

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 991 404**

51 Int. Cl.:

G06F 16/2455 (2009.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **03.08.2015** **E 18175669 (3)**

97 Fecha y número de publicación de la concesión europea: **02.10.2024** **EP 3392788**

54 Título: **Manejo de peticiones de datos**

45 Fecha de publicación y mención en BOPI de la
traducción de la patente:
03.12.2024

73 Titular/es:

AMADEUS S.A.S. (100.0%)
485 route du Pin Montard, Sophia Antipolis
06410 Biot, FR

72 Inventor/es:

LEGRAND, GUILLAUME

74 Agente/Representante:

ELZABURU, S.L.P

ES 2 991 404 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

Manejo de peticiones de datos

La presente invención se refiere en general a la tecnología de bases de datos. Más específicamente, se dirige a un mecanismo para manejar peticiones de datos entrantes que tiene en cuenta la exactitud de los resultados preparados que se devolverán a una petición de datos.

Antecedentes de la técnica

Un problema común en la tecnología de bases de datos es garantizar tiempos de respuesta cortos a consultas o peticiones de bases de datos que requieren un procesamiento complejo, largo y que consume muchos recursos, p. ej., debido a grandes volúmenes de datos. Por ejemplo, dicho procesamiento que consume mucha potencia de cálculo debe realizarse en respuesta a las denominadas "consultas abiertas" que contienen solo poca información de entrada (p. ej., solo se especifican uno o dos parámetros de una docena de parámetros posibles y/o los intervalos de valores especificados de los parámetros son amplios) y, en consecuencia, dan lugar a una gran cantidad de resultados en general. Las posibilidades de acelerar el procesamiento de datos al aumentar el rendimiento del hardware son limitadas. Por tanto, se llama la atención sobre la mejora de los mecanismos subyacentes al procesamiento complejo, largo y que requiere muchos recursos para responder a dichas consultas.

Un planteamiento general para acortar los tiempos de respuesta es preparar los resultados que se devolverán en respuesta a las peticiones previstas (p. ej., calcular o recopilar previamente dichos resultados) y mantener los resultados de la consulta correspondiente en un grupo de resultados preparados. En realidad, las peticiones y respuestas no se procesan a partir de la gran base de datos original, sino que se dirigen al grupo de resultados preparados.

Sin embargo, otro problema que surge con estos planteamientos de preparación de resultados es mantener los resultados preparados actualizados para garantizar que los resultados preparados devueltos en respuesta a las peticiones de datos reflejen correctamente el resultado de un procesamiento complejo, largo y que requiere muchos recursos. En caso de que la base de datos original subyacente cambie, los resultados preparados quedan desactualizados y responder a las peticiones a partir del grupo de resultados preparados entregaría unos resultados preparados incorrectos. Por tanto, se emplean estrategias de actualización para mantener actualizado el grupo de resultados preparados.

En la técnica anterior son conocidas diversas estrategias de actualización relativamente sencillas como, por ejemplo, volver a calcular frecuentemente todo el dominio de resultados preparados, establecer y mantener manualmente calendarios de recálculo y volver a calcular los resultados preparados cuando son demasiado antiguos.

Se han desarrollado estrategias de actualización algo más sofisticadas, como las descritas, p. ej., en el documento WO 01/33472 y WO 02/25557.

El documento WO 01/33472 se refiere a un sistema de disponibilidad utilizado en un sistema de planificación de viajes. El sistema incluye una memoria caché que contiene entradas con información sobre disponibilidad de asientos de aerolíneas. Un gestor de la memoria caché gestiona la información de entrada en la memoria caché para mantener la información en la memoria caché correcta, actual, completa o lo más útil posible. En respuesta a una consulta dirigida a la memoria caché, el gestor de la memoria caché determina si una respuesta almacenada está obsoleta y, si es así, envía una consulta de disponibilidad a una fuente de información de disponibilidad. Las entradas de memoria caché que se modificarán se obtienen mediante notificaciones asíncronas de sistemas externos y se determinan mediante un modelo determinista, predictivo o estadístico.

