(19) OFICINA ESPAÑOLA DE PATENTES Y MARCAS

ESPAÑA

(11) Número de publicación: **2 862 428**

(51) Int. CI.:

H04L 12/24 (2006.01)
G06F 11/20 (2006.01)
H04L 29/06 (2006.01)
G06F 11/14 (2006.01)
G06F 11/18 (2006.01)

(12) TRADUCCIÓN DE PATENTE EUROPEA T3

(86) Fecha de presentación y número de la solicitud internacional: **18.03.2019 PCT/CN2019/078549**

(87) Fecha y número de publicación internacional: **31.05.2019 WO19101244**

(96) Fecha de presentación y número de la solicitud europea: **18.03.2019 E 19725901 (3)**

(97) Fecha y número de publicación de la concesión europea: **20.01.2021 EP 3580913**

(54) Título: **Recuperación de tiempo de inactividad de sistema de consenso**

(45) Fecha de publicación y mención en BOPI de la traducción de la patente:
07.10.2021

(73) Titular/es:

ADVANCED NEW TECHNOLOGIES CO., LTD. (100.0%)
Cayman Corporate Centre, 27 Hospital Road
George Town, Grand Cayman KY1-9008, KY

(72) Inventor/es:

YANG, DAYI

(74) Agente/Representante:

VALLEJO LÓPEZ, Juan Pedro

DESCRIPCIÓN

Recuperación de tiempo de inactividad de sistema de consenso

Campo técnico

Esta solicitud se refiere en general a procedimientos y dispositivos para un sistema y método de consenso, y en particular, a un sistema y método de consenso de Tolerancia Práctica a Fallos Bizantinos (PBFT).

Antecedentes

La Tolerancia Práctica a Fallos Bizantinos (PBFT) es un tipo de mecanismo de consenso que puede implementarse en sistemas distribuidos tales como sistemas de cadena de bloques. El mecanismo de consenso de PBFT posibilita que un sistema distribuido alcance un consenso suficiente con seguridad y ejecución, a pesar de que pueden fallar ciertos nodos del sistema (por ejemplo, debido a conexión de red pobre o que se vuelvan defectuosos de otra manera) o se propague información incorrecta a otros clientes (por ejemplo, actuando de manera maliciosa). El objetivo de tal mecanismo es defenderse contra fallos de sistema catastróficos mitigando la influencia de los nodos que no funcionan en la función correcta del sistema y en el consenso alcanzado mediante nodos que funcionan (por ejemplo, nodos no defectuosos y honestos) en el sistema.

El mecanismo de consenso de PBFT se enfoca en proporcionar una replicación práctica de máquina de estado bizantino que tolera los fallos bizantinos (por ejemplo, nodos que no funcionan) a través de una suposición de que hay fallos de nodo independientes y mensajes manipulados propagados mediante nodos específicos e independientes. En este mecanismo de consenso de PBFT, por ejemplo, todos los nodos en un sistema de cadena de bloques se ordenan en una secuencia siendo un nodo el nodo primario (también conocido como el nodo líder o maestro) y denominados los otros como los nodos de respaldo (también conocidos como nodos seguidores). Todos los nodos en el sistema se comunican entre sí y el objetivo es para que todos los nodos honestos entren en un acuerdo/consenso en un estado del sistema.

Por ejemplo, para que funcione el mecanismo de consenso de PBFT, la suposición es que la cantidad de nodos que no funcionan en un sistema de cadena de bloques no pueden equivaler de manera simultánea o superar un tercio de los nodos globales en el sistema en una ventana de vulnerabilidad dada. El método proporciona eficazmente tanto la ejecución como la seguridad siempre que la mayoría de los nodos F sean nodos que no funcionan al mismo tiempo. En otras palabras, en algunas implementaciones, el número F de nodos que no funcionan que pueden tolerarse mediante el mecanismo de consenso de PBFT equivale a (N-1)/3, redondeado hacia abajo hasta el número entero más cercano, en donde N designa el número total de nodos en el sistema. En algunas implementaciones, un sistema de cadena de bloques que implementa el mecanismo de consenso de PBFT puede manejar hasta F fallos bizantinos donde hay al menos 3F+1 nodos en total.

El mecanismo de consenso de PBFT comprende en general un protocolo de operación normal (también conocido como el protocolo de triple etapa) y un protocolo de cambio de vista, en donde se proporciona el protocolo de operación normal para asegurar la seguridad del mecanismo, mientras que se proporciona el protocolo de cambio de vista para asegurar la ejecución del mecanismo. El protocolo de etapa normal incluye principalmente tres fases en orden, es decir, una fase de preparación previa, una fase de preparación y una fase de confirmación. Todas las fases están activadas por mensaje, es decir, se activa una siguiente fase en el protocolo obteniendo un número suficiente de mensajes en una fase actual. La totalidad del proceso bajo el protocolo de operación normal se avanza en gran medida dependiendo de un número suficiente de mensajes recibidos de manera consecutiva en cada fase. Incluso en el protocolo de cambio de vista, el proceso se avanza basándose en los mensajes de preparación en el protocolo de operación normal. Por lo tanto, puede observarse que el mecanismo de consenso de PBFT se basa enormemente en que funcionen los mensajes de consenso. Si uno o más nodos se vuelven no funcionales (por ejemplo, experimentan tiempo de inactividad y reinicio), los mensajes almacenados en la memoria se perderán, lo que afecta al proceso de consenso total, incluso incurriendo en inconsistencia. El documento "Enhancing Bitcoin Security and Performance with Strong Consistency via Collective Signing" (Eleftherios Kokoris Kogias *et al*) introduce ByzCoin, un protocolo de consenso bizantino novedoso que aprovecha la firma colectiva escalable para confirmar transacciones de Bitcoin de manera irreversible en segundos.

Sumario

Los aspectos de la invención proporcionan un método de consenso basado en tolerancia práctica a fallos bizantinos implementado por ordenador, un aparato de consenso y un sistema de consenso como se expone en las reivindicaciones adjuntas.

Breve descripción de los dibujos

La Figura 1 ilustra una red, de acuerdo con diversas realizaciones.

La Figura 2A ilustra un protocolo de operación normal de PBFT.

La Figura 2B ilustra un protocolo de operación normal de PBFT con un fallo

La Figura 2C ilustra un protocolo de operación normal y un protocolo de cambio de vista de PBFT.

La Figura 3A ilustra un diagrama de flujo de las etapas de un protocolo de operación normal de PBFT.

La Figura 3B ilustra un diagrama de flujo de las etapas de un protocolo de cambio de vista de PBFT.

La Figura 3C ilustra un diagrama de flujo de las etapas de un protocolo de operación normal de un sistema de consenso, de acuerdo con diversas realizaciones.

La Figura 4A-Figura 4D cada una ilustra un diagrama de flujo de etapas de consenso, de acuerdo con diversas realizaciones.

La Figura 5A ilustra un diagrama de flujo de un método de consenso, de acuerdo con diversas realizaciones.

La Figura 5B ilustra un diagrama de flujo de un método de consenso, de acuerdo con diversas realizaciones.

La Figura 6A ilustra un diagrama de bloques de un sistema de consenso, de acuerdo con diversas realizaciones.

La Figura 6B ilustra un diagrama de bloques de un sistema de consenso, de acuerdo con diversas realizaciones.

La Figura 7 ilustra un diagrama de bloques de un sistema informático en el que puede implementarse cualquiera de las realizaciones descritas en el presente documento.

Descripción detallada

Las realizaciones desveladas en el presente documento incluyen, pero sin limitación, sistemas de recuperación de tiempo de inactividad de PBFT, métodos, y un medio legible por ordenador no transitorio. En diversas realizaciones, un sistema de red distribuido tal como un sistema de cadena de bloques puede comprender una pluralidad de nodos. El sistema de cadena de bloques puede implementar un mecanismo de consenso de PBFT, con uno de la pluralidad de nodos designados como un nodo primario y los otros nodos como nodos de respaldo. De acuerdo con algunas realizaciones, para cada ronda de verificación de consenso ejecutada en el sistema de cadena de bloques, únicamente se almacena una parte de los mensajes de consenso, en lugar de todos ellos. Por ejemplo, se almacena un mensaje de preparación previa y un número suficiente de mensajes de preparación durante el protocolo de operación normal. En algunas realizaciones, únicamente se almacena el mensaje de preparación previa y (Q-1) mensajes de preparación. Q indica quórum y es (N+F+1)/2 redondeado hacia arriba hasta el número entero más cercano, y F es (N-1)/3 redondeado hacia abajo hasta el número entero más cercano. De esta manera, es posible reanudar y avanzar el proceso de verificación de consenso de manera eficaz y de manera eficiente de cualquier interrupción (por ejemplo, una avería a nivel de sistema), con menos consumo de almacenamiento de sistema y sin provocar resultados de consenso inconsistentes y ramificación a la cadena de bloques. Similar a PBFT, los sistemas, métodos, y el medio legible por ordenador no transitorio desvelados pueden aplicarse a otros protocolos de consenso tales como SecureRing, Byzantine Paxos, Q/U, HQ, Zyzzvyva, ABsTRACTs, RBFT, Adapt, Tangaroa, CheapBFT, MinBFT, FastBFT, etc. Puede hacerse referencia para diversos aspectos de PBFT a M. Castro, B. Liskov, "Practical Byzantine Fault Tolerance", Proceedings of the Third Symposium on Operating Systems Design and Implementation, (Feb 1999), que se incorpora por referencia en el presente documento en su totalidad.

La Figura 1 muestra una red 120, de acuerdo con diversas realizaciones. Los componentes presentados a continuación se pretende que sean ilustrativos. Como se muestra, la red 120 puede comprender un sistema de red distribuido 112 tal como un sistema de cadena de bloques. El sistema de red 112 puede comprender uno o más nodos (por ejemplo, Nodo 0, Nodo 1, Nodo 2, Nodo 3, Nodo 4... Nodo i, ..., etc.) implementados en uno o más dispositivos informáticos tales como servidores, ordenadores, teléfonos móviles, etc. El sistema de red 112 puede estar instalado con software apropiado (por ejemplo, programa de consenso) y/o hardware (por ejemplo, cables, conexiones inalámbricas) para acceder a otros dispositivos de la red 120 o sistemas adicionales. Cada nodo puede incluir uno o más procesadores y una o más memorias acopladas al uno o más procesadores. Por ejemplo, la una o más memorias son no transitorias y legibles por ordenador, y están configuradas con instrucciones ejecutables por uno o más procesadores para hacer que el uno o más procesadores realicen operaciones descritas en el presente documento. Aunque se muestran los nodos como únicos componentes en esta figura, se apreciará que estos nodos pueden implementarse como dispositivos únicos o múltiples dispositivos acoplados juntos. En general, los nodos pueden comunicarse entre sí y con otros dispositivos fuera del sistema de red 112. Por ejemplo, pueden comunicarse datos a través de una o más redes alámbricas o inalámbricas (por ejemplo, Internet).

En diversas realizaciones, el sistema de red 112 puede implementarse como un sistema de cadena de bloques que comprende una pluralidad de nodos. Por ejemplo, como se muestra en la Figura 1, el sistema de cadena de bloques

comprende una pluralidad de nodos de cadena de bloques (por ejemplo, Nodo 0, Nodo 1, Nodo 2, Nodo 3, Nodo 4, ...nodo i, ... etc.). Los nodos pueden formar una red (por ejemplo, red entre pares), comunicándose un nodo de cadena de bloques con otro. El orden y el número de los nodos de cadena de bloques como se muestran son simplemente ejemplos y por simplicidad de ilustración. Los nodos de cadena de bloques pueden implementarse en servidores, ordenadores, etc. Cada nodo de cadena de bloques puede corresponder a uno o más dispositivos de hardware físicos o dispositivos virtuales acoplados juntos mediante diversos tipos de métodos de comunicación tales como TCP/IP. Dependiendo de las clasificaciones, los nodos de cadena de bloques pueden comprender nodos totales, nodos Geth, nodos de consenso, etc.

En diversas realizaciones, el sistema de cadena de bloques puede interactuar con otros sistemas y dispositivos tales como el nodo A y el nodo B (por ejemplo, nodos ligeros). Las interacciones pueden implicar la transmisión y recepción de datos para el fin de, por ejemplo, recibir una solicitud y devolver un resultado de la ejecución de la solicitud. En un ejemplo, el usuario A puede desear transaccionar con el usuario B a través de la cadena de bloques. La transacción puede implicar transferir algún activo en la cuenta del usuario A a la cuenta del usuario B. El usuario A y el usuario B pueden usar respectivos nodo A y nodo B de los dispositivos instalados con un software de cadena de bloques apropiado (por ejemplo, monedero de criptomoneda) para la transacción. El nodo A puede acceder a la cadena de bloques a través de la comunicación con el Nodo 0, y el nodo B puede acceder a la cadena de bloques a través de la comunicación con el Nodo 1. Por ejemplo, el Nodo A puede enviar una solicitud de transacción a la cadena de bloques a través del Nodo 0, y el Nodo B puede enviar una solicitud de ejecución de contrato inteligente a la cadena de bloques a través de Nodo 1. Fuera de la cadena de bloques, el nodo A y el nodo B pueden tener otros canales de comunicación (por ejemplo, comunicación de internet normal sin pasar a través de los Nodos 0 y 1).

Los nodos de cadena de bloques pueden cada uno comprender o acoplarse a una memoria. En algunas realizaciones, la memoria puede almacenar una base de datos de agrupación. La base de datos de agrupación puede ser accesible para la pluralidad de nodos de cadena de bloques de una manera distribuida. Por ejemplo, la base de datos de agrupación puede almacenarse respectivamente en las memorias de los nodos de cadena de bloques. La base de datos de agrupación puede almacenar una pluralidad de transacciones enviadas mediante el uno o más dispositivos de usuario tales como los nodos A y B operados por los usuarios.

Los nodos de cadena de bloques forman una red (por ejemplo, una red P2P) que, a través de consenso, registra transacciones en un libro mayor distribuido conocido como cadena de bloques. Los participantes de una red P2P pueden denominarse como nodos, que mantienen la cadena de bloques. En una red P2P de cadena de bloques, cada nodo participa en las verificaciones de consenso y almacena una copia de libro mayor completa de la cadena de bloques. Cada nodo confirma lotes de transacciones mediante un algoritmo de consenso de cadena de bloques para asegurar que todos los nodos tienen resultados de confirmación consistentes y por lo tanto copias consistentes de la cadena de bloques.

Uno de los algoritmos de consenso de cadena de bloques es la Tolerancia Práctica a Fallos Bizantinos (PBFT). La tolerancia a fallos bizantinos se origina a partir del problema general bizantino. Para un sistema de red P2P, siempre que el número de nodos que no funcionen esté dentro de un cierto límite, el sistema puede continuar funcionando de manera apropiada. Tal sistema se denomina sistema tolerante a fallos bizantinos. PBFT es un ejemplo de una optimización de la capacidad de red de tolerancia a fallos bizantinos. PBFT proporciona a la red con una máquina de estado bizantina, que copia servidores y que sincroniza interacciones de cliente con copias de servidor.

En el centro de la operación PBFT está el mantenimiento de la vista global consistente de la información registrada en la cadena de bloques, que forma la red troncal para posibilitar que los usuarios interactúen entre sí de una manera descentralizada. La seguridad del mecanismo de consenso de PBFT es crítica para un sistema de cadena de bloques. Las dos propiedades clave de un modelo de consenso son: 1) seguridad o consistencia: todos los nodos honestos producen la misma salida válida; y 2) ejecución: todos los nodos honestos en consenso producen eventualmente un valor sin quedarse estancados en una etapa intermedia. Un mecanismo de consenso de PBFT seguro y de consenso necesita tolerar una amplia diversidad de comportamientos bizantinos, que incluyen caídas de nodos, partición de la red, retardo de mensaje, entrega de mensaje fuera de orden, corrupción de mensaje y similares y alcanza consenso en nodos siempre que el número de nodos que no funcionen en el sistema esté limitado. Para este fin, el modelo PBFT funciona bajo cualquiera de uno de dos protocolos mutuamente exclusivos: protocolo de operación normal/consistencia y protocolo de cambio de vista que se describen adicionalmente a continuación. En esta memoria descriptiva, que no funciona significa defectuoso y/o malicioso, y que funciona significa no defectuoso y honesto. Posibles actos defectuosos y/o maliciosos incluyen: fallo al entregar mensaje, retardo de entrega de mensaje, retardo de mensaje fuera de orden, fallos bizantinos (entregar mensajes arbitrarios a diferentes nodos, violar el protocolo), etc.

En algunas realizaciones, un sistema de cadena de bloques que implementa el mecanismo de Tolerancia Práctica a Fallos Bizantinos (PBFT) puede comprender un número total de N nodos, actuando uno de los N nodos como un nodo primario y actuando el otro de los N nodos como nodos de respaldo. La designación de nodo primario puede no estar fijada a un nodo particular, ya que puede elegirse otro nodo para que se vuelva un nuevo nodo primario a través del protocolo de cambio de vista. Por ejemplo, puede elegirse el nodo primario a través de una operación módulo, en la que un nodo que funciona con el número de serie más bajo (número de vista de modulo) se vuelve el nuevo nodo primario. La vista actual y el número total de nodos N puede determinar el id de nodo primario = (vista+1) mod N. En

PBFT, la vista se cambia cada vez que se elige un nuevo nodo primario. Por ejemplo, con cada cambio de vista, la vista aumenta de manera monotónica desde cero. Es decir, la vista puede cambiar con un cambio en el nodo primario.

En algunas realizaciones, el nodo primario está funcionando en la vista v, y se ejecuta el protocolo de operación normal. Para la operación normal, el nodo primario y/o los nodos de respaldo pueden recibir solicitudes asociadas con transacciones no verificadas desde uno o más clientes. Por ejemplo, el nodo A como un cliente puede enviar una solicitud al nodo primario y/o a los nodos de respaldo. La solicitud puede incluir una transacción no verificada (por ejemplo, una transacción que va a añadirse en un nuevo bloque en la cadena de bloques mediante verificación de consenso). Las transacciones no verificadas pueden incluir, por ejemplo, transacciones financieras basadas en cadena de bloques, transacciones de implementación o ejecución de contratos inteligentes, etc. Los nodos primario y de respaldo pueden realizar o no alguna verificación preliminar de las transacciones. Los nodos de respaldo que reciben las solicitudes pueden reenviar las solicitudes recibidas al nodo primario. Una vez que las solicitudes con transacciones no verificadas en el nodo primario alcanzan un cierto nivelo o cumplen de otra manera con una condición de activación, el nodo primario puede iniciar una ronda de verificación de consenso y proponer un resultado de verificación para las transacciones no verificadas. Los nodos de respaldo pueden responder al consenso y confirmar la propuesta para alcanzar un consenso. Los requisitos para los nodos son que son deterministas y empiezan en el mismo estado. El resultado final es que todos los nodos honestos entran en un consenso en el orden del registro y pueden aceptarlo o rechazarlo. Una vez verificadas por consenso, las transacciones pueden empaquetarse en un nuevo bloque de la cadena de bloques y añadirse a las copias de la cadena de bloques locales mantenidas mediante los nodos. También, son notificados los clientes (por ejemplo, el nodo A) que enviaron originalmente las solicitudes.

Como se ha indicado anteriormente, para conservar la seguridad, el mecanismo de consenso de PBFT comprende principalmente tres fases para el protocolo de operación normal: una fase de preparación previa, una fase de preparación y una fase de confirmación. Haciendo referencia a la Figura 2A a la Figura 2C, un ejemplo de un sistema de cadena de bloques que implementa el mecanismo de consenso de PBFT comprende cuatro réplicas (siendo la réplica otro término para nodo): Réplica 0, Réplica 1, Réplica 2 y Réplica 3. Los números 0 a 3 son números de serie de réplica que pueden usarse para determinar un nuevo nodo primario. La réplica 0 puede corresponder al nodo primario 0, y las réplicas 1, 2, y 3 pueden corresponder a los nodos de respaldo 1, 2, y 3. Las réplicas pueden implementarse, por ejemplo, en nodos correspondientes del sistema de red 112 anteriormente descrito. Un protocolo de operación normal se muestra en la Figura 2A con ningún nodo que no funcione presente, y otro protocolo de operación normal se muestra en la Figura 2B siendo la réplica 3 un nodo que no funciona. Para ambas situaciones, el protocolo de operación normal puede comprender adicionalmente dos fases: una fase de solicitud y una fase de respuesta, además de la fase de preparación previa, la fase de preparación y la fase de confirmación. Se muestra un diagrama de flujo de las etapas que corresponden a la Figura 2A en la Figura 3A.

Haciendo referencia a la Figura 2A, la Figura 2B, y a la Figura 3A, el protocolo de operación normal comienza en la fase de solicitud cuando un cliente envía una solicitud (mensaje) al nodo primario (Réplica 0) que es responsable para defender la solicitud. La solicitud puede comprender información del cliente, una operación de solicitud (por ejemplo, una o más transacciones para la verificación de consenso), y una indicación de tiempo de solicitud. El cliente (también denominado como un nodo cliente) puede implementarse, por ejemplo, en el nodo A anteriormente descrito. El nodo A puede ser un nodo ligero (por ejemplo, implementado en un teléfono móvil). Adicionalmente o como alternativa, el cliente puede enviar la solicitud a un nodo de respaldo, que reenvía la solicitud al nodo primario antes de la fase de preparación previa. Independientemente de si el nodo primario o de respaldo recibe la solicitud, el correspondiente nodo puede multidifundir la solicitud recibida a los otros nodos en la red. Por lo tanto, el nodo primario puede acabar obteniendo las solicitudes pendientes enviadas mediante los clientes a la red de consenso de una manera u otra (etapa 311).

