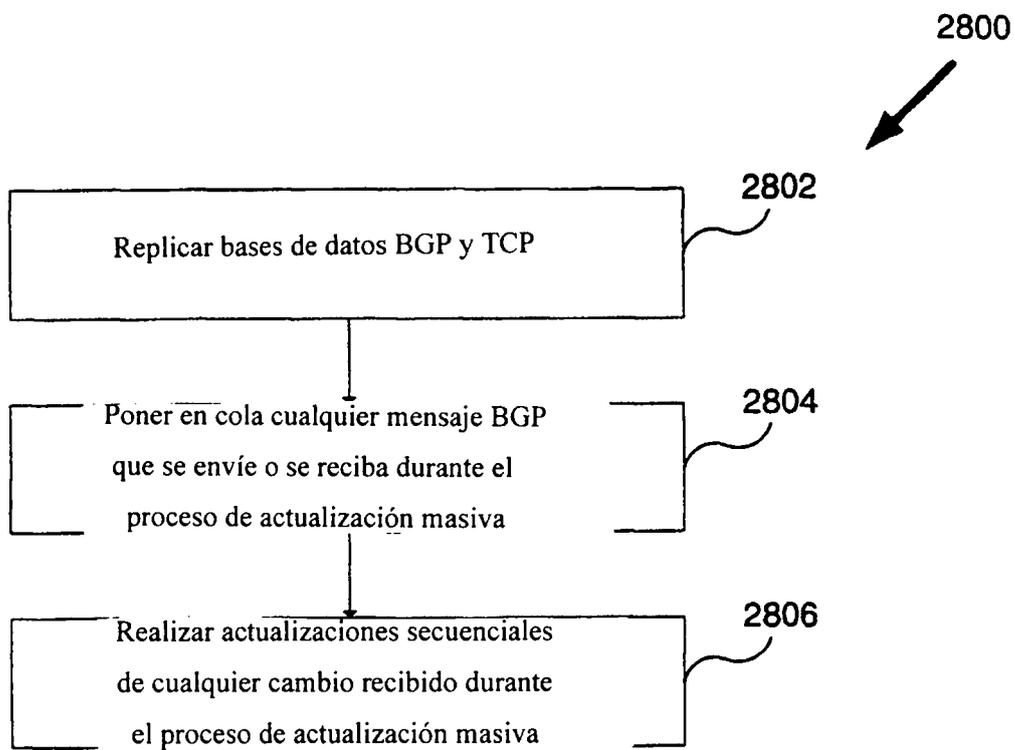


FIG. 28

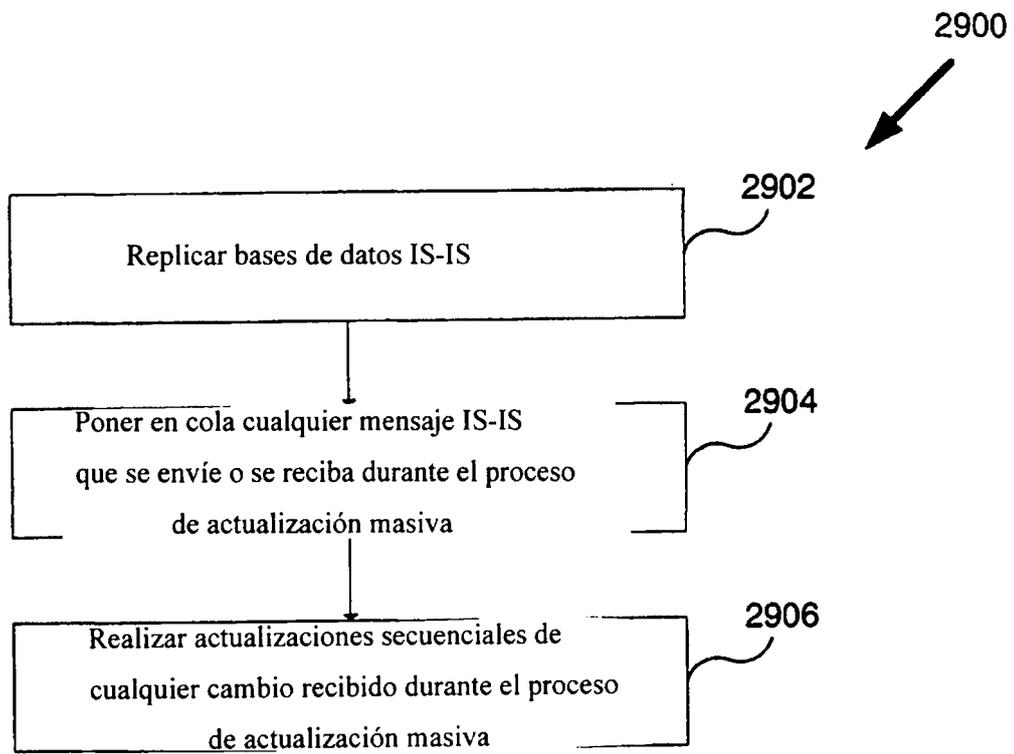


