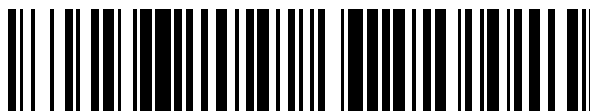


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 626 655**

51 Int. Cl.:

G06F 15/16 (2006.01)

G06F 11/20 (2006.01)

H04L 29/08 (2006.01)

H04L 29/14 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **07.09.2012 PCT/US2012/054038**

87 Fecha y número de publicación internacional: **14.03.2013 WO13036697**

96 Fecha de presentación y número de la solicitud europea: **07.09.2012 E 12829430 (3)**

97 Fecha y número de publicación de la concesión europea: **15.03.2017 EP 2754059**

54 Título: **Conmutación por error de clientes agrupados**

30 Prioridad:

09.09.2011 US 201113228732

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
25.07.2017

73 Titular/es:

**MICROSOFT TECHNOLOGY LICENSING, LLC
(100.0%)**

**One Microsoft Way
Redmond, Washington 98052-6399, US**

72 Inventor/es:

**KRUSE, DAVID, M.;
FATHALLA, DIAA;
PINKERTON, JAMES, T.;
GEORGE, MATHEW;
PRAHALAD, PRASHANTH y
JOLLY, THOMAS, E.**

74 Agente/Representante:

CARPINTERO LÓPEZ, Mario

ES 2 626 655 T3

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín Europeo de Patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre Concesión de Patentes Europeas).

DESCRIPCIÓN

Conmutación por error de clientes agrupados

Antecedentes

Los entornos agrupados, por ejemplo, los entornos donde se distribuye la carga de trabajo a través de múltiples máquinas, se usan habitualmente para proporcionar conmutación por error y alta disponibilidad de información a los clientes. Los entornos agrupados permiten a los clientes acceder a recursos a través del uno o más nodos que son una parte del entorno. Un entorno agrupado puede actuar como un cliente, un servidor, o ambos. En un servidor de grupo de clientes, una aplicación puede residir en cualquiera de los nodos que forman el grupo. La aplicación puede emitir solicitudes de recursos que se almacenan localmente en el grupo de clientes o se almacenan a distancia. Si se produce un error en el nodo, el cliente realiza una conmutación por error, o migra, a un nodo diferente del grupo. Sin embargo, cuando el cliente solicita de nuevo el acceso a un recurso con el que estaba trabajando en el momento del error, el recurso puede estar acotado o bloqueado por el servidor para el nodo cliente anterior en el que residía la aplicación.

Se han hecho realizaciones con respecto a estas y otras consideraciones. Además, aunque se han discutido problemas relativamente específicos, debe entenderse que las realizaciones no deben limitarse a resolver los problemas específicos identificados en los antecedentes.

El documento EP 1 643 406 A2 se refiere a la reutilización del identificador de registro. Se proporciona una función para generar y usar un único identificador como una clave para identificar una única instancia de agente de usuario entre los múltiples agentes de usuario del mismo usuario. La instalación genera un identificador para una primera instancia de agente de usuario, que es una instancia de un usuario en una instancia de aplicación. La función usa el identificador para un registro de la primera instancia de agente de usuario con un servidor lógico y asocia el identificador con la primera instancia de agente de usuario y el par de registro de servidor lógico. A continuación, la función usa el identificador para un registro posterior de la primera instancia de agente de usuario con el servidor lógico.

El documento US 7 664 991 B1 se refiere a un sistema y un procedimiento para la recuperación de E/S de sistemas de archivos distribuidos. Las realizaciones pueden detectar una pérdida de acceso a un servidor en la red de almacenamiento y recuperar las solicitudes de E/S de aplicación en tiempo real una vez que se restablece el acceso al servidor. Las realizaciones pueden detectar fallos en el servidor y/o en la red y almacenar solicitudes de E/S fallidas y nuevas. Puede detectarse la recuperación del fallo (por ejemplo, reconexión de la red, reinicio del nodo servidor o conmutación por error, si se trata de un entorno agrupado) y, después de que se detecte la recuperación, puede enviarse al servidor cualquier solicitud de E/S fallida y nueva almacenada. En una realización, para detectar la recuperación del fallo, una solicitud de E/S fallida puede volver a emitirse repetidamente hasta que la solicitud de E/S tenga éxito. Las realizaciones pueden implementarse en una diversidad de entornos de almacenamiento, incluyendo entornos donde los clientes emiten E/S directas a través de una red de almacenamiento para almacenar y controlar las E/S a través de una red a un servidor.

Sumario

El presente sumario se proporciona para presentar una selección de conceptos en una forma simplificada que se describen adicionalmente a continuación en la sección Descripción Detallada. El presente sumario no pretende identificar las características clave o las características esenciales del objeto reivindicado, ni pretende usarse como ayuda para determinar el ámbito del objeto reivindicado.

En el presente documento se desvelan sistemas y procedimientos que proporcionan una aplicación o un procedimiento con acceso continuo a un recurso después de que la aplicación migre a un nuevo nodo en un entorno de cliente agrupado. Una aplicación o un procedimiento residente en un nodo en un grupo de clientes envía una solicitud a un servidor para acceder a un recurso. En las realizaciones, se usa un único identificador de instancia de aplicación para identificar una aplicación que solicita un recurso. El único identificador de aplicación puede estar provisto de la solicitud. Cuando el cliente accede a un recurso, el identificador de instancia de aplicación se asocia con el recurso solicitado.

Antes de que la aplicación o el procedimiento complete sus operaciones en el recurso, el nodo en el que reside el cliente en el entorno agrupado puede experimentar un error que hace que falle o se pierda el acceso de otro modo al recurso antes de que la aplicación libere correctamente el recurso. En tales circunstancias, el recurso puede permanecer en un estado acotado o bloqueado en el servidor por la solicitud del cliente anterior. Tras fallar en un nodo diferente en el grupo de clientes, la aplicación en el nuevo nodo cliente puede restablecer una conexión con el servidor que gestiona el recurso y realizar una segunda solicitud del recurso al que la aplicación tenía acceso anteriormente en el momento del error. La segunda solicitud puede incluir el identificador de instancia de aplicación que se envió con la primera solicitud. Aunque la segunda solicitud del recurso puede recibirse de un nodo diferente en el entorno agrupado, el identificador de instancia de aplicación permite que el servidor que gestiona la solicitud determine que la segunda solicitud pertenece a la misma aplicación o procedimiento que había bloqueado previamente el recurso. Si lo hace, el servidor puede invalidar el recurso y conceder la segunda solicitud del cliente

para acceder al recurso mientras garantiza que no se produzca una situación de conflicto.

Las realizaciones pueden implementarse como un procedimiento informático, un sistema informático o como un artículo de fabricación tal como un producto de programa informático o un medio legible por ordenador. El producto de programa informático puede ser un medio de almacenamiento informático legible por un sistema informático y que codifica un programa informático de instrucciones para ejecutar un procedimiento informático. El producto de programa informático también puede ser una señal propagada en un soporte legible por un sistema informático y que codifica un programa informático de instrucciones para ejecutar un procedimiento informático.

Breve descripción de los dibujos

Se describen realizaciones no limitantes y no exhaustivas con referencia a las siguientes figuras.

- La figura 1 ilustra un sistema que puede usarse para implementar las realizaciones descritas en el presente documento.
- La figura 2 es un diagrama de bloques que ilustra un entorno de software que puede usarse para implementar las realizaciones desveladas en el presente documento.
- La figura 3 es una realización de un procedimiento que un cliente puede realizar para obtener acceso continuo a un recurso en un entorno agrupado.
- La figura 4 es una realización de un procedimiento realizado por un nodo en un entorno agrupado para proporcionar acceso continuo a un recurso.
- La figura 5 ilustra un diagrama de bloques de un entorno informático adecuado para implementar las realizaciones.

Descripción detallada

A continuación, se describen más detalladamente diversas realizaciones con referencia a los dibujos adjuntos, que forman parte de la misma, y que muestran realizaciones específicas a modo de ejemplo. Sin embargo, las realizaciones pueden implementarse en muchas formas diferentes y no deben interpretarse como limitadas a las realizaciones expuestas en el presente documento; por el contrario, estas realizaciones se proporcionan de manera que la presente divulgación será minuciosa y completa y transmitirá con todo detalle el ámbito de las realizaciones a los expertos en la materia. Las realizaciones pueden ponerse en práctica como procedimientos, sistemas o dispositivos. En consecuencia, las realizaciones pueden adoptar la forma de una implementación de hardware, una implementación completamente de software o una implementación que combina aspectos de software y hardware. Por lo tanto, la siguiente descripción detallada no debe tomarse en un sentido limitante.

Las realizaciones de la presente divulgación están relacionadas con la provisión de mecanismos de conmutación por error de cliente agrupado que permiten a un solicitante volver a obtener acceso a un recurso después de un suceso de conmutación por error. En las realizaciones, un solicitante puede ser un procedimiento, una aplicación o uno o más procedimientos secundarios de una aplicación. Un recurso puede ser un archivo, un objeto, datos o cualquier otro tipo de recurso en un entorno informático. En las realizaciones, un recurso puede residir en un servidor independiente o puede residir en un entorno agrupado. En las realizaciones desveladas en el presente documento, un entorno agrupado puede incluir uno o más nodos (por ejemplo, dispositivos de cliente y/o de servidor).

En una realización a modo de ejemplo, una aplicación que reside en un nodo en un entorno agrupado puede solicitar acceso a un recurso específico. En las realizaciones, el recurso puede almacenarse localmente (por ejemplo, en el nodo cliente), en un dispositivo remoto (por ejemplo, un servidor remoto o un nodo diferente en el entorno agrupado de cliente), o en un entorno agrupado (por ejemplo, un entorno que contiene múltiples nodos) que es diferente del entorno agrupado de cliente. Por ejemplo, en las realizaciones el entorno agrupado puede ser un grupo de clientes o de servidores; sin embargo, los expertos en la materia apreciarán que los sistemas y los procedimientos desvelados en el presente documento pueden emplearse en cualquier otro tipo de entorno, tal como, pero sin limitarse a, una red virtual.