Por consiguiente, el nodo primario actúa como un líder y conduce a los nodos de respaldo a verificar la transacción/transacciones asociadas con la solicitud. El nodo primario es responsable de ordenar la ejecución de solicitudes dentro de su vista. En la fase de preparación previa, el nodo primario puede obtener una pluralidad de solicitudes, validar las solicitudes obtenidas y proponer un número de secuencia para cada una de las solicitudes. Por lo tanto, a cada una de las solicitudes puede asignarse un número de secuencia creciente y por lo tanto ponerse en orden. Adicionalmente, el mensaje de preparación previa puede comprender una altura de bloque. La altura de bloque puede estar basada en una altura actual de la cadena de bloques. Por ejemplo, si la cadena de bloques tiene actualmente 1000 bloques, la altura de bloque puede ser 1000 que indica que ya existen 1000 bloques en la cadena de bloques, o puede ser 1001 que indica que la transacción/transacciones asociadas con la solicitud se propone que se empaqueten en el bloque de orden 1001 de la cadena de bloques, que ya ha de verificarse mediante otros nodos. El nodo primario puede reenviar una solicitud del cliente junto con el correspondiente número de secuencia y/o la altura de bloque. Por ejemplo, después de obtener las solicitudes, el nodo primario puede disponer las solicitudes en un orden para ejecutar las correspondientes transacciones asignando los números de secuencia y almacenándolos en una lista. El nodo primario puede enviar un mensaje de preparación previa a cada nodo de respaldo (réplica 1 a réplica 3) en el sistema de cadena de bloques (etapa 312). Como se muestra en la Figura 2A, el nodo primario puede multidifundir la lista en o junto con el mensaje de preparación previa a los nodos de respaldo. Como se muestra en la Figura 2B, incluso si un nodo de respaldo (Replica 3) no está funcionando y el nodo primario no tiene conocimiento de ello, el nodo primario puede aún enviar el mensaje de preparación previa (etapa 313). Cada nodo de respaldo acepta el mensaje de preparación previa siempre que sea válido. El mensaje de preparación previa puede contener un número

de vista, un número de secuencia, una firma mediante el nodo primario, un resumen (d), otros metadatos y similares, que permiten la determinación de la validez del mensaje de preparación previa.

En la fase de preparación, si un nodo de respaldo acepta el mensaje de preparación previa, puede realizar un seguimiento mediante multidifusión de un mensaje de preparación a otros nodos en el sistema de cadena de bloques que incluyen el nodo primario (etapa 314). La multidifusión del mensaje de preparación indica que el nodo emisor está de acuerdo con el mensaje de preparación previa. Cada mensaje de preparación se acepta por el nodo de recepción siempre que sea válido. La validez del mensaje de preparación puede determinarse de manera similar basándose en el número de vista, el número de secuencia, la firma del correspondiente nodo de respaldo, un resumen (d), otros metadatos y similares. Se prepara un nodo de respaldo, si ha recibido un mensaje de preparación previa válido desde el nodo primario, y ha obtenido (Q-1) o más mensajes de preparación distintos, válidos y consistentes desde otros nodos (etapa 315), en donde quórum (Q) designa el número de réplicas/nodos requeridos para asegurar todos los requisitos de consistencia de datos y tolerancia a fallos de la réplica/nodo. En algunas realizaciones, el sistema de cadena de bloques que implementa el sistema de PBFT tiene un número de al menos 3F+1 réplicas/nodos, en donde F designa el número de nodos de fallos/que no funcionan bizantinos que el PBFT puede tolerar para funcionar con seguridad y ejecución, y quórum (Q) equivale a 2F+1. En este caso, puede almacenarse un mensaje de preparación previa y al menos 2F mensajes. Los 2F mensajes de preparación pueden incluir el mensaje de preparación de multidifusión. En este punto, son necesarios Q-1 (en este caso, 2F) en lugar de Q mensajes de preparación puesto que el mensaje de preparación previa puede tratarse como un equivalente de un mensaje de preparación del nodo primario (aunque el nodo primario puede no enviar el mensaje de preparación propiamente dicho). Si se cuenta el mensaje de preparación previa como un mensaje de preparación más, entonces habría al menos Q (por ejemplo, 2F+1) mensajes de preparación distintos y válidos que indican que al menos Q (por ejemplo, 2F+1) de todos los nodos aceptaron el mensaje de preparación previa, de los cuales pueden tolerarse hasta F nodos que no funcionan. Por lo tanto, la fase de preparación previa a la de preparación asegura que al menos F+1 nodos que no funcionan (2F+1 nodos preparados previamente pero que contabilizan hasta F nodos que no funcionan) están de acuerdo que si se ejecuta una solicitud en la vista v, se ejecutarán con su número de secuencia.

La fase de preparación asegura la ordenación consistente tolerante a fallos de cada solicitud dentro de las vistas.

En algunas realizaciones, después de recibir el mensaje de preparación previa y los (Q-1) mensajes de preparación, el nodo de respaldo puede verificar el orden y comparar el resultado de verificación con un resultado de verificación propuesto escrito por el nodo primario en el mensaje de preparación previa. Puede haber un número de maneras para verificar el orden. Por ejemplo, el resultado de verificación propuesto puede comprender una raíz de Merkle Patricia Trie escrita en el resumen (d). El nodo de respaldo puede disponer las transacciones asociadas con las solicitudes de acuerdo con el orden y calcular una raíz de Merkle Patricia Trie para comparar con la raíz de Merkle Patricia Trie propuesta. El cálculo puede requerir también cierta información existente tal como la función de troceo de nodo de los bloques existentes en la cadena de bloques. La comparación produce un resumen (D(m)) calculado mediante el nodo de respaldo. Si el resumen (D(m)) es consistente con el resumen (d), la verificación es satisfactoria. Una vez verificado, el nodo de respaldo puede estar de acuerdo con la ordenación de las solicitudes (por ejemplo, el orden para empaquetar las transacciones asociadas con las solicitudes en un nuevo bloque de la cadena de bloques). De manera similar, el nodo de respaldo puede verificar si los mensajes de confirmación (descritos a continuación con respecto a la fase de confirmación) que recibe comprenden el mismo resumen D(m) para determinar si otros nodos también están de acuerdo con la ordenación de las solicitudes. Si un nodo preparado ha obtenido Q (por ejemplo, 2F+1) mensajes de confirmación y se han ejecutado todas las solicitudes con números de secuencia inferiores, el nodo puede ejecutar la solicitud.

En algunas realizaciones, el mensaje de preparación previa puede comprender un resumen (d) del nuevo bloque o información de otra manera relacionada con la ejecución de las solicitudes (por ejemplo, las transacciones asociadas con las solicitudes). El resumen (d) (por ejemplo, un valor de función de troceo) puede ser el resultado numérico de la aplicación de un algoritmo de función de troceo a los datos tales como las transacciones. El nodo de respaldo puede ejecutar las transacciones para confirmar el resumen (d). Para una pluralidad de solicitudes, el nodo de respaldo puede ejecutar las solicitudes de acuerdo con el orden (es decir, el número de secuencias de las solicitudes) para obtener un resumen D(m). Si D(m) y d son consistentes, el nodo de respaldo multidifunde un mensaje de confirmación (descrito a continuación con respecto a la fase de confirmación) que indica que el nodo de respaldo está de acuerdo con el resultado de validación del nodo primario. Para una solicitud pendiente de un cierto número de secuencia, si un nodo preparado previamente ha obtenido Q (por ejemplo, 2F+1) mensajes de confirmación y se han ejecutado todas las solicitudes con números de secuencia inferiores, el nodo puede ejecutar la solicitud.

En la fase de confirmación, si se prepara un nodo, puede multidifundir un mensaje de confirmación a otros nodos (etapa 316). El nodo puede recibir mensajes de confirmación de otros nodos. Cada nodo acepta el mensaje de confirmación siempre que sea válido. El mensaje de confirmación puede contener un número de vista, un número de secuencia, una firma, un resumen, otros metadatos y similares, que permiten la determinación de la validez del mensaje. En algunas realizaciones, si un nodo ha obtenido al menos Q mensajes de confirmación distintos, válidos y consistentes, indica que se ha confirmado un quórum de nodos (es decir, se han preparado al menos (Q-F) nodos honestos) y se ha alcanzado el consenso (etapa 317). Los al menos Q mensajes de confirmación válidos pueden incluir el mensaje de confirmación de multidifusión. Por lo tanto, la fase de preparación para confirmación asegura que

al menos (Q-F) nodos que funcionan están de acuerdo (Q mensajes de confirmación pero que contabilizan hasta F nodos que no funcionan) con que se ejecutará eventualmente una solicitud en la vista v con su número de secuencia. Puesto que los nodos pueden confirmar en diferentes vistas (por ejemplo, cuando algunos nodos ya han entrado en una nueva vista y algunos otros nodos permanecen en la vista anterior), los mensajes de confirmación recibidos pueden corresponder a confirmaciones realizadas en diferentes vistas. La fase de confirmación asegura la ordenación consistente tolerante a fallos de cada solicitud a través de las vistas ya que los nodos que funcionan están de acuerdo con el número de secuencia de cada solicitud.

En algunas realizaciones, si un nodo ha obtenido al menos Q mensajes de confirmación distintos, válidos y consistentes, el nodo puede ejecutar la o las correspondientes solicitudes. Por ejemplo, una vez que se han obtenido Q mensajes de confirmación, significa que el nuevo bloque está verificado en consenso. Por lo tanto, el nodo puede empaquetar el nuevo bloque en la copia de la cadena de bloques mantenida localmente. De otra manera, el nodo de respaldo puede activar directamente el protocolo de cambio de vista.

En la fase de respuesta, después de la ejecución de la solicitud o solicitudes, el nodo envía un mensaje de respuesta directamente al cliente. Para una transacción empaquetada en la cadena de bloques, el mensaje de respuesta puede comprender una dirección de la transacción en la cadena de bloques. Puesto que están permitidos hasta F fallos, el cliente espera las (Q-F) respuestas con firmas válidas desde diferentes nodos y con la misma indicación de tiempo de solicitud y el mismo resultado de ejecución antes de aceptar el resultado. Para el sistema de red de PBFT mostrado en la Figura 2A y en la Figura 2B, hay cuatro nodos totales, por lo que puede tolerase como máximo un (F=1) nodo que no funcione. Por lo tanto, incluso no funcionando la réplica 3, puede alcanzarse aún el consenso en la Figura 2B.

Para conservar la ejecución, puede sustituirse el nodo primario en un protocolo de cambio de vista si ha pasado una cantidad de tiempo específica sin que el nodo primario multidifunda la solicitud. Por ejemplo, el nodo de respaldo puede mantener un temporizador. El nodo de respaldo inicia el temporizador cuando recibe una solicitud y el temporizador ya no está ejecutándose. El nodo de respaldo detiene el temporizador cuando ya no está esperando ejecutar la solicitud (es decir, se ejecuta la solicitud), pero reinicia el temporizador si está esperando ejecutar una o más otras solicitudes. Si se agota el temporizador, el nodo de respaldo puede determinar que el nodo primario no está funcionando. Por lo tanto, el nodo de respaldo puede multidifundir un mensaje de cambio de vista a otros nodos. Para otro ejemplo, el nodo de respaldo puede determinar que el nodo primario es malicioso. Por lo tanto, el nodo de respaldo puede multidifundir un mensaje de cambio de vista. Para otro ejemplo, el cliente puede usar un temporizador para determinar si ha pasado demasiado tiempo después de que un cliente envía la solicitud al nodo primario sin recibir una respuesta. Cuando se agota este temporizador, el cliente envía su solicitud a todos los nodos. Si un nodo ya tiene conocimiento acerca de la solicitud, se ignora la re-difusión. Si el nodo no tiene conocimiento acerca de la solicitud, iniciará un temporizador. En el tiempo de espera del temporizador del nodo, el nodo inicia el proceso de cambio de vista multidifundiendo el mensaje de cambio de vista a otros nodos de respaldo basándose en la sospecha de que el nodo primario es defectuoso (etapa 321). El mensaje de cambio de vista incluye el estado del sistema (en forma de mensajes archivados que incluyen el mensaje de sí mismo durante la operación normal anterior), de modo que otros nodos tendrán conocimiento de que el nodo emisor no ha fallado.

Una súper-mayoría de nodos que funcionan pueden decidir si un nodo primario no está funcionando y retirarlo con el siguiente nodo primario en línea como el sustituyente. El cambio de vista tiene lugar cuando suficientes nodos creen que el nodo primario ha fallado. Una porción de la Figura 2C muestra el protocolo de cambio de vista, y un diagrama de flujo de las etapas que corresponden al protocolo de cambio de vista se muestra en la Figura 3B. Haciendo referencia a la Figura 2C y a la Figura 3B, bajo la fase de cambio de vista, si la vista actual es v, el nodo p = (v+1) mod N espera la obtención de Q mensajes de cambio de vista válidos para que se vuelva el nuevo nodo primario, donde p es el número de serie de réplica/nodo, v es el número de vista, N es el número total de réplicas/nodos (etapa 322). Los Q mensajes de cambio de vista Q pueden incluir el mensaje de cambio de vista de multidifusión. Puesto que la vista anterior es v, los mensajes de cambio de vista puede cada uno comprender una nueva vista *v+1*. Una vez que el nodo primario p ha obtenido Q mensajes de cambio de vista, multidifunde un nuevo mensaje de vista (etapa 323). Este mensaje contiene todos los mensajes de cambio de vista válidos recibidos así como un conjunto de todas las solicitudes que pueden no haberse completado aún debido al fallo de nodo primario. El nuevo nodo primario puede decidir sobre el punto de comprobación último y asegurar, entre otras cosas, que los nodos no defectuosos estén al día con los últimos estados, que puede implicar solicitudes anteriores de re-confirmación (por ejemplo, solicitudes preparadas, confirmadas pero no ejecutadas) en la nueva vista. Aunque está teniendo lugar el cambio de vista, no se aceptan nuevas solicitudes. Después de que un nodo recibe un nuevo mensaje de vista válido que incluye los Q mensajes de cambio de vista, entra en la vista *v*+1 y procesa el conjunto de solicitudes no completadas. Posteriormente, el protocolo de operación normal continúa, y los nodos rehacen las solitudes entre el número de secuencia del último punto de comprobación estable y el número más alto en un mensaje de preparación, pero evitan la re-ejecución de solicitudes. Los nodos de respaldo pueden establecer un temporizador para el nuevo nodo primario (etapa 324).

La Figura 3C es similar a la Figura 3B, excepto una adición de una fase de almacenamiento. Es decir, las etapas 331-337 son similares a las etapas 311-317 respectivamente, excepto que la etapa 399 se realiza adicionalmente entre las etapas 335 y 336. En algunas realizaciones, como se muestra en la Figura 3C, entre la fase de preparación (el nodo de respaldo o primario obtiene los (Q-1) mensajes de preparación) y la fase de confirmación (el nodo de respaldo o

primario multidifunde el mensaje de confirmación), el mensaje de preparación previa y pueden almacenarse al menos (Q-1) mensajes de preparación en la fase de almacenamiento. Se describen detalles adicionales a continuación con referencia a la Figura 4A a la Figura 6B.

La Figura 4A ilustra un diagrama de flujo de etapas de consenso 410a realizadas mediante un nodo primario, de acuerdo con diversas realizaciones de esta memoria descriptiva. La Figura 4B ilustra un diagrama de flujo de etapas de consenso 410b realizadas mediante un nodo de respaldo, de acuerdo con diversas realizaciones de esta memoria descriptiva. Las dos figuras muestran un sistema de cadena de bloques que implementa el mecanismo de consenso de PBFT donde están incluidos al menos 3F+1 nodos. Sin embargo, la presente memoria descriptiva no está limitada a esto. El sistema de cadena de bloques puede tener otro número de nodos distinto de "al menos 3F+1", siempre que haya un quórum de nodos en el sistema para mantener un proceso de consenso válido y satisfacer los requisitos de seguridad y ejecución. En algunas realizaciones, las etapas de consenso 410a se realizan mediante un nodo primario en la vista *v* como se muestra en la Figura 4A, y las etapas de consenso 410b se realizan mediante un nodo de respaldo en la vista v como se muestra en la Figura 4B, sin activar un cambio de vista. La vista indica cuál de los N nodos se considera como el nodo primario, donde N designa el número de los nodos en el sistema de red. Las etapas 410a y 410b puede implementarse cada una mediante uno o más componentes del sistema 100 de la Figura 1 (por ejemplo, Nodo 0, Nodo 1, Nodo 2, ... , o Nodo i anteriormente descritos o un dispositivo similar, o una combinación de cualquiera de los nodos y un dispositivo adicional (por ejemplo, nodo A)). En esta figura, el Nodo A (por ejemplo, un nodo ligero anteriormente descrito) es el cliente, y del Nodo 0 al Nodo 3 son nodos en el sistema de red 112. En la vista actual v, el Nodo 0 actúa como el nodo primario y de los Nodos 1 a 3 actúan como nodos de respaldo. Las etapas 410a y 410b pueden implementarse cada una mediante un sistema o dispositivo de consenso (por ejemplo, ordenador, servidor) que comprende diversa máquina de hardware y/o software. Por ejemplo, el sistema o dispositivo de consenso puede comprender uno o más procesadores y uno o más medios de almacenamiento legibles por ordenador no transitorios (por ejemplo, una o más memorias) acoplados al uno o más procesadores, y configurados con instrucciones ejecutables por el uno o más procesadores para hacer que el sistema o dispositivo (por ejemplo, el procesador) realice las etapas 410a o 410b. Se pretende que las operaciones que se presentan a continuación sean ilustrativas. Dependiendo de la implementación, las operaciones pueden incluir etapas adicionales, menos o alternativas realizadas en diversos órdenes o en paralelo.

En la dirección vertical de las Figuras 4A y 4B, las diversas etapas corresponden a las fases de "Solicitud", "Preparación previa", "Preparación", "Almacenamiento", "Confirmación" y "Respuesta", que puede hacerse referencia a las descripciones anteriores con referencia a la Figura 1 a la Figura 3C. Se muestra la disposición de diversas fases por claridad, y puede no tener requisitos de secuencia estrictos. Por ejemplo, la fase de almacenamiento puede comenzar antes de que finalice la fase de preparación y/o finalizar después de que empiece la fase de confirmación. Como se muestra en la Figura 4A, por ejemplo, la etapa opcional 498 puede realizarse adicionalmente entre la etapa 415 y la etapa 417, cuando tiene lugar una interrupción (por ejemplo, situación de tiempo de inactividad), como se describe a continuación. El nodo primario y los nodos de respaldo pueden ser aquellos definidos en el mecanismo de consenso de PBFT.

Las etapas 410a de la Figura 4A y las etapas 410b de la Figura 4B pueden aplicarse a una ronda de verificación de consenso de una o más solicitudes. Por ejemplo, una ronda de verificación de consenso puede procesar una o más solicitudes de transacción. Si es satisfactoria, las correspondientes transacciones se empaquetan en un nuevo bloque de la cadena de bloques. La descripción a continuación hace referencia a cualquiera de la Figura 4A o la Figura 4B ya que ciertas etapas están entrelazadas, a menos que se indique específicamente. Las etapas 411a y 412a se hallan en la Figura 4A únicamente, mientras que las etapas 411b y 412b se hallan en la Figura 4B únicamente. Las etapas 413, 414, 415, 416, y 417 se muestran tanto en la Figura 4A como en la Figura 4B.

En la etapa 411a, como se muestra en la Figura 4A, en la fase de solicitud, el nodo primario puede obtener una solicitud de un cliente. Por ejemplo, la solicitud puede obtenerse directamente mediante el nodo primario (Nodo 0) del cliente (Nodo A) o de un nodo de respaldo (por ejemplo, Nodo de respaldo 1, 2, o 3) que reenvió la solicitud al nodo primario, como se muestra mediante las líneas discontinuas. En algunas realizaciones, la solicitud puede implicar una transacción/transacciones (con o sin un contrato inteligente) para verificación de consenso. La verificación de consenso puede realizarse durante la ejecución del protocolo de operación normal. Como alternativa, las solicitudes pueden corresponder a otras operaciones.

En la etapa 412a, en la fase de preparación previa, el nodo primario (Nodo 0) multidifunde un mensaje de preparación previa junto con la solicitud a los nodos de respaldo (Nodos 1, 2 y 3). En algunas realizaciones, después de obtener múltiples solicitudes, el nodo primario puede multidifundir el mensaje de preparación previa y las múltiples solicitudes a cada uno de los nodos de respaldo. El mensaje de preparación previa puede incluir un orden para las solicitudes (por ejemplo, un orden para transacciones asociadas con las solicitudes).

