

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 713 459**

51 Int. Cl.:

G06F 17/30 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **26.05.2016** E 16171563 (6)

97 Fecha y número de publicación de la concesión europea: **28.11.2018** EP 3249551

54 Título: **Un sistema de control para controlar el funcionamiento de una unidad de procesamiento de datos**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
21.05.2019

73 Titular/es:

**IG KNOWHOW LIMITED (100.0%)
Cannon Bridge House 25 Dowgate Hill
London EC4R 2YA, GB**

72 Inventor/es:

**LEVINE, BORIS y
CLOKE, JAMIE ROLAND**

74 Agente/Representante:

LINAGE GONZÁLEZ, Rafael

ES 2 713 459 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

Un sistema de control para controlar el funcionamiento de una unidad de procesamiento de datos

5 **Campo técnico**

La presente invención se refiere a un sistema de control para controlar el funcionamiento de una unidad de procesamiento de datos que comprende un grupo de módulos de procesamiento interconectados, y más particularmente pero no exclusivamente, se refiere a un sistema de control y un método no invasivo para el control automatizado de una unidad de procesamiento de datos compleja que usa la medición de la latencia del procesamiento de registros dentro de la unidad de procesamiento de datos, y que toma medidas basándose en los resultados de las mediciones de latencia para minimizar la latencia de la unidad de procesamiento de datos.

15 **Antecedentes**

Existe la necesidad de una medición e informe precisos y en tiempo real de la información de latencia para una red de procesamiento de datos compleja que recibe una pluralidad de flujos de datos de diferentes proveedores de datos, y comprende una red de unidades de procesamiento de datos interconectadas (en lo sucesivo, también denominadas indistintamente como cajas de procesador) a través de la cual se transmiten los flujos de datos, antes de enviarse a uno o más receptores de datos. Cada flujo de datos corresponde a un flujo continuo de registros de datos recibidos de los proveedores de datos a lo largo del tiempo y en la figura 3 se pueden ver ejemplos de la estructura de datos de algunos de los registros en estos flujos de datos. Los registros de datos recibidos por tales redes de procesamiento de datos típicamente tienen una tasa de llegada muy alta; en algunos casos, una red puede recibir miles o incluso cientos de miles de registros de datos por segundo para su procesamiento.

Para añadir una capa de complejidad adicional, no solo hay varias caja de procesador interconectados dentro de una red de procesamiento de datos compleja, sino que cada caja de procesador también puede comprender múltiples subprocesadores interconectados (en lo sucesivo, también denominados indistintamente como módulos de procesamiento) que leen, analizan y procesan los registros de datos que pasan a través de las caja de procesador. La información transportada por los registros de datos en diferentes flujos de datos (recibidos de diferentes proveedores de datos) puede variar y los pasos exactos de procesamiento y análisis que se requieren para un registro de datos en un flujo de datos pueden diferir de los requeridos para otro registro de datos en un flujo de datos diferente. Por lo tanto, puede verse que la ruta tomada por un registro de datos particular dentro de tal red variará basándose en el procesamiento y análisis que se lleve a cabo para ese registro.

En redes tan complejas, es importante identificar lo antes posible cualquier problema que surja, lo que podría aumentar la latencia del sistema. También es importante poder diagnosticar y abordar la causa del aumento de la latencia; esto requiere una localización precisa de la porción particular de la red que está experimentando un retraso en el procesamiento en un momento determinado, a fin de permitir que se tomen las medidas adecuadas para intentar minimizar la latencia de la red en su conjunto.

Se conoce un sistema para medir y reportar la información de latencia mencionada anteriormente para una red de procesamiento de datos compleja. La implementación de este conocido sistema de medición de latencia se ilustra en la figura 1 en contexto con una unidad de procesamiento de datos de ejemplo (donde varias unidades de procesamiento de datos están comprendidas dentro de la red de procesamiento de datos) y se describe a continuación.

La figura 1 muestra una caja 1 de procesador de la técnica anterior (que comprende una parte de la red de procesamiento de datos global que está bajo análisis) que recibe y procesa flujos 2 de datos entrantes de diversos proveedores de datos. Los flujos 3 de datos procesados se envían posteriormente a otras cajas de procesador en la red de procesamiento de datos, o a uno o más receptores de datos si la caja de procesador actual es el componente final en la red de procesamiento de datos. La caja de procesador está en comunicación operativa con un almacén 4 de datos que almacena datos de un módulo de registro de datos. Puede proporcionarse un módulo de análisis de latencia fuera de línea 5 para analizar los datos para determinar si se ha producido algún caso de latencia incrementada dentro de la caja 1 de procesador.

La caja 1 de procesador comprende múltiples componentes a través de los cuales pasan los registros de datos que forman los flujos de datos entrantes, a saber, una entrada 6 de flujo de datos, un bloque 8 de procesamiento principal y un módulo 10 de registro. Cada registro de datos entrante es analizado por la entrada 6 de flujo de datos y, posteriormente, reenviado a una entrada apropiada (no mostrada) del bloque 8 de procesamiento principal basándose en ciertas propiedades de ese registro de datos. El bloque 8 de procesamiento principal comprende cientos o miles de subprocesadores individuales o módulos de procesamiento (que tampoco se muestran en esta figura) que leen, procesan y escriben datos en los registros de datos que pasan. Una descripción más detallada de cómo los registros de datos interactúan con los subprocesadores o módulos de procesamiento dentro del bloque 8 de procesamiento principal se proporcionará más adelante con referencia a la figura 2. Al salir del bloque 8 de procesamiento principal, los registros de datos se pasan al módulo 10 de registro (otro componente del sistema de

medición de latencia), donde se extrae (registra) la información relacionada con el procesamiento realizado para cada registro de datos, junto con la información de marca de tiempo correspondiente para cada evento que ocurrió.

5 La información extraída por el módulo 10 de registro se almacena en el almacén 4 de datos en tiempo real. Sin embargo, el módulo 5 de análisis de latencia que está en comunicación operativa con el almacén 4 de datos analiza los datos almacenados fuera de línea (no en tiempo real). El módulo 5 de análisis de latencia fuera de línea utiliza esta información extraída para calcular el tiempo que le llevó a cada registro de datos recorrer el bloque de procesamiento principal y, por lo tanto, obtener una medición de la latencia asociada con esa ruta tomada por el registro de datos. Esta información de latencia se compara con la información de datos históricos (o contra los valores de latencia esperados para una ruta en particular) que se recupera del almacén 4 de datos, para identificar cualquier problema de latencia que surja a lo largo de rutas específicas dentro del bloque 8 de procesamiento principal. Si se identifica algún caso de mayor latencia, el módulo 4 de análisis de latencia fuera de línea puede proporcionar comandos 11 de realimentación a la entrada 6 de flujo de datos que le indica que realice ciertas acciones destinadas a disminuir la latencia (por ejemplo, seleccionar una entrada de flujo de datos diferente para disminuir la frecuencia con la que los registros de datos se introducen en el bloque de procesamiento principal). Alternativamente, como es más frecuente el caso, el módulo 5 de análisis de latencia fuera de línea genera una alarma de tal manera que un operador humano pueda revisar los resultados y, si es posible, tomar medidas correctivas.

20 Se apreciará que puede haber miles de registros de datos que pasan por el bloque 8 de procesamiento principal en un momento dado; cada subprocesador puede configurarse para escribir una marca de tiempo y un identificador único asociado con ese subprocesador en el registro de datos para proporcionar un registro de todas las interacciones entre los subprocesadores y el registro de datos, y el momento en que ocurrió la interacción. Como los registros de datos a menudo son procesados por múltiples subprocesadores, cada registro de datos almacena una lista de marcas de tiempo e identificadores que deben analizarse para determinar la latencia. El resultado acumulativo de miles de registros de datos procesados por decenas o incluso cientos de subprocesadores por segundo significa que el volumen de información que se almacena en el almacén de datos y luego se procesa mediante el módulo de análisis de latencia fuera de línea es enorme y, por lo tanto, se requiere una gran capacidad de procesamiento y almacenamiento. Por lo tanto, el módulo 5 de análisis de latencia sin conexión y su almacén de datos asociado 4 se ubican de manera remota desde la caja 1 de procesador (por ejemplo, en un servidor separado), para evitar que el procesamiento de los registros de datos se ralentice a medida que pasan por la caja de procesador. Por lo tanto, es necesario que el análisis y los cálculos de latencia se realicen siempre fuera de línea (es decir, no en tiempo real) ya que la cantidad de tiempo requerido para procesar ese volumen de datos es enorme.

35 Se debe tener en cuenta que como una red de procesamiento de datos compleja típica generalmente comprende múltiples cajas de procesador interconectadas, es posible que el sistema de medición de latencia se implemente de tal manera que cada caja de procesador tenga un módulo de análisis de latencia asociado. Adicional o alternativamente, se puede proporcionar un módulo de análisis de latencia general al final de la red que analiza la información acumulada extraída de los registros de datos una vez que han pasado a través de toda la red, proporcionando una imagen general de la latencia de la red. Tales módulos de análisis de latencia general sufrirán aún más los tiempos de procesamiento retrasados debido a los enormes volúmenes de datos analizados.

45 Ahora se proporcionará una descripción más detallada del bloque de procesamiento principal con referencia a la figura 2.

El bloque 8 de procesamiento principal comprende una pasarela 12 de entrada para recibir registros de datos, una pasarela 14 de salida para enviar registros de datos a los componentes aguas abajo de la red de procesamiento de datos, así como múltiples subprocesadores o módulos 16a-e de procesamiento (etiquetados de A a E en la figura) entre las dos pasarelas que leen, extraen y analizan la información de los registros de datos entrantes; la información también puede escribirse en los registros de datos por estos subprocesadores 16a-e. Cada uno de los subprocesadores 16a-e corresponde efectivamente a una máquina de estados y hay múltiples permutaciones de mapeo entre los registros de datos de entrada para cada subprocesador y los registros de datos de salida (si hay alguno) de ese subprocesador, dependiendo de las propiedades de los registros de datos de entrada y el tipo de procesamiento que lleva a cabo cada subprocesador.

55 Los registros de datos entrantes típicamente se proporcionan en uno de tres formatos, dependiendo del proveedor de datos que generó el registro de datos y el tipo de información que se pretende que esté contenida en el registro de datos. Los ejemplos de los diferentes formatos de registros de datos se muestran en la figura 3 y están etiquetados como A, B y C. Cada registro 20a, b, c de datos comprende al menos la siguiente información: un identificador 22a, b, c de formato (ID) que permite la identificación del formato del registro de datos y una pluralidad de campos 24a, b, c de datos que se rellenan con diferentes valores de parámetros dependiendo del formato del registro de datos. Algunos de los campos de datos pueden ser comunes en todos los formatos de registro de datos, mientras que otros pueden ser específicos de un formato de registro de datos en particular.

65 La ruta recorrida a través del bloque 8 de procesamiento principal por un registro de datos particular, representada en la figura 2 por las líneas dibujadas entre los subprocesadores 16a-e, puede ser determinada total o parcialmente

por el formato del registro de datos, y/o por las condiciones dinámicas que están vinculadas a los valores contenidos en los registros de datos en sí mismas o se basan en condiciones externas (por ejemplo, la hora del día o la presencia de flujos de datos de entrada externos). Por ejemplo, un registro de datos con un primer formato puede entrar en el bloque 8 de procesamiento principal y ser procesado por los subprocesadores A y C antes de ser transmitido al módulo 10 de registro (como lo indica la ruta de la línea continua entre los componentes relevantes); otro registro de datos con un segundo formato puede recorrer una de dos rutas dentro del bloque 8 de procesamiento principal, ya sea entre los subprocesadores A, B y D o entre los subprocesadores A, B y C (ambas rutas están indicadas por líneas discontinuas), donde la elección entre los subprocesadores C y D es arbitraria.

5
10 Se observará que no hay una correspondencia uno a uno entre el número de registros de datos introducidos y el número de registros de datos que salen de cada subprocesador 16a-e y, de hecho, del bloque 8 de procesamiento principal: las rutas recorridas por algunos de los registros de datos a través del bloque de procesamiento principal terminan antes de la pasarela 14 de salida y, por lo tanto, algunos registros de datos de entrada se incluyen en el bloque 8 de procesamiento principal y nunca se envían al módulo 10 de registro. En consecuencia, el número de flujos de datos de entrada siempre es diferente al número de flujos de datos de salida.

Por lo tanto, puede verse que la ruta de cualquier registro de datos a través del bloque 8 de procesamiento principal es complicada y cada registro de datos puede verse afectado por retrasos que ocurren en varios subprocesadores diferentes. La ruta exacta recorrida por cada registro de datos tampoco es completamente predecible ya que puede haber varios subprocesadores capaces de realizar una función de procesamiento particular para un formato específico de registro de datos (este diseño es para fines de redundancia), pero solo un subconjunto de estos será utilizado para cualquier registro de datos dado. Para evaluar con precisión la latencia de todo el bloque de procesamiento principal, localizar la causa de cualquier latencia incrementada en subcomponentes específicos y, por lo tanto, determinar las medidas apropiadas que deben tomarse, el módulo 5 de análisis de latencia y su almacén 4 de datos asociado deben mantener un registro de todas las rutas posibles que cada registro de datos puede tomar a través del bloque 8 de procesamiento principal. Esto incluye mantener un registro de todas las posibles combinaciones de subprocesadores 16a-e que pueden interactuar con un formato particular de registro de datos, y analizar toda la información resultante extraída de cada registro de datos para determinar qué ruta se recorrió y, por lo tanto, localizar la causa raíz del retraso en un registro de datos.

De forma alternativa o adicional, cada subprocesador 16a-e puede aplicar una marca de tiempo a cada registro que procesa; las marcas de tiempo para cada registro se extraen mediante el módulo 5 de análisis de latencia y se analizan para trazar el curso del registro de datos a través del bloque 8 de procesamiento principal, y la latencia asociada con cada subprocesador individual 16a-e.

Como se mencionó anteriormente, está claro que el volumen de datos que debe ser procesado por el módulo de análisis de latencia y almacenado por el almacén de datos es enorme, lo que significa que el análisis de latencia debe realizarse fuera de línea y el módulo 5 de análisis de latencia y el almacén 4 de datos deben ubicarse de manera remota de la caja de procesador. Además, debido al tamaño típico (en el orden de varios terabytes de información) de los conjuntos de datos resultantes que deben analizarse para obtener una imagen precisa de la latencia en cada punto y a lo largo de cada ruta recorrida en el bloque de procesamiento principal, el tiempo de análisis requerido es extremadamente largo. Por lo tanto, los análisis de latencia no podrían realizarse lo suficientemente rápido para efectuar cambios en tiempo real y abordar los problemas de latencia a medida que surgen. Como ejemplo ilustrativo, una unidad de procesamiento de datos conocida (caja de procesador) que comprende al menos 20 módulos de procesamiento individuales (o subprocesadores) recibe 5,2 mil millones de registros de datos en un solo día. Cada módulo de procesamiento escribe al menos 10 bytes de marca de tiempo y datos de identificación en cada registro de datos con el que interactúa. El volumen total de datos que deben procesarse todos los días durante el análisis de latencia es, por lo tanto, del orden de varios terabytes para cada unidad de procesamiento de datos (o caja de procesador) dentro de la red.

En otras palabras, los sistemas actuales conocidos para medir y abordar los problemas de latencia en las redes de procesamiento de datos son reaccionarios solamente y, por lo tanto, siempre intentarán solucionar los problemas después de que hayan surgido. Dichos sistemas solo podrán llevar a cabo efectivamente la "limitación de daños". Además, tales sistemas de medición de latencia son intrusivos: todos los registros de datos salientes son procesados por el módulo 10 de registro para extraer marcas de tiempo e información de ruta de procesamiento, y por lo tanto existe el riesgo de que se puedan introducir retrasos adicionales en la red mediante una acumulación de procesamiento en el módulo de registro. Esto también se debe en parte al hecho de que muchas de las unidades de procesamiento de datos (caja de procesador) usadas en las redes de procesamiento de datos son sistemas heredados; el procesamiento se realiza de manera efectiva dentro de las "cajas negras" de tal manera que no es posible modificar los métodos de procesamiento o de escribir información en los registros de datos para disminuir el tiempo de análisis requerido. Además, el procesamiento de los registros de datos en el módulo 10 de registro puede dar lugar a la introducción de un error en los registros de datos; este error se puede propagar posteriormente a los usuarios finales, los receptores de datos, si no se verifica. También será necesario llevar a cabo un riguroso proceso de evaluación de la calidad, lo que aumentará el tiempo de procesamiento.

