

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 658 188**

51 Int. Cl.:

**G06F 12/06** (2006.01)

**G06F 3/06** (2006.01)

**G06F 12/02** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **27.12.2012 PCT/CN2012/087743**

87 Fecha y número de publicación internacional: **03.07.2014 WO14101070**

96 Fecha de presentación y número de la solicitud europea: **27.12.2012 E 12890935 (5)**

97 Fecha y número de publicación de la concesión europea: **20.12.2017 EP 2921964**

54 Título: **Método y aparato de extensión de particiones**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:  
**08.03.2018**

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

72 Inventor/es:  
**ZHANG, FENG**

74 Agente/Representante:  
**LEHMANN NOVO, María Isabel**

ES 2 658 188 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

Método y aparato de extensión de particiones

### CAMPO TÉCNICO

5 La presente invención se refiere al campo de las tecnologías de almacenamiento y, en particular, a un método y un aparato de extensión de particiones.

### ANTECEDENTES

10 En un sistema de almacenamiento distribuido, una tabla de hash distribuida (DHT por sus siglas en inglés) se utiliza generalmente para determinar una posición de almacenamiento de datos. Es decir, un valor Hash se calcula utilizando una función hash y los datos se asignan al espacio Hash de acuerdo con el valor Hash, en que el espacio Hash es un anillo conectado de la cabeza a la cola. Un anillo Hash puede dividirse en varias partes iguales, cada una de las partes se denomina partición y las particiones se distribuyen por igual a los nodos físicos. Además, se establece una relación de mapeo entre datos y una partición y se establece una relación de mapeo entre una partición y un nodo físico y se almacena en una DHT. La lectura y escritura de datos, la expansión de capacidad de un clúster distribuido, la resolución de problemas y similares dependen de un cambio de la DHT. Por lo tanto, DHT necesita alcanzar el objetivo de que la carga sea pareja, la fiabilidad sea buena, una manera de almacenamiento de los datos en un nodo facilite la migración de datos y la expansibilidad sea buena (es decir, la expansión de la capacidad se puede realizar múltiples veces para un sistema).

15 En la técnica anterior, la expansión de la capacidad del sistema puede implementarse añadiendo un nuevo nodo físico en el sistema de almacenamiento distribuido, es decir, aumentando una cantidad de nodos físicos. Un método de expansión de la capacidad específico puede incluir: agregar un nuevo nodo físico al sistema y mover una partición que está en un nodo físico original al nuevo nodo físico, para realizar la expansión de la capacidad del sistema de almacenamiento distribuido. Las cantidades de particiones en los nodos físicos son prácticamente las mismas.

20 Durante un proceso de implementación de la expansión de la capacidad del sistema anterior, la técnica anterior tiene al menos los siguientes problemas: si la expansión de la capacidad se realiza múltiples veces para el sistema de almacenamiento distribuido, una cantidad de particiones en un nodo físico está disminuyendo, y la expansión de la capacidad del sistema se ve afectada (por ejemplo, cuando sólo existen cuatro particiones en cada nodo físico, la capacidad del sistema sólo se puede expandir a un máximo de cuatro veces la capacidad original), lo que causa una expansibilidad deficiente del sistema de almacenamiento distribuido.

25 El documento US 2012/323852 A1 se refiere a la redistribución de la base de datos utilizando particiones virtuales. Se describe que cuando al sistema se añade un nuevo nodo, se crea(n) partición(es) en el nuevo nodo y los datos se redistribuyen desde los nodos existentes a la nueva partición del nuevo nodo.

### SUMARIO

30 Realizaciones de la presente invención proporcionan un método y un aparato de extensión de particiones, de modo que puede aumentarse una cantidad de particiones en un nodo físico durante la expansión de la capacidad de un sistema de almacenamiento distribuido, garantizando así la expansibilidad del sistema de almacenamiento distribuido.

En las reivindicaciones adjuntas se proporcionan soluciones bajo la presente invención.

35 Con el método y el aparato de extensión de particiones proporcionados en las realizaciones de la presente invención se puede adquirir un valor medio de cantidades de particiones en todos los nodos físicos en un sistema de almacenamiento distribuido; si el valor medio de las cantidades de las particiones es menor que un umbral preestablecido, la multiplicación se puede realizar en una cantidad de particiones en cada nodo físico; y una relación de mapeo entre una partición recién agregada y un nodo físico en el que se encuentra la nueva partición agregada se puede añadir a una tabla de hash distribuida DHT pre-almacenada. Comparado con la técnica anterior, en la que la expansibilidad de un sistema de almacenamiento distribuido es deficiente debido a una pequeña cantidad de particiones en un nodo físico, con el método y el aparato de extensión de partición proporcionados en las realizaciones de la presente invención, si una cantidad de particiones es menor que un umbral preestablecido, la multiplicación en la cantidad de particiones se puede realizar durante la expansión de la capacidad del sistema de almacenamiento distribuido, asegurando con ello una expansibilidad del sistema de almacenamiento distribuido.

40 **BREVE DESCRIPCIÓN DE LOS DIBUJOS**

5 Para describir las soluciones técnicas en las realizaciones de la presente invención o en la técnica anterior más claramente, lo siguiente presenta brevemente los dibujos adjuntos requeridos para describir las realizaciones o la técnica anterior. Aparentemente, los dibujos que se acompañan en la siguiente descripción muestran meramente algunas realizaciones de la presente invención, y una persona medianamente experta en la técnica todavía puede derivar otros dibujos de estos dibujos adjuntos sin esfuerzos creativos.

La FIG. 1 es un diagrama de flujo de un método de extensión de particiones de acuerdo con la Realización 1 de la presente invención;

la FIG. 2 es un diagrama de flujo de un método de extensión de particiones de acuerdo con la Realización 2 de la presente invención;

10 la FIG. 3 es un diagrama estructural esquemático de un aparato de extensión de particiones de acuerdo con la Realización 3 de la presente invención;

la FIG. 4 es otro diagrama estructural esquemático de un aparato de extensión de particiones de acuerdo con la Realización 3 de la presente invención; y

15 la FIG. 5 es un diagrama estructural esquemático de un aparato de extensión de particiones de acuerdo con la Realización 4 de la presente invención.

### DESCRIPCIÓN DE REALIZACIONES

20 Lo siguiente describe de forma clara y completa las soluciones técnicas en las realizaciones de la presente invención con referencia a los dibujos adjuntos en las realizaciones de la presente invención. Aparentemente, las realizaciones descritas son meramente algunas, pero no todas las realizaciones de la presente invención. Todas las demás realizaciones obtenidas por una persona medianamente experta en la técnica basada en las realizaciones de la presente invención sin esfuerzos creativos caerán dentro del alcance de protección de la presente invención.