Como se muestra en la Figura 4B, que ilustra las etapas realizadas mediante un nodo de respaldo (por ejemplo, Nodo 1, 2 o 3) bajo el protocolo de operación normal, el nodo de respaldo obtiene un mensaje de preparación previa junto con una solicitud en la fase de preparación previa en la etapa 411b. La solicitud puede incluir transacciones asociadas para la verificación de consenso. En algunas realizaciones, el mensaje de preparación previa y la solicitud pueden obtenerse desde el nodo primario. En algunas realizaciones, el mensaje de preparación previa puede obtenerse desde

el nodo primario, y la solicitud puede obtenerse desde el cliente, el nodo primario, y/o cualquier otro nodo de respaldo. Si el nodo primario no está funcionando, puede activarse el protocolo de cambio de vista.

En la etapa 412b, en la fase de preparación, el nodo de respaldo multidifunde un mensaje de preparación a otros nodos en el sistema, si el mensaje de preparación previa es válido.

En la etapa 413, en la fase de preparación, el nodo primario o el nodo de respaldo reciben mensajes de preparación enviados desde otros nodos. La obtención de (Q-1) mensajes de preparación válidos puede ser una condición a cumplirse antes de que el proceso de consenso entre en la siguiente fase de confirmación. En las realizaciones mostradas en las Figuras 4A y 4B, por ejemplo, (Q-1) es 2F, y se requieren 2F o más mensajes de preparación. Los 2F o más mensajes de preparación pueden incluir el propio mensaje de preparación del nodo de respaldo o primario. Para un nodo de respaldo, los 2F o más mensajes de preparación pueden incluir el mensaje de preparación en la etapa 412b (es decir, el mensaje de preparación multidifundido mediante el mismo nodo de respaldo en la etapa 412b).

En la etapa 414, el nodo primario o de respaldo puede almacenar el mensaje de preparación previa y al menos (Q-1) mensajes de preparación. Por ejemplo, si se obtiene un número de 3F mensajes de preparación mediante un nodo, puede almacenarse mediante el nodo el mensaje de preparación previa y un número de mensajes de preparación entre 2F y 3F inclusive. En algunas realizaciones, únicamente se almacena el mensaje de preparación previa y Q-1 mensajes de preparación. En algunas realizaciones, únicamente se almacena el mensaje de preparación previa y los 2F mensajes de preparación. Por ejemplo, si se obtienen 3F mensajes de preparación, el mensaje de preparación previa y 2F mensajes de preparación pueden ser la cantidad mínima de mensajes de consenso que necesitan almacenarse para que un proceso de consenso válido se reanude y avance de manera eficaz y eficiente después de que la totalidad del sistema se recupere de una interrupción (por ejemplo, una avería de sistema), sin consumir demasiados recursos de almacenamiento de sistema. En algunas realizaciones, almacenar el mensaje de preparación previa y el (Q-1) o más mensajes de preparación comprende: almacenar únicamente el mensaje de preparación previa y el (Q-1) o más mensajes de preparación, que significa que no se almacena ningún otro mensaje distinto del mensaje de preparación previa y los al menos 2F mensajes de preparación. Por ejemplo, para cada ronda de consenso-verificación, no se almacenan mensajes de confirmación. Lo mismo puede aplicarse cuando se realizan múltiples rondas de verificaciones de consenso.

La etapa 413 y la etapa 414 pueden realizarse en secuencia, simultáneamente o de cualquier otra manera. En algunas realizaciones, el almacenamiento del mensaje de preparación previa y los al menos (Q-1) mensajes de preparación puede realizarse únicamente cuando se obtienen (Q-1) o más mensajes de preparación. En otras realizaciones, el almacenamiento del mensaje de preparación previa y los al menos (Q-1) mensajes de preparación puede realizarse en cualquier momento después de que se obtienen el respectivo mensaje.

En algunas realizaciones, el mensaje de preparación previa y los al menos (Q-1) mensajes de preparación pueden almacenarse de diversas maneras siempre que los mensajes almacenados sean recuperables después de que el sistema se recupere de la interrupción. Por ejemplo, el mensaje de preparación previa y los al menos (Q-1) mensajes de preparación pueden almacenarse en un almacenamiento persistente que asegura que el almacenamiento no se vea afectado por averías y reinicios de sistema.

En algunas realizaciones, si no hay interrupción (por ejemplo, tiempo de inactividad provocado por una avería de sistema) para la operación de sistema, puede realizarse la etapa 415. En una realización, la fase de confirmación en la etapa 415 se realiza después de que se almacenan al menos el mensaje de preparación previa y los (Q-1) mensajes de preparación. Para la etapa 415, en la fase de confirmación, el nodo primario y los nodos de respaldo cada uno multidifunde un mensaje de confirmación a otros nodos. Cada nodo puede recibir también mensajes de confirmación multidifundidos mediante otros nodos. En la etapa 416, el nodo primario o de respaldo puede obtener al menos un quórum (Q) de mensajes de confirmación (en este caso, 2F+1). Para el nodo de respaldo o primario, como se muestra en las Figuras 4A y 4B, los Q mensajes de confirmación pueden incluir el mensaje de confirmación en la etapa 415 (es decir, el mensaje de confirmación multidifundido por el mismo nodo de respaldo o primario en la etapa 415). En la etapa 417, si un nodo observa que se han confirmado suficientes nodos (por ejemplo, Q nodos), el nodo puede ejecutar la solicitud de acuerdo con el orden y notificar al cliente (Nodo A) mediante un mensaje de respuesta.

En algunas realizaciones, si hay una interrupción (por ejemplo, tiempo de inactividad provocado por una avería de sistema) a la operación del sistema después de que se multidifunde un mensaje de confirmación, puede realizarse una etapa opcional 498 después de la etapa 415 y antes de la etapa 417. En la etapa 498, el nodo primario o de respaldo puede realizar un reinicio de sistema, y cargar el mensaje de preparación previa y los al menos (Q-1) mensajes de preparación que el nodo almacenó una vez en la etapa 414. En algunas realizaciones, el sistema puede reiniciarse después de la interrupción, ya sea voluntaria o involuntariamente. A continuación, pueden seguir el resto de las etapas 416 a 417 o la etapa 417.

En algunas realizaciones, si hay una interrupción (por ejemplo, tiempo de inactividad provocado por una avería de sistema) a la operación del sistema antes de que se multidifunda un mensaje de confirmación, puede realizarse una etapa opcional 499 después de la etapa 414 y antes de la etapa 415. En la etapa 499, el nodo primario o de respaldo puede cargar el mensaje de preparación previa y los al menos (Q-1) mensajes de preparación una vez almacenados

en la fase de almacenamiento (etapa 414). En algunas realizaciones, el sistema puede reiniciarse después de la interrupción, ya sea voluntaria o involuntariamente. Si se realiza la etapa 499, puede activarse el protocolo de cambio de vista bajo ciertas circunstancias (por ejemplo, si el nodo que no está funcionando es el nodo primario, y el nodo primario no reanuda su estado de funcionamiento dentro del periodo de tiempo de espera). Sin embargo, si no se cumple la condición de tiempo de espera (por ejemplo, se completa la etapa 499 antes de activar la condición de tiempo de espera), puede no activarse el protocolo de cambio de vista, como se muestra en la Figura 4A y en la Figura 4B. Por lo tanto, puede no activarse el protocolo de cambio de vista si el nodo primario que no funciona reanuda su estado de funcionamiento lo suficientemente rápido para evitar la condición de tiempo de espera, y pueden seguir las etapas 415 a 417 en el protocolo. Si se cumple la condición de tiempo de espera (por ejemplo, la etapa 499 no se completa antes de que se active la condición de tiempo de espera), puede activarse el protocolo de cambio de vista como se describe a continuación con referencia a la Figura 4C y a la Figura 4D.

La Figura 4C ilustra un diagrama de flujo de etapas de consenso 420a mediante un nodo de respaldo en la vista v que se vuelve un nuevo nodo primario en la vista *v*+1, de acuerdo con diversas realizaciones de esta memoria descriptiva. La Figura 4D ilustra un diagrama de flujo de etapas de consenso 420b mediante un nodo de respaldo en la vista v que permanece como un nodo de respaldo en la vista *v*+1, de acuerdo con diversas realizaciones de esta memoria descriptiva. Las etapas 420a y 420b pueden implementarse mediante uno o más componentes del sistema 100 de la Figura 1 (por ejemplo, Nodo 0, Nodo 1, Nodo 2, ... , o Nodo i anteriormente descritos o un dispositivo similar, o una combinación de cualquiera de los nodos y un dispositivo adicional (por ejemplo, Nodo A)). En esta figura, el Nodo A (por ejemplo, un nodo ligero anteriormente descrito) es un cliente, y del Nodo 0 al Nodo 3 son nodos de cadena de bloques. Como se describe en la Figura 4A y en la Figura 4B, en la vista v, el Nodo 0 actuó como el nodo primario, pero para la vista *v*+1 en la Figura 4C y en la Figura 4D, el Nodo 1 se vuelve un nuevo nodo primario, permaneciendo los Nodos 2 a 3 como nodos de respaldo. Las etapas 420a y 420b puede implementarse cada una mediante uno o más nodos del sistema de red distribuido (por ejemplo, un sistema de cadena de bloques). Las etapas 420a y 420b pueden implementarse cada una mediante un sistema o dispositivo de consenso (por ejemplo, ordenador, servidor) que comprende diversa máquina de hardware y/o software. Por ejemplo, el sistema o dispositivo de consenso puede comprender uno o más procesadores y uno o más medios de almacenamiento legibles por ordenador no transitorios (por ejemplo, una o más memorias) acoplados al uno o más procesadores, y configurados con instrucciones ejecutables por el uno o más procesadores para hacer que el sistema o dispositivo (por ejemplo, el procesador) realicen las etapas 420a y 420b. Se pretende que las operaciones que se presentan a continuación sean ilustrativas. Dependiendo de la implementación, las operaciones pueden incluir etapas adicionales, menos o alternativas realizadas en diversos órdenes o en paralelo.

Como se ha descrito anteriormente, si se activa el cambio de vista después de la etapa 414 y antes de la etapa 415 de la Figura 4B, se realizan las etapas mostradas en la Figura 4C y en la Figura 4D. Por brevedad, no se reproducen las etapas antes de la etapa 499 (las etapas hasta la etapa 414 mostrada en la Figura 4B) en la Figura 4C y en la Figura 4D.

En algunas realizaciones, las etapas de consenso 420a y 420b como se muestra en la Figura 4C y en la Figura 4D pueden corresponder a la situación de cambio de vista de activación. El nodo primario en la vista v (por ejemplo, el Nodo 0) puede volverse defectuoso o no funcionar de otra manera. Para la Figura 4C, un nodo de respaldo como en la vista v (por ejemplo, el Nodo 1) que se vuelve un nuevo nodo primario en la vista *v*+1 puede realizar las etapas 499, 423, 424a, 425a, 426a, 425, 426, y 427. Un nodo de respaldo en la vista v (por ejemplo, el Nodo 2 o 3) que permanece como un nodo de respaldo en la vista *v*+1 puede realizar las etapas 499, 423, 424b, 425b, 426b, 425, 426, y 427. En la dirección vertical de las dos figuras, las diversas etapas corresponden a las fases de "cambio de vista", "nueva vista", "preparación", "confirmación" y "respuesta", que puede hacerse referencia a las descripciones anteriormente con referencia a la Figura 1 a la Figura 3C. Se muestra la disposición de diversas fases por claridad y puede no tener requisitos de secuencia estrictos. El nodo primario y los nodos de respaldo pueden ser aquellos definidos en el mecanismo de consenso de PBFT. Las descripciones a continuación hacen referencia a cualquiera de la Figura 4C o la Figura 4D ya que ciertas etapas están entrelazadas.

En algunas realizaciones, como se muestra en la etapa 499, aún en la vista v, el nodo primario (Nodo 0) y algunos de los nodos de respaldo (Nodo 1, 2, y/o 3) puede cada uno cargar el mensaje de preparación previa y los al menos (Q-1) mensajes de preparación que se almacenan respectivamente en la etapa 414. Si los mensajes se almacenaron desde un almacenamiento persistente, pueden ahora cargarse desde el almacenamiento persistente. El reinicio de sistema puede realizarse en respuesta a una interrupción a la operación normal (por ejemplo, tiempo de inactividad provocado por una avería de sistema).

En una realización, bajo la sospecha de que el nodo primario puede no estar funcionando, un nodo de respaldo (por ejemplo, el Nodo 1, 2 o 3) puede multidifundir un mensaje de cambio de vista, que puede comprender el mensaje de preparación previa cargado y los al menos (Q-1) mensajes de preparación cargados, como se muestra en la etapa 423. Bajo el protocolo de cambio de vista, uno de los nodos de respaldo puede volverse un nuevo nodo primario, y el resto pueden permanecer como nodos de respaldo. La selección del nuevo nodo primario se ha descrito anteriormente. Por ejemplo, como se muestra, puede elegirse que el Nodo 1 sea el nuevo nodo primario, mientras que los Nodos 2 y 3 pueden permanecer como nodos de respaldo.

En la etapa 424a, cuando un nodo de respaldo ha obtenido al menos Q mensajes de cambio de vista de otros nodos indicando cada uno que el correspondiente nodo está de acuerdo con el mensaje de cambio de vista, puede seleccionarse un nuevo nodo primario (por ejemplo, Nodo 1). Los al menos Q mensajes de cambio de vista pueden incluir el mensaje de cambio de vista multidifundido mediante el nodo de respaldo propiamente dicho. En la etapa 425a, el nuevo nodo primario (por ejemplo, Nodo 1) multidifunde a al menos alguno de los nodos de respaldo un nuevo mensaje de vista que comprende los al menos Q mensajes de cambio de vista.

Como se muestra en la Figura 4D, en la etapa 424b, en el proceso bajo el protocolo de cambio de vista, un nodo de respaldo puede obtener desde el nuevo nodo primario un nuevo mensaje de vista que indica que el nuevo nodo primario ha recibido Q o más mensajes de cambio de vista indicando cada uno que el correspondiente nodo está de acuerdo con el mensaje de cambio de vista. En la etapa 425b, el nodo de respaldo multidifunde otro mensaje de preparación que indica una aceptación del nuevo mensaje de vista. El otro mensaje de preparación puede diferir del mensaje de preparación de la Figura 4A y la Figura 4B al menos en términos del número de vista.

Haciendo referencia a la Figura 4C, en la etapa 426a, el nuevo nodo primario (Nodo 1) puede obtener otros (Q-1) o más mensajes de preparación. En la etapa 426b, los nodos de respaldo restantes pueden cada uno obtener otros (Q-1) o más mensajes de preparación. La fase de preparación de la Figura 4C y la Figura 4D es similar a la fase de preparación de la Figura 4A y la Figura 4B, excepto que los contenidos del mensaje de preparación pueden diferir después del cambio de vista y algunos nodos pueden haber confirmado algunas de las solicitudes. Para distinguir, el mensaje de preparación para la fase de preparación de la Figura 4C y la Figura 4D se denomina como otro mensaje de preparación u otra cantidad de mensajes de preparación.

Las etapas 425 a 427 bajo el protocolo de cambio de vista son similares a las etapas 415 a 417 bajo el protocolo de operación normal, pero pueden diferir en los siguientes aspectos: (1) el número de vista, (2) solicitudes confirmadas no necesitan volver a confirmarse en el correspondiente nodo, (3) que el Nodo 0 que no funciona puede no realizar las etapas 425 a 427, o no realizar de manera honesta las etapas 425 a 427.

Los métodos desvelados pueden asegurar las funciones apropiadas del sistema de cadena de bloques con menos demanda del consumo de almacenamiento. En un ejemplo, en un sistema de cadena de bloques con un número total de al menos 3F+1 nodos, cuando los al menos F+1 nodos han multidifundido los mensajes de confirmación, significa que se han preparado al menos 2F+1 nodos, y el mensaje de preparación previa y los al menos 2F mensajes de preparación se encuentran en persistencia. En algunas realizaciones, el mensaje de preparación previa y los al menos 2F mensajes de preparación se almacenan mediante los respectivos nodos en la fase de almacenamiento. Por ejemplo, el nodo primario y/o algunos nodos de respaldo han almacenado el mensaje de preparación previa y los mensajes de preparación. Como tal, incluso si uno o más o, en el peor de los casos, todos los nodos experimentan una avería y reinicio de sistema, a diferencia del proceso sin la etapa de almacenamiento, se almacena el mensaje de preparación previa y los al menos 2F mensajes una vez almacenados en la fase de almacenamiento. Como resultado, incluso si hay F nodos (que pueden haber multidifundido o no los mensajes de confirmación) que no reinician y reanudan la funcionalidad, puesto que se almacena y carga el mensaje de preparación previa y los al menos 2F mensajes, la totalidad del proceso de verificación de consenso puede reanudarse y avanzarse de manera eficaz con menos demanda de consumo de almacenamiento, y sin verse afectado por la avería de sistema que puede provocar de otra manera inconsistencia y/o ramificación o afectar a la seguridad y/o ejecución del sistema.

En algunas realizaciones, si el nodo primario no está entre los nodos que se reiniciaron, puede activarse el cambio de vista si finaliza periodo de tiempo de espera. Puesto que se han preparado al menos Q nodos e incluso si se han confirmado F de ellos y no realizan el reinicio, (Q-F) nodos pueden realizar el reinicio de sistema y cargar los mensajes de preparación previa y preparación almacenados. El mensaje de cambio de vista multidifundido mediante los nodos reiniciados (Q-F) llevaría los mensajes de preparación previa y preparación desde antes de la avería, lo que asegura que el nuevo mensaje de vista multidifundido por el nuevo nodo primario llevará los mismos. Por lo tanto, se evitan resultados de consenso inconsistentes y ramificación de cadena de bloques.

En otras realizaciones, si el nodo primario está entre los Q nodos que reiniciaron, el nodo primario puede intentar reanudar el protocolo de operación normal o proponer otras operaciones. Si el reinicio no es suficientemente rápido, puesto que al menos (Q-F) nodos están bloqueados mediante los mensajes de preparación previa y preparación, no responderán al nodo primario. Por consiguiente, no puede alcanzarse consenso, y puede activarse el cambio de vista para elegir un nuevo nodo primario. El resto puede seguir las realizaciones de cambio de vista anteriormente descritas.

La Figura 5A ilustra un diagrama de flujo de un método de consenso 510, de acuerdo con diversas realizaciones de esta memoria descriptiva. El método 510 puede implementarse mediante uno o más componentes del sistema 100 de la Figura 1 (por ejemplo, Nodo 0, Nodo 1, Nodo 2, ... , o Nodo i anteriormente descritos o un dispositivo similar, o una combinación de cualquiera de los nodos y uno o más dispositivos adicionales (por ejemplo, el nodo A)). El método 510 puede implementarse mediante uno o más nodos de cadena de bloques (por ejemplo, un nodo de respaldo). El método 510 puede implementarse mediante un sistema o dispositivo de consenso (por ejemplo, ordenador, servidor) que comprende diversa máquina de hardware y/o software. Por ejemplo, el sistema o dispositivo de consenso puede comprender uno o más procesadores y uno o más medios de almacenamiento legibles por ordenador no transitorios (por ejemplo, una o más memorias) acoplados al uno o más procesadores, y configurados con instrucciones

ejecutables por el uno o más procesadores para hacer que el sistema o dispositivo (por ejemplo, el procesador) realicen el método 510. Se pretende que las operaciones del método 510 que se presentan a continuación sean ilustrativas. Dependiendo de la implementación, el método 510 puede incluir etapas adicionales, menos o alternativas realizadas en diversos órdenes o en paralelo. Los diversos bloques descritos a continuación pueden no tener que ejecutarse en la secuencia mostrada en la figura a menos que se especifique de otra manera. Por ejemplo, el bloque 512 puede empezar después de que se inicie el bloque 513 y finalizar antes de que se finalice el bloque 513. De manera similar, el bloque 515 puede empezar después de que se inicie el bloque 516 y finalizar antes de que se finalice el bloque 516. Para otro ejemplo, los bloques 513 y 514 pueden realizarse en secuencia o en paralelo.

En algunas realizaciones, el método 510 puede implementarse en una cadena de bloques mantenida mediante un número (N) de nodos, en donde uno de los nodos actúa como un nodo primario y los otros (N-1) nodos actúan como nodos de respaldo, y se realiza el método 510 mediante uno de los nodos de respaldo. N puede ser cualquier número entero no menor que cuatro. En algunas realizaciones, N equivale 3F+1, en donde F designa el número de nodos que no funcionan que el sistema puede tolerar en el mecanismo de consenso de PBFT. Los nodos primario y de respaldo pueden ser aquellos definidos en el mecanismo de consenso de PBFT. El método 510 puede aplicarse a una ronda de verificación de consenso para una o más solicitudes (por ejemplo, solicitudes de transacción de cadena de bloques). Las etapas del método 510 pueden realizarse mediante un nodo de respaldo en una vista actual, que puede permanecer como un nodo de respaldo o volverse un nuevo nodo primario en caso de que tenga lugar un cambio de vista. La vista de acuerdo con el mecanismo de consenso de PBFT puede cambiar o no durante la implementación del método 510. Puede hacerse referencia a detalles del método 510 en la Figura 1 a la Figura 4B y en descripciones anteriormente relacionadas.