65

El documento WO 2015/126301 A1 describe un método en una unidad de mezcla para mezclar los paquetes de datos recibidos de una pluralidad de flujos de datos, en el que el método incluye los pasos de recibir paquetes de datos decodificados y su información de tiempo, cuya información de tiempo comprende al menos un tiempo de espera máximo determinado para cada paquete de datos decodificado, optimizar un tiempo de espera para una mezcla posterior de paquetes de datos basándose en la información de tiempo, y mezclar al menos un subconjunto de los paquetes de datos recibidos decodificados basándose en el tiempo de espera optimizado.

El documento US 2009/287628 A1 describe un método y sistema para acelerar el hardware de varias operaciones de procesamiento de datos en un sistema de toma de decisiones basado en reglas tal como un motor de reglas de negocio, un procesador de flujo de evento, y un procesador de flujo de evento complejo. Preferiblemente, los flujos de datos entrantes son verificados por una pluralidad de condiciones de reglas. Entre las operaciones de procesamiento de datos que son aceleradas por hardware incluyen operaciones de verificación de condición de regla, operaciones de filtrado, y operaciones de combinación de ruta. Las operaciones de verificación de condición de regla generan resultados de verificación de condición de regla para los flujos de datos procesados, en los que los resultados de verificación de condición de regla son indicativos de cualquier condición de regla que los flujos de datos han satisfecho. La generación de tales resultados con un bajo grado de latencia proporciona iniciativas con la habilidad para realizar la toma de decisiones oportunas basándose en los datos presentes en los flujos de datos recibidos.

El documento US 2011/280137 A1 describe un aparato, un medio legible por máquina, y métodos para probar una red. El aparato puede incluir una máquina de estadísticas acoplada a una memoria. La máquina de estadísticas puede extraer un identificador de grupo de paquetes (PGID) de un paquete recibido de la red y recupera estadísticas de tráfico asociadas con el PGID de la memoria. Las estadísticas de tráfico recuperadas pueden incluir un recuento de paquetes por intervalo y un umbral de paquetes por intervalo, indicando el recuento de paquetes por intervalo un número de paquetes recuperado en un intervalo de tiempo de una secuencia de intervalos de tiempo. Al inicio de cada nuevo intervalo de tiempo, la máquina de estadísticas puede establecer un indicador que indica si el recuento de paquetes por intervalo para el intervalo de tiempo previo era igual o no o mayor que el umbral de paquetes por intervalo. El indicador puede ser guardado en la memoria como parte de las estadísticas de tráfico asociadas con el PGID extraído.

La presente invención busca abordar o al menos reducir sustancialmente los problemas descritos anteriormente con los sistemas de la técnica anterior.

Resumen de la invención

De acuerdo con un aspecto de la presente invención, se proporciona un sistema de control para controlar el funcionamiento de una unidad de procesamiento de datos, la unidad de procesamiento de datos recibiendo una primera pluralidad de flujos de entrada de registros de datos de diferentes proveedores de datos y comprendiendo un grupo de módulos de procesamiento interconectados a través de los cuales pasan los flujos de entrada para generar una segunda pluralidad de flujos de salida de registros de datos procesados, en el que las pluralidades primera y segunda son diferentes, el sistema de control comprendiendo: un módulo de control de entrada para introducir cada uno de los flujos de entrada recibidos en la unidad de procesamiento de datos, cada flujo de entrada teniendo un formato de datos definido de forma remota que está asociado con un subconjunto específico de los módulos de procesamiento de la unidad de procesamiento de datos, el módulo de control de entrada comprendiendo: un generador de registros ficticios para generar registros de datos ficticios para insertar en los flujos de entrada; el generador de registros ficticios estando dispuesto para generar registros de datos ficticios con un identificador conocido en un campo de datos inmutables del registro de datos y que tiene un formato seleccionado de uno de los formatos de datos definidos de forma remota; y un módulo de inyección de flujo de datos para inyectar la primera pluralidad de flujos de entrada en la unidad de procesamiento de datos, en el que el módulo de inyección de flujo de datos está dispuesto para intercalar un registro ficticio identificable de forma única en los registros de datos de la primera pluralidad de flujos de entrada que tienen el mismo formato definido de forma remota que el registro ficticio y para repetir el intercalado de otros registros ficticios en intervalos de tiempo predeterminados regulares según lo determinado por un reloj de referencia; un filtro de registro de datos para leer el campo de datos inmutables de cada registro de datos en cada flujo de salida y para identificar y separar registros ficticios de los flujos de salida que tienen el identificador conocido en el campo de datos inmutables; un módulo de análisis de latencia dispuesto para: recibir registros de datos ficticios del filtro de registro de datos; determinar un valor de latencia actual del registro de datos ficticios a través de la unidad de procesamiento de datos usando el conocimiento de cuándo el registro de datos ficticios se intercaló en el flujo de datos de entrada y la hora de recepción del registro de datos ficticios usando el reloj de referencia; comparar el valor de latencia actual con un valor de latencia determinado previamente de un registro de datos ficticios anterior para determinar si hay un cambio significativo en el valor de latencia; usar el formato del registro ficticio para determinar a qué subconjunto de módulos de procesamiento de la unidad de procesamiento de datos le preocupa la latencia; y determinar y emitir un comando de realimentación apropiado para que el módulo de control de entrada tome medidas para abordar el cambio significativo en la latencia detectada; en el que el módulo de control de entrada comprende un procesador de control dispuesto para recibir el comando de realimentación del módulo de análisis de latencia y para regular el caudal de flujos de entrada de datos de acuerdo con el comando de realimentación.

El módulo de control de entrada puede comprender además un generador de marca de tiempo acoplado al reloj de referencia, estando dispuesto el generador de marca de tiempo para generar una marca de tiempo para la inserción en el registro ficticio en un campo de datos inmutable adicional del registro de datos.

5 Cuando un formato original de la marca de tiempo es incompatible con el formato del campo de datos inmutables adicional, el generador de marcas de tiempo se puede organizar para convertir la marca de tiempo en un nuevo formato compatible con el campo de datos inmutables adicional y el módulo de análisis de latencia se puede
10 organizar para reconvertir la marca de tiempo de lectura del campo de datos inmutable adicional a su formato original.

15 El módulo de análisis de latencia puede estar dispuesto para comparar la latencia de diferentes registros ficticios relacionados con diferentes subconjuntos de módulos de procesamiento; para determinar dónde tienen una mayor latencia los múltiples subconjuntos de módulos de procesamiento y para determinar y generar un comando de realimentación dirigido a abordar la latencia de los múltiples subconjuntos de módulos de procesamiento. Además, cuando los subconjuntos se superponen, el módulo de análisis de latencia puede estar dispuesto para usar esto para determinar con mayor precisión la ubicación de cualquier aumento en la latencia.

20 El módulo de análisis de latencia puede estar dispuesto para almacenar los valores de latencia de cada registro ficticio en un almacén de datos de resultados agrupados de acuerdo con el formato en el que se recibió el registro de datos ficticios.

25 El módulo de análisis de latencia puede estar dispuesto para comparar el valor de latencia actual con un valor de latencia determinado previamente de un registro de datos ficticios anterior para determinar si hay un cambio significativo en el valor de latencia después de que haya transcurrido un período de tiempo predeterminado o después de que se haya analizado un número predeterminado de registros de datos ficticios.

30 Preferiblemente, el módulo de análisis de latencia está dispuesto para generar un comando de realimentación que hace que el procesador de control acelere un formato particular de registro de datos.

35 En algunas realizaciones, el generador de registros ficticios y el módulo de control de entrada están dispuestos para generar registros de datos ficticios e insertarlos en los flujos de entrada de acuerdo con un período de tiempo predeterminado que es diferente para diferentes flujos de entrada. En otras realizaciones, el generador de registros ficticios y el módulo de control de entrada están dispuestos para generar registros de datos ficticios e insertarlos en los flujos de entrada en diferentes desfases de tiempo entre sí.

Preferiblemente, cada registro de datos ficticios comprende un identificador de formato de registro de datos.

40 El sistema puede comprender además un almacén de datos de formato acoplado al módulo de control de entrada para almacenar un conjunto de formatos definidos de forma remota y un conjunto de reglas que determinan cómo el generador de registros ficticios genera los registros de datos ficticios a partir de los formatos de datos almacenados.

45 El sistema de control puede estar dispuesto para controlar una pluralidad de unidades de procesamiento de datos dispuestas en una red interconectada y el sistema comprende además una pluralidad de filtros de registro de datos que se proporcionan cada uno a la salida de una unidad de procesamiento de datos respectiva.

50 Preferiblemente, el módulo de análisis de latencia está dispuesto para enviar comandos de realimentación a las salidas de las unidades de procesamiento de datos para controlar la salida de los flujos de salida de los registros de datos procesados.

55 De acuerdo con otro aspecto de la presente invención, se proporciona un método para controlar el funcionamiento de una unidad de procesamiento de datos, la unidad de procesamiento de datos recibiendo una primera pluralidad de flujos de entrada de registros de datos de diferentes proveedores de datos y comprendiendo un grupo de módulos de procesamiento interconectados a través de los cuales los flujos de entrada pasan para generar una
60 segunda pluralidad de flujos de salida de registros de datos procesados, en el que las pluralidades primera y segunda son diferentes, el método comprendiendo: introducir, mediante un módulo de control de entrada, cada uno de los flujos de entrada recibidos en la unidad de procesamiento de datos, cada flujo de entrada teniendo un formato de datos definido de forma remota que está asociado con un subconjunto específico de los módulos de procesamiento de la unidad de procesamiento de datos, comprendiendo el paso de entrada: generar registros de
65 datos ficticios para insertar en los flujos de entrada; el paso de generación de registro de datos ficticios generando registros de datos ficticios con un identificador conocido en un campo de datos inmutables del registro de datos y que tiene un formato seleccionado de uno de los formatos de datos definidos de manera remota; e inyectar la primera pluralidad de flujos de entrada en la unidad de procesamiento de datos, en el que el paso de inyección comprende intercalar un registro ficticio identificable de forma única en los registros de datos de dicha primera pluralidad de flujos de entrada que tienen el mismo formato definido de forma remota que el registro ficticio y repetir el intercalado de otros registros ficticios a intervalos de tiempo predeterminados regulares según lo determinado por

un reloj de referencia; leer, mediante un filtro de registro de datos, el campo de datos inmutables de cada registro de datos en cada flujo de salida e identificar y separar registros ficticios de los flujos de salida que tienen el identificador conocido en el campo de datos inmutables; analizar los registros de datos ficticios, el paso de análisis comprendiendo: recibir registros de datos ficticios del filtro de registro de datos; determinar un valor de latencia actual del registro de datos ficticios a través de la unidad de procesamiento de datos usando el conocimiento de cuándo el registro de datos ficticios se intercaló en el flujo de datos de entrada y la hora de recepción del registro de datos ficticios usando el reloj de referencia; comparar el valor de latencia actual con un valor de latencia determinado previamente de un registro de datos ficticios anterior para determinar si hay un cambio significativo en el valor de latencia; usar el formato del registro ficticio para determinar a qué subconjunto de módulos de procesamiento de la unidad de procesamiento de datos le preocupa la latencia; y determinar y emitir un comando de realimentación apropiado para que el módulo de control de entrada tome medidas para abordar el cambio significativo en la latencia detectada; en el que el paso de entrada comprende recibir el comando de realimentación y regular el caudal de flujos de entrada de datos de acuerdo con el comando de realimentación.

15 **Breve descripción de los dibujos**

la figura 1 es un diagrama de bloques esquemático que muestra un sistema para la medición de latencia en una red que maneja y procesa flujos de datos y una porción del entorno en el que funciona el sistema;

20 la figura 2 es un diagrama de bloques esquemático que muestra detalles adicionales del componente de procesamiento principal del sistema de la figura 1;

la figura 3 es un diagrama esquemático que ilustra estructuras de ejemplo de diferentes formatos de registros de datos que pueden estar comprendidos dentro de los flujos de datos manejados por el sistema de la figura 1;

25 la figura 4 es un diagrama de bloques esquemático de un sistema de control de acuerdo con una realización de la presente invención y el entorno en el que funciona el sistema de control;

30 la figura 5 es un diagrama de bloques esquemático de una porción del sistema de control de la figura 4, que muestra detalles adicionales de un módulo de control de entrada de ejemplo que se puede usar para controlar y generar datos de entrada;

35 la figura 6 es un diagrama esquemático que ilustra estructuras de ejemplo de diferentes formatos de registros de datos ficticios generados por el módulo de control de entrada de la figura 6, de acuerdo con una realización de la presente invención;

40 la figura 7 es un diagrama de bloques esquemático de una porción del sistema de control de la figura 4, que muestra detalles adicionales de un módulo de análisis de latencia de ejemplo que se puede usar para calcular la latencia del sistema;

la figura 8 es un diagrama de flujo que muestra una descripción general del método general de funcionamiento del sistema de control de la figura 4;

45 la figura 9 es un diagrama de flujo que muestra detalles de un método para generar registros de datos ficticios utilizados por el módulo de control de entrada de la figura 6, de acuerdo con una realización de la presente invención;

50 la figura 10 es un diagrama de flujo que muestra detalles de un método para distinguir entre registros de datos genuinos y ficticios usados por el sistema de control de la figura 4;

la figura 11 es un diagrama de flujo que muestra detalles de un método para medir la latencia y determinar si se deben tomar medidas, usado por el sistema de control de la figura 4; y

55 la figura 12 es un diagrama de bloques esquemático que muestra una visión general del sistema de control de la figura 4 implementado en una red de procesamiento de datos a gran escala, de acuerdo con una realización de la presente invención.

Descripción detallada

60 Con referencia a la figura 4, ahora se describirá un sistema 30 de control para medir la latencia de acuerdo con una realización de la presente invención. El sistema de la presente realización tiene una configuración similar a la descrita en el sistema de la técnica anterior (que se muestra en la figura 1), aunque funciona de una manera muy diferente. El objetivo general del sistema es el mismo que el del sistema de la técnica anterior descrito anteriormente: medir la latencia de una red de procesamiento de datos en la que se implementa este sistema y tomar medidas de control en consecuencia si se detectan casos de latencia incrementada, y el paso y el procesamiento de los registros de datos a través de la red se mantiene prácticamente sin cambios. Sin embargo, el

mecanismo real mediante el cual se realizan las mediciones de latencia y la configuración e implementación detalladas del sistema de control de la figura 4 son diferentes. Para evitar repeticiones innecesarias, la siguiente descripción se centrará en estas diferencias.

