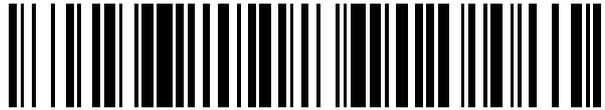


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 628 611**

51 Int. Cl.:

| | |
|-------------------|-----------|
| G06F 9/06 | (2006.01) |
| G06F 9/44 | (2006.01) |
| G06F 13/00 | (2006.01) |
| G06F 11/34 | (2006.01) |
| G06F 3/06 | (2006.01) |

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

- 86 Fecha de presentación y número de la solicitud internacional: **12.11.2012 PCT/US2012/064587**
- 87 Fecha y número de publicación internacional: **23.05.2013 WO13074426**
- 96 Fecha de presentación y número de la solicitud europea: **12.11.2012 E 12850595 (5)**
- 97 Fecha y número de publicación de la concesión europea: **05.04.2017 EP 2780799**

54 Título: **E / S del disco de restricción usando el modelo de simulación de la unidad de disco**

30 Prioridad:
17.11.2011 US 201113298993

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
03.08.2017

73 Titular/es:
**MICROSOFT TECHNOLOGY LICENSING, LLC
(100.0%)
One Microsoft Way
Redmond, Washington 98052-6399, US**

72 Inventor/es:
**DENG, YIMIN;
CHAU, HO YUEN;
ZUO, YUE y
FOLTZ, FORREST CURTIS**

74 Agente/Representante:
CARPINTERO LÓPEZ, Mario

ES 2 628 611 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

E / S del disco de restricción usando el modelo de simulación de la unidad de disco

Antecedentes

5 Los sistemas de red a gran escala son plataformas comunes que se emplean en una variedad de configuraciones para ejecutar aplicaciones y mantener datos para las funciones empresariales y operativas. Por ejemplo, un centro de datos (por ejemplo, una infraestructura informático física en nube) puede proporcionar una variedad de servicios (por ejemplo, aplicaciones web, servicios de correo electrónico, servicios de motores de búsqueda, etc.) A una pluralidad de clientes simultáneamente. Estos sistemas de red a gran escala suelen incluir un gran número de recursos distribuidos a través del centro de datos, en el que cada recurso se asemeja a una máquina física o a una máquina virtual (VM) que se ejecuta en un nodo o un servidor físico. Cuando el centro de datos aloja múltiples inquilinos (por ejemplo, programas de cliente), estos recursos son asignados óptimamente desde el mismo centro de datos a los diferentes inquilinos.

15 A menudo, varias máquinas virtuales se ejecutan simultáneamente en el mismo nodo físico dentro de una red informática o en el centro de datos. Estas máquinas virtuales que comparten un nodo físico común se pueden asignar a los diferentes inquilinos y pueden requerir diferentes cantidades de recursos en diferentes momentos. Por ejemplo, los recursos pueden incluir un disco físico (por ejemplo, un controlador de disco duro) asociado con un nodo físico, en el que el disco físico tiene una cantidad restringida de accesibilidad durante un cierto periodo de tiempo. Cuando un inquilino requiere un mayor uso del disco físico para realizar una tarea en particular, las máquinas virtuales del inquilino que se ejecutan en el nodo físico pueden sobrecargar potencialmente el disco físico enviando una gran cantidad de solicitudes, evitando que otras máquinas virtuales que se están ejecutando en el nodo físico compartan adecuadamente los recursos del mismo.

25 En la actualidad, puesto que el rendimiento de entrada / salida de disco (E / S) (por ejemplo, el tiempo de acceso a un segmento de datos) típicamente está retardado con respecto al rendimiento de la CPU (por ejemplo, la velocidad del procesador) y al rendimiento de la red (por ejemplo, el uso de ancho de banda) de las máquinas virtuales que se ejecutan en el nodo físico, existen enfoques básicos para medir las solicitudes enviadas desde las máquinas virtuales al disco físico. Estos enfoques miden las solicitudes intentando describir el rendimiento de E / S del disco utilizando ya sea la velocidad de almacenamiento (MB / segundo) o, alternativamente, la velocidad de actividad (solicitudes / segundo). En el caso de utilizar MB / segundo, este enfoque considera un tamaño de datos que se leen o se escriben, pero no cómo los datos están dispuestos dentro del disco físico. Por otro lado, en el caso de usar solicitudes / segundo, este enfoque considera un número de veces que se pide a un cabezal de disco que se mueva sobre el disco físico dentro de un intervalo de tiempo dado. Es decir, el uso del enfoque de solicitudes / segundo no considera el tamaño de los datos a los que se accede con el movimiento del cabezal del disco.

35 En consecuencia, debido a que el enfoque basado en MB / segundo (relevante sólo para el acceso secuencial) y el enfoque basado en las solicitudes / segundo (relevante sólo para el acceso aleatorio) están orientados cada uno en un aspecto específico diferente del rendimiento de E / S del disco, estos enfoques existentes no puede tener en cuenta tanto el acceso secuencial como el acceso aleatorio en el disco físico. Por ejemplo, el enfoque basado en MB / segundo está orientado a accesos secuenciales, asumiendo así el riesgo de ser superado por una avalancha de accesos aleatorios y generando un retardo de solicitudes. Por otra parte, el enfoque basado en las solicitudes / segundo está orientado a accesos aleatorios, asumiendo así el riesgo de encontrar una serie de accesos secuenciales y limitando innecesariamente la cantidad de rendimiento al disco físico. En consecuencia, al no tener en cuenta el acceso secuencial y el acceso aleatorio, ni reconocer las grandes diferencias en el rendimiento de E / S del disco resultante entre ellos, hace que estos enfoques sean ineficaces.

45 El documento US 2010/0011182 desvela un dispositivo de almacenamiento y un planificador, determinando el planificador una ejecución de ventanas desplazables correspondientes al momento límite solicitado, y de esta manera las ventanas son ejecutadas de forma ordenada.

Sumario

50 Este sumario se proporciona para introducir conceptos en una forma simplificada que se describen más adelante en la Descripción Detallada. El presente sumario no tiene como objeto identificar las características clave o las características esenciales de la materia objeto reivindicado, ni está destinado a utilizarse como ayuda para determinar el alcance de la materia objeto reivindicado.

55 Las realizaciones de la presente invención proporcionan un mecanismo que regula las solicitudes emitidas desde aplicaciones de servicio que se ejecutan en máquinas virtuales (VM), que residen en una máquina física común o nodo de un centro de datos, de acuerdo con un modelo de unidad de disco. Tal como se utiliza en la presente memoria descriptiva, la frase "modelo de unidad de disco" se refiere en general a una representación de disco blando de una unidad de disco duro (unidad de disco duro), en la que el modelo de unidad de disco simula las característi-

cas de rendimiento de la unidad de disco duro cuando calcula los tiempos de ejecución proyectados para realizar las solicitudes de aplicación de servicio. En una realización ejemplar, el modelo de unidad de disco simula las características de rendimiento de la unidad de disco duro considerando al menos los dos factores de la unidad de disco duro que siguen: un conjunto de parámetros que corresponden a un conjunto de parámetros reales que existen actualmente en la unidad de disco duro; y un estado actual del cabezal de disco de la unidad de disco duro. Como se explica con más detalle a continuación, el estado actual del cabezal de disco se puede usar para calcular un tiempo consumido moviendo el cabezal de disco desde una primera posición que satisface una solicitud que se está ejecutando actualmente a una segunda posición que satisface una solicitud pendiente.

Los enfoques anteriores para cuantificar el rendimiento de una unidad de disco duro se limitan a considerar ya sea MB / segundo (cantidad de información transferida) o solicitudes / segundo (velocidad de operación). El enfoque empleado en las realizaciones de la presente invención se basa en el modelo de unidad de disco que cuantifica con precisión el rendimiento de E / S del disco en base a diversos factores, incluyendo los dos factores que se han mencionado más arriba. Además, el enfoque en las realizaciones de la presente invención aplica el modelo de unidad de disco para determinar cómo se deben manejar las solicitudes de E / S de lectura o escritura de disco (en lo sucesivo, "solicitudes"), tales como limitar o "restringir", estando basadas las solicitudes en el hecho de la manera en la que la unidad de disco duro está funcionando en relación con el modelo de unidad de disco. La restricción adecuada de la E / S del disco es ventajosa en un entorno informático en nube por varias razones, tales como proporcionar una forma para desacoplar el rendimiento de disco de VM del hardware de almacenamiento físico subyacente (por ejemplo, la unidad de disco duro) y proporcionar la capacidad de asignar un umbral de rendimiento dinámicamente a una máquina virtual individual. Como se explica más detalladamente a continuación, el umbral de rendimiento se calcula comparando un tiempo de ejecución proyectado de una solicitud (utilizando el modelo de unidad de disco) y un tiempo de ejecución real de la solicitud (midiendo el tiempo que tarda la unidad de disco duro en realizar la solicitud).

En una realización ejemplar, el modelo de unidad de disco se utiliza con el fin de restringir las solicitudes de E / S del disco. Inicialmente, para cada operación de E / S del disco que se solicita, se calcula un tiempo de ejecución proyectado (es decir, el tiempo utilizado por el modelo de unidad de disco para manejar la operación de E / S del disco). En operación, el tiempo de ejecución proyectado utilizado por el modelo de unidad de disco se puede usar como una barra de línea de base para la decisión de restricción.

Una vez pasado a la unidad de disco duro, se mide un tiempo de ejecución real (es decir, el tiempo empleado por la unidad de disco duro para realizar la operación de E / S del disco). Si el tiempo de ejecución real de la unidad de disco duro es menor que el tiempo de ejecución proyectado del modelo de unidad de disco, la unidad de disco duro es superada. Cuando la unidad de disco duro supera el modelo de unidad de disco en cierta medida, las operaciones de E / S del disco subsiguientes se pueden ejecutar de forma retardada, o, en otras palabras, las solicitudes subsiguientes pueden ser restringidas. A modo de ejemplo, la restricción puede implicar la adición de tiempo al programar las solicitudes posteriores cuando se supera la unidad de disco duro. O bien, la restricción puede implicar el procesamiento inmediato de las solicitudes posteriores cuando la unidad de disco duro rinde de manera inferior. De esta manera, el uso del modelo de unidad de disco para determinar cuándo colocar selectivamente restricciones en el consumo de solicitudes puede mejorar la capacidad de la unidad de disco duro o afectar negativamente la calidad de servicio garantizada a los inquilinos del centro de datos en el que residen las máquinas virtuales, maximizando al mismo tiempo el rendimiento de las operaciones de E / S del disco. Además, en la práctica, la restricción se puede usar cuando la unidad de disco duro no está excesivamente sobrecargada. En un ejemplo, si se asignan diferentes niveles de rendimiento a diferentes tipos de máquinas virtuales, la restricción puede ayudar a reforzar estos niveles de rendimiento. En otro ejemplo, si la máquina virtual de un cliente en particular está garantizada con un nivel de rendimiento independientemente del lugar en el que se despliegue (por ejemplo, en una máquina física rápida o en una máquina más lenta), la restricción puede ayudar a asegurar que se garantiza el nivel de rendimiento garantizado.