FIG. 29

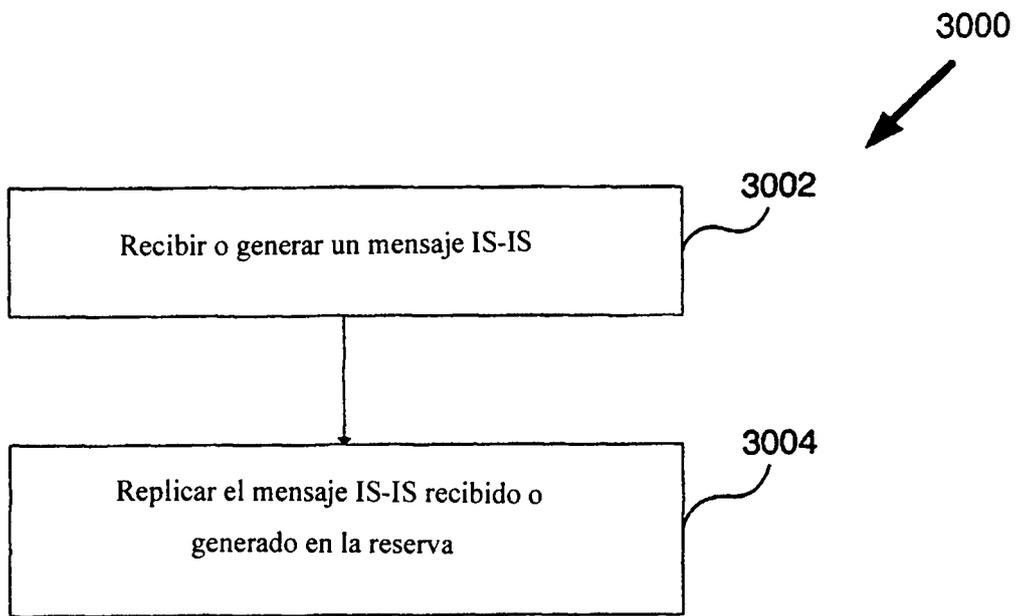


FIG. 30