En dichos entornos, los recursos pueden compartirse entre clientes y aplicaciones. Cuando una aplicación accede a un recurso, el recurso puede estar acotado o bloqueado, prohibiendo de este modo que otras aplicaciones accedan al recurso hasta que la aplicación de acceso libere el recurso. El acotamiento o bloqueo del recurso puede emplearse para proteger contra un conflicto, es decir, proteger contra la modificación del recurso por otra aplicación antes de que la aplicación de acceso haya realizado sus operaciones en el recurso. Sin embargo, si falla el nodo en un entorno de cliente agrupado, la aplicación que accede al recurso puede no liberar correctamente el recurso de un estado acotado o bloqueado. Por ejemplo, el nodo cliente que accede al recurso en nombre de la aplicación puede perder una conexión de red, puede colgarse, o puede perder de otro modo el acceso al recurso antes de que la aplicación complete sus operaciones y libere adecuadamente el recurso. Por lo tanto, el recurso puede permanecer en un estado en el que no está disponible para otros clientes o aplicaciones. Pueden emplearse mecanismos que liberen automáticamente un recurso de un estado acotado o bloqueado, evitando de este modo que el recurso se bloquee permanentemente. Sin embargo, estos mecanismos a menudo esperan un período de tiempo antes de liberar un recurso acotado o bloqueado.

En algunos casos, cuando la aplicación realiza una conmutación por error para migrar desde el nodo cliente fallido a un nodo cliente diferente en el grupo de clientes, la aplicación puede intentar restablecer su conexión anterior con el servidor y reanudar su o sus operaciones en el recurso a través del nodo cliente diferente. Sin embargo, debido a que el recurso no se ha liberado correctamente por el nodo cliente fallido, que ha accedido previamente al recurso en nombre de la aplicación, debido al error, la aplicación que tenía acceso previo al recurso puede no ser capaz de reanudar su acceso al recurso hasta que el servidor libere el recurso de su estado acotado o bloqueado. Sin embargo, debido a que ahora un nodo diferente está intentando acceder al recurso en nombre de la aplicación, el servidor puede no ser capaz de identificar la aplicación como la misma aplicación que estableció previamente el bloqueo en el recurso. Sin embargo, debido a que la misma aplicación está intentando acceder al recurso, no existe una situación de conflicto. En estas situaciones, esperar que el servidor libere el bloqueo anterior en el recurso puede provocar un retraso inaceptable para la aplicación.

Tal como se ha descrito, debido a que la aplicación está operando en un entorno de cliente agrupado, cuando la aplicación solicita acceder al recurso una segunda vez, la solicitud de acceso al recurso puede realizarse desde una localización diferente, tal como un nodo diferente en el entorno de cliente agrupado. Por lo tanto, la segunda solicitud puede proceder de una localización o dirección IP diferente. Debido a que la solicitud puede realizarse desde una localización diferente, un servidor puede tener dificultades para garantizar que el cliente o la aplicación que intenta volver a acceder al recurso sea, en realidad, el mismo cliente que accedió previamente al recurso. Los sistemas y procedimientos desvelados en el presente documento proporcionan mecanismos para identificar situaciones donde la misma aplicación está intentando acceder a un recurso, evitando de este modo dicho retraso y proporcionando un acceso continuo de la aplicación al recurso.

La figura 1 ilustra un sistema 100 que puede usarse para implementar algunas de las realizaciones desveladas en el presente documento. El sistema 100 incluye un grupo 102 de clientes y un grupo 106 de servidores. El grupo de clientes incluye múltiples nodos, tales como los clientes 102A y 102B. Los clientes 102A y 102B pueden ser un dispositivo o aplicación que reside en el grupo 102 de clientes. El grupo 102 de clientes puede comunicarse con el grupo 106 de servidores a través de la red 108. En las realizaciones, la red puede ser Internet, una WAN, una LAN o cualquier otro tipo de red conocida en la técnica. El grupo 106 de servidores almacena los recursos a los que acceden las aplicaciones del grupo 102 de clientes (por ejemplo, las aplicaciones que residen en el cliente 102A o el cliente 102B). En las realizaciones, un cliente (por ejemplo, el cliente 102A) puede establecer una sesión con el grupo 106 para acceder a los recursos del grupo 106 en nombre de una aplicación que reside en el cliente. Aunque en la figura 1, el grupo 102 de clientes solo incluye dos clientes (por ejemplo, el cliente 102A y el cliente 102B), los expertos en la materia apreciarán que puede incluirse cualquier número de clientes en el grupo 102 de clientes.

Como se muestra en la figura 1, el grupo 106 de servidores incluye los servidores 106A, 106B y 106C, que proporcionan una alta disponibilidad y redundancia para la información almacenada en el grupo 106. En las realizaciones, el grupo 106 puede tener un sistema de archivos, una base de datos u otra información a la que acceden los clientes 102 y 104. Aunque en la figura 1 se muestran tres servidores, en otras realizaciones el grupo 106 puede incluir más de tres servidores, o menos de tres servidores. Además, aunque las realizaciones descritas en el presente documento se refieren a un cliente que se comunica con un servidor que es parte de un grupo de servidores, los expertos en la materia apreciarán que las realizaciones desveladas en el presente documento también pueden realizarse usando un servidor independiente.

En las realizaciones, el grupo 102 de clientes proporciona mecanismos de conmutación por error que permiten a un cliente migrar desde un primer nodo cliente a un segundo nodo cliente en caso de que se produzca un error o un fallo en el primer nodo cliente. Los expertos en la materia apreciarán que puede emplearse cualquier tipo de mecanismo de conmutación por error con los sistemas y procedimientos desvelados en el presente documento. Los procedimientos y sistemas desvelados en el presente documento pueden emplearse para evitar un retraso indebido cuando una aplicación intenta volver a obtener el acceso a un recurso que migra de un cliente a otro (por ejemplo, del cliente 102A al cliente 102B) en el caso de una conmutación por error. En las realizaciones, un identificador de instancia de aplicación que identifica la aplicación que accede al recurso puede asociarse con el recurso. El identificador de instancia de aplicación puede ser un identificador único global (GUID) que está asociado con una aplicación, una acción realizada por una aplicación o un procedimiento secundario de una aplicación. Por ejemplo, en una realización, una aplicación puede asociarse con un identificador de instancia de aplicación que es un GUID. En otra realización, un identificador de instancia de aplicación puede asociarse con una operación o acción específica realizada por una aplicación. Por ejemplo, si la aplicación emite dos solicitudes abiertas diferentes para dos archivos diferentes, cada solicitud abierta puede tener su propio identificador de instancia de aplicación. En otra realización más, un identificador de instancia de aplicación puede asociarse con uno o más procedimientos secundarios de la aplicación. Como será evidente para los expertos en la materia a partir de las realizaciones descritas en el presente documento, la asociación del identificador de instancia de aplicación de una aplicación con sus uno o más procedimientos secundarios permitirá que los procedimientos secundarios accedan al recurso si el recurso se coloca en un estado bloqueado o acotado que pertenece a la aplicación. En las realizaciones, el identificador de instancia de aplicación puede enviarse por el cliente en el momento de, o después de, enviar una solicitud para un recurso.

De acuerdo con otra realización, además de almacenar la información a la que acceden los clientes que forman parte del grupo 102 de clientes, el grupo 106 de servidores también proporciona un mecanismo de conmutación por

error que permite el acceso continuo de un recurso en caso de fallo de un nodo servidor. De nuevo, los expertos en la materia apreciarán que puede emplearse cualquier tipo de mecanismo de conmutación por error con los sistemas y procedimientos desvelados en el presente documento.

En las realizaciones, cuando un cliente solicita acceso a un recurso en nombre de una aplicación, el identificador de instancia de aplicación de la aplicación se envía con la solicitud. El servidor que recibe la solicitud puede asociar el identificador de instancia de aplicación con el recurso. Por ejemplo, el grupo de servidores puede almacenar el identificador de instancia de aplicación en una tabla o caché localizada en uno o más nodos (por ejemplo, servidores tales como los servidores 106A, 106B y/o 106C) localizados en el grupo 106 de servidores de tal manera que el identificador de instancia de aplicación se asocia con el recurso. Antes de que el cliente se haga con el recurso, el cliente puede experimentar un error que lo obligará a perder la conexión con el recurso. Por ejemplo, el cliente que aloja la aplicación o que realiza solicitudes u operaciones en nombre de la aplicación puede perder su conexión de red al grupo de servidores, el cliente puede colgarse o puede producirse cualquier otro tipo de error que interfiera con el uso de las aplicaciones del recurso. Tras experimentar el error, la aplicación puede conmutar por error a un nuevo nodo cliente en el grupo 102 de clientes. El nuevo nodo cliente puede volver a conectarse al grupo de servidores y enviar una segunda solicitud para acceder al recurso en nombre de la aplicación. En las realizaciones, el cliente puede volver a conectar con el mismo nodo en el grupo 106 de servidores o un nodo diferente. La segunda solicitud para acceder al recurso puede incluir el identificador de instancia de aplicación de la aplicación. Tras recibir la segunda solicitud, el servidor (por ejemplo, un servidor 106A del grupo 106 de servidores) compara el identificador de instancia de aplicación de la segunda solicitud con el identificador de instancia de aplicación asociado con el recurso. Si coinciden los dos identificadores de instancia de aplicación, el grupo de servidores invalida el recurso. En las realizaciones, la invalidación del recurso puede comprender cerrar un archivo, eliminar un bloqueo del recurso o realizar de otro modo cualquier acción que libere el recurso para su uso. A continuación, el nodo servidor puede conceder la segunda solicitud de la aplicación para acceder al recurso. Si el identificador de instancia de aplicación del segundo nodo no coincide con el identificador de instancia de aplicación asociado con el recurso, el servidor no permitirá el acceso al recurso hasta que se libere el recurso.

Para ilustrar una realización, un solicitante (por ejemplo, un procedimiento, una aplicación, etc.) en el cliente 102A en el grupo 102 de clientes puede solicitar que el cliente 102A establezca una sesión con un servidor del grupo 106 de servidores. Por ejemplo, el cliente 102A puede establecer una sesión con el servidor 106A para acceder a una base de datos almacenada que está en el servidor 106A o que forma parte del grupo 106 de servidores, en el que el servidor 106A puede acceder a la base de datos. A continuación, el cliente 102A envía una solicitud de un recurso en nombre del solicitante. Un identificador de instancia de aplicación que identifica al solicitante está asociado con la solicitud. En las realizaciones, la solicitud puede incluir el identificador de instancia de aplicación o el identificador de instancia de aplicación puede enviarse por separado de una manera tal que el servidor 106A pueda determinar que el identificador de instancia de aplicación se asocia con la solicitud. En otra realización más, el servidor 106A o el grupo 106A de servidores puede tener ya la información necesaria para asociar el identificador de instancia de aplicación con la solicitud sin tener que recibir el identificador de instancia de aplicación junto con la solicitud. A continuación, el servidor 106A concede al solicitante acceso al recurso, permitiendo de este modo al solicitante realizar operaciones sobre o acceder de otro modo al recurso. Cuando se concede al solicitante acceso al recurso, el servidor 106A asocia un identificador de instancia de aplicación con el recurso de una manera tal que indica que el solicitante está actualmente accediendo al recurso. A continuación, el recurso puede acotarse o bloquearse para que otros clientes o aplicaciones no puedan acceder o modificar el recurso hasta que el cliente 102 haya completado su operación.