Breve descripción de los dibujos

Las realizaciones de la presente invención se describen en detalle a continuación haciendo referencia a las figuras de los dibujos adjuntos, en las que:

la figura 1 es un diagrama de bloques de un entorno informático ejemplar adecuado para su uso en la implementación de realizaciones de la presente invención;

la figura 2 es un diagrama de bloques que ilustra una plataforma en nube ejemplar, adecuado para su uso en la implementación de realizaciones de la presente invención, que está configurado para emplear un modelo de unidad de disco para calcular un tiempo de ejecución proyectado de una solicitud entrante;

la figura 3 es un diagrama de bloques de un entorno informático distribuido ejemplar que representa las cosas dentro de un nodo físico que comunica con un controlador de filtro, de acuerdo con una realización de la presente invención;

la figura 4 es un diagrama de flujo que muestra un procedimiento ejemplar para restringir las solicitudes de una o más máquinas virtuales (VM) a una unidad de disco duro (HDD) utilizando el controlador de filtro, de acuerdo con una realización de la presente invención; y

5 la figura 5 es un diagrama de flujo que muestra un procedimiento ejemplar para restringir una velocidad de solicitudes que circulan a la HDD, de acuerdo con una realización de la presente invención.

Descripción detallada

La materia objeto de las realizaciones de la presente invención se describe con especificidad en la presente memoria descriptiva para cumplir con los requisitos legales. Sin embargo, la misma descripción no pretende restringir el alcance de esta patente. Por el contrario, los inventores han considerado que la materia reivindicada también podría estar incorporada de otras maneras, incluir diferentes pasos o combinaciones de pasos similares a los descritos en la presente memoria descriptiva, conjuntamente con otras tecnologías presentes o futuras.

10 Las realizaciones de la presente invención se refieren a procedimientos, sistemas informáticos y medios legibles por ordenador para aplicar un restrictor a una fuente (por ejemplo, máquinas virtuales (VM) instanciadas en un nodo físico) de solicitudes de operaciones de E / S del disco que deben ser manejadas por una unidad de disco duro (unidad de disco duro). El restrictor aplicado puede medir, en operación, las solicitudes de las máquinas virtuales de acuerdo con una comparación entre un tiempo de ejecución proyectado de una solicitud (utilizando el modelo de unidad de disco) y un tiempo de ejecución real de la solicitud (midiendo el tiempo que tarda la unidad de disco duro en realizar la solicitud). En una realización ejemplar, la medición implica retardar la entrega de aquellas solicitudes que son cuestionables después de identificar que el tiempo de ejecución proyectado excede al tiempo de ejecución real. A modo de ejemplo, la medición puede ser implementada por un controlador de filtro que hace que el servidor almacene, o "ponga en cola", las solicitudes retardadas emitidas por las máquinas virtuales.

15 En un aspecto, las realizaciones de la presente invención se refieren a uno o más medios legibles por ordenador que tienen instrucciones ejecutables por ordenador incorporadas sobre el mismo que, cuando se ejecutan, realizan un procedimiento para restringir solicitudes de una o más máquinas virtuales a la unidad de disco duro usando el controlador de filtro. Inicialmente, el procedimiento puede comenzar en el controlador de filtro que recibe una solicitud del sujeto de una aplicación de servicio que se ejecuta en una o más máquinas virtuales. Al analizar la solicitud, se accede al modelo de unidad de disco. En una realización ejemplar, el modelo de unidad de disco se comporta sustancialmente de manera similar al disco duro. En un caso particular, el proceso de acceso incluye leer la simulación del modelo de unidad de disco de las características de rendimiento de la unidad de disco duro y leer la simulación del modelo de unidad de disco de un estado actual de un cabezal de disco de la unidad de disco duro.

20 El procedimiento realizado por el controlador de filtro puede incluir además el paso de calcular un tiempo de ejecución proyectado para realizar la solicitud del sujeto como una función de las características de rendimiento y del estado actual del cabezal de disco. Además, el tiempo de ejecución real para realizar la solicitud del sujeto se puede medir cuando el controlador de filtro permite que la solicitud del sujeto pase a la unidad de disco duro. Utilizando una comparación del tiempo de ejecución proyectado y el tiempo de ejecución real, el tráfico de las solicitudes de las máquinas virtuales, que son emitidas después de la solicitud del sujeto a la unidad de disco duro, se restringe.

25 En otro aspecto, las realizaciones de la presente invención se refieren a un sistema informático capaz de restringir una velocidad a la que las solicitudes de una aplicación de servicio son ejecutadas por la unidad de disco duro, en el que la restricción es gestionada de acuerdo con el modelo de unidad de disco. El sistema informático incluye al menos una máquina física (por ejemplo, un nodo de un centro de datos) configurado con un medio de almacenamiento informático que tiene una pluralidad de componentes de software de ordenador que se ejecutan en el mismo. Inicialmente, los componentes del software del ordenador incluyen una máquina virtual, un bus de máquina virtual y un controlador de sistema de archivos que se ejecuta en la partición de servidor. En operación, la máquina virtual está configurada para alojar una porción de la aplicación de servicio (por ejemplo, una instancia de una función de aplicación de servicio o un programa de componentes). Esta porción de la aplicación de servicio es capaz de emitir solicitudes para leer datos o escribir datos en un archivo de disco duro virtual (VHD) asociado con la máquina virtual. Normalmente, una copia del archivo de disco duro virtual se mantiene en un disco duro de la unidad de disco duro.

30 En realizaciones adicionales del sistema informático, el controlador de filtro está configurado para comparar un tiempo de ejecución proyectado de una primera solicitud con un tiempo de ejecución real de la primera solicitud. En general, el "tiempo de ejecución proyectado" se calcula utilizando el modelo de unidad de disco que simula las características de rendimiento de la unidad de disco duro, mientras que el "tiempo de ejecución real" se mide al completar una solicitud en la unidad de disco duro. El controlador de filtro está configurado además para afectar una temporización para realizar la solicitud subsiguiente en base a un resultado de la comparación. El bus de máquina virtual está configurado para entregar la solicitud desde la aplicación de servicio al controlador de filtro y la cola o las colas están configuradas para mantener solicitudes emitidas desde la máquina virtual que fueron retardadas por el controlador de filtro. Normalmente, las solicitudes en estas colas se procesarán y se enviarán a la unidad de disco duro en un momento posterior.

En todavía otro aspecto adicional, las realizaciones de la presente invención se refieren a un procedimiento informatizado para restringir una velocidad de solicitudes que circulan a una unidad de disco duro. En las realizaciones, el procedimiento implica emplear el modelo de unidad de disco para calcular el tiempo de ejecución proyectado para que la unidad de disco duro realice una primera solicitud y medir el tiempo de ejecución real utilizado por la unidad de disco duro para realizar la primera solicitud. Cuando se calcula el tiempo de ejecución proyectado, el modelo de unidad de disco considera un tiempo para que un cabezal de disco (utilizado para leer datos o escribir datos en un disco duro de la unidad de disco duro) se desplace desde el último sector de datos orientados por la primera solicitud y al primer sector de datos orientados por una segunda solicitud.

En algún momento, el tiempo de ejecución proyectado se compara con el tiempo de ejecución real y la velocidad de las solicitudes a las que se permite pasar a la unidad de disco duro se gestiona en función de un resultado de la comparación. En un caso, el proceso de gestión incluye imponer un retardo a la solicitud antes de pasar la solicitud a la unidad de disco duro cuando el tiempo de ejecución proyectado excede el tiempo de ejecución real por un valor mayor que un umbral de rendimiento. En otro caso, el proceso de gestión incluye pasar la solicitud a la unidad de disco duro al recibir la solicitud cuando el tiempo de ejecución proyectado no excede el tiempo de ejecución real por un valor mayor que el umbral de rendimiento.

Habiendo descrito brevemente una descripción general de las realizaciones de la presente invención, se describe a continuación un entorno operativo ejemplar adecuado para implementar realizaciones de la presente invención. En particular, un entorno operativo ejemplar soporta la funcionalidad de un mecanismo (por ejemplo, el controlador de filtro 230 de la figura 2) que permite a las máquinas virtuales que residen en un nodo físico común enviar solicitudes de E / S del disco imponiendo restricciones en el consumo de recursos en el disco duro; por lo tanto, mejorar una condición tensionada de sobrecarga de la unidad de disco duro con solicitudes pendientes o afectar negativamente el rendimiento de la unidad de disco duro innecesariamente.

Entorno operativo

Haciendo referencia inicialmente a la figura 1 en particular, se muestra un entorno operativo ejemplar para implementar realizaciones de la presente invención y se designa en general como dispositivo informático 100. El dispositivo informático 100 es sólo un ejemplo de un entorno informático adecuado y no pretende sugerir ninguna restricción en cuanto al alcance de uso o funcionalidad de la invención. Tampoco se debe interpretar el dispositivo informático 100 como que tiene cualquier dependencia o requerimiento relacionado con cualquiera de los componentes ilustrados o combinación de los mismos.

La invención se puede describir en el contexto general de código informático o de las instrucciones utilizables por la máquina, que incluyen instrucciones ejecutables por ordenador, tales como módulos de programa, que son ejecutadas por un ordenador u otra máquina, tal como un asistente de datos personales u otro dispositivo portátil. Generalmente, los módulos de programa que incluyen rutinas, programas, objetos, componentes, estructuras de datos, etc., se refieren al código que realiza tareas particulares o implementa determinados tipos de datos abstractos. La invención se puede practicar en una variedad de configuraciones de sistema, incluyendo dispositivos portátiles, electrónica de consumo, ordenadores de propósito general, otros dispositivos informático especializados, etc. La invención también se puede practicar en entornos informáticos distribuidos en los que las tareas son realizadas por dispositivos de procesamiento a distancia que están enlazados a través de una red de comunicaciones.