- 5 Al igual que con el sistema de la técnica anterior descrito anteriormente, los flujos 2 de datos entrantes (que comprenden un flujo de registros de datos recibidos a lo largo del tiempo) se proporcionan a la caja 32 de procesador; los registros de datos son procesados por el bloque 34 de procesamiento principal (que es sustancialmente el mismo que el bloque 8 de procesamiento principal que se muestra en la figura 2) y luego se envían al siguiente componente de la red (por ejemplo, a otras caja de procesador o a una salida de red general).
- 10 Sin embargo, el sistema de la figura 4 comprende un módulo 36 de control de entrada que está configurado para generar registros 37 de datos ficticios que son capaces de pasar a través del bloque 34 de procesamiento principal de manera similar a los registros 2 de datos genuinos recibidos de los proveedores de datos sin que el bloque 34 de procesamiento principal sea capaz de distinguir entre los dos tipos de registros. El módulo 36 de control de entrada genera diferentes formatos de registros 37 de datos ficticios para que coincidan con los formatos de los registros 2 de datos entrantes (descritos en la figura 2); en lo sucesivo, los registros 2 de datos entrantes de los proveedores de datos se denominarán registros de datos genuinos para distinguirlos de los registros de datos ficticios generados por el sistema de control. Los registros 37 de datos ficticios generados se inyectan a intervalos predeterminados regulares en las rutas de caudal de sus flujos de datos correspondientes (por ejemplo, los registros de datos ficticios pueden inyectarse en los flujos de datos correspondientes una vez por segundo). Por lo tanto, los registros de datos ficticios se intercalan con los registros de datos genuinos del formato correspondiente a medida que se transmiten a la caja de procesador para producir un flujo 39 de datos de entrada mixto de cada formato. Esto garantiza que los dos tipos 36, 37 de registros de datos no serán distinguibles entre sí, y que las fluctuaciones de latencia experimentadas por los registros de datos genuinos también serán registradas por los registros de datos ficticios.
- 25 Los registros de datos ficticios entran al bloque 34 de procesamiento principal a través de las mismas entradas que los registros de datos genuinos que tienen un formato correspondiente al flujo de datos en el que están colocados, y siguen rutas similares a través del bloque 34 de procesamiento principal. Más detalles de la generación de estos registros ficticios se proporcionarán más adelante con referencia a las figuras 5 a 7. Para facilitar la referencia, las rutas de caudal de los registros 2 de datos genuinos se ilustran en esta figura (así como en todas las figuras posteriores) mediante flechas de línea continua mientras que las rutas de caudal de los registros 37 de datos ficticios se ilustran usando flechas de línea discontinua. En algunos casos, donde los registros de datos ficticios y genuinos siguen las mismas rutas de caudal (donde los registros de datos genuinos y ficticios se intercalan entre sí), las rutas de caudal efectivas del flujo 38 de datos mixto que comprenden los dos tipos de registros se ilustran usando flechas de línea de puntos y rayas.
- 35 La caja de procesador de la figura 4 también comprende un filtro 39 de registro de datos que es capaz de distinguir entre los registros 37 de datos ficticios y los registros 2 de datos genuinos recibidos del bloque 34 de procesamiento principal. El filtro 39 de registro de datos tiene dos salidas separadas: una salida 40 de registro de datos genuinos y una salida 42 de registro de datos ficticios. El filtro 38 de registro de datos está configurado para identificar los registros 37 de datos ficticios y extraerlos de los flujos 38 de datos entrantes; los registros 2 de datos genuinos se envían posteriormente a los componentes descendentes a través de la salida 40 de registro de datos genuinos y los registros 37 de datos ficticios se envían al módulo 44 de análisis de latencia a través de la salida 42 de registro ficticio. Posteriormente, los registros 37 de datos ficticios se analizan mediante el módulo 44 de análisis de latencia y se calculan sus valores de latencia asociados. Basándose en los resultados de estos análisis de latencia, los comandos 45 de realimentación se envían al módulo 36 de control de entrada para tomar las medidas apropiadas. Tales acciones pueden incluir cambiar automáticamente a un flujo de datos con una tasa de llegada de datos más baja, o disminuir automáticamente la velocidad de salida de uno o más flujos de datos de entrada desde el módulo 36 de control de entrada a la caja 32 de procesador.
- 50 Por lo tanto, se puede observar que el sistema de control de la figura 4 ha disminuido la cantidad de datos que necesita procesar el módulo 44 de análisis de latencia; en lugar de tener que procesar cada registro de datos que pasa por el sistema, solo los registros 37 de datos ficticios son procesados. Para los sistemas que reciben miles de registros 2 de datos por segundo de cada proveedor de datos, el procesamiento de registros de datos ficticios (que solo se pueden generar una vez por segundo) representa una disminución de mil veces en la potencia de procesamiento y el espacio de almacenamiento requerido por el módulo de análisis de latencia. Por lo tanto, el análisis de latencia se puede llevar a cabo mucho más rápidamente y, por lo tanto, el sistema de control puede detectar y responder a los aumentos de latencia mucho más rápidamente.
- 60 Los inventores también descubrieron que los registros de datos de un formato en particular solo pasarán típicamente a través de una cierta subsección del bloque 34 de procesamiento principal; por ejemplo, los registros de datos del Formato A solo pueden pasar a través de la Sección 1 del bloque 34 de procesamiento principal; los registros de datos del Formato B solo pueden pasar a través de la Sección 2 del bloque 34 de procesamiento principal y los registros de datos del Formato C solo pueden pasar a través de la Sección 3 del bloque 34 de procesamiento principal. Por lo tanto, cada formato de registro de datos ficticios podrá proporcionar información de latencia sobre una sección particular del sistema y, por lo tanto, la identificación del formato de registro de datos ficticios que
- 65

muestra una latencia incrementada permite al módulo 44 de análisis de latencia localizar fácilmente la sección del bloque 34 de procesamiento principal que está experimentando retrasos del procesamiento.

5 Se debe tener en cuenta que si bien la cobertura acumulada de todas las secciones abarca todos los subprocesadores dentro del bloque de procesamiento principal completo, las secciones destacadas en el bloque 34 de procesamiento principal de la figura 4 no están separadas físicamente y se excluyen mutuamente, pero pueden tener cierta superposición; algunos de los subprocesadores 16a-e (o módulos de procesamiento) pueden pertenecer a más de una sección, ya que pueden ser capaces de procesar más de un formato de registro de datos. Por lo tanto, es posible que la latencia asociada con algunos de los subprocesadores se mida usando dos o más formatos de registro de datos ficticios. Por lo tanto, una ventaja adicional que se logra con esta superposición de secciones en el bloque 34 de procesamiento principal (y sus componentes constituyentes) es que la fuente de cualquier problema de latencia se puede localizar con mayor precisión.

15 También se ha encontrado que los análisis de latencia por el sistema de control actual, que destacan secciones particulares de un bloque 34 de procesamiento principal con latencia incrementada, pueden lograr un nivel de precisión suficiente para permitir que las latencias de la red se gestionen de manera efectiva. La precisión adicional provista al analizar cada una de las posibles rutas de procesamiento de datos a través de una unidad de procesamiento de datos particular no hace una diferencia apreciable en la efectividad del manejo de la latencia, y el enorme ahorro de tiempo y espacio logrado por el sistema de la presente invención más que compensar cualquier diferencia en la eficacia.

25 El módulo 36 de control de entrada y el módulo 44 de análisis de latencia se muestran con mayor detalle en las figuras 5 y 7 respectivamente, y los ejemplos de los registros de datos ficticios que se crean en el módulo 36 de control de entrada y se analizan mediante el módulo 44 de análisis de latencia se muestran en la figura 6.

30 El módulo 36 de control de entrada de la figura 5 comprende un módulo 50 de entrada de flujo de datos para recibir flujos 2 de datos entrantes de los proveedores de datos, y un generador 52 de registros ficticios para generar los registros 37 de datos ficticios. El módulo 50 de entrada de flujo de datos y el generador 52 de registros ficticios están ambos por separado en comunicación operativa con un módulo 54 de salida de flujo de datos. De acuerdo con una realización de la presente invención, al llegar al módulo 54 de salida del flujo de datos, los formatos de registro de datos en cada flujo 2 de datos entrantes se identifican y los registros de datos se envían posteriormente a través de pasarelas de salida separadas (no mostradas) al bloque 34 de procesamiento principal. El módulo 36 de control de entrada también comprende un procesador 56 de control que está en comunicación operativa con los módulos 50, 54 de entrada y salida de flujo de datos; el procesador 56 de control está configurado para aceptar los comandos 45 de realimentación del módulo 44 de análisis de latencia y para tomar las medidas apropiadas para disminuir la latencia. En algunos casos, el procesador 56 de control puede configurarse para enviar un comando al módulo 50 de entrada de flujo de datos para que deje de aceptar ciertos formatos de datos entrantes, o para que reciba datos entrantes de un proveedor de datos diferente con una tasa de llegada de datos más baja; en otros casos, el procesador 56 de control está configurado para enviar un comando al módulo 54 de salida de flujo de datos para acelerar parte o la totalidad de la salida de datos al bloque 34 de procesamiento principal (es decir, para disminuir la frecuencia con la que los registros de datos de un formato específico se proporcionan al bloque 34 de procesamiento principal).

45 El generador 52 de registros ficticios está configurado para generar registros 37 de datos ficticios de diferentes formatos (correspondientes a los formatos de los registros de datos entrantes), y para proporcionar registros de datos ficticios al módulo de salida de flujo de datos a intervalos regulares, para que puedan ser inyectados en el flujo de datos con el formato correspondiente y enviados como flujos 38 de datos intercalados de los formatos relevantes a través de la pasarela de salida correspondiente al bloque 34 de procesamiento principal.

50 En algunas realizaciones, la frecuencia con la que se generan registros de datos ficticios de diferentes formatos difiere (por ejemplo, los registros del formato A pueden generarse a intervalos de un segundo mientras que los registros del formato B se generan a intervalos de dos segundos. Adicional o alternativamente, la frecuencia de generación de registros de datos ficticios puede ser la misma para todos los formatos de datos, pero los ciclos de tiempo de generación pueden estar fuera de sincronización o desfasados entre sí en una cierta cantidad (por ejemplo, los registros de formato A se generan cada minuto en el minuto, mientras que los registros de datos del formato B se generan en la marca de treinta segundos).

60 El generador 52 de registros ficticios está en comunicación operativa con un almacén 58 de datos y un generador 60 de marca de tiempo. El almacén 58 de datos contiene las plantillas almacenadas 61 a, b, c para cada formato de registro de datos ficticios que se va a generar y las reglas 62 relacionadas con la generación de los registros de datos ficticios. Estas plantillas 61a, b, c son precargadas en el generador 52 de registros ficticios, de modo que se puede generar rápida y fácilmente un registro de datos ficticios de un formato particular; sin embargo, en el caso de que se requiera un nuevo formato de registro de datos ficticios (por ejemplo, si se va a introducir un nuevo flujo de datos a la caja de procesador), el generador 52 de registros ficticios puede recuperar los formatos y reglas necesarios del almacén 58 de datos. El generador 60 de marca de tiempo también está en comunicación operativa con un reloj 63 de referencia y está configurado para generar una marca de tiempo (basada en la hora obtenida del

reloj 63 de referencia) cada vez que se genera un nuevo registro de datos ficticios, y para proporcionar esta marca de tiempo al generador 52 de registros ficticios para la inserción en el registro de datos ficticios recién generado.

Las estructuras de datos de ejemplo correspondientes a los diferentes formatos de registros de datos ficticios que pueden generarse se muestran en la figura 6. Cada registro de datos ficticios 64a, b, c está destinado a imitar los registros de datos genuinos con el formato correspondiente, y cada registro 64a, b, c de datos ficticios comprende la siguiente información:

(a) un identificador 66a, b, c de formato proporcionado en un campo de datos inmutables que es usado por el módulo 54 de salida de flujo de datos y el bloque 34 de procesamiento principal para identificar el formato del registro de datos ficticios y, de ese modo, garantizar que el registro de datos ficticios se transmita por las mismas rutas y se procese de la misma manera que los registros de datos genuinos con el mismo formato;

(b) una pluralidad de campos 68a, b, c de datos 'útiles', correspondientes a los campos de datos que se requieren para que el bloque 34 de procesamiento principal trate el registro de datos ficticios de la misma manera que un registro de datos genuino: los valores en estos subprocesadores 16a-e leen los campos (etiquetados como Parámetros 1 a 3, X a Z y A a C en la figura) en el bloque 34 de procesamiento principal y se usan durante el procesamiento;

(c) al menos un campo 70a, b, c de datos 'no interactivo' o más inmutable correspondiente a un campo de datos que debe estar presente para que el registro de datos imite con precisión la estructura del registro de datos genuino, pero para los cuales los valores nunca son leídos o procesados por el bloque 34 de procesamiento principal;

(d) uno o más campos 72a, b, c de datos 'no útiles' que están presentes en registros de datos genuinos de un formato correspondiente pero no son necesarios para los fines del análisis de latencia (es decir, no es necesario que interactúen ni sean leídos por el bloque 34 de procesamiento principal para que sea tratado como un registro de datos genuino).

El generador 52 de registros ficticios llena los campos 68a, b, c de datos 'útiles' con valores de parámetros representativos (por ejemplo, valores de parámetros dentro de un cierto rango que serán aceptados por el bloque 34 de procesamiento principal como dentro de un rango esperado de valores posibles), y o bien no llena los campos 72a, b, c de datos 'no útiles' con ningún valor de parámetro o los llena con valores nulos, ya que no estarán involucrados en el procesamiento realizado por el bloque de procesamiento principal. El generador 60 de marca de tiempo proporciona (empuja) una marca de tiempo al generador 52 de registros ficticios que inserta la marca 74a, b, c de tiempo en el campo 70a, b, c de datos "no interactivo" en lugar de aquellos parámetros con los cuales se rellenaría el campo de datos normalmente. Es ventajoso insertar la marca 74a, b, c de tiempo en un campo 70a, b, c de datos no interactivo, ya que se sabe que dicho campo es inmutable porque siempre estará presente en los registros de datos ficticios pero el bloque 34 de procesamiento principal no leerá ni procesará la información de la marca de tiempo y, por lo tanto, la información de la marca de tiempo pasará por el sistema sin distorsión.

El módulo 44 de análisis de latencia que se muestra en la figura 7 comprende una entrada 80 para recibir los registros 37 de datos ficticios entrantes del filtro 39 de registros de datos, y un analizador 82 de latencia configurado para calcular y evaluar la latencia de cada registro de datos ficticios entrantes. El analizador 82 de latencia está en comunicación operativa con un almacén 84 de datos, un módulo 86 de control y un reloj 88 de referencia. Para cada registro de datos ficticios entrantes, el analizador 82 de latencia está configurado para extraer la información 74a, b, c de marca de tiempo, comparar esa información con la hora obtenida del reloj 88 de referencia y calcular el valor de latencia asociado con ese registro de datos ficticios. El valor de latencia calculado se almacena entonces en el almacén 84 de datos, junto con los valores 85a, b, c de latencia obtenidos de registros de datos ficticios recibidos previamente del mismo formato.

El analizador 82 de latencia también está configurado para recuperar y analizar periódicamente los valores de latencia almacenados e identificar tendencias que indiquen aumentos de latencia potencialmente indeseables. En caso de que se detecten incrementos en la latencia, se envía un comando 45 de realimentación a través del módulo 86 de control al procesador 56 de control del módulo 36 de control de entrada para tomar las medidas apropiadas para disminuir la latencia sistémica.

Se apreciará que para llevar a cabo los cálculos de latencia con precisión, el reloj 88 de referencia del módulo 44 de análisis de latencia debe estar sincronizado con el reloj 63 de referencia utilizado por el módulo 36 de control de entrada. Preferiblemente, el módulo 36 de control de entrada (y su generador de registros ficticios constituyente) está en comunicación operativa con el reloj 88 de referencia del módulo 44 de análisis de latencia y, por lo tanto, los valores de tiempo pueden compararse directamente con el analizador 82 de latencia sin tener que compensar ningún desfase de tiempo entre diferentes relojes. Esto es especialmente importante cuando el módulo 36 de control de entrada y el módulo 44 de análisis de latencia están ubicados en diferentes porciones de la red.

La figura 8 proporciona una descripción general del método de análisis de latencia usado por el sistema de la figura 4. El proceso 100 comienza con la generación en el paso 105 de un registro de datos ficticios por el módulo 36 de

control de entrada; este registro de datos ficticios se proporciona luego en el paso 110 al bloque 34 de procesamiento principal mediante el módulo 36 de control de entrada; como se mencionó anteriormente, los registros 37 de datos ficticios se introducen al bloque 34 de procesamiento principal a través de las mismas entradas que los registros 2 de datos genuinos que tienen un formato correspondiente, y los dos tipos de registros de datos son virtualmente indistinguibles entre sí posteriormente por el bloque 34 de procesamiento principal.

Después de que los registros 38 de datos hayan pasado por el bloque 34 de procesamiento principal, los registros 37 de datos ficticios se identifican y se separan de los registros 2 de datos genuinos mediante el filtro 39 de registro de datos en el paso 115 y los primeros se transmiten al módulo 44 de análisis de latencia. Luego, se extrae la información de la marca de tiempo de cada registro de datos ficticios y se calcula un valor de latencia (es decir, se determina el tiempo necesario para que el registro ficticio pase a través del bloque 34 de procesamiento principal) en el paso 120. Los valores de latencia calculados para los registros de datos ficticios recibidos previamente se almacenan en el almacén 84 de datos, por ejemplo, en forma de una tabla de valores de latencia obtenidos a lo largo del tiempo, y estos registros 85a, b, c se actualizan en el paso 125 con los valores de latencia calculados de nuevo. Después de que haya transcurrido un cierto período de tiempo o de que se haya almacenado un cierto número de valores de latencia, se realiza un análisis automatizado en el paso 130 usando los valores de latencia almacenados para determinar si ha habido algún cambio en los valores de latencia (es decir, un aumento en los valores de latencia calculados sobre un valor de latencia umbral) que indicarían cualquier retraso de procesamiento que ocurra dentro del bloque de procesamiento principal y también permitirían ubicar la fuente de estos retrasos.