#### Realización 1

Una realización de la presente invención proporciona un método de extensión de particiones. Tal como se muestra en la FIG. 1, el método incluye los siguientes pasos:

25 101. Un aparato de extensión de particiones adquiere un valor medio de cantidades de particiones en todos los nodos físicos en un sistema de almacenamiento distribuido.

30 Cuando se utiliza una tabla hash distribuida (Distributed Hash Table, DHT) para almacenar datos, el sistema de almacenamiento distribuido puede calcular, de acuerdo con un identificador de datos diferentes y utilizando una función hash, un valor hash Hash que es de los datos y corresponde al identificador de datos. Los datos se asignan al espacio hash de acuerdo con el valor hash, en que el espacio hash es un espacio circular conectado de la cabeza a la cola. Un anillo hash se puede dividir en varias partes iguales, a cada parte se la alude como una partición, y las particiones se distribuyen por igual en los nodos físicos. Un nodo físico puede ser una dirección física del sistema de almacenamiento distribuido. El sistema de almacenamiento distribuido puede dividir las direcciones físicas del sistema de almacenamiento distribuido en múltiples segmentos de direcciones, y cada segmento de direcciones físicas es un nodo físico.

35 Un método para adquirir, mediante el aparato de extensión de particiones, el valor medio de las cantidades de las particiones en los nodos físicos en el sistema de almacenamiento distribuido puede incluir: adquirir periódicamente, por el aparato de extensión de particiones, el valor medio de las cantidades de las particiones en todos los nodos físicos en el sistema de almacenamiento distribuido; o responder, mediante el aparato de extensión de particiones, a una instrucción de activación de un usuario, para adquirir el valor medio de las cantidades de las particiones en todos los nodos físicos en el sistema de almacenamiento distribuido. Específicamente, el aparato de extensión de particiones puede preestablecer un período de adquisición, y adquirir espontáneamente el valor medio de las cantidades de las particiones en todos los nodos físicos en el sistema de almacenamiento distribuido de acuerdo con el período; o cuando se realiza la capacidad de expansión del sistema de almacenamiento distribuido, el aparato de extensión de particiones recibe la instrucción de activación del usuario y responde a la instrucción de activación para adquirir el valor medio de las cantidades de las particiones en todos los nodos físicos en el sistema de almacenamiento distribuido.

102. Si el valor medio de las cantidades de las particiones es menor que un umbral preestablecido, el aparato de extensión de particiones realiza la multiplicación en una cantidad de particiones en cada nodo físico.

50 El umbral preestablecido puede establecerse en un valor que satisfaga el valor medio de las cantidades de las particiones en todos los nodos físicos en el sistema de almacenamiento distribuido cuando la expansión de la capacidad se realiza para el sistema de almacenamiento distribuido al menos una vez. Después de que el aparato de extensión de particiones realiza la multiplicación en la cantidad de particiones en cada nodo físico, el método en

esta realización de la presente invención incluye, además: adquirir, mediante el aparato de extensión de particiones, un valor hash de datos en cada partición en cada nodo físico; adquirir, mediante el aparato de extensión de particiones, un valor de módulo que se obtiene al realizar una operación de módulo en el valor hash de los datos y una cantidad que se obtiene después de la extensión y es de particiones en el nodo físico en el que se encuentran los datos; y determinar, mediante el aparato de extensión de particiones de acuerdo con el valor de módulo de los datos, una partición de destino a la que han de migrar los datos, y si la partición de destino es diferente de la partición en la que se encuentran actualmente los datos, migrar, mediante el aparato de extensión de particiones, los datos a la partición de destino.

103. El aparato de extensión de particiones agrega una relación de mapeo entre una partición recién agregada y un nodo físico en el cual la partición recién agregada se ubica en una DHT pre-almacenada.

En el sistema de almacenamiento distribuido, la DHT puede usarse para determinar una posición de almacenamiento de datos. Es decir, la DHT puede almacenar una relación de mapeo entre un identificador de datos y una partición y una relación de mapeo entre una partición y un nodo físico. Por lo tanto, después de realizar la multiplicación de la cantidad de particiones en cada nodo físico, el aparato de extensión de particiones necesita agregar la relación de mapeo entre la partición recién agregada y el nodo físico en el que se encuentra la partición recién agregada en la DHT pre-almacenada.

En el método de extensión de particiones proporcionado en esta realización de la presente invención, se puede adquirir un valor medio de cantidades de particiones en todos los nodos físicos en un sistema de almacenamiento distribuido; si el valor medio de las cantidades de las particiones es menor que un umbral preestablecido, la multiplicación se puede realizar en una cantidad de particiones en cada nodo físico; y una relación de mapeo entre una partición recién agregada y un nodo físico en el que está ubicada la partición recién agregada se puede agregar a una tabla de hash distribuida DHT pre-almacenada. Comparado con la técnica anterior en la cual la expansibilidad de un sistema de almacenamiento distribuido es deficiente debido a una pequeña cantidad de particiones en un nodo físico, en el método de extensión de particiones proporcionado en esta realización de la presente invención, si una cantidad de particiones es menor que un valor preestablecido, la multiplicación en la cantidad de las particiones se puede realizar durante la expansión de la capacidad del sistema de almacenamiento distribuido, asegurando con ello la expansibilidad del sistema de almacenamiento distribuido.

## Realización 2

Una realización de la presente invención proporciona un método de extensión de particiones. Tal como se muestra en la FIG. 2, el método incluye los siguientes pasos:

201. Un aparato de extensión de particiones adquiere un valor medio de cantidades de particiones en todos los nodos físicos en un sistema de almacenamiento distribuido.

A modo de ejemplo, un método para adquirir, mediante el aparato de extensión de particiones, el valor medio de las cantidades de las particiones en todos los nodos físicos en el sistema de almacenamiento distribuido puede ser específicamente el siguiente: las consultas del aparato de extensión de particiones, en una DHT pre-almacenada, una cantidad de nodos físicos en el sistema de almacenamiento distribuido y una cantidad total de las particiones en todos los nodos físicos; y el aparato de extensión de particiones calcula el valor medio de las cantidades de las particiones en los nodos físicos de acuerdo con la cantidad de nodos físicos y la cantidad total de las particiones en todos los nodos físicos.

La cantidad de nodos físicos es una cantidad total de nodos físicos en el sistema de almacenamiento distribuido. El método para calcular, mediante el aparato de extensión de particiones, el valor medio de las cantidades de las particiones en los nodos físicos de acuerdo con la cantidad de nodos físicos y la cantidad total de las particiones en todos los nodos físicos puede ser específicamente como sigue : usar un cociente que se obtiene mediante el aparato de extensión de particiones dividiendo la cantidad total de las particiones en todos los nodos físicos por la cantidad de nodos físicos como el valor medio de las cantidades de las particiones en los nodos físicos. Por ejemplo, si la cantidad de nodos físicos es A, la cantidad total de las particiones en todos los nodos físicos es B, y el valor medio de las cantidades de las particiones en los nodos físicos es C, la cantidad de nodos físicos, la cantidad total de las particiones en todos los nodos físicos y el valor medio de las cantidades de las particiones en los nodos físicos debe satisfacer la fórmula  $C = A/B$ .