De forma similar, el documento WO 02/25557 se refiere a un sistema de recuperación de información en el que la información recibida de fuentes de información se almacena en memoria caché para su uso futuro, tal como por ejemplo para futuras peticiones de clientes. Se pueden generar consultas proactivas para rellenar una memoria caché y/o actualizar información almacenada actualmente en memoria caché. En un sistema de información de una aerolínea, las consultas proactivas se ordenan sobre la base de datos estadísticos o indicaciones predictivas tales como la proximidad de la hora de salida, la antigüedad de los datos almacenados en memoria caché, los asientos restantes en una aeronave, las vacaciones o eventos especiales o el tipo de equipo. Además, se reciben actualizaciones mediante notificaciones externas de las aerolíneas tales como los mensajes AVS.

Además, el documento WO 99/22315 describe un mecanismo para actualizar automáticamente documentos en una memoria caché utilizando un modelo probabilístico basado en datos estadísticos. Para cada documento, la memoria caché determina una probabilidad $Psi(t)$ de que un objeto almacenado en memoria caché i esté desactualizado en un momento particular t (es decir, el servidor ha cambiado ese objeto) y una probabilidad $Pri(h)$ de que un usuario solicite el objeto i en el momento de petición h . La memoria caché actualiza aquellos objetos con el mayor producto $Pi = Psi(t) \times Pri(h)$, es decir, la probabilidad de que se devuelva un objeto desactualizado al usuario en la próxima petición. Para mantener estos valores de probabilidad, la memoria caché mantiene y rastrea datos estadísticos históricos de los objetos almacenados en memoria caché, tales como una estimación del intervalo medio entre actualizaciones de EUI en el servidor. El EUI de un objeto se actualiza, p. ej., cuando el servidor actualiza el objeto mismo o cuando el objeto no se actualiza una vez transcurrido su tiempo de actualización medio estimado.

El documento US2009/0204753 A1 se refiere a la actualización de una memoria caché en función de las respuestas de consultas provistas por un sistema de búsqueda en respuesta a consultas. Se proporciona una entrada de memoria caché para cada consulta única si hay espacio disponible en la memoria caché. Se asigna un valor de temperatura a cada entrada de memoria caché en función de la frecuencia de aparición de la consulta correspondiente. Se asigna un valor de antigüedad a cada entrada de memoria caché en función del momento de la última actualización o creación de la respuesta de la consulta correspondiente. La antigüedad de las entradas de memoria caché se actualiza periódicamente y la temperatura de una entrada de memoria caché se actualiza cuando se produce una consulta correspondiente. Si hay recursos del sistema disponibles, la respuesta de la consulta de una entrada de memoria caché se actualiza en función de la temperatura y la antigüedad de la entrada de memoria caché. La actualización es limitada si los recursos no están disponibles.

El artículo "Synchronizing a database to Improve Freshness" de J. Cho et al. estudia las posibilidades de actualizar una copia local de una fuente de datos autónoma para mantener la copia actualizada. Entre otras cosas, el proceso de Poisson se describe como un planteamiento para modelar la evolución probabilística de un elemento.

Además, el documento WO 99/22315 A1 proporciona un procedimiento y un sistema para actualizar automáticamente documentos en una memoria caché. Para cada documento, la memoria caché estima una distribución de probabilidad de tiempos para las peticiones de clientes para ese documento y una distribución de probabilidad de tiempos para los cambios por parte del servidor en ese documento. La memoria caché estima una distribución de probabilidad de tiempos para las peticiones de clientes con ese documento y determina una distribución de probabilidad acumulativa que refleja la tasa de llamada marginal estimada en el límite de almacenamiento de la memoria caché y la ventaja marginal de añadir almacenamiento a la memoria caché.

En la documentación "Oracle Cache. Administration Deployment Release, 1.0.2.3", de Deborah Steiner et al., se describe cómo utilizar Oracle Web Cache para almacenar en memoria caché contenido estático y generado dinámicamente desde uno o más servidores web de aplicaciones.

Compendio de la invención

La presente invención se define en las reivindicaciones independientes.