El bloque 511 incluye la obtención de un mensaje de preparación previa desde el nodo primario. En algunas realizaciones, antes de obtener el mensaje de preparación previa desde el nodo primario, el método 510 comprende adicionalmente obtener una o más solicitudes de transacción desde al menos uno de: uno o más clientes, el nodo primario, o uno o más de los otros nodos de respaldo. La transacción en la expresión "solicitud de transacción" puede implementarse mediante un sistema de cadena de bloques y registrarse en la cadena de bloques. La transacción puede incluir, por ejemplo, una transacción financiera, una transacción de contrato de cadena de bloques para implementar o invocar un contrato de cadena de bloques, una transacción que actualiza un estado (por ejemplo, estado mundial) de la cadena de bloques, etc. La transacción no tiene que implicar un intercambio financiero. Las solicitudes de transacción pueden comprender transacciones de cadena de bloques que van a añadirse a la cadena de bloques mediante verificación de consenso. En una realización, el mensaje de preparación previa comprende un orden de la una o más transacciones que corresponden a la una o más solicitudes de transacción. El orden puede proponerse mediante el nodo primario que multidifunde el mensaje de preparación previa para la ejecución de las solicitudes de transacción. El orden puede corresponder a una identificación de valor de función de troceo único de un nuevo bloque propuesto que contiene las transacciones. El nodo primario y los nodos de respaldo verificarán el orden propuesto e intentarán alcanzar un consenso. Como alternativa, la solicitud puede comprender otra instrucción a uno o más dispositivos informáticos para proporcionar información o realizar otra función.

El bloque 512 incluye multidifundir un mensaje de preparación para al menos alguno del nodo primario y los otros (N-2) nodos de respaldo, indicando el mensaje de preparación una aceptación del mensaje de preparación previa. Multidifundir significa difundir a uno o más de todos los otros nodos en el sistema de PBFT. Cada nodo de respaldo que funciona puede multidifundir el mensaje de preparación.

El bloque 513 incluye obtener (Q-1) o más mensajes de preparación respectivamente desde (Q-1) o más de los nodos de respaldo, en donde Q (quórum) es (N+F+1)/2 redondeado hacia arriba hasta el número entero más cercano, y F es (N-1)/3 redondeado hacia abajo hasta el número entero más cercano. Por ejemplo, el nodo que realiza el método 510 es uno de los N nodos. Los (Q-1) mensajes de preparación pueden ser desde distintos nodos y son válidos y consistentes, que indica que los al menos (Q-1) nodos de respaldo y el nodo primario están de acuerdo con el mensaje de preparación previa.

El bloque 514 incluye almacenar el mensaje de preparación previa y el (Q-1) o más mensajes de preparación. Por ejemplo, si (Q-1) es 2F y si se obtienen 3F mensajes de preparación en la etapa anterior, en este punto, puede almacenarse el mensaje de preparación previa y un número de mensajes de preparación entre 2F y 3F inclusive. En algunas realizaciones, únicamente se almacena el mensaje de preparación previa y los (Q-1) mensajes de preparación. Por ejemplo, si (Q-1) es 2F y si se obtienen 3F mensajes de preparación en la etapa anterior, en este punto, únicamente puede almacenarse el mensaje de preparación previa y 2F mensajes de preparación. En algunas realizaciones, almacenar el mensaje de preparación previa y el (Q-1) o más mensajes de preparación comprende: almacenar únicamente el mensaje de preparación previa y el (Q-1) o más mensajes de preparación. Por ejemplo, únicamente se almacena el mensaje de preparación previa y los (Q-1) mensajes de preparación. No se almacena mensaje distinto del mensaje de preparación previa y el (Q-1) o más mensajes de preparación. Por ejemplo, para cada ronda de consenso-verificación, no se almacenan mensajes de confirmación. Lo mismo puede aplicarse cuando se realizan múltiples rondas de verificaciones de consenso.

En algunas realizaciones, el mensaje de preparación previa y los al menos (Q-1) mensajes de preparación pueden almacenarse de diversas maneras siempre que los datos almacenados sean recuperables después de la recuperación

de tiempo de inactividad de sistema, tal como un reinicio de sistema. Por ejemplo, el mensaje de preparación previa y el (Q-1) o más mensajes de preparación pueden almacenarse en una base de datos relacional, una base de datos no relacional, un sistema de documento, etc. Por ejemplo, el mensaje de preparación previa y el (Q-1) o más mensajes de preparación pueden almacenarse en un almacenamiento persistente. La etapa de almacenamiento y las otras etapas descritas en el presente documento pueden no estar limitadas por el lenguaje de programación.

En algunas realizaciones, el bloque 514 puede realizarse únicamente cuando se satisface el bloque 513, es decir, únicamente cuando se obtienen (Q-1) o más mensajes de preparación. En otras realizaciones, cada mensaje de preparación previa o preparación puede almacenarse tan pronto como se reciba.

En algunas realizaciones, después de almacenar el mensaje de preparación previa y el (Q-1) o más mensajes de preparación (bloque 514) y antes de multidifundir el mensaje de confirmación (bloque 515), el método comprende adicionalmente: realizar un reinicio de sistema; y cargar el mensaje de preparación previa almacenado y el (Q-1) o más mensajes de preparación almacenados. El reinicio de sistema puede realizarse en respuesta a una interrupción a la operación normal (por ejemplo, avería de sistema, corte de energía, etc.). La interrupción puede ocurrir en uno o más o todos los nodos en el sistema de PBFT. En algunas realizaciones, hasta todos los N nodos experimentan una avería, y al menos Q de los N nodos realizan el reinicio de sistema y cargan el correspondiente mensaje previamente preparado almacenado y el (Q-1) o más mensajes de preparación almacenados respectivamente. A continuación, el protocolo de cambio de vista puede activarse o no.

En una realización, el protocolo de cambio de vista puede no activarse si el reinicio es lo suficientemente rápido para evitar la activación del tiempo de espera, y por lo tanto el reinicio de sistema evita activar el cambio de vista. Es decir, la realización del reinicio de sistema comprende: realizar el reinicio de sistema sin activar un cambio de vista. Por consiguiente, puede seguir el resto de las etapas del método 510 desde el bloque 515.

De otra manera, puede activarse el protocolo de cambio de vista. En una realización, después de almacenar el mensaje de preparación previa y el (Q-1) o más mensajes de preparación y antes de multidifundir el mensaje de confirmación, el método comprende adicionalmente: multidifundir un mensaje de cambio de vista que comprende el mensaje de preparación previa cargado y el (Q-1) o más mensajes de preparación cargados. Otros nodos de respaldo pueden multidifundir también el mensaje de cambio de vista. Uno de los nodos de respaldo puede elegirse para que se vuelva un nuevo nodo primario, que puede ser o no el nodo de respaldo que realizó las etapas anteriores.

En algunas realizaciones, si el nodo de respaldo que realizó las etapas anteriores no se elige que sea un nuevo nodo primario, puede permanecer como el nodo de respaldo y realizar las siguientes etapas durante el cambio de vista. Después de almacenar el mensaje de preparación previa y el (Q-1) o más mensajes de preparación y antes de multidifundir el mensaje de confirmación, el método comprende adicionalmente: obtener desde un nuevo nodo primario un nuevo mensaje de vista que indica que el nuevo nodo primario ha recibido Q o más mensajes de cambio de vista indicando cada uno que el correspondiente nodo está de acuerdo con el mensaje de cambio de vista; multidifundir otro mensaje de preparación a al menos alguno de los nodos de respaldo que incluye el nuevo nodo primario, indicando el otro mensaje de preparación una aceptación del nuevo mensaje de vista; y obtener otros (Q-1) o más mensajes de preparación respectivamente desde (Q-1) o más de los nodos de respaldo, en donde los otros (Q-1) o más mensajes de preparación incluyen multidifundir otro mensaje de preparación.

En otras realizaciones, si el nodo que realizó las etapas anteriores se elige que sea el nuevo nodo primario, puede volverse el nuevo nodo primario y realizar las siguientes etapas durante el cambio de vista. Después de almacenar el mensaje de preparación previa y el (Q-1) o más mensajes de preparación y antes de multidifundir el mensaje de confirmación, el método comprende adicionalmente: obtener, respectivamente desde Q o más de los nodos de respaldo, Q o más mensajes de cambio de vista indicando cada uno que el correspondiente nodo está de acuerdo con el mensaje de cambio de vista, en donde el Q o más mensajes de cambio de vista incluyen el mensaje de cambio de vista de multidifusión; multidifundir a al menos alguno de los nodos de respaldo un nuevo mensaje de vista que indica que el nodo de respaldo que actúa como un nuevo nodo primario ha recibido el Q o más mensajes de cambio de vista; y obtener otros (Q-1) o más mensajes de preparación respectivamente desde (Q-1) o más de los nodos de respaldo, en donde los otros (Q-1) o más mensajes de preparación incluyen el otro mensaje de preparación de multidifusión.

Pueden realizarse los bloques 515 y 516 y las siguientes etapas mientras están en la misma vista que los bloques 511-514 si no tiene lugar cambio de vista, o en una nueva vista si el cambio de vista tiene lugar antes del bloque 515.

El bloque 515 incluye multidifundir un mensaje de confirmación a al menos alguno del nodo primario y los otros nodos de respaldo, indicando el mensaje de confirmación que el nodo de respaldo está de acuerdo con el (Q-1) o más mensajes de preparación. En algunas realizaciones, el mensaje de confirmación indica que el nodo de respaldo está de acuerdo con el mensaje de preparación previa y ha obtenido el (Q-1) o más mensajes de preparación. En algunas realizaciones, pueden realizarse etapas de verificación para estar de acuerdo con multidifundir el mensaje de confirmación. Por ejemplo, como se ha descrito anteriormente, puede determinarse el resumen D(m) de acuerdo con el orden para verificación contra el resumen d. Si es consistente, puede multidifundirse el mensaje de confirmación.

En algunas realizaciones, hasta F de los (Q-1) o más de los nodos de respaldo en el bloque 513 son defectuosos o

no funcionan de otra manera después de multidifundir los mensajes de confirmación respectivamente y no realizan el reinicio de sistema. Por ejemplo, los F nodos que se han confirmado pueden experimentar una avería de sistema y no reiniciar para reanudar la función. A pesar de eso, la verificación de consenso puede llevarse a cabo de manera apropiada sin provocar resultados inconsistentes y la ramificación a la cadena de bloques.

El bloque 516 incluye obtener, respectivamente desde Q o más nodos entre el nodo primario y los nodos de respaldo, Q o más mensajes de confirmación indicando cada uno que el correspondiente nodo está de acuerdo con (Q-1) o más mensajes de preparación recibidos mediante el correspondiente nodo. En algunas realizaciones, el mensaje de confirmación indica que el correspondiente nodo que multidifunde el mensaje de confirmación está de acuerdo con el mensaje de preparación previa y ha obtenido (Q-1) o más mensajes de preparación. Los Q mensajes de confirmación pueden ser desde nodos distintos y son válidos y consistentes, que indica que están preparados Q nodos para ejecutar las solicitudes en el orden. Por lo tanto, se alcanza un consenso por una mayoría de los nodos, y puede realizarse la siguiente etapa de ejecución.

En algunas realizaciones, después de multidifundir el mensaje de confirmación (bloque 515) y antes de ejecutar las solicitudes, el método comprende adicionalmente: realizar un reinicio de sistema, y cargar el mensaje de preparación previa almacenado y el (Q-1) o más mensajes de preparación almacenados. El reinicio de sistema puede realizarse de manera voluntaria o involuntaria. El reinicio de sistema puede provocarse por una interrupción al sistema o a la función del dispositivo, tal como una avería de sistema.

En algunas realizaciones, el método 510 puede incluir adicionalmente empaquetar la una o más transacciones en una copia local de la cadena de bloques mantenida mediante el nodo de respaldo de acuerdo con el orden. Por ejemplo, las solicitudes pueden verificarse en consenso como los al menos (Q-F) nodos honestos (Q mensajes de confirmación pero contabilizando para como máximo F nodos que no funcionan) que han verificado el resumen d en sus mensajes de confirmación (o para el nodo primario, puede no tener que realizar la verificación puesto que propuso el resumen d). Como resultado, si suficientes nodos han verificado las correspondientes transacciones, las transacciones pueden empaquetarse en la cadena de bloques. Puede notificarse al cliente o los clientes (por ejemplo, el nodo A) que enviaron originalmente la solicitud o solicitudes.

La Figura 5B ilustra un diagrama de flujo de un método de consenso 520, de acuerdo con diversas realizaciones de esta memoria descriptiva. El método 520 puede implementarse mediante uno o más componentes del sistema 100 de la Figura 1 (por ejemplo, Nodo 0, Nodo 1, Nodo 2, ... , o Nodo i anteriormente descritos o un dispositivo similar, o una combinación de cualquiera de los nodos y uno o más dispositivos adicionales (por ejemplo, el nodo A)). El método 520 puede implementarse mediante uno o más nodos de cadena de bloques (por ejemplo, un nodo primario). El método 520 puede implementarse mediante un sistema o dispositivo de consenso (por ejemplo, ordenador, servidor) que comprende diversa máquina de hardware y/o software. Por ejemplo, el sistema o dispositivo de consenso puede comprender uno o más procesadores y uno o más medios de almacenamiento legibles por ordenador no transitorios (por ejemplo, una o más memorias) acoplados al uno o más procesadores, y configurados con instrucciones ejecutables por el uno o más procesadores para hacer que el sistema o dispositivo (por ejemplo, el procesador) realicen el método 520. Se pretende que las operaciones del método 520 que se presentan a continuación sean ilustrativas. Dependiendo de la implementación, el método 520 puede incluir etapas adicionales, menos o alternativas realizadas en diversos órdenes o en paralelo. Los diversos bloques descritos a continuación pueden no tener que ejecutarse en la secuencia mostrada en la figura a menos que se especifique de otra manera. Por ejemplo, el bloque 521 puede empezar después de que se inicie el bloque 522 y finalizar antes de que se finalice el bloque 522. De manera similar, el bloque 524 puede empezar después de que se inicie el bloque 525 y finalizar antes de que se finalice el bloque 525. Para otro ejemplo, los bloques 522 y 523 pueden realizarse en secuencia o en paralelo.

En algunas realizaciones, el método 520 puede implementarse en una cadena de bloques mantenida mediante un número (N) de nodos, en donde uno de los nodos actúa como un nodo primario y los otros (N-1) nodos actúan como nodos de respaldo, y se realiza el método 520 mediante el nodo primario. Los nodos primario y de respaldo pueden ser aquellos definidos en el modelo de PBFT. El método 520 puede aplicarse a una ronda de verificación de consenso para una o más solicitudes (por ejemplo, solicitudes de transacción de cadena de bloques). Puede hacerse referencia a detalles del método 520 en la Figura 1 a la Figura 4B y en descripciones anteriormente relacionadas.

El bloque 521 incluye multidifundir un mensaje de preparación previa a al menos alguno de los nodos de respaldo. En algunas realizaciones, antes de multidifundir el mensaje de preparación previa a al menos alguno de los nodos de respaldo, el método 520 comprende adicionalmente obtener una o más solicitudes de transacción desde al menos uno de: uno o más clientes (por ejemplo, nodos ligeros) o uno o más de los nodos de respaldo. Las solicitudes de transacción pueden comprender transacciones de cadena de bloques que van a añadirse a la cadena de bloques mediante verificación de consenso. En una realización, el mensaje de preparación previa comprende un orden de la una o más transacciones que corresponden a la una o más solicitudes de transacción. El orden puede proponerse mediante el nodo primario que multidifunde el mensaje de preparación previa para la ejecución de las solicitudes de transacción. El orden puede corresponder a una identificación de valor de función de troceo único de un nuevo bloque propuesto que contiene las transacciones. El nodo primario y los nodos de respaldo verificarán el orden propuesto e intentarán alcanzar un consenso. Como alternativa, la solicitud puede comprender otra instrucción a uno o más dispositivos informáticos para proporcionar información o realizar otra función.

Los bloques 522 a 525 pueden ser similares a los bloques 513 a 516 y las descripciones anteriores relacionadas, excepto que si el nodo primario se vuelve como que no funciona, se activa el cambio de vista y se elige un nuevo nodo primario.

El bloque 522 incluye obtener (Q-1) o más mensajes de preparación respectivamente desde (Q-1) o más de los nodos de respaldo, en donde los mensajes de preparación indica cada uno una aceptación del mensaje de preparación previa mediante el correspondiente nodo de respaldo, Q (quórum) es (N+F+1)/2 redondeado hacia arriba hasta el número entero más cercano, y F es (N-1)/3 redondeado hacia abajo hasta el número entero más cercano. Los nodos de respaldo pueden haber multidifundido los mensajes de preparación respectivamente. En algunas realizaciones, F representa un número máximo de nodos que no funcionan permitidos entre los N nodos para mantener funcionando un sistema de consenso de los N nodos. Por ejemplo, el nodo que realiza el método 520 es uno de los N nodos. El (Q-1) o más mensajes de preparación pueden ser desde distintos modos y ser válidos y consistentes, que indica que (Q-1) o más nodos de respaldo y el nodo primario están de acuerdo con el mensaje de preparación previa.

El bloque 523 incluye almacenar el mensaje de preparación previa y el (Q-1) o más mensajes de preparación. Por ejemplo, si (Q-1) es 2F y si se obtienen 3F mensajes de preparación en la etapa anterior, en este punto, puede almacenarse el mensaje de preparación previa y un número de mensajes de preparación entre 2F y 3F inclusive. En algunas realizaciones, únicamente se almacena el mensaje de preparación previa y los (Q-1) mensajes de preparación. Por ejemplo, si (Q-1) es 2F y si se obtienen 3F mensajes de preparación en la etapa anterior, en este punto, únicamente puede almacenarse el mensaje de preparación previa y 2F mensajes de preparación. En algunas realizaciones, almacenar el mensaje de preparación previa y el (Q-1) o más mensajes de preparación comprende: almacenar únicamente el mensaje de preparación previa y el (Q-1) o más mensajes de preparación. Por ejemplo, únicamente se almacena el mensaje de preparación previa y los (Q-1) mensajes de preparación. No se almacena mensaje distinto del mensaje de preparación previa y el (Q-1) o más mensajes de preparación. Por ejemplo, para cada ronda de consenso-verificación, no se almacenan mensajes de confirmación. Lo mismo puede aplicarse cuando se realizan múltiples rondas de verificaciones de consenso.

En algunas realizaciones, el mensaje de preparación previa y el (Q-1) o más mensajes de preparación pueden almacenarse de diversas maneras siempre que los datos almacenados sean recuperables después de la recuperación de tiempo de inactividad de sistema, tal como un reinicio de sistema. Por ejemplo, el mensaje de preparación previa y el (Q-1) o más mensajes de preparación pueden almacenarse en una base de datos relacional, una base de datos no relacional, un sistema de documento, etc. Por ejemplo, el mensaje de preparación previa y el (Q-1) o más mensajes de preparación pueden almacenarse en un almacenamiento persistente. La etapa de almacenamiento y las otras etapas descritas en el presente documento pueden no estar limitadas por el lenguaje de programación.

En algunas realizaciones, el bloque 523 puede realizarse únicamente cuando se satisface el bloque 522, es decir, únicamente cuando se obtienen (Q-1) o más mensajes de preparación. En otras realizaciones, cada mensaje de preparación previa o preparación puede almacenarse tan pronto como se reciba.

En algunas realizaciones, después de almacenar el mensaje de preparación previa y el (Q-1) o más mensajes de preparación (bloque 523) y antes de multidifundir el mensaje de confirmación (bloque 524), el método comprende adicionalmente: realizar un reinicio de sistema; y cargar el mensaje de preparación previa almacenado y el (Q-1) o más mensajes de preparación almacenados. El reinicio de sistema puede realizarse en respuesta a una interrupción a la operación normal (por ejemplo, avería de sistema, corte de energía, etc.). La interrupción puede ocurrir en uno o más o todos los nodos en el sistema de PBFT. En algunas realizaciones, hasta todos los N nodos experimentan una avería, y al menos Q de los N nodos realizan el reinicio de sistema y cargan el correspondiente mensaje previamente preparado almacenado y el (Q-1) o más mensajes de preparación almacenados respectivamente. A continuación, el protocolo de cambio de vista puede activarse o no.

En una realización, el protocolo de cambio de vista puede no activarse si el reinicio es lo suficientemente rápido para evitar la activación del tiempo de espera, y por lo tanto el reinicio de sistema evita activar el cambio de vista. Por consiguiente, puede seguir el resto de las etapas del método 520 desde el bloque 524. En otras realizaciones, puede activarse el protocolo de cambio de vista, y pueden no seguir el resto de las etapas del método 520 desde el bloque 524.

El bloque 524 incluye multidifundir un mensaje de confirmación a al menos alguno de los nodos de respaldo, indicando el mensaje de confirmación que el nodo primario está de acuerdo con el (Q-1) o más mensajes de preparación. En algunas realizaciones, el mensaje de confirmación indica que el nodo primario ha obtenido el (Q-1) o más mensajes de preparación. En algunas realizaciones, hasta F de los (Q-1) o más de los nodos de respaldo en el bloque 522 son defectuosos o no funcionan de otra manera después de multidifundir los mensajes de confirmación respectivamente y no realizan el reinicio de sistema. Por ejemplo, los F nodos que se han confirmado pueden experimentar una avería de sistema y no reiniciar para reanudar la función. A pesar de eso, la verificación de consenso puede llevarse a cabo de manera apropiada sin provocar resultados inconsistentes y la ramificación a la cadena de bloques.