Además, en un escenario de aplicación de esta realización, cuando el sistema de almacenamiento distribuido usa la DHT para almacenar datos, las cantidades de las particiones en los nodos físicos son aproximadamente las mismas. Por lo tanto, el aparato de extensión de particiones también puede consultar, en la DHT pre-almacenada, una cantidad de particiones en cualquier nodo físico en el sistema de almacenamiento distribuido, y usar la cantidad de particiones en el nodo físico como el valor medio de las cantidades de las particiones en todos los nodos físicos en el sistema de almacenamiento distribuido.

202. El aparato de extensión de particiones determina si el valor medio de las cantidades de las particiones es menor que un umbral preestablecido; si el valor medio de las cantidades de las particiones es menor que el umbral preestablecido, realizar el paso 203; o si el valor medio de las cantidades de las particiones no es menor que el umbral preestablecido, realizar el paso 211.

5 203. El aparato de extensión de particiones realiza la multiplicación en una cantidad de particiones en cada nodo físico.

El que el aparato de extensión de particiones realice la multiplicación en una cantidad de particiones en cada nodo físico puede incluir: multiplicar, por el aparato de extensión de particiones, la cantidad de las particiones en cada nodo físico a al menos dos veces la cantidad de las particiones.

10 Específicamente, el aparato de extensión de particiones puede modificar la cantidad de las particiones en cada nodo físico en la DHT pre-almacenada a al menos dos veces la cantidad de las particiones en cada nodo físico.

15 Debe señalarse que, generalmente, cuando el aparato de extensión de particiones realiza la multiplicación en la cantidad de las particiones en cada nodo físico, los tiempos para la multiplicación en la cantidad de las particiones son 2. Ciertamente, el aparato de extensión de particiones también puede realizar la extensión tres veces o incluso más veces en la cantidad de particiones en cada nodo físico. El que el aparato de extensión de particiones realice la extensión dos veces en la cantidad de particiones en cada nodo físico se usa como ejemplo en esta realización de la presente invención. Sin embargo, los tiempos para la multiplicación realizada por el aparato de extensión de particiones en la cantidad de particiones en cada nodo físico no están limitados.

204. El aparato de extensión de particiones adquiere un valor hash de datos en cada partición en cada nodo físico.

20 Específicamente, el aparato de extensión de particiones puede consultar, en la DHT pre-almacenada y de acuerdo con un identificador de datos en cada partición, el valor hash de los datos en cada partición en cada nodo físico. Los valores hash de diferentes datos en una partición son diferentes. Por ejemplo, una partición A incluye datos 1, datos 2 y datos 3. Un identificador de datos de los datos 1 es a y un valor hash es Hash1, un identificador de datos de los datos 2 es b y un valor hash es Hash2, y un identificador de datos de los datos 3 es c y un valor de hash es Hash3.

25 205. El aparato de extensión de particiones adquiere un valor de módulo que se obtiene al realizar la operación de módulo en el valor hash de los datos y una cantidad que se obtiene después de la extensión y es de particiones en el nodo físico en el que se ubican los datos.

30 La cantidad que se obtiene después de la extensión y es de las particiones en el nodo físico en el que se ubican los datos es un producto obtenido al multiplicar una cantidad que está antes de la extensión y es de particiones en el nodo físico en el que se ubican los datos por los tiempos para la multiplicación realizada por el aparato de extensión de particiones en la cantidad de particiones en cada nodo físico.

35 A modo de ejemplo, si un valor hash de datos K es  $Hash_k$ , una cantidad de particiones en un nodo físico en el que se ubican los datos K es N, y el aparato de extensión de particiones realiza la multiplicación por n veces la cantidad de las particiones en cada nodo físico, una cantidad que se obtiene después de la extensión y es de particiones en el nodo físico en el que se ubican los datos K es  $n \times N$ , y un valor de módulo  $x_k$ , que se obtiene al realizar la operación de módulo en el valor hash de los datos K y la cantidad que se obtiene después de la extensión y es de las particiones en el nodo físico en el que se ubica los datos K, se puede obtener mediante cálculo de acuerdo con una fórmula  $x_k = Hash_k \% (n \times N)$ .

40 206. El aparato de extensión de particiones determina, de acuerdo con el valor de módulo de los datos, una partición de destino a la que han de migrar los datos, y si la partición de destino es diferente de la partición en la que se encuentran actualmente los datos, los datos migran a la partición de destino.

45 Cuando el sistema de almacenamiento distribuido almacena datos, se puede calcular un valor hash Hash de los datos a ser almacenados de acuerdo con una función hash, los datos a ser almacenados se distribuyen, de acuerdo con un valor hash de cada parte de los datos a ser almacenados y la cantidad total de las particiones en el sistema de almacenamiento distribuido, a cada partición para el almacenamiento, y luego las particiones en el sistema de almacenamiento distribuido se distribuyen uniformemente en cada nodo físico.

50 A modo de ejemplo, si los valores hash Hash de algunos datos son 45, 46, 47, 48, ..., 75, respectivamente, y una cantidad de particiones es N, datos cuyos resultados se obtienen de la operación módulo en un valor hash de cada parte de los datos y la cantidad de las particiones son las mismas se distribuye en una misma partición, por ejemplo en una fórmula  $Hash \% N$ , si  $N = 3$ , los restos obtenidos de la operación módulo en los valores hash de los datos y la cantidad de las particiones son 0, 1, 2, 0, 1, 2, 0, 1, 2, 0, 1, 2 ... 0, 1, 2, 0, respectivamente. Los datos pueden distribuirse de acuerdo con los restos obtenidos de la operación de módulo, a las siguientes tres particiones:

partición 1 (un resto es 0): 45, 48, 51, 54, 57, 60, 63, 66, 69, 72, 75;

partición 2 (un resto es 1): 46, 49, 52, 55, 58, 61, 64, 67, 70, 73; y

partición 3 (un resto es 2): 47, 50, 53, 56, 59, 62, 65, 68, 71, 74.

5 A modo de ejemplo, la partición 1 anterior se usa como ejemplo para realizar la extensión de particiones. En este caso, la cantidad N de las particiones es 3. Cuando los tiempos n para la multiplicación de la cantidad de particiones es  $n = 2$ , una cantidad que se obtiene después de la extensión y es de particiones en un nodo físico en el que la partición 1 está ubicada es  $2 \times 3$ , y los valores del módulo de todos los datos en la partición 1 son  $x_{45} = 45\%(2 \times 3) = 3$ ,  $x_{48} = 48\%(2 \times 3) = 0$ ,  $x_{51} = 51\%(2 \times 3) = 3$ ,  $x_{54} = 54\%(2 \times 3) = 0$ ,  $x_{57} = 57\%(2 \times 3) = 3$  ...  $x_{72} = 72\%(2 \times 3) = 0$ , and  $x_{75} = 75\%(2 \times 3) = 3$ , respectivamente.