Antes de que el solicitante complete sus operaciones en el recurso, se produce un error que hace que el cliente 102A falle o pierda de otro modo su conexión al recurso. Debido a que el solicitante del cliente ha completado su operación, no ha liberado el control del recurso. Por lo tanto, el recurso puede permanecer en un estado acotado o bloqueado. El grupo 102 de solicitantes o clientes puede emplear un mecanismo de conmutación por error para hacer migrar el solicitante del cliente 102A al cliente 102B. Una vez completada la operación de conmutación por error, el cliente 102B puede volver a conectarse al grupo 106 de servidores en nombre del solicitante. El cliente 102B puede volver a conectarse al servidor 106A o establecer una nueva conexión con cualquier otro servidor en el grupo 106 de servidores (por ejemplo, el servidor 106B o 106C). En una situación a modo de ejemplo, el cliente 102B vuelve a conectarse al servidor 106A. Tras volver a conectarse, el cliente 102B puede enviar una segunda solicitud para acceder al recurso en nombre del solicitante. Como se ha indicado anteriormente, debido a que el solicitante no ha liberado el control del recurso, el recurso aún puede estar en un estado bloqueado o acotado. Con el fin de acceder al recurso sin esperar que el servidor cambie automáticamente el estado del recurso, por ejemplo, a través de una operación de tiempo de espera, el solicitante puede proporcionar de nuevo su identificador de instancia de aplicación con la segunda solicitud. El servidor 106A compara el identificador de instancia de aplicación provisto de la segunda solicitud al identificador de instancia de aplicación asociado con el recurso. Por ejemplo, comparando el identificador de instancia de aplicación recibido o asociado de otro modo con la segunda solicitud a un identificador de instancia de aplicación que el servidor 106A ha asociado con el recurso. El identificador de instancia de aplicación asociado puede almacenarse en una caché o tabla local del servidor 106A, o puede almacenarse en otro lugar en el grupo 106 de servidores. Si el identificador de instancia de aplicación almacenado en la caché coincide con el identificador de instancia de aplicación que está asociado con el recurso, el servidor 106A invalida o libera de otro modo el recurso y permite al cliente 102B acceder de nuevo al recurso en nombre del

solicitante sin esperar a que se libere el recurso por algún otro mecanismo (por ejemplo, por el tiempo de espera de estado acotado o bloqueado). Si los identificadores de instancia de aplicación no coinciden, el cliente 102B tendrá que esperar a que el recurso se libere antes de acceder al mismo.

Aunque en el ejemplo anterior, el cliente 102B se ha vuelto a conectar al mismo servidor 106A, también es posible, en otras realizaciones, que el cliente se conecte a otro nodo en el grupo 106 de servidores. Por ejemplo, el cliente 102B puede volver a conectarse al servidor 106B y presentar una segunda solicitud para volver a obtener el acceso al recurso en nombre del solicitante. La segunda solicitud puede asociarse de nuevo con el identificador de instancia de aplicación del solicitante, por ejemplo, al incluirse en la segunda solicitud o asociarse de otro modo con la segunda solicitud. En este ejemplo, el servidor 106B puede no tener el identificador de instancia de aplicación asociado con el recurso almacenado en su caché local debido a que el acceso original del recurso estaba en el servidor 106A. En tal situación, el servidor 106B puede contactar con los otros servidores del grupo 106 de servidores para determinar si tienen un identificador de aplicación asociado con el recurso. Si el identificador de aplicación asociado con el recurso se almacena en un nodo diferente del grupo de servidores (por ejemplo, el servidor 106A), el identificador de instancia de aplicación en el otro nodo del grupo de servidores se compara con el identificador de instancia de aplicación provisto de la segunda solicitud. Si coinciden, el servidor 106B puede enviar una solicitud al servidor 106A para invalidar el recurso y, a continuación, el servidor 106B puede permitir al solicitante (ahora en el cliente 102B) acceder al recurso. Si los identificadores de instancia de aplicación no coinciden, el cliente 102B tendrá que esperar a que se libere el recurso.

Basándose en los ejemplos anteriores, los expertos en la materia apreciarán que cualquier nodo cliente en el grupo 102 de clientes puede solicitar acceso y, a continuación, proporcionar acceso, a un solicitante en el grupo 102 de clientes. Además, cualquier nodo servidor en un grupo de servidores (por ejemplo, cualquier servidor en el grupo 106 de servidores) es capaz de determinar si el solicitante tenía acceso previamente al recurso incluso si el acceso se produjo en un nodo servidor diferente en el grupo de servidores. Los expertos en la materia apreciarán que la siguiente descripción es simplemente un ejemplo de cómo puede operar la realización mostrada en la figura 1 y de que existen otras realizaciones. Por ejemplo, en lugar de acceder a recursos en un servidor remoto o en un grupo de servidores, los nodos cliente pueden realizar las realizaciones descritas en el presente documento para proporcionar a los solicitantes (por ejemplo, aplicaciones o procedimientos) acceso continuo a recursos que residen en el entorno agrupado (por ejemplo, en los mismos o diferentes nodos de grupo de clientes que forman el grupo de clientes). Como se describe con mayor detalle a continuación, las realizaciones descritas en el presente documento pueden implicar diversas etapas u operaciones diferentes. Además, las realizaciones descritas en el presente documento pueden implementarse usando cualquier componente o módulo de software o hardware apropiado.

Volviendo ahora a la figura 2, la figura ilustra un diagrama de bloques de un entorno 200 de software que muestra un grupo 201 de nodos cliente con múltiples nodos cliente (por ejemplo, los clientes 202 y 204) y un grupo 206 de nodos servidor con múltiples nodos servidor (por ejemplo, el nodo 1 208 y el nodo 2 216). En las realizaciones, el cliente 202 solicita el acceso a un recurso, tal como el recurso 226, en un entorno 206 de grupo de servidores en nombre de un solicitante. El grupo 201 de nodos cliente puede ser un grupo de clientes, tal como el grupo 102 de clientes (figura 1). Aunque no se ilustra, el grupo de clientes puede contener más de dos clientes. El grupo 206 de nodos servidor, puede ser un grupo de servidores, tal como el grupo 106 de servidores (figura 1) o puede ser cualquier otro tipo de entorno agrupado, tal como, pero sin limitarse a, una red virtual. El recurso 226 puede almacenarse en un almacén 228 de datos que es parte del entorno agrupado. Aunque no se muestra, en unas realizaciones alternativas, el almacén 228 de datos puede no ser parte del entorno agrupado, pero puede conectarse al entorno agrupado a través de una red. Ejemplos de dicha red incluyen, pero sin limitarse a, Internet, una WAN, una LAN, o cualquier otro tipo de red conocida en la técnica. En otras realizaciones más, el almacén de datos puede ser parte de un nodo (por ejemplo, un dispositivo) que es una parte del grupo 206.

El grupo 206 de nodos servidor puede incluir uno o más nodos, tales como el nodo 1 208 y el nodo 2 216. Aunque en la figura 2 solo se ilustran dos nodos, puede incluirse cualquier número de grupos de nodos en el entorno 206 agrupado. En las realizaciones, los grupos 208 y 216 de nodos son capaces de recibir una solicitud para realizar una operación en y/o conceder el acceso al recurso 226. En las realizaciones, el recurso 226 puede ser un archivo, un objeto, una aplicación, datos o cualquier otro tipo de recurso almacenado en o accesible para un nodo en el grupo 206 de nodos o para un servidor independiente.

En las realizaciones, un cliente envía una solicitud 222 inicial al entorno 206 agrupado. Como se ilustra en la figura 2, la solicitud 222 inicial puede enviarse por el cliente 202 y recibirse por el nodo 1 208. Sin embargo, en realizaciones alternativas, la solicitud 222 inicial puede enviarse por el cliente o cualquier otro nodo cliente en el grupo 201 de clientes y recibirse por el nodo 2 216 o cualquier otro nodo en el grupo 206 de servidores. Un ejemplo de solicitudes incluye, pero sin limitarse a, solicitudes para crear, abrir o acceder de otro modo a un archivo. La solicitud 222 puede transmitirse desde el cliente al grupo de nodos a través de una red, tal como, pero sin limitarse a, Internet, una WAN, una LAN o cualquier otro tipo de red conocida en la técnica. La solicitud 222 inicial puede incluir una solicitud de acceso a un recurso, tal como el recurso 226. En las realizaciones, la solicitud 222 también puede incluir un identificador de instancia de aplicación que identifica el solicitante en cuyo nombre está haciendo la solicitud el cliente 202. En las realizaciones, la solicitud 222 inicial puede consistir en uno o más mensajes. Por ejemplo, la solicitud 222 puede ser un único mensaje que contiene tanto la solicitud como un identificador de instancia de aplicación. En otra realización, la solicitud 222 pueden ser múltiples mensajes que incluyen una o más

solicitudes, así como uno o más identificadores de instancia de aplicación. En las realizaciones, el cliente 202 puede incluir una caché 214 de instancia de aplicación que se usa para almacenar y/o generar uno o más identificadores de instancia de aplicación que pueden transmitirse con la solicitud 222.

Como se muestra en la figura 2, el nodo 1 208 puede recibir la solicitud 222 y un identificador de instancia de aplicación del cliente 202. Si el recurso 226 solicitado está disponible, por ejemplo, no acotado o bloqueado por otro cliente o aplicación, el nodo 1 puede conceder a la solicitud del cliente (por ejemplo, el cliente 202) el acceso al recurso 226 en nombre de un solicitante que se está ejecutando en el cliente. Tras conceder acceso al recurso 226, el controlador 210 de filtro puede asignar o crear de otro modo una asociación entre el cliente 202 y el recurso 226 almacenando el identificador de instancia de aplicación que ha recibido del cliente 202. En las realizaciones, la asociación puede almacenarse como un objeto en la caché 212 de instancia de aplicación que es una parte del nodo 1. Aunque la realización ilustrada muestra la caché 212 de instancia de aplicación como una parte del nodo 1 208, en las realizaciones, la caché 212 de instancia de aplicación puede almacenarse en otro lugar como una parte del grupo 206 de nodos. Los expertos en la materia apreciarán que el grupo 206 de nodos puede incluir una o más cachés de instancia de aplicación, tal como la caché 220 de instancia de aplicación en el nodo 2 216. En las realizaciones, cuando hay más de una caché de instancia de aplicación, los datos almacenados en las múltiples cachés de instancia de aplicación pueden replicarse a través de todas las cachés de instancia de aplicación o cada caché de instancia de aplicación puede almacenar datos por separado.