El bloque 525 incluye obtener, respectivamente desde Q o más nodos entre el nodo primario y los nodos de respaldo,

Q o más mensajes de confirmación indicado cada uno que el correspondiente nodo está de acuerdo con los (Q-1 o más mensajes de preparación recibidos mediante el correspondiente nodo, en donde el Q o más mensajes de confirmación incluyen el mensaje de confirmación de multidifusión. En algunas realizaciones, el mensaje de confirmación indica que el correspondiente nodo que multidifunde el mensaje de confirmación está de acuerdo con el mensaje de preparación previa y ha obtenido (Q-1) o más mensajes de preparación. Los Q o más mensajes de confirmación pueden ser desde nodos distintos y son válidos y consistentes, que indica que están preparados Q o más nodos para ejecutar las solicitudes en el orden. Por lo tanto, se alcanza un consenso por una mayoría de los nodos, y puede realizarse la siguiente etapa de ejecución.

En algunas realizaciones, después de multidifundir el mensaje de confirmación (bloque 525) y antes de ejecutar las solicitudes, el método comprende adicionalmente: realizar un reinicio de sistema, y cargar el mensaje de preparación previa almacenado y el (Q-1) o más mensajes de preparación almacenados. El reinicio de sistema puede realizarse de manera voluntaria o involuntaria. El reinicio de sistema puede provocarse por una interrupción al sistema o a la función del dispositivo, tal como una avería de sistema.

En algunas realizaciones, el método 520 puede incluir adicionalmente empaquetar la una o más transacciones en una copia local de la cadena de bloques mantenida mediante el nodo primario de acuerdo con el orden. Por ejemplo, las solicitudes pueden verificarse en consenso como los al menos (Q-F) nodos honestos (Q mensajes de confirmación pero contabilizando para como máximo F nodos que no funcionan) que han verificado el resumen d en sus mensajes de confirmación (o para el nodo primario, puede no tener que realizar la verificación puesto que propuso el resumen d). Como resultado, si suficientes nodos han verificado las correspondientes transacciones, las transacciones pueden empaquetarse en la cadena de bloques. Puede notificarse al cliente o los clientes (por ejemplo, el nodo A) que enviaron originalmente la solicitud o solicitudes.

La Figura 6A ilustra un diagrama de bloques de un sistema de consenso 610, de acuerdo con diversas realizaciones. El sistema de consenso 610 (por ejemplo, un sistema informático) puede ser un ejemplo de la implementación del nodo 0, nodo 1, nodo 2, ... , o nodo i anteriormente descritos o un dispositivo similar, o una combinación de cualquiera de los nodos y un dispositivo adicional (por ejemplo, el nodo A). El método 510 puede implementarse mediante el sistema de consenso 610. El sistema de consenso 610 puede comprender uno o más procesadores y uno o más medios de almacenamiento legibles por ordenador no transitorios (por ejemplo, una o más memorias) acoplados al uno o más procesadores, y configurados con instrucciones ejecutables por el uno o más procesadores para hacer que el sistema o dispositivo (por ejemplo, el procesador) realicen el método 510. El sistema de consenso 610 puede implementarse en un nodo de respaldo. El sistema de consenso 610 puede comprender diversas unidades/módulos que corresponden a las instrucciones (por ejemplo, instrucciones de software).

En algunas realizaciones, el sistema de consenso 610 puede denominarse como un aparato de consenso. El aparato de consenso puede ser para mantener una cadena de bloques, en donde un número (N) de nodos mantienen la cadena de bloques actuando uno de los N nodos como un nodo primario y actuando los otros (N-1) nodos como nodos de respaldo, actuando el aparato de consenso como uno de los (N-1) nodos de respaldo y que comprende uno o más procesadores y una o más memorias legibles por ordenador no transitorias acopladas al uno o más procesadores y configuradas con instrucciones ejecutables por el uno o más procesadores para hacer que el aparato realice las operaciones. El sistema de consenso puede comprender diversas unidades/módulos que corresponden a las instrucciones (por ejemplo, instrucciones de software). El aparato de consenso puede comprender un primer módulo de obtención 611 para obtener un mensaje de preparación previa desde el nodo primario; un primer módulo de multidifusión 612 para multidifundir un mensaje de preparación a al menos alguno del nodo primario y los otros (N-2) nodos de respaldo, indicando el mensaje de preparación una aceptación del mensaje de preparación previa; un segundo módulo de obtención 613 para obtener (Q-1) o más mensajes de preparación respectivamente desde (Q-1) o más de los nodos de respaldo, en donde Q (quórum) es (N+F+1)/2 redondeado hacia arriba hasta el número entero más cercano, F es (N-1)/3 redondeado hacia abajo hasta el número entero más cercano, y el (Q-1) o más mensajes de preparación incluyen el mensaje de preparación de multidifusión; un módulo de almacenamiento 614 para almacenar el mensaje de preparación previa y el (Q-1) o más mensajes de preparación; un segundo módulo de multidifusión 615 para multidifundir un mensaje de confirmación a al menos alguno del nodo primario y a los otros nodos de respaldo, indicando el mensaje de confirmación que el nodo de respaldo está de acuerdo con el (Q-1) o más mensajes de preparación; y un tercer módulo de obtención 616 para obtener, respectivamente desde Q o más nodos entre el nodo primario y los nodos de respaldo, Q o más mensajes de confirmación indicando cada uno que el correspondiente nodo está de acuerdo con (Q-1) o más mensajes de preparación recibidos mediante el correspondiente nodo, en donde el Q o más mensajes de confirmación incluyen el mensaje de confirmación de multidifusión.

En algunas realizaciones, el primer módulo de obtención 611 u otro módulo también es para obtener una o más solicitudes de transacción desde al menos uno de: uno o más clientes, el nodo primario, o uno o más de los otros nodos de respaldo. El aparato de consenso puede comprender adicionalmente un módulo de empaquetamiento 617 para empaquetar la una o más transacciones en una copia local de la cadena de bloques mantenida mediante el nodo de respaldo de acuerdo con el orden.

La Figura 6B ilustra un diagrama de bloques de un sistema de consenso 620, de acuerdo con diversas realizaciones.

El sistema de consenso 620 (por ejemplo, un sistema informático) puede ser un ejemplo de una implementación del nodo 0, nodo 1, nodo 2, ... , o nodo i anteriormente descritos o un dispositivo similar, o una combinación de cualquiera de los nodos y un dispositivo adicional (por ejemplo, el nodo A). El método 520 puede implementarse mediante el sistema de consenso 620. El sistema de consenso 620 puede comprender uno o más procesadores y uno o más medios de almacenamiento legibles por ordenador no transitorios (por ejemplo, una o más memorias) acoplados al uno o más procesadores, y configurados con instrucciones ejecutables por el uno o más procesadores para hacer que el sistema o dispositivo (por ejemplo, el procesador) realicen el método 520. El sistema de consenso 620 puede implementarse en un nodo primario. El sistema de consenso 620 puede comprender diversas unidades/módulos que corresponden a las instrucciones (por ejemplo, instrucciones de software).

En algunas realizaciones, el sistema de consenso 620 puede denominarse como un aparato de consenso. El aparato de consenso puede ser para mantener una cadena de bloques, en donde un número (N) de nodos mantienen la cadena de bloques actuando uno de los N nodos como un nodo primario y actuando los otros (N-1) nodos como nodos de respaldo, actuando el aparato de consenso como el nodo primario y que comprende uno o más procesadores y una o más memorias legibles por ordenador no transitorias acopladas al uno o más procesadores y configuradas con instrucciones ejecutables por el uno o más procesadores para hacer que el aparato realice las operaciones. El sistema de consenso puede comprender diversas unidades/módulos que corresponden a las instrucciones (por ejemplo, instrucciones de software). El aparato de consenso puede comprender un primer módulo de multidifusión 621 para multidifundir un mensaje de preparación previa a al menos alguno de los nodos de respaldo; un primer módulo de obtención 622 para obtener (Q-1) o más mensajes de preparación respectivamente desde (Q-1) o más de los nodos de respaldo, en donde los mensajes de preparación indica cada uno una aceptación del mensaje de preparación previa mediante el correspondiente nodo de respaldo, Q (quórum) es (N+F+1)/2 redondeado hacia arriba hasta el número entero más cercano, y F es (N-1)/3 redondeado hacia abajo hasta el número entero más cercano; un módulo de almacenamiento 623 para almacenar el mensaje de preparación previa y el (Q-1) o más mensajes de preparación; un segundo módulo de multidifusión 624 para multidifundir un mensaje de confirmación a al menos alguno de los nodos de respaldo, indicando el mensaje de confirmación que el nodo primario está de acuerdo con el (Q-1) o más mensajes de preparación; y un segundo módulo de obtención 625 para obtener, respectivamente desde Q o más nodos entre el nodo primario y los nodos de respaldo, Q o más mensajes de confirmación indicando cada uno que el correspondiente nodo está de acuerdo con (Q-1) o más mensajes de preparación recibidos mediante el correspondiente nodo, en donde el Q o más mensajes de confirmación incluyen el mensaje de confirmación de multidifusión.

En algunas realizaciones, el aparato de consenso puede comprender adicionalmente un tercer módulo de obtención 626 para obtener una o más solicitudes de transacción desde al menos uno de: uno o más clientes o uno o más de los nodos de respaldo. El aparato de consenso puede comprender adicionalmente un módulo de empaquetamiento 627 para empaquetar la una o más transacciones en una copia local de la cadena de bloques mantenida mediante el nodo primario de acuerdo con el orden.

Las técnicas descritas en este documento se implementan mediante uno o más dispositivos informáticos de uso especial. Los dispositivos informáticos de uso especial pueden ser sistemas informáticos de escritorio, sistemas informáticos servidores, sistemas informáticos portátiles, dispositivos portátiles, dispositivos de red o cualquier otro dispositivo o combinación de dispositivos que incorpora lógica por cable y/o de programa para implementar las técnicas. Los dispositivos informáticos de fin especial pueden implementarse como ordenadores personales, portátiles, teléfonos celulares, teléfonos de cámara, teléfonos inteligentes, asistentes digitales personales, reproductores de medios, dispositivos de navegación, dispositivos de correo electrónico, consolas de juegos, ordenadores de tableta, dispositivos llevables, o una combinación de los mismos. El o los dispositivos informáticos generalmente están controlados y coordinados por el software del sistema operativo. Los sistemas operativos convencionales controlan y programan procesos informáticos para su ejecución, realizan gestión de memoria, proporcionan sistema de archivos, interconexión en red, servicios de E/S y proporcionan una funcionalidad de interfaz de usuario, tal como una interfaz gráfica de usuario ("GUI"), entre otras cosas. Los diversos sistemas, aparatos, medios de almacenamiento, módulos, y unidades descritas en el presente documento pueden implementarse en los dispositivos informáticos de fin especial, o en uno o más chips informáticos del uno o más dispositivos informáticos de fin especial. En algunas realizaciones, las instrucciones descritas en el presente documento pueden implementarse en una máquina virtual en el dispositivo informático de fin especial. Cuando se ejecutan, las instrucciones pueden hacer que el dispositivo informático de fin especial realice diversos métodos descritos en el presente documento. La máquina virtual puede incluir un software, hardware o una combinación de los mismos. Por ejemplo, la máquina virtual puede incluir un software de Máquina Virtual de Ethereum (EVM) que proporciona el entorno de tiempo de ejecución para contratos inteligentes en Ethereum.

La Figura 7 es un diagrama de bloques que ilustra un sistema informático 700 en el que se puede implementar cualquiera de las realizaciones descritas en el presente documento. El sistema 700 puede realizar cualquiera de los métodos descritos en el presente documento (por ejemplo, el método de consenso 510, el método de consenso 520). El sistema 700 puede implementarse en cualquiera de los sistemas descritos en el presente documento (por ejemplo, el sistema de consenso 610, el sistema de consenso 620). El sistema 700 puede implementarse en cualquiera de los nodos descritos en el presente documento y configurarse para realizar correspondientes etapas para implementar el contrato de cadena de bloques. El sistema informático 700 incluye un bus 702 u otro mecanismo de comunicación para comunicar información, y uno o más procesador o procesadores de hardware 704 acoplados con el bus 702 para procesar información. El procesador o procesadores de hardware 704 pueden ser, por ejemplo, uno o más

microprocesadores de fin general.

El sistema informático 700 también incluye una memoria principal 706, tal como una memoria de acceso aleatorio (RAM), caché y/u otros dispositivos de almacenamiento dinámico, acoplados al bus 702 para almacenar información e instrucciones ejecutables por el procesador o procesadores 704. La memoria principal 706 puede usarse también para almacenar variables temporales u otra información intermedia durante la ejecución de instrucciones ejecutables por el procesador o los procesadores 704. Tales instrucciones, cuando se almacenan en medios de almacenamiento accesibles al procesador o procesadores 704, representan el sistema informático 700 en una máquina de uso especial que se personaliza para realizar las operaciones especificadas en las instrucciones. El sistema informático 700 incluye adicionalmente una memoria de sólo lectura (ROM) 708 u otro dispositivo de almacenamiento estático acoplado al bus 702 para almacenar información estática e instrucciones para el procesador o procesadores 704. Se proporciona un dispositivo de almacenamiento 710, tal como un disco magnético, un disco óptico o una unidad de memoria USB (unidad flash), etc., y se acopla al bus 702 para almacenar información e instrucciones.

El sistema informático 700 puede implementar las técnicas descritas en el presente documento usando lógica de cableado permanente personalizada, uno o más ASIC o FPGA, firmware y/o lógica de programa que, en combinación con el sistema informático, provoca o programa al sistema informático 700 para que sea una máquina de fin especial. De acuerdo con una realización, las operaciones, métodos y procesos descritos en el presente documento son realizados por el sistema informático 700 en respuesta a que el procesador o procesadores 704 ejecuten una o más secuencias de una o más instrucciones contenidas en la memoria principal 706. Tales instrucciones pueden leerse en la memoria principal 706 desde otro medio de almacenamiento, tal como el dispositivo de almacenamiento 710. La ejecución de las secuencias de instrucciones contenidas en la memoria principal 706 provoca que el procesador o procesadores 704 realicen las etapas del proceso descritas en el presente documento. En realizaciones alternativas, puede usarse circuitería de cableado permanente en lugar de o en combinación con instrucciones de software.

La memoria principal 706, la ROM 708, y/o el almacenamiento 710 pueden incluir medios de almacenamiento no transitorios. La expresión "medios no transitorios" y términos similares, como se usan en el presente documento, se refieren a medios que almacenan datos y/o instrucciones que hacen que una máquina funcione de una manera específica, los medios excluyen las señales transitorias. Tales medios no transitorios pueden comprender medios no volátiles y/o medios volátiles. Los medios no volátiles incluyen, por ejemplo, discos ópticos o magnéticos, tales como un dispositivo de almacenamiento 710. Los medios volátiles incluyen memoria dinámica, tal como la memoria principal 706. Las formas comunes de medios no transitorios incluyen, por ejemplo, un disco flexible, un disco flexible, un disco duro, disco de estado sólido, cinta magnética, o cualquier otro medio de almacenamiento de datos magnético, un CD-ROM, cualquier otro medio de almacenamiento de datos óptico, cualquier medio físico con patrones de orificios, una RAM, una PROM y EPROM, una FLASH-EPROM, NVRAM, cualquier otro chip o cartucho de memoria y versiones en red de los mismos.

El sistema informático 700 también incluye una interfaz de comunicación 718 acoplada al bus 702. La interfaz de red 718 proporciona un acoplamiento de comunicación de datos bidireccional a uno o más enlaces de red que están conectados a una o más redes locales. Por ejemplo, la interfaz de red 718 puede ser una tarjeta de red digital de servicios integrados (ISDN), módem por cable, módem por satélite o un módem para proporcionar una conexión de comunicación de datos a un correspondiente tipo de línea telefónica. Como otro ejemplo, la interfaz de red 718 puede ser una tarjeta de red de área local (LAN) para proporcionar una conexión de comunicación de datos a una LAN compatible (o componente WAN para comunicarse con una WAN). También pueden implementarse enlaces inalámbricos. En cualquier implementación de este tipo, la interfaz de red 718 envía y recibe señales eléctricas, electromagnéticas u ópticas que transportan flujos de datos digitales que representan diversos tipos de información.

El sistema informático 700 puede enviar mensajes y recibir datos, incluyendo código de programa, a través de la red o redes, el enlace de red y la interfaz de comunicación 718. En el ejemplo de Internet, un servidor podría transmitir un código solicitado para un programa de aplicación a través de Internet, el ISP, la red local y la interfaz de red 718.

El código recibido puede ejecutarse mediante el procesador o procesadores 704 como se recibió, y/o almacenarse en el dispositivo de almacenamiento 710, u otro almacenamiento no volátil para su ejecución posterior.

Cada uno de los procesos, métodos y algoritmos descritos en las secciones anteriores puede estar incorporado y automatizado total o parcialmente por módulos de código ejecutados por uno o más sistemas informáticos o procesadores informáticos que comprenden hardware informático. Los procesos y algoritmos pueden implementarse parcial o totalmente en circuitería específica de la aplicación.

Las diversas características y procesos anteriormente descritos pueden usarse de manera independiente unos de otros, o pueden combinarse de diversas maneras. Se pretende que todas las posibles combinaciones y subcombinaciones entren dentro del alcance de la presente memoria descriptiva. Además, ciertos bloques de métodos o proceso pueden omitirse en algunas implementaciones. Los métodos y procesos descritos en el presente documento tampoco están limitados a ninguna secuencia particular y los bloques o estados relacionados con los mismos pueden realizarse en otras secuencias que sean apropiadas. Por ejemplo, los bloques o estados descritos pueden realizarse en un orden distinto del específicamente desvelado o múltiples bloques o estados pueden combinarse en un único

bloque o estado. Los ejemplos de bloques o estados pueden realizarse en serie, en paralelo o en alguna otra manera. Los bloques o estados pueden añadirse a o retirarse de las realizaciones desveladas. Los ejemplos de sistemas y componentes descritos en el presente documento pueden configurarse de manera diferente a lo descrito. Por ejemplo, pueden añadirse elementos, eliminarse, o reorganizarse en comparación con las realizaciones de desveladas.

Las diversas operaciones de los métodos descritos en el presente documento pueden realizarse, al menos parcialmente, por uno o más procesadores que están configurados temporalmente (por ejemplo, por software) o configurados permanentemente para realizar las operaciones relevantes. Ya sea que estén configurados de forma temporal o permanente, dichos procesadores pueden constituir motores implementados por procesadores que operan para realizar una o más operaciones o funciones descritas en el presente documento.

De manera similar, los métodos descritos en el presente documento pueden implementarse al menos parcialmente con un procesador, siendo un procesador o procesadores particulares un ejemplo de hardware. Por ejemplo, al menos algunas de las operaciones de un método pueden realizarse por uno o más procesadores o motores implementados por procesador. Además, el uno o más procesadores también pueden operar para soportar la realización de las operaciones relevantes en un entorno de "informática en la nube" o como un "software como servicio" (SaaS). Por ejemplo, al menos algunas de las operaciones pueden realizarse por un grupo de ordenadores (como ejemplos de máquinas que incluyen procesadores), siendo estas operaciones accesibles a través de una red (por ejemplo, Internet) y a través de una o más interfaces apropiadas (por ejemplo, una Interfaz de Programa de Aplicación (API)).

El rendimiento de algunas de las operaciones puede distribuirse entre los procesadores, que no solo se encuentran dentro de una sola máquina, sino que están desplegados en varias máquinas. En algunas realizaciones, los procesadores o motores implementados por procesador pueden ubicarse en una única ubicación geográfica (por ejemplo, dentro de un entorno doméstico, un entorno de oficina o una granja de servidores). En otras realizaciones, los procesadores o motores implementados por procesador pueden distribuirse a través de varias ubicaciones geográficas.

A lo largo de la presente memoria descriptiva, varias instancias pueden implementar componentes, operaciones o estructuras descritas como una única instancia. Aunque las operaciones individuales de uno o más métodos se ilustran y describen como operaciones separadas, una o más de las operaciones individuales pueden realizarse simultáneamente y nada requiere que las operaciones se realicen en el orden ilustrado. Las estructuras y la funcionalidad presentadas como componentes separados en las configuraciones se pueden implementar como una estructura o componente combinado. De forma similar, las estructuras y la funcionalidad presentadas como un único componente se pueden implementar como componentes separados. Estas y otras variaciones, modificaciones, adiciones y mejoras entran dentro del alcance de la materia objeto del presente documento. Adicionalmente, los términos relacionados (tales como "primero", "segundo", "tercero" etc.) usados en el presente documento no indican cualquier orden, altura, o importancia, sino que en su lugar se usan para distinguir un elemento de otro elemento. Adicionalmente, los términos "un" "una" y "pluralidad" no indican una limitación de cantidad en el presente documento, sino que en su lugar indican la presencia de al menos uno de los artículos mencionados.