10 Específicamente, si el tiempo n para la multiplicación en la cantidad de particiones es  $n = 2$ , el aparato de extensión de particiones puede migrar datos cuyo valor de módulo es 3 a una partición recién agregada en un nodo físico en el que se encuentran los datos y puede retener datos cuyo valor de módulo es 0 en una partición original. Por ejemplo, el aparato de extensión de particiones puede migrar datos correspondientes a  $x_{45} = 45\%(2 \times 3) = 3$ ,  $x_{51} = 51\%(2 \times 3) = 3$ ,  $x_{57} = 57\%(2 \times 3) = 3$  ...  $x_{75} = 75\%(2 \times 3) = 3$  a una partición recién agregada en un nodo físico en el que se encuentran los datos, y puede retener los datos correspondientes a  $x_{48} = 48\%(2 \times 3) = 0$ ,  $x_{54} = 54\%(2 \times 3) = 0$  ...  $x_{72} = 72\%(2 \times 3) = 0$  en una partición original. Alternativamente, el aparato de extensión de particiones puede migrar datos cuyo valor de módulo es 0 a una partición recién añadida en un nodo físico en el que se encuentran los datos, y retener los datos cuyo valor de módulo es 3 en una partición original. Por ejemplo, el aparato de extensión de particiones puede migrar datos correspondientes a  $x_{48} = 48\%(2 \times 3) = 0$ ,  $x_{54} = 54\%(2 \times 3) = 0$  ...  $x_{72} = 72\%(2 \times 3) = 0$  a una partición recién agregada en un nodo físico en el que se encuentran los datos, y puede retener los datos correspondientes a  $x_{45} = 45\%(2 \times 3) = 3$ ,  $x_{51} = 51\%(2 \times 3) = 3$ ,  $x_{57} = 57\%(2 \times 3) = 3$  ...  $x_{75} = 75\%(2 \times 3) = 3$  en una partición original.

25 Debe señalarse que, la partición 1 anterior se usa como un ejemplo. Cuando el tiempo n para la multiplicación en la cantidad de particiones es  $n = 3$ , los valores de módulo de los datos en la partición 1 son  $x_{45} = 45\%(3 \times 3) = 0$ ,  $x_{48} = 48\%(3 \times 3) = 3$ ,  $x_{51} = 51\%(3 \times 3) = 6$ ,  $x_{54} = 54\%(3 \times 3) = 0$ ,  $x_{57} = 57\%(3 \times 3) = 3$  ...  $x_{72} = 72\%(3 \times 3) = 0$  and  $x_{75} = 75\%(3 \times 3) = 3$ , respectivamente. El aparato de extensión de particiones puede migrar, de acuerdo con los valores de módulo 0, 3 y 6 de datos, los datos correspondientes a los valores de módulo. Un método para migrar, mediante el aparato de extensión de particiones, datos a una partición de destino es similar al método para multiplicar la cantidad de particiones por dos veces, y los detalles no se describen de nuevo en esta realización.

30 207. El aparato de extensión de particiones agrega una relación de correspondencia de mapeo entre una partición recién agregada y un nodo físico en el que la partición recién agregada está ubicada en una DHT pre-almacenada.

208. El aparato de extensión de particiones adquiere un valor medio de cantidades que se obtienen después de la extensión y son de particiones en todos los nodos físicos en el sistema de almacenamiento distribuido.

35 Después de realizar la multiplicación en la cantidad de particiones en todos los nodos físicos en el sistema de almacenamiento distribuido y de completar la migración de los datos correspondientes en las particiones, el aparato de extensión de particiones puede adquirir el valor medio de las cantidades que se obtienen después de la extensión y son de las particiones en todos los nodos físicos en el sistema de almacenamiento distribuido y determina, de acuerdo con el valor medio adquirido, si la cantidad de particiones en el sistema de almacenamiento distribuido cumple una cantidad de particiones requeridas por el sistema de almacenamiento distribuido para realizar la expansión al menos una vez.

45 Debe señalarse que el método para adquirir, mediante el aparato de extensión de particiones, el valor medio de las cantidades que se obtienen después de la extensión y son de las particiones en todos los nodos físicos en el sistema de almacenamiento distribuido es similar al método para adquirir, mediante el aparato de extensión de particiones, el valor medio de las cantidades de las particiones en todos los nodos físicos en el sistema de almacenamiento distribuido en esta realización, y los detalles no se describen de nuevo en esta realización.

209. El aparato de extensión de particiones determina si el valor medio de las cantidades de las particiones después de la extensión es menor que el umbral preestablecido; si el valor medio de las cantidades de las particiones después de la extensión es menor que el umbral preestablecido, realizar el paso 210; o si el valor medio de las cantidades de las particiones después de la extensión no es inferior al umbral preestablecido, realizar el paso 211.

50 210. El aparato de extensión de particiones realiza de nuevo una multiplicación en la cantidad de particiones en cada nodo físico, en donde los tiempos de multiplicación son los mismos.

Cuando el valor medio de las cantidades de las particiones después de la extensión es menor que el umbral preestablecido, el aparato de extensión de particiones puede realizar de nuevo la multiplicación en la cantidad de las particiones en cada nodo físico. Los tiempos de multiplicación para la multiplicación realizada de nuevo por el

aparato de extensión de particiones en la cantidad de particiones en cada nodo físico son los mismos que los tiempos de multiplicación para la multiplicación realizada por primera vez por el aparato de extensión de partición en la cantidad de particiones.

5 Después de que el aparato de extensión de particiones realice de nuevo la multiplicación en la cantidad de particiones en cada nodo físico, el método de esta realización puede incluir, además: agregar, mediante el aparato de extensión de particiones, la relación de mapeo entre la partición recién agregada y el nodo físico en el que la partición recién agregada está ubicada en la DHT pre-almacenada.

10 Debe señalarse que el aparato de extensión de particiones puede realizar la multiplicación en la cantidad de particiones en los nodos físicos al menos una vez, los tiempos de multiplicación cada vez son iguales y los tiempos para la multiplicación son mayores que o iguales a 2.

211. Fin del proceso de extensión de particiones.

15 Debe señalarse que el aparato de extensión de particiones puede realizar, de acuerdo con el valor medio de las cantidades de las particiones en los nodos físicos en el sistema de almacenamiento distribuido, la extensión en las particiones en los nodos físicos en el sistema de almacenamiento distribuido al menos una vez. Una cantidad de veces para la extensión de particiones no está limitada en esta realización de la presente invención. Un método para realizar, mediante el aparato de extensión de particiones, la extensión dos o más veces en las particiones en los nodos físicos en el sistema de almacenamiento distribuido es similar al método de extensión de particiones en esta realización, y los detalles no se describen de nuevo en esta realización.