Los ejemplos de análisis que se pueden llevar a cabo en los valores de latencia almacenados incluyen el cálculo de agregados o de promedios móviles para tener en cuenta las actualizaciones constantes de los valores de latencia almacenados, así como para permitir que se destaque cualquier tendencia relativamente a largo plazo. Esto tendría la ventaja de permitir que el módulo 44 de análisis de latencia también detecte tendencias de latencia que aumentan lentamente, lo que podría indicar problemas de latencia potencialmente graves que surgirán en el futuro si no se resuelven. Por lo tanto, el sistema podrá tomar medidas preventivas para diagnosticar y abordar estos problemas antes de que se manifiesten como largos retrasos en el procesamiento y crear acumulaciones de procesamiento dentro de la red. En algunas realizaciones, el módulo de análisis de latencia también puede configurarse para generar automáticamente una representación gráfica de todos los valores de latencia almacenados durante ese período de tiempo, ya sea para que un operador de la red los revise, a fin de permitir cambios en la latencia durante un cierto período de tiempo (por ejemplo, en el transcurso de un día) para ser monitorizados, y para evaluar las tendencias de latencia a largo plazo, o para el análisis automatizado a más largo plazo, de manera que las tendencias problemáticas se puedan revertir al tomar algunas medidas correctivas automatizadas.

Una vez que se ha detectado un retraso en el procesamiento, el módulo 44 de análisis de latencia se configura en el paso 135 para determinar la causa raíz del problema, decidir el curso de acción más efectivo para disminuir la latencia y enviar comandos 45 de realimentación al módulo 36 de control de entrada para tomar medidas en consecuencia. El módulo 44 de análisis de latencia puede indicar al módulo 36 de control de entrada que cambie a flujos de datos de un proveedor de datos diferente que tenga una tasa de llegada de registros de datos más baja; alternativamente, si la latencia dentro del bloque 34 de procesamiento principal en su conjunto es muy grande (es decir, todos los registros de datos ficticios experimentan retrasos en el procesamiento), el módulo 44 de análisis de latencia puede indicar al módulo 36 de control de entrada que deje de enviar registros de datos adicionales al bloque 34 de procesamiento principal hasta que se haya eliminado la acumulación de registros de datos a procesar. Además, como el formato de los registros de datos corresponde a una subsección determinada del bloque 34 de procesamiento principal, si se encuentra que solo un formato de registro de datos ficticios tiene una latencia de media mayor, lo que indica que una sección particular del bloque 34 de procesamiento principal está sobrecargada o funciona mal, el módulo 44 de análisis de latencia puede indicar al módulo 36 de control de entrada que solo tome medidas con respecto a los flujos de datos que corresponden a ese formato en particular. Las medidas que puede indicar el módulo 44 de análisis de latencia pueden, por lo tanto, dirigirse a porciones específicas del bloque 34 de procesamiento principal y a flujos de datos particulares.

En una realización, se sabe que algunos formatos o tipos de registro de datos son más importantes que otros; puede ser más importante (incluso vital) proporcionar un cierto formato de registro de datos al receptor de datos que otros formatos de registros de datos. En estas realizaciones, el módulo 44 de análisis de latencia puede utilizar este hecho para seleccionar un formato particular de registro de datos para la aceleración que tiene una alta tasa de rendimiento de datos, pero se sabe que es de menor importancia. Además, debido a la superposición de secciones antes mencionada dentro del bloque 34 de procesamiento principal, la aceleración de un formato particular de registro de datos de esta manera puede mejorar la latencia en varias secciones.

Ahora se proporcionará una descripción más detallada del proceso usado por el módulo 36 de control de entrada para generar los registros de datos ficticios con referencia a la figura 9.

El proceso 200 comienza con la selección en el paso 205 del formato del registro de datos ficticios que se generará. En algunas realizaciones, el generador 52 de registros ficticios puede programarse con instrucciones para recorrer todos los formatos posibles de registro de datos ficticios a intervalos regulares. En otras palabras, el generador 52 de registros ficticios puede iniciar la generación de registros de datos ficticios en formato A, B y C, a su vez, a

intervalos predeterminados que están desfasados entre sí. Por ejemplo, los registros del formato A se pueden generar cada minuto en el minuto, mientras que los registros de datos del formato B se pueden generar en la marca de vigésimo segundo de cada minuto y los registros de datos del formato C se pueden generar en la marca de cuarenta segundos de cada minuto. Alternativamente, en algunas realizaciones, el generador 52 de registros ficticios
 5 puede generar los registros de datos ficticios de cada formato de forma independiente y en paralelo, pero cada formato de registro de datos ficticios puede tener una frecuencia de generación diferente.

Además, en algunas realizaciones, la frecuencia de generación de ciertos formatos de registros de datos ficticios puede variar a lo largo del curso del funcionamiento del sistema. Por ejemplo, si más registros de un formato particular pasan a través de una sección particular del bloque 34 de procesamiento principal, la frecuencia de generación de registros de datos ficticios con el formato correspondiente puede aumentarse para brindar una imagen más precisa de la latencia asociada con esa sección del sistema. Se debe tener en cuenta que, como la proporción del número de registros 37 de datos ficticios con respecto a los registros 2 de datos genuinos que pasan por el sistema en un momento dado es tan pequeña, la sobrecarga de procesamiento del sistema asociada con la generación de un mayor número de registros de datos ficticios y el aumento de la frecuencia de su entrada es mínima, y es prácticamente indetectable a nivel global del sistema.
 10
 15

Una vez que se ha decidido el formato de un registro de datos ficticios que se generará, en el paso 210, el generador 52 de registros ficticios sigue instrucciones programadas previamente para generar el registro de datos ficticios, correspondientes al formato seleccionado, con los campos de datos requeridos y el identificador de formato. En algunas realizaciones, el generador 52 de registros ficticios genera el registro de datos ficticios en tiempo real cada vez; en otras realizaciones, el generador 52 de registros ficticios recupera en el paso 215 una plantilla 61a, b, c correspondiente al formato de registro de datos ficticios apropiado del almacén de datos. Adicional o alternativamente, estas plantillas se cargan previamente en el generador 52 de registros ficticios y el generador 52 de registros ficticios simplemente selecciona en el paso 215 la plantilla con el identificador de formato apropiado cada vez que se requiere generar un nuevo registro de datos ficticios.
 20
 25

El generador 52 de registros ficticios rellena a continuación en el paso 220 los campos del registro de datos ficticios con los valores de parámetros apropiados: los campos 68a, b, c de datos 'útiles' se rellenan con valores de parámetros representativos que los subprocesadores del bloque 34 de procesamiento principal no podrán distinguir de los valores de los parámetros encontrados en registros de datos genuinos. Estos valores de parámetros pueden seleccionarse al azar de un rango de valores posibles esperados y pueden variar cada vez que se produce un registro de datos ficticios. Alternativamente, si las plantillas de registro de datos ficticios se cargan previamente en el generador 52 de registros ficticios, como suele ser el caso, es posible que estos campos de datos ya hayan sido rellenos previamente con valores de parámetros representativos que permanecen iguales para cada registro de datos ficticios de ese formato que se genera. Los campos 74a, b, c de datos 'no útiles' se dejan en blanco o se pueden rellenar aleatoriamente (por ejemplo, con valores nulos o extremadamente grandes, por conveniencia) ya que estos campos de datos no interactuarán con el bloque 34 de procesamiento principal mientras el registro de datos ficticios pasa a través de él y, por lo tanto, su valor exacto no es importante.
 30
 35
 40

Una vez que se ha generado el registro de datos ficticios y se han relleno sus campos de datos, el generador 52 de registros ficticios obtiene en el paso 225 una marca 74a, b, c de tiempo del generador 60 de marca de tiempo y lo inserta en un campo de datos particular que se ha indicado previamente como un campo 70a, b, c de datos 'no interactivo' o más inmutable (es decir, los valores contenidos en este campo de datos no se extraerán ni utilizarán para el análisis, ni se escribirán valores adicionales en este campo de datos, a lo largo del curso de la ruta del registro a través del bloque de procesamiento principal). Esto garantiza que la información 74a, b, c de marca de tiempo que debe estar contenida dentro de este campo de datos no se distorsione por ninguna interacción dentro del bloque de procesamiento principal, y el módulo 44 de análisis de latencia puede recuperarla de manera confiable del registro de datos ficticios una vez que haya pasado por el bloque 34 de procesamiento principal.
 45
 50

En algunas realizaciones, los parámetros que originalmente estaban destinados a ser usados para rellenar el campo 70a, b, c de datos 'no interactivo' elegido se expresan típicamente en un formato que es incompatible con un formato de tiempo (por ejemplo, un campo de datos elegido solo puede aceptar números de punto flotante, y los valores de marca de tiempo expresados en minutos y segundos serían incompatibles con este formato). Como los registros de datos genuinos en los que se basan los registros de datos ficticios son producidos por los proveedores de datos y, por lo tanto, su formato está fuera del control de la red de procesamiento de datos y del sistema 30 de control, no es posible para el generador 52 de registros ficticios para modificar fácil y rápidamente el formato del campo de datos en el que se insertará la marca de tiempo. En tales casos, la información de la marca de tiempo debe ser convertida o reexpresada/encryptada para generar un número de punto flotante representativo para la inserción en el campo de datos elegido.
 55
 60

En algunas realizaciones, puede darse el caso de que un formato particular de registro de datos no contenga ningún campo 70a, b, c de datos no interactivo. En tales casos, el módulo 36 de control de entrada estará al tanto de las transformadas de coordenadas a las que estarán sujetos los parámetros en un campo elegido, y el generador 52 de registros ficticios puede, por lo tanto, llevar a cabo un paso de premanipulación (ingeniería inversa) aplicando la transformación de coordenadas inversa a la marca de tiempo antes de insertarla en el campo de datos elegido.
 65

Cuando el analizador 82 de latencia recupera posteriormente las marcas 74a, b, c de tiempo después de haber experimentado la transformación de coordenadas pertinente, la marca 74a, b, c de tiempo estará en el formato deseado y no requerirá ningún procesamiento posterior por parte de analizador 82 de latencia.

5 Cuando se completa el proceso 210 de generación del registro de datos ficticios, el generador 52 de registros ficticios reenvía el registro de datos ficticios al módulo 54 de salida del flujo de datos, donde se lee el identificador de formato del registro de datos ficticios y el registro de datos ficticios se inserta en el paso 230 en el flujo 38 de datos del formato correspondiente que se está enviando al bloque 34 de procesamiento principal. El registro de datos ficticios abandona el módulo 36 de control de entrada y entra en el bloque 34 de procesamiento principal, donde
10 pasa a través de la sección correspondiente del bloque 34 de procesamiento principal de acuerdo con su formato, interactuando con los respectivos subprocesadores 16a-e (o módulos de procesamiento) a lo largo del camino. Después de proporcionar el registro de datos ficticios al módulo 54 de salida del flujo de datos, el generador 52 de registros ficticios espera en el paso 235 un período de tiempo predeterminado antes de generar un nuevo registro de datos del mismo formato.

15 Este proceso 200 se repite para cada formato de registro de datos ficticios generado por el módulo de control de entrada.

20 Todos los registros 38 de datos (tanto ficticios como genuinos) que hayan pasado por el bloque 34 de procesamiento principal deben pasar por el filtro 39 de registro de datos antes de que puedan abandonar la caja 32 de procesador; esto es para permitir que el sistema 30 de control separe los registros 37 de datos ficticios del registro 2s de datos genuinos, para garantizar que el análisis de latencia solo se realice usando el tipo anterior de registro.

25 El proceso por el cual el filtro 39 de registro de datos separa los registros de datos ficticios de los registros 2 de datos genuinos se describirá ahora con referencia a la figura 10. El proceso 300 comienza con el filtro 39 de registro de datos recibiendo en el paso 305 un registro de datos entrantes de los flujos 38 de datos mixtos e identificando en el paso 310 un campo de datos particular que se usa para distinguir entre los registros de datos genuinos y ficticios. La información en este campo de datos identificado se verificará en el paso 315 para determinar si contiene un indicador que indica que el registro en cuestión es un registro de datos ficticios; por ejemplo, el campo de datos
30 puede contener un valor o parámetro específico que indica que el registro es un registro de datos ficticios. En algunas realizaciones, el generador 52 de registros ficticios puede haber establecido que el valor de uno de los campos de datos no útiles sea un valor específico (que permanece constante para cada registro de datos ficticios que se genera y puede ser constante en todos los formatos de registro de datos ficticios, aunque introducido en diferentes campos de datos 'no útiles' para cada formato de registro de datos). Este valor debería ser fácilmente reconocible por el filtro 39 de registro de datos, e idealmente no sería un valor que pudiera estar contenido en ese campo de datos para un registro 2 de datos genuino. Alternativamente, como es más frecuente el caso, los registros de datos ficticios (como en todos los registros de datos) tienen cada uno un identificador único asociado a ellos (es decir, un ID de registro). Los identificadores de registro de cada registro de datos ficticios que se generan son conocidos por el filtro 39 de registro de datos, que por lo tanto puede identificar el campo de datos relevante de
35 'identificador de registro', extraer el identificador de registro y compararlo con los identificadores de registros de datos ficticios conocidos para identificar los registros de datos ficticios.

45 Si el filtro 39 de registro de datos determina en el paso 315 que el campo de datos seleccionado contiene un parámetro correspondiente al valor especificado predefinido (por ejemplo, el ID del registro de datos coincide con un valor que se sabe corresponde con el de un registro de datos ficticios), el filtro 39 de registro de datos identifica este registro como un registro de datos ficticios y lo extrae del flujo 38 de datos entrantes en el paso 320, remitiéndolo al módulo 44 de análisis de latencia. Si el filtro 39 de registro de datos determina en el paso 315 que el campo de datos seleccionado no contiene el valor de parámetro de identificación esperado (es decir, el ID de registro no coincide con ninguno de los ID de registro predefinidos que se sabe que están asociados con registros de datos ficticios), el filtro
50 39 de registro de datos identifica ese registro de datos como genuino y envía el registro de datos en el paso 325 a los componentes aguas abajo posteriores. Estas pueden ser otras unidades de procesamiento de datos o caja de procesador en la red, o los usuarios de extremo finales de los flujos de datos (es decir, los receptores de datos).

55 El proceso de análisis de latencia llevado a cabo por el módulo 44 de análisis de latencia se describirá ahora con mayor detalle con referencia a la figura 11. Este proceso 400 comienza con el recibo en el paso 405 del registro de datos ficticios del filtro 39 de registro de datos. El analizador 82 de latencia está programado para extraer en el paso 410 la información 74a, b, c de marca de tiempo del campo requerido del registro de datos ficticios: el analizador 82 de latencia contiene instrucciones que permiten al analizador 82 de latencia identificar el campo 70a, b, c de datos relevantes no interactivos en el registro de datos ficticios de un formato particular y, posteriormente, para extraer los datos 74a, b, c de marca de tiempo. En aquellas realizaciones donde la marca de tiempo se extrae en formato
60 cifrado (la marca de tiempo se modificó para que coincida con el formato de los datos requeridos para ese campo de datos), el analizador 82 de latencia también contiene instrucciones para descifrar la información de la marca de tiempo en un formato de análisis fácil de leer y comparable.

65 El analizador 82 de latencia obtiene entonces una marca 74a, b, c de tiempo del reloj 88 de referencia y compara en el paso 415 este valor con la marca 74a, b, c de tiempo que se extrajo del registro de datos ficticios para calcular en

el paso 420 el tiempo tomado por el registro de datos ficticios para recorrer la caja 32 de procesador (es decir, la latencia asociada con la ruta recorrida por ese registro de datos ficticios). Este valor de latencia se almacena luego en el paso 425 en el almacén 84 de datos asociado, junto con la información de identificación relevante (por ejemplo, el formato del registro de datos ficticios a partir del cual se obtuvo el valor de latencia y la hora del día en que se obtuvo ese valor de latencia).

Periódicamente, el analizador 82 de latencia recupera un subconjunto de los valores 85a, b, c de latencia almacenados del almacén 84 de datos y realiza un análisis en el paso 430 de los datos almacenados para trazar las tendencias de latencia a lo largo del tiempo y determinar si hay alguna indicación de latencias potencialmente crecientes. Por ejemplo, en una realización, el analizador 82 de latencia puede realizar un análisis de media móvil en los valores de latencia almacenados para identificar aumentos de latencia, ya que dichos análisis tienden a descontar picos o fluctuaciones de puntos de datos únicos anómalos, pero permitirán que el módulo 44 de análisis de latencia identifique cualquier caso de mayor latencia que se manifiesta en dos o tres registros de datos ficticios sucesivos. En otra realización, el analizador 82 de latencia puede comparar las frecuencias de entrada de los registros de datos ficticios (los valores para estas frecuencias de entrada se obtienen de las marcas de tiempo extraídas, o pueden programarse previamente en el analizador 82 de latencia) con las frecuencias de salida de registros de datos ficticios (obtenidos de las marcas de tiempo generadas en el módulo 44 de análisis de latencia por el reloj 88 de referencia). Estos análisis podrían permitir que el sistema de control tome medidas preventivas en el paso 435 para abordar los problemas relacionados con la latencia antes de que comiencen a afectar seriamente el rendimiento y la tasa de procesamiento de datos de la caja 32 de procesador.