En una realización, el identificador de instancia de aplicación recibido del cliente, un identificador de instancia de aplicación que identifica un solicitante (por ejemplo, una aplicación o un procedimiento), puede almacenarse en una estructura `_NETWORK_APP_INSTANCE_ECP_CONTEXT`. La estructura `_NETWORK_APP_INSTANCE_ECP_CONTEXT` puede definirse de la siguiente manera:

```
typedef struct _NETWORK_APP_INSTANCE_ECP_CONTEXT {
    USHORT Size;
    USHORT Reserved;
    GUID AppInstanceID;
} _NETWORK_APP_INSTANCE_ECP_CONTEXT,
*PNETWORK_APP_INSTANCE_ECP_CONTEXT;
```

En tales realizaciones, el tamaño variable puede almacenar información relacionada con el tamaño de la estructura y el `AppInstanceID` variable puede ser un identificador de instancia de aplicación único para una aplicación cliente de grupo de conmutación por error, tal como un solicitante que ejecuta en el cliente 202. En las realizaciones, la `_NETWORK_APP_INSTANCE_ECP_CONTEXT`, u otro objeto o variable que contiene el identificador de instancia de aplicación del solicitante pueden almacenarse en la caché 214 de identificador único global (GUID). En las realizaciones, la estructura `_NETWORK_APP_INSTANCE_ECP_CONTEXT` puede enviarse desde un cliente a un servidor en asociación con una solicitud de acceso a un recurso (por ejemplo, una solicitud de creación o de apertura). En una realización, el identificador de instancia de aplicación del solicitante puede almacenarse en la caché GUID del nodo cliente que el solicitante está ejecutando en el entorno 201 de cliente agrupado. En otra realización, aunque no se muestra en la figura 2, el grupo 201 de nodo cliente puede tener un depósito central que almacena identificadores de instancia de aplicación. En tal realización, múltiples nodos cliente en el grupo 201 de nodos cliente pueden acceder al depósito centralizado. En otra realización más, los identificadores de instancia de aplicación pueden almacenarse a través de múltiples cachés GUID (por ejemplo, el caché 214 GUID y el caché 216 GUID). En tales realizaciones, el grupo 201 de nodos cliente puede emplear un algoritmo de replicación para garantizar que las múltiples cachés GUID contengan los mismos identificadores de instancia de aplicación.

Como se ha descrito anteriormente, el identificador de instancia de aplicación puede asociarse con el recurso 226 mientras que el cliente 202 accede al recurso 226 en nombre de un solicitante. Un nodo 206 servidor puede almacenar dicha asociación en una o más cachés de instancia de aplicación que forman parte del grupo 206 de nodos servidor, tales como las cachés 212 y 220 de instancia de aplicación. En una realización, el identificador de instancia de aplicación puede asociarse con el recurso añadiéndolo a una lista de parámetros extra creados (ECP) del recurso 226. La lista de ECP puede almacenarse en una caché de instancia de aplicación que forma parte del grupo 206 de nodos servidor, tales como las cachés 212 y 220 de instancia de aplicación. En las realizaciones, cuando un ECP se recibe por un servidor, el servidor extrae un identificador de instancia de aplicación del ECP y lo añade a una caché para asociarse con un recurso, un manipulador de recursos, etc. Como se ha descrito con respecto al almacenamiento de identificadores de instancia de aplicación en el grupo 201 de clientes, los identificadores de instancia de aplicación asociados con un nodo pueden almacenarse en una caché de instancia de aplicación individual en un nodo en el grupo 206 de nodos servidor, en un depósito central en el grupo 206 de servidores o replicarse a través de múltiples cachés de instancia de aplicación en múltiples nodos en el grupo 206 de nodos servidor.

En las realizaciones, el recurso 226 está acotado o bloqueado mientras que un solicitante que ejecuta en el cliente 202 tiene acceso al recurso 226, impidiendo de este modo que otros clientes o aplicaciones accedan al recurso 226 y evitando cualquier conflicto potencial. En las realizaciones, antes de que el solicitante complete su operación en el recurso 226, el cliente 202 experimenta un error que hace que pierda la conexión con el recurso. Por ejemplo, el cliente puede colgarse, desconectarse o perder su conexión de red con el nodo 208 servidor. En tales casos, el

recurso 226 puede estar todavía en un estado acotado o bloqueado debido a que el solicitante no ha liberado un bloqueo en el recurso, evitando de este modo que otros clientes accedan al recurso 226.

Cuando el error se produce en el cliente 202, el solicitante puede utilizar un mecanismo 232 de conmutación por error de cliente para migrar a un nuevo nodo cliente (por ejemplo, el cliente 204) en el grupo 201 de clientes. Los expertos en la materia apreciarán que puede emplearse cualquier tipo de mecanismo de conmutación por error en la conmutación 232 por error de cliente. En las realizaciones, el mecanismo 232 de conmutación por error también puede incluir la migración del identificador de instancia de aplicación del solicitante que puede haberse almacenado en la caché 214 GUID en el cliente 202 fallido. Una vez completada la migración, el solicitante puede intentar volver a obtener el acceso al recurso 226. En las realizaciones, el cliente 204 puede enviar una segunda solicitud 224 al nodo 1 para solicitar acceso al recurso 226 en nombre del solicitante. Sin embargo, sin las realizaciones de acceso continuo desveladas en el presente documento, cuando el nodo 1 208 recibe una solicitud para acceder al recurso 226 en nombre del cliente 204 (el remitente de la segunda solicitud 224), puede denegar la solicitud debido a que el recurso 226 está todavía en un estado acotado o bloqueado desde el acceso previo que el cliente 202 hizo en nombre del recurso. Sin las realizaciones desveladas en el presente documento, el nodo 1 208 reconocería que la segunda solicitud de acceso al recurso 226 fue desde una localización diferente (por ejemplo, el cliente 204). El nodo 1 208 no podría determinar que la solicitud es para el mismo solicitante que mantiene el bloqueo del recurso 226 y, por lo tanto, determinaría que la concesión de la solicitud daría como resultado un conflicto. Sin embargo, si el mismo solicitante está intentando acceder al recurso 226, no hay ningún motivo de conflicto y obligar al cliente a esperar que el recurso se libere por el sistema puede dar como resultado retrasos indebidos.

El identificador de instancia de aplicación puede usarse para resolver este problema. En las realizaciones, la segunda solicitud 224 también puede incluir que el identificador de instancia de aplicación identifique que el solicitante ha migrado al cliente 204 durante la conmutación por error mostrada en 232. En las realizaciones, el identificador de instancia de aplicación del solicitante puede estar presente en la caché 228 GUID del cliente 204 antes de la migración del solicitante durante la conmutación 232 por error de cliente. Por ejemplo, puede haberse empleado un mecanismo de replicación para replicar el identificador de instancia de aplicación del solicitante a través de los nodos en el grupo 201 de clientes. En otra realización, el solicitante 203 puede almacenar su identificador de instancia de aplicación. En otra realización más, puede hacerse migrar el identificador de instancia de aplicación del solicitante 203 durante la conmutación 232 por error de cliente.

Tal como se ha descrito con respecto a la solicitud 222, el identificador de instancia de aplicación puede transmitirse en el mismo mensaje que la segunda solicitud 224 o la segunda solicitud 224 puede estar compuesta por un número de mensajes diferentes. Cuando la segunda solicitud se recibe en el grupo 206 de nodos, o un nodo individual en el grupo, tal como el nodo 1 208, y el servidor de recepción determina que el recurso está acotado o bloqueado, se determina si el identificador de instancia de aplicación en la segunda solicitud 224 es el mismo que el identificador de instancia de aplicación asociado con el recurso 226. En las realizaciones, el nodo 2 216 comparará el identificador de instancia de aplicación recibido con la segunda solicitud 222 con el identificador de instancia de aplicación que está asociado con el recurso 226. El identificador de instancia de aplicación asociado con el recurso 226 puede almacenarse en la caché 212 de instancia de aplicación del nodo 1 212. En las realizaciones donde existen múltiples cachés de instancia de aplicación en el grupo 206 de nodos, la determinación puede comprobar más de una caché de instancia de aplicación en el grupo 206 de nodos. En tales realizaciones, si un identificador de instancia de aplicación coincidente no está localizado en la caché 212 de instancia de aplicación, el nodo 1 216 puede enviar una solicitud al nodo 2 212 para determinar si un identificador de instancia de aplicación coincidente está localizado en la caché 220 de instancia de aplicación.

En una realización, si el identificador de instancia de aplicación recibido en la segunda solicitud 224 no coincide con el identificador de instancia de aplicación asociado con el recurso 226 (que puede almacenarse en la caché 212 y/o 220 de instancia de aplicación), la segunda solicitud 224 puede no concederse hasta que el recurso 226 esté libre. Sin embargo, si se encuentra una coincidencia, el servidor de recepción (por ejemplo, el nodo 1 208) y/o el grupo 206 de nodos servidor realizan acciones para conceder acceso al recurso 226 sin provocar un retraso indebido al cliente 204 y al solicitante 203. En tales casos, el grupo 206 de nodos puede invalidar el recurso 226, sacando de este modo el recurso 226 de un estado acotado o bloqueado. En las realizaciones, invalidar un acceso previo puede comprender cualquier acción que saque un recurso fuera de un estado acotado o bloqueado. Un ejemplo no limitante es cerrar un archivo abierto (por ejemplo, si el recurso 226 es un archivo). Una vez que el acceso anterior está invalidado, puede concederse la segunda solicitud 224 para acceder al recurso 226, proporcionando de este modo un acceso continuo al solicitante 203.

En una realización, el nodo que recibe la segunda solicitud 224, tal como el nodo 1 208 en la figura 2, puede realizar las acciones necesarias para invalidar el acceso previo del recurso 226 si un nodo diferente (por ejemplo, el nodo 2 216) tiene acceso y/o permiso para invalidar el acceso previo. Sin embargo, en algunos casos, el nodo que recibe la solicitud puede no tener acceso o permiso para invalidar el acceso previo. Por ejemplo, tal instancia puede producirse si la solicitud 222 original se hizo al nodo 2 216, en cuyo caso, el nodo 2 216 puede tener control sobre el recurso. En tales casos, el nodo que recibe la segunda solicitud 224 puede enviar una solicitud al nodo de control para invalidar el acceso previo. Una vez que el nodo de control ha invalidado el acceso previo, el nodo que recibe la segunda solicitud 224 puede conceder la segunda solicitud 224. En otras realizaciones, el nodo que recibe la segunda solicitud 224 puede enviar una solicitud a un nodo diferente para conceder al cliente 204 y/o al solicitante

203 (ahora residente en el cliente 204) el acceso al recurso 226.

El procedimiento descrito evita un retraso indebido en la concesión de una segunda solicitud 224 para acceder a un recurso 226 de un solicitante 203 que ha accedido previamente y aún mantiene un bloqueo en el recurso 226 a través del uso de identificadores de instancia de aplicación. Además, los identificadores de instancia de aplicación ofrecen la ventaja de garantizar que cualquier solicitud concedida no cree un conflicto en el recurso 226. Por ejemplo, si la solicitud se ha recibido de una aplicación diferente, la solicitud incluirá un identificador de instancia de aplicación que es diferente del identificador de instancia de aplicación asociado con el recurso que daría como resultado que se rechazara la solicitud. Debido a que los identificadores de instancia de aplicación son identificadores globalmente únicos, el identificador de instancia de aplicación para diferentes aplicaciones no será el mismo.

La figura 3 es una realización de un procedimiento 300 que un solicitante puede emplear para obtener acceso continuo a un recurso en un entorno agrupado de cliente. Por ejemplo, un solicitante puede ser un cliente, tal como el cliente 202 (figura 2), que emplea el procedimiento 300 para acceder a un recurso (por ejemplo, el recurso 226). En las realizaciones, el recurso puede residir en una máquina remota, tal como un servidor. El servidor puede ser un servidor independiente o parte de un entorno agrupado, tal como el grupo 206 de servidores (figura 2). El flujo comienza en la operación 302 donde se envía una solicitud de un recurso a un servidor. En las realizaciones, la solicitud puede ser acceder a un recurso. En las realizaciones, el acceso a un recurso puede comprender abrir un archivo, crear un archivo, o acceder o realizar de otro modo una operación en un recurso que puede ser remoto para un cliente. En las realizaciones, un solicitante puede operar en un entorno agrupado de cliente. En tales realizaciones, la solicitud enviada en la operación 302 puede enviarse desde un primer cliente en el entorno agrupado de cliente.

El flujo continúa hacia la operación 304 donde se envía un identificador de instancia de aplicación, por ejemplo, a un servidor (por ejemplo, un servidor independiente o un nodo en un entorno agrupado). En una realización, el primer cliente que envió la solicitud también puede enviar el identificador de instancia de aplicación en nombre del solicitante. Como se ha descrito anteriormente, un identificador de instancia de aplicación es un GUID que identifica al solicitante (por ejemplo, una aplicación, un cliente o un procedimiento secundario de una aplicación que solicita acceso a un recurso). En una realización, el identificador de instancia de aplicación puede enviarse en un mensaje transmitido a través de una red. El identificador de instancia de aplicación puede transmitirse en el mismo mensaje que contiene la solicitud en la operación 302 o puede transmitirse en un mensaje diferente. En dichas realizaciones, un objeto que contiene el identificador de instancia de aplicación, tal como, pero sin limitarse a, la `_NETWORK_APP_INSTANCE_ECP_CONTEXT` descrita con respecto a la figura 2, puede enviarse en la operación 302.

En una realización, puede usarse una interfaz para enviar el identificador de instancia de aplicación en la operación 304. La interfaz puede ser una interfaz de nivel de núcleo localizada en un cliente o disponible para un cliente que opera en un entorno agrupado de cliente. En las realizaciones, la interfaz de nivel de núcleo puede usarse por el solicitante y/o el cliente para enviar un identificador de instancia de aplicación a un servidor. El siguiente es un ejemplo no limitante de una interfaz de nivel de núcleo que puede emplearse en la operación 304 para enviar un identificador de instancia de aplicación:

```
# if(NTDDI_VERSION >= NTDDI_WIN8)

//

// Contexto de ECP para una aplicación para proporcionar su ID de instancia,

//

typedef struct _NETWORK_APP_INSTANCE_ECP_CONTEXT {

//

// Esto debe establecerse para el tamaño de esta estructura,

//

USHORT Size;

//
```

```

// Esto debe establecerse a cero,
//
USHORT Reserved;
//
// El llamante coloca un GUID que siempre debe ser único para una única
// instancia de la aplicación.
//
GUID AppInstanceID;
} NETWORK_APP_INSTANCE_ECP_CONTEXT,
*PNETWORK_APP_INSTANCE_ECP_CONTEXT;
//
// El GUID usado para la estructura APP_INSTANCE_ECP_CONTEXT.
//
// {6AA6BC45-A7EF-4af7-9008-FA462E144D74}
//
DEFINE_GUID(GUID_ECP_NETWORK_APP_INSTANCE, 0x6aa6bc45, 0xa7ef, 0x4af7, 0x90,
0x8, 0xfa, 0x46, 0x2e, 0x14, 0x4d, 0x74);
# endif // NTDDI_VERSION >= NTDDI_WIN8

```

Aunque se proporciona una interfaz de nivel de núcleo específica, los expertos en la materia apreciarán que pueden emplearse otras interfaces de nivel de núcleo en la operación 304 para enviar el identificador de instancia de aplicación.

En otra realización, puede emplearse una interfaz de usuario de aplicación (API) en la operación 304 para enviar un identificador de instancia de aplicación. En dicha realización, el solicitante y/o el cliente pueden enviar un identificador de instancia de aplicación haciendo una llamada a la API. La API puede alojarse en el cliente que realiza la operación 304 (por ejemplo, el primer cliente en un grupo de servidores) o la API puede alojarse en otro dispositivo y acceder a la misma por el solicitante u otra aplicación o procedimiento. El siguiente es un ejemplo no limitante de una API que puede emplearse en la operación 304 para enviar un identificador de instancia de aplicación:

```

NTSTATUS RegisterAppInstance (
    _in PGUID AppInstance
);

```

Aunque se proporciona una API específica, los expertos en la materia apreciarán que pueden emplearse otras API en la operación 304. Además, aunque la operación 304 se ilustra como una operación discreta, los expertos en la materia apreciarán que el envío del identificador de instancia de aplicación puede realizarse simultáneamente con el envío de la solicitud en la operación 302.

Cuando el recurso solicitado no está bloqueado, la solicitud enviada en la operación 302 se concede y el flujo continúa hacia la operación 306 donde se accede al recurso. Como se ha descrito anteriormente, el servidor o dispositivo que controla el recurso puede colocar el recurso en un estado acotado o bloqueado mientras el solicitante accede al recurso en la operación 306. En algún momento, mientras se accede al recurso, se produce un error, tal como los errores descritos con referencia a la figura 2, que hace que el cliente falle o pierda de otro modo la conexión al recurso. El error puede hacer que el cliente (por ejemplo, el primer cliente en el grupo de servidores) pierda el acceso al recurso antes de que el solicitante complete su uso del recurso. En tales circunstancias, el recurso puede no liberarse de su estado acotado o bloqueado.

El flujo continúa hacia la operación 308, donde se realiza una operación de conmutación por error. En las realizaciones, la operación de conmutación por error puede comprender la clonación del solicitante y su estado a un cliente diferente en el grupo de nodos cliente (por ejemplo, un segundo cliente). En las realizaciones, el estado del solicitante puede clonarse en el segundo y el solicitante puede ejecutarse en el segundo cliente de una manera tal que pueda reanudar la ejecución desde el punto donde falló el primer cliente. En otra realización, el solicitante puede estar en comunicación con el primer cliente (en lugar de ejecutarse en el mismo) en el momento de la primera conmutación por error de los clientes. En tales realizaciones, la operación de conmutación por error puede comprender que el solicitante establezca comunicaciones con un segundo cliente en el grupo de clientes.

En las realizaciones, la información de estado, incluyendo pero sin limitarse al identificador de instancia de aplicación de solicitantes, puede transferirse desde el primer cliente al segundo cliente. En una realización, el primer cliente puede enviar un mensaje que incluye el identificador de instancia de aplicación del solicitante y/o la información de estado del solicitante. El identificador de instancia de aplicación y/o el estado pueden enviarse durante el procedimiento de conmutación por error o, en las realizaciones, pueden enviarse antes de que el primer cliente fracase, tal como durante un procedimiento de replicación que clona información a través de los clientes en un entorno agrupado de cliente. En otra realización, el identificador de instancia de aplicación del solicitante y/o la información de estado pueden almacenarse en una localización o depósito central en la red agrupada de cliente. En dichas realizaciones, el procedimiento de conmutación por error puede proporcionar al segundo cliente la localización del identificador de instancia de aplicación del solicitante y/o información de estado. En otra realización más, el solicitante puede mantener su identificador de instancia de aplicación. En tales realizaciones, la operación de conmutación por error de cliente puede comprender volver a localizar o establecer de otro modo una conexión entre el solicitante y un segundo cliente.

En las realizaciones, después de la operación de conmutación por error de cliente, el flujo continúa hacia la operación 310. En la operación 310, se envía una segunda solicitud para el mismo recurso al entorno agrupado. En las realizaciones, la segunda solicitud se envía por el segundo cliente en el grupo de clientes en nombre del solicitante. La segunda solicitud puede enviarse usando la misma manera que se ha descrito con respecto al primer recurso en la operación 302. Con el fin de mantener un acceso continuo al recurso y evitar un retraso indebido, el flujo continúa hacia la operación 312 donde el identificador de instancia de aplicación se envía nuevamente al entorno agrupado. El identificador de instancia de aplicación puede enviarse en la operación 308 de acuerdo con una de las realizaciones descritas con respecto a la operación 304. En las realizaciones, debido a que un cliente diferente (por ejemplo, el segundo cliente) está enviando la segunda solicitud, el servidor que recibe la solicitud puede no ser capaz de identificar la segunda solicitud como perteneciente al mismo solicitante que mantiene un bloqueo en el recurso (por ejemplo, debido a que la solicitud se realiza desde una máquina diferente, una dirección diferente, etc.). Sin embargo, enviando los identificadores de instancia de aplicación en las operaciones 304 y 308, el servidor será capaz de identificar las solicitudes como pertenecientes al mismo solicitante y concederá acceso continuo al recurso como se ha descrito anteriormente con respecto a las figuras 1 y 2. El flujo continúa hacia la operación 314 y el solicitante reanuda el acceso al recurso. En las realizaciones, el segundo cliente puede recibir una respuesta a la segunda solicitud desde el servidor indicando que el servidor ha concedido la segunda solicitud. En las realizaciones, tras recibir la indicación, el segundo cliente puede acceder al recurso en nombre del solicitante.

La figura 4 es una realización de un procedimiento 400 realizado por un nodo en un entorno agrupado de servidor para proporcionar acceso continuo a un recurso. Las realizaciones del procedimiento 400 pueden realizarse por un nodo tal como el nodo 1 208 (figura 2) en un entorno agrupado, tal como el grupo 206 de nodos (figura 2). En las realizaciones, el procedimiento 400 puede realizarse por un nodo que tiene acceso a un recurso. El flujo comienza en la operación 402 donde el nodo recibe una solicitud de un recurso. En las realizaciones, un recurso puede ser un archivo, un objeto, un procedimiento, datos o cualquier otro tipo de recurso que esté bajo el control de y/o al que pueda accederse por el nodo que realiza la operación 400. Un identificador de instancia de aplicación puede recibirse con la solicitud en la operación 402.