20 Además, debe señalarse que el método de multiplicación proporcionado en esta realización no puede aplicarse sólo a un proceso de extensión de particiones del sistema de almacenamiento distribuido, sino que también se aplica a un proceso de multiplicación de un nodo de almacenamiento en otro sistema de almacenamiento, y un el alcance de aplicación del método de la presente invención no está limitado en esta realización.

25 En el método de extensión de particiones proporcionado en esta realización de la presente invención, se puede adquirir un valor medio de cantidades de particiones en todos los nodos físicos en un sistema de almacenamiento distribuido; si el valor medio de las cantidades de las particiones es menor que un umbral preestablecido, la multiplicación se puede realizar en una cantidad de particiones en cada nodo físico; y una relación de mapeo entre una partición recién agregada y un nodo físico en el que está ubicada la nueva partición agregada se puede agregar a una tabla DHT de hash distribuida pre-almacenada. Comparado con la técnica anterior en la cual la expansibilidad de un sistema de almacenamiento distribuido es deficiente debido a una pequeña cantidad de particiones en un  
30 nodo físico, en el método de extensión de particiones proporcionado en esta realización de la presente invención, si una cantidad de particiones es menor que un valor preestablecido, la multiplicación en la cantidad de particiones se puede realizar durante la expansión de la capacidad del sistema de almacenamiento distribuido, asegurando con ello la expansibilidad del sistema de almacenamiento distribuido.

### Realización 3

35 Esta realización de la presente invención proporciona un aparato de extensión de particiones. Tal como se muestra en la FIG. 3, el aparato incluye: una primera unidad de adquisición 31, una primera unidad de multiplicación 32 y una unidad de mapeo 33.

La primera unidad de adquisición 31 está configurada para adquirir un valor medio de cantidades de particiones en todos los nodos físicos en un sistema de almacenamiento distribuido.

40 La primera unidad de multiplicación 32 está configurada para: si el valor medio de las cantidades de las particiones es menor que un umbral preestablecido, realizar la multiplicación en una cantidad de particiones en cada nodo físico.

La unidad de mapeo 33 está configurada para agregar una relación de mapeo entre una partición recién agregada y un nodo físico en el que la partición recién agregada está ubicada en una tabla de hash distribuida DHT pre-almacenada.

45 Tal como se muestra en la FIG. 4, el aparato de extensión de particiones incluye, además: una segunda unidad de adquisición 34, una tercera unidad de adquisición 35 y una unidad de migración 36.

La segunda unidad de adquisición 34 está configurada para: después de que la primera unidad de multiplicación 32 realiza la multiplicación en la cantidad de las particiones en cada nodo físico, adquirir un valor hash de datos en cada partición en cada nodo físico.

La tercera unidad de adquisición 35 está configurada para adquirir un valor de módulo que se obtiene realizando una operación de módulo en el valor hash de los datos y una cantidad que se obtiene después de la extensión y es de particiones en el nodo físico en el que están ubicados los datos.

5 La unidad de migración 36 está configurada para determinar, de acuerdo con el valor de módulo de los datos, una partición de destino a la que han de migrar los datos, y si la partición de destino es diferente de la partición en la que están ubicados actualmente los datos, migrar los datos a la partición de destino.

La primera unidad de multiplicación 32 está configurada además para multiplicar la cantidad de las particiones en cada nodo físico por al menos dos veces la cantidad de las particiones.

10 El aparato de extensión de particiones incluye, además: una cuarta unidad de adquisición 37 y una segunda unidad de multiplicación 38.

15 La cuarta unidad de adquisición 37 está configurada para: después de que la unidad de mapeo 33 agregue la relación de mapeo entre la partición recién agregada y el nodo físico en el que la nueva partición agregada se encuentra en la tabla DHT de hash distribuida pre-almacenada, adquirir un valor medio de cantidades que se obtienen después de la extensión y son de particiones en todos los nodos físicos en el sistema de almacenamiento distribuido.

La segunda unidad multiplicadora 38 está configurada para: si el valor medio de las cantidades de las particiones después de la extensión es menor que el umbral preestablecido, realizar de nuevo la multiplicación de la cantidad de particiones en cada nodo físico, en que los tiempos para la multiplicación son cada vez los mismos.

20 La primera unidad de adquisición 31 está configurada, además, para adquirir periódicamente el valor medio de las cantidades de las particiones en todos los nodos físicos en el sistema de almacenamiento distribuido, o responder a una instrucción de activación de un usuario, para adquirir el valor medio de las cantidades de las particiones en todos los nodos físicos en el sistema de almacenamiento distribuido.

25 Debe señalarse que, para la descripción específica de algunos módulos funcionales en el aparato de extensión de particiones proporcionado en esta realización de la presente invención, se remite al contenido correspondiente en las realizaciones del método, y los detalles no se describen de nuevo en esta realización.

30 El aparato de extensión de particiones proporcionado en esta realización de la presente invención puede adquirir un valor medio de cantidades de particiones en todos los nodos físicos en un sistema de almacenamiento distribuido; si el valor medio de las cantidades de las particiones es menor que un umbral preestablecido, realizar la multiplicación en una cantidad de particiones en cada nodo físico; y agregar una relación de mapeo entre una partición recién agregada y un nodo físico en el que la partición recién agregada está ubicada en una tabla DHT de hash distribuida pre-almacenada. Comparado con la técnica anterior, en la que la expansibilidad de un sistema de almacenamiento distribuido es deficiente debido a una pequeña cantidad de particiones en un nodo físico, si una cantidad de particiones es menor que un valor preestablecido, el aparato de extensión de particiones proporcionado en esta realización de la presente invención puede realizar la multiplicación en la cantidad de las particiones durante la expansión de capacidad del sistema de almacenamiento distribuido, asegurando con ello la expansibilidad del sistema de almacenamiento distribuido.

#### Realización 4

Una realización de la presente invención proporciona un aparato de extensión de particiones. Tal como se muestra en la FIG. 5, el aparato puede incluir: un procesador 41 y una memoria 42.

40 El procesador 41 está configurado para adquirir un valor medio de cantidades de particiones en todos los nodos físicos en un sistema de almacenamiento distribuido; si el valor medio de las cantidades de las particiones es menor que un umbral preestablecido, realizar la multiplicación en una cantidad de particiones en cada nodo físico; y agregar una relación de mapeo entre una partición recién agregada y un nodo físico en el que la partición recién agregada está ubicada en una tabla DHT de hash distribuida pre-almacenada.

45 La memoria 42 está configurada para almacenar la relación de mapeo entre la partición recién agregada y el nodo físico en el que está ubicada la partición recién agregada, que es agregada por el procesador 41.