REIVINDICACIONES

1. Un método de consenso basado en Tolerancia Práctica a Fallos Bizantinos implementado por ordenador para implementarse en una cadena de bloques mantenida mediante un número (N) de nodos, en donde uno de los nodos actúa como un nodo primario y los otros (N-1) nodos actúan como nodos de respaldo, y el método se realiza mediante el nodo primario, comprendiendo el método:

multidifundir (412a, 512) un mensaje de preparación previa a al menos alguno de los nodos de respaldo;
obtener (413, 513) Q-1 o más mensajes de preparación respectivamente desde Q-1 o más de los nodos de respaldo, en donde cada uno de los mensajes de preparación indica una aceptación del mensaje de preparación previa mediante el correspondiente nodo de respaldo, en donde Q es un quórum igual a (N+F+1)/2 redondeado hacia arriba hasta el número entero más cercano, y F es (N-1)/3 redondeado hacia abajo hasta el número entero más cercano;
almacenar (414, 514) al menos una cantidad mínima de mensajes de consenso para la recuperación después de que se averíen uno o más de los N nodos, en donde los mensajes de consenso en la cantidad mínima de mensajes de consenso comprenden el mensaje de preparación previa y al menos Q-1 de los Q-1 o más mensajes de preparación;
multidifundir (415, 515) un mensaje de confirmación a al menos alguno de los nodos de respaldo, indicando el mensaje de confirmación que el nodo primario está de acuerdo con los Q-1 o más mensajes de preparación; y
obtener (416, 516), respectivamente desde Q o más nodos entre el nodo primario y los nodos de respaldo, Q o más mensajes de confirmación indicando cada uno que un correspondiente nodo de los Q o más nodos está de acuerdo con Q-1 o más mensajes de preparación recibidos mediante el correspondiente nodo.

2. El método de la reivindicación 1, en donde:
antes de multidifundir el mensaje de preparación previa a al menos alguno de los nodos de respaldo, el método comprende además obtener una o más solicitudes de transacción desde al menos uno de: uno o más clientes o uno o más de los nodos de respaldo.

3. El método de la reivindicación 2, en donde:

el mensaje de preparación previa comprende un orden de una o más transacciones que corresponden a la una o más solicitudes de transacción; y
el mensaje de confirmación indica que el correspondiente nodo que envió el mensaje de confirmación está de acuerdo con el orden.

4. El método de la reivindicación 3, que comprende adicionalmente:
empaquetar la una o más transacciones en una copia local de la cadena de bloques mantenida mediante el nodo primario de acuerdo con el orden.

5. El método de cualquier reivindicación anterior, en donde:
el Q o más mensajes de confirmación incluyen el mensaje de confirmación de multidifusión.

6. El método de cualquier reivindicación anterior, en donde almacenar (414) al menos una cantidad mínima de mensajes de consenso para la recuperación después de que se averíen uno o más de los N nodos comprende:
almacenar únicamente el mensaje de preparación previa y los al menos Q-1 de los Q-1 o más mensajes de preparación.

7. El método de cualquier reivindicación anterior, después de multidifundir el mensaje de confirmación, que comprende además:

realizar un reinicio de sistema; y
cargar al menos la cantidad mínima almacenada de mensajes de consenso.

8. El método de cualquiera de las reivindicaciones 1-6, después de almacenar al menos la cantidad mínima de mensajes de consenso y antes de multidifundir el mensaje de confirmación, que comprende además:

realizar un reinicio de sistema; y
cargar al menos la cantidad mínima almacenada de mensajes de consenso.

9. El método de la reivindicación 8, en donde la realización del reinicio de sistema comprende:
realizar el reinicio de sistema sin activar un cambio de vista.

10. Un sistema de consenso que actúa como el nodo primario para mantener la cadena de bloques, que comprende:

uno o más procesadores; y
una o más memorias legibles por ordenador acopladas al uno o más procesadores y que tienen instrucciones

almacenadas en las mismas que son ejecutables mediante el uno o más procesadores para realizar las etapas del método de cualquiera de las reivindicaciones 1 a 9.

11. Un aparato de consenso que actúa como el nodo primario para mantener la cadena de bloques, comprendiendo el aparato una pluralidad de módulos para realizar las etapas del método de cualquiera de las reivindicaciones 1 a 9.

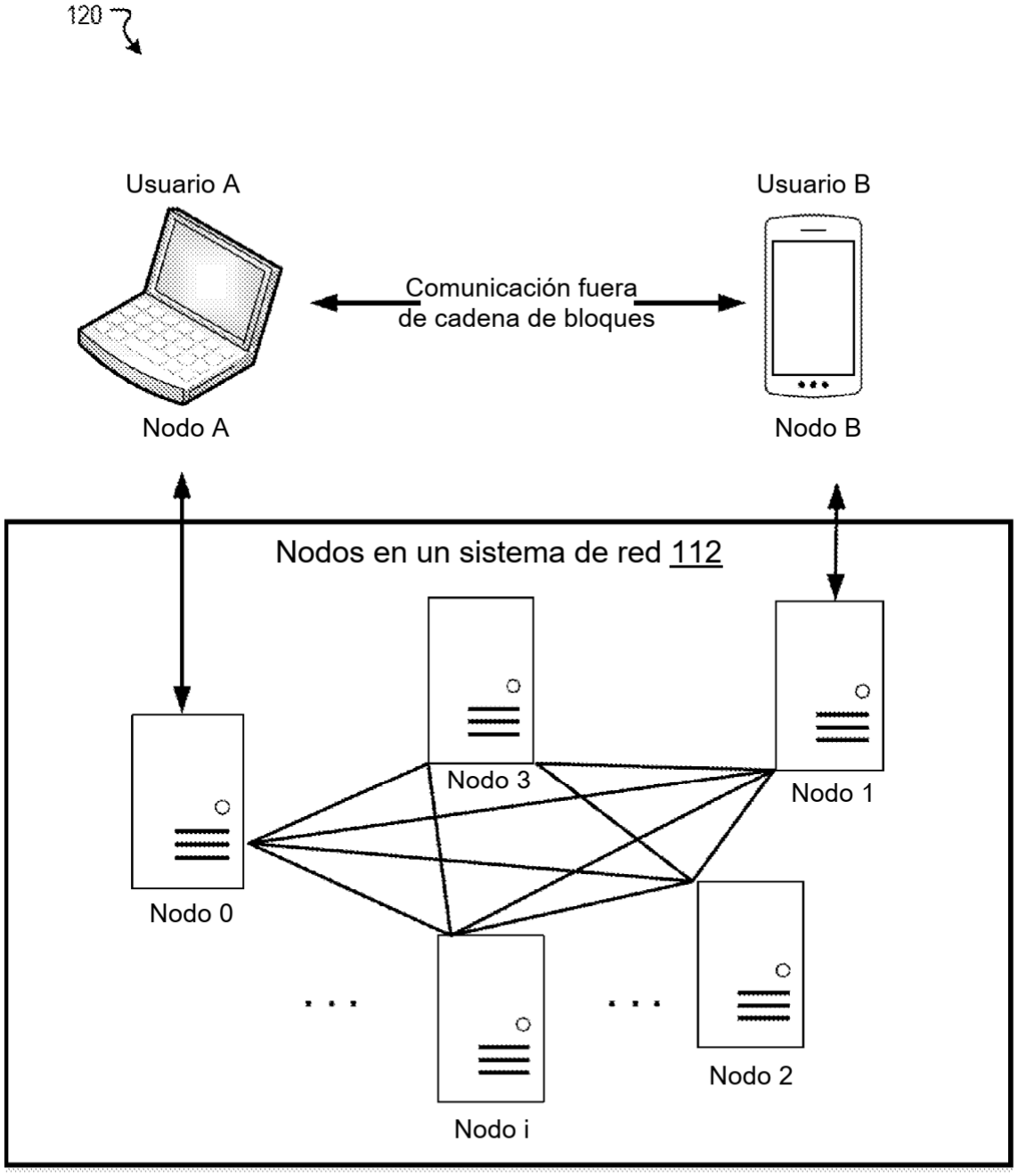
120
Usuario A
Usuario B
Comunicación fuera
de cadena de bloques
Nodo A
Nodo B
Nodos en un sistema de red 112
Nodo 3
Nodo 1
Nodo 0
Nodo 2
Nodo i

FIG. 1

Protocolo de operación normal

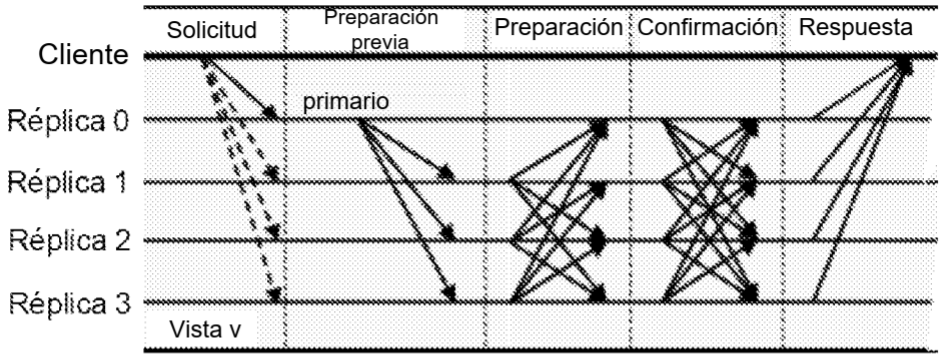

FIG. 2A

Protocolo de operación normal con un nodo de respaldo defectuoso

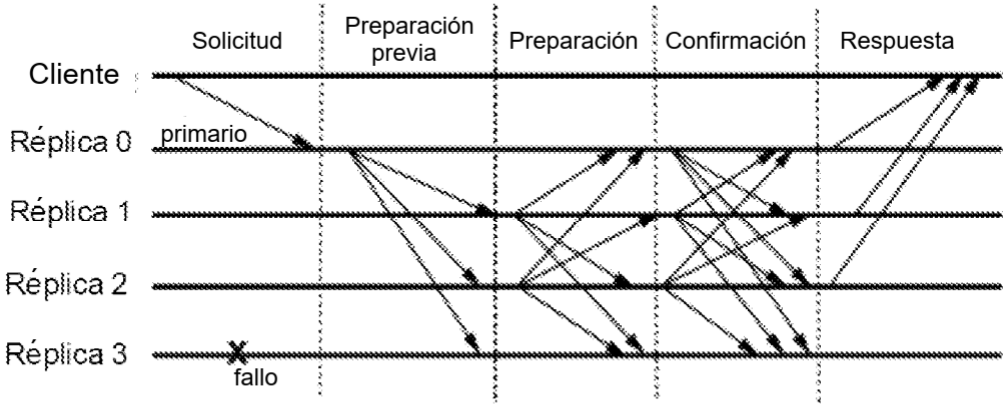

FIG. 2B

Protocolo de operación normal **Protocolo de cambio de vista**

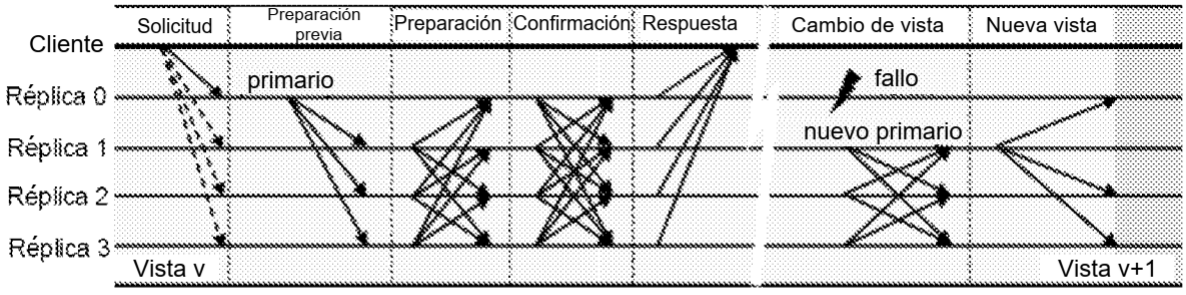

FIG. 2C

Protocolo de operación normal

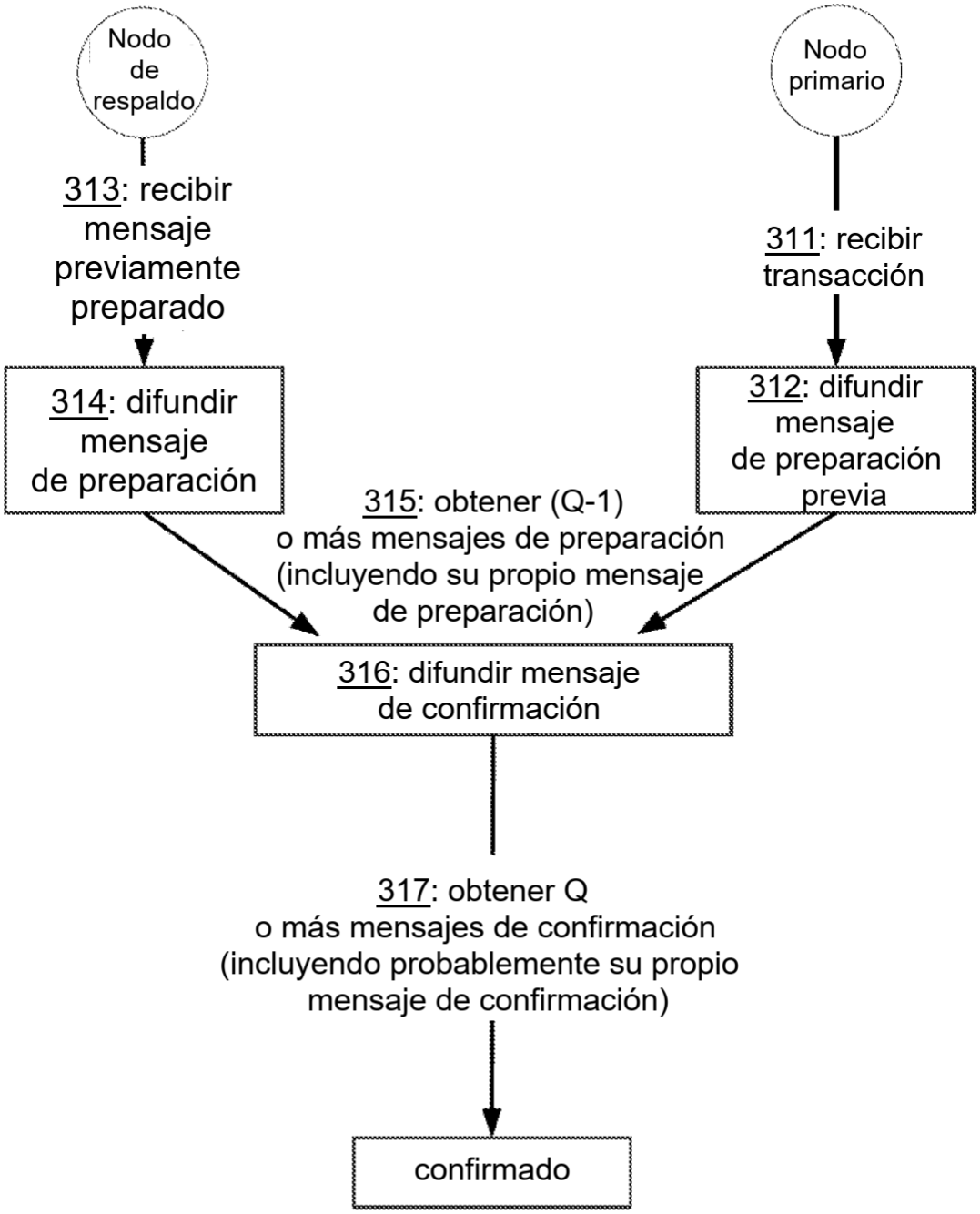

FIG. 3A

Protocolo de cambio de vista

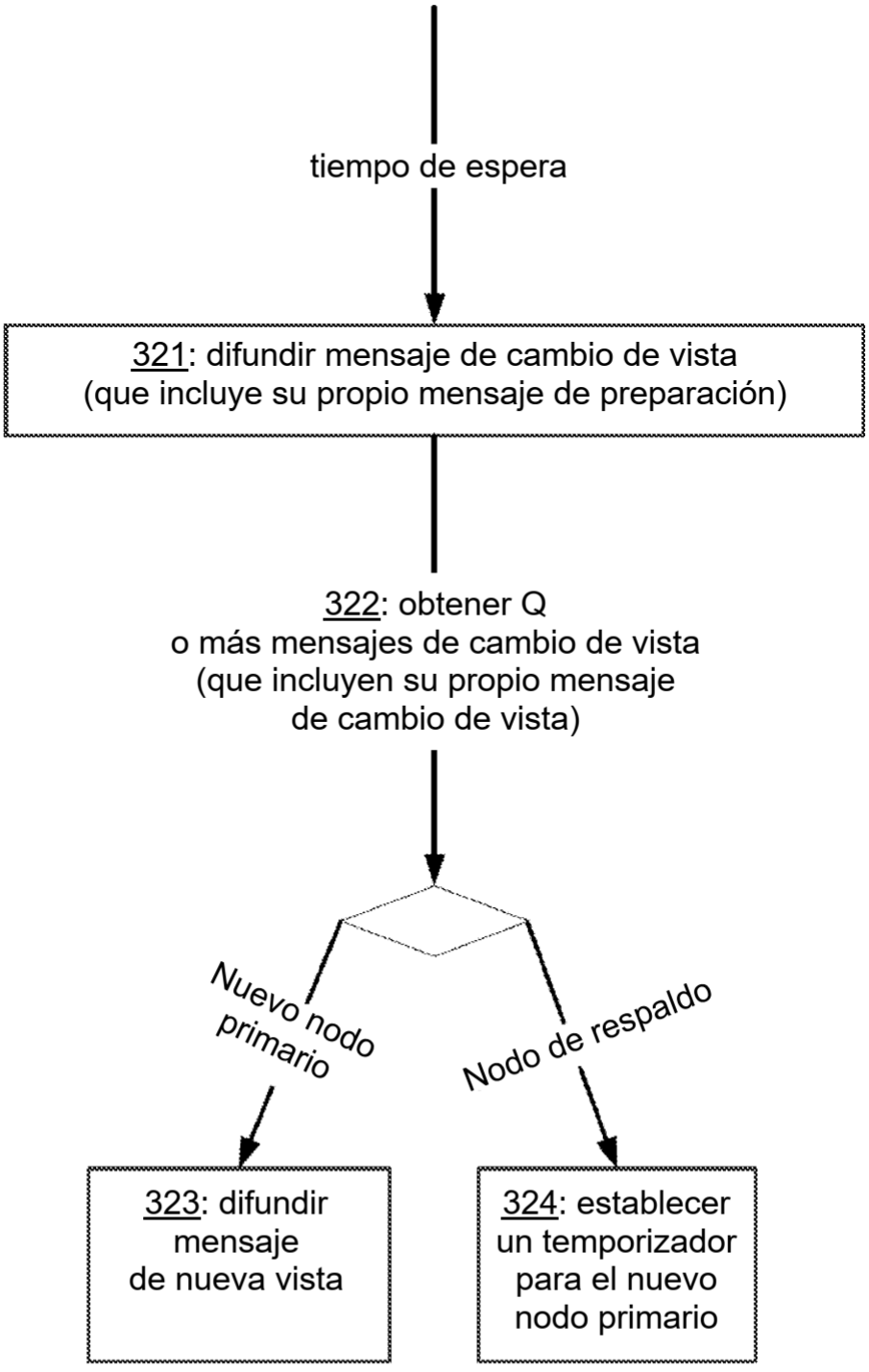

FIG. 3B

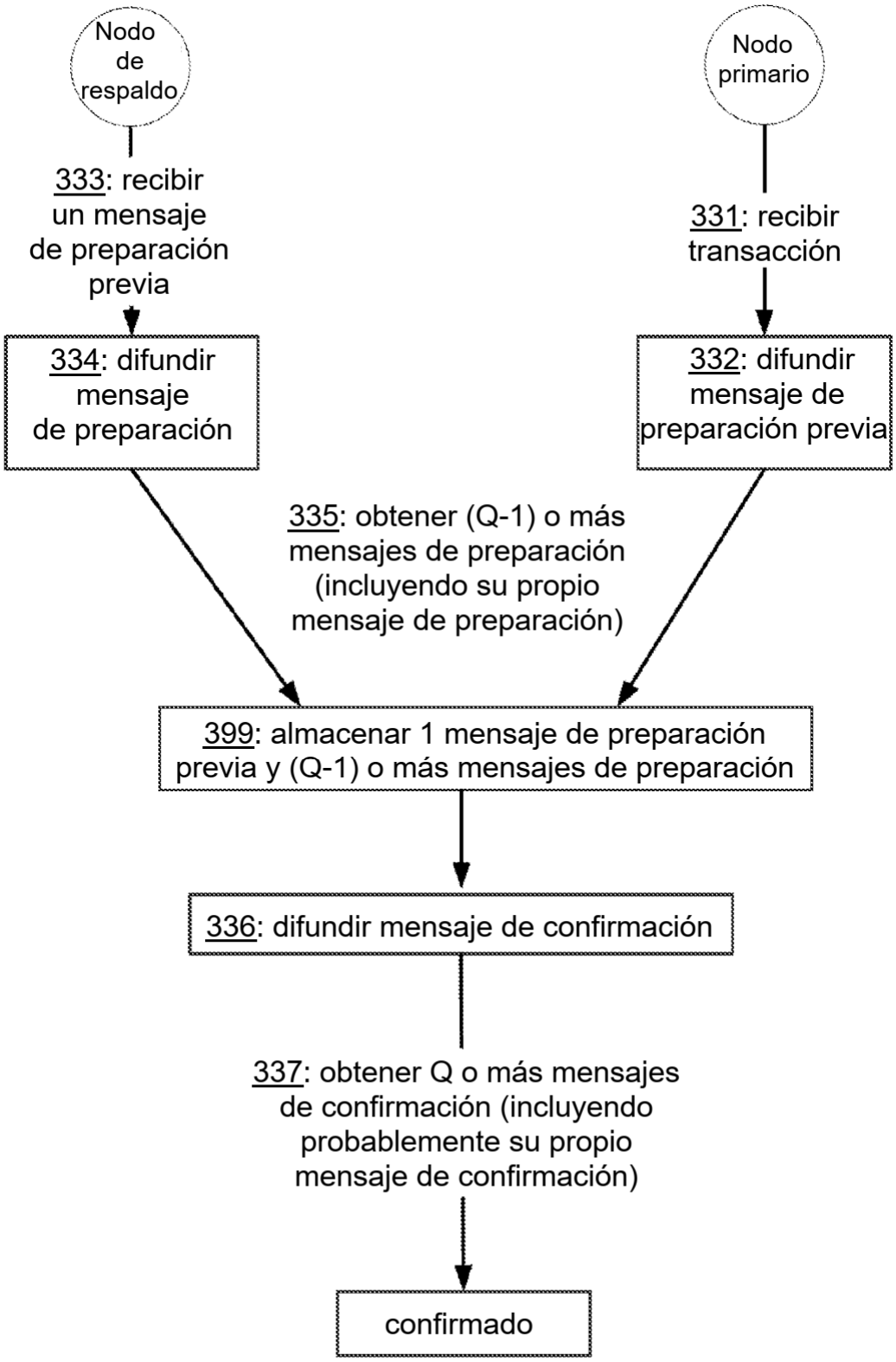
Nodo de respaldo
Nodo primario
333: recibir un mensaje de preparación previa
331: recibir transacción
334: difundir mensaje de preparación
332: difundir mensaje de preparación previa
335: obtener (Q-1) o más mensajes de preparación (incluyendo su propio mensaje de preparación)
399: almacenar 1 mensaje de preparación previa y (Q-1) o más mensajes de preparación
336: difundir mensaje de confirmación
337: obtener Q o más mensajes de confirmación (incluyendo probablemente su propio mensaje de confirmación)
confirmado