El flujo continúa hacia la operación de decisión 404 donde se determina si el recurso está en un estado acotado o bloqueado. Los expertos en la materia apreciarán que en la operación 404 puede emplearse cualquier manera de determinar si un recurso está acotado o bloqueado. Si el recurso no está en un estado acotado o bloqueado, el flujo se ramifica hacia NO en la operación 412 donde se concede la solicitud del recurso. En las realizaciones, la concesión de la solicitud puede comprender permitir al solicitante el acceso al recurso, realizar una operación en el recurso en nombre del solicitante, o permitir cualquier tipo de acceso o modificación al recurso. Por ejemplo, la concesión de la solicitud en la operación 412 puede incluir abrir un archivo o crear un archivo.

Si el recurso está en un estado acotado o bloqueado, el flujo se ramifica hacia SÍ desde la operación 404 a la operación de decisión 406. En la operación de decisión 406, el identificador de instancia de aplicación recibido con la solicitud en la operación 402 se compara con un identificador de instancia de aplicación que está asociado con el recurso. Por ejemplo, como se describe con respecto a la figura 2, un nodo puede asociar un identificador de instancia de aplicación con un recurso cuando un cliente o aplicación accede a un recurso. Como se ha descrito anteriormente, la asociación del identificador de instancia de aplicación de un solicitante que accede a un recurso puede almacenarse en un nodo, por ejemplo, en una caché de instancia de aplicación, como se describe en varias realizaciones expuestas en la figura 2. En las realizaciones, el identificador de instancia de aplicación que se proporciona en un ECP enviado con una solicitud para un recurso, por ejemplo, en una estructura

_NETWORK_APP_INSTANCE_ECP_CONTEXT, puede añadirse a una lista de ECP asociada con el recurso.

En una realización, la asociación del recurso de instancia de aplicación puede residir localmente en el nodo que realiza el procedimiento 400. En tales casos, la comparación puede realizarse en una caché de instancia de aplicación local residente en el servidor. Sin embargo, tal como se ha expuesto con respecto a la figura 2, un entorno agrupado puede contener un número de cachés de instancia de aplicación distribuidas a través de diferentes nodos. Además, cada una de las diferentes cachés de instancia de aplicación puede almacenar datos separados y/o diferentes. El identificador de aplicación asociado con el recurso acotado o bloqueado puede almacenarse en un nodo diferente en el entorno agrupado. En tales casos, la operación 406 puede incluir el envío de una solicitud a un nodo diferente para realizar la comparación en la operación 406. La solicitud puede incluir el identificador de instancia de aplicación recibido en la operación 402.

Si el identificador de instancia de aplicación recibido no es el mismo que el identificador de instancia de aplicación asociado con el recurso, el flujo se ramifica hacia NO en la operación 410. En la operación 410 se deniega la solicitud de acceso al recurso recibido en la operación 402. En las realizaciones, la solicitud puede denegarse con el fin de evitar un conflicto de recursos. Debido a que el identificador de aplicación recibido no es el mismo que el identificador de instancia de aplicación asociado, la solicitud de acceso al recurso recibido en la operación 402 es de un solicitante o aplicación diferente. La concesión de una solicitud al cliente o aplicación diferente, como puede ser en este caso, puede provocar una situación de conflicto que interfiera con la aplicación que actualmente accede al recurso. Por ejemplo, la aplicación diferente puede modificar el recurso de una manera que modifica o interfiere de otro modo con las operaciones realizadas en el recurso por el solicitante que actualmente mantiene un bloqueo en el recurso.

Sin embargo, recibir un identificador de aplicación con la solicitud 402 que es el mismo que el identificador de aplicación asociado con los recursos acotados o bloqueados indica que puede haber ocurrido un error que provocó que el solicitante que accedía al recurso perdiera su acceso al recurso sin liberar adecuadamente el recurso. Por ejemplo, el solicitante puede operar en un grupo de nodos cliente. El cliente específico en el que el solicitante estaba operando podría haber perdido la conexión al servidor o fallado antes de que el solicitante completara sus operaciones sobre el recurso. Con el fin de proporcionar acceso continuo al recurso, es decir, para permitir al solicitante volver a obtener el acceso al recurso sin experimentar retrasos indebidos o inaceptables, el flujo se ramifica hacia SÍ en la operación 408.

En la operación 408, el recurso está invalidado. Como se ha descrito anteriormente, la invalidación del recurso puede incluir el cambio del estado acotado del recurso o eliminar de otro modo un bloqueo en el recurso. Por ejemplo, si el recurso es un archivo, invalidar el recurso puede incluir el cierre del archivo. Los expertos en la materia apreciarán que en la operación 408 puede emplearse cualquier procedimiento de liberación de un recurso acotado o bloqueado.

Haciendo referencia de nuevo a la figura 2, en las realizaciones, el acceso a un recurso puede estar bajo el control de un nodo en el entorno agrupado diferente del nodo que recibe la solicitud de acceso al recurso en la operación 402. Por ejemplo, un manipulador del recurso puede residir en un nodo diferente en el entorno agrupado. En tales realizaciones, invalidar el recurso puede incluir enviar una solicitud al nodo que controla el acceso al recurso para invalidar el recurso. En respuesta al envío de la solicitud, el nodo remoto puede invalidar el recurso.

Después de que el recurso se invalide, el flujo continúa hacia la operación 412 donde se concede la solicitud de acceso al recurso. La concesión de la solicitud puede comprender permitir al solicitante el acceso al recurso, realizar una operación en el recurso en nombre del solicitante, o permitir cualquier tipo de acceso o modificación al recurso. Por ejemplo, la concesión de la solicitud en la operación 412 puede incluir abrir un archivo o crear un archivo. La concesión de dicho acceso puede realizarse por el nodo que recibe la solicitud en la operación 402, o por otro nodo en el entorno agrupado.

Los procedimientos 300 y 400 son simplemente algunos ejemplos de flujos operativos que pueden realizarse de acuerdo con las realizaciones. Las realizaciones no están limitadas a la descripción específica proporcionada anteriormente con respecto a las figuras 3-6 y pueden incluir operaciones adicionales. Además, las etapas operativas representadas pueden combinarse en otras etapas y/o reordenarse. Además, pueden usarse más o menos etapas con respecto a las empleadas con los procedimientos descritos en las figuras 3-4.

La figura 5 ilustra un sistema 500 informático general, que puede usarse para implementar las realizaciones descritas en el presente documento. El sistema 500 informático es solo un ejemplo de un entorno informático y no pretende sugerir ninguna limitación en cuanto al ámbito de uso o funcionalidad de las arquitecturas informáticas y de red. Tampoco debe interpretarse que el sistema 500 informático tenga ninguna dependencia o requisito relacionado con uno cualquiera o una combinación de los componentes ilustrados en el sistema 500 informático a modo de ejemplo. En las realizaciones, el sistema 500 puede usarse como los clientes y/o servidores descritos anteriormente con respecto a las figuras 1 y 2.

En su configuración más básica, el sistema 500 incluye habitualmente al menos una unidad 502 de procesamiento y una memoria 504. Dependiendo de la configuración exacta y del tipo de dispositivo informático, la memoria 504

puede ser volátil (tal como RAM), no volátil (tal como ROM, memoria flash, etc.) o alguna combinación. Esta configuración más básica se ilustra en la figura 5 por la línea 506 discontinua. La memoria 504 de sistema almacena las instrucciones 520, tales como las instrucciones para realizar los procedimientos de disponibilidad continua desvelados en el presente documento y los datos 522, tales como los identificadores de instancia de aplicación que pueden almacenarse en un sistema de almacenamiento de archivos con un almacenamiento tal como el almacenamiento 508.

La expresión medio legible por ordenador tal como se usa en el presente documento puede incluir medios de almacenamiento informático. Los medios de almacenamiento informático pueden incluir medios volátiles y no volátiles, extraíbles y no extraíbles, implementados en cualquier procedimiento o tecnología para el almacenamiento de información, tales como instrucciones legibles por ordenador, estructuras de datos, módulos de programa u otros datos. La memoria 504 de sistema, el almacenamiento extraíble y el almacenamiento 508 no extraíble son todos ejemplos de medios de almacenamiento informático (por ejemplo, el almacenamiento de memoria). Los medios de almacenamiento informático pueden incluir, pero sin limitarse a, RAM, ROM, memoria de solo lectura borrrable eléctricamente (EEPROM), memoria flash u otra tecnología de memoria, CD-ROM, discos versátiles digitales (DVD) u otro almacenamiento óptico, cinta magnética, almacenamiento en disco magnético u otros dispositivos de almacenamiento magnético, o cualquier otro medio que pueda usarse para almacenar información y al que pueda accederse por el dispositivo 500 informático. Cualquiera de tales medios de almacenamiento informático puede ser parte del dispositivo 500. El dispositivo 500 informático también puede tener un dispositivo(s) (514) de entrada, tal como un teclado, un ratón, un puntero, un dispositivo de entrada de sonido, un dispositivo de entrada táctil, etc. También puede incluirse un dispositivo(s) 516 de salida, tal como una pantalla, unos altavoces, una impresora, etc. Los dispositivos mencionados anteriormente son ejemplos y pueden usarse otros.

La expresión medio legible por ordenador tal como se usa en el presente documento también puede incluir medios de comunicación. Los medios de comunicación pueden incorporarse mediante instrucciones legibles por ordenador, estructuras de datos, módulos de programa u otros datos en una señal de datos modulada, tal como una onda portadora u otro mecanismo de transporte, e incluyen cualquier medio de suministro de información. La expresión "señal de datos modulada" puede describir una señal que tiene una o más características establecidas o cambiadas de tal manera que codifican información en la señal. A modo de ejemplo, y sin limitación, los medios de comunicación pueden incluir medios cableados, tales como una red cableada o una conexión directa por cable, y medios inalámbricos, tales como medios acústicos, de radiofrecuencia (RF), infrarrojos y otros medios inalámbricos.

Las realizaciones de la invención pueden ponerse en práctica a través de un sistema en un chip (SOC) donde cada uno o muchos de los componentes ilustrados en la figura 5 pueden integrarse en un único circuito integrado. Un dispositivo SOC de este tipo puede incluir una o más unidades de procesamiento, unidades gráficas, unidades de comunicaciones, unidades de virtualización de sistemas y diversas funcionalidades de aplicación, todas las cuales están integradas (o "quemadas") sobre el sustrato de chip como un único circuito integrado. Cuando se opera a través de un SOC, la funcionalidad, descrita en el presente documento, con respecto a proporcionar acceso continuo a un recurso puede operar a través de una lógica de aplicación específica integrada con otros componentes del dispositivo/sistema 500 informático en el único circuito integrado (chip).