Tras la identificación de tales problemas de latencia que surgen, el analizador 82 de latencia envía información con respecto a estos problemas de latencia al módulo 86 de control. El módulo 86 de control está programado con reglas e instrucciones relacionadas con las medidas que deben tomarse basándose en los resultados del análisis de latencia, y genera comandos de acuerdo con estas reglas que se envían al módulo 36 de control de entrada para tomar medidas para disminuir la latencia del sistema.

Se apreciará que el módulo 44 de análisis de latencia realiza un análisis de latencia separado para cada formato de registro de datos ficticios que se identifica. Sin embargo, los resultados del análisis para los diferentes formatos de registro de datos pueden evaluarse en combinación para permitir la determinación de las medidas más apropiadas que se realizarán para el sistema en su totalidad.

La figura 12 muestra una vista general de la implementación del sistema de control en una red 500 que comprende múltiples cajas 32a, b, c, d de procesador de acuerdo con una realización de la presente invención. En esta figura, las líneas discontinuas indican las rutas seguidas por los registros 37 de datos ficticios; las líneas continuas indican las rutas seguidas por los registros 2 de datos genuinos; y las líneas de puntos y rayas indican aquellas rutas seguidas tanto por registros de datos ficticios como genuinos (es decir, por el flujo 38 de datos intercalados/mezclados).

En la realización de la figura 12, se supone que la caja 1 de procesador es la primera caja en la red 500 para recibir los flujos 2 de datos entrantes de los proveedores de datos, y la caja 3 de procesador es la caja final en la red 500 para procesar los flujos 3 de datos antes de que se envíen a los receptores de datos. Sin embargo, debe entenderse que esto es simplemente un ejemplo ilustrativo y que se pueden utilizar muchas otras cajas de procesador dentro de la red, entre las cajas de procesador 1 y 3, además de (o en lugar de) la caja 2 de procesador y el grupo 32d de caja de procesador.

Los flujos 2 de datos entrantes primero entran al módulo 36 de control de entrada donde se generan registros 37 de datos ficticios correspondientes al formato de los registros de datos entrantes (como se describe con referencia a la figura 9); estos registros de datos ficticios se insertan luego en el flujo 38 de datos correspondiente cuando se envían a la caja 1 de procesador. Los registros 38 de datos son procesados por el bloque de procesamiento principal dentro de la caja 1 de procesador, y el filtro 39a de registro de datos de la caja 1 de procesador separa los registros 37 de datos ficticios de los registros 2 de datos genuinos, como se describió con referencia a la figura 10.

En algunas realizaciones, algunos o todos los registros 37 de datos ficticios se transmiten desde el filtro 39a de registro de datos al módulo 44 de análisis de latencia para su análisis. Este es un paso opcional, y en otras realizaciones, todos los registros 37 de datos ficticios pueden enviarse junto con los registros de datos genuinos a la siguiente caja de procesador en la red. La extracción de un subconjunto de registros de datos ficticios para su análisis después del procesamiento por cada caja de procesador puede permitir obtener una imagen más clara de las latencias asociadas con esa caja de procesador en particular.

Los registros de datos de salida se pasan posteriormente a uno de la caja 2 de procesador o al grupo 32d de caja de procesador (que a su vez comprende las cajas de procesador A a D). La caja 2 de procesador y el grupo 32d de la caja de procesador realizan el mismo análisis en los registros 38 de datos entrantes y están presentes en la red para redundancia y resiliencia, de modo que si uno de ellos funciona mal o si el rendimiento de datos de la caja 1 de procesador es demasiado alto para una caja de procesador para hacer frente por sí sola, algunos o todos los datos pueden transmitirse a las otras caja de procesador. Alternativamente, en algunas realizaciones, las caja de

procesador A a D pueden procesar cada una uno o más formatos diferentes de registros de datos; por ejemplo, la caja de procesador A solo puede procesar los registros de datos del formato A, mientras que la caja de procesador B puede procesar los registros de datos del formato A y del formato B.

- 5 Los registros de datos se procesan en los bloques de procesamiento principales respectivos de la caja 2 de procesador y las caja de procesador A a D, y los filtros de registro de datos respectivos (no mostrados para las caja de procesador A a D) pueden filtrar algunos o todos los registros de datos ficticios para su transmisión al módulo de análisis de latencia. Tras la salida de la caja 2 de procesador o del grupo 32d de la caja de procesador, los registros de datos entran a la caja 3 de procesador donde se procesan y filtran nuevamente. En el ejemplo ilustrado, como la
- 10 caja 3 de procesador es la última caja de procesador en la red 500 de procesamiento de datos, todos los registros 37 de datos ficticios se extraen de los flujos 38 de datos entrantes y se transmiten al módulo 44 de análisis de latencia; solo los registros de datos genuinos se propagan hacia adelante fuera de la red como flujos 3 de datos de salida.
- 15 Se apreciará que aunque las cajas de procesador 1, 2 y 3 se muestran cada una en la figura como cajas de un solo procesador, esto es por simplicidad y para mejorar la claridad de la figura solamente. Es posible que cada una de estas cajas de procesador se corresponda con un grupo de cajas de procesador (como las del grupo 2 de cajas de procesador). Además, también se debe tener en cuenta que las flechas simples se usan en toda la figura para indicar las rutas que pueden tomar los registros de datos de todos los formatos diferentes (para indicar las rutas
- 20 posibles seguidas por formatos de registros de datos individuales, el diagrama sería innecesariamente complicado); sin embargo, no todos los registros de datos seguirán los mismos caminos cada vez, las flechas simplemente indican todas las rutas posibles que podrían ser recorridas.

25 El módulo 44 de análisis de latencia lleva a cabo un análisis de latencia de los registros 37 de datos ficticios recibidos sustancialmente como se describe con referencia a la figura 11, y determina si se detectan casos de latencia aumentada que justifiquen el envío de un comando 45 de realimentación al módulo 36 de control de entrada para tomar medidas. Ahora se detallarán ejemplos de medidas que se pueden tomar en diferentes escenarios como resultado del análisis realizado por el módulo 44 de análisis de latencia.

30 En un ejemplo, el módulo 44 de análisis de latencia puede detectar que ha surgido un caso de latencia aumentada en la caja de procesador A y que la caja de procesador está sobrepasada o sobrecargada por la cantidad de datos entrantes. En respuesta a esto, se puede enviar un comando 45 de realimentación desde el módulo 44 de análisis de latencia al módulo 36 de control de entrada (o a la caja 1 de procesador, o incluso a la caja de procesador A) para disminuir la cantidad de datos que se envían a la caja de procesador A y para redistribuir estos registros de

35 datos entre las caja de procesador B a D, según corresponda. En algunas realizaciones, donde los problemas de latencia que surgen en la caja de procesador A son el resultado de un mal funcionamiento del sistema, esto puede implicar apagar completamente la entrada de datos a la caja de procesador A hasta que se pueda solucionar el problema.

40 En otro ejemplo, si se detectan casos de latencia aumentada en más de una caja de procesador (por ejemplo, si los registros de datos que pasan a través de las cajas de procesador A y B experimentan una latencia incrementada), se puede considerar que simplemente redistribuyendo los flujos de datos entrantes para las otras caja de procesador en el grupo 32d de la caja de procesador será insuficiente para abordar la mayor latencia observada y, de hecho, puede sobrecargar las dos caja de procesador restantes. En este caso, se puede enviar un comando de

45 realimentación al módulo 36 de control de entrada o a la caja 1 de procesador para cambiar la salida de todos los flujos de datos salientes desde el grupo 32c de la caja de procesador a la caja 2 de procesador (o viceversa, si se observa una mayor latencia en la salida de registros de datos de la caja 2 de procesador en su lugar).

50 En algunos casos, si se considera que se observan casos de mayor latencia en el primer componente de la red (es decir, caja 1 de procesador) se puede enviar un comando 45 de realimentación al módulo 36 de control de entrada para cambiar a la entrada de un flujo de datos a la caja 1 de procesador que tiene una tasa de llegada de datos más baja. Alternativa o adicionalmente, el módulo 36 de control de entrada puede simplemente disminuir la tasa de rendimiento de datos (aceleración) a la caja 1 de procesador. En algunos casos, el módulo 44 de análisis de latencia puede determinar que solo un subconjunto de registros de datos (correspondientes a un formato particular de

55 registros de datos ficticios recibidos) que pasan por la caja 1 de procesador se están retrasando. Esto indicaría un problema de mayor latencia en solo una sección de la caja 1 de procesador y, por lo tanto, las medidas anteriores solo deben tomarse con respecto al flujo de datos que comprende los registros de datos del formato requerido para pasar a través de la sección afectada.

60 Una ventaja del sistema de control de la presente invención es que permite que la interacción compleja y las interacciones entre las unidades del subprocesador (cada una con diferentes efectos en las latencias de los registros de datos que pasan a través del sistema) se traduzcan en un conjunto de condiciones empíricas. Por lo tanto, es posible generar fácilmente reglas lógicas simples y de alto nivel (por ejemplo, en el caso de que se considere que los registros del formato A tienen latencias incrementadas, acelerar los flujos de datos del formato A) se implementan

65 posteriormente en tiempo real para abordar automáticamente los problemas de latencia a medida que surgen.

Por ejemplo, se prevén otras realizaciones de la presente invención en las que se puede asociar un módulo de análisis de latencia separado con cada caja de procesador individual o unidad de procesamiento de datos en la red de procesamiento de datos a gran escala.

5

REIVINDICACIONES

- 1.- Un sistema (30) de control para controlar el funcionamiento de una unidad (32) de procesamiento de datos, la unidad (32) de procesamiento de datos recibiendo una primera pluralidad de flujos (2) de entrada de registros de datos de diferentes proveedores de datos y comprendiendo un grupo de módulos (16a-e) de procesamiento interconectado a través de los cuales pasan los flujos de entrada para generar una segunda pluralidad de flujos (3) de salida de registros de datos procesados, en el que las pluralidades primera y segunda son diferentes, el sistema de control comprendiendo:
- 5
- 10 a) un módulo (36) de control de entrada para introducir cada uno de los flujos (2) de entrada recibidos en la unidad (32) de procesamiento de datos, teniendo cada flujo de entrada un formato de datos definido de forma remota que está asociado con un subconjunto específico de los módulos (16a-e) de procesamiento de la unidad (32) de procesamiento de datos, comprendiendo el módulo de control de entrada:
- 15 i) un generador (52) de registros ficticios para generar registros (37) de datos ficticios para insertar en los flujos (2) de entrada; estando dispuesto el generador (52) de registros ficticios para generar registros de datos ficticios con un identificador conocido en un campo de datos inmutables del registro de datos y que tiene un formato seleccionado de uno de los formatos de datos definidos de manera remota; y
- 20 ii) un módulo (54) de inyección de flujo de datos para inyectar la primera pluralidad de flujos (2) de entrada en la unidad de procesamiento de datos, en el que el módulo (54) de inyección de flujo de datos está dispuesto para intercalar un registro ficticio identificable de forma única en los registros de datos de la primera pluralidad de flujos (2) de entrada que tienen el mismo formato definido de manera remota que el registro ficticio y para repetir el intercalado de otros registros ficticios en intervalos de tiempo predeterminados regulares según lo determinado por un reloj (63, 88) de referencia;
- 25
- b) un filtro (39) de registro de datos para leer el campo de datos inmutables de cada registro de datos en cada flujo de salida y para identificar y separar registros ficticios de los flujos de salida que tienen el identificador conocido en el campo de datos inmutables;
- 30
- c) un módulo (44) de análisis de latencia dispuesto para:
- a. recibir registros de datos ficticios del filtro (39) de registro de datos;
- 35 b. determinar un valor de latencia actual del registro de datos ficticios a través de la unidad (32) de procesamiento de datos usando el conocimiento de cuándo el registro de datos ficticios se intercaló en el flujo (38) de datos de entrada y la hora de recepción del registro de datos ficticios usando el reloj de referencia;
- c. comparar el valor de latencia actual con un valor de latencia determinado previamente de un registro de datos ficticios anterior para determinar si hay un cambio significativo en el valor de latencia;
- 40
- d. usar el formato del registro ficticio para determinar a qué subconjunto de módulos (16a-e) de procesamiento de la unidad (32) de procesamiento de datos se refiere la latencia; y
- 45 e. determinar y enviar un comando (45) de realimentación apropiado para que el módulo (36) de control de entrada tome medidas para abordar el cambio significativo en la latencia detectada;
- en el que el módulo (36) de control de entrada incluye un procesador (56) de control dispuesto para recibir el comando (45) realimentación desde el módulo (44) de análisis de latencia y para regular el caudal de flujos de entrada de datos de acuerdo con el comando de realimentación.
- 50
- 2.- El sistema de control de la reivindicación 1, en el que el módulo (36) de control de entrada comprende además un generador (60) de marca de tiempo acoplado al reloj (63) de referencia, estando dispuesto el generador (60) de marca de tiempo para generar una marca (74a, b, c) de tiempo para la inserción en el registro ficticio en un campo (70a, b, c) de datos más inmutable del registro de datos.
- 55
- 3.- El sistema de control de la reivindicación 2, en el que un formato original de la marca (74a, b, c) de tiempo es incompatible con el formato del campo (70a, b, c) de datos inmutable adicional y el generador (60) de marca de tiempo está dispuesto para convertir la marca de tiempo en un nuevo formato compatible con el campo de datos inmutable adicional y el módulo (44) de análisis de latencia está dispuesto para reconvertir la marca de tiempo de lectura del campo de datos inmutable adicional nuevamente a su formato original.
- 60
- 4.- El sistema de control de cualquier reivindicación precedente, en el que el módulo (44) de análisis de latencia está dispuesto para comparar la latencia de diferentes registros ficticios relacionados con diferentes subconjuntos de módulos (16a-e) de procesamiento; para determinar dónde múltiples subconjuntos de módulos de procesamiento
- 65

tienen una mayor latencia y para determinar y generar un comando (45) de realimentación dirigido a abordar la latencia de los múltiples subconjuntos de módulos de procesamiento.