50 El procesador 41 está configurado, además, para adquirir un valor hash de datos en cada partición en cada nodo físico; adquirir un valor de módulo que se obtiene al realizar la operación de módulo en el valor hash de los datos y una cantidad que se obtiene después de la extensión y es de particiones en el nodo físico en el que están ubicados los datos; y determinar, de acuerdo con el valor de módulo de los datos, una partición de destino a la que han de migrar los datos, y si la partición de destino es diferente de la partición en la que están ubicados actualmente los datos, migrar los datos a la partición de destino.

El procesador 41 está configurado, además, para multiplicar la cantidad de las particiones en cada nodo físico al menos dos veces la cantidad de las particiones.

5 El procesador 41 está configurado para, además: después de que la relación de mapeo entre la nueva partición correspondiente a la partición y el nodo físico en el que se ubica la partición se agregue a DHT, adquirir un valor medio de las cantidades que se obtienen después de la extensión y son de particiones en todos los nodos físicos en el sistema de almacenamiento distribuido; y si el valor medio de las cantidades de las particiones después de la extensión es menor que el umbral preestablecido, realizar de nuevo la multiplicación en la cantidad de particiones en cada nodo físico, en que los tiempos para la multiplicación cada vez son los mismos.

10 El procesador 41 está configurado además para adquirir periódicamente el valor medio de las cantidades de las particiones en todos los nodos físicos en el sistema de almacenamiento distribuido, o responder a una instrucción de activación de un usuario, para adquirir el valor medio de las cantidades de las particiones en todos los nodos físicos en el sistema de almacenamiento distribuido.

15 Debe señalarse que, para la descripción específica de algunos módulos funcionales en el aparato de extensión de particiones proporcionado en esta realización de la presente invención, se remite al contenido correspondiente en las realizaciones del método, y los detalles no se describen de nuevo en esta realización.

20 El aparato de extensión de particiones proporcionado en esta realización de la presente invención puede adquirir un valor medio de cantidades de particiones en todos los nodos físicos en un sistema de almacenamiento distribuido; si el valor medio de las cantidades de las particiones es menor que un umbral preestablecido, realizar la multiplicación en una cantidad de particiones en cada nodo físico; y agregar una relación de asignación entre una partición recién agregada y un nodo físico en el que la partición recién agregada está ubicada en una tabla DHT de hash distribuida pre-almacenada. Comparado con la técnica anterior en la que la expansibilidad de un sistema de almacenamiento distribuido es deficiente debido a una pequeña cantidad de particiones en un nodo físico, si la cantidad de particiones es menor que un valor preestablecido, el aparato de extensión de particiones proporcionado en esta  
25 realización de la presente invención puede realizar la multiplicación en la cantidad de las particiones durante la expansión de capacidad del sistema de almacenamiento distribuido, asegurando con ello la expansibilidad del sistema de almacenamiento distribuido.

30 En base a las descripciones anteriores de las maneras de implementación, una persona experta en la técnica puede entender claramente que la presente invención se puede implementar mediante software además del hardware universal necesario o solo por hardware. En la mayoría de las circunstancias, la primera es una manera de implementación preferida. En base a tal entendimiento, las soluciones técnicas de la presente invención esencialmente o la parte que contribuye a la técnica anterior pueden implementarse en forma de un producto de software. El producto de software se almacena en un medio de almacenamiento legible, tal como un disquete, un disco duro o un disco óptico de un ordenador, e incluye varias instrucciones para instruir a un dispositivo informático (que puede ser un ordenador personal, un servidor o un dispositivo de red) para realizar los métodos descritos en las  
35 realizaciones de la presente invención.

40 Las descripciones anteriores son meramente realizaciones específicas de la presente invención, pero no están destinadas a limitar el alcance de protección de la presente invención. Cualquier variación o reemplazo que la persona experta en la técnica descubra fácilmente dentro del alcance técnico divulgado en la presente invención estará dentro del alcance de protección de la presente invención. Por lo tanto, el alcance de protección de la presente invención estará sujeto al alcance de protección de las reivindicaciones.

## REIVINDICACIONES

1. Un método de extensión de particiones, que comprende:

adquirir (101, 201) un valor medio de cantidades de particiones en todos los nodos físicos en un sistema de almacenamiento distribuido, en donde una partición es una parte de un espacio hash cuando el espacio hash se divide en varias partes iguales, y el espacio Hash es un espacio circular conectado de la cabeza a la cola;

realizar (102, 203) la multiplicación en una cantidad de particiones en cada nodo físico, si el valor medio de las cantidades de las particiones es menor que un umbral preestablecido, en donde el umbral preestablecido es un valor que satisface el valor medio de las cantidades de las particiones en todos los nodos físicos en el sistema de almacenamiento distribuido cuando la expansión de la capacidad se realiza al menos una vez para el sistema de almacenamiento distribuido;

adquirir (204) un valor hash de datos en cada partición en cada nodo físico;

adquirir (205) un valor de módulo que se obtiene al realizar la operación de módulo en el valor hash de los datos y una cantidad que se obtiene después de la extensión y es de particiones en el nodo físico en el que se ubican los datos; y

determinar (206), de acuerdo con el valor de módulo de los datos, una partición de destino a la que han de migrar los datos; y si la partición de destino es diferente de la partición en la que se encuentran actualmente los datos, migrar los datos a la partición de destino; y

agregar (103, 207) una relación de mapeo entre una partición recién agregada después de la multiplicación y un nodo físico en el que la partición recién agregada está ubicada en una tabla de hash distribuida DHT, pre-almacenada.

2. El método de extensión de particiones de acuerdo con la reivindicación 1, en el que realizar la multiplicación en una cantidad de particiones en cada nodo físico comprende:

multiplicar (210) la cantidad de las particiones en cada nodo físico por al menos dos veces la cantidad de las particiones.

3. El método de extensión de particiones de acuerdo con la reivindicación 1, en el que adquirir un valor medio de cantidades de particiones en todos los nodos físicos en un sistema de almacenamiento distribuido comprende:

adquirir periódicamente el valor medio de las cantidades de las particiones en todos los nodos físicos en el sistema de almacenamiento distribuido; o

responder a una instrucción de activación de un usuario, para adquirir el valor medio de las cantidades de las particiones en todos los nodos físicos en el sistema de almacenamiento distribuido.