FIG. 3C

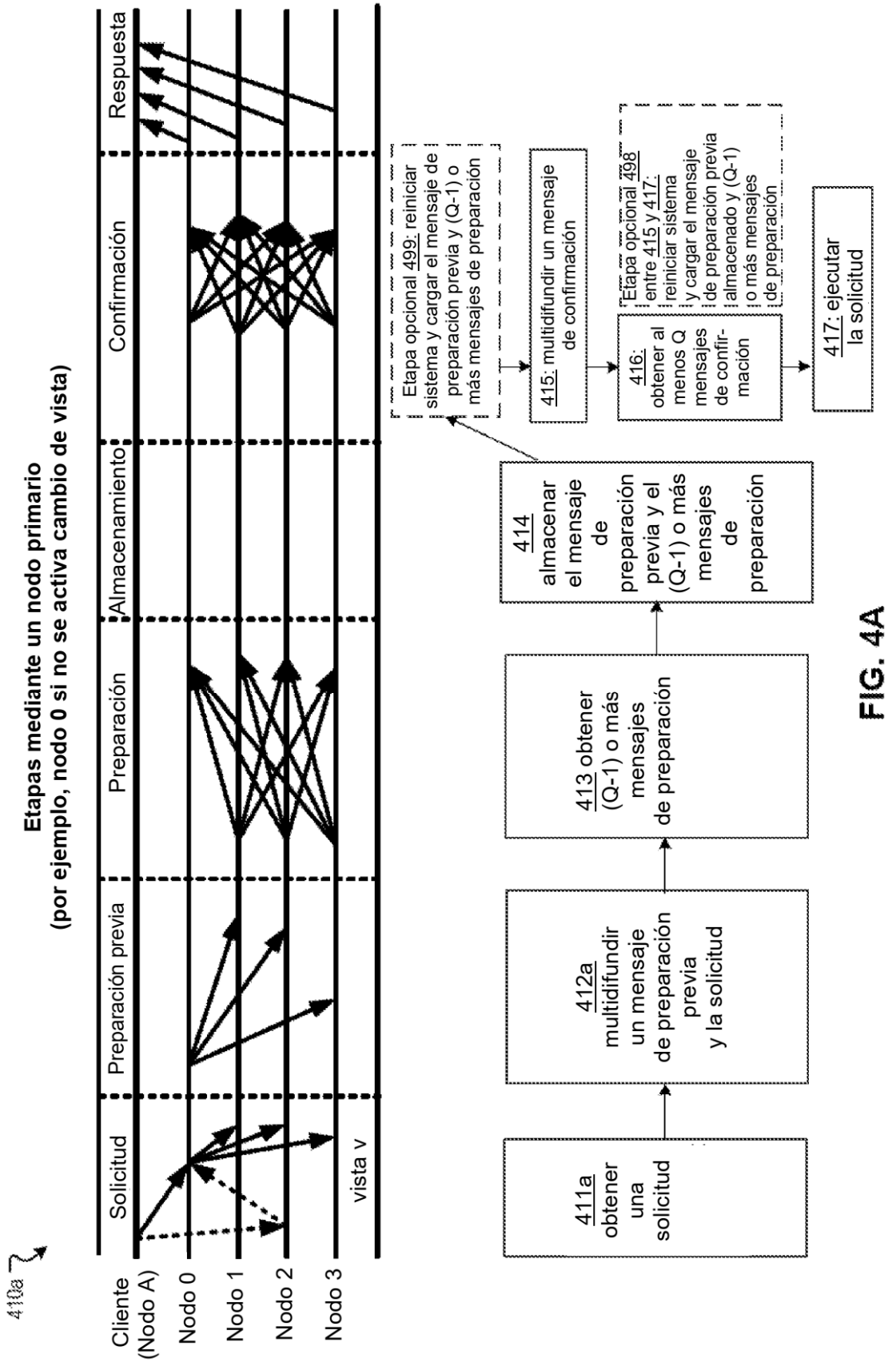
411a
Etapas mediante un nodo primario
(por ejemplo, nodo 0 si no se activa cambio de vista)
Solicitud
Preparación previa
Preparación
Almacenamiento
Confirmación
Respuesta
Cliente (Nodo A)
Nodo 0
Nodo 1
Nodo 2
Nodo 3
vista v
411a obtener una solicitud
412a multidifundir un mensaje de preparación previa y la solicitud
413 obtener (Q-1) o más mensajes de preparación
414 almacenar el mensaje de preparación previa y el (Q-1) o más mensajes de preparación
Etapa opcional 499: reiniciar sistema y cargar el mensaje de preparación previa y (Q-1) o más mensajes de preparación
415: multidifundir un mensaje de confirmación
416: obtener al menos Q mensajes de confirmación
Etapa opcional 498 entre 415 y 417: reiniciar sistema y cargar el mensaje de preparación previa almacenado y (Q-1) o más mensajes de preparación
417: ejecutar la solicitud

FIG. 4A

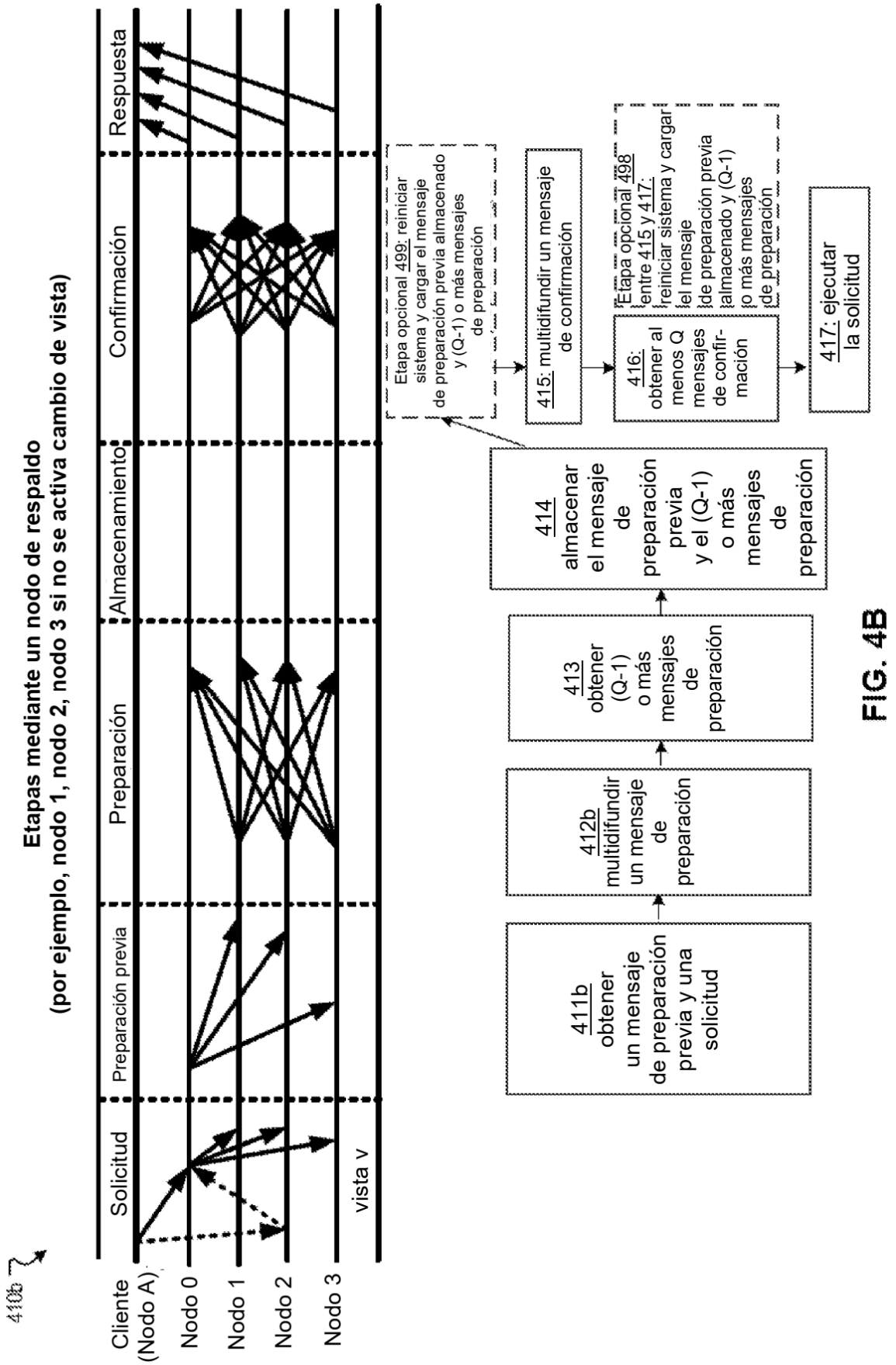
410b
Etapas mediante un nodo de respaldo
(por ejemplo, nodo 1, nodo 2, nodo 3 si no se activa cambio de vista)
Solicitud
Preparación previa
Preparación
Almacenamiento
Confirmación
Respuesta
Cliente (Nodo A)
Nodo 0
Nodo 1
Nodo 2
Nodo 3
vista v
411b obtener un mensaje de preparación previa y una solicitud
412b multidifundir un mensaje de preparación
413 obtener (Q-1) o más mensajes de preparación
414 almacenar el mensaje de preparación previa y el (Q-1) o más mensajes de preparación
Etapa opcional 499: reiniciar sistema y cargar el mensaje de preparación previa almacenado y (Q-1) o más mensajes de preparación
415: multidifundir un mensaje de confirmación
416: obtener al menos Q mensajes de confirmación
Etapa opcional 498 entre 415 y 417: reiniciar sistema y cargar el mensaje de preparación previa almacenado y (Q-1) o más mensajes de preparación
417: ejecutar la solicitud

FIG. 4B

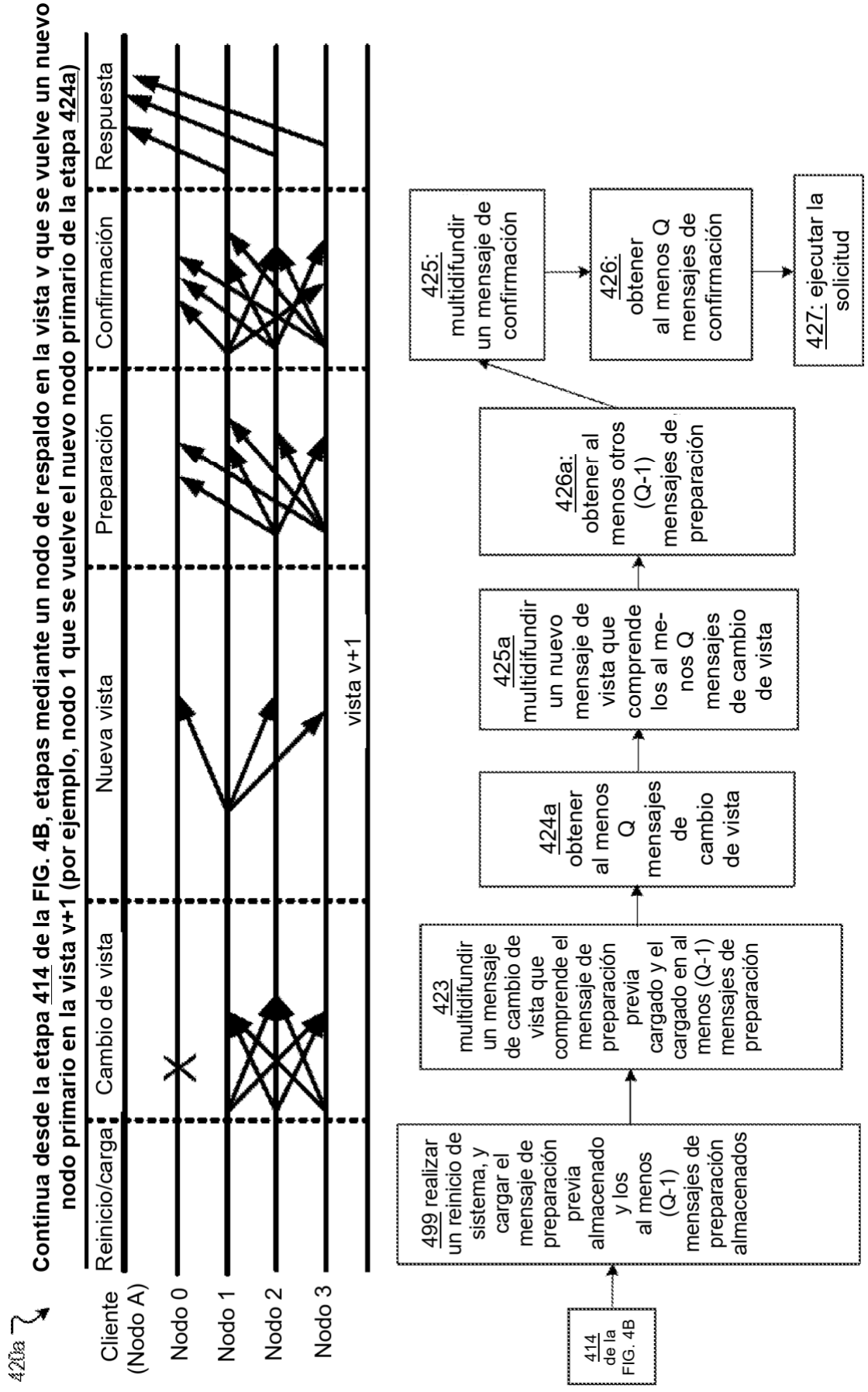
421a
Continua desde la etapa 414 de la FIG. 4B, etapas mediante un nodo de respaldo en la vista v que se vuelve un nuevo nodo primario en la vista v+1 (por ejemplo, nodo 1 que se vuelve el nuevo nodo primario de la etapa 424a)
Reinicio/carga
Cambio de vista
Nueva vista
Preparación
Confirmación
Respuesta
Cliente (Nodo A)
Nodo 0
Nodo 1
Nodo 2
Nodo 3
vista v+1
414 de la FIG. 4B
499 realizar un reinicio de sistema, y cargar el mensaje de preparación previa almacenado y los al menos (Q-1) mensajes de preparación almacenados
423 multidifundir un mensaje de cambio de vista que comprende el mensaje de preparación previa cargado y el cargado en al menos (Q-1) mensajes de preparación
424a obtener al menos Q mensajes de cambio de vista
425a multidifundir un nuevo mensaje de vista que comprende los al menos Q mensajes de cambio de vista
426a: obtener al menos otros (Q-1) mensajes de preparación
425: multidifundir un mensaje de confirmación
426: obtener al menos Q mensajes de confirmación
427: ejecutar la solicitud

FIG. 4C

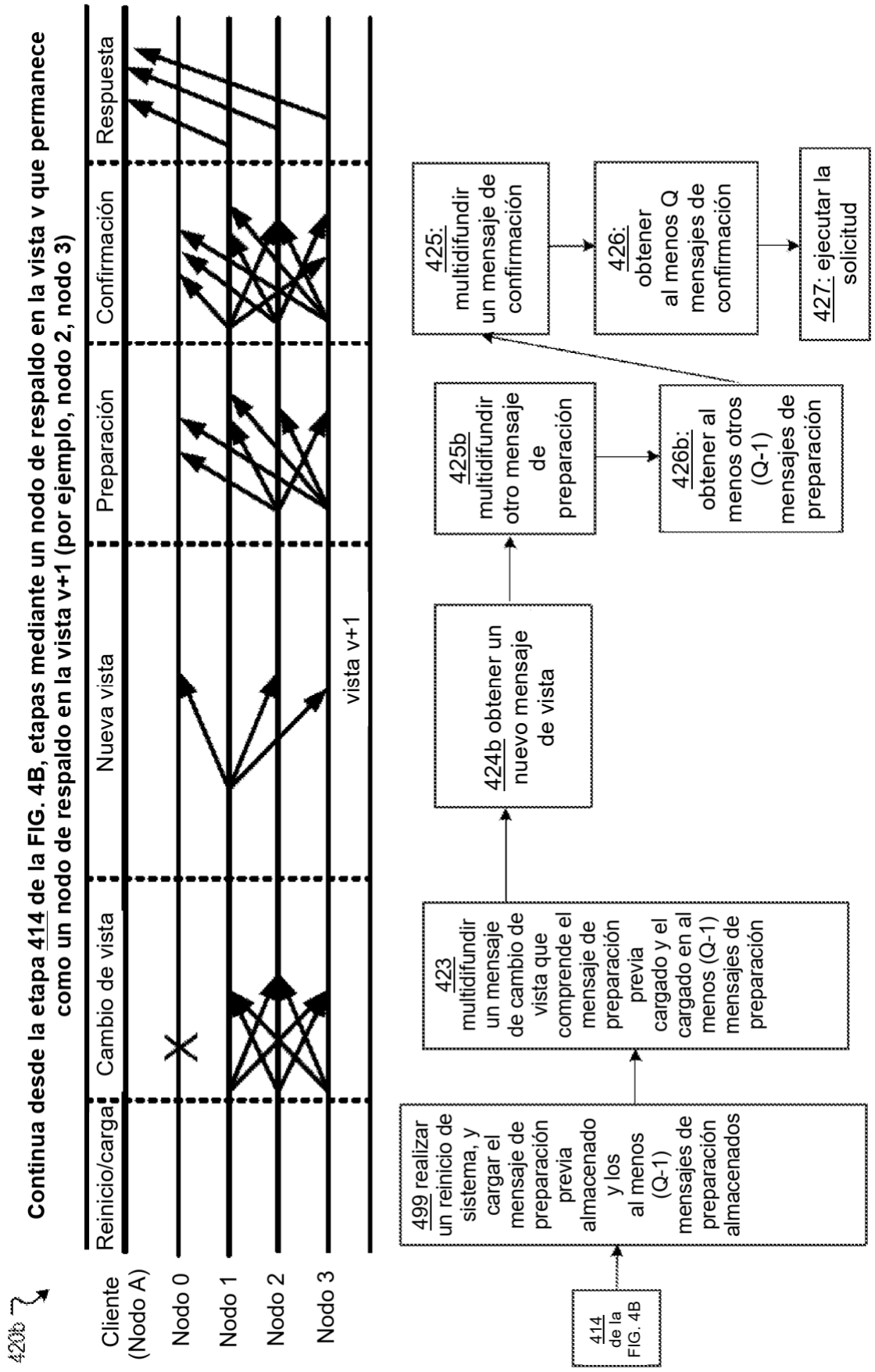
420b
Continua desde la etapa 414 de la FIG. 4B, etapas mediante un nodo de respaldo en la vista v que permanece como un nodo de respaldo en la vista v+1 (por ejemplo, nodo 2, nodo 3)
Reinicio/carga
Cambio de vista
Nueva vista
Preparación
Confirmación
Respuesta
Cliente (Nodo A)
Nodo 0
Nodo 1
Nodo 2
Nodo 3
vista v+1
414 de la FIG. 4B
499 realizar un reinicio de sistema, y cargar el mensaje de preparación previa almacenado y los al menos (Q-1) mensajes de preparación almacenados
423 multidifundir un mensaje de cambio de vista que comprende el mensaje de preparación previa cargado y el cargado en al menos (Q-1) mensajes de preparación
424b obtener un nuevo mensaje de vista
425b multidifundir otro mensaje de preparación
426b: obtener al menos otros (Q-1) mensajes de preparación
425: multidifundir un mensaje de confirmación
426: obtener al menos Q mensajes de confirmación
427: ejecutar la solicitud