- 5 5.- El sistema de control de la reivindicación 4, en el que los subconjuntos se superponen y, por lo tanto, el módulo (44) de análisis de latencia está dispuesto para usar esto para determinar con mayor precisión la ubicación de cualquier aumento de latencia.
- 10 6.- El sistema de control de cualquier reivindicación precedente, en el que el módulo (44) de análisis de latencia está dispuesto para almacenar los valores de latencia de cada registro ficticio en un almacén (84) de datos de resultados agrupados de acuerdo con el formato en el que se recibió el registro de datos ficticios.
- 15 7.- El sistema de control de cualquier reivindicación precedente, en el que el módulo (44) de análisis de latencia está dispuesto para comparar el valor de latencia actual con un valor de latencia determinado previamente de un registro de datos ficticios anterior para determinar si hay un cambio significativo en el valor de latencia después de que un período de tiempo predeterminado ha transcurrido o después de que se haya analizado un número predeterminado de registros de datos ficticios.
- 20 8.- El sistema de control de cualquier reivindicación precedente, en el que el módulo (44) de análisis de latencia está dispuesto para generar un comando (45) de realimentación que hace que el procesador (56) de control acelere un formato particular de registro de datos.
- 25 9.- El sistema de control de cualquier reivindicación precedente, en el que el generador (52) de registros ficticios y el módulo (36) de control de entrada están dispuestos para generar registros de datos ficticios e insertarlos en los flujos (2) de entrada de acuerdo con un período de tiempo predeterminado que es diferente para diferentes flujos de entrada.
- 30 10.- El sistema de control de cualquier reivindicación precedente, en el que el generador (52) de registros ficticios y el módulo (36) de control de entrada están dispuestos para generar registros de datos ficticios e insertarlos en los flujos de entrada en diferentes desfases de tiempo entre sí.
- 35 11.- El sistema de control de cualquier reivindicación precedente, en el que cada registro de datos ficticios comprende un identificador (66a, b, c) de formato de registro de datos.
- 40 12.- El sistema de control de cualquier reivindicación precedente, que comprende además un almacén (58) de datos de formato acoplado al módulo (36) de control de entrada para almacenar un conjunto de formatos definidos de forma remota (61 a, b, c) y un conjunto de reglas (62) que determinan cómo se generan los registros de datos ficticios a partir de los formatos de datos almacenados por el generador (52) de registros ficticios.
- 45 13.- El sistema de control de cualquier reivindicación precedente, en el que el sistema de control está dispuesto para controlar una pluralidad de unidades (32a, b, c) de procesamiento de datos dispuestas en una red interconectada y el sistema (30) comprende además una pluralidad de filtros (39a, b, c) de registro de datos cada uno siendo proporcionado a la salida de una unidad de procesamiento de datos respectiva.
- 50 14.- El sistema de control de la reivindicación 13, en el que el módulo (44) de análisis de latencia está dispuesto para enviar comandos (45) de realimentación a las salidas de las unidades (32a, b, c) de procesamiento de datos para controlar la salida de los flujos de salida de registros de datos procesados.
- 55 15.- Un método para controlar el funcionamiento de una unidad (32) de procesamiento de datos, la unidad (32) de procesamiento de datos recibiendo una primera pluralidad de flujos (2) de entrada de registros de datos de diferentes proveedores de datos y comprendiendo un grupo de módulos (16a-e) de procesamiento interconectados a través de los cuales los flujos de entrada pasan para generar una segunda pluralidad de flujos (3) de salida de registros de datos procesados, en el que las pluralidades primera y segunda son diferentes, el método comprendiendo:
- 60 a) introducir, mediante un módulo (36) de control de entrada, cada uno de los flujos de entrada recibidos en la unidad (32) de procesamiento de datos, teniendo cada flujo de entrada un formato de datos definido de manera remota que está asociado con un subconjunto específico de los módulos (16a-e) de procesamiento de la unidad (32) de procesamiento de datos, el paso de entrada comprendiendo:
- 65 iii) generar registros de datos ficticios para insertar en los flujos de entrada; el paso de generación de registro de datos ficticios generando registros de datos ficticios con un identificador conocido en un campo de datos inmutables del registro de datos y que tiene un formato seleccionado de uno de los formatos de datos definidos de manera remota; e
- iv) inyectar la primera pluralidad de flujos de entrada en la unidad (32) de procesamiento de datos, en el que el paso de inyección comprende intercalar un registro ficticio identificable de forma única en los registros de datos de la

primera pluralidad de flujos de entrada que tienen el mismo formato definido de manera remota que el registro ficticio y repetir el intercalado de otros registros ficticios en intervalos de tiempo predeterminados regulares según lo determinado por un reloj (63, 88) de referencia;

- 5 d) leer, mediante un filtro (39) de registro de datos, el campo de datos inmutables de cada registro de datos en cada flujo de salida e identificar y separar registros ficticios de los flujos de salida que tienen el identificador conocido en el campo de datos inmutables;
- 10 e) analizar los registros de datos ficticios, comprendiendo el paso de análisis:
- a. recibir registros de datos ficticios del filtro (39) de registro de datos;
- 15 b. determinar un valor de latencia actual del registro de datos ficticios a través de la unidad (32) de procesamiento de datos usando el conocimiento de cuándo el registro de datos ficticios se intercaló en el flujo de datos de entrada y la hora de recepción del registro de datos ficticios usando el reloj de referencia (88);
- c. comparar el valor de latencia actual con un valor de latencia determinado previamente de un registro de datos ficticios anterior para determinar si hay un cambio significativo en el valor de latencia;
- 20 d. usar el formato del registro ficticio para determinar a qué subconjunto de módulos (16a-e) de procesamiento de la unidad (32) de procesamiento de datos se refiere la latencia; y
- e. determinar y emitir un comando (45) de realimentación apropiado para que el módulo (36) de control de entrada tome medidas para abordar el cambio significativo en la latencia detectada;
- 25 en el que el paso de entrada comprende recibir el comando (45) de realimentación y regular el caudal de flujos de entrada de datos de acuerdo con el comando de realimentación.

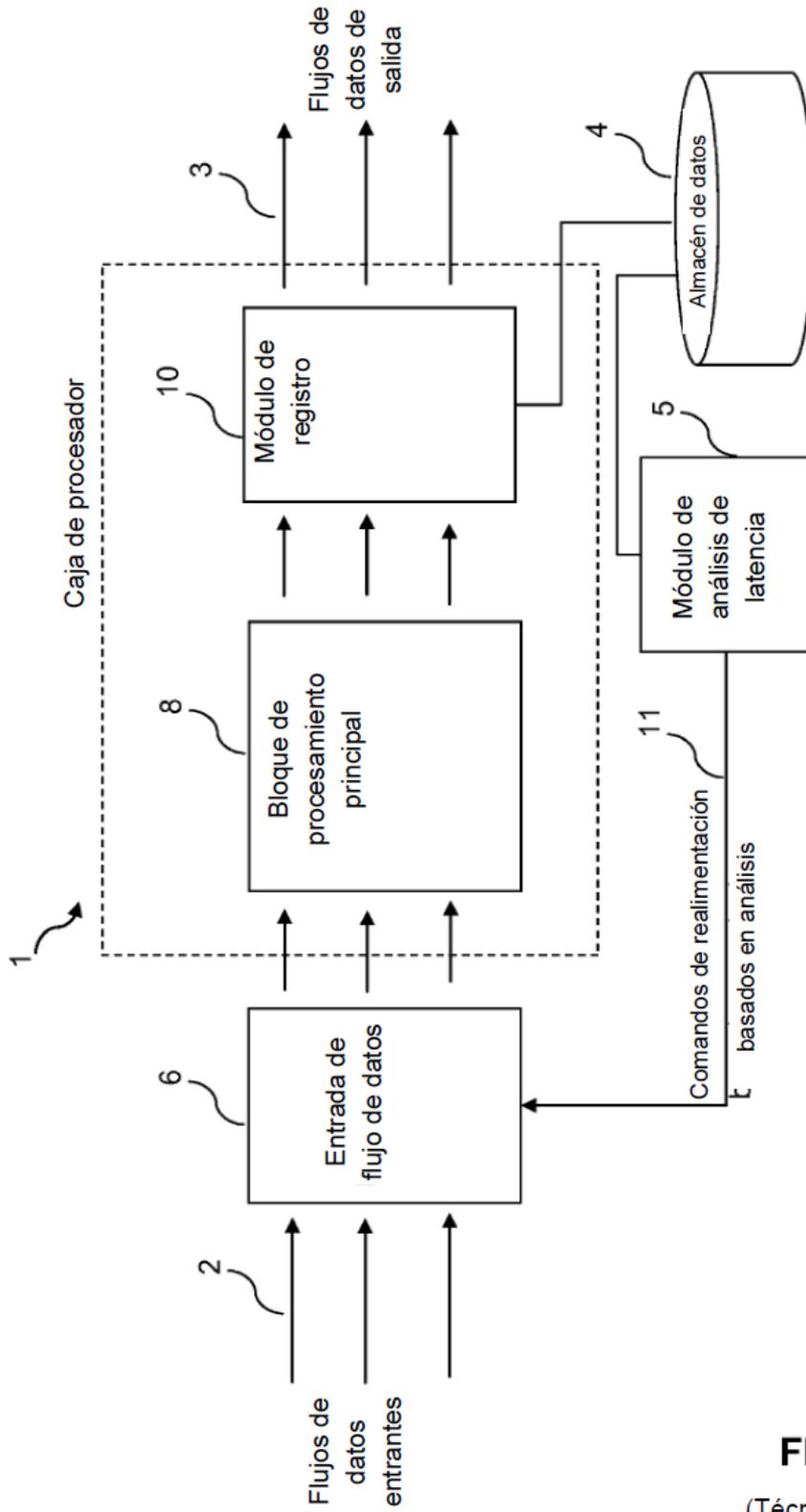


FIG. 1

(Técnica anterior)

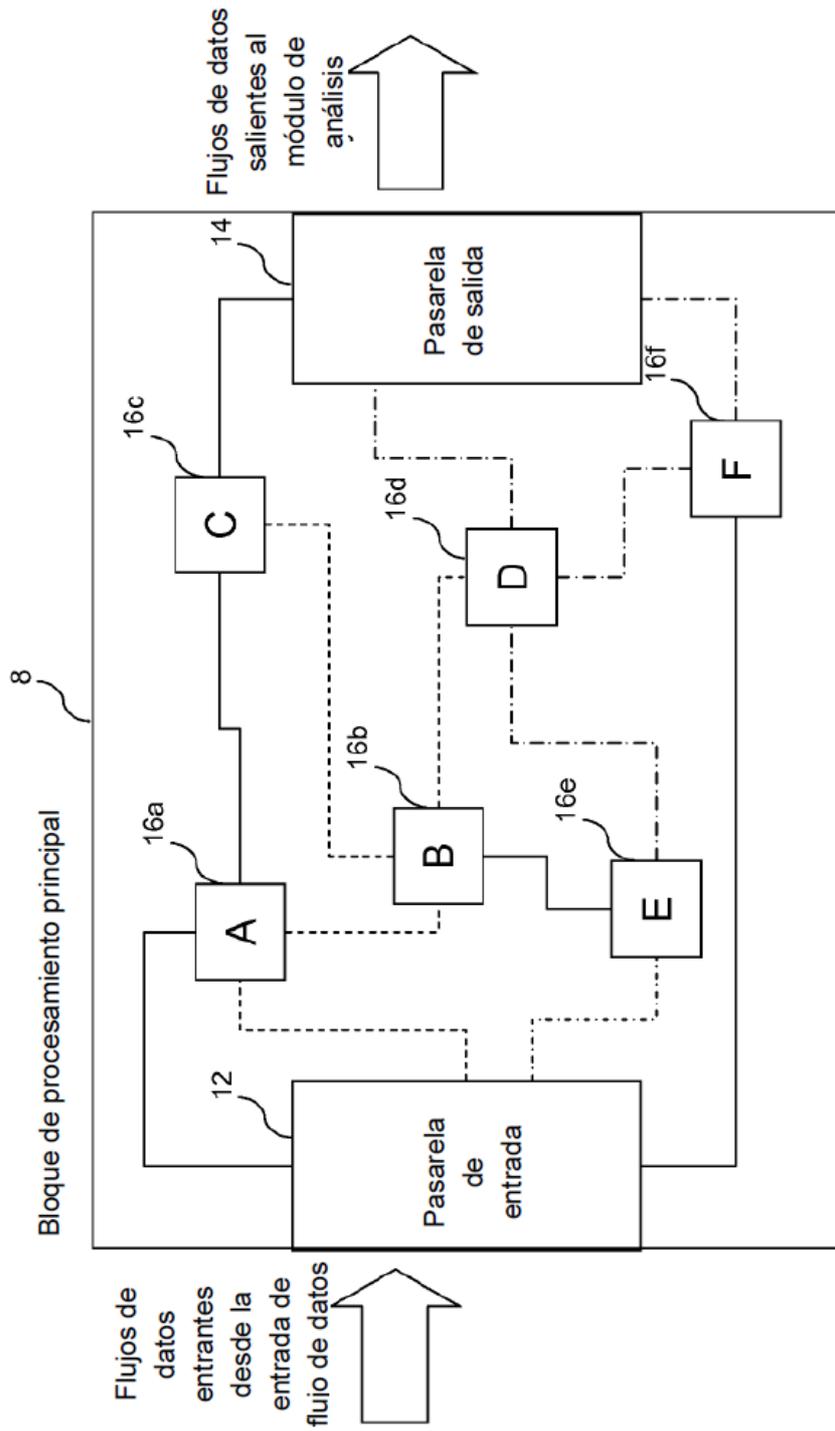


FIG. 2

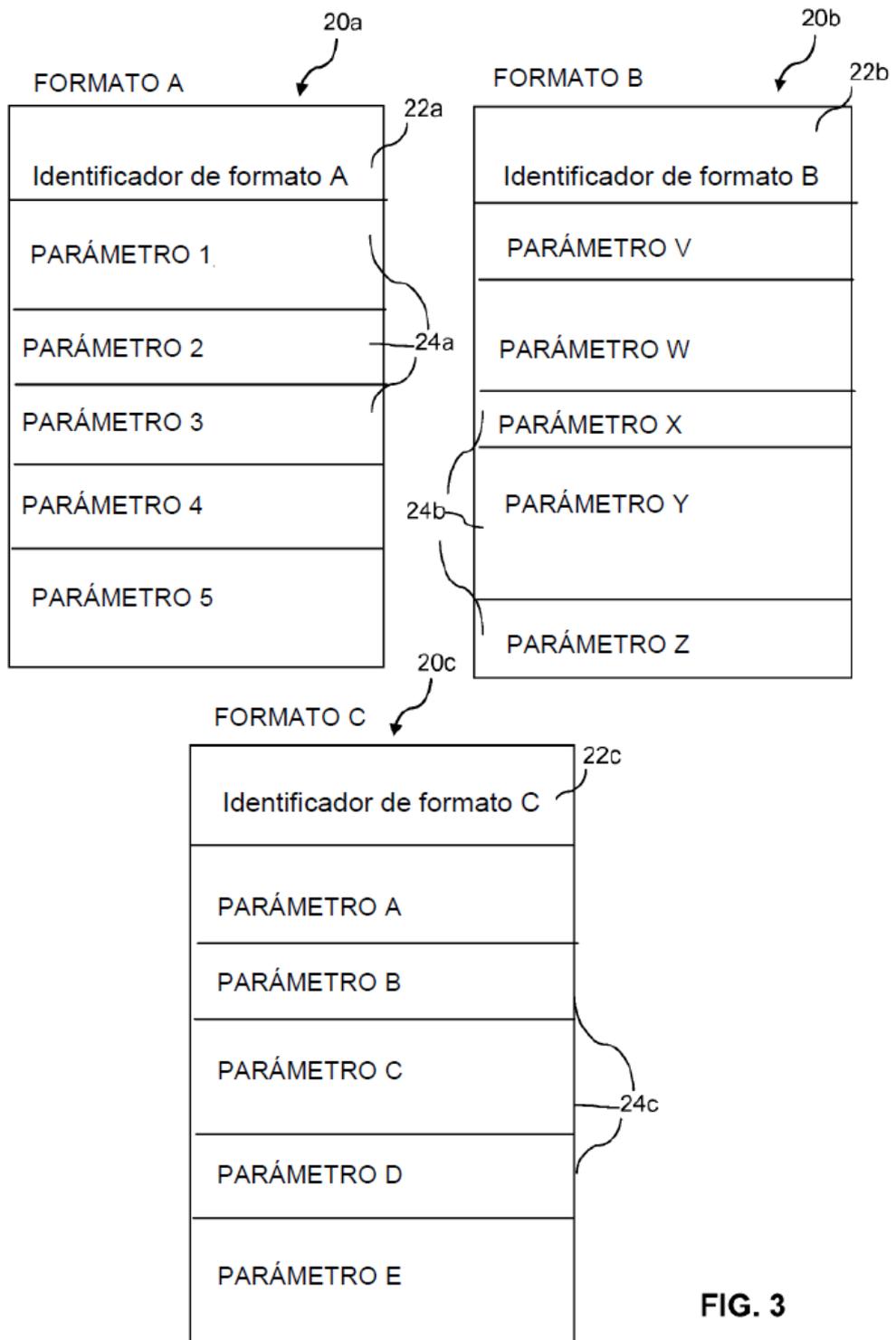


FIG. 3

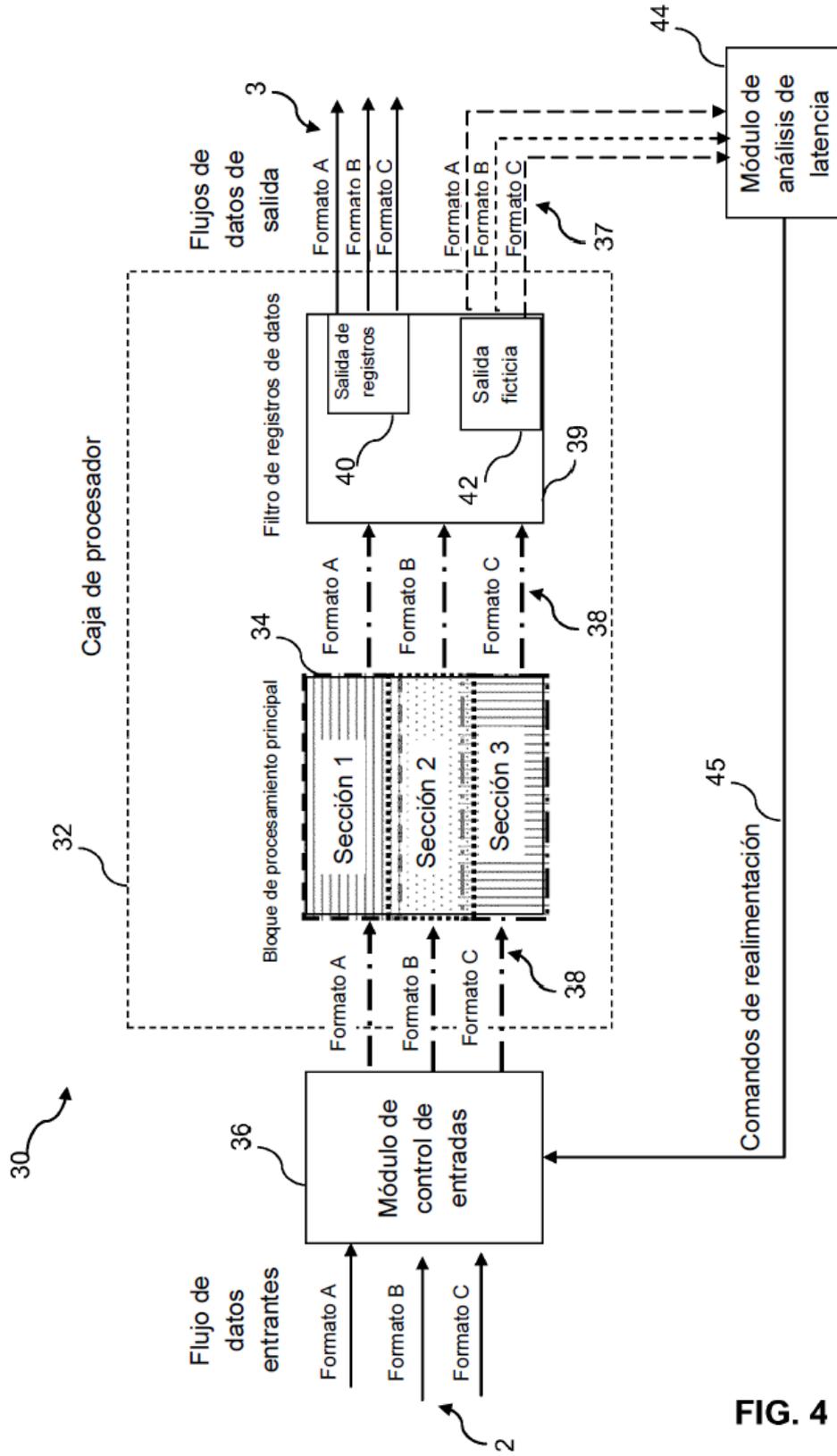


FIG. 4

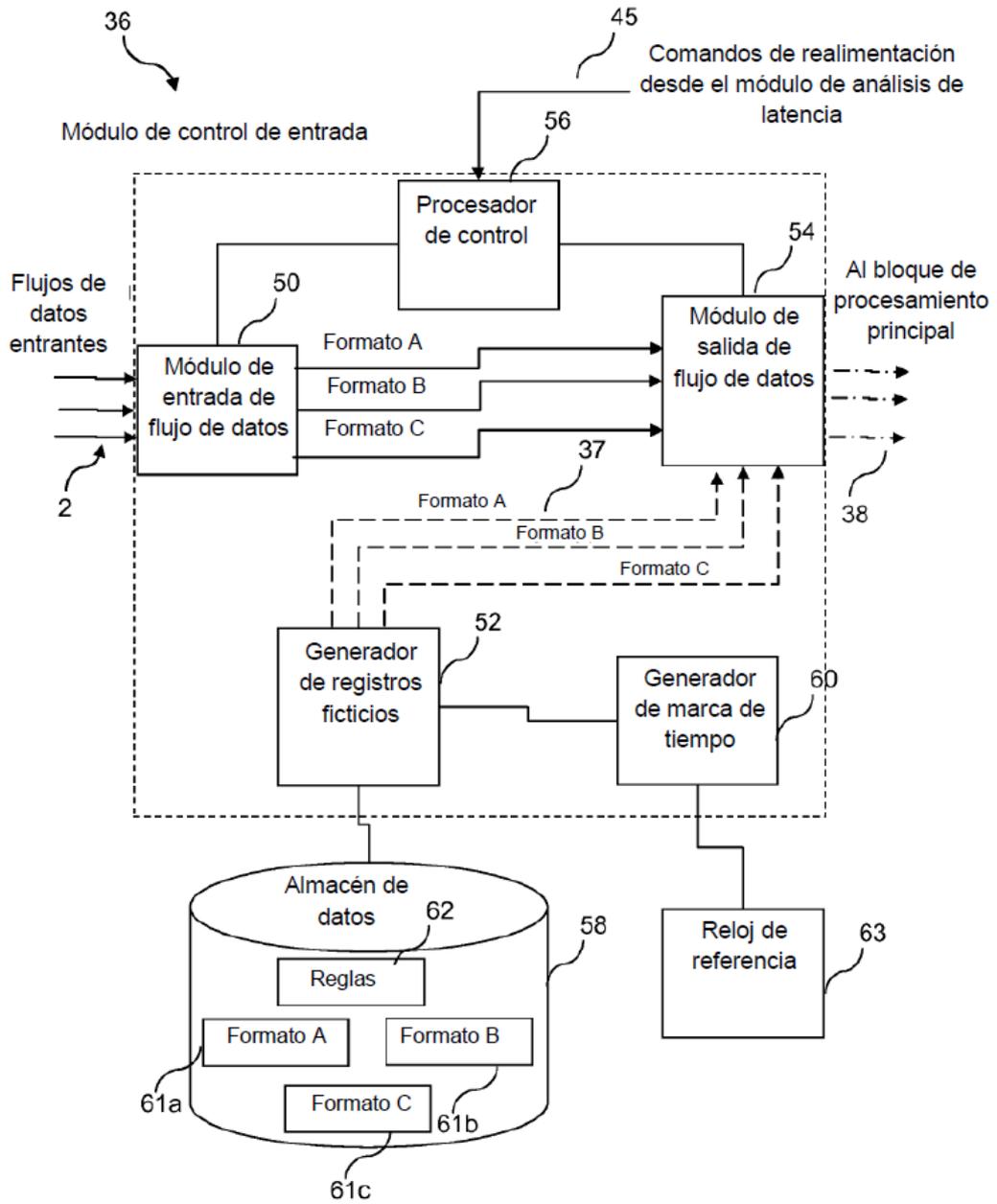


FIG. 5

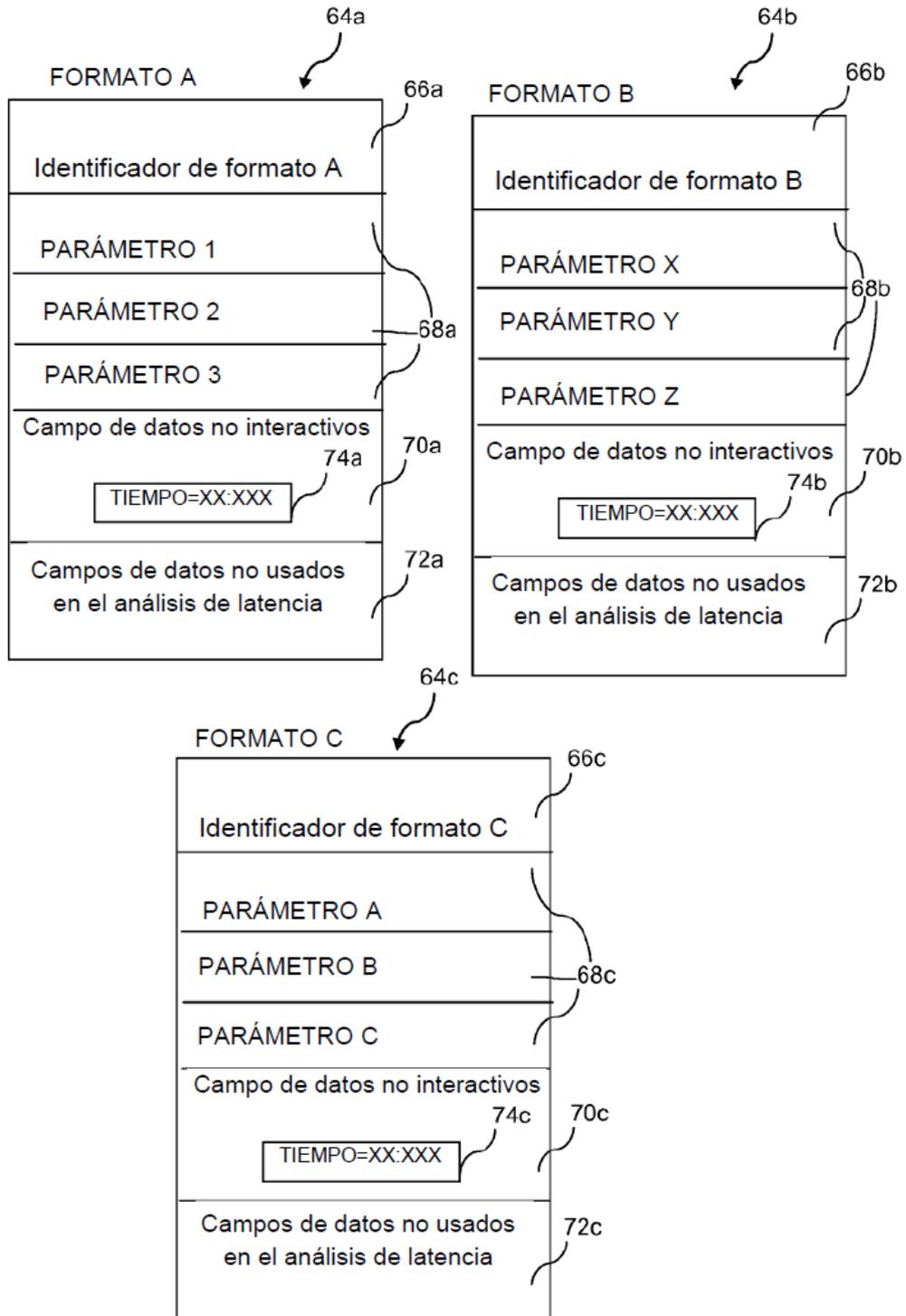


FIG. 6

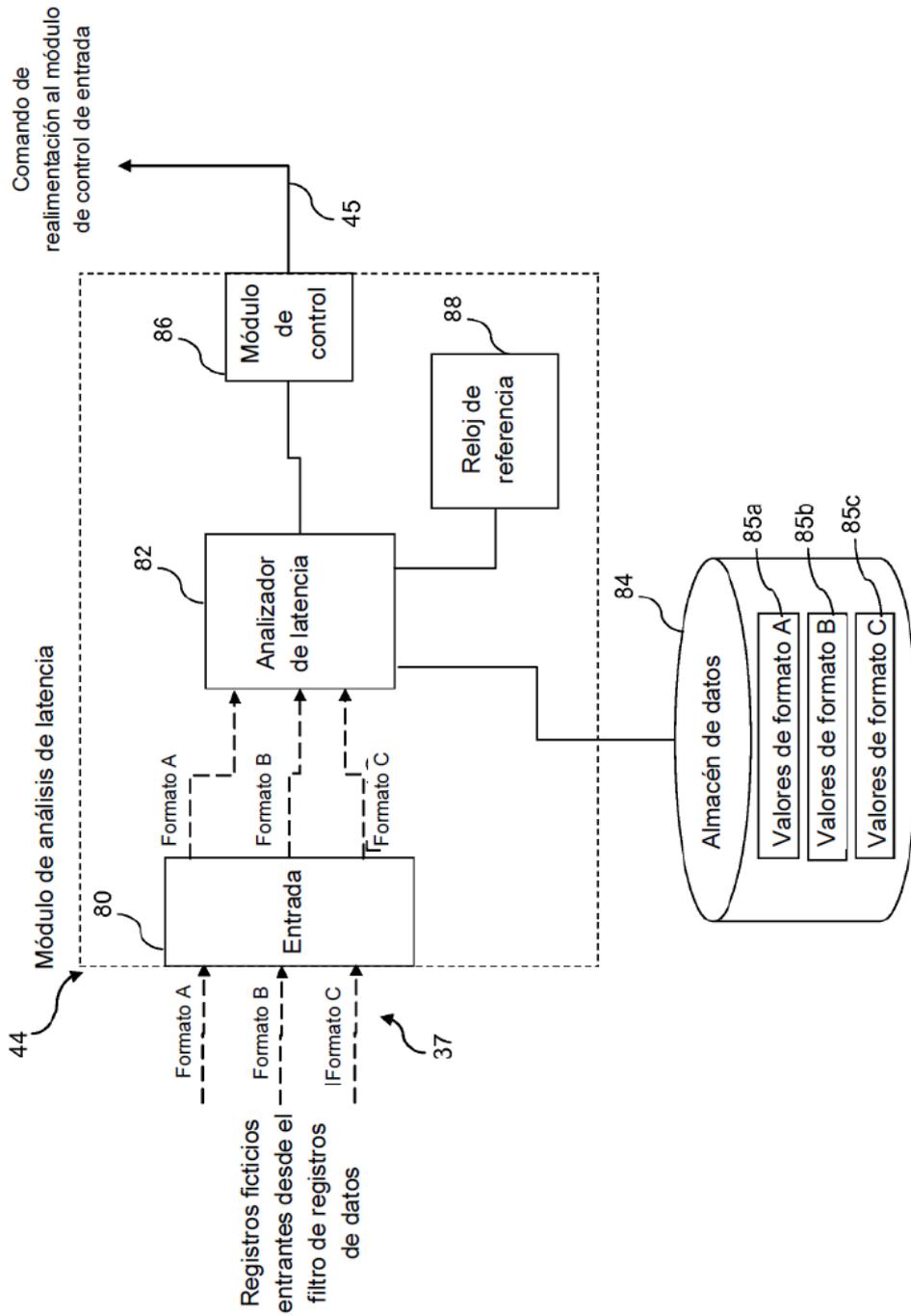