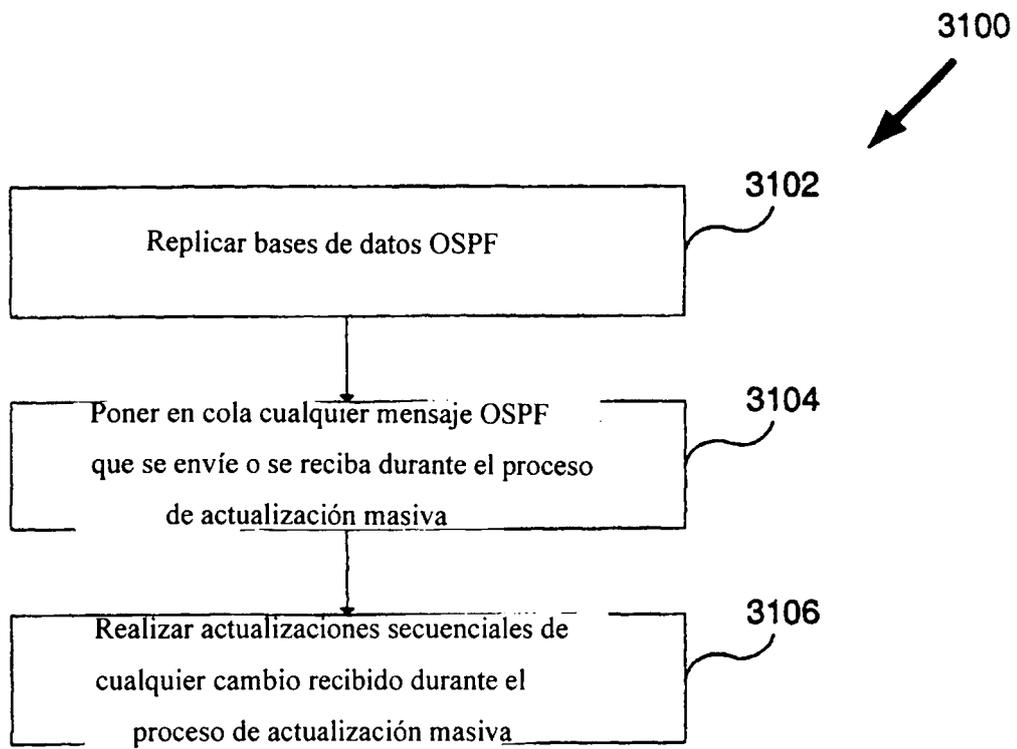


FIG. 31

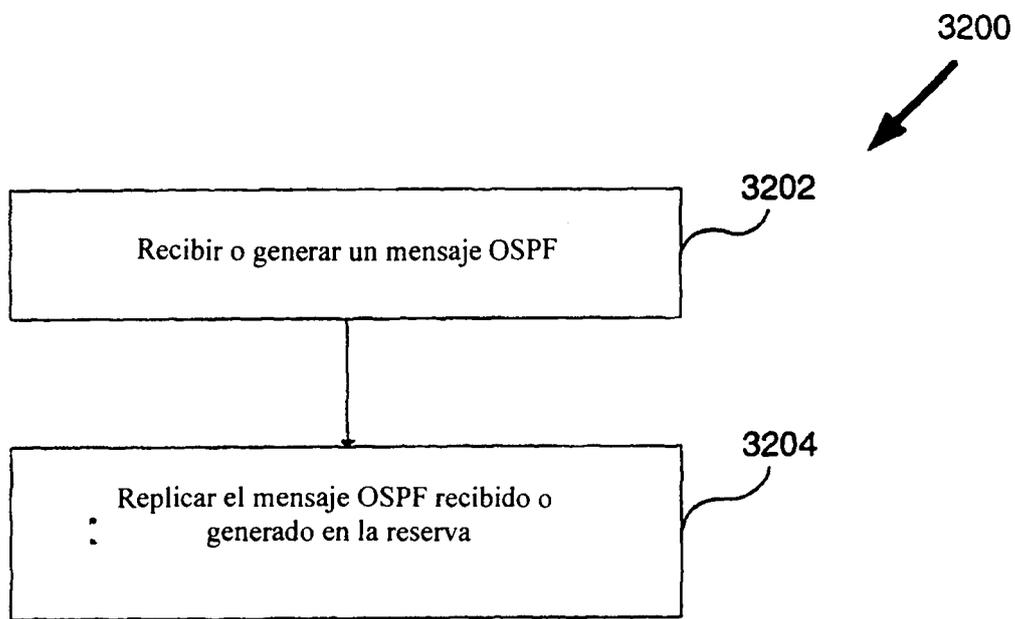


FIG. 32