Se ha hecho referencia a lo largo de la presente memoria descriptiva a "una realización", lo que significa que un rasgo, estructura o característica descrita específica se incluye en al menos una realización. Por lo tanto, el uso de tales frases puede referirse a más de una realización. Además, los rasgos, estructuras o características descritas pueden combinarse de cualquier manera adecuada en una o más realizaciones.

Sin embargo, los expertos en la materia pueden reconocer que las realizaciones pueden ponerse en práctica sin uno o más de los detalles específicos, o con otros procedimientos, recursos, materiales, etc. En otros casos, no se han mostrado o descrito en detalle estructuras, recursos u operaciones bien conocidas, simplemente para evitar complicar aspectos de las realizaciones.

Aunque se han ilustrado y descrito realizaciones y aplicaciones a modo de ejemplo, debe entenderse que las realizaciones no se limitan a la configuración y los recursos precisos descritos anteriormente. Pueden hacerse diversas modificaciones, cambios y variaciones evidentes para los expertos en la materia en la disposición, operación y detalles de los procedimientos y sistemas desvelados en el presente documento.

REIVINDICACIONES

1. Un procedimiento para proporcionar acceso continuo a un recurso, comprendiendo el procedimiento:

recibir en un grupo (106) de servidores, que incluye al menos un servidor (106A, 106B, 106C; 208, 216), una primera solicitud (222) para acceder a un recurso (226) desde una aplicación (203) cliente, ejecutándose la aplicación cliente en un primer cliente (102A; 202), en el que la primera solicitud se recibe desde el primer cliente en un grupo (106; 201) de clientes;
asociar en el grupo de servidores un primer identificador de instancia de aplicación con el recurso;
permitir por el grupo de servidores el primer acceso solicitado al recurso;
recibir en el grupo de servidores una segunda solicitud (224) para el recurso desde la aplicación cliente después de recuperarse de un fallo, en el que la segunda solicitud se recibe desde un segundo cliente (102B; 204) diferente del primer cliente;
recibir en el grupo de servidores un segundo identificador de instancia de aplicación asociado con la segunda solicitud;
determinar si el primer identificador de instancia de aplicación y el segundo identificador de instancia de aplicación son los mismos; y
cuando los identificadores de instancia de aplicación primero y segundo son los mismos, realizar por el grupo de servidores las etapas que comprenden:
invalidar la primera solicitud; y
conceder la segunda solicitud de acceso al recurso.

2. El procedimiento de la reivindicación 1, en el que el primer identificador de instancia de aplicación está asociado con una instancia de aplicación de una solicitud abierta.

3. El procedimiento de la reivindicación 1, en el que el primer identificador de instancia de aplicación está asociado con un procedimiento.

4. El procedimiento de la reivindicación 1, en el que el primer identificador de instancia de aplicación está asociado con al menos un procedimiento secundario de una aplicación.

5. El procedimiento de la reivindicación 1, en el que la asociación del primer identificador de instancia de aplicación comprende recibir el primer identificador de instancia de aplicación en una estructura `_NETWORK_APP_INSTANCE_ECP_CONTEXT`.

6. Un procedimiento para proporcionar una conmutación por error de clientes agrupados para hacer migrar una aplicación (203) cliente desde un primer cliente (102A; 202) a un segundo cliente (102B; 204), comprendiendo el procedimiento:

recibir, desde el primer cliente (102A; 202) en el segundo cliente (102B; 204), un identificador de instancia de aplicación para la aplicación (203) cliente;
enviar, desde el segundo cliente a un grupo (106) de servidores que incluye al menos un servidor (106A, 106B, 106C; 208, 216), una solicitud (224) para acceder al recurso en nombre de la aplicación cliente;
enviar, desde el segundo cliente al grupo de servidores, el identificador de instancia de aplicación para la aplicación cliente; y
acceder, por el segundo cliente, cuando se permite por el grupo (106) de servidores, al recurso en nombre de la aplicación cliente.

7. El procedimiento de la reivindicación 6, en el que un servidor que previamente ha concedido una solicitud para acceder al recurso desde el primer cliente en nombre de la aplicación cliente es el servidor que ha permitido el acceso al recurso por el segundo cliente.

8. El procedimiento de la reivindicación 6, en el que el segundo cliente envía la solicitud en respuesta a una conmutación por error de aplicación cliente.

9. Un sistema para facilitar la conmutación por error de aplicación cliente en un entorno agrupado, comprendiendo el sistema:

un grupo (106) de servidores que incluye al menos un servidor (106A, 106B, 106C; 208, 216) que comprende:

al menos un procesador (502) configurado para ejecutar instrucciones ejecutables por ordenador;
al menos un medio de almacenamiento legible por ordenador que almacena las instrucciones ejecutables por ordenador que cuando se ejecutan por el al menos un procesador se adaptan para realizar las siguientes etapas:

recibir una primera solicitud (222) para acceder a un recurso (226) desde un primer cliente (102A; 202) en nombre de una aplicación (203) cliente;

- asociar un primer identificador de instancia de aplicación con el recurso;
permitir el acceso de la aplicación cliente al recurso;
recibir una segunda solicitud (224) para el recurso desde un segundo cliente (102B; 204), en el que el
segundo cliente es diferente del primer cliente;
5 recibir, desde el segundo cliente, un segundo identificador de instancia de aplicación asociado con la
segunda solicitud;
determinar si el primer identificador de instancia de aplicación y el segundo identificador de instancia de
aplicación son los mismos;
10 cuando los identificadores de instancia de aplicación primero y segundo son los mismos, realizar las
etapas que comprenden:
invalidar la primera solicitud; y
conceder la segunda solicitud de acceso al recurso.
10. El sistema de la reivindicación 9, en el que el sistema comprende además:
el primer cliente, que comprende:
15 al menos un procesador configurado para ejecutar instrucciones ejecutables por ordenador;
al menos un medio de almacenamiento legible por ordenador que almacena las instrucciones ejecutables por
ordenador que cuando se ejecutan por el al menos un procesador se adaptan para realizar las siguientes
etapas:
20 enviar la primera solicitud;
enviar el identificador de instancia de aplicación a un segundo cliente.