4. Un aparato de extensión de particiones, que comprende:

una primera unidad de adquisición (31), configurada para adquirir un valor medio de las cantidades de las particiones en todos los nodos físicos en un sistema de almacenamiento distribuido, en donde una partición es una parte de un espacio hash cuando el espacio hash se divide en varias partes iguales, y el espacio Hash es un espacio circular conectado de la cabeza a la cola;

una primera unidad de multiplicación (32), configurada para: si el valor medio de las cantidades de las particiones es menor que un umbral preestablecido, realizar la multiplicación en una cantidad de particiones en cada nodo físico, en donde el umbral preestablecido es un valor que satisface el valor medio de las cantidades de las particiones en todos los nodos físicos en el sistema de almacenamiento distribuido cuando la expansión de la capacidad se realiza al menos una vez para el sistema de almacenamiento distribuido; y

una unidad de mapeo (33), configurada para agregar una relación de mapeo entre una partición recién agregada después de la multiplicación y un nodo físico, en que la partición recién agregada está ubicada en una tabla de hash distribuida DHT pre-almacenada;

una segunda unidad de adquisición (34), configurada para: después de que la primera unidad de multiplicación realiza la multiplicación en la cantidad de las particiones en cada nodo físico, adquirir un valor hash de datos en cada partición en cada nodo físico;

una tercera unidad de adquisición (35), configurada para adquirir un valor de módulo que se obtiene realizando una operación de módulo en el valor hash de los datos y una cantidad que se obtiene después de la extensión y es de particiones en el nodo físico en el que están ubicados los datos; y

una unidad de migración (36), configurada para determinar, de acuerdo con el valor de módulo de los datos, una partición de destino a la que han de migrar los datos, y si la partición de destino es diferente de la partición en la que están ubicados actualmente los datos, migrar los datos a la partición de destino.

5. El aparato de extensión de particiones de acuerdo con la reivindicación 4, en donde la primera unidad de multiplicación (31) está configurada además para multiplicar la cantidad de las particiones en cada nodo físico por al menos dos veces la cantidad de las particiones.
- 5 6. El aparato de extensión de particiones de acuerdo con la reivindicación 4, en donde la primera unidad de adquisición está configurada, además, para adquirir periódicamente el valor medio de las cantidades de las particiones en todos los nodos físicos en el sistema de almacenamiento distribuido, o responder a una instrucción de activación de un usuario, para adquirir el valor medio de las cantidades de las particiones en todos los nodos físicos en el sistema de almacenamiento distribuido.