Haciendo referencia a la figura 1, el dispositivo informático 100 incluye un bus 110 que acopla directa o indirectamente los siguientes dispositivos: memoria 112, uno o más procesadores 114, uno o más componentes de presentación 116, puertos de entrada / salida (E / S) 118, componentes de entrada / salida 120, y una fuente de alimentación ilustrativa 122. El bus 110 representa lo que pueden ser uno o más buses (tales como un bus de dirección, un bus de datos, o una combinación de los mismos). Aunque los diversos bloques de la figura 1 se muestran con líneas en aras de la claridad, en realidad, delinear varios componentes no es tan claro, y metafóricamente, las líneas serían más precisamente grises y difusas. Por ejemplo, se puede considerar que un componente de presentación tal como un dispositivo de visualización es un componente de E / S. Además, los procesadores tienen memoria. Los inventores reconocen que esta es la naturaleza de la técnica, y reiteran que el diagrama de la figura 1 es meramente ilustrativo de un dispositivo informático ejemplar que se puede usar en conexión con una o más realizaciones de la presente invención. No se hace distinción entre categorías tales como "estación de trabajo", "servidor", "ordenador portátil", "dispositivo portátil", etc., ya que todos están contemplados dentro del alcance de la figura 1 y se hace referencia a "dispositivo informático."

El dispositivo informático 100 incluye típicamente una variedad de medios legibles por ordenador. El medio legible por ordenador puede ser cualquier medio disponible al que pueda acceder el dispositivo informático 100 e incluye medios volátiles y no volátiles, medios removibles y no removibles. A modo de ejemplo y no de limitación, los medios legibles por ordenador pueden comprender medios de almacenamiento informático y medios de comunicación. Los medios de almacenamiento informático incluyen medios volátiles y no volátiles, removibles y no removibles implementados en cualquier procedimiento o tecnología para el almacenamiento de información, tales como instrucciones legibles por ordenador, estructuras de datos, módulos de programa u otros datos. Los medios de almacenamiento

informático incluyen, pero no se limitan a, RAM, ROM, EEPROM, memoria flash u otra tecnología de memoria, CD - ROM, discos versátiles digitales (DVD) u otro almacenamiento en disco óptico, cintas magnéticas, otros dispositivos de almacenamiento magnético, o cualquier otro medio que se pueda usar para almacenar la información deseada y a la que se puede acceder por el dispositivo informático 100. Los medios de comunicación típicamente incorporan instrucciones legibles por ordenador, estructuras de datos, módulos de programa u otros datos en una señal de datos modulada tal como una onda portadora u otro mecanismo de transporte e incluyen cualquier medio de suministro de información. El término "señal de datos modulada" significa una señal que tiene una o más de sus características establecidas o cambiadas de tal manera que codifican información en la señal. A modo de ejemplo, y sin restricción, los medios de comunicación incluyen medios cableados tales como una red cableada o una conexión directa por cable, y medios inalámbricos tales como medios acústicos, RF, infrarrojos y otros medios inalámbricos. Las combinaciones de cualquiera de lo anterior también deben incluirse dentro del alcance de los medios legibles por ordenador.

La memoria 112 incluye medios de almacenamiento informático en forma de memoria volátil y / o no volátil. La memoria puede ser removible, no removible, o una combinación de las mismas. Ejemplos de dispositivos de hardware incluyen memoria de estado sólido, discos duros, unidades de discos ópticos, etc. El dispositivo informático 100 incluye uno o más procesadores que leen datos de varias entidades tales como la memoria 112 o componentes de E / S 120. El o los componentes de presentación 116 presentan indicaciones de datos a un usuario u otro dispositivo. Los componentes de presentación ejemplares incluyen un dispositivo de visualización, un altavoz, un componente de impresión, un componente vibrante, etc.

Los puertos de E / S 118 permiten que el dispositivo informático 100 esté acoplado lógicamente a otros dispositivos que incluyen componentes de E / S 120, algunos de los cuales pueden ser incorporados. Los componentes ilustrativos incluyen un micrófono, una palanca de juegos, una almohadilla para juegos, una antena parabólica, un escáner, una impresora, un dispositivo inalámbrico, etc.

Sistema de implementación

Haciendo referencia a las figuras 1 y 2, un nodo físico 215 puede ser implementado por el dispositivo informático ejemplar 100 de la figura 1. Una o más de las máquinas virtuales (VM) 201, 202, 203 y 204 puede incluir porciones de la memoria 112 de la figura 1 y / o porciones de los procesadores 114 de la figura 1. En las realizaciones, cada una de las máquinas virtuales 201 - 204 puede representar una partición invitada respectiva del nodo 215, en la que la partición invitada está habilitada para alojar una aplicación de servicio o, al menos, una parte de la misma. Generalmente, la aplicación de servicio es propiedad de un inquilino (por ejemplo, un cliente) de una plataforma informática en nube que aloja la aplicación de servicio de una manera distribuida. El nodo 215 puede soportar además un servidor 210 con varios componentes 220, 225, 230, 235, 240 y 245 que se ejecutan sobre el mismo. En las realizaciones, el servidor 210 representa una partición raíz del nodo 215 en la que la partición raíz tiene la tarea de administrar en general solicitudes de las particiones invitadas, tales como solicitudes para implementar operaciones de E / S del disco en la unidad de disco duro (HDD) 260.

Volviendo a continuación a la figura 2, se ilustra un diagrama de bloques, de acuerdo con una realización de la presente invención, que muestra una plataforma informática en nube ejemplar 200 que está configurada para emplear un modelo de unidad de disco 235 para calcular un tiempo de ejecución proyectado de una solicitud entrante 270. Como se ilustra, la solicitud 270 se emite desde una instancia de rol 211 de una aplicación de servicio que está asignada a la primera VM 201. Se emplea un bus de VM 205 para transportar la solicitud 270 desde la primera VM 201 al servidor 210. En las realizaciones, el bus de VM 205 se utiliza para transportar tráfico de solicitudes desde las máquinas virtuales 201 - 204 en el nodo 215 al servidor 210 para la restricción antes del procesamiento en la unidad de disco duro 260.

Aunque sólo se ilustra una porción de aplicación de servicio (instancia de rol 211) en la figura 2, se debe entender que se pueden asignar y colocar a las máquinas virtuales 201 - 204 varios otros programas de componentes de otras aplicaciones de servicio, tal como se explica más detalladamente a continuación. Es decir, las máquinas virtuales 201 - 204 dentro del nodo 215 pueden estar ocupadas por diversas aplicaciones de servicio. A modo de ejemplo, una porción de la aplicación de servicio A puede ocupar la máquina virtual 202 mientras que porciones de la aplicación de servicio B pueden ocupar las máquinas virtuales 203 y 204, en la que cada una de las máquinas virtuales 202 - 204 se instancia en el nodo físico 215 para soportar la funcionalidad de las solicitudes de servicio A y B.

Se debe entender y apreciar que la plataforma informática en nube 200 que se muestra en la figura 2 es meramente un ejemplo de un entorno de sistema informático adecuado y no pretende sugerir ninguna restricción en cuanto al alcance de uso o funcionalidad de las realizaciones de la presente invención. Por ejemplo, la plataforma informática en nube 200 puede ser una nube pública, una nube privada o una nube dedicada. Tampoco se debería interpretar que la plataforma informática en nube 200 tenga cualquier dependencia o requisito relacionado con cualquier componente único o combinación de componentes ilustrados en la misma. Además, aunque los diversos bloques de la figura 2 se muestran con líneas en aras de la claridad, en realidad, delinear varios componentes no es tan claro, y metafóricamente, las líneas serían con más precisión grises y difusas. Además, se puede emplear cualquier número

de máquinas físicas, máquinas virtuales, centros de datos, puntos finales o combinaciones de los mismos para conseguir la funcionalidad deseada dentro del alcance de las realizaciones de la presente invención.

La plataforma informática en nube 200 incluye típicamente centros de datos configurados para alojar y soportar la operación de las máquinas virtuales 201 - 204 que alojan porciones (por ejemplo, instancia de rol 211 u otros puntos extremos) de aplicaciones de servicios distribuidos propiedad de inquilinos / clientes de la plataforma informática en nube 200. La expresión "aplicación de servicio", tal como se usa en la presente memoria descriptiva, se refiere ampliamente a cualquier software o porciones de software que se ejecuta en la parte superior de, o accede a las localizaciones de almacenamiento dentro de, la plataforma informática en nube 200. En una realización, la instancia de rol 211 puede representar una porción de software o un programa de componentes que participa en la funcionalidad de soporte de la aplicación de servicio. Se comprenderá y apreciará que la instancia de rol 211 que se muestra en la figura 2 es meramente un ejemplo de una parte adecuada para soportar una aplicación de servicio y no pretende sugerir ninguna restricción en cuanto al alcance de uso o funcionalidad de las realizaciones de la presente invención.

En general, las máquinas virtuales 201 - 204 se instancian y están designadas para una o más aplicaciones de servicio, respectivamente, en base a las demandas (por ejemplo, la cantidad de carga de procesamiento) colocadas sobre la aplicación de servicio. Como se usa en la presente memoria descriptiva, la frase "máquina virtual" no pretende ser limitativa, y se puede referir a cualquier software, aplicación, sistema operativo o programa que sea ejecutado por una unidad de procesamiento para subyacer la funcionalidad de la aplicación de servicio.

Además, las máquinas virtuales 201 - 204 pueden solicitar capacidad de procesamiento, acceso a ubicaciones de almacenamiento, recursos y otros activos (por ejemplo, agente servidores) dentro del nodo 215 para soportar adecuadamente las aplicaciones de servicio. En respuesta, las máquinas virtuales 201 - 204 pueden ser recursos asignados dinámicamente (por ejemplo, ancho de banda de red, capacidad de procesamiento de CPU o E / S del disco en la unidad de disco duro) para satisfacer una carga de trabajo actual. Como se ha indicado más arriba, el servidor 210 puede ayudar a asignar recursos a las máquinas virtuales 201 - 204. Específicamente, el controlador de filtro 230 del servidor 210 puede gestionar la restricción de solicitudes emitidas desde las máquinas virtuales 201 - 204 que son dirigidas a la unidad de disco duro 260 para su procesamiento.

Como se ha explicado más arriba, las máquinas virtuales 201 - 204 pueden ser recursos asignados dinámicamente dentro del nodo 215. Por realizaciones de la presente invención, el nodo 215 puede representar cualquier forma de dispositivo de cálculo, tal como, por ejemplo, un ordenador personal, un ordenador de sobremesa, un ordenador portátil, un dispositivo móvil, un dispositivo electrónico de consumo, un servidor o servidores, el dispositivo informático 100 de la figura 1, y otros similares. En una instancia, el nodo 215 soporta las operaciones de las máquinas virtuales 201 - 204, respectivamente, mientras aloja simultáneamente otras máquinas virtuales (no mostradas) preparadas para soportar otros inquilinos de la plataforma informática en nube 200, en la que los inquilinos incluyen puntos finales de otras aplicaciones de servicio pertenecientes a diferentes clientes.

Típicamente, el nodo 215 incluye o está conectado a alguna forma de una unidad informática (por ejemplo, una unidad de procesamiento central, un microprocesador, etc.) para soportar las operaciones de las máquinas virtuales 201 - 204 y / o componentes 220, 225, 230, 235, 240 y 245 que se ejecutan sobre las mismas. Tal como se utiliza en la presente memoria descriptiva, la frase "unidad informática" se refiere en general a un dispositivo informático dedicado con potencia de procesamiento y memoria de almacenamiento, que soporta uno o más sistemas operativos u otro software subyacente. En un caso, la unidad informática está configurada con elementos de hardware tangibles, o máquinas, que están integradas o acopladas operativamente al nodo 215 para permitir que cada punto extremo realice una variedad de procesos y operaciones. En otro caso, la unidad informática puede incluir un procesador (no mostrado) acoplado al medio legible por ordenador alojado por el nodo 215. Generalmente, el medio legible por ordenador almacena, al menos temporalmente, una pluralidad de componentes de software informático (por ejemplo, los componentes 220, 225, 230, 235, 240 y 245) que son ejecutables por el procesador. Tal como se utiliza en la presente memoria descriptiva, el término "procesador" no pretende ser limitativo y puede abarcar cualquier elemento de la unidad informática que actúe en una capacidad computacional. En tal capacidad, el procesador puede configurarse como un artículo tangible que procesa instrucciones. En una realización ejemplar, el procesamiento puede implicar la búsqueda, decodificación / interpretación, ejecución y escritura de instrucciones.

En un aspecto, las máquinas virtuales 201 - 204 operan dentro del contexto de la plataforma informática en nube 200 y, en consecuencia, se pueden comunicar internamente dentro del nodo 215, se pueden comunicar a través de nodos físicos por medio de conexiones realizadas dinámicamente que atraviesan un centro de datos y externamente a través de una topología de red física a recursos de una red remota (por ejemplo, red privada de empresa). Las conexiones pueden implicar la interconexión de máquinas virtuales distribuidas a través de recursos físicos del centro de datos por medio de una nube de red (no mostrada). La nube de red interconecta estos recursos de tal manera que una máquina virtual puede reconocer una posición de otra máquina virtual para establecer una comunicación entre ellas. Por ejemplo, la nube de red puede establecer esta comunicación a través de canales que conectan a máquinas virtuales que alojan programas de componente de una aplicación de servicio común. A modo de ejemplo, los canales pueden incluir, sin restricción, una o más redes de área local (LAN) y / o redes de área extensa (WAN).

Tales entornos de red son comunes en las oficinas, redes de ordenadores en la empresa, intranets y el Internet. Por consiguiente, la red no se describe adicionalmente en la presente memoria descriptiva.

La funcionalidad general de los componentes 220, 225, 230, 235, 240 y 245 que se ejecutan en el servidor 210 se describirá a continuación con respecto a la solicitud 270 emitida desde la instancia de rol 211. En realizaciones, el gestor de E / S 220 recibe inicialmente la solicitud 270 y analiza la solicitud para obtener una comprensión de las instrucciones en la misma. En una realización ejemplar, las instrucciones de la solicitud 270 pueden incluir un comando para realizar una operación de lectura o una operación de escritura que consume E / S del disco en la unidad de disco duro 260. El gestor de E / S 220 pasa entonces la solicitud 270 al gestor de filtro 225 para determinar si se debe gestionar la solicitud 270 de inmediato o imponer un retardo a la solicitud 270. El gestor de filtro 225 puede tener una pluralidad de tipos de filtros a su disposición para ayudar a gestionar adecuadamente el tráfico de solicitudes de las máquinas virtuales 201 - 204. En un caso, el gestor de filtro puede consultar el controlador de filtro 230 para acelerar el tráfico de solicitudes de las máquinas virtuales 201 - 204, de acuerdo con un modelo de unidad de disco 235 y un estado actual de la unidad de disco duro 260. Las realizaciones de la restricción se describirán más completamente a continuación con respecto a la figura 3. Una vez que el gestor de filtros 225 considera que se puede procesar la solicitud 270, el sistema de ficheros 240 interpreta la solicitud 270, determina el fichero del disco duro virtual (VHD) apropiado para acceder e identifica una posición del fichero de VHD. A continuación, la pila de controladores de almacenamiento 245 implementa la solicitud 270 controlando la unidad de disco duro 260 para leer o escribir en el archivo de disco duro virtual en un disco duro de la unidad de disco duro 260.

En realizaciones, el archivo de disco duro virtual en general representa un archivo de disco duro virtual con un formato particular aplicado al mismo. Este formato de archivo puede mantener el almacenamiento adecuado de datos en el disco duro 260, tales como las particiones de disco y un sistema de archivos. Estas particiones de disco pueden incluir datos estructurados, tales como archivos y carpetas. En operación, el archivo de disco duro virtual se puede utilizar como la memoria de disco duro de una máquina virtual.

A continuación se explicará una configuración general de la unidad de disco duro 260. Inicialmente, la unidad de disco duro 260 representa un dispositivo de almacenamiento magnético digital de acceso aleatorio no volátil. Incluido dentro de la unidad de disco duro 260 hay un plato rígido rotativo (por ejemplo, un disco duro) que es rotado por un husillo accionado por motor dentro de un recinto protector. Los datos son leídos magnéticamente del, y son escritos en el, disco duro por al menos un cabezal de disco. En una realización, la unidad de disco duro 260 emplea un mecanismo que mueve el cabezal del disco a una posición correcta y lee un número de bytes del disco duro. Mover el cabezal a la posición correcta consume un cierto tiempo que es considerado por el controlador de filtro 230 cuando se asignan tareas a la unidad de disco duro 260.

A menudo, la solicitud 270 pedirá a la unidad de disco duro 260 que se oriente a una porción particular de datos en el disco duro como parte de una operación de lectura o de escritura. Los datos orientados pueden estar situados dentro de un primer sector del disco duro mientras que otra porción de datos orientados por una solicitud subsiguiente puede estar situada dentro del segundo sector del disco duro. Además, el primer sector y el segundo sector pueden estar situados en pistas diferentes (por ejemplo, círculos concéntricos de información almacenada magnéticamente) grabadas en el disco duro. Por consiguiente, al completar la solicitud 270 y comenzar la realización de la solicitud subsiguiente, se consume un cierto tiempo cuando el cabezal del disco se desplaza una distancia lateral desde el primer sector al segundo sector del disco duro. En las realizaciones, los sectores que contienen datos sobre la superficie magnética del disco duro representan subdivisiones de tamaño micrométrico o dominios magnéticos asignados a un archivo o carpeta particular (por ejemplo, un archivo de disco duro virtual).

Al pasar de realizar la solicitud 270 a realizar la solicitud subsiguiente, la unidad de disco duro 260 realiza varios procesos que consumen tiempo y, por tanto, afectan a la programación de las solicitudes de las máquinas virtuales 201 - 204. Estos procesos que consumen mucho tiempo de la unidad de disco duro 260 incluyen la latencia del controlador, la velocidad de búsqueda, la latencia rotacional y la velocidad de transferencia de datos. La latencia del controlador se refiere a la sobrecarga que el controlador de disco (por ejemplo, el gestor de E / S 220) impone en la interpretación y ejecución de la solicitud 270. La velocidad de búsqueda, como se ha mencionado más arriba, se refiere al tiempo consumido para mover el cabezal del disco desde una pista del disco duro que mantiene el primer sector de datos (orientado por una solicitud pendiente) a otra pista en el cabezal del disco que mantiene el segundo sector de datos (orientado por una solicitud posterior). La latencia rotacional se refiere al tiempo consumido para hacer rotar el disco duro al inicio del segundo sector de datos una vez que el cabezal del disco se mueve a la pista apropiada. La velocidad de transferencia de datos se refiere a la velocidad a la que se hace rotar el disco duro y en la que se leen o escriben los datos en el disco duro (por ejemplo, el número de bloques de archivos contiguos transferidos desde o hacia el disco duro).

En una realización ejemplar, estos procesos consumidores de tiempo (por ejemplo, latencia del controlador, velocidad de búsqueda, latencia rotacional y velocidad de transferencia de datos) se cuantifican como un conjunto de parámetros reales que afectan al rendimiento de E / S del disco de la unidad de disco duro 260. Es decir, este conjunto de parámetros reales expone las características de operación de la unidad de disco duro 260 que se pueden emplear para influir en la restricción de las solicitudes dirigidas a la unidad de disco duro 260. Por ejemplo, se puede

incorporar un conjunto de parámetros estimados (que se corresponden sustancialmente con el conjunto de parámetros reales de la unidad de disco duro 260) que intentan especificar las características de rendimiento de la unidad de disco duro 260 dentro del modelo de unidad de disco 235.

5 Tal como se usa en la presente memoria descriptiva, la frase "modelo de unidad de disco" no pretende ser limitativa, sino que puede abarcar cualquier lógica, heurística o reglas basadas en software que ayudan al controlador de filtro 230 a gestionar el flujo del tráfico orientados hacia la unidad de disco duro 260. En una realización, el modelo de unidad de disco 235 es construido por el controlador de filtro 230 para imitar sustancialmente el comportamiento de la unidad de disco duro 260 programando el conjunto de parámetros estimados (que se ha explicado más arriba) dentro del modelo de unidad de disco 235. Además, el modelo de unidad de disco 235 se puede programar para
10 considerar un estado actual de la unidad de disco duro 260, que especifica la posición actual del cabezal del disco y el tiempo total consumido para satisfacer la solicitud pendiente 270. El estado actual de la unidad de disco duro 260 se puede usar para determinar el tiempo requerido para que el cabezal del disco se desplace a la siguiente posición en el disco duro orientado por la solicitud subsiguiente (por ejemplo, tiempo consumido para cubrir una distancia de desplazamiento al actualizar la posición del cabezal del disco). Por lo tanto, para la solicitud pendiente 270, el tiempo consumido se estima en base al tipo de acceso (lectura o escritura), la posición de los datos y la cantidad de procesamiento requerida para completar la solicitud 270. En consecuencia, el tiempo de ejecución proyectado determinado usando el modelo de unidad de disco 235, utilizado por el controlador de filtro 230 para gobernar la restricción de las solicitudes, considera tanto el conjunto de parámetros reales (captura de las características de rendimiento actuales de la unidad de disco duro 260) así como el tiempo consumido para la transición de una solicitud a otra (utilizando el cálculo de la distancia entre una posición de cabezal de disco actual y una posición de cabezal de disco siguiente).

A modo de antecedentes, el acceso secuencial implica leer los sectores del disco duro consecutivamente dentro de una pista particular. Es decir, el acceso secuencial comprende leer un sector adyacente después de otro sector físicamente próximo en el disco duro sin trasladar el cabezal del disco. Típicamente, hay una cantidad mínima de tiempo consumida para mover el cabezal del disco al siguiente sector físicamente próximo en el disco duro. El acceso aleatorio implica la lectura de datos en diferentes sectores que están en diversas ubicaciones (por ejemplo, diferentes pistas en el disco duro) en relación unas con las otras. Típicamente, hay una cantidad significativa de tiempo consumido cuando el cabezal del disco se desplaza de un sector a otro. Como se ha mencionado más arriba, la consideración del modelo de unidad de disco del conjunto de parámetros estimados a partir de la unidad de disco duro 260 y el estado actual / futuro de la posición del cabezal de disco representa el consumo de tiempo implicado tanto en el acceso secuencial como en el acceso aleatorio, respectivamente.