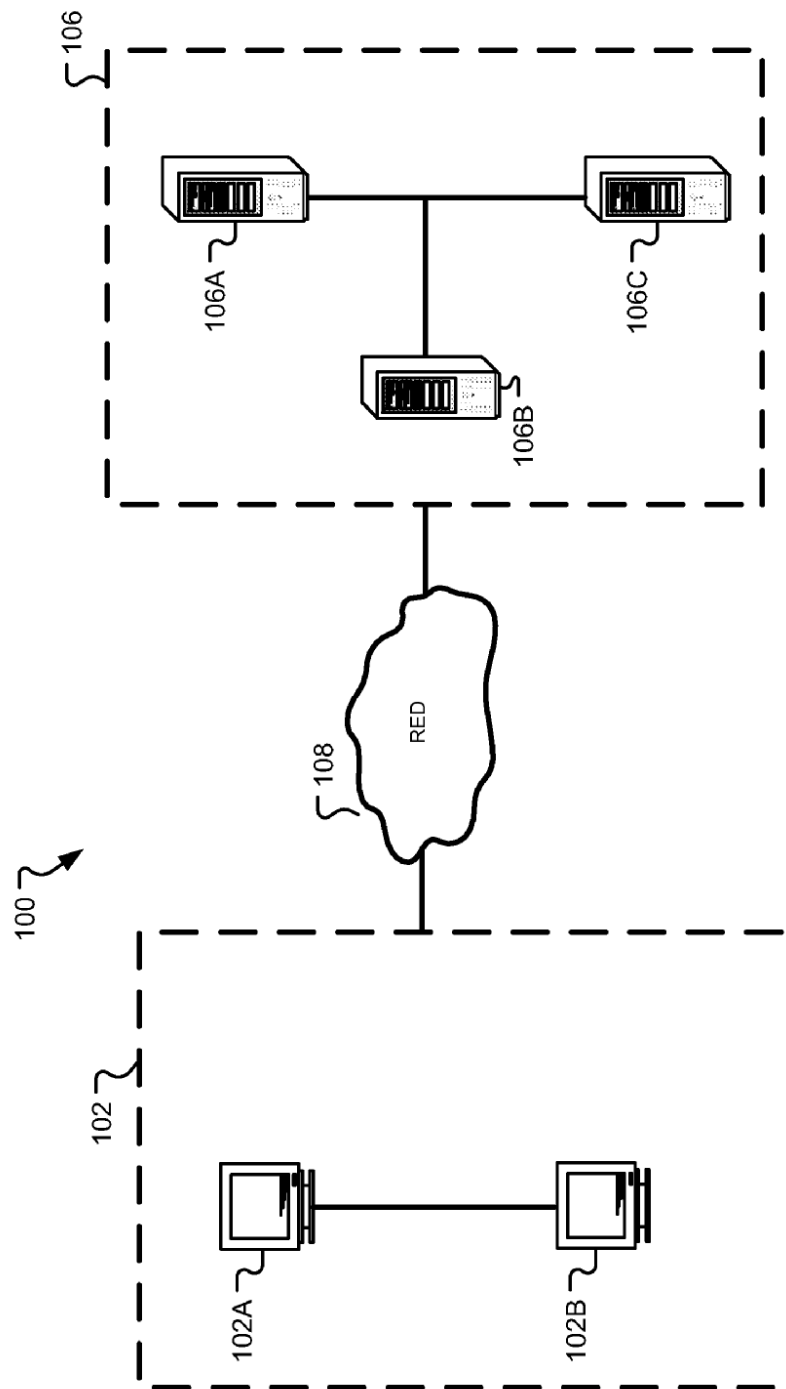


FIG. 1

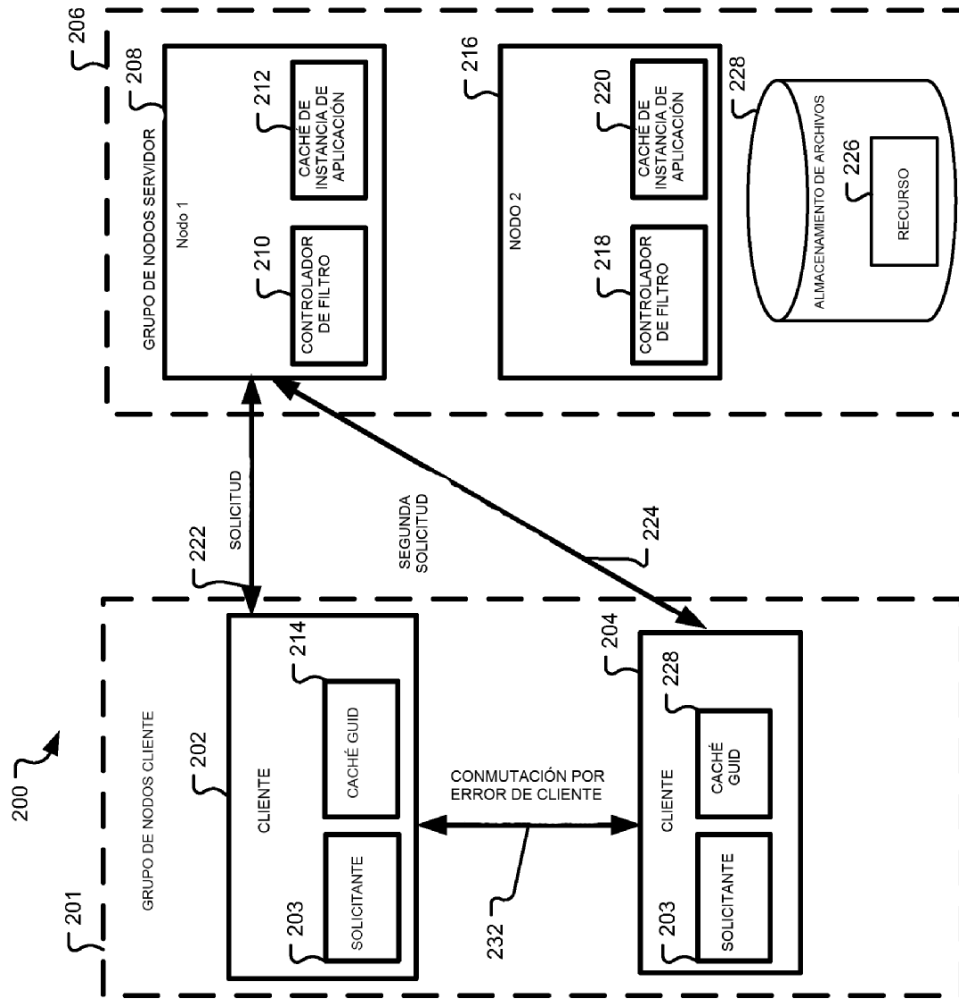


FIG. 2

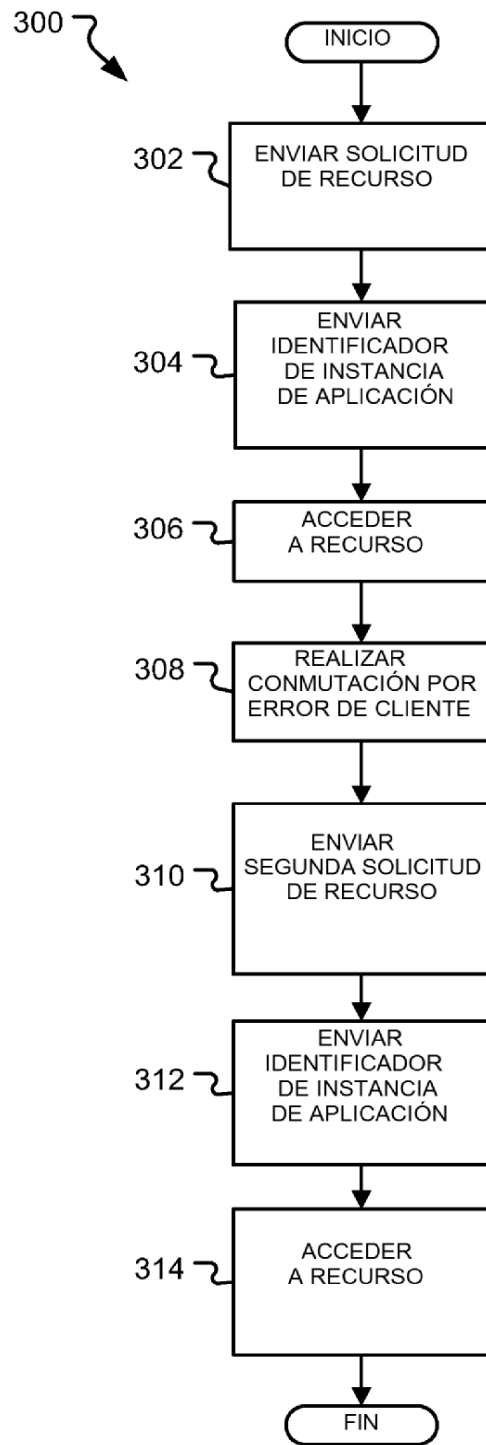


FIG. 3

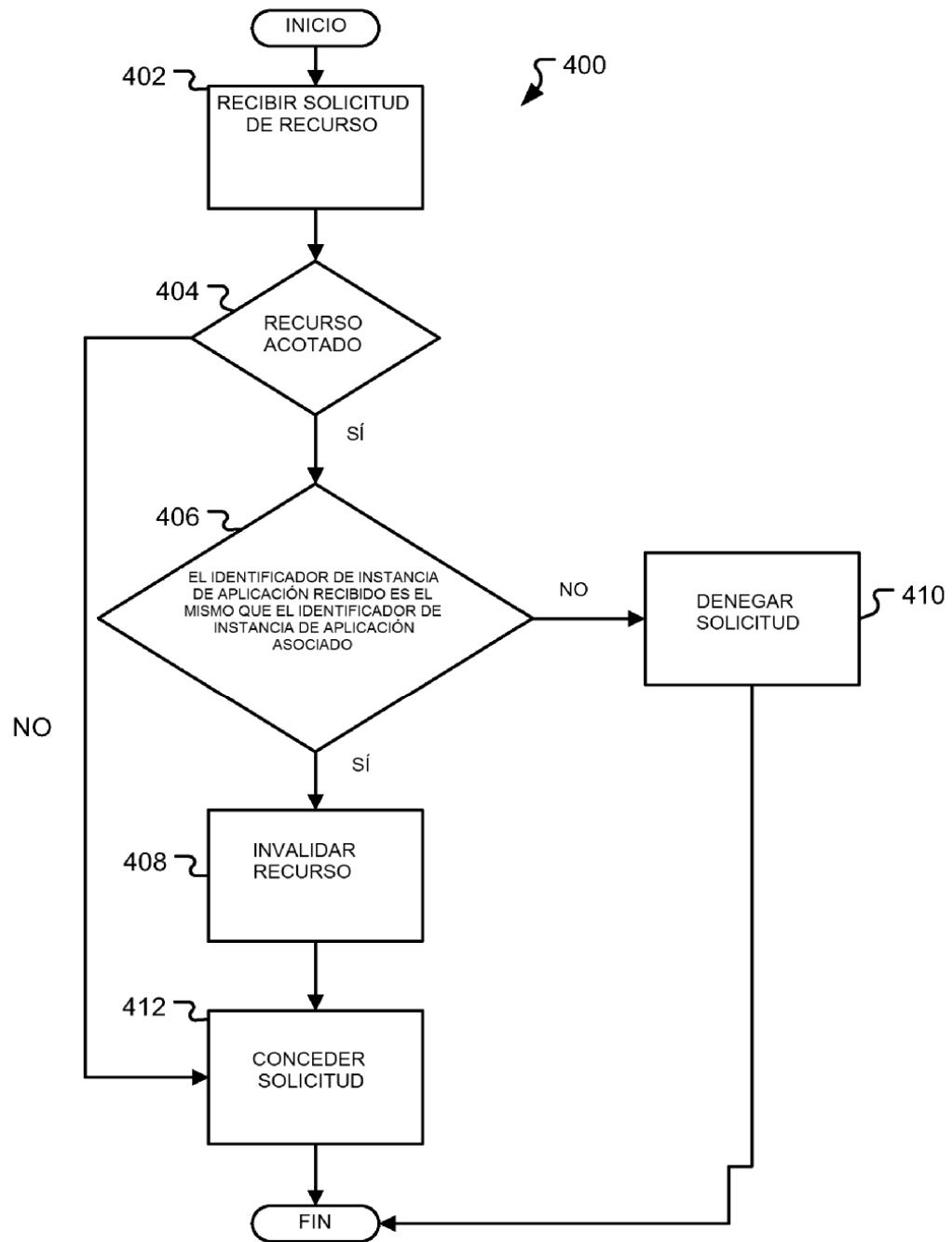


FIG. 4

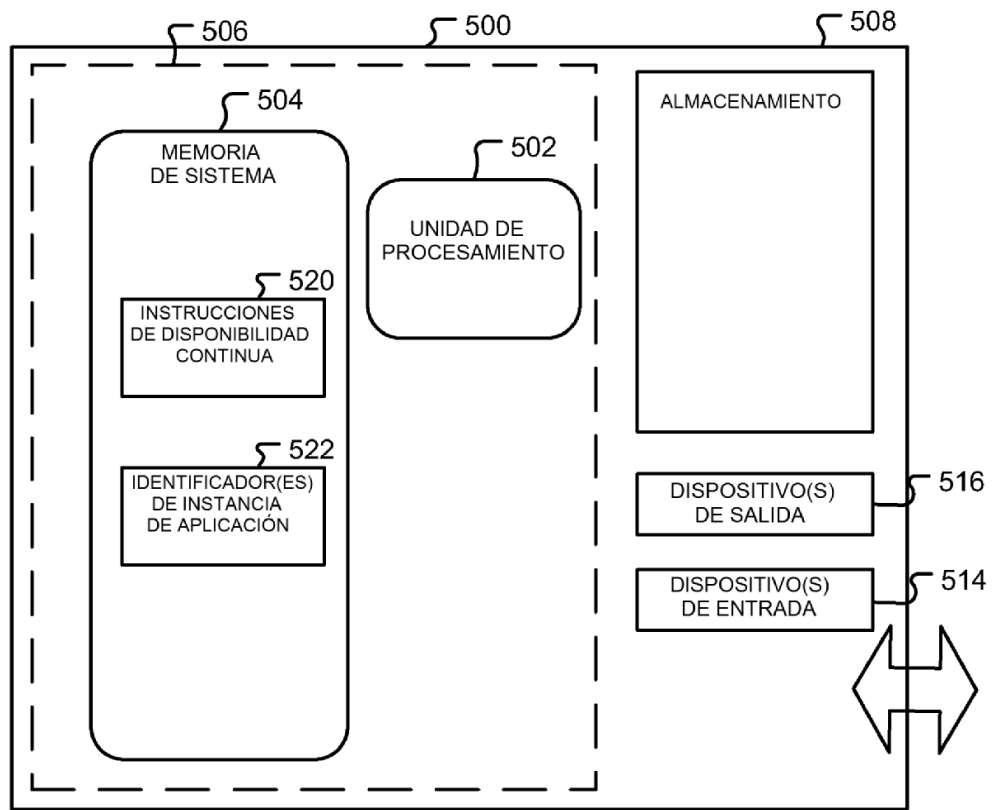


FIG. 5