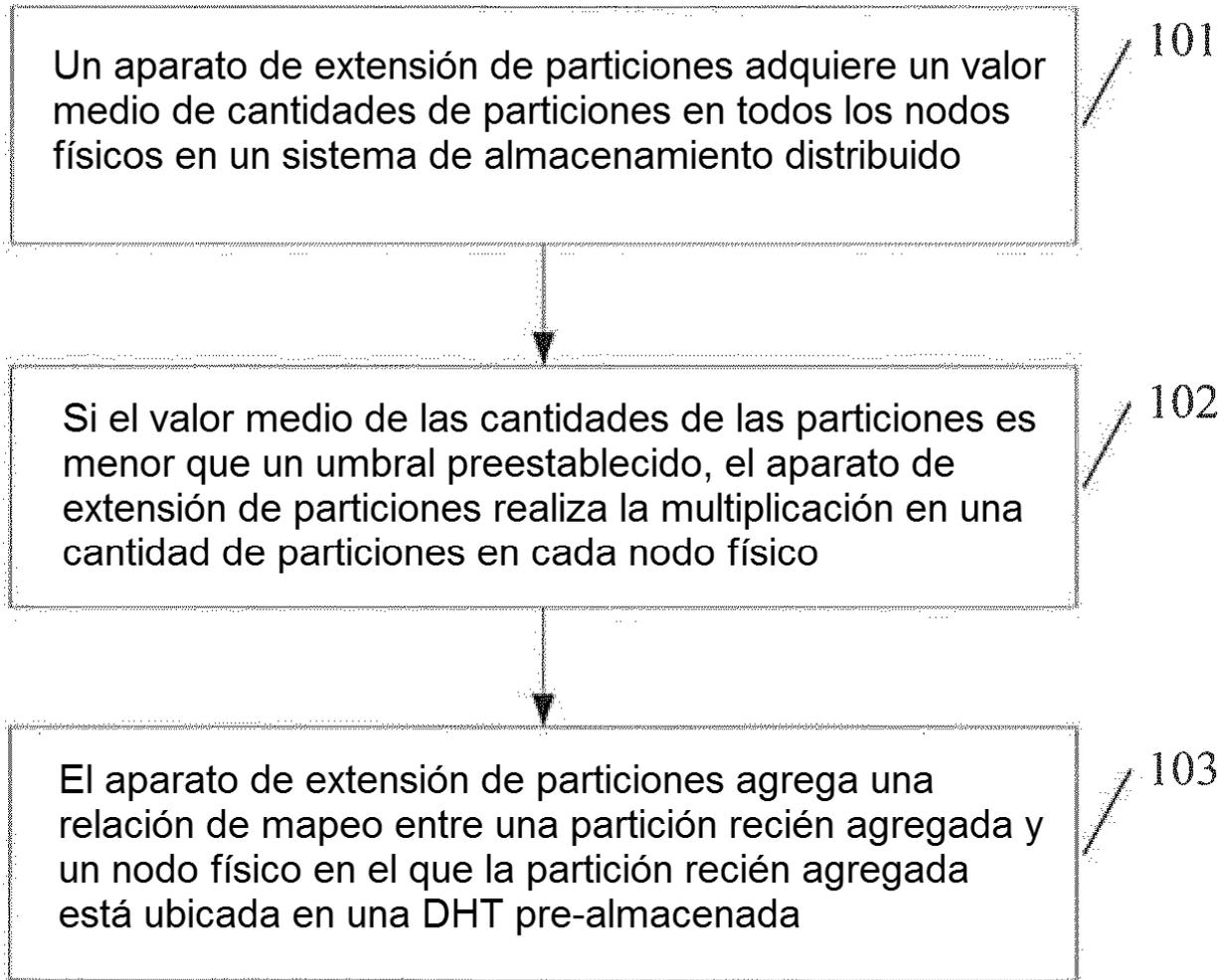


FIG. 1

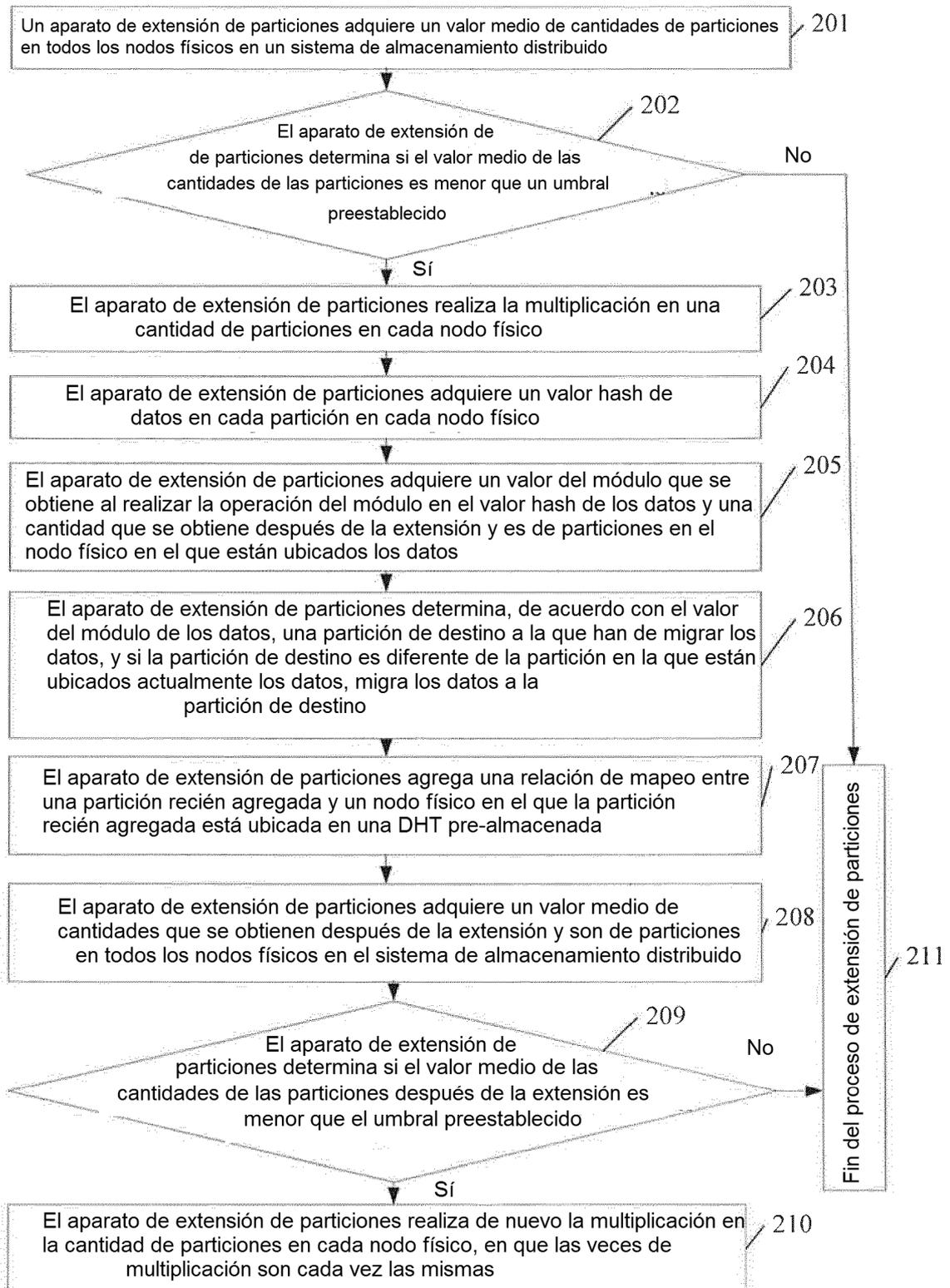


FIG. 2

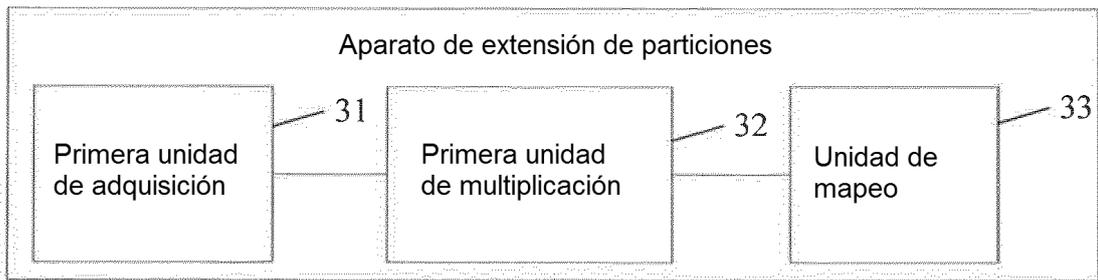


FIG. 3

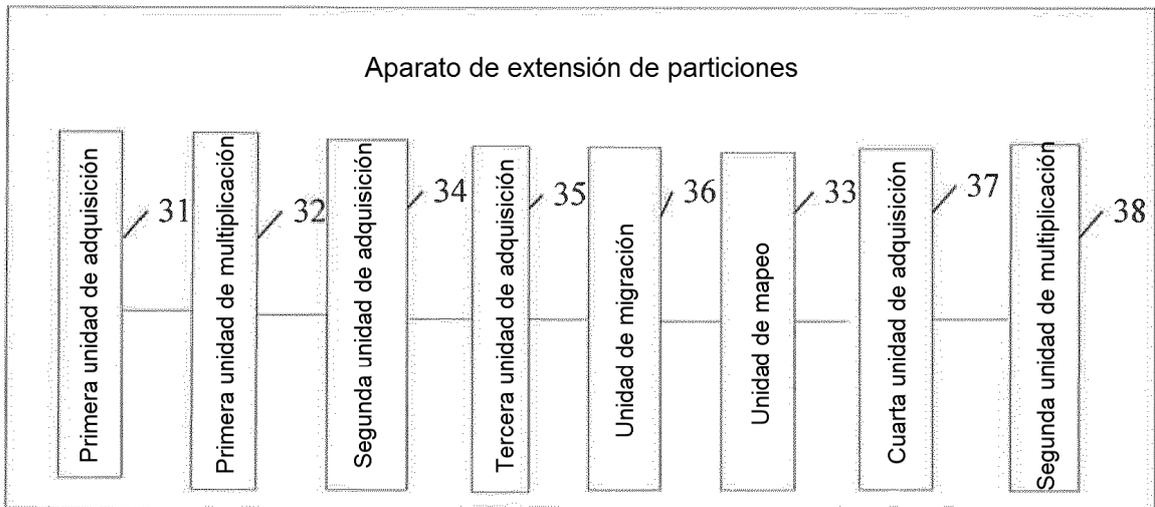


FIG. 4

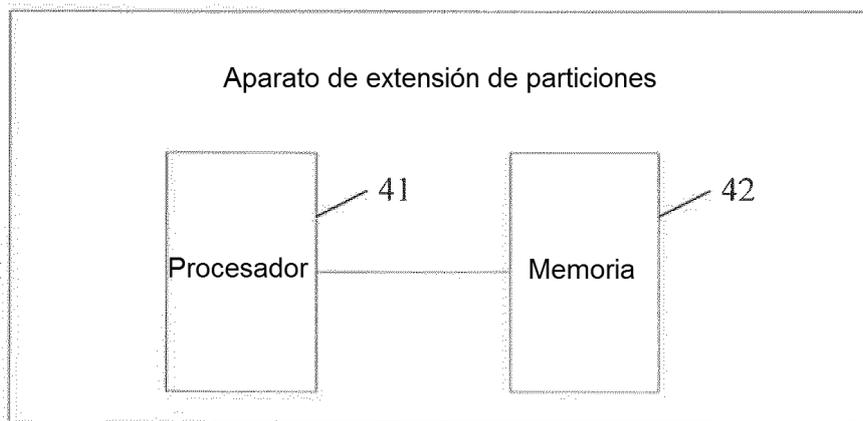


FIG. 5