Operación del controlador del filtro

Volviendo a la figura 3, se ilustra un diagrama de bloques que muestra un entorno informático distribuido ejemplar 300 que representa las colas 301 - 304 dentro de uno o más nodos físicos que comunican con el controlador de filtro 230, de acuerdo con una realización de la presente invención. En las realizaciones, las máquinas virtuales 201 - 204 gestionan las colas 301 - 304, respectivamente, aunque cada una de las máquinas virtuales 201 - 204 puede tener múltiples colas para mantener temporalmente solicitudes pendientes. Una vez liberadas de las colas 301 - 304, las solicitudes son atendidas por el sistema de archivos en coordinación con la unidad de disco duro 260. Se debe hacer notar que cada una de las colas 301 - 304 puede ser gestionada por separado, en el que a cada una se le asigna un porcentaje del disco de software.

Inicialmente, las aplicaciones de servicio que se ejecutan dentro de las máquinas virtuales 201 - 204 pueden emitir solicitudes para realizar una operación de lectura y / o escritura en la unidad de disco duro 260. En un caso, estas operaciones de lectura y / o escritura pueden estar dirigidas a un servidor a través de un bus de máquina virtual y pueden dirigirse a los respectivos archivos de disco duro virtual de interés que residen en la unidad de disco duro 260. Cuando alcanza el servidor, el controlador de filtro 230 intercepta cada operación de lectura y escritura y las procesa antes de pasarlas al sistema de archivos. En general, el controlador de filtro 230 es responsable de la gestión de las solicitudes de E / S del disco restrictor.

Durante el procesamiento, el controlador de filtro 230 calcula el tiempo requerido para que el modelo de unidad de disco procese la operación de lectura o escritura que satisface la solicitud de la aplicación de servicio. Este cálculo considera al menos las dos entradas siguientes: un estado actual del modelo de unidad de disco; y atributos de la operación de lectura o escritura ejecutada actualmente (por ejemplo, tamaño de datos y su posición). Por ejemplo, el tiempo empleado para procesar la operación de lectura o escritura por el modelo de unidad de disco se puede calcular como la suma de las latencias enumeradas a continuación: latencia del controlador (por ejemplo, tiempo estimado para que el controlador de disco analice la solicitud); tiempo de búsqueda (por ejemplo, tiempo estimado para que el cabezal de disco se mueva desde una posición actual en una primera pista de la unidad de disco duro hasta una posición proyectada en una segunda pista de la unidad de disco duro en la que se encuentra localizado el dato a acceder); latencia rotacional (por ejemplo, tiempo estimado para que el sector solicitado rote a una posición bajo el cabezal del disco) y tiempo de transferencia de datos (por ejemplo, tiempo estimado para leer o escribir los datos de un tamaño dado). Por lo tanto, al considerar estas diversas entradas, el cálculo de un tiempo de ejecución proyecta-

do (utilizando el modelo de unidad de disco) explica tanto el acceso secuencial como el acceso aleatorio. Por ejemplo, incorporar el tiempo de búsqueda dentro del tiempo de ejecución proyectado ayuda a tener en cuenta un retardo incurrido durante el acceso aleatorio de datos, en el que el tiempo de búsqueda puede ser insignificante en valor cuando la solicitud se puede realizar mediante acceso secuencial de datos.

- 5 Al determinar el tiempo de ejecución proyectado, se mide el tiempo de ejecución real empleado por la pila de controlador de almacenamiento y el hardware de disco físico (por ejemplo, la unidad de disco duro 260) para manejar la operación de lectura y escritura dirigida al archivo de disco duro virtual. En un caso, el controlador de filtro 230 puede registrar el tiempo real consumido por el hardware de disco físico para procesar una solicitud del sujeto.

- 10 Al calcular el tiempo de ejecución proyectado y medir el tiempo de ejecución real, el controlador de filtro 230 puede implementar un análisis que determina la diferencia de tiempo entre las dos métricas anteriores (tiempo de ejecución proyectado y tiempo de ejecución real). Esta diferencia de tiempo indica si el hardware del disco físico está superando o es inferior al modelo de unidad de disco. Si el hardware físico de disco supera al modelo de unidad de disco, entonces la solicitud se puede retardar manteniéndola en una cola asociada con el archivo de disco duro virtual de la máquina virtual solicitante. Típicamente, hay una cola separada para mantener las solicitudes retardadas para cada disco duro virtual. A una máquina virtual se le puede asignar más de un disco duro virtual, por lo tanto, la máquina virtual puede estar asociada con más de una cola.

- 15 En un ejemplo específico, si la diferencia entre el tiempo de ejecución real y el tiempo de ejecución proyectado es mayor que un umbral de rendimiento predefinido asociado con la máquina virtual solicitante, las operaciones siguientes del disco duro 260 se ejecutarán de manera retardada o, en otras palabras, restringida. Haciendo referencia a la figura 3, por ejemplo, si las solicitudes de la cola 301 se procesan más rápidamente que el modelo de unidad de disco, entonces el controlador de filtro 230 puede elegir retardar las solicitudes de servicio emitidas desde la máquina virtual 201. En otra realización, si el hardware de disco físico supera al modelo de unidad de disco, la cantidad de retardo de la solicitud se puede basar en la diferencia de la comparación (por ejemplo, si el hardware del disco físico funciona mucho más rápido, se impone un retardo más largo). Por consiguiente, el controlador de filtro intenta sincronizar los tiempos predichos por el modelo de unidad de disco y el tiempo realmente tomado por la unidad de disco duro para procesar la solicitud.

- 20 Sin embargo, si la máquina virtual solicitante está recibiendo un mejor rendimiento que los límites predefinidos impuestos por el esquema de restricción actual (retardando innecesariamente el sistema de archivos), el controlador de filtro 230 levanta los retardos en las solicitudes. De este modo, el controlador de filtro 230 puede acelerar o disminuir el flujo de solicitudes para imitar las características de rendimiento de la unidad de disco duro. Por ejemplo, el controlador de filtro 230 está equipado para adaptarse a un nivel de rendimiento de la unidad de disco duro 260, tanto si la unidad de disco duro 260 representa hardware de disco físico en un escritorio personal o si la unidad de disco duro 260 representa un disco de grado comercial empleado en un centro de datos.

- 25 Al realizar sus funciones como gestor del flujo de solicitudes a la unidad de disco duro 260, el controlador de filtro 230 puede realizar varias rutinas de devolución de llamada para implementar estas funciones. Una de las rutinas de devolución de llamada implica una devolución de llamada de preoperación que es invocada por el gestor de filtros 230 para cada solicitud de E / S del disco antes de que la solicitud sea procesada por la unidad de disco duro 260. La devolución de llamada de preoperación, inicialmente determina si la solicitud se debe manejar inmediatamente. Esta decisión se basa en la diferencia del tiempo total consumido por el hardware de disco físico y el tiempo total proyectado para manejar las mismas solicitudes utilizando el modelo de unidad de disco. Si la decisión indica que la solicitud debe ser manejada inmediatamente, el controlador de filtro 230 pasa la solicitud al sistema de archivos y a la pila de almacenamiento del controlador (véanse los números de referencia 240 y 245 de la figura 2). Si la solicitud se debe manejar de una manera retardada, el controlador de filtro 230 inserta la solicitud en una cola apropiada y programa una rutina de temporizador que activa el procesamiento de la cola en un momento posterior. Como se ha mencionado más arriba, puede existir una cola separada para cada archivo de disco duro virtual que está siendo limitado. Además, puede haber uno o más archivos de disco duro virtual asociados con una única máquina virtual.

- 30 Otra de las rutinas de devolución de llamada implica una devolución de llamada de postoperación que es invocada por el gestor de filtro 230 para cada solicitud de E / S del disco después de que se procese la solicitud. La devolución de llamada de postoperación en las realizaciones registra el tiempo consumido por el hardware del disco físico para manejar la solicitud y actualiza el tiempo total consumido por el hardware del disco físico. Además, la devolución de llamada de postoperación calcula el tiempo de ejecución proyectado para manejar la solicitud accediendo al modelo de unidad de disco. El tiempo de ejecución proyectado se utiliza para actualizar el tiempo total utilizado por el modelo de unidad de disco. Además, el estado actual de la unidad de disco duro 260, tal como la posición del cabezal de disco, se puede almacenar para referencia futura en el modelo de unidad de disco.

55 Flujos de procesos ejemplares

Volviendo a continuación a la figura 4, se ilustra un diagrama de flujo que muestra un procedimiento ejemplar 400 para restringir solicitudes de una o más máquinas virtuales a la unidad de disco duro usando el controlador de filtro,

de acuerdo con una realización de la presente invención. Aunque los términos "paso" y / o "bloque" se pueden usar en la presente memoria descriptiva para connotar diferentes elementos de procedimientos empleados, los términos no deben ser interpretados como implicando un orden particular en medio de o entre varios pasos descritos en la presente memoria descriptiva salvo y excepto cuando el orden de los pasos individuales se describa explícitamente.

5 Inicialmente, el procedimiento 400 puede comenzar cuando el controlador de filtro recibe una solicitud del sujeto desde una aplicación de servicio que se ejecuta en las máquinas virtuales, como se indica en el bloque 410. A modo de ejemplo, la solicitud del sujeto puede representar una operación de lectura o una operación de escritura orientados a un archivo de disco duro virtual que mantiene datos almacenados en un disco duro de la unidad de disco duro, en el que el archivo de disco duro virtual está asociado con al menos una de las máquinas virtuales. Normalmente,

10 un cabezal de disco de la unidad de disco duro está configurado para leer o escribir en el disco duro en respuesta a la operación de lectura o a la operación de escritura, respectivamente.

Al analizar la solicitud en el controlador de filtro, se accede al modelo de unidad de disco, como se indica en el bloque 412. En una realización ejemplar, el modelo de unidad de disco se comporta sustancialmente de manera similar al disco duro y ejemplariza una representación de disco blando de la unidad de disco duro. En un caso particular, el

15 proceso de acceso incluye leer la simulación del modelo de unidad de disco de las características de rendimiento de la unidad de disco duro (véase el bloque 414) y leer la simulación del modelo de unidad de disco de un estado actual de un cabezal de disco de la unidad de disco duro (véase el bloque 416). Como se ha mencionado más arriba, las características de rendimiento simuladas se especifican mediante un conjunto de parámetros estimados que corresponden a un conjunto de parámetros reales existentes en la unidad de disco duro. El conjunto de parámetros estimados incluye al menos lo siguiente: una latencia de un controlador; una velocidad a la que busca el cabezal del disco; una latencia de rotación; y una velocidad de transferencia de datos. La latencia del controlador incluye la latencia que el controlador de disco impone para interpretar y ejecutar un comando de solicitud. La velocidad de búsqueda incluye una velocidad a la cual el cabezal del disco se desplaza a través de las pistas en un disco duro. La latencia de rotación incluye un tiempo para hacer rotar un disco duro a una posición de datos orientados por la solicitud del sujeto. Y la velocidad de transferencia de datos incluye una velocidad a la cual el cabezal de disco lee datos orientados por la solicitud del sujeto desde un disco duro.