Según un aspecto, se proporciona un procedimiento para manejar peticiones de datos dirigidas a un entorno de base de datos. El entorno de la base de datos tiene por lo menos una primera plataforma que proporciona resultados originales que se almacenarán en una segunda plataforma como resultados preparados. La segunda plataforma mantiene un grupo de resultados preparados por medio de la primera plataforma que se devolverán a las peticiones de datos. El entorno de base de datos también dispone de una unidad de control para procesar las peticiones de datos dirigidas al entorno de base de datos. Cada resultado preparado que se mantiene en el grupo de la segunda plataforma está asociado con un indicador de actualización que es una medida de que el resultado preparado asociado que se mantiene en el grupo de la segunda plataforma debe actualizarse. La unidad de control recibe una petición de datos. La unidad de control determina por lo menos un resultado preparado correspondiente a la petición de datos. Las unidades de control comparan el indicador de actualización del resultado preparado determinado correspondiente a la petición de datos con un valor umbral. Si la comparación indica el requisito de actualizar el resultado preparado, la unidad de control recupera una versión actualizada del por lo menos un resultado preparado de la primera plataforma. La unidad de control actualiza el resultado preparado en el grupo de la segunda plataforma y el indicador de actualización asociado en función de la versión actualizada del por lo menos un resultado preparado al almacenar una nueva marca de tiempo de la actualización del por lo menos un resultado preparado. La unidad de control devuelve la versión actualizada de por lo menos un resultado. Por otro lado, si la comparación no indica el requisito de actualizar el por lo menos un resultado preparado, la unidad de control devuelve el por lo menos un resultado preparado determinado. El indicador de actualización se define por $(1 - acc) \cdot t$, en el que acc es una probabilidad de que el resultado preparado asociado sea válido y t es una antigüedad del resultado preparado asociado, en el que $acc = e^{-\lambda t}$, en el que la tasa de validez λ es un indicador de la frecuencia con la que cambia el resultado preparado.

De esta manera, se mejora la exactitud de los resultados preparados (es decir, los resultados preparados todavía están en línea con los datos de respuesta originales) a largo plazo, como se explicará en detalle a continuación.

Según un segundo aspecto, se proporciona una unidad de control dispuesta en consecuencia.

Según un tercer aspecto, se proporciona un programa informático que se ejecutará mediante una unidad de control y que habilitará a la unidad de control con la funcionalidad respectiva.

En las reivindicaciones dependientes se describen aspectos adicionales.

Breve descripción de las figuras

La descripción posterior de realizaciones se basa en el conjunto de figuras adjuntas en el que números de referencia similares se refieren a elementos y mensajes similares y en el que:

la FIG. 1 muestra una descripción esquemática de alto nivel de la arquitectura del sistema como se describe en la presente memoria;

las FIG. 2A y 2B representan un flujo de mensajes de alto nivel para procesar peticiones de datos mediante el uso de indicadores de actualización;

5 la FIG. 3 representa un efecto de disminución de la probabilidad de validez de los resultados preparados;

la FIG. 4 presenta una arquitectura de sistema ejemplar del entorno de base de datos;

las FIG. 5A y 5B representan un ejemplo de flujo de mensajes más específico para procesar peticiones de datos mediante el uso de indicadores de actualización;

10 la FIG. 6A y 6B representan otro ejemplo de flujo de mensajes para procesar peticiones de datos mediante el uso de indicadores de actualización;

la FIG. 7 visualiza un bucle de realimentación entre la primera plataforma y la unidad de control para adaptar dinámicamente el valor umbral;

la FIG. 8 muestra un ejemplo de una estructura interna de una unidad de control.

Descripción detallada

15 Antes de pasar a la descripción de las realizaciones sobre la base de las FIG. 4 a 8, se explican primero algunos aspectos más generales con referencia a las FIG. 1 a 3.

Como ya se ha explicado al principio, las metodologías descritas en la presente memoria se refieren a sistemas de bases de datos que ofrecen resultados preparados a clientes que solicitan datos de un sistema de base de datos. El objetivo técnico de preparar los resultados antes de que los clientes los soliciten es en general disminuir los tiempos de respuesta a dichas peticiones de datos. De aquí en adelante, el término "preparado" se utiliza para referirse a cualquier tipo de procesamiento, cálculo y recopilación previos de los resultados de forma asíncrona, es decir, independientemente y antes de una petición de datos de un cliente. Ejemplos de preparación de resultados son los rastreadores de Internet que recopilan o copian el contenido de los servidores web de Internet, así como los resultados generados en respuesta a peticiones de datos anteriores que se mantienen en una memoria caché para una recuperación rápida, pero también los resultados de búsqueda a partir de datos subyacentes que requieren cálculos complejos y que requieren mucho tiempo, como se describe, p. ej., para las recomