

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 625 690**

51 Int. Cl.:

**G06F 17/30** (2006.01)

**G06F 3/06** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **12.12.2012 PCT/CN2012/086413**

87 Fecha y número de publicación internacional: **19.06.2014 WO14089767**

96 Fecha de presentación y número de la solicitud europea: **12.12.2012 E 12878640 (7)**

97 Fecha y número de publicación de la concesión europea: **22.02.2017 EP 2765524**

54 Título: **Método y dispositivo de procesamiento de datos en un sistema de clúster**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:  
**20.07.2017**

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

72 Inventor/es:  
**LIU, QIANG;  
SUN, QUANCHENG;  
LIU, XIAOBO;  
YOU, JUN;  
YANG, HUADI;  
ZHOU, DAN y  
HUANG, YAN**

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

ES 2 625 690 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 dispositivo de procesamiento de datos en un sistema de clúster.

Campo técnico

5 La presente invención se refiere a tecnologías de almacenamiento, y, en particular, a un método y aparato de procesamiento de datos en un sistema de clúster.

Antecedentes

10 A la deduplicación de datos (para abreviar, deduplicación) también se la denomina compresión inteligente o almacenamiento de instancia única, y constituye una tecnología de almacenamiento capaz de buscar datos duplicados automáticamente, de sólo reservar una copia única para datos iguales, y de utilizar un puntero que apunta a una copia única para reemplazar otras copias duplicadas, con el fin de cumplir requerimientos para eliminar datos redundantes y reducir capacidad de almacenamiento.

15 La deduplicación de datos de clúster (para abreviar, deduplicación de clúster) se refiere a una tecnología que organiza múltiples nodos físicos de deduplicación para mejorar el rendimiento y la capacidad de la deduplicación. En la tecnología de deduplicación de clúster de la técnica anterior, como regla general, un nodo físico que recibe un flujo de datos divide el flujo de datos en distintos bloques de datos, agrupa los bloques de datos obtenidos, y para cada grupo, realiza un muestreo de una parte de la información de metadatos a partir de información de metadatos de bloques de datos en el grupo y envía la parte de información de metadatos a todos los nodos físicos en un sistema de clúster para realizar una consulta; cada nodo físico en el sistema de clúster almacena un bloque de datos conocido e información de metadatos correspondiente, compara la información de metadatos muestreada con la información de metadatos almacenada en cada nodo físico, obtiene un nodo físico objetivo que tiene la mayor parte de bloques de datos duplicados de un resultado de consulta, y a continuación envía toda la información de bloques de datos de un grupo de datos correspondiente a la información de metadatos muestreada hacia el nodo físico objetivo para realizar una consulta de datos duplicados.

20 La patente de Estados Unidos US 2008/270729 describe una solución para enrutar un segmento hacia un nodo de clúster para deduplicación, cuando se recibe el flujo de datos, el flujo de datos se divide en segmentos, cada uno de los cuales incluye sub-segmentos; se obtiene una huella para cada sub-segmento del segmento; los nodos de clúster están asociados a un segmento basado en un cálculo realizado en el contenido del segmento; y el cliente envía el segmento y las huellas para sub-segmentos del segmento a un nodo de clúster asociado al segmento para realizar una deduplicación.

25 La técnica anterior presenta las siguientes desventajas: En la tecnología de deduplicación de clúster de la técnica anterior, es preciso enviar la información de metadatos muestreada a todos los nodos físicos para realizar una consulta, lo cual genera una gran cantidad de tiempos de interacción entre los nodos físicos en un proceso de deduplicación, y en el caso de que existan muchos nodos físicos en un sistema de clúster, cuando cada nodo físico realiza una deduplicación, aumenta la cantidad de cálculo con un incremento del número de nodos físicos en el sistema de clúster, lo cual da lugar a la degradación del rendimiento de deduplicación del sistema.

Compendio

La presente invención ofrece un método y aparato de procesamiento de datos en un sistema de clúster, con el fin de mejorar, de forma eficaz, el rendimiento de deduplicación del sistema de clúster.

40 Para conseguir el objetivo de la presente invención, como primer aspecto la presente invención ofrece un método de procesamiento de datos en un sistema de clúster, en el que el método incluye:

dividir un flujo de datos recibido por un nodo físico actual en m bloques de datos, obtener una huella de cada bloque de datos, y obtener n primeros valores de boceto que representan el flujo de datos de acuerdo con un primer algoritmo, donde m es un número entero mayor o igual que 1, y n es un número entero mayor o igual que 1 y menor o igual que m;

45 identificar, de acuerdo con un segundo algoritmo, un primer nodo físico que corresponde a cada primer valor de boceto en los n primeros valores de boceto y que está en el sistema de clúster, y enviar cada primer valor de boceto al correspondiente primer nodo físico para realizar una consulta;

50 recibir al menos un mensaje de respuesta de primeros nodos físicos correspondientes a los n primeros valores de boceto, y obtener una dirección de almacenamiento correspondiente al primer valor de boceto a partir del mensaje de respuesta;

obtener una primera dirección de almacenamiento a partir de todas las direcciones de almacenamiento recibidas, comparar una huella de un bloque de datos almacenado en una región de almacenamiento a la que apunta la primera dirección de almacenamiento con la huella de cada bloque de datos obtenida dividiendo el flujo de datos para obtener un bloque de datos no duplicados como un nuevo bloque de datos; y

obtener un segundo valor de boceto que representa el nuevo bloque de datos, identificar, de acuerdo con el segundo algoritmo, un segundo nodo físico correspondiente al segundo valor de boceto, almacenar la correspondencia entre el segundo valor de boceto y una dirección de almacenamiento del nuevo bloque de datos representada por el segundo valor de boceto en el segundo nodo físico.

5 En una posible primera manera del primer aspecto, obtener una primera dirección de almacenamiento a partir de todas las direcciones de almacenamiento recibidas incluye: obtener, del mensaje de respuesta, la dirección de almacenamiento correspondiente al primer valor de boceto, recoger estadísticas sobre la cantidad de accesos de cada dirección de almacenamiento que lleva el mensaje de respuesta, y seleccionar s direcciones de almacenamiento que tienen la mayor cantidad de accesos como la primera dirección de almacenamiento, donde s es mayor o igual que 1.

10 En una segunda posible manera, el método además incluye: de acuerdo con una política de almacenamiento pre-establecida, seleccionar una dirección de almacenamiento en la cual se escriben el nuevo bloque de datos y una huella del nuevo bloque de datos, y cuando se cumple una condición de almacenamiento pre-establecida, escribir el nuevo bloque de datos y la huella del nuevo bloque de datos en una región de almacenamiento a la que apunta la dirección de almacenamiento seleccionada.

15 En una posible tercera manera, y bajo la condición de que se cumpla una condición de almacenamiento pre-establecida, el método incluye: el nuevo bloque de datos es almacenado en una memoria caché del nodo físico actual, y cuando los datos en la memoria caché del nodo físico actual alcanzan un segundo umbral pre-establecido, se cumple la condición de almacenamiento pre-establecida.

20 Como segundo aspecto la presente invención ofrece un aparato de procesamiento de datos en un sistema de clúster, en el que el aparato incluye:

una unidad divisora, configurada para dividir un flujo de datos recibido por un nodo físico actual en m bloques de datos, y obtener una huella de cada bloque de datos, donde m es un número entero mayor o igual que 1;

25 una unidad de obtención de valor de boceto, configurada para obtener n primeros valores de boceto que representan el flujo de datos de acuerdo con un primer algoritmo donde n es un número entero mayor o igual que 1 y menor o igual que m;

una unidad de identificación, configurada para identificar, de acuerdo con un segundo algoritmo, un primer nodo físico que corresponde a cada valor de boceto en los n primeros valores de boceto y que está en el sistema de clúster;

30 una unidad de envío, configurada para enviar cada primer valor de boceto al correspondiente primer nodo físico para realizar una consulta;

una unidad de recepción, configurada para recibir un mensaje de respuesta del primer nodo físico y obtener una primera dirección de almacenamiento a partir del mensaje de respuesta;

35 una unidad de consulta, configurada para comparar una huella de un bloque de datos en una región de almacenamiento a la que apunta la primera dirección de almacenamiento con la huella de cada bloque de datos obtenida dividiendo el flujo de datos recibido para obtener un bloque de datos no duplicados como un nuevo bloque de datos; y

40 una unidad de actualización, configurada para obtener un segundo valor de boceto que representa el nuevo bloque de datos, identificar, de acuerdo con el segundo algoritmo, un segundo nodo físico correspondiente al segundo valor de boceto, y almacenar la correspondencia entre el segundo valor de boceto y una dirección de almacenamiento del nuevo bloque de datos representada por el segundo valor de boceto en el segundo nodo físico.

En una segunda posible manera de implementación, el aparato además incluye:

45 una unidad de obtención de dirección de escritura, configurada para: de acuerdo con una política de almacenamiento pre-establecida, seleccionar una dirección de almacenamiento en la cual se escriben el nuevo bloque de datos y una huella del nuevo bloque de datos; y

una unidad de escritura, configurada para: cuando se cumple una condición de almacenamiento pre-establecida, escribir el nuevo bloque de datos y la huella del nuevo bloque de datos en una región de almacenamiento correspondiente a la dirección de almacenamiento seleccionada.

50 En una tercera posible manera de implementación, la unidad de escritura está específicamente configurada para almacenar el nuevo bloque de datos en una memoria caché del nodo físico actual, y cuando los datos en la memoria caché del nodo físico actual alcanzan un segundo umbral pre-establecido, y se cumple la condición de almacenamiento pre-establecida, escribir el nuevo bloque de datos y la huella del nuevo bloque de datos en la región de almacenamiento correspondiente a la dirección de almacenamiento seleccionada.

Como otra forma de implementación del segundo aspecto, la presente invención además ofrece un aparato de procesamiento de datos en un sistema de clúster, donde el aparato incluye un procesador, una memoria, un bus, donde el procesador y la memoria se comunican entre sí mediante el bus;

la memoria está configurada para almacenar un programa; y

5 el procesador está configurado para ejecutar el programa en la memoria;

el procesador está configurado para ejecutar el programa en la memoria; donde

el programa incluye:

una unidad divisora, configurada para dividir un flujo de datos recibido por un nodo físico actual en  $m$  bloques de datos, y obtener una huella de cada bloque de datos, donde  $m$  es un número entero mayor o igual que 1;

10 una unidad de identificación, configurada para identificar, de acuerdo con un segundo algoritmo, un primer nodo físico que corresponde a cada valor de boceto en los  $n$  primeros valores de boceto y que está en el sistema de clúster;

una unidad de envío, configurada para enviar cada primer valor de boceto al correspondiente primer nodo físico para realizar una consulta;

15 una unidad de recepción, configurada para recibir un mensaje de respuesta del primer nodo físico y obtener una primera dirección de almacenamiento a partir del mensaje de respuesta;

una unidad de consulta, configurada para comparar una huella de un bloque de datos en una región de almacenamiento a la que apunta la primera dirección de almacenamiento con la huella de cada bloque de datos obtenida dividiendo el flujo de datos recibido, y realizar una consulta a un bloque de datos duplicados, donde un  
20 bloque de datos no duplicados obtenido se utiliza como un nuevo bloque de datos; y

una unidad de actualización, configurada para obtener al menos un segundo valor de boceto que representa el nuevo bloque de datos, identificar, de acuerdo con el segundo algoritmo, un segundo nodo físico correspondiente a cada segundo valor de boceto en el segundo valor de boceto, y almacenar la correspondencia entre cada segundo valor de boceto y una dirección de almacenamiento de un bloque de datos representada por cada segundo valor de  
25 boceto en el correspondiente segundo nodo físico.

En esta forma de implementación, la unidad de recepción puede incluir

una subunidad de recogida de estadísticas, configurada para obtener una dirección de almacenamiento correspondiente al primer valor de boceto a partir del mensaje de respuesta, y recoger estadísticas sobre el número de accesos de cada dirección de almacenamiento que lleva el mensaje de respuesta; y

30 una subunidad de obtención de dirección, configurada para seleccionar, de acuerdo con un resultado estadístico de la subunidad de recogida de estadísticas,  $s$  direcciones de almacenamiento que tienen el mayor número de accesos como la primera dirección de almacenamiento, donde  $s$  es mayor o igual que 1.

En combinación con el segundo aspecto y la primera posible manera del segundo aspecto, en una posible segunda manera de implementación, el aparato además incluye:

35 una unidad de obtención de dirección de escritura, configurada para: de acuerdo con una política de almacenamiento pre-establecida, seleccionar una dirección de almacenamiento en la cual se escriben el nuevo bloque de datos y una huella del nuevo bloque de datos; y

40 una unidad de escritura, configurada para: cuando se cumple una condición de almacenamiento pre-establecida, escribir el nuevo bloque de datos y la huella del nuevo bloque de datos en una región de almacenamiento correspondiente a la dirección de almacenamiento seleccionada.

Asimismo, en esta forma de implementación según el segundo aspecto de la presente invención, la unidad de escritura puede estar específicamente configurada para almacenar el nuevo bloque de datos en una memoria caché del nodo físico actual, y cuando los datos en la memoria caché del nodo físico actual alcanzan un segundo umbral pre-establecido, y se cumple la condición de almacenamiento pre-establecida, escribir el nuevo bloque de datos y la  
45 huella del nuevo bloque de datos en la región de almacenamiento correspondiente a la dirección de almacenamiento seleccionada.

Como tercer aspecto, la presente invención ofrece un producto de programa informático para almacenamiento de datos, donde el producto de programa informático incluye un medio de almacenamiento legible por ordenador que almacena un código de programa, y una instrucción incluida en el código de programa se utiliza para:

dividir un flujo de datos recibido por un nodo físico actual en  $m$  bloques de datos, obtener una huella de cada bloque de datos, y obtener, de acuerdo con un primer algoritmo,  $n$  primeros valores de boceto que representan el flujo de datos, donde  $m$  es un número entero mayor o igual que 1, y  $n$  es un número entero mayor o igual que 1 y menor o igual que  $m$ ;

- 5 identificar, de acuerdo con un segundo algoritmo, un primer nodo físico que corresponde a cada primer valor de boceto en los  $n$  primeros valores de boceto y está en el sistema de clúster, y enviar cada primer valor de boceto al correspondiente primer nodo físico para realizar una consulta;

- 10 recibir al menos un mensaje de respuesta de primeros nodos físicos correspondientes a los  $n$  primeros valores de boceto, y obtener una dirección de almacenamiento correspondiente al primer valor de boceto a partir del mensaje de respuesta;

- 15 obtener una primera dirección de almacenamiento a partir de todas las direcciones de almacenamiento recibidas, comparar una huella de un bloque de datos almacenado en una región de almacenamiento a la que apunta la primera dirección de almacenamiento con la huella de cada bloque de datos obtenida dividiendo el flujo de datos, y realizar una consulta a un bloque de datos duplicados, donde un bloque de datos no duplicados obtenido se utiliza como un nuevo bloque de datos; y

obtener al menos un segundo valor de boceto que representa el nuevo bloque de datos, identificar, de acuerdo con el segundo algoritmo, un segundo nodo físico correspondiente a cada segundo valor de boceto en el segundo valor de boceto, y almacenar la correspondencia entre cada segundo valor de boceto y una dirección de almacenamiento de un bloque de datos representada por cada segundo valor de boceto en el correspondiente segundo nodo físico.

- 20 De acuerdo con la presente invención, cuando se realiza una consulta de datos duplicados en el flujo de datos recibidos, el primer nodo físico que corresponde a cada primer valor de boceto y que está en el sistema de clúster se identifica de acuerdo con el primer valor de boceto que representa el flujo de datos y, a continuación, el primer valor de boceto que representa el flujo de datos se envía al nodo físico identificado para realizar la consulta de datos duplicados, la cantidad de nodos físicos identificados para realizar la consulta de datos duplicados no aumenta con un incremento de la cantidad de nodos en el sistema de clúster; por lo tanto, una cantidad de cálculo de cada nodo no aumenta con un incremento del número de nodos en el sistema de clúster, y el rendimiento de deduplicación del sistema de clúster se mejora de manera eficaz.

#### Breve descripción de las figuras

- 30 Para describir las soluciones técnicas de la presente invención de manera más clara, a continuación se describen brevemente los dibujos que acompañan esta memoria y que describen realizaciones de la presente invención o de la técnica anterior. Al parecer, los dibujos que acompañan la presente invención en la siguiente descripción muestran algunas realizaciones de la presente invención únicamente.

La Fig. 1 es un diagrama de flujo de una realización de un método de procesamiento de datos en un sistema de clúster de acuerdo con la presente invención;

- 35 La Fig. 2 es un diagrama esquemático de muestreo y división de bloques realizado por un nodo físico después de que se ha recibido un flujo de datos según una realización de la presente invención;

La Fig. 3 es un diagrama esquemático de una estructura interna de un nodo físico según una realización de la presente invención;

- 40 La Fig. 4 es un diagrama de flujo de una realización de otro método de procesamiento de datos en un sistema de clúster de acuerdo con la presente invención;

La Fig. 5 es un diagrama de flujo de una realización de otro método de procesamiento de datos en un sistema de clúster de acuerdo con la presente invención;

La Fig. 6 es un diagrama esquemático de una estructura interna de otro nodo físico según una realización de la presente invención;

- 45 La Fig. 7 es un diagrama de flujo de una realización de otro método de procesamiento de datos en un sistema de clúster de acuerdo con la presente invención;

La Fig. 8 es un diagrama esquemático de migración de datos según una realización de la presente invención;

La Fig. 9 es un diagrama estructural de un aparato de procesamiento de datos en un sistema de clúster según una realización de la presente invención; y

- 50 La Fig. 10 es un diagrama estructural de otro aparato de procesamiento de datos en un sistema de clúster según una realización de la presente invención.

Descripción de las realizaciones

5 Con el propósito de esclarecer los objetivos, soluciones técnicas y ventajas de la presente invención, a continuación se describen con claridad las soluciones técnicas de la presente invención con referencia a los dibujos de realizaciones de dicha presente invención que la acompañan. Aparentemente, las realizaciones que han de ser descritas son simplemente algunas pero no todas las realizaciones de la presente invención.

Las realizaciones de la presente invención pueden aplicarse en un sistema de clúster para almacenamiento, donde el sistema de clúster incluye múltiples nodos físicos. Un nodo físico que posee un motor de deduplicación puede utilizarse como materia de ejecución de las realizaciones de la presente invención, y lleva a cabo el método en las realizaciones de la presente invención después de realizar una tarea de deduplicación.

10 La Fig. 1 es un diagrama de flujo de una realización de un método de procesamiento de datos en un sistema de clúster de acuerdo con la presente invención, y tal y como se muestra en la Fig. 1, el método en esta realización puede incluir:

Etapa 10: Dividir un flujo de datos recibido por un nodo físico actual en  $m$  bloques de datos y obtener una huella de cada bloque de datos, donde  $m$  es un número entero mayor o igual que 1.

15 En el sistema de clúster, un nodo físico puede utilizarse como un nodo para realizar una tarea de deduplicación de clúster siempre y cuando se instale en el nodo físico un aparato para llevar a cabo el método de procesamiento de datos en la realización de la presente invención.

Después de recibir el flujo de datos, el nodo físico actual divide los datos en el flujo de datos en  $m$  bloques de datos y obtiene la huella de cada bloque de datos.

20 Etapa 11: Obtener, de acuerdo con un primer algoritmo,  $n$  primeros valores de boceto que representan el flujo de datos, donde  $n$  es un número entero mayor o igual que 1 y menor o igual que  $m$ .

25 Los  $n$  primeros valores de boceto que representan el flujo de datos se obtienen mediante un algoritmo pre-establecido, y para facilitar la descripción, a un algoritmo para obtener un primer valor de boceto que representa el flujo de datos se lo denomina el primer algoritmo, y a un valor de boceto que representa el flujo de datos recibidos se lo denomina el primer valor de boceto.

30 El primer algoritmo no se limita en la realización de la presente invención siempre y cuando con el primer algoritmo se pueda obtener el primer valor de boceto que representa el flujo de datos. Por ejemplo, los bloques de datos obtenidos mediante división son muestreados directamente, y se calcula un valor de boceto de los  $n$  bloques de datos muestreados, y el valor de boceto calculado se utiliza como un valor de boceto que representa el flujo de datos recibidos.

Opcionalmente, para hacer que los datos muestreados sean más uniformes, en la realización de la presente invención, obtener los  $n$  primeros valores de boceto que representan el flujo de datos de acuerdo con el primer algoritmo puede incluir las siguientes etapas:

35 Las huellas de los bloques de datos obtenidos de acuerdo con el flujo de datos obtenidos se agrupan en  $n$  primeros grupos de huellas, en donde cada primer grupo de huellas incluye huellas de al menos dos bloques de datos.

Para cada primer grupo de huellas, se obtiene un primer valor de boceto que representa cada grupo de huellas, con el fin de obtener los  $n$  primeros valores de boceto.

40 Si  $n$  es igual que 1, indica que las huellas correspondientes a los bloques de datos obtenidos a partir del flujo de datos recibidos se agrupan en un primer grupo de huellas, y una huella se muestrea directamente como primer valor de boceto del primer grupo de huellas a partir de las huellas correspondientes a los bloques de datos obtenidos mediante división.

En la realización de la presente invención, se pueden agrupar múltiples bloques de datos consecutivos en un primer grupo de huellas y, seguramente, existen múltiples métodos para dividir bloques de datos con el fin de obtener un primer grupo de huellas, y no están limitados en esta realización.

45 En la realización de la presente invención, si un primer valor de boceto se obtiene dividiendo un flujo de datos en bloques de datos y a continuación agrupando las huellas correspondientes a los bloques de datos, de la misma manera, también pueden existir múltiples métodos para obtener un primer valor de boceto a partir de cada grupo de huellas siempre y cuando el primer valor de boceto pueda representar un correspondiente primer grupo de huellas. Por ejemplo, puede seleccionarse como el primer valor de boceto una huella mayor entre las huellas de los bloques de datos, o se selecciona aleatoriamente información de huellas de uno o más bloques de datos como el primer valor de boceto a partir de un primer grupo de huellas, lo cual no está limitado en la realización de la presente invención.

50

En esta realización, también se puede seleccionar un valor de una huella más pequeña, es decir, la más pequeña, a partir de información de huellas de todos los bloques de datos en el primer grupo de huellas, como un primer valor de boceto que representa un grupo de huellas al cual pertenece la huella más pequeña.

5 Etapa 12: Identificar, de acuerdo con un segundo algoritmo, un primer nodo físico que corresponde a cada primer valor de boceto en los n primeros valores de boceto y que está en el sistema de clúster.

10 En la realización de la presente invención, para facilitar la descripción, a un nodo físico correspondiente al primer valor de boceto identificado se lo denomina el primer nodo físico, y a un algoritmo para identificar el primer nodo físico se lo denomina el segundo algoritmo; y si un primer valor de boceto corresponde a un primer nodo físico, los n primeros valores de boceto pueden corresponder a n primeros nodos físicos, a lo cual la realización de la presente invención definitivamente no está limitada.

También puede haber múltiples segundos algoritmos. Por ejemplo, la extracción de raíces y operaciones de redondeo se realizan sobre el primer valor de boceto hasta que el valor es menor que m, y se redondea un resultado final para obtener el correspondiente primer nodo físico.

15 En esta realización, el segundo algoritmo adoptado puede además ser: Se realiza una operación de módulo sobre el total de nodos físicos en el sistema de clúster por el primer valor de boceto, y se obtiene el primer nodo físico que corresponde al primer valor de boceto y que está en el sistema de clúster.

Etapa 13: Enviar cada primer valor de boceto al primer nodo físico correspondiente para realizar una consulta, recibir al menos un mensaje de respuesta de los primeros nodos físicos correspondientes a los n primeros valores de boceto, y obtener una primera dirección de almacenamiento a partir del mensaje de respuesta.

20 En la realización de la presente invención, en el sistema de clúster, cada uno o la mayoría de los nodos físicos almacenan una tabla de índices, donde la correspondencia entre un valor de boceto y una dirección de almacenamiento de un bloque de datos almacenado representada por el valor de boceto se almacena en la tabla de índices, y la tabla de índices se almacena en un nodo físico en el sistema de clúster según una política pre-establecida, y un bloque de datos y la información de huella correspondiente al bloque de datos están almacenados en una región de almacenamiento a la que apunta una dirección de almacenamiento diferente.

25 Cabe destacar que una tabla de índices en cada nodo físico puede ser una tabla de índices completa e incluye correspondencia entre todos los valores de boceto de datos almacenados y direcciones de almacenamiento de bloques de datos almacenados representados por los valores de boceto en el sistema de clúster, y una parte de una tabla de índices completa también puede almacenarse en cada nodo físico de acuerdo con una política pre-establecida, con el fin de reducir una cantidad de cálculo para consultar a un bloque duplicado en un nodo físico. Es posible adoptar esta última manera en esta realización. A continuación se describe en detalle cómo almacenar, en un nodo físico, cada entrada de índice en la tabla de índices.

30 En la realización de la presente invención, la tabla de índices se almacena en un nodo físico según una política establecida, no hay una relación directa entre los datos en una tabla de índices de un nodo físico específico y datos en el nodo físico, y la tabla de índices está solo almacenada en el nodo físico específico según una política de asignación pre-establecida.

35 En una implementación específica, puesto que una región de almacenamiento correspondiente a una dirección de almacenamiento puede tener datos de múltiples grupos, y si se selecciona un valor de boceto de cada grupo, se puede dar un caso en el cual una dirección de almacenamiento corresponda a múltiples valores de boceto diferentes; por lo tanto, una misma dirección de almacenamiento en la tabla de índices puede corresponder a múltiples valores de boceto diferentes; no obstante, un mismo valor de boceto corresponde a una dirección de almacenamiento. Cuando se consultan múltiples primeros valores de boceto en la tabla de índices, se pueden obtener múltiples direcciones de almacenamiento correspondientes, y cuando se selecciona la primera dirección de almacenamiento a partir de una dirección de almacenamiento correspondiente al primer valor de boceto, se puede pre-establecer una regla para seleccionar la primera dirección de almacenamiento o ésta puede ser establecida por un usuario de acuerdo con una situación real, lo cual no está limitado en esta realización.

40 Por ejemplo, es posible obtener la dirección de almacenamiento correspondiente al primer valor de boceto a partir del mensaje de respuesta, se recogen estadísticas sobre la cantidad de accesos de cada dirección de almacenamiento que lleva el mensaje de respuesta, y se seleccionan s direcciones de almacenamiento que tienen la mayor cantidad de accesos como la primera dirección de almacenamiento, donde s es mayor o igual que 1.

45 En una implementación específica, cuando se recibe una solicitud de consulta que lleva el primer valor de boceto, el primer nodo físico coincide con un valor de boceto igual que el primer valor de boceto de acuerdo con una tabla de índices almacenada localmente, y transmite como respuesta, a través de un mensaje de respuesta, una dirección de almacenamiento correspondiente al valor de boceto coincidente a un nodo físico que envía la solicitud de consulta. Cada primer valor de boceto corresponde a un primer nodo físico; por lo tanto, múltiples primeros nodos físicos reciben una solicitud de consulta, y realizan localmente una consulta de acuerdo con la solicitud de consulta. Si se encuentra la dirección de almacenamiento mediante la consulta, la dirección de almacenamiento se devuelve a un

nodo que envía la solicitud de consulta, y el nodo que envía la solicitud de consulta puede recibir múltiples direcciones de almacenamiento devueltas por primeros nodos físicos diferentes, y si la dirección de almacenamiento no se encuentra mediante la consulta, se transmite como respuesta un valor nulo, por ejemplo, 0 o no se transmite ninguna respuesta, y un usuario puede establecer cómo representar específicamente que la dirección de almacenamiento no se ha encontrado mediante la consulta.

En un caso en el que la dirección de almacenamiento transmitida como respuesta por el primer nodo físico no se recibe después de que se envía cada primer valor de boceto al correspondiente primer nodo físico para realizar una consulta, un bloque de datos obtenido mediante la división del flujo de datos se utiliza como un bloque de datos no duplicados, y entonces se lleva a cabo la etapa 15.

Etapa 14: Comparar una huella de un bloque de datos en una región de almacenamiento a la que apunta la primera dirección de almacenamiento con la huella de cada bloque de datos obtenida dividiendo el flujo de datos recibido, y realizar una consulta a un bloque de datos duplicados, donde un bloque de datos no duplicados obtenido se utiliza como un nuevo bloque de datos.

En una implementación específica, el nodo físico actual almacena bloques de datos e información de huellas después de procesar el flujo de datos recibidos en una memoria caché, y puede adoptar múltiples métodos para comparar las huellas de los bloques de datos. Por ejemplo, en el método 1, un bloque de datos en la región de almacenamiento correspondiente a la primera dirección de almacenamiento se carga en la memoria caché del nodo físico actual, y se realiza una comparación con la huella de cada bloque de datos correspondiente al flujo de datos recibidos para realizar una consulta de bloque duplicado; en el método 2, se envía una instrucción de consulta al nodo físico al que apunta la primera dirección de almacenamiento, donde las huellas de los m datos obtenidos mediante la división del flujo de datos se llevan en la instrucción de consulta, y se recibe un resultado de consulta devuelto por el nodo físico al que apunta la primera dirección de almacenamiento. En el método 1, la información de huellas almacenada en la memoria caché del nodo actual puede utilizarse en una consulta de datos duplicados, lo cual puede mejorar aún más una velocidad de deduplicación.

Cuando se inicializa una entrada de índice en la tabla de índices, un usuario puede también pre-almacenar, en la tabla de índices y de acuerdo con datos almacenados en el sistema, la correspondencia entre un valor de boceto conocido y una dirección de almacenamiento de un bloque de datos representada por el valor de boceto conocido, y no se limita una manera específica en la realización de la presente invención. La tabla de índices puede también ser nula durante la inicialización, y durante la consulta de datos duplicados, cuando un nuevo bloque de datos obtenido se almacena en un sistema de almacenamiento, se insertan en la tabla de índices un valor de boceto obtenido y una dirección de almacenamiento del nuevo bloque de datos representada por el valor de boceto, con el fin de que la tabla de índices se actualice constantemente; por lo tanto, la realización de la presente invención además incluye:

Etapa 15: Obtener al menos un segundo valor de boceto que representa el nuevo bloque de datos, identificar, de acuerdo con el segundo algoritmo, un segundo nodo físico correspondiente a cada segundo valor de boceto en el segundo valor de boceto, y almacenar la correspondencia entre cada segundo valor de boceto y una dirección de almacenamiento de un bloque de datos representada por cada segundo valor de boceto en el correspondiente segundo nodo físico.

En la realización de la presente invención, el bloque de datos no duplicados encontrado mediante la consulta se utiliza como el nuevo bloque de datos, y cuando el nuevo bloque de datos que se encuentra mediante la consulta alcanza cierto tamaño y precisa ser almacenado, es necesario indicar en la tabla de índices una ubicación de almacenamiento del nuevo bloque de datos en el sistema; por lo tanto, la tabla de índices necesita ser constantemente actualizada. Cuando se obtiene la dirección de almacenamiento para el nuevo bloque de datos, se obtiene un valor de boceto que representa el nuevo bloque de datos, y se denomina segundo valor de boceto para facilitar la descripción, el segundo nodo físico se identifica de acuerdo con el segundo algoritmo, y la correspondencia entre cada segundo valor de boceto y una dirección de almacenamiento de un bloque de datos representada por cada segundo valor de boceto se almacena en el correspondiente segundo nodo físico, de este modo implementándose una actualización constante de una entrada de índice en el sistema de clúster.

Un método para obtener el segundo valor de boceto puede ser el mismo que un algoritmo para obtener el primer valor de boceto. Por ejemplo, puede adoptarse el primer algoritmo, y en un caso en el que se asegura que un conjunto de primeros valores de boceto es un subconjunto de un conjunto de segundos valores de boceto, o un conjunto de segundos valores de boceto es un conjunto de primeros valores de boceto, también se pueden adoptar otros algoritmos.

La solución en esta realización se describe en detalle a continuación con referencia a un ejemplo de una manera de implementación específica. La Fig. 2 es un diagrama esquemático de muestreo y división de bloques realizado por un nodo físico actual después de que se ha recibido un flujo de datos. En esta realización, múltiples bloques de datos pueden considerarse como un súper bloque de datos (súper fragmento); por lo tanto, en una implementación específica, para una agrupación más precisa, se agrupan huellas correspondientes a bloques de datos, un bloque de datos correspondiente a huellas en un grupo es un súper bloque de datos, y se obtiene un primer valor de boceto a partir de cada grupo de huellas, y el primer valor de boceto también puede denominarse un valor de boceto de súper

bloque de datos (ID de súper fragmento, SID), y seguramente, también se pueden agrupar directamente bloques de datos obtenidos mediante división de flujo de datos. Con referencia a la Figura 2, después de recibir el flujo de datos, el nodo físico actual divide los datos recibidos en bloques de datos, diversos bloques de datos consecutivos se agrupan en un grupo, es decir, un súper fragmento en el dibujo que acompaña la presente invención, y una manera de agrupación incluye una manera como, por ejemplo, dividir bloques de longitud variable o dividir bloques de longitud fija, lo cual no está limitado en esta realización. Un bloque de datos se obtiene a partir de cada súper fragmento, y una huella correspondiente al bloque de datos obtenido se utiliza como un SID que representa a cada súper fragmento.

Al tomar un nodo físico en un sistema de clúster como ejemplo, se puede hacer referencia a la Fig. 3 en la que se representa un diagrama esquemático de ejemplo de una estructura interna del nodo físico. Se pueden incluir una memoria caché y una región de almacenamiento en cada nodo físico, y en una implementación específica, con el fin de mejorar el rendimiento de la consulta, se almacena una tabla de índices en la memoria caché del nodo físico, donde la correspondencia entre un segundo valor de boceto y una dirección de almacenamiento de un bloque de datos almacenado representada por el segundo valor de boceto se almacena en la tabla de índices. Tal y como se describe más arriba, una tabla de índices en cada nodo físico puede ser una tabla de índices completa e incluye la correspondencia entre todos los segundos valores de boceto y direcciones de almacenamiento de huellas de bloques de datos almacenados representadas por los segundos valores de boceto en el sistema de clúster, y una parte de una tabla de índices completa también puede almacenarse en cada nodo físico de acuerdo con una política pre-establecida. En una implementación específica, con el fin de reducir el uso de una memoria, es posible adoptar esta última manera, y una parte de una tabla de índices completa se almacena en cada nodo físico de acuerdo con una política pre-establecida. Por ejemplo, tomando un valor de boceto A como ejemplo, un nodo físico se identifica realizando una operación de módulo sobre el total de nodos físicos en el sistema de clúster por A, y la correspondencia entre A y una dirección de almacenamiento de un bloque de datos almacenado representada por A se almacena en el nodo físico identificado; y durante una consulta de datos duplicados, se realiza una operación de módulo sobre el total de nodos físicos en el sistema de clúster por un primer valor de boceto obtenido con el fin de obtener un primer nodo físico correspondiente al primer valor de boceto.

Cada nodo físico incluye un aparato de almacenamiento, que permite que cada nodo físico tenga una función de almacenamiento de datos durante un largo plazo, el aparato de almacenamiento puede ser un disco magnético y también puede ser otro aparato de almacenamiento, por ejemplo, una tarjeta SSD, y el aparato de almacenamiento en cada nodo físico se denomina un repositorio de instancia única (SIR, repositorio de instancia única, por su sigla en inglés), y en la Figura FIG. 3, se toma como ejemplo un disco magnético. Puede haber diversas regiones de almacenamiento en un aparato de almacenamiento en un nodo físico, y en una implementación específica, cada región de almacenamiento puede considerarse de manera figurada como un contenedor (contenedor) para almacenar datos, cada contenedor de almacenamiento tiene un único número de serie en el sistema de clúster, que puede denominarse un número de serie de contenedor de almacenamiento (ID de contenedor, CID), y el número de serie de contenedor indica una ubicación del contenedor de almacenamiento en el sistema de clúster, por ejemplo, una región de almacenamiento en un nodo físico en el sistema de clúster. Por lo tanto, en una implementación específica, la dirección de almacenamiento precedente del bloque de datos almacenados se expresa como un CID, que indica en qué región de almacenamiento y nodo físico se almacena el bloque de datos, y en una implementación específica, la correspondencia entre un valor de boceto y una dirección de almacenamiento de un bloque de datos almacenados representada por el valor de boceto en la tabla de índices precedente puede indicarse como correspondencia entre un SID y un CID, y además de un bloque de datos de almacenamiento, es posible almacenar adicionalmente en cada región de almacenamiento información de huellas correspondiente a un bloque de datos. Adicionalmente es posible incluir una región de almacenamiento en búfer (búfer de contenedor) en una memoria caché de cada nodo físico y se configura para almacenar temporalmente un nuevo bloque de datos obtenido mediante identificación.

En la técnica anterior, se envía información de huellas muestreada a la totalidad de los nodos para realizar una consulta, y, de esta forma, en un caso en el cual aumenta el número de nodos en el sistema de clúster, la cantidad de cálculo se incrementa cada vez más. Un experimento muestra que, cuando la cantidad de nodos en un sistema de clúster excede 16, el rendimiento de la deduplicación del sistema se degrada en gran medida. En la realización de la presente invención, a diferencia de la técnica anterior en la que un valor de boceto muestreado se envía a la totalidad de los nodos para realizar una consulta, cuando se realiza una consulta de datos duplicados en un flujo de datos recibidos, el primer nodo físico que corresponde a cada primer valor de boceto y que está en el sistema de clúster se identifica de acuerdo con el primer valor de boceto que representa el flujo de datos y, a continuación, el primer valor de boceto que representa el flujo de datos se envía al nodo físico identificado para realizar la consulta de datos duplicados, y un procedimiento de la consulta de datos duplicados no varía con un incremento del número de nodos en el sistema de clúster; por lo tanto, la cantidad de cálculo de cada nodo no aumenta con un incremento del número de nodos en el sistema de clúster.

Con referencia a la Figura 4, la presente invención también ofrece una realización de otro método de procesamiento de datos de clúster, y difiere respecto del método de procesamiento de datos en un sistema de clúster correspondiente a la Figura 1 en que, después de finalizada una consulta de datos duplicados en el sistema de clúster, se requiere determinar si existe un nuevo bloque de datos en un flujo de datos recibidos y almacenar el

nuevo bloque de datos, y otras etapas coinciden con las de la realización correspondiente a la Figura 1; por lo tanto, según la realización correspondiente a la Figura 1, el método puede asimismo incluir:

Etapa 46: De acuerdo con una política de almacenamiento pre-establecida, seleccionar una dirección de almacenamiento en la cual se escriben el nuevo bloque de datos y una huella del nuevo bloque de datos.

- 5 La política de almacenamiento del nuevo bloque de datos puede estar pre-establecida por un usuario, y puede haber múltiples políticas. Por ejemplo, en la Política 1, primero se obtiene información de carga de cada nodo físico en el sistema de clúster, se selecciona un nodo físico al cual se migran datos de acuerdo con la información de carga, se obtienen la dirección de almacenamiento en la cual se escriben el nuevo bloque de datos y la huella del nuevo bloque de datos a partir del nodo físico seleccionado al cual se migran datos, y de esta forma, es posible conseguir un equilibrio de carga entre todos los nodos físicos; en la Política 2, es posible que no se determine la información de carga, se obtiene la dirección de almacenamiento en la cual se escriben el nuevo bloque de datos y la huella del nuevo bloque de datos a partir del nodo físico actual, y de esta forma, es posible reducir el intercambio entre nodos. Por lo tanto, un usuario puede establecer una política de almacenamiento específica de acuerdo con una situación real, la cual no está limitada en la realización de la presente invención.
- 10
- 15 Si se adopta la Política 1, puede haber múltiples maneras para escribir datos de acuerdo con la información de carga. Por ejemplo, es posible pre-establecer un umbral y puede denominarse un primer umbral pre-establecido, y cuando un valor de carga promedio de todos los nodos físicos en el sistema de clúster supera el primer umbral pre-establecido, un nodo en el que se escriben datos se selecciona a partir de un nodo físico cuyo valor de carga es menor que el valor de carga promedio, la dirección de almacenamiento en la que se escriben el nuevo bloque de datos y la huella del nuevo bloque de datos se obtiene del nodo en el que se escriben los datos, y cuando el valor de carga promedio de todos los nodos físicos del sistema de clúster es menor que o igual al primer umbral pre-establecido, la dirección de almacenamiento en la que se escriben el nuevo bloque de datos y la huella del nuevo bloque de datos se obtiene del nodo físico actual.
- 20

Otro método para realizar la migración de datos de acuerdo con un valor de carga puede incluir:

- 25 A1: Obtener información de carga de los nodos físicos en el sistema de clúster, y cuando una diferencia de carga entre cualesquiera dos nodos físicos del sistema de clúster es mayor que el primer umbral, un nodo al cual se migran datos se selecciona a partir de un nodo físico cuya carga es menor, y obtener la dirección de almacenamiento en la cual se escriben el nuevo bloque de datos y la huella del nuevo bloque de datos a partir del nodo al cual se migran datos.
- 30 A2: Cuando una diferencia de carga entre cualesquiera dos nodos físicos del sistema de clúster es menor o igual que el primer umbral, obtener, a partir del nodo físico actual, la dirección de almacenamiento en la cual se escriben el nuevo bloque de datos y la huella del nuevo bloque de datos.

Es posible obtener la información de carga mediante el nodo físico actual, y también es posible obtener directamente la información de carga mediante el nodo físico actual a partir de un tercero después de que el tercero obtiene la información de carga.

35

Etapa 47: Cuando se cumple una condición de almacenamiento pre-establecida, escribir el nuevo bloque de datos y la huella del nuevo bloque de datos en una región de almacenamiento correspondiente a la dirección de almacenamiento seleccionada.

- 40 Opcionalmente, que se cumpla la condición de almacenamiento pre-establecida puede implicar que: El nuevo bloque de datos se almacena en una memoria caché del nodo físico actual, y cuando los datos en la memoria caché del nodo físico actual alcanzan un segundo umbral pre-establecido, se cumple la condición de almacenamiento pre-establecida.

- 45 Tal y como se describe en el ejemplo de la Figura 3, se puede incluir un búfer de contenedor en cada nodo físico y configurar para almacenar temporalmente un nuevo bloque de datos obtenido mediante identificación, y cuando un tamaño de datos almacenados en el búfer de contenedor supera un segundo umbral, se puede considerar que se cumple la condición de almacenamiento pre-establecida, y el usuario puede establecer el segundo umbral según una situación real, que no está limitada en esta realización.

- 50 En la realización de la presente invención, después de obtener la dirección de almacenamiento del nuevo bloque de datos, no se limita en la presente invención una ocasión específica para almacenar la correspondencia entre el segundo valor de boceto y la dirección de almacenamiento del bloque de datos representada por el segundo valor de boceto en el correspondiente segundo nodo físico.

En el método de procesamiento de datos en un sistema de clúster ofrecido por la realización correspondiente a la Figura 4, se implementa el equilibrio de carga y mejora el rendimiento del sistema cuando se almacena el nuevo bloque de datos.

De acuerdo con el método de procesamiento de datos correspondiente a la Figura 1, con referencia a la Figura 5, para un caso en que un nodo virtual se incluye en el sistema de clúster, la presente invención ofrece otra realización, al menos un nodo virtual se obtiene lógicamente mediante la división de un nodo físico en el sistema de clúster, y la correspondencia (VPT) entre todos los nodos virtuales y todos los nodos físicos en el sistema de clúster se incluye en cada nodo físico, y en la realización de la presente invención, se supone que al menos un nodo virtual se obtiene lógicamente mediante la división de cada nodo físico, y un método de consulta de datos duplicados incluye:

Etapa 50: Procesar un flujo de datos recibidos mediante un nodo físico actual para obtener m bloques de datos, y obtener una huella de cada bloque de datos.

Etapa 51: Obtener, de acuerdo con un primer algoritmo, n primeros valores de boceto que representan el flujo de datos, donde n es un número entero mayor o igual que 1 y menor o igual que m.

Una forma de procesar el flujo de datos y una forma de obtener los primeros valores de boceto en las etapas 50 y 51 son iguales que las de la realización correspondiente a la Figura 1, y no se vuelven a describir en esta etapa.

Etapa 52: Identificar, según un segundo algoritmo, un primer nodo virtual que corresponde a cada valor de boceto en los n primeros valores de boceto y que está en el sistema de clúster, y consultando la correspondencia entre un nodo virtual en el nodo físico actual y un nodo físico, obtener un primer nodo físico correspondiente al primer nodo virtual, donde cada primer valor de boceto corresponde a un primer nodo virtual, y un primer nodo virtual puede corresponder a distintos primeros valores de boceto al mismo tiempo.

Para facilitar la descripción, un nodo virtual correspondiente identificado según el primer valor de boceto se denomina el primer nodo virtual, es posible obtener al menos un nodo virtual mediante la división de cada nodo físico en un sistema de almacenamiento de clúster, y en una solución de implementación específica, se pueden obtener dos o más nodos virtuales mediante la división de cada nodo físico, cada nodo virtual tiene un único número de serie, y la correspondencia entre un nodo virtual y un nodo físico en un sistema de clúster puede almacenarse en cada nodo físico, que indica en qué nodo físico se ubica el nodo virtual.

Etapa 53: Enviar cada primer valor de boceto al primer nodo virtual en el primer nodo físico correspondiente para realizar una consulta, recibir un mensaje de respuesta del primer nodo físico, obtener una dirección de almacenamiento correspondiente al primer valor de boceto a partir del mensaje de respuesta, y seleccionar una primera dirección de almacenamiento a partir de la dirección de almacenamiento correspondiente al primer valor de boceto.

En la realización de la presente invención, cada uno o la mayoría de los nodos físicos en el sistema de clúster almacena una tabla de índices, donde la correspondencia entre un valor de boceto obtenido de antemano y una dirección de almacenamiento de un bloque de datos almacenados representada por el valor de boceto se almacena en la tabla de índices, y la tabla de índices se almacena en el nodo físico en el sistema de clúster según una política pre-establecida, y después de que se obtiene un nodo virtual mediante la división de un nodo físico, la tabla de índices puede almacenarse en el nodo virtual obtenido mediante la división del nodo físico. A continuación se describe en detalle cómo almacenar, en el nodo virtual, cada entrada de índice en la tabla de índices.

En un caso en el cual no hay nodo virtual en el sistema de clúster de bloques de datos, una dirección de almacenamiento en la tabla de índices indica específicamente qué región de almacenamiento en qué nodo físico, y en un caso en el cual hay un nodo virtual en el sistema de clúster, una dirección de almacenamiento indica específicamente qué región de almacenamiento en qué nodo virtual.

De manera similar, una tabla de índices en cada nodo virtual puede ser una tabla de índices completa e incluye la correspondencia entre todos los valores de boceto y direcciones de almacenamiento de huellas de bloques de datos almacenados representadas por los valores de boceto en el sistema de clúster; y una parte de una tabla de índices completa también puede almacenarse en cada nodo virtual de acuerdo con una política establecida, con el fin de reducir el uso de una memoria.

El primer nodo virtual que corresponde a cada valor de boceto en los n primeros valores de boceto y que está en el sistema de clúster se identifica según el segundo algoritmo, de manera similar, hay también múltiples segundos algoritmos siempre y cuando se habilite un primer valor de boceto para que corresponda a un único nodo virtual en el sistema de clúster. Por ejemplo, tomando un primer valor de boceto A como ejemplo, cuando se realiza una consulta de datos duplicados en un bloque de datos, se realiza una operación de módulo en la totalidad de nodos virtuales en el sistema de clúster por el primer valor de boceto obtenido para obtener un primer nodo virtual correspondiente al primer valor de boceto y a continuación se identifica un primer nodo físico correspondiente al primer nodo virtual según la correspondencia entre un nodo virtual y un nodo físico. Cuando se encuentra el primer nodo físico, se encuentra el primer nodo virtual en el primer nodo físico.

En cada nodo virtual se incluyen muchas regiones de almacenamiento, y en una implementación específica, cada región de almacenamiento puede considerarse, de manera figurada, como un contenedor para almacenar datos, teniendo cada contenedor un único número de serie CID en el sistema de clúster, y el número indica en qué región de almacenamiento en qué nodo virtual en el sistema de clúster se ubica la región de almacenamiento. Además de

un bloque de datos, también se puede almacenar en cada región de almacenamiento información de huellas correspondiente al bloque de datos.

Adicionalmente es posible incluir una región de almacenamiento en búfer (búfer de contenedor) en una memoria caché de cada nodo físico y se configura para almacenar temporalmente un nuevo bloque de datos obtenido mediante identificación.

De manera similar a la realización correspondiente a la Figura 1, en una implementación específica, una misma dirección de almacenamiento en la tabla de índices puede corresponder a múltiples valores de boceto diferentes; no obstante, un valor de boceto corresponde a una dirección de almacenamiento. Cuando se consultan múltiples primeros valores de boceto en la tabla de índices, se pueden obtener múltiples direcciones de almacenamiento correspondientes, y cuando se selecciona la primera dirección de almacenamiento a partir de la dirección de almacenamiento correspondiente al primer valor de boceto, se puede pre-establecer una regla para seleccionar la primera dirección de almacenamiento o ésta puede ser establecida por un usuario de acuerdo con una situación real, y se puede hacer referencia a la descripción de seleccionar la primera dirección de almacenamiento en el método correspondiente a la Figura 1, lo cual no está limitado en esta realización.

En la realización de la presente invención, después de recibir una solicitud de consulta, un nodo físico realiza una consulta en un nodo virtual local, y si se encuentra la dirección de almacenamiento mediante la consulta, el nodo físico transmite como respuesta un resultado de consulta a un nodo que envía la solicitud de consulta. En una implementación específica, después de recibir la solicitud de consulta, el nodo físico realiza la consulta en el nodo virtual local, y transmite como respuesta, colocando el resultado de consulta en un mensaje de respuesta, el resultado de consulta al nodo físico que envía la solicitud de consulta, si se encuentra la dirección de almacenamiento mediante la consulta, la dirección de almacenamiento encontrada mediante la consulta se coloca en el mensaje de respuesta, y si no se encuentra la dirección de almacenamiento mediante la consulta, se transmite como respuesta un valor nulo en el mensaje de respuesta, o en una implementación específica, no se devuelve ninguna respuesta en un caso en el cual no se encuentra la dirección de almacenamiento mediante la consulta. Un usuario puede establecer una manera de implementación específica según una situación real.

Si el nodo físico actual no recibe la dirección de almacenamiento devuelta por los primeros nodos físicos correspondientes a los n primeros valores de boceto, un bloque de datos obtenido mediante la división del flujo de datos se utiliza como un bloque de datos no duplicados.

Etapa 54: Comparar una huella de un bloque de datos almacenado en una región de almacenamiento a la que apunta la primera dirección de almacenamiento con la huella de cada bloque de datos obtenida dividiendo el flujo de datos recibidos, y realizar una consulta a un bloque de datos duplicados, donde un bloque de datos no duplicados obtenido se utiliza como un nuevo bloque de datos.

En una implementación específica, el nodo físico actual almacena el bloque de datos y la información de huellas después de procesar el flujo de datos recibidos en la memoria caché, se pueden adoptar múltiples métodos para comparar las huellas de los bloques de datos, y se puede hacer referencia a la descripción en el método correspondiente a la Figura 1.

Etapa 55: Obtener al menos un segundo valor de boceto que representa el nuevo bloque de datos, identificar, de acuerdo con el segundo algoritmo, un segundo nodo virtual correspondiente a cada segundo valor de boceto en el segundo valor de boceto, identificar, de acuerdo con la correspondencia entre un nodo virtual y un nodo físico, un segundo nodo físico donde se ubica el segundo nodo virtual, y almacenar la correspondencia entre cada segundo valor de boceto y una dirección de almacenamiento de un bloque de datos representada por el segundo valor de boceto en el segundo nodo virtual del correspondiente segundo nodo físico.

Al tomar como ejemplo que dos nodos virtuales se obtienen mediante la división de un nodo físico en un sistema de clúster, se puede hacer referencia a la Figura 6 en la que se representa un diagrama esquemático de ejemplo de una estructura interna del nodo físico. En la Figura 6, dos nodos virtuales (nodo virtual, VN) se obtienen mediante la división de un nodo físico (nodo físico, PN), y en un nodo virtual se incluyen una tabla de índices y un aparato de almacenamiento que tiene capacidad para almacenar datos durante un largo período. Una tabla de correspondencia (VPT) entre un nodo virtual y un nodo físico se almacena en cada nodo físico al mismo tiempo.

En la técnica anterior, la información de huellas muestreadas se envía a todos los nodos para realizar una consulta. Se entiende que una cantidad de cálculo de realizar una consulta en un nodo es z, y cuando hay m nodos físicos, una cantidad de cálculo es z multiplicada por m. En el método de deduplicación en un sistema de clúster que tiene un nodo virtual provisto por esta realización, cuando se realiza una consulta de datos duplicados en el flujo de datos recibidos, el primer nodo físico que corresponde a cada primer valor de boceto y que está en el sistema de clúster se identifica según el primer valor de boceto que representa el flujo de datos, a continuación el primer valor de boceto que representa el flujo de datos se envía al nodo virtual identificado para realizar la consulta de datos duplicados, el número de nodos virtuales identificados no aumenta con un incremento del número de nodos físicos en el sistema de clúster, y un procedimiento de la consulta de datos duplicados no varía con un incremento del número de nodos

en el sistema de clúster, con el fin de mejorar de forma eficaz el rendimiento de deduplicación del sistema de clúster.

5 Con referencia a la Figura 7, la presente invención además ofrece una realización de un método de procesamiento de datos en un sistema de clúster, y difiere del método de procesamiento de datos en un sistema de clúster correspondiente a la Figura 5 en que, después de finalizar una consulta de datos duplicados en el sistema de clúster, se requiere determinar si existe un nuevo bloque de datos en un flujo de datos recibidos y almacenar el nuevo bloque de datos, y otras etapas coinciden con las de la realización correspondiente a la Figura 5; por lo tanto, según la realización correspondiente a la Figura 5, el método puede además incluir:

10 Etapa 76: De acuerdo con una política de almacenamiento pre-establecida, seleccionar una dirección de almacenamiento en la cual se escriben el nuevo bloque de datos y una huella del nuevo bloque de datos.

15 La política de almacenamiento del nuevo bloque de datos puede estar pre-establecida por un usuario, y puede haber múltiples políticas de almacenamiento. Se puede hacer referencia a la descripción sobre cómo seleccionar la dirección de almacenamiento en la cual se escriben el nuevo bloque de datos y la huella correspondiente al nuevo bloque de datos en la realización correspondiente a la Figura 4, y la diferencia radica en que, después de que se determina un nodo físico al cual se migran datos, se considera una carga de un nodo virtual dentro del nodo físico, y la dirección de almacenamiento en la cual se escriben el nuevo bloque de datos y la huella del nuevo bloque de datos se obtiene a partir de un nodo virtual cuya carga es menor en el nodo físico seleccionado al cual se migran datos.

20 Etapa 77: Cuando se cumple una condición de almacenamiento pre-establecida, escribir el nuevo bloque de datos y la huella del nuevo bloque de datos en una región de almacenamiento correspondiente a la dirección de almacenamiento seleccionada.

25 Tal y como se describe en el ejemplo de la Figura 6, se puede incluir un búfer de contenedor en cada nodo físico y configurar para almacenar temporalmente un nuevo bloque de datos obtenido a mediante identificación, y cuando un tamaño de los datos almacenados en el búfer de contenedor supera un segundo umbral, se puede considerar que se cumple la condición de almacenamiento pre-establecida, y el usuario puede establecer el segundo umbral según una situación real, que no está limitada en esta realización.

30 Después de obtener la dirección de almacenamiento del nuevo bloque de datos, no se limita en la presente invención una ocasión específica para almacenar la correspondencia entre el segundo valor de boceto y la dirección de almacenamiento del bloque de datos representada por el segundo valor de boceto en el correspondiente segundo nodo físico.

En el método de procesamiento de datos en un sistema de clúster provisto por la realización correspondiente a la Figura 7, se actualiza constantemente una tabla de índices en un nodo virtual, para aumentar aún más una probabilidad de encontrar datos duplicados.

35 Con referencia a la Figura 8, sobre la base de que se incluye un nodo virtual en un sistema de clúster, se ofrece una solución para la migración de datos en un nodo físico en el sistema de clúster. Por ejemplo, en la Figura 8, se describe un diagrama esquemático de migración de datos en un caso en que un nodo físico se extiende en dos nodos físicos.

40 Cuando se cumple una condición de migración de datos, uno de los nodos virtuales en un nodo físico PN1 se mueve íntegramente hacia otro nodo físico PN2, que incluye: mover una tabla de índices en un nodo virtual VN1 y datos en un repositorio de instancia única que pertenece al VN1 del PN2; actualizar correspondencia entre un nodo virtual migrado y un nodo físico en el sistema de clúster, y notificar a otro nodo físico en el sistema de clúster de la actualización de la correspondencia entre un nodo migrado y un nodo físico. La correspondencia actualizada entre todos los nodos virtuales y todos los nodos físicos en el sistema de clúster se almacena en un PN2 recién agregado.

45 La condición de migración de datos puede ser que los datos necesitan migrarse a un nodo físico recién agregado durante la expansión de capacidad de un nodo físico en el sistema de clúster; y también puede ser que, para equilibrar cargas, los datos en un nodo físico cuya carga es alta se migran a un nodo físico cuya carga es baja, y la condición de migración de datos no está limitada en la realización de la presente invención.

50 Cuando no se obtiene un nodo virtual mediante la división en un sistema de clúster y se migran datos en un nodo físico, debido a que se cambian un bloque de datos y una dirección de almacenamiento de una huella del bloque de datos, se debe modificar la correspondencia entre una dirección lógica del bloque de datos y un CID, y los datos de un CID involucrado pueden almacenarse en varios nodos físicos, de manera que la cantidad de cálculos es bastante amplia; sin embargo, al utilizar el sistema de clúster provisto en la realización de la presente invención, durante la migración de datos, debido a que los datos en un nodo virtual se mueven íntegramente, un CID en la tabla de índices indica una ubicación donde los datos están almacenados en el nodo virtual; por lo tanto, mientras que no cambie una identidad del nodo virtual, no es necesario que se modifique el CID, y sólo se modifica la correspondencia entre un nodo virtual involucrado y un nodo físico, lo que simplifica un procedimiento de trabajo durante la extensión de nodo y reduce la cantidad de cálculos en gran medida.

La realización de la presente invención se aplica a un sistema de clúster para almacenamiento y, en el sistema de clúster, se incluyen múltiples nodos de almacenamiento. Un aparato de procesamiento de datos en un sistema de clúster provisto en una realización de la presente invención está configurado para llevar a cabo el método de procesamiento de datos precedente en un sistema de clúster, y el aparato se puede situar en un nodo físico en un sistema de clúster o en un administrador en un sistema de clúster, o también se puede situar como un nodo separado en el sistema de clúster. La realización de la presente invención no limita una ubicación específica donde se sitúa el aparato de procesamiento de datos.

Con referencia a la Figura 9, una realización de la presente invención ofrece un aparato de procesamiento de datos 900 en un sistema de clúster, donde el aparato incluye:

Una unidad divisora 90 está configurada para dividir un flujo de datos recibido por un nodo físico actual en  $m$  bloques de datos, y obtener una huella de cada bloque de datos, donde  $m$  es un número entero mayor o igual que 1.

Una unidad de obtención de valor de boceto 91 está configurada para obtener, de acuerdo con un primer algoritmo,  $n$  primeros valores de boceto que representan el flujo de datos, donde  $n$  es un número entero mayor o igual que 1 y menor o igual que  $m$ .

Se puede hacer referencia a la descripción en el método anterior sobre el primer algoritmo para obtener el primer valor de boceto, que no se vuelve a describir en la presente descripción.

Una unidad de identificación 92 está configurada para identificar, de acuerdo con un segundo algoritmo, un primer nodo físico que corresponde a cada valor de boceto en los  $n$  primeros valores de boceto y que está en el sistema de clúster.

La unidad de identificación 92 puede identificar el primer nodo físico de diversas maneras. Por ejemplo, la extracción de raíces y operaciones de redondeo se realizan sobre el primer valor de boceto hasta que el valor sea menor a  $m$ , y se redondea un resultado final para obtener el correspondiente primer nodo físico; y también se puede realizar una operación de módulo sobre la cantidad de todos los nodos físicos en el sistema de clúster por cada primer valor de boceto para obtener el primer nodo físico que corresponde a cada primer valor de boceto y que está en el sistema de clúster; y en la realización de la presente invención, el último puede adoptarse para que:

la unidad de identificación 92 esté específicamente configurada para realizar una operación de módulo sobre la cantidad total de nodos físicos en el sistema de clúster por cada primer valor de boceto para obtener el primer nodo físico que corresponde a cada primer valor de boceto y que está en el sistema de clúster.

Una unidad de envío 93 está configurada para enviar cada primer valor de boceto al correspondiente primer nodo físico para realizar una consulta.

Después de recibir una solicitud de consulta enviada por la unidad de envío 93, otro nodo en el sistema de clúster consulta, según una tabla de índices almacenada de forma local, si hay una dirección de almacenamiento que corresponda al primer valor de boceto. En la realización de la presente invención, cada uno o la mayoría de los nodos físicos en el sistema de clúster almacena una tabla de índices, donde la correspondencia entre un valor de boceto y una dirección de almacenamiento de un bloque de datos almacenado representada por el valor de boceto se almacena en la tabla de índices; la tabla de índices se almacena en un nodo físico en el sistema de clúster según una política pre-establecida, y un bloque de datos y la información de huella correspondiente al bloque de datos están almacenados en una región de almacenamiento a la que apunta una dirección de almacenamiento diferente.

Una unidad de recepción 94 está configurada para recibir al menos un mensaje de respuesta del primer nodo físico correspondiente a los  $n$  primeros valores de boceto, y obtener una primera dirección de almacenamiento a partir del mensaje de respuesta.

La unidad de recepción 94 puede incluir: una subunidad de recogida de estadísticas 941, configurada para obtener una dirección de almacenamiento correspondiente al primer valor de boceto a partir del mensaje de respuesta, y recoger estadísticas sobre el número de accesos de cada dirección de almacenamiento que lleva el mensaje de respuesta; y una subunidad de obtención de dirección 942, configurada para seleccionar, de acuerdo con un resultado estadístico de la subunidad de recogida de estadísticas 941,  $s$  direcciones de almacenamiento que tienen el mayor número de accesos como la primera dirección de almacenamiento, donde  $s$  es mayor o igual que 1.

Una unidad de consulta 95 está configurada para comparar una huella de un bloque de datos almacenada en una región de almacenamiento a la que apunta la primera dirección de almacenamiento con la huella de cada bloque de datos obtenida dividiendo el flujo de datos recibido, y realizar una consulta a un bloque de datos duplicados, donde un bloque de datos no duplicados obtenido mediante la unidad de consulta 95 se utiliza como un nuevo bloque de datos.

Una unidad de actualización 96 está configurada para obtener al menos un segundo valor de boceto que representa el nuevo bloque de datos, identificar, de acuerdo con el segundo algoritmo, un segundo nodo físico correspondiente

a cada segundo valor de boceto en el segundo valor de boceto, y almacenar la correspondencia entre cada segundo valor de boceto y una dirección de almacenamiento de un bloque de datos representada por cada segundo valor de boceto en el correspondiente segundo nodo físico.

5 Un algoritmo para obtener el segundo valor de boceto que presenta el nuevo bloque de datos de la unidad de actualización 96 puede ser el mismo que el primer algoritmo para obtener el primer valor de boceto, que no está limitado en la presente descripción.

10 La correspondencia entre cada segundo valor de boceto y una dirección de almacenamiento de un bloque de datos representada por cada segundo valor de boceto se utiliza como una entrada de índice en la tabla de índices. En la realización de la presente invención, la unidad de actualización 96 almacena la tabla de índices en un nodo físico según una política establecida, no hay una relación directa entre los datos en una tabla de índices de un nodo físico específico y datos en el nodo físico, y la tabla de índices solo está almacenada en el nodo físico particular según una política de asignación establecida. Se puede hacer referencia a la descripción en el método anterior sobre la descripción de la tabla de índices, y los detalles no se vuelven a describir en la presente memoria.

15 Con referencia a la descripción del método precedente, en una implementación específica, puede haber múltiples maneras de que la unidad de obtención de valor de boceto 91 obtenga, de acuerdo con el primer algoritmo, los n primeros valores de boceto que representan el flujo de datos. Por ejemplo, la unidad de obtención de valor de boceto 91 puede incluir: una subunidad de agrupamiento de huellas 911 y una subunidad de obtención 912, donde:

20 la subunidad de agrupamiento 911 está configurada para agrupar huellas de los m bloques de datos del flujo de datos en n primeros grupos de huellas, donde cada primer grupo de huellas incluye huellas de al menos dos bloques de datos; y

la subunidad de obtención 912 está configurada para obtener un primer valor de boceto que representa cada grupo de huellas, con el fin de obtener los n primeros valores de boceto.

25 La subunidad de obtención 912 puede estar específicamente configurada para seleccionar una huella más pequeña de cada grupo de huellas como un primer valor de boceto que representa un grupo de huellas al que pertenece la huella más pequeña.

30 La unidad de recepción 94 del aparato de procesamiento de datos en un sistema de clúster puede además estar configurada para: por ejemplo, en un caso en que la dirección de almacenamiento transmitida como respuesta por el primer nodo físico no se recibe después de que el primer valor de boceto se envía al correspondiente primer nodo físico para realizar una consulta, utilizar un bloque de datos obtenido mediante la división del flujo de datos como un bloque de datos no duplicados, y luego activar la unidad de actualización 96.

El aparato provisto por la realización de la presente invención puede estar situado en un controlador de un nodo de almacenamiento en el sistema de clúster, y está configurado para realizar el método de procesamiento de datos precedente. Se puede hacer referencia a la descripción en la realización de método para la descripción detallada de las funciones de todas las unidades, y los detalles no se vuelven a describir en la presente memoria.

35 Mediante el aparato de procesamiento de datos en un sistema de clúster provisto por la realización de la presente invención, cuando se realiza una consulta de datos duplicados en el flujo de datos recibidos, el primer nodo físico que corresponde a cada primer valor de boceto y que está en el sistema de clúster se identifica de acuerdo con el primer valor de boceto que representa el flujo de datos y, a continuación, el primer valor de boceto que representa el flujo de datos se envía al nodo físico identificado para la consulta de datos duplicados, y un procedimiento de la consulta de datos duplicados no cambia con un incremento de la cantidad de nodos en el sistema de clúster; por lo tanto, una cantidad de cálculos de cada nodo no aumenta con un incremento de la cantidad de nodos en el sistema de clúster.

45 Cuando se encuentran, mediante la consulta, datos no duplicados, la unidad de consulta 95 puede considerar los datos no duplicados como un nuevo bloque de datos y almacenar el nuevo bloque de datos en un repositorio de instancia única; por lo tanto, la realización de la presente invención puede además incluir:

Una unidad de obtención de dirección de escritura 97, configurada para: de acuerdo con una política de almacenamiento pre-establecida, seleccionar una dirección de almacenamiento en la cual se escriben el nuevo bloque de datos y una huella del nuevo bloque de datos.

50 Una unidad de escritura 98 configurada para: cuando se cumple una condición de almacenamiento pre-establecida, escribir el nuevo bloque de datos y la huella del nuevo bloque de datos en una región de almacenamiento correspondiente a la dirección de almacenamiento seleccionada.

Puede haber múltiples políticas de almacenamiento del nuevo bloque de datos de la unidad de obtención de dirección de escritura 97. Por ejemplo, la unidad de obtención de dirección de escritura 97 incluye:

una subunidad de obtención de información de carga 971, configurada para obtener información de carga de cada nodo físico en el sistema de clúster; y

5 una subunidad de selección de dirección 972, configurada para seleccionar, según la información de carga, un nodo físico al cual se migran los datos, y obtener, del nodo físico seleccionado al cual se migran los datos, la dirección de almacenamiento en la que se escriben el nuevo bloque de datos y la huella del nuevo bloque de datos.

Hay múltiples métodos mediante los cuales la subunidad de selección de dirección 972 selecciona, según la información de carga, la dirección de almacenamiento en la que se escriben los nuevos datos. Por ejemplo, cuando se identifica, según la información de carga obtenida por la subunidad de obtención de información de carga 971, que un valor de carga promedio de todos los nodos físicos en el sistema de clúster supera un primer umbral pre-establecido, un nodo en el que se escriben los datos se selecciona de un nodo físico cuyo valor de carga es menor que el valor de carga promedio, y la dirección de almacenamiento en la que se escriben el nuevo bloque de datos y la huella del nuevo bloque de datos se obtiene del nodo en el que los datos se escriben, y cuando el valor de carga promedio de todos los nodos físicos del sistema de clúster es menor o igual que el primer umbral pre-establecido, la dirección de almacenamiento en la que se escriben el nuevo bloque de datos y la huella del nuevo bloque de datos se obtiene del nodo físico actual.

Ciertamente, en la política de almacenamiento del nuevo bloque de datos de la unidad de obtención de dirección de escritura 97, puede que la información de carga tampoco sea identificada, y la dirección se selecciona directamente del nodo físico actual; por lo tanto, la unidad de obtención de dirección de escritura 97 puede configurarse específicamente para obtener, del nodo físico actual, la dirección de almacenamiento en la que se escriben el nuevo bloque de datos y la huella del nuevo bloque de datos.

La unidad de escritura 98 puede estar configurada para, por ejemplo, almacenar el nuevo bloque de datos en una memoria caché del nodo físico actual, y cuando los datos en la memoria caché del nodo físico actual alcanzan un segundo umbral pre-establecido, y se cumple la condición de almacenamiento pre-establecida, escribir el nuevo bloque de datos y la huella del nuevo bloque de datos en una región de almacenamiento correspondiente a la dirección de almacenamiento seleccionada.

Para un caso en el que un nodo virtual está incluido en el sistema de clúster, al menos un nodo virtual se obtiene de forma lógica mediante la división de un nodo físico en el sistema de clúster, y la correspondencia (VPT) entre todos los nodos virtuales y todos los nodos físicos en el sistema de clúster se incluye en cada nodo físico, y en la realización de la presente invención, se supone que al menos un nodo virtual se obtiene de forma lógica mediante la división de cada nodo físico, de manera que:

la unidad de identificación 92 pueda estar específicamente configurada para identificar, de acuerdo con el segundo algoritmo, un primer nodo virtual que corresponde a cada valor de boceto en los n primeros valores de boceto y que está en el sistema de clúster, y obtener, al consultar correspondencia entre un nodo virtual en el nodo físico actual y un nodo físico, un primer nodo físico correspondiente al primer nodo virtual, donde se utiliza un bloque de datos no duplicados obtenido como un nuevo bloque de datos; y

la unidad de actualización 96 esté específicamente configurada para obtener al menos un segundo valor de boceto que representa el nuevo bloque de datos, identificar, de acuerdo con el segundo algoritmo, un segundo nodo virtual correspondiente a cada segundo valor de boceto en el segundo valor de boceto, identificar, de acuerdo con la correspondencia entre un nodo virtual y un nodo físico, un segundo nodo físico donde se ubica el segundo nodo virtual, y, después de obtener una dirección de almacenamiento del nuevo bloque de datos, almacenar la correspondencia entre cada segundo valor de boceto y una dirección de almacenamiento de un bloque de datos representada por el segundo valor de boceto en el segundo nodo virtual del correspondiente segundo nodo físico.

Cuando se implemente la función de identificación, según el segundo algoritmo, un primer nodo virtual que corresponde a cada valor de boceto en los n primeros valores de boceto y que está en el sistema de clúster, la unidad de identificación 92 puede realizar una operación de módulo sobre la cantidad de todos los nodos virtuales en el sistema de clúster por cada primer valor de boceto para obtener el primer nodo virtual que corresponde a cada primer valor de boceto y que está en el sistema de clúster.

En un caso en que un nodo virtual está incluido en el sistema de clúster, cuando se realiza una solución de migración de datos en un nodo físico en el sistema de clúster, la unidad de escritura 98 está además configurada para: cuando se cumple una condición de migración de datos, migrar íntegramente al menos un nodo virtual en un nodo físico cuyos datos necesitan ser migrados a un nodo físico objetivo; y la unidad de actualización 96 está además configurada para actualizar la correspondencia entre un nodo virtual migrado en el nodo físico actual y el nodo físico, y notificar a otro nodo físico en el sistema de clúster de la actualización de la correspondencia entre el nodo migrado y el nodo físico.

La condición de migración de datos puede estar pre-establecida por un usuario, puede ser que los datos necesitan migrarse a un nodo físico recién agregado durante la expansión de capacidad de un nodo físico en el sistema de

clúster; y también puede ser que, para equilibrar la carga, los datos de un nodo físico cuya carga es alta son migrados a un nodo físico cuya carga es baja, y la condición de migración de datos no está limitada en la realización de la presente invención.

5 El aparato de procesamiento de datos en un sistema de clúster puede implementar una migración integral de un nodo virtual en un nodo físico, mientras que un número de serie de un nodo virtual migrado hacia un nuevo nodo físico no cambia, donde una dirección de almacenamiento de datos correspondiente a un valor de boceto en la tabla de índices representa una ubicación donde se almacenan los datos en el nodo virtual; por lo tanto, mientras que no cambie una identidad del nodo virtual, no es necesario modificar la dirección de almacenamiento de datos, y solo se modifica la correspondencia entre un nodo virtual involucrado y un nodo físico, lo que simplifica un procedimiento de trabajo durante la extensión de nodo y reduce en gran medida la cantidad de cálculos.

10 Con referencia a la Figura 10, una realización de la presente invención además ofrece un aparato de procesamiento de datos 100 en un sistema de clúster, donde el aparato puede estar situado en un nodo de almacenamiento en el sistema de clúster e incluye un procesador 101, una memoria 102, un controlador de red 103, y un controlador de disco magnético 104, donde el procesador 101, la memoria primaria 102, el controlador de red 103, y el controlador de disco magnético 104 están conectados a través de un bus 105.

El controlador de almacenamiento 102 está configurado para almacenar un programa 1021.

El procesador 101 está configurado para llevar a cabo el programa 1021 en la memoria primaria 102.

20 El procesador 101 se comunica con otro nodo en el sistema de clúster a través del controlador de red 103, y se comunica, a través del controlador de disco magnético 104, con un aparato de almacenamiento que tiene una capacidad de almacenamiento de datos a largo plazo en un nodo local, por ejemplo, un disco magnético en la Figura 10. Un aparato de almacenamiento controlado por el controlador de disco magnético 104 no se limita a un disco magnético, y puede ser también un aparato formado por otro medio de almacenamiento.

En una implementación específica, la memoria 102 puede ser una memoria primaria, por ejemplo, una memoria, y una forma de implementación específica no está limitada en la realización de la presente invención.

25 Específicamente, el programa 1021 puede incluir un código de programa, donde el código de programa incluye una instrucción informática operativa.

El procesador 102 puede ser una unidad de procesamiento central CPU o un circuito integrado de aplicaciones específicas ASIC (circuito integrado de aplicaciones específicas), o estar configurado con uno o más circuitos integrados para implementar las realizaciones de la presente invención.

30 Con referencia a la Figura 9, el programa 1021 puede incluir:

Una unidad divisora 90 está configurada para dividir un flujo de datos recibido por un nodo físico actual en  $m$  bloques de datos, y obtener una huella de cada bloque de datos, donde  $m$  es un número entero mayor o igual que 1.

35 Una unidad de obtención de valor de boceto 91 está configurada para obtener, de acuerdo con un primer algoritmo,  $n$  primeros valores de boceto que representan el flujo de datos, donde  $n$  es un número entero mayor o igual que 1 y menor o igual que  $m$ .

Se puede hacer referencia a la descripción en el método precedente para el primer algoritmo para obtener el primer valor de boceto, y los detalles no se vuelven a describir en la presente memoria.

40 Una unidad de identificación 92 está configurada para identificar, de acuerdo con un segundo algoritmo, un primer nodo físico que corresponde a cada valor de boceto en los  $n$  primeros valores de boceto y que está en el sistema de clúster.

45 La unidad de identificación 92 puede identificar el primer nodo físico de diversas maneras. Por ejemplo, la extracción de raíces y operaciones de redondeo se realizan sobre el primer valor de boceto hasta que el valor sea menor a  $m$ , y se redondea un resultado final para obtener el correspondiente primer nodo físico; y también se puede realizar una operación de módulo sobre la cantidad de todos los nodos físicos en el sistema de clúster por cada primer valor de boceto para obtener el primer nodo físico que corresponde a cada primer valor boceto y que está en el sistema de clúster; y en la realización de la presente invención, se adopta el último, de tal manera que:

la unidad de identificación 92 esté específicamente configurada para realizar una operación de módulo sobre la cantidad total de nodos físicos en el sistema de clúster por cada primer valor de boceto para obtener el primer nodo físico que corresponde a cada primer valor de boceto y que está en el sistema de clúster.

50 Una unidad de envío 93 está configurada para enviar cada primer valor de boceto al correspondiente primer nodo físico para realizar una consulta.

- Después de recibir una solicitud de consulta enviada por la unidad de envío 93, otro nodo en el sistema de clúster consulta, según una tabla de índices almacenada de forma local, si hay una dirección de almacenamiento correspondiente al primer valor de boceto. En la realización de la presente invención, cada uno o la mayoría de los nodos físicos en el sistema de clúster almacena una tabla de índices, donde la correspondencia entre un valor de boceto y una dirección de almacenamiento de un bloque de datos almacenado representada por el valor de boceto se almacena en la tabla de índices, la tabla de índices se almacena en un nodo físico en el sistema de clúster según una política pre-establecida, y un bloque de datos y la información de huella correspondiente al bloque de datos están almacenados en una región de almacenamiento a la que apunta una dirección de almacenamiento diferente.
- Una unidad de recepción 94 está configurada para recibir al menos un mensaje de respuesta de los primeros nodos físicos correspondientes a los n primeros valores de boceto, y obtener una primera dirección de almacenamiento a partir del mensaje de respuesta, y
- la unidad de recepción 94 puede incluir: una subunidad de recogida de estadísticas 941, configurada para obtener la dirección de almacenamiento correspondiente al primer valor de boceto a partir del mensaje de respuesta, y recoger estadísticas sobre el número de accesos de cada dirección de almacenamiento que lleva el mensaje de respuesta; y una subunidad de obtención de dirección 942, configurada para seleccionar, de acuerdo con un resultado estadístico de la subunidad de recogida de estadísticas 941, s direcciones de almacenamiento que tienen el mayor número de accesos como la primera dirección de almacenamiento, donde s es mayor o igual que 1.
- Una unidad de consulta 95 está configurada para comparar una huella de un bloque de datos almacenada en una región de almacenamiento a la que apunta la primera dirección de almacenamiento con la huella de cada bloque de datos obtenida dividiendo el flujo de datos recibido, y consultar un bloque de datos duplicados, donde un bloque de datos no duplicados obtenido se utiliza como un nuevo bloque de datos.
- Una unidad de actualización 96 está configurada para obtener al menos un segundo valor de boceto que representa el nuevo bloque de datos, identificar, de acuerdo con el segundo algoritmo, un segundo nodo físico correspondiente a cada segundo valor de boceto en el segundo valor de boceto, y almacenar la correspondencia entre cada segundo valor de boceto y una dirección de almacenamiento de un bloque de datos representada por cada segundo valor de boceto en el correspondiente segundo nodo físico.
- Un algoritmo para obtener el segundo valor de boceto que presenta el nuevo bloque de datos de la unidad de actualización 96 puede ser igual al primer algoritmo para obtener el primer valor de boceto, el cual no está limitado en la presente descripción.
- La correspondencia entre cada segundo valor de boceto y una dirección de almacenamiento de un bloque de datos representada por cada segundo valor de boceto se utiliza como una entrada de índice en la tabla de índices. En la realización de la presente invención, la unidad de actualización 96 almacena la tabla de índices en un nodo físico según una política establecida, no hay una relación directa entre los datos en una tabla de índices de un nodo físico específico y los datos en el nodo físico, y la tabla de índices solo está almacenada en el nodo físico específico según una política de asignación establecida. Se puede hacer referencia a la descripción en el método precedente sobre la descripción de la tabla de índices, y los detalles no se vuelven a describir en la presente memoria.
- Con referencia a la descripción en el método precedente, en una implementación específica, puede haber múltiples maneras de que la unidad de obtención de valor de boceto 91 obtenga, según el primer algoritmo, los n primeros valores de boceto que representan el flujo de datos. Por ejemplo, la unidad de obtención de valor de boceto 91 puede incluir: una subunidad de agrupamiento de huellas 911 y una subunidad de obtención 912, donde:
- la subunidad de agrupamiento 911 está configurada para agrupar huellas de los m bloques de datos del flujo de datos en n primeros grupos de huellas, donde cada primer grupo de huellas incluye huellas de al menos dos bloques de datos; y
- la subunidad de obtención 912 está configurada para obtener un primer valor de boceto que representa cada grupo de huellas, con el fin de obtener los n primeros valores de boceto.
- La subunidad de obtención 912 puede estar específicamente configurada para seleccionar una huella más pequeña de cada grupo de huellas como un primer valor de boceto que representa un grupo de huellas al que pertenece la huella más pequeña.
- La unidad de recepción 94 del aparato de procesamiento de datos en un sistema de clúster puede además estar configurada para: por ejemplo, en un caso en que la dirección de almacenamiento transmitida como respuesta por el primer nodo físico no se recibe después de que el primer valor de boceto se envía al correspondiente primer nodo físico para realizar una consulta, utilizar un bloque de datos obtenido mediante la división del flujo de datos como un bloque de datos no duplicados, y luego activar la unidad de actualización 96.
- El aparato provisto por la realización de la presente invención puede estar situado en un controlador de un nodo de almacenamiento en el sistema de clúster, y está configurado para realizar el método de procesamiento de datos

precedente. Se puede hacer referencia a la descripción en la realización de método para la descripción detallada de las funciones de todas las unidades, y los detalles no se vuelven a describir en la presente memoria.

- 5 Mediante el aparato de procesamiento de datos en un sistema de clúster provisto por la realización de la presente invención, cuando se realiza una consulta de datos duplicados en el flujo de datos recibidos, un valor de boceto muestreado se envía a un nodo físico identificado para la consulta de datos duplicados, y un procedimiento de la consulta de datos duplicados no cambia con un incremento de la cantidad de nodos en el sistema de clúster; por lo tanto, una cantidad de cálculos de cada nodo no aumenta con un incremento de la cantidad de nodos en el sistema de clúster.
- 10 Cuando se encuentran, mediante la consulta, datos no duplicados, la unidad de consulta 95 puede considerar los datos no duplicados como un nuevo bloque de datos y almacenar el nuevo bloque de datos en un repositorio de instancia única; por lo tanto, la realización de la presente invención puede además incluir:
- 15 Una unidad de obtención de dirección de escritura 97, configurada para: de acuerdo con una política de almacenamiento pre-establecida, seleccionar una dirección de almacenamiento en la cual se escriben el nuevo bloque de datos y una huella del nuevo bloque de datos.
- 15 Una unidad de escritura 98, configurada para: cuando se cumple una condición de almacenamiento pre-establecida, escribir el nuevo bloque de datos y la huella del nuevo bloque de datos en una región de almacenamiento correspondiente a la dirección de almacenamiento seleccionada.
- Puede haber múltiples políticas de almacenamiento del nuevo bloque de datos de la unidad de obtención de dirección de escritura 97. Por ejemplo, la unidad de obtención de dirección de escritura 97 incluye:
- 20 una subunidad de obtención de información de carga 971, configurada para obtener información de carga de cada nodo físico en el sistema de clúster; y
- una subunidad de selección de dirección 972, configurada para seleccionar, según la información de carga, un nodo físico al cual se migran los datos, y obtener, del nodo físico seleccionado al cual se migran los datos, la dirección de almacenamiento en la que se escriben el nuevo bloque de datos y la huella del nuevo bloque de datos.
- 25 Hay múltiples métodos mediante los cuales la subunidad de selección de dirección 972 selecciona, según la información de carga, la dirección de almacenamiento en la que se escriben los nuevos datos. Por ejemplo, cuando se identifica, según la información de carga obtenida por la subunidad de obtención de información de carga 971, que un valor de carga promedio de todos los nodos físicos en el sistema de clúster supera un primer umbral pre-establecido, un nodo en el que se escriben los datos se selecciona de un nodo físico cuyo valor de carga es menor
- 30 que el valor de carga promedio, y la dirección de almacenamiento en la que se escriben el nuevo bloque de datos y la huella del nuevo bloque de datos se obtiene del nodo en el que los datos se escriben, y cuando el valor de carga promedio de todos los nodos físicos del sistema de clúster es menor o igual que el primer umbral pre-establecido, la dirección de almacenamiento en la que se escriben el nuevo bloque de datos y la huella del nuevo bloque de datos se obtiene del nodo físico actual.
- 35 Ciertamente, en la política de almacenamiento del nuevo bloque de datos de la unidad de obtención de dirección de escritura 97, puede que la información de carga tampoco sea identificada, y la dirección se selecciona directamente del nodo físico actual; por lo tanto, la unidad de obtención de dirección de escritura 97 puede estar específicamente configurada para obtener, del nodo físico actual, la dirección de almacenamiento en la que se escriben el nuevo bloque de datos y la huella del nuevo bloque de datos.
- 40 La unidad de escritura 98 puede estar configurada para, por ejemplo, almacenar el nuevo bloque de datos en una memoria caché del nodo físico actual, y cuando los datos en la memoria caché del nodo físico actual alcanzan un segundo umbral pre-establecido, y se cumple la condición de almacenamiento pre-establecida, escribir el nuevo bloque de datos y la huella del nuevo bloque de datos en una región de almacenamiento correspondiente a la dirección de almacenamiento seleccionada.
- 45 Para un caso en que un nodo virtual está incluido en el sistema de clúster, al menos un nodo virtual se obtiene de forma lógica mediante la división de un nodo físico en el sistema de clúster, y la correspondencia (VPT) entre todos los nodos virtuales y los nodos físicos en el sistema de clúster está incluida en cada nodo físico, y en la realización de la presente invención, se entiende que al menos un nodo virtual se obtiene de forma lógica mediante la división de cada nodo físico, de manera que:
- 50 la unidad de identificación 92 pueda estar específicamente configurada para identificar, de acuerdo con el segundo algoritmo, un primer nodo virtual que corresponde a cada valor de boceto en los n primeros valores de boceto y que está en el sistema de clúster, y obtener, al consultar correspondencia entre un nodo virtual en el nodo físico actual y un nodo físico, un primer nodo físico correspondiente al primer nodo virtual, donde se utiliza un bloque de datos no duplicados obtenido como un nuevo bloque de datos; y

la unidad de actualización 96 esté específicamente configurada para obtener al menos un segundo valor de boceto que representa el nuevo bloque de datos, identificar, de acuerdo con el segundo algoritmo, un segundo nodo virtual correspondiente a cada segundo valor de boceto en el segundo valor de boceto, identificar, de acuerdo con la correspondencia entre un nodo virtual y un nodo físico, un segundo nodo físico donde se ubica el segundo nodo virtual, y, después de obtener una dirección de almacenamiento del nuevo bloque de datos, almacenar la correspondencia entre cada segundo valor de boceto y una dirección de almacenamiento de un bloque de datos representada por el segundo valor de boceto en el segundo nodo virtual del correspondiente segundo nodo físico.

Cuando se implemente la función de identificación, según el segundo algoritmo, un primer nodo virtual que corresponde a cada valor de boceto en los  $n$  primeros valores de boceto y que está en el sistema de clúster, la unidad de identificación 92 puede realizar una operación de módulo sobre el número de todos los nodos virtuales en el sistema de clúster por cada primer valor de boceto para obtener el primer nodo virtual que corresponde a cada primer valor de boceto y que está en el sistema de clúster.

En un caso en el que un nodo virtual está incluido en el sistema de clúster, cuando se realiza una solución de migración de datos en un nodo físico en el sistema de clúster, la unidad de escritura 98 está además configurada para: cuando se cumple una condición de migración de datos, migrar íntegramente al menos un nodo virtual en un nodo físico cuyos datos necesitan ser migrados a un nodo físico objetivo; y la unidad de actualización 96 está además configurada para actualizar la correspondencia entre un nodo virtual migrado en el nodo físico actual y el nodo físico, y notificar a otro nodo físico en el sistema de clúster de la actualización de la correspondencia entre el nodo migrado y el nodo físico.

La condición de migración de datos puede estar pre-establecida por un usuario, puede ser que los datos necesitan migrarse hacia un nodo físico recién agregado durante la expansión de capacidad de un nodo físico en el sistema de clúster; y también puede ser que, para equilibrar la carga, los datos de un nodo físico cuya carga es alta son migrados a un nodo físico cuya carga es baja, y la condición de migración de datos no está limitada en la realización de la presente invención.

El aparato de procesamiento de datos en un sistema de clúster puede implementar una migración integral de un nodo virtual en un nodo físico, mientras que un número de serie de un nodo virtual migrado hacia un nuevo nodo físico no cambia, donde una dirección de almacenamiento de datos correspondiente a un valor de boceto en la tabla de índices representa una ubicación donde se almacenan datos en el nodo virtual; por lo tanto, mientras que no cambie una identidad del nodo virtual, no es necesario modificar la dirección de almacenamiento de datos, y solo se modifica la correspondencia entre un nodo virtual involucrado y un nodo físico, lo que simplifica un procedimiento de trabajo durante la extensión de nodo y reduce, en gran medida, la cantidad de cálculos.

Una realización de la presente invención además ofrece un producto de programa informático para procesamiento de datos, donde el producto de programa informático incluye un medio de almacenamiento legible por ordenador que almacena un código de programa, donde una instrucción incluida en el código de programa se utiliza para:

dividir un flujo de datos recibido por un nodo físico actual en  $m$  bloques de datos, obtener una huella de cada bloque de datos, y obtener, de acuerdo con un primer algoritmo,  $n$  primeros valores de boceto que representan el flujo de datos, donde  $m$  es un número entero mayor o igual que 1, y  $n$  es un número entero mayor o igual que 1 y menor o igual que  $m$ .

identificar, de acuerdo con un segundo algoritmo, un primer nodo físico que corresponde a cada primer valor de boceto en los  $n$  primeros valores de boceto y que está en el sistema de clúster, y enviar cada primer valor de boceto al correspondiente primer nodo físico para realizar una consulta;

recibir al menos un mensaje de respuesta de primeros nodos físicos correspondientes a los  $n$  primeros valores de boceto, y obtener una primera dirección de almacenamiento a partir del mensaje de respuesta; comparar una huella de un bloque de datos almacenado en una región de almacenamiento a la que apunta la primera dirección de almacenamiento con la huella de cada bloque de datos obtenida mediante la división del flujo de datos, y consultar un bloque de datos duplicados, donde se utiliza un bloque de datos no duplicados obtenido como un nuevo bloque de datos; y

obtener al menos un segundo valor de boceto que representa el nuevo bloque de datos, identificar, de acuerdo con el segundo algoritmo, un segundo nodo físico correspondiente a cada segundo valor de boceto en el segundo valor de boceto, y almacenar la correspondencia entre cada segundo valor de boceto y una dirección de almacenamiento de un bloque de datos representada por cada segundo valor de boceto en un correspondiente segundo nodo físico.

El producto de programa informático para procesamiento de datos provisto por la realización de la presente invención incluye el medio de almacenamiento legible por ordenador que almacena el código de programa, donde la instrucción incluida en el código de programa puede utilizarse para llevar a cabo el método en las realizaciones de

método precedentes, pudiéndose hacer referencia a las realizaciones de método para implementaciones específicas, y los detalles no se vuelven a describir en la presente memoria.

5 Los expertos en la técnica han de comprender claramente que, en aras de una descripción conveniente y breve, se puede hacer referencia a los procesos correspondientes en las realizaciones de método para procesos de trabajo específicos de los sistemas, aparatos, y unidades precedentes, y los detalles no se vuelven a describir en la presente memoria.

10 En las diversas realizaciones proporcionadas en la presente solicitud, se ha de comprender que el sistema, aparato y método descritos pueden implementarse de otras maneras. Por ejemplo, las realizaciones de aparato descritas son meramente ejemplos. Por ejemplo, la división de las unidades es simplemente un tipo de función lógica de división, y en la implementación real puede haber otras formas de división. Por ejemplo, se pueden combinar o integrar en otro sistema múltiples unidades o componentes, o algunas características se pueden ignorar o no llevar a cabo. Además, los acoplamientos mutuos mostrados o descritos o los acoplamientos directos o conexiones de comunicación se pueden implementar a través de algunas interfaces de comunicación. Los acoplamientos indirectos o conexiones de comunicación entre los aparatos o unidades se pueden implementar de forma electrónica, mecánica o de otra forma.

15 Las unidades descritas como partes separadas pueden o no estar físicamente separadas, y las partes mostradas como unidades pueden o no ser unidades físicas, pueden estar ubicadas en una posición, o también pueden estar distribuidas en una pluralidad de unidades de red. Algunas o todas las unidades se pueden seleccionar según las necesidades reales para lograr los objetivos de las soluciones de las realizaciones.

20 Además, las unidades funcionales en las realizaciones de la presente invención pueden estar integradas en una unidad de procesamiento, cada una de las unidades también puede existir físicamente por sí sola, y dos o más unidades también están integradas en una unidad.

25 Cuando las funciones se implementan en una forma de una unidad funcional de software y se venden o usan como un producto independiente, las funciones se pueden almacenar en un medio de almacenamiento legible por ordenador. Partiendo de esta premisa, las soluciones técnicas de la presente invención, en esencia, o la parte que contribuye a la técnica anterior, o parte de las soluciones técnicas, se pueden realizar en la forma de un producto de software. El producto de software informático está almacenado en un medio de almacenamiento, e incluye diversas instrucciones para indicar a un dispositivo informático (que puede ser un ordenador personal, un servidor o un dispositivo de red, y así sucesivamente) que realice parte o todas las etapas del método descrito en la realización de la presente invención. El medio de almacenamiento anterior incluye: cualquier medio capaz de almacenar un código de programa, tal y como una disco flash USB, un disco duro retirable, una memoria de solo lectura (ROM, memoria de sólo lectura), una memoria de acceso aleatorio (RAM, memoria de acceso aleatorio), un disco magnético, o un disco óptico.

35 La descripción anterior meramente se refiere a realizaciones específicas de la presente invención, pero no está concebida para limitar la presente invención.

**REIVINDICACIONES**

1. Un método de procesamiento de datos un sistema de clúster, que comprende:

5 dividir (10) un flujo de datos recibido por un nodo físico actual en m bloques de datos, obtener una huella de cada bloque de datos, y obtener n primeros valores de boceto que representan el flujo de datos de acuerdo con un primer algoritmo, en donde m es un número entero mayor o igual que 1, y n es un número entero mayor o igual que 1 y menor o igual que m;

identificar (12), de acuerdo con un segundo algoritmo, un primer nodo físico que corresponde a cada primer valor de boceto en los n primeros valores de boceto y que está en el sistema de clúster, y enviar cada primer valor de boceto al correspondiente primer nodo físico para realizar una consulta;

10 recibir (13) al menos un mensaje de respuesta de los primeros nodos físicos correspondientes a los n primeros valores de boceto, y obtener una primera dirección de almacenamiento a partir del mensaje de respuesta;

comparar (14) una huella de un bloque de datos almacenado en una región de almacenamiento a la que apunta la primera dirección de almacenamiento con la huella de cada bloque de datos obtenida mediante la división del flujo de datos para obtener un bloque de datos no duplicados como un nuevo bloque de datos; y

15 obtener (15) un segundo valor de boceto que representa el nuevo bloque de datos, identificar, de acuerdo con el segundo algoritmo, un segundo nodo físico correspondiente al segundo valor de boceto, y almacenar la correspondencia entre el segundo valor de boceto y una dirección de almacenamiento del nuevo bloque de datos representada por el segundo valor de boceto en el segundo nodo físico.

20 2. El método de acuerdo con la reivindicación 1, en donde al menos un nodo virtual se obtiene de forma lógica mediante la división de cada nodo físico en el sistema de clúster, y la correspondencia entre un nodo virtual y un nodo físico en el sistema de clúster se comprende en cada nodo físico;

la etapa de identificar, de acuerdo con un segundo algoritmo, un primer nodo físico que corresponde a cada primer valor de boceto en los n primeros valores de boceto y que está en el sistema de clúster comprende:

25 identificar, de acuerdo con el segundo algoritmo, un primer nodo virtual que corresponde a cada primer valor de boceto en los n primeros valores de boceto y que está en el sistema de clúster, y obtener, al consultar correspondencia entre un nodo virtual y un nodo físico, un primer nodo físico correspondiente al primer nodo virtual; y

30 la etapa de identificar, de acuerdo con el segundo algoritmo, un segundo nodo físico correspondiente al segundo valor de boceto, y almacenar la correspondencia entre el segundo valor de boceto y una dirección de almacenamiento del nuevo bloque de datos representada por el segundo valor de boceto en el segundo nodo físico comprende:

35 identificar, de acuerdo con el segundo algoritmo, un segundo nodo virtual correspondiente al segundo valor de boceto, identificar, de acuerdo con la correspondencia entre un nodo virtual y un nodo físico, un segundo nodo físico donde se ubica el segundo nodo virtual, y almacenar la correspondencia entre el segundo valor de boceto y una dirección de almacenamiento del nuevo bloque de datos representada por el segundo valor de boceto en el segundo nodo virtual del segundo nodo físico.

3. El método según la reivindicación 2, en donde la etapa de identificar, de acuerdo con el segundo algoritmo, un primer nodo virtual que corresponde a cada primer valor de boceto en los n primeros valores de boceto y que está en el sistema de clúster comprende:

40 realizar una operación de módulo sobre la cantidad total de nodos virtuales en el sistema de clúster por cada primer valor de boceto para obtener el primer nodo virtual que corresponde a cada primer valor de boceto y que está en el sistema de clúster.

4. El método según la reivindicación 2 o 3, que además comprende:

45 cuando una condición de migración de datos se cumple, migrar íntegramente al menos un nodo virtual en un nodo físico cuyos datos necesitan migrarse a un nodo físico objetivo; y

actualizar la correspondencia entre un nodo virtual migrado en el nodo físico actual y un nodo físico, y notificar a otro nodo físico en el sistema de clúster de la actualización de la correspondencia entre un nodo migrado y un nodo físico.

5. El método según la reivindicación 1, en donde la etapa de identificar, de acuerdo con un segundo algoritmo, un primer nodo físico que corresponde a cada primer valor de boceto en los n primeros valores de boceto y que está en el sistema de clúster comprende:
- 5 realizar una operación de módulo sobre la cantidad total de nodos físicos en el sistema de clúster por cada primer valor de boceto para obtener el primer nodo físico que corresponde a cada primer valor de boceto y que está en el sistema de clúster.
6. El método según cualquiera de las reivindicaciones 1 a 5, en donde la etapa de obtener un segundo valor de boceto que representa el nuevo bloque de datos comprende: obtener, de acuerdo con el primer algoritmo, el segundo valor de boceto que representa el nuevo bloque de datos.
- 10 7. El método según cualquiera de las reivindicaciones 1 a 6, en donde la obtención, de acuerdo con un primer algoritmo, de n primeros valores de boceto que representan el flujo de datos comprende:
- agrupar huellas de los m bloques de datos del flujo de datos en n primeros grupos de huellas, en donde cada primer grupo de huellas comprende huellas de al menos dos bloques de datos; y
- 15 obtener un primer valor de boceto que representa cada grupo de huellas, con el fin de obtener los n primeros valores de boceto.
8. El método según cualquiera de las reivindicaciones 1 a 5, que además comprende:
- seleccionar (46), de acuerdo con una política de almacenamiento pre-establecida, una dirección de almacenamiento en la cual se escriben el nuevo bloque de datos y una huella del nuevo bloque de datos, cuando se cumple una
- 20 condición de almacenamiento pre-establecida, escribir (47) el nuevo bloque de datos y la huella del nuevo bloque de datos en una región de almacenamiento a la que apunta la dirección de almacenamiento seleccionada.
9. El método según la reivindicación 8, en donde la etapa de seleccionar, de acuerdo con una política de almacenamiento pre-establecida, una dirección de almacenamiento en la cual se escriben el nuevo bloque de datos y una huella del nuevo bloque de datos comprende:
- 25 obtener información de carga de un nodo físico en el sistema de clúster, seleccionar, de acuerdo con la información de carga, un nodo físico al cual se migran los datos, y obtener, del nodo físico al cual se migran los datos, la dirección de almacenamiento en la que se escriben el nuevo bloque de datos y la huella del nuevo bloque de datos.
10. El método según la reivindicación 8, en donde la etapa de seleccionar, de acuerdo con una política de almacenamiento pre-establecida, una dirección de almacenamiento en la cual se escriben el nuevo bloque de datos
- 30 y una huella del nuevo bloque de datos comprende:
- obtener, del nodo físico actual, la dirección de almacenamiento en la cual se escriben el nuevo bloque de datos y la huella del nuevo bloque de datos.
11. El método según la reivindicación 7, en donde la etapa de obtener un primer valor de boceto que representa cada grupo de huellas comprende: seleccionar, de cada grupo de huellas, una huella más pequeña como un primer
- 35 valor de boceto que representa un grupo de huellas al que pertenece la huella más pequeña.
12. Un aparato de procesamiento de datos en un sistema de clúster, que comprende:
- una unidad divisora (90), configurada para dividir un flujo de datos recibido por un nodo físico actual en m bloques de datos, y obtener una huella de cada bloque de datos, donde m es un número entero mayor o igual que 1;
- 40 una unidad de obtención de valor de boceto (91), configurada para obtener n primeros valores de boceto que representan el flujo de datos de acuerdo con un primer algoritmo, en donde n es un número entero mayor o igual que 1 y menor o igual que m;
- una unidad de identificación, configurada para identificar, de acuerdo con un segundo algoritmo, un primer nodo físico que corresponde a cada primer valor de boceto en los n primeros valores de boceto y que está en el sistema
- 45 de clúster;
- una unidad de envío (93), configurada para enviar cada primer valor de boceto al correspondiente primer nodo físico para realizar una consulta;

una unidad de recepción (94), configurada para recibir al menos un mensaje de respuesta de los primeros nodos físicos correspondientes a los n primeros valores de boceto, y obtener una primera dirección de almacenamiento a partir del mensaje de respuesta;

5 una unidad de consulta (95), configurada para comparar una huella de un bloque de datos almacenado en una región de almacenamiento a la que apunta la primera dirección de almacenamiento con la huella de cada bloque de datos obtenida mediante la división del flujo de datos para obtener un bloque de datos no duplicados como un nuevo bloque de datos; y

10 una unidad de actualización (96), configurada para obtener un segundo valor de boceto que representa el nuevo bloque de datos, identificar, de acuerdo con el segundo algoritmo, un segundo nodo físico correspondiente al segundo valor de boceto, y almacenar la correspondencia entre el segundo valor de boceto y una dirección de almacenamiento del nuevo bloque de datos representada por el segundo valor de boceto en el segundo nodo físico.

13. El aparato según la reivindicación 12, que además comprende:

15 una unidad de obtención de dirección de escritura (97), configurada para: de acuerdo con una política de almacenamiento pre-establecida, seleccionar una dirección de almacenamiento en la cual se escriben el nuevo bloque de datos y una huella del nuevo bloque de datos; y

una unidad de escritura (98), configurada para: cuando se cumple una condición de almacenamiento pre-establecida, escribir el nuevo bloque de datos y la huella del nuevo bloque de datos en una región de almacenamiento correspondiente a la dirección de almacenamiento seleccionada.

20 14. El aparato según la reivindicación 13, en donde al menos un nodo virtual se obtiene de forma lógica mediante la división de cada nodo físico en el sistema de clúster, y la correspondencia entre un nodo virtual y un nodo físico en el sistema de clúster se comprende en cada nodo físico;

25 la unidad de identificación está específicamente configurada para identificar, de acuerdo con el segundo algoritmo, un primer nodo virtual que corresponde a cada valor de boceto en los n primeros valores de boceto y que está en el sistema de clúster, y obtener, al consultar correspondencia entre un nodo virtual y un nodo físico, un primer nodo físico correspondiente al primer nodo virtual, en donde el bloque de datos no duplicados obtenido por la unidad de consulta se utiliza como un nuevo bloque de datos; y

30 la unidad de actualización (96) está específicamente configurada para obtener el segundo valor de boceto que representa el nuevo bloque de datos, identificar, de acuerdo con el segundo algoritmo, un segundo nodo virtual correspondiente al segundo valor de boceto, identificar, de acuerdo con la correspondencia entre un nodo virtual y un nodo físico, un segundo nodo físico donde se ubica el segundo nodo virtual, y, almacenar la correspondencia el segundo valor de boceto y una dirección de almacenamiento del bloque de datos representada por el segundo valor de boceto en el segundo nodo virtual del segundo nodo físico.

35 15. El aparato según la reivindicación 14, en donde la unidad de escritura (98) está además configurada para: cuando se cumple una condición de migración de datos, migrar íntegramente al menos un nodo virtual en un nodo físico cuyos datos necesitan ser migrados a un nodo físico objetivo; y

la unidad de actualización (96) está además configurada para actualizar la correspondencia entre un nodo virtual migrado en el nodo físico actual y un nodo físico, y notificar a otro nodo físico en el sistema de clúster de la actualización de la correspondencia entre un nodo migrado y un nodo físico.

40 16. El aparato según la reivindicación 12, en donde la unidad de identificación está específicamente configurada para realizar una operación de módulo sobre la cantidad total de nodos físicos en el sistema de clúster por cada primer valor de boceto para obtener el primer nodo físico que corresponde a cada primer valor de boceto y que está en el sistema de clúster.

45 17. El aparato según cualquiera de las reivindicaciones 12 a 15, en donde la unidad de obtención de valor de boceto (91) comprende:

una subunidad de agrupamiento de huellas (911), configurada para agrupar huellas de los m bloques de datos del flujo de datos en n primeros grupos de huellas, en donde cada primer grupo de huellas comprende huellas de al menos dos bloques de datos; y

50 una subunidad de obtención (912), configurada para obtener un primer valor de boceto que representa cada grupo de huellas, con el fin de obtener los n primeros valores de boceto.

18. El aparato según cualquiera de las reivindicaciones 13 a 15, en donde la unidad de obtención de dirección de escritura (97) comprende:

una subunidad de obtención de información de carga (971), configurada para obtener una información de carga del nodo físico en el sistema de clúster; y

5 una subunidad de selección de dirección (972), configurada para seleccionar, según la información de carga obtenida por la subunidad de obtención de información de carga, un nodo físico al cual se migran los datos, y obtener, del nodo físico al cual se migran los datos, la dirección de almacenamiento en la que se escriben el nuevo bloque de datos y la huella del nuevo bloque de datos.

10 19. El aparato según cualquiera de las reivindicaciones 13 a 15, en donde la unidad de obtención de dirección de escritura (97) está específicamente configurada para obtener, del nodo físico actual, la dirección de almacenamiento en la que se escriben el nuevo bloque de datos y la huella del nuevo bloque de datos.