20

25

El procedimiento 400 realizado por el controlador de filtro puede incluir además el paso de calcular un tiempo de ejecución proyectado para realizar la solicitud del sujeto como una función de las características de rendimiento y el estado actual del cabezal de disco, como se indica en el bloque 418. En una realización ejemplar, el cálculo de un

30 tiempo de ejecución proyectado para realizar la solicitud del sujeto implica al menos los siguientes pasos: identificar una primera posición del cabezal de disco cuando se realiza una solicitud precedente; extrapolar desde el modelo de unidad de disco una segunda posición del cabezal de disco cuando se realiza la solicitud del sujeto; y calcular una distancia de desplazamiento entre la primera posición y la segunda posición.

Además, el tiempo de ejecución real para realizar la solicitud del sujeto puede medirse cuando el controlador de filtro permite que la solicitud del sujeto pase a la unidad de disco duro, como se indica en el bloque 420. Usando una comparación del tiempo de ejecución proyectado y el tiempo de ejecución real, el tráfico de las solicitudes de las máquinas virtuales, que se emiten después de la solicitud del sujeto a la unidad de disco duro, se regulan, como se indica en el bloque 422. En operación, la restricción del tráfico de las solicitudes de las máquinas virtuales a la unidad de disco duro usando una comparación del tiempo de ejecución proyectado y el tiempo de ejecución real implica

35

40 determinar una cantidad de tiempo que la ejecución proyectada excede el tiempo de ejecución real. En una instancia, la cantidad de tiempo excedente se compara con un umbral de rendimiento asignado a la máquina virtual que emite la solicitud para determinar si retardar y / o poner en cola la solicitud. Una vez que la solicitud del sujeto se libera de la cola, se transporta a la unidad de disco duro para el cumplimiento de la solicitud del sujeto.

Volviendo a la figura 5, se ilustra un diagrama de flujo que muestra un procedimiento ejemplar 500 para restringir una velocidad de solicitudes que circulan a la unidad de disco duro, de acuerdo con una realización de la presente invención. En las realizaciones, como se indica en el bloque 510, el procedimiento 500 implica el empleo de un modelo de unidad de disco para calcular el tiempo de ejecución proyectado para la unidad de disco duro para realizar una primera solicitud. El procedimiento 500 puede implicar entonces medir el tiempo de ejecución real utilizado por la unidad de disco duro para realizar la primera solicitud, como se indica en el bloque 512. Cuando se calcula el

45

50 tiempo de ejecución proyectado, el modelo de unidad de disco considera un tiempo para un cabezal de disco (utilizado para leer datos o escribir datos en un disco duro de la unidad de disco duro) para desplazarse desde un primer sector de datos orientado por la primera solicitud y un segundo sector de datos orientado por una segunda solicitud.

Como se indica en los bloques 514 y 516, el tiempo de ejecución proyectado se compara con el tiempo de ejecución real y la velocidad de las solicitudes permitidas para pasar al disco duro se gestiona en función del resultado de la comparación. En un caso, como se indica en el bloque 518, el proceso de gestión incluye imponer un retardo a la segunda solicitud antes de pasar la segunda solicitud a la unidad de disco duro cuando el tiempo de ejecución proyectado excede el tiempo de ejecución real por un valor mayor que un umbral de rendimiento. En otro caso, como se indica en el bloque 520, el proceso de gestión incluye pasar la segunda solicitud a la unidad de disco duro al recibir la segunda solicitud cuando el tiempo de ejecución proyectado no excede el tiempo de ejecución real por un valor

55

60 mayor que el umbral de rendimiento.

Se han descrito realizaciones de la presente invención en relación con realizaciones particulares, que pretenden ser en todos los aspectos ilustrativas en lugar de restrictivas. Realizaciones alternativas resultarán evidentes a los expertos en la técnica a las que pertenecen las realizaciones de la presente invención sin apartarse de su alcance.

- 5 De lo que antecede, se verá que esta invención está bien adaptada para obtener todos los extremos y objetos expuestos más arriba, junto con otras ventajas que son evidentes e inherentes al sistema y procedimiento. Se comprenderá que ciertas características y subcombinaciones son de utilidad y se pueden emplear sin referencia a otras características y subcombinaciones. Esto está contemplado y se encuentra dentro del alcance de las reivindicaciones.

REIVINDICACIONES

- 5 1. Uno o más medios legibles por ordenador que tienen instrucciones ejecutables por ordenador incorporadas en los mismos que, cuando se ejecutan, realizan un procedimiento para restringir solicitudes de una o más máquinas virtuales (VM) a una unidad de disco duro utilizando un controlador de filtro, comprendiendo el procedimiento:
- recibir una solicitud del sujeto desde una aplicación de servicio que se ejecuta en las una o más máquinas virtuales (410);
- acceder a un modelo de unidad de disco que se comporta sustancialmente de manera similar a la unidad de disco duro (412), en el que el proceso de acceso comprende:
- 10 (a) leer la simulación del modelo de unidad de disco del rendimiento de las características de la unidad de disco duro; y
- (b) leer la simulación del modelo de unidad de disco de un estado actual de un cabezal de disco de la unidad de disco duro;
- 15 calcular un tiempo de ejecución proyectado para realizar la solicitud del sujeto como una función de las características de rendimiento y del estado actual del cabezal de disco (418);
- medir un tiempo de ejecución real para realizar la solicitud del sujeto al pasar la solicitud del sujeto a la unidad de disco duro; y
- restringir el tráfico de solicitudes de las máquinas virtuales a la unidad de disco duro utilizando una comparación del tiempo de ejecución proyectado y el tiempo de ejecución real (422)
- 20 2. Los uno o más medios legibles por ordenador de la reivindicación 1, en los que el procedimiento comprende, además, enviar la solicitud del sujeto a una unidad de disco duro para el cumplimiento de la solicitud del sujeto, en la que la solicitud del sujeto representa una operación de lectura o una operación de escritura orientados a un archivo de disco duro virtual (VHD) que mantiene los datos almacenados en un disco duro de la unidad de disco duro, y en el que el cabezal de disco está configurada para leer o escribir en el disco duro en respuesta a la operación de lectura o la operación de escritura, respectivamente.
- 25 3. Los uno o más medios legibles por ordenador de la reivindicación 2, comprendiendo el procedimiento, además, construir el modelo de unidad de disco, en el que el modelo de unidad de disco incorpora una representación de disco blando de la unidad de disco duro, que soporta archivos de disco duro virtual asociados con la una o más máquinas virtuales.
- 30 4. Los uno o más medios legibles por ordenador de la reivindicación 1, en los que el cálculo de un tiempo de ejecución proyectado para realizar la solicitud del sujeto comprende:
- identificar una primera posición del cabezal de disco cuando se realiza una solicitud precedente;
- extrapolar desde el modelo de unidad de disco una segunda posición del cabezal de disco cuando se realiza la solicitud del sujeto; y
- 35 calcular un tiempo para que el cabezal de disco se desplace entre la primera posición y la segunda posición.
5. Los uno o más medios legibles por ordenador de la reivindicación 1, en los que las características de rendimiento simuladas se especifican mediante un conjunto de parámetros estimados que corresponden a un conjunto de parámetros reales existentes en la unidad de disco duro, en los que el conjunto de parámetros estimados comprende una latencia de un controlador, y en el que la latencia del controlador incluye un tiempo consumido por el controlador para interpretar y ejecutar la solicitud del sujeto.
- 40 6. Los uno o más medios legibles por ordenador de la reivindicación 5, en los que el conjunto de parámetros estimados comprende una velocidad a la cual el cabezal de disco busca, y en la que la velocidad de búsqueda incluye una velocidad a la cual el cabezal de disco se desplaza a través de pistas en un disco duro.
- 45 7. Los uno o más medios legibles por ordenador de la reivindicación 5, en los que el conjunto de parámetros estimados comprende una latencia de rotación, y en el que la latencia rotacional incluye un tiempo para hacer rotar un disco duro hasta una posición de datos orientados por la solicitud del sujeto.
8. Los uno o más medios legibles por ordenador de la reivindicación 5, en los que el conjunto de parámetros estimados comprende una velocidad de transferencia de datos y en el que la velocidad de transferencia de datos

incluye una velocidad a la cual el cabezal de disco lee o escribe datos orientados por la solicitud del sujeto a o desde un disco duro.

- 5 9. Los uno o más medios legibles por ordenador de la reivindicación 1, en los que el tráfico de restricción de las solicitudes de las máquinas virtuales a la unidad de disco duro usando una comparación del tiempo de ejecución proyectado y el tiempo de ejecución real comprende determinar la cantidad de tiempo que la ejecución proyectada excede al tiempo de ejecución real, comprendiendo el procedimiento, además, asignar a una de las máquinas virtuales un umbral de rendimiento que se basa en la cantidad de tiempo excedente.
- 10 10. Los uno o más medios legibles por ordenador de la reivindicación 9, retardando el procedimiento adicionalmente en una o más colas las solicitudes de las máquinas virtuales a la unidad de disco duro en proporción a un valor del umbral de rendimiento.
- 15 11. Un sistema informático (200) capaz de restringir una velocidad a la que las solicitudes de una aplicación de servicio son ejecutadas por una unidad de disco duro que utiliza un modelo de unidad de disco, comprendiendo el sistema informático al menos una máquina física configurada con un medio de almacenamiento informático (260) que tiene una pluralidad de componentes de software de ordenador que se ejecutan en el mismo, comprendiendo los componentes de software de ordenador:
- una máquina virtual VM (201) para alojar una porción de la aplicación de servicio, en la que la aplicación de servicio emite una segunda solicitud para leer datos o escribir datos en un archivo de disco duro virtual asociado con la máquina virtual;
- 20 un controlador de filtro (230) para comparar un tiempo de ejecución proyectado de una primera solicitud, que se calcula utilizando el modelo de unidad de disco (235) que simula las características de operación de la unidad de disco duro, frente a un tiempo de ejecución real de la primera solicitud, que se mide al realizar la primera solicitud en la unidad de disco duro, y para afectar un tiempo para realizar la segunda solicitud en base a un resultado de la comparación.
- 25 12. El sistema informático de la reivindicación 11, que comprende, además, un bus de máquina virtual para entregar la solicitud desde la aplicación de servicio al controlador de filtro y una o más colas para mantener las solicitudes emitidas desde la máquina virtual y retardadas por el controlador de filtro, en el que las una o más colas son alimentadas a la unidad de disco duro.
- 30 13. El sistema informático de la reivindicación 12, en el que la primera solicitud representa una operación de lectura o escritura que orienta datos sobre un primer sector de la unidad de disco duro, en la que la segunda solicitud representa una operación de lectura o escritura que orienta datos sobre un segundo sector del disco duro y en el que un cabezal de disco de la unidad de disco duro está configurado para leer o escribir en el disco duro en respuesta a la operación de lectura o escritura.
- 35 14. Un procedimiento informático para restringir una velocidad de solicitudes que circulan a una unidad de disco duro, comprendiendo el procedimiento:
- emplear un modelo de unidad de disco para calcular un tiempo de ejecución proyectado para la unidad de disco duro para realizar una primera solicitud;
- medir un tiempo de ejecución real utilizado por la unidad de disco duro para realizar la primera solicitud;
- comparar el tiempo de ejecución proyectado con el tiempo de ejecución real (514); y
- 40 gestionar la velocidad de las solicitudes a las que se permite pasar a la unidad de disco duro en función de un resultado de la comparación, comprendiendo el proceso de gestión:
- (a) cuando el tiempo de ejecución proyectado excede el tiempo de ejecución real por un valor mayor que un umbral de rendimiento, (518), imponer un retardo a una segunda solicitud antes de pasar la segunda solicitud a la unidad de disco duro, en el que el retardo es impuesto hasta que la diferencia entre el tiempo de ejecución real y el tiempo de ejecución proyectado satisface el umbral de rendimiento; y
- 45 (b) en caso contrario, pasar la segunda solicitud a la unidad de disco duro al recibir la segunda solicitud (520).
- 50 15. El procedimiento informático de la reivindicación 14, en el que la unidad de disco duro incluye un cabezal de disco para leer datos o escribir datos en un disco duro, y en el que el modelo de unidad de disco considera un tiempo para que el cabezal de disco se desplace desde un primer sector de datos orientados por la primera solicitud a un segundo sector de datos orientados por la segunda solicitud.

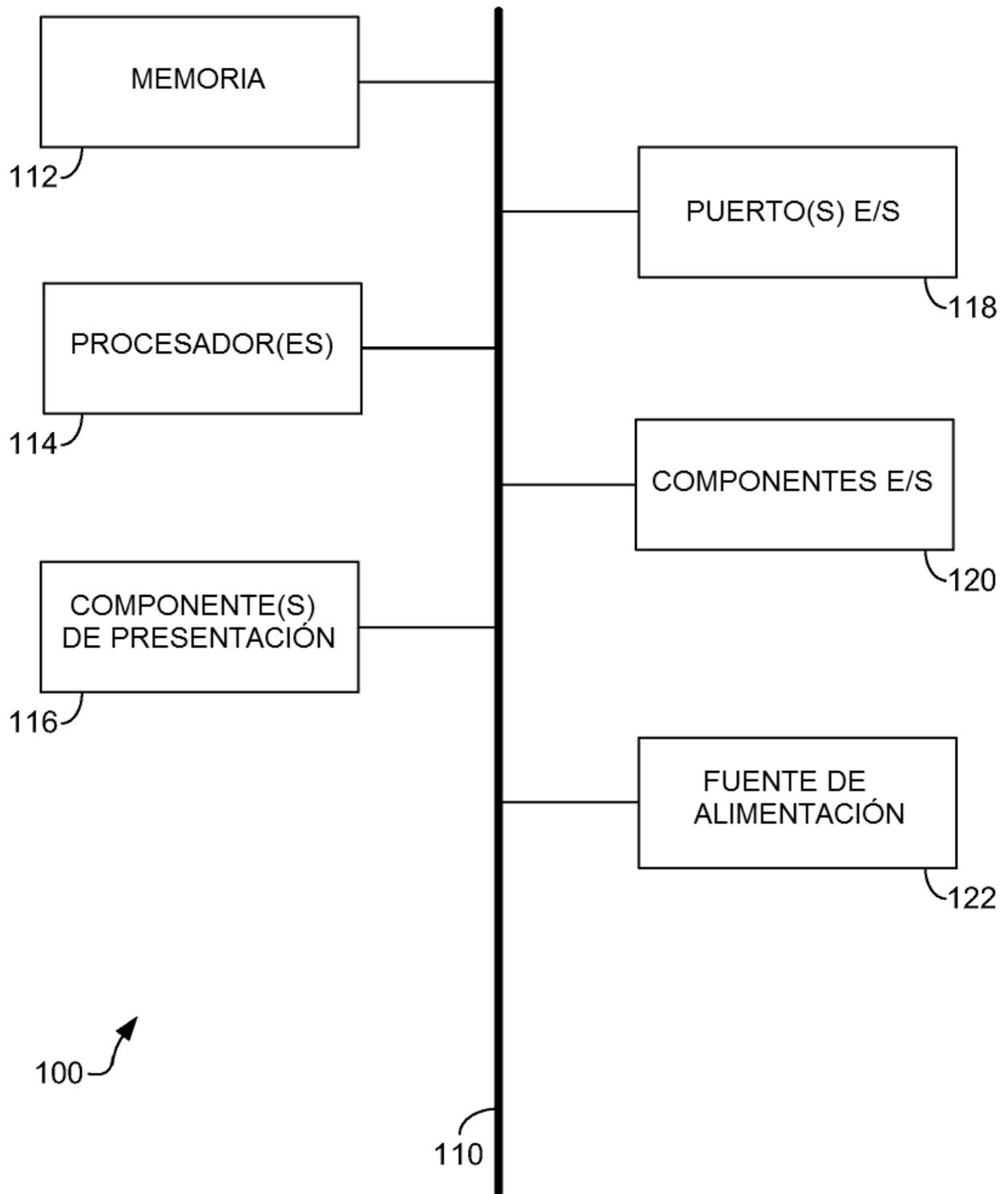


FIG. 1.

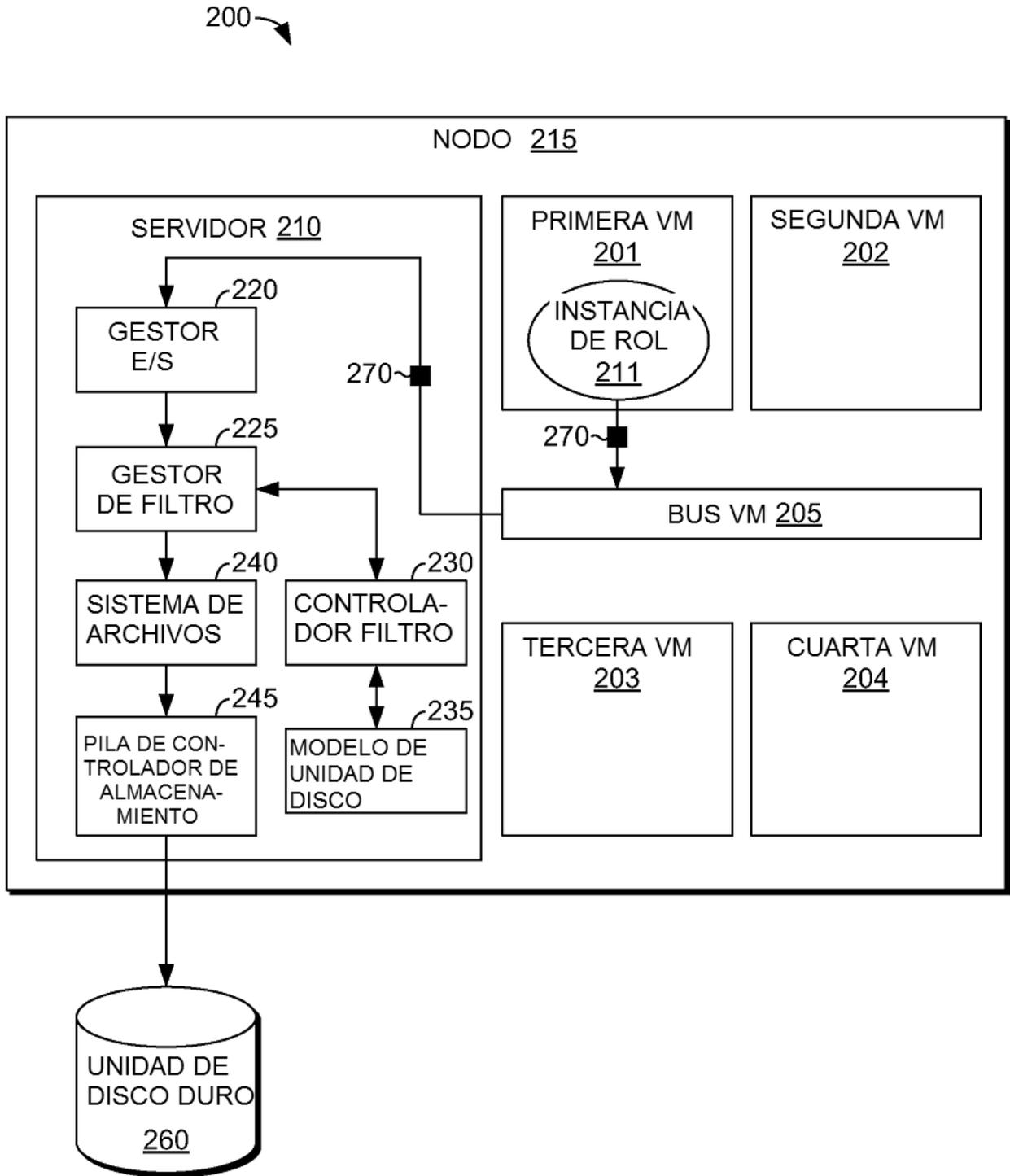


FIG. 2.

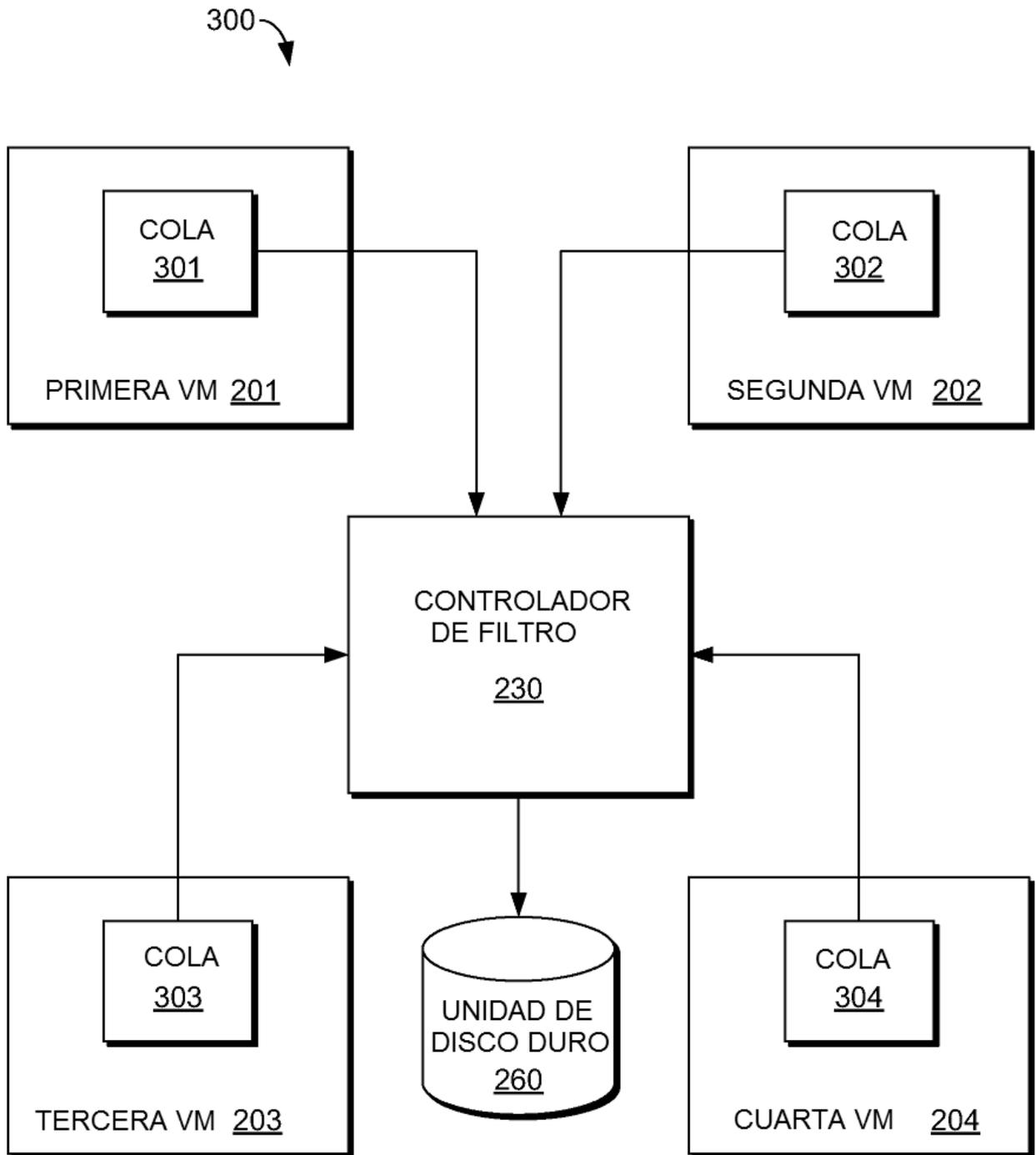


FIG. 3.

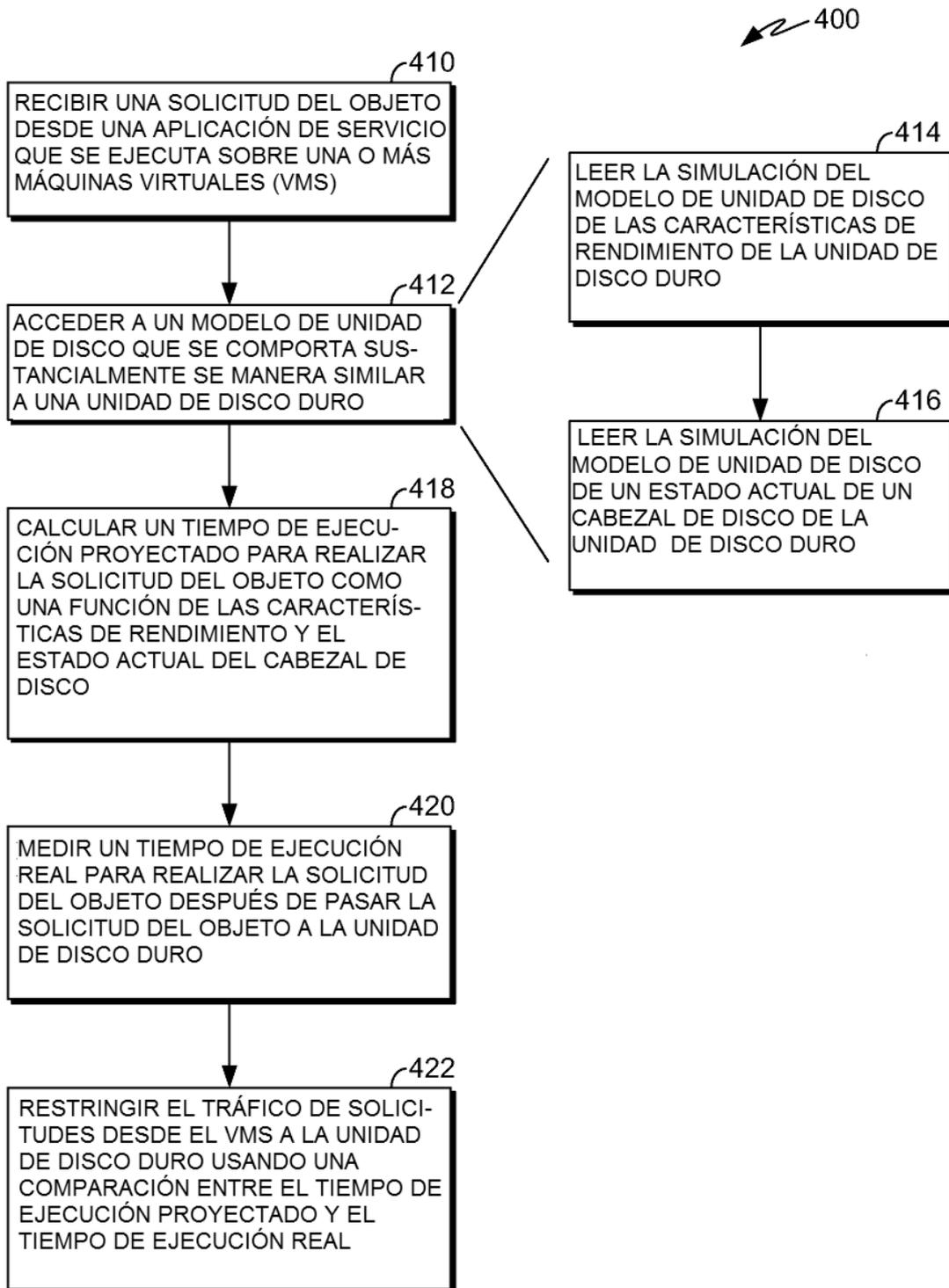


FIG. 4.

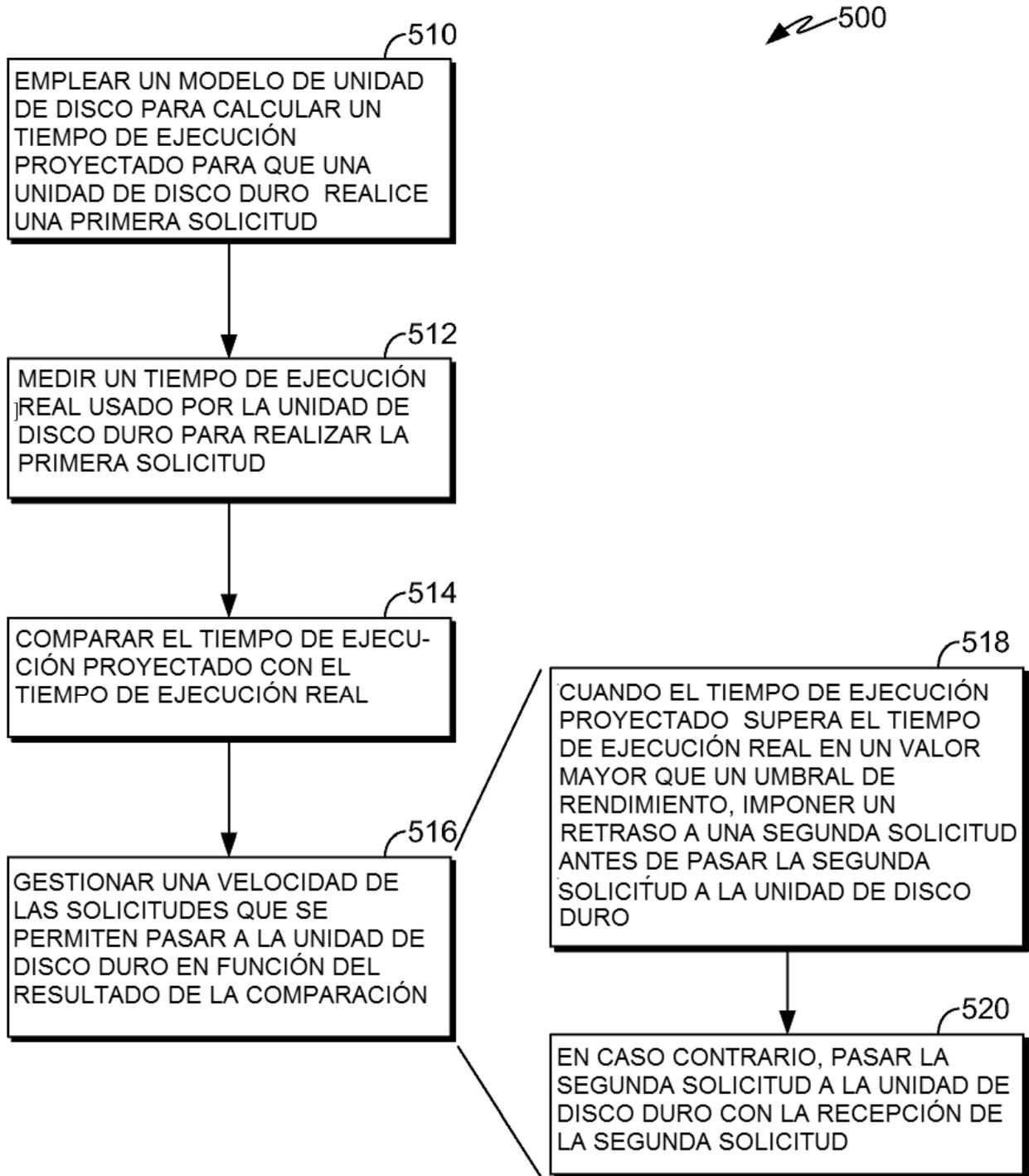


FIG. 5.