FIG. 4D

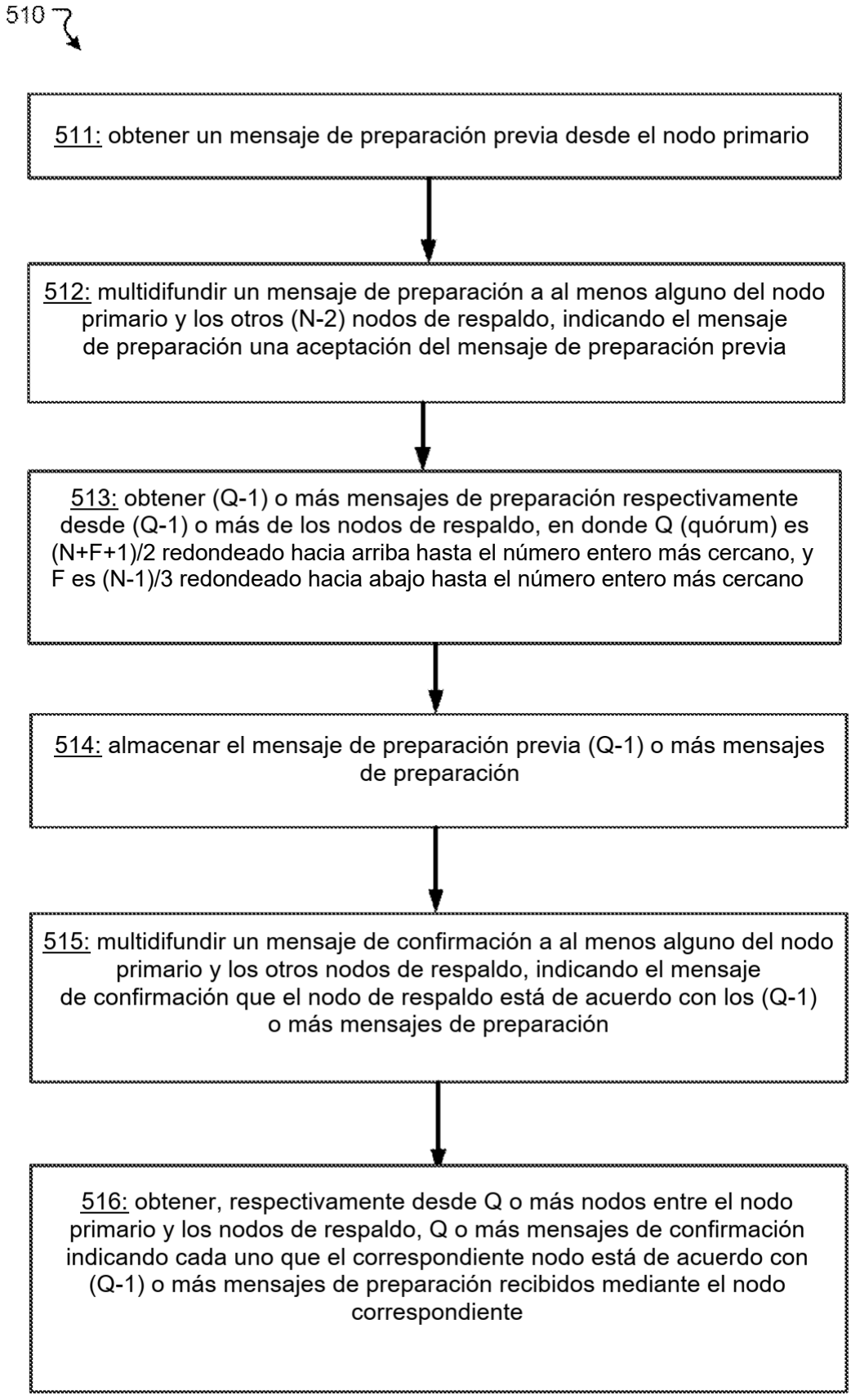
510
511: obtener un mensaje de preparación previa desde el nodo primario
512: multidifundir un mensaje de preparación a al menos alguno del nodo primario y los otros (N-2) nodos de respaldo, indicando el mensaje de preparación una aceptación del mensaje de preparación previa
513: obtener (Q-1) o más mensajes de preparación respectivamente desde (Q-1) o más de los nodos de respaldo, en donde Q (quórum) es (N+F+1)/2 redondeado hacia arriba hasta el número entero más cercano, y F es (N-1)/3 redondeado hacia abajo hasta el número entero más cercano
514: almacenar el mensaje de preparación previa (Q-1) o más mensajes de preparación
515: multidifundir un mensaje de confirmación a al menos alguno del nodo primario y los otros nodos de respaldo, indicando el mensaje de confirmación que el nodo de respaldo está de acuerdo con los (Q-1) o más mensajes de preparación
516: obtener, respectivamente desde Q o más nodos entre el nodo primario y los nodos de respaldo, Q o más mensajes de confirmación indicando cada uno que el correspondiente nodo está de acuerdo con (Q-1) o más mensajes de preparación recibidos mediante el nodo correspondiente

FIG. 5A

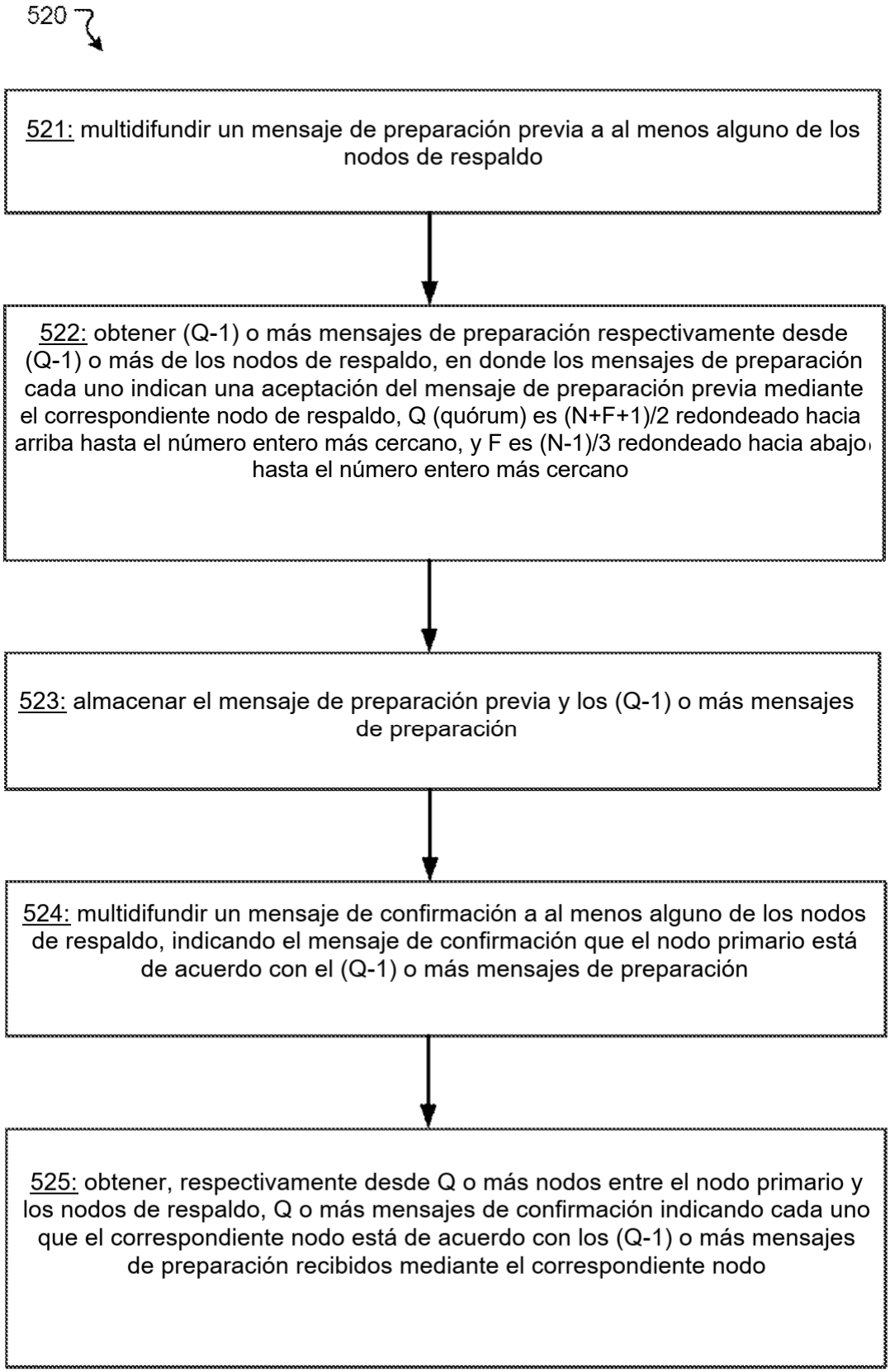
520
521: multidifundir un mensaje de preparación previa a al menos alguno de los nodos de respaldo
522: obtener (Q-1) o más mensajes de preparación respectivamente desde (Q-1) o más de los nodos de respaldo, en donde los mensajes de preparación cada uno indican una aceptación del mensaje de preparación previa mediante el correspondiente nodo de respaldo, Q (quórum) es (N+F+1)/2 redondeado hacia arriba hasta el número entero más cercano, y F es (N-1)/3 redondeado hacia abajo hasta el número entero más cercano
523: almacenar el mensaje de preparación previa y los (Q-1) o más mensajes de preparación
524: multidifundir un mensaje de confirmación a al menos alguno de los nodos de respaldo, indicando el mensaje de confirmación que el nodo primario está de acuerdo con el (Q-1) o más mensajes de preparación
525: obtener, respectivamente desde Q o más nodos entre el nodo primario y los nodos de respaldo, Q o más mensajes de confirmación indicando cada uno que el correspondiente nodo está de acuerdo con los (Q-1) o más mensajes de preparación recibidos mediante el correspondiente nodo

FIG. 5B

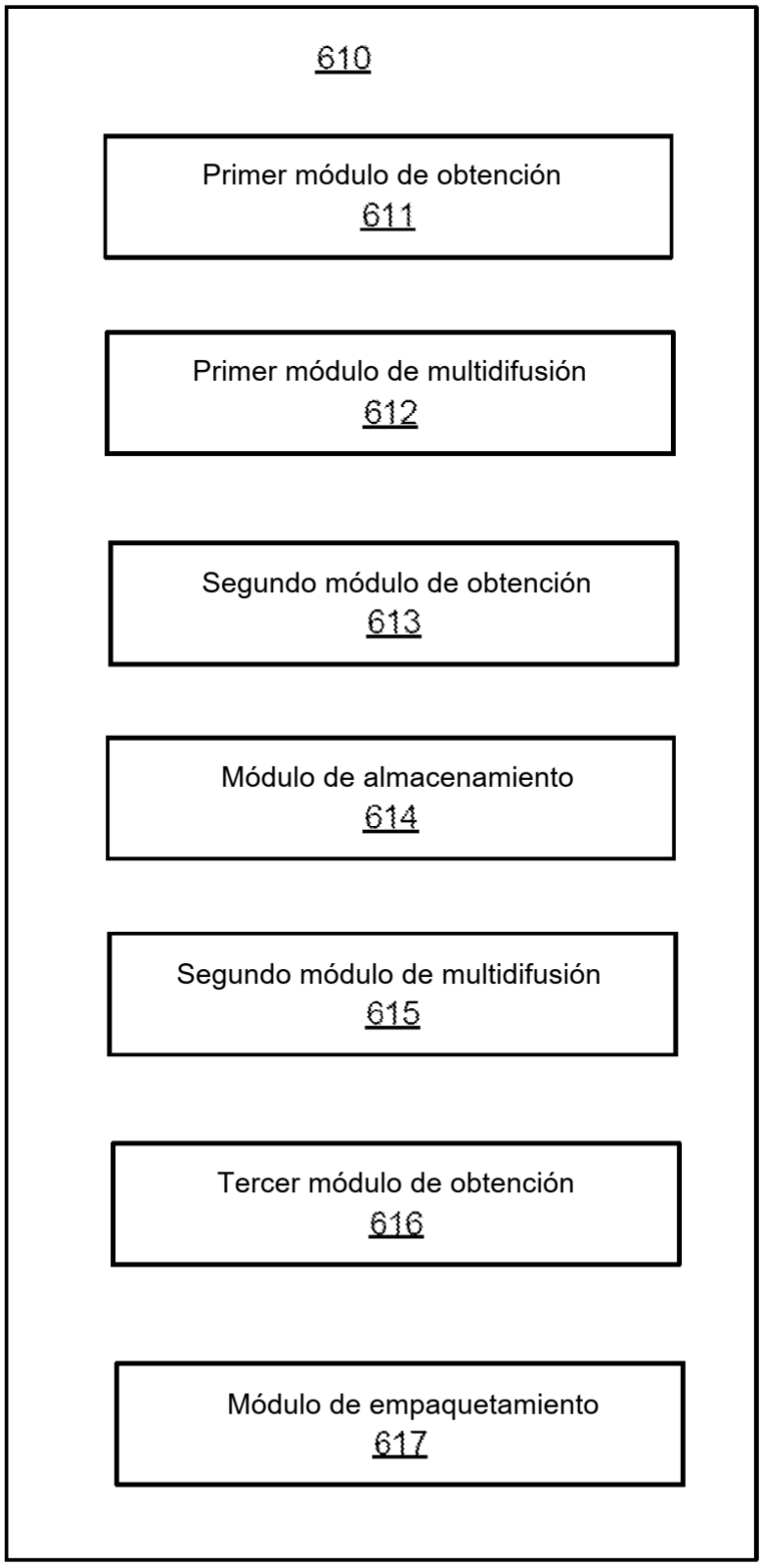
610
Primer módulo de obtención
611
Primer módulo de multidifusión
612
Segundo módulo de obtención
613
Módulo de almacenamiento
614
Segundo módulo de multidifusión
615
Tercer módulo de obtención
616
Módulo de empaquetamiento
617

FIG. 6A

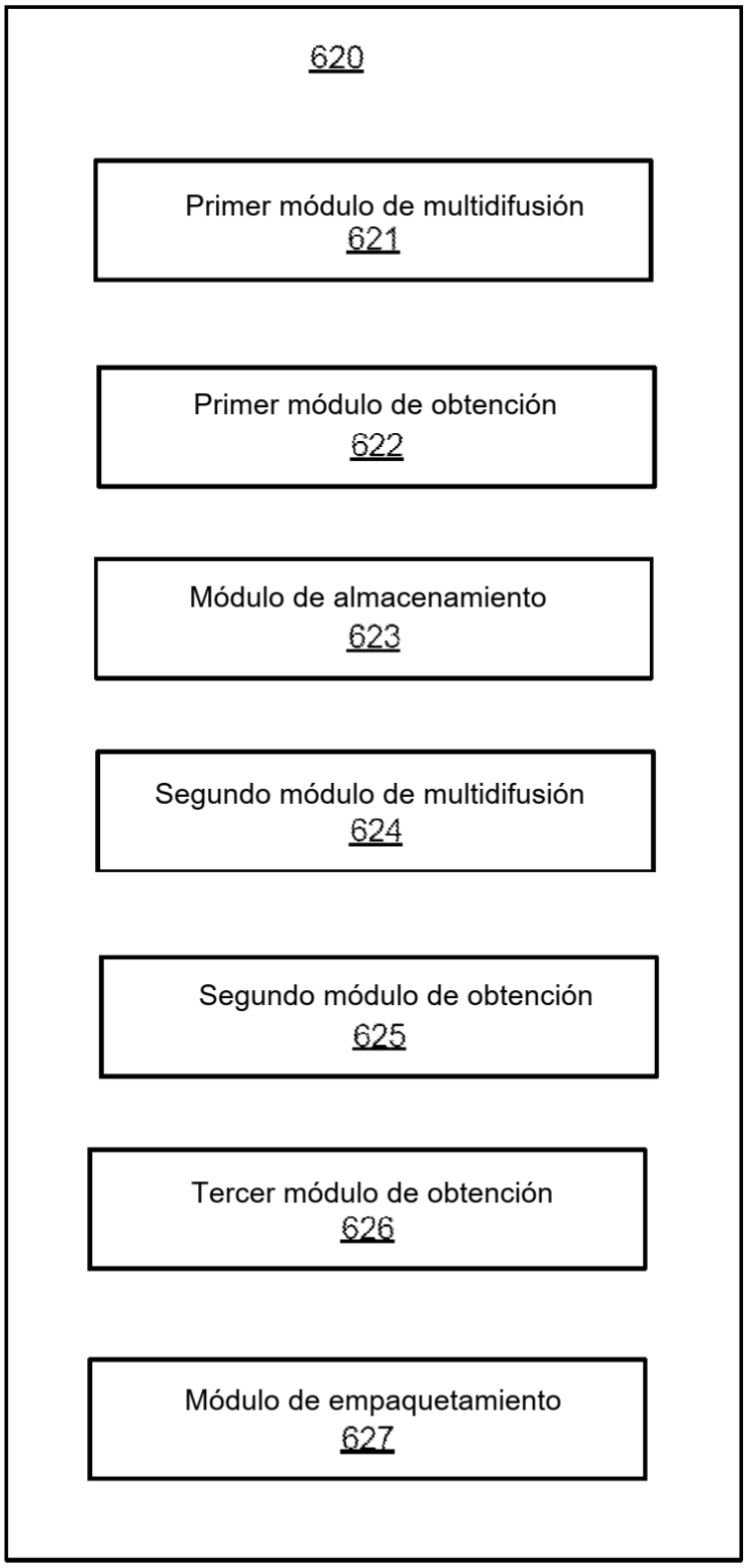
620
Primer módulo de multidifusión
621
Primer módulo de obtención
622
Módulo de almacenamiento
623
Segundo módulo de multidifusión
624
Segundo módulo de obtención
625
Tercer módulo de obtención
626
Módulo de empaquetamiento
627

FIG. 6B

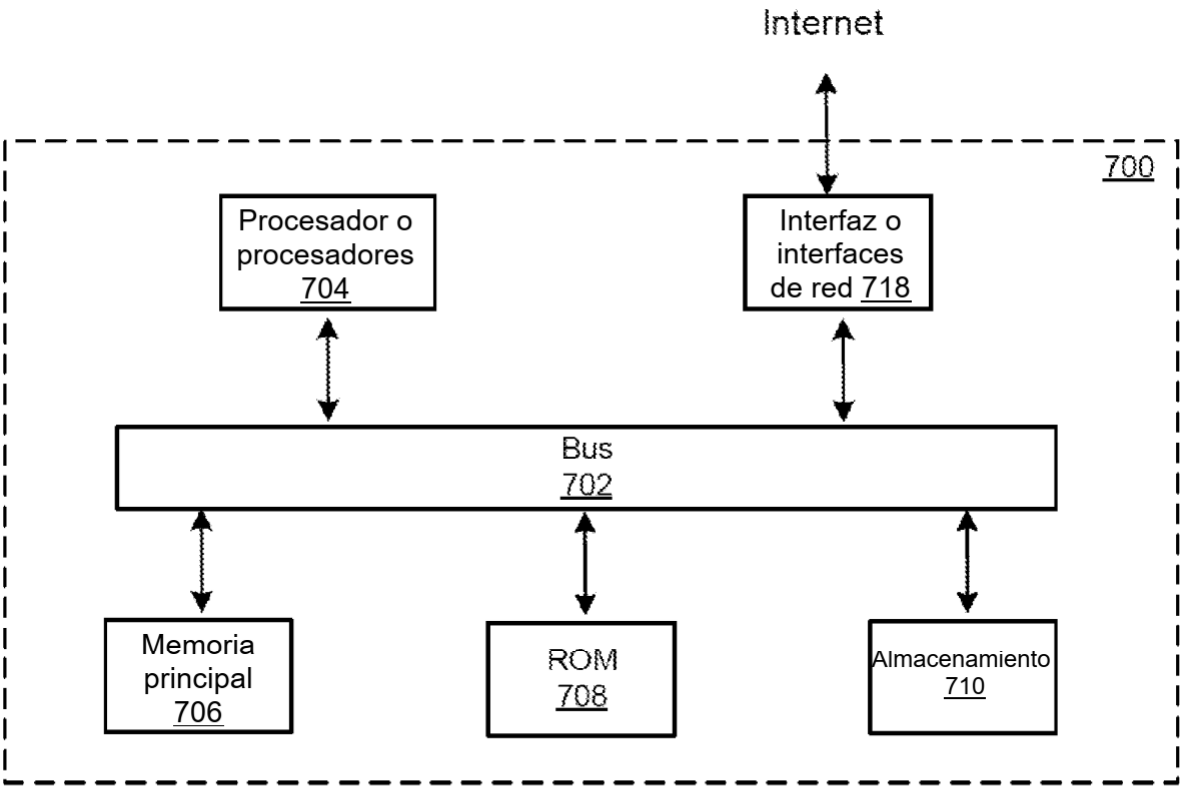
Internet
700
Procesador o procesadores 704
Interfaz o interfaces de red 718
Bus 702
Memoria principal 706
ROM 708
Almacenamiento 710

FIG. 7