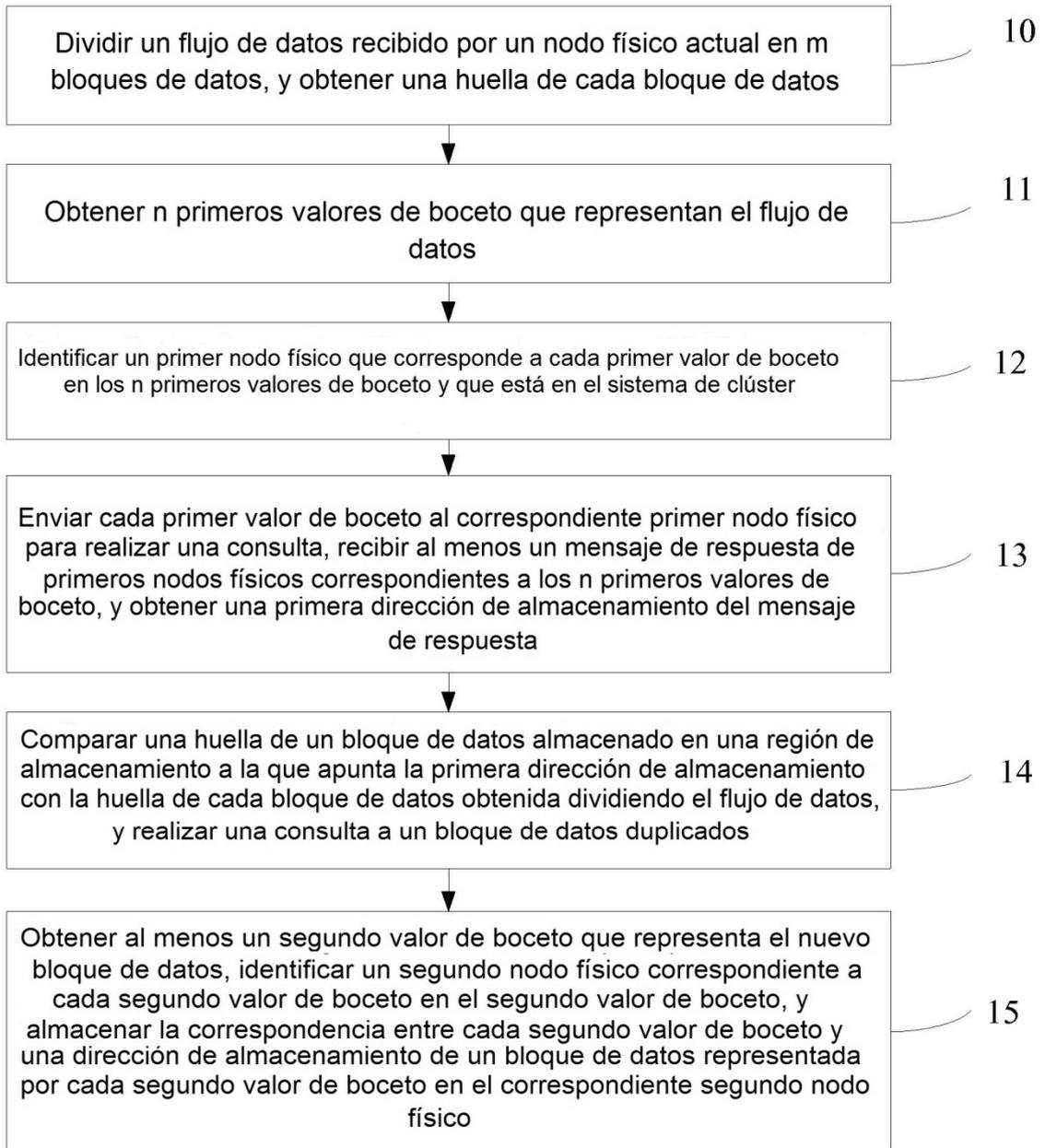


FIG. 1

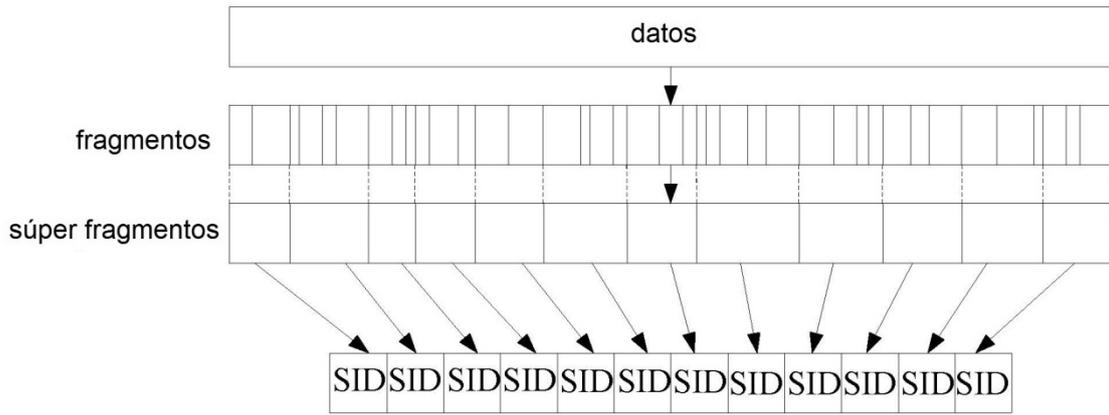


FIG. 2

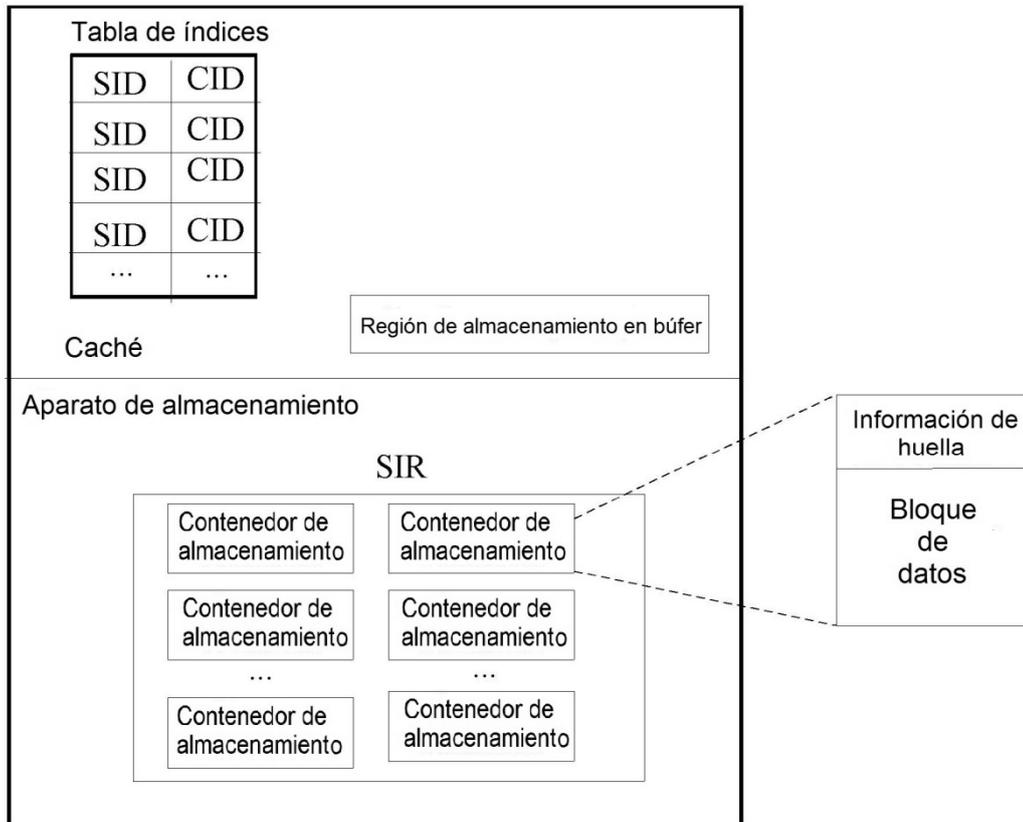


FIG. 3

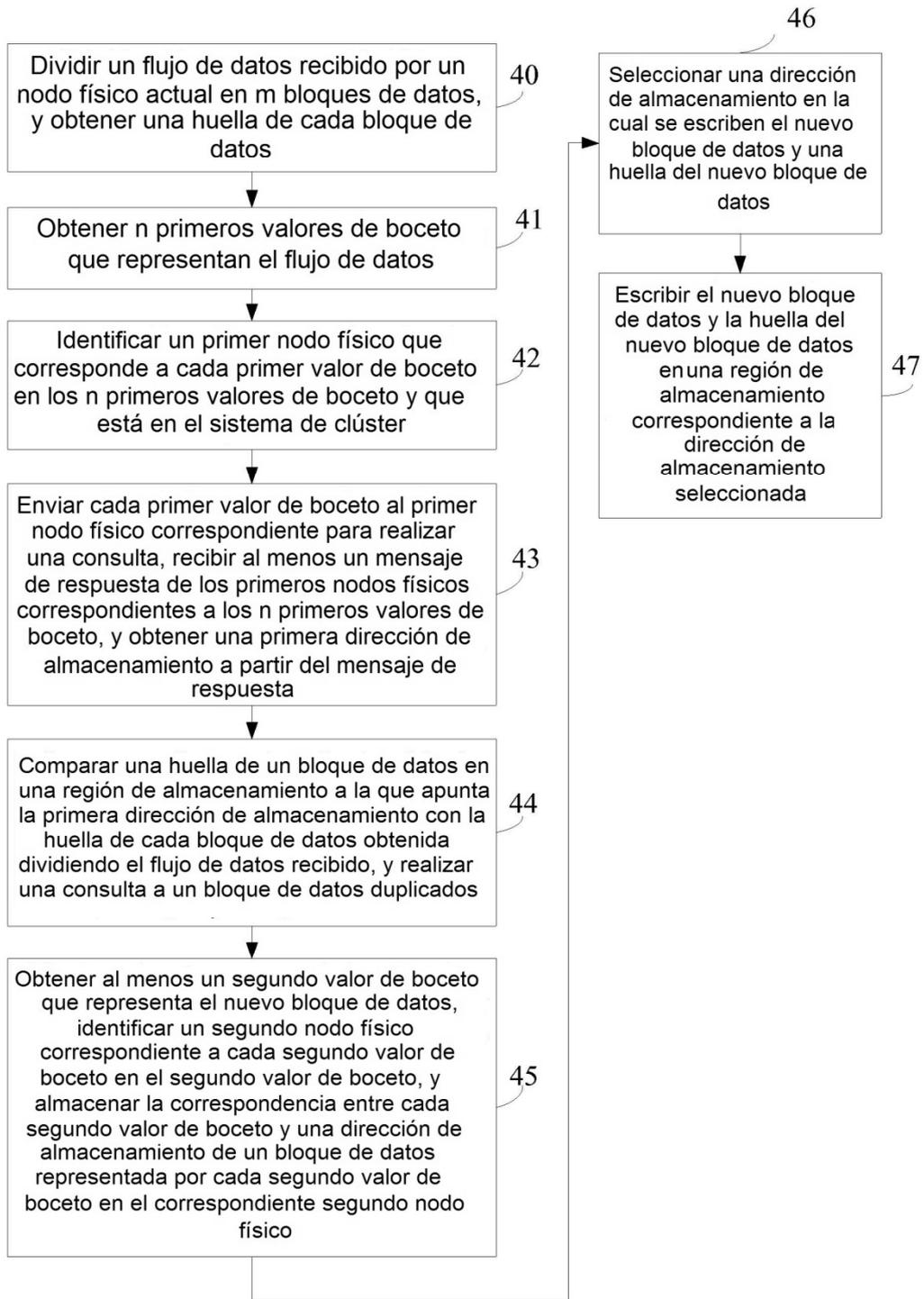


FIG. 4

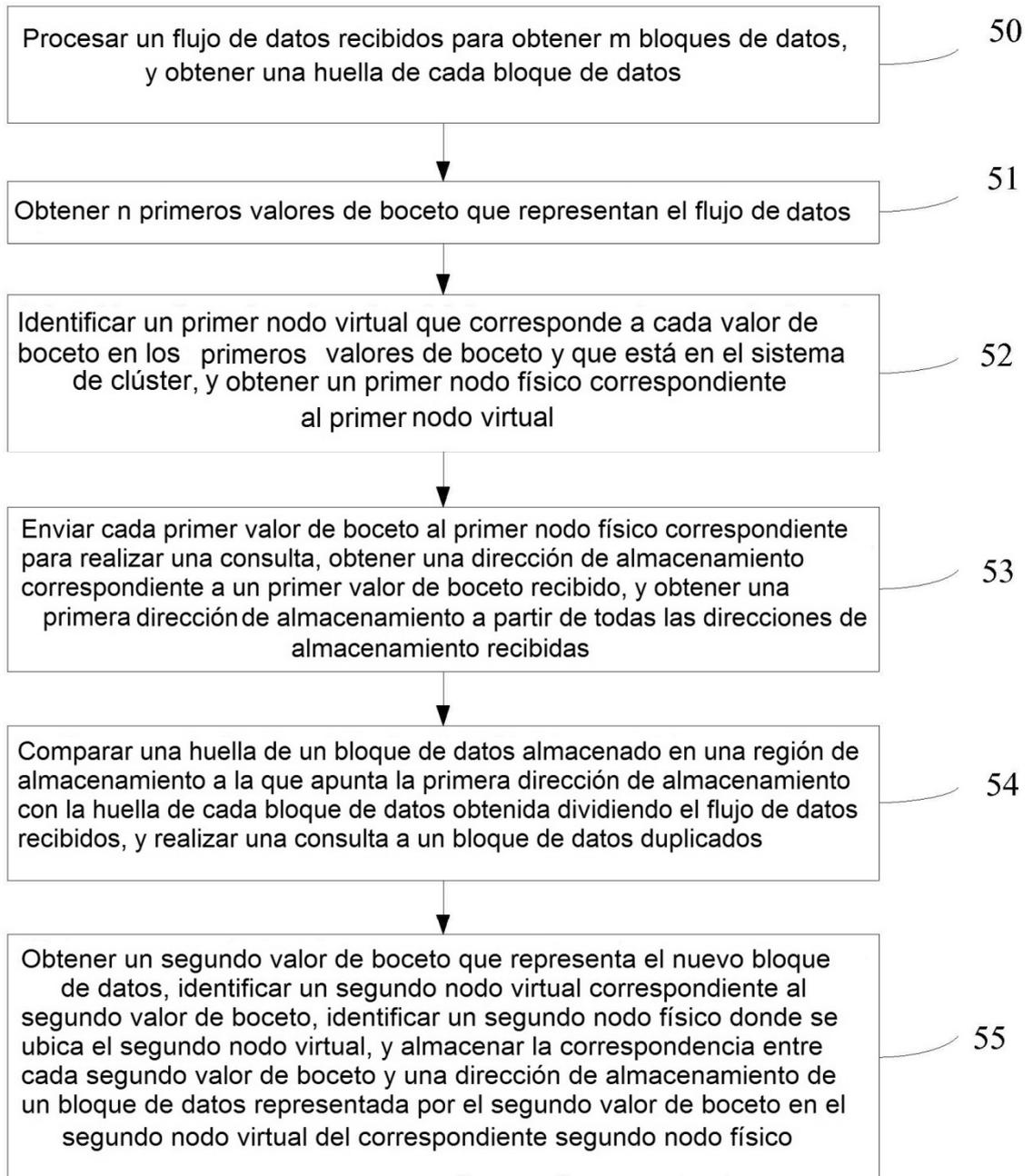


FIG. 5

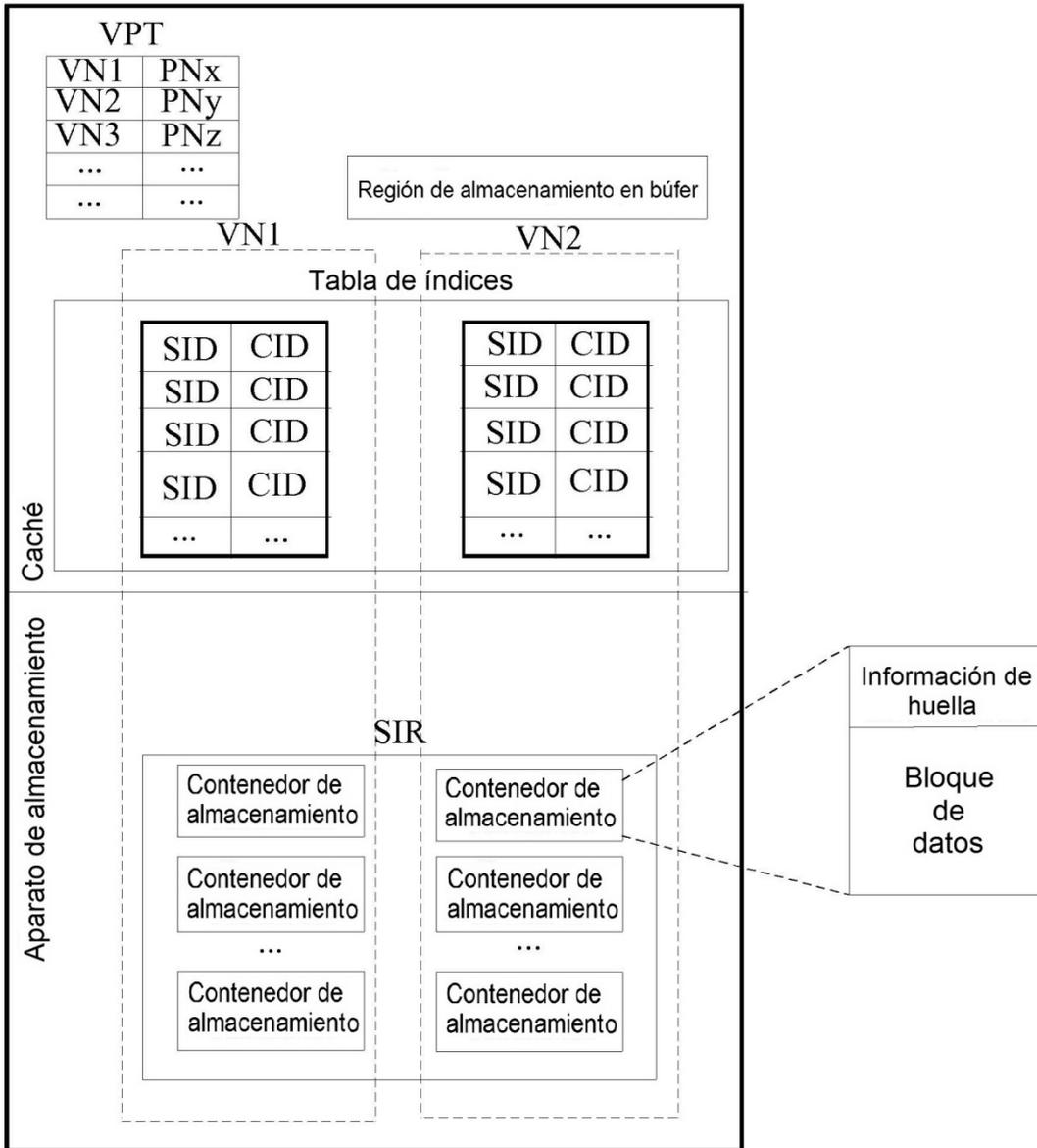


FIG. 6

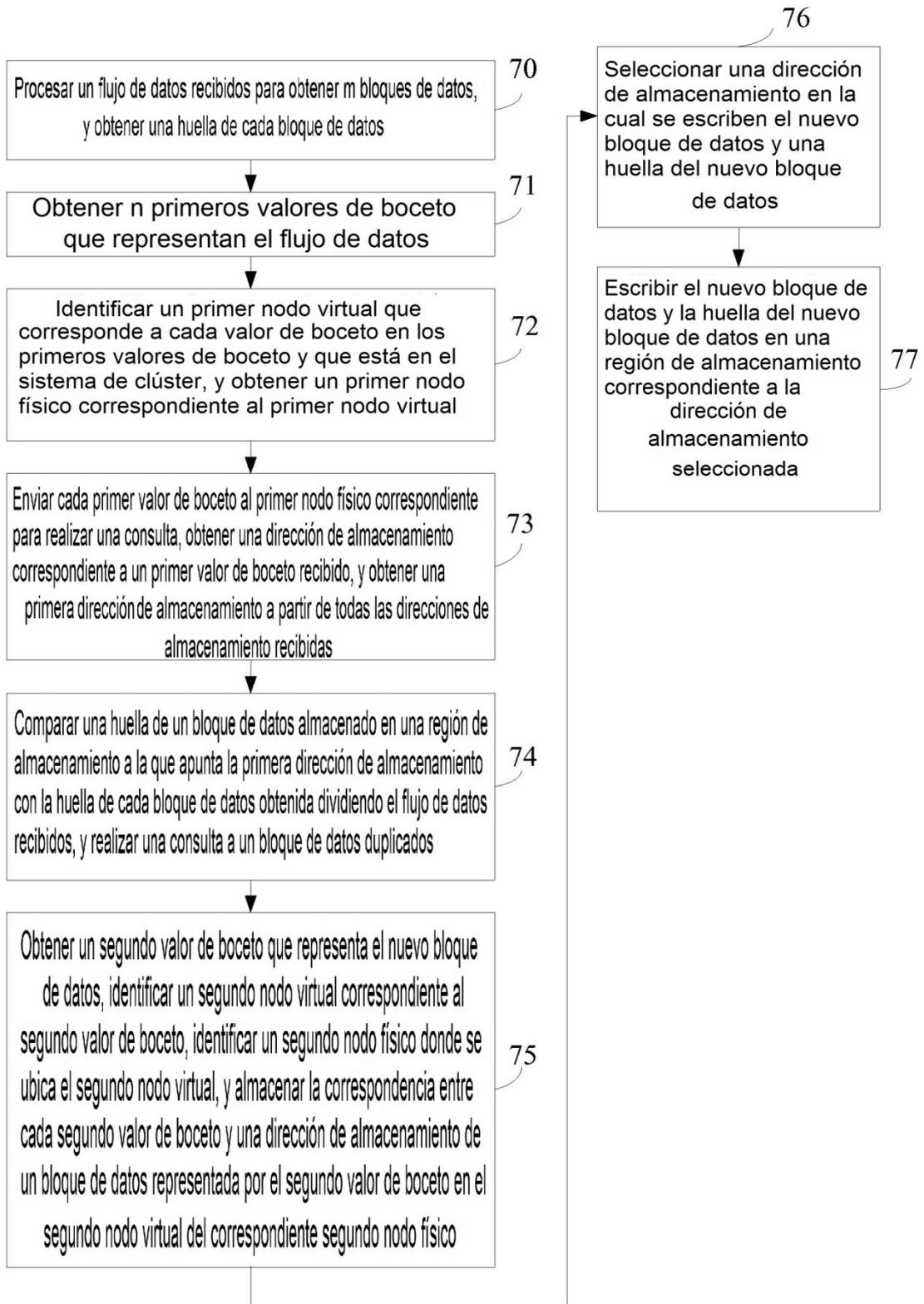


FIG. 7

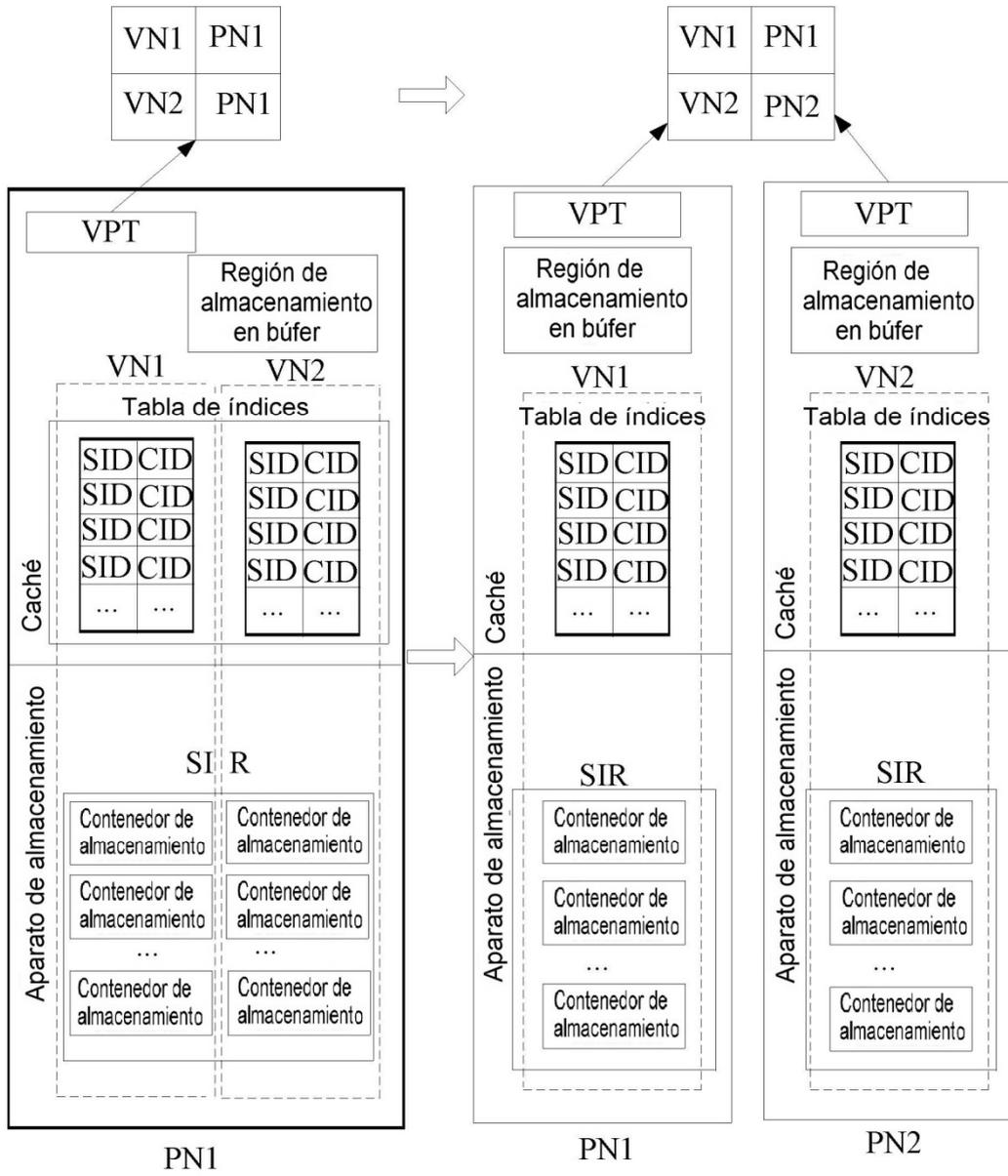


FIG. 8

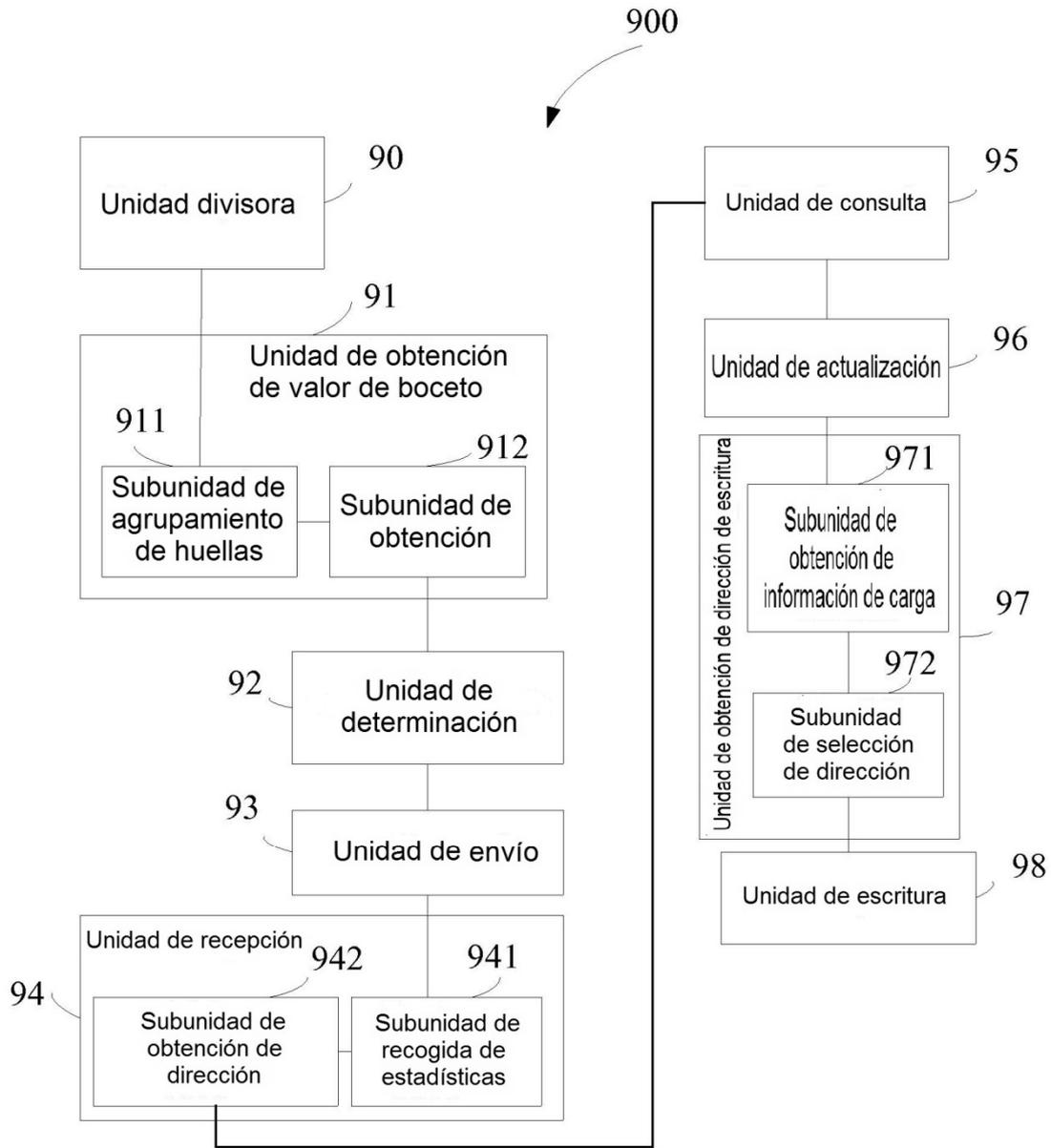


FIG. 9

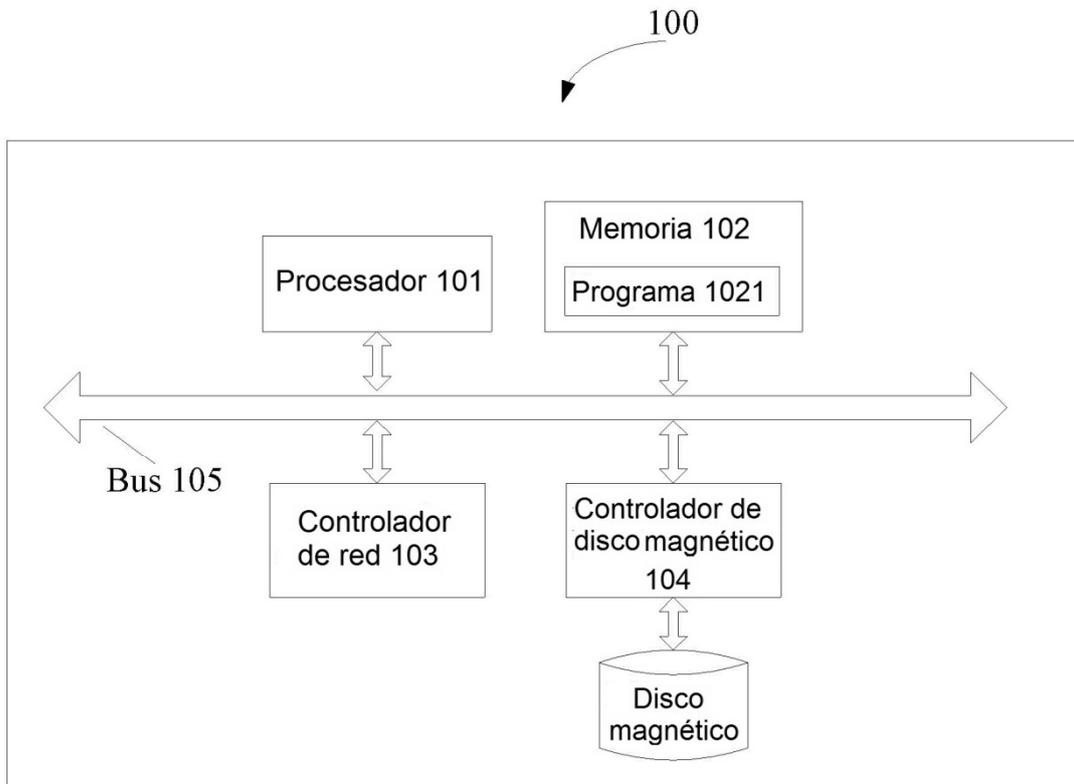


FIG. 10