FIG. 7

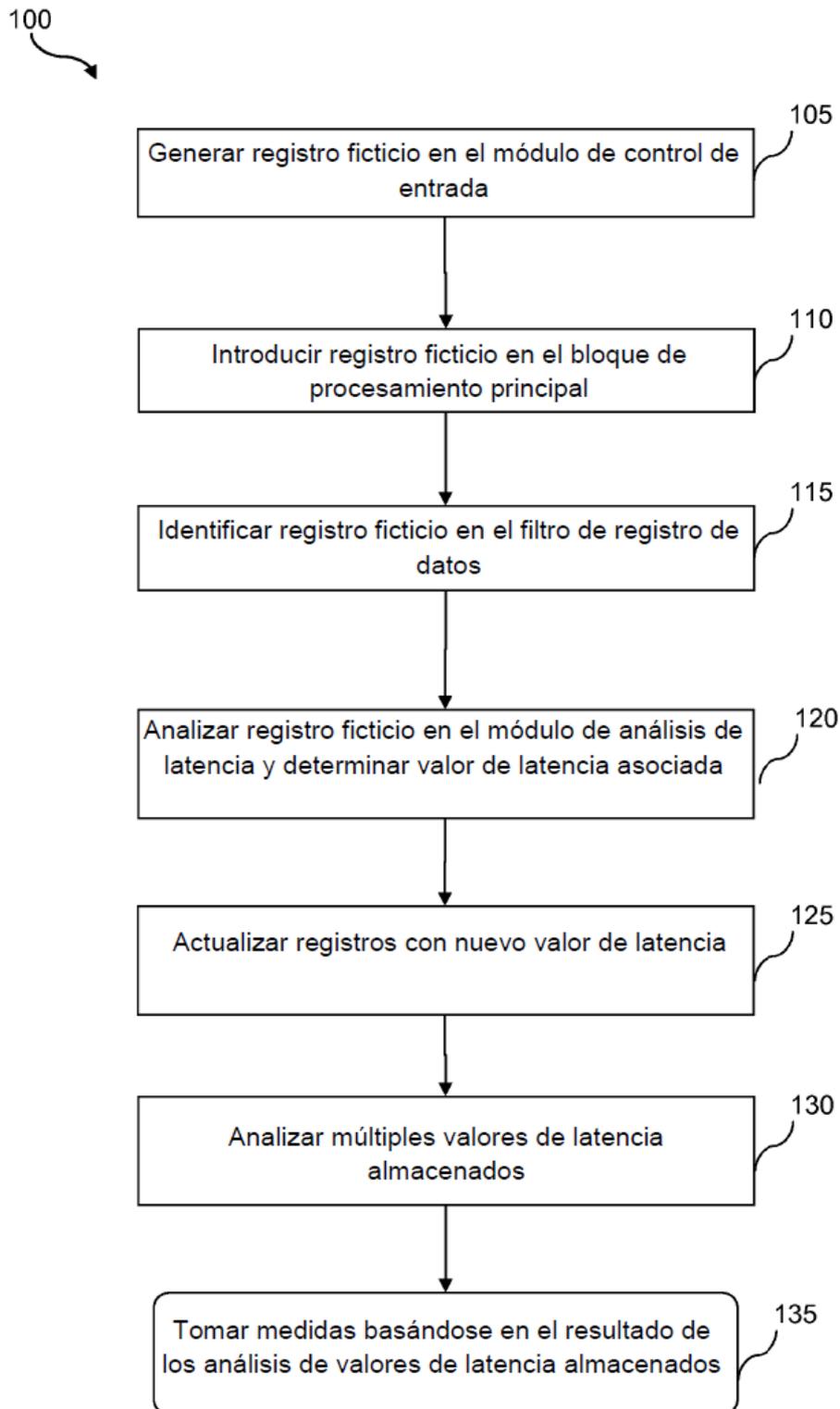


FIG. 8

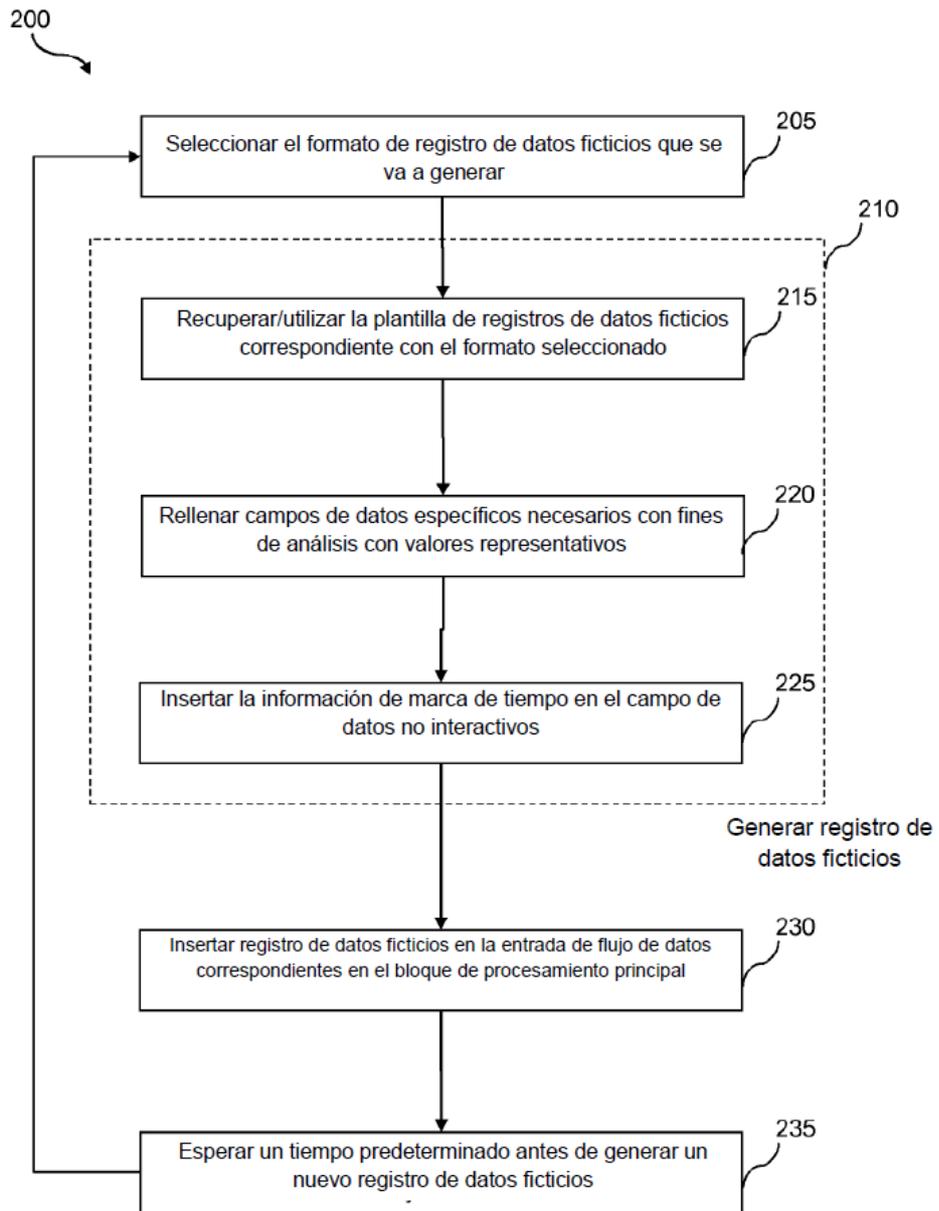


FIG. 9

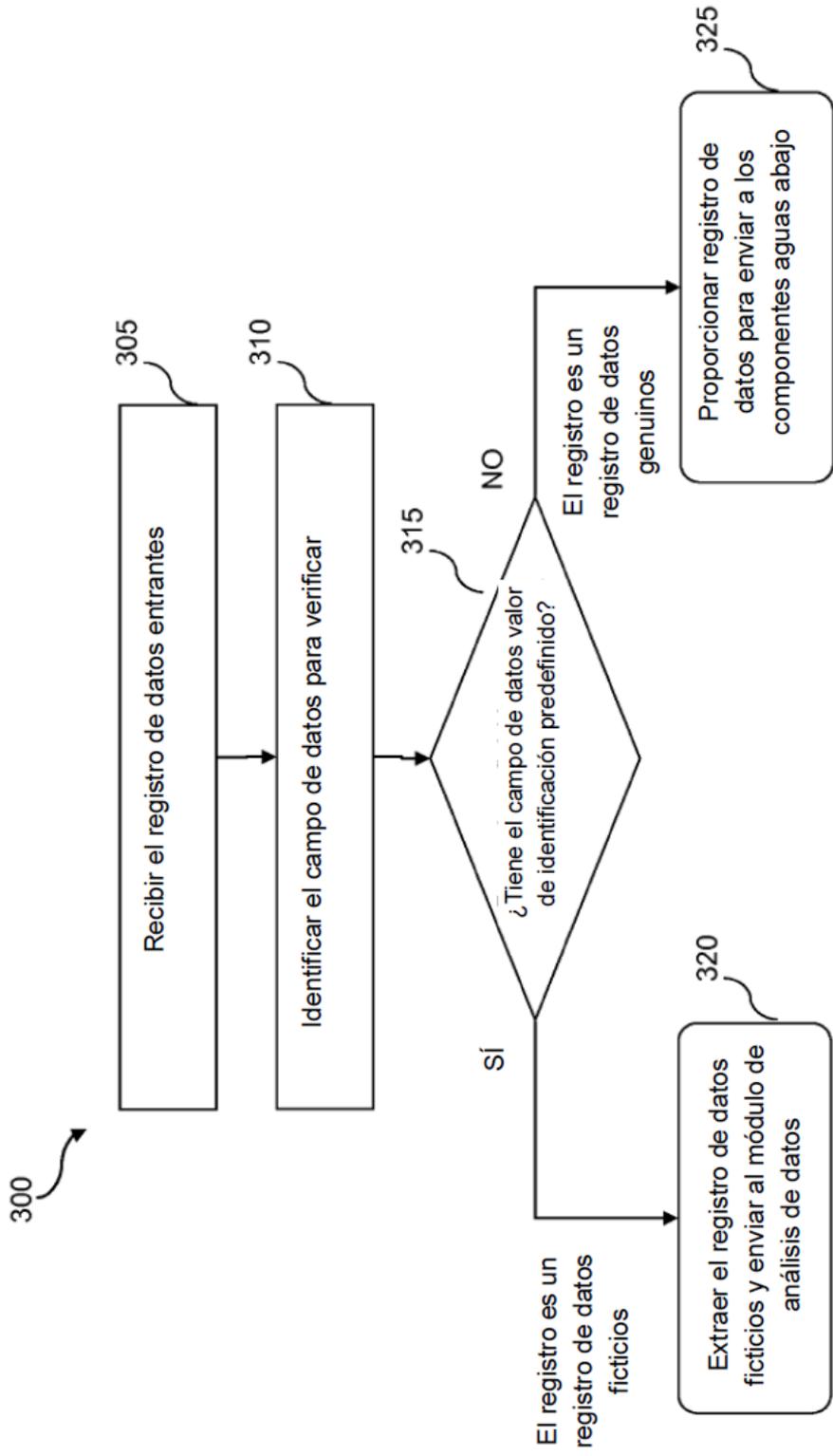


FIG. 10

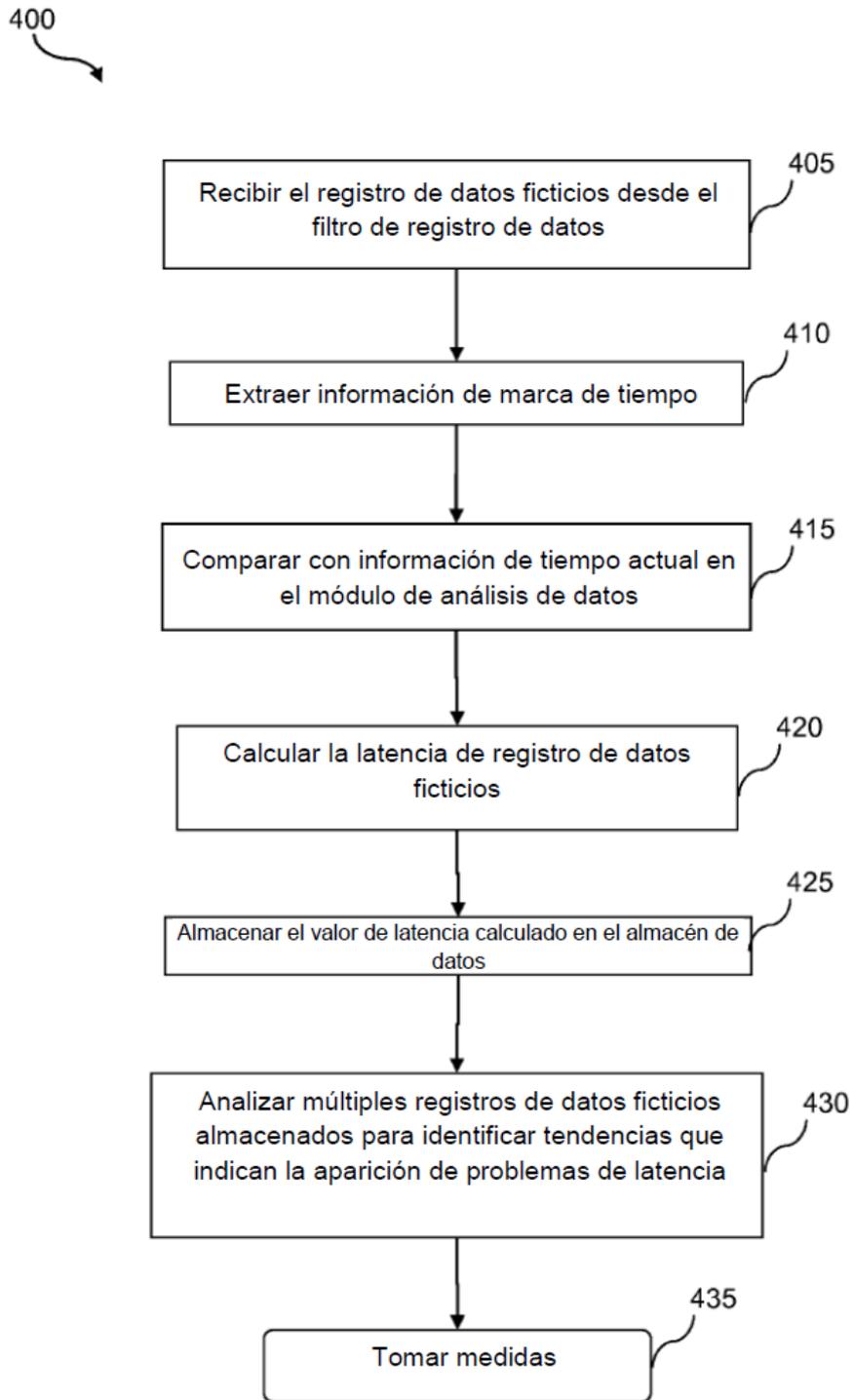


FIG. 11

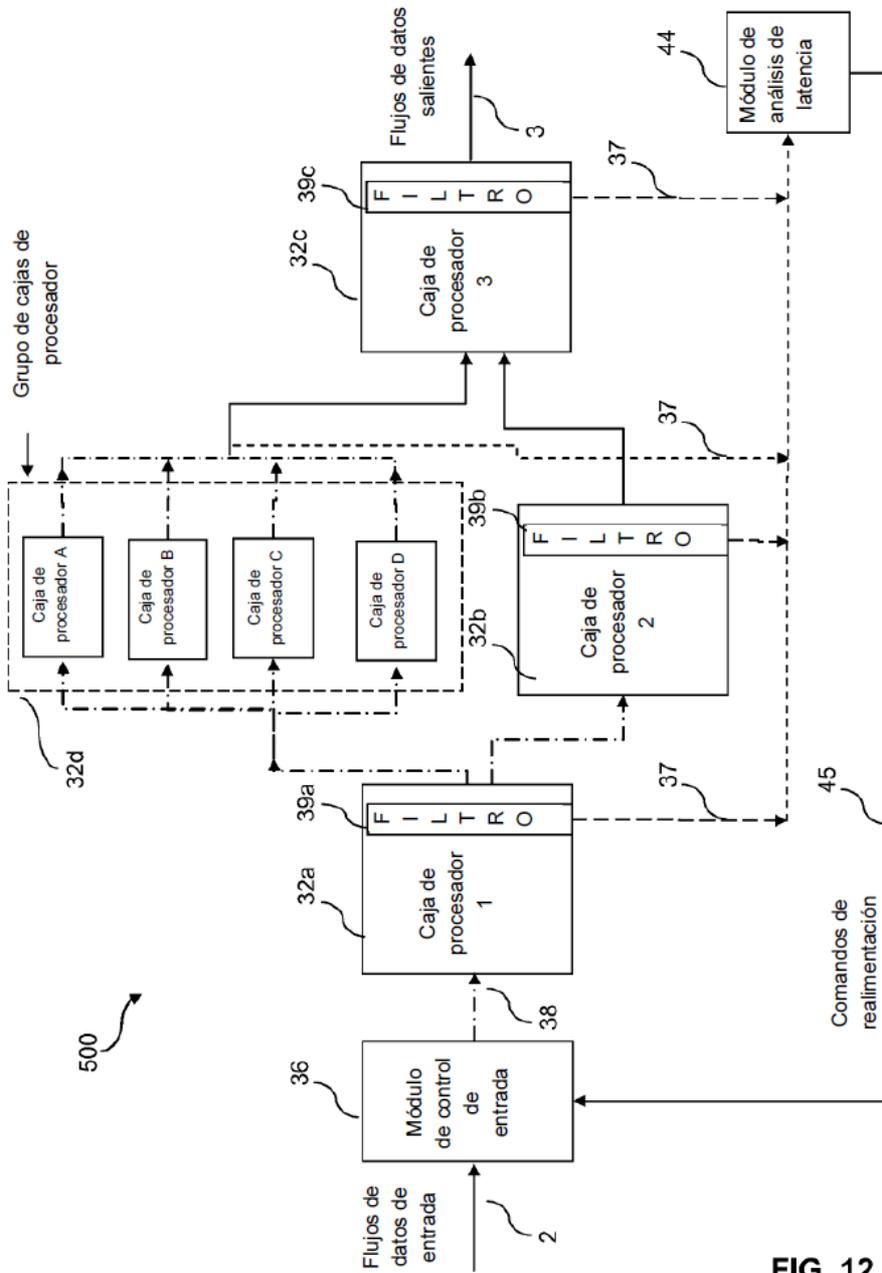


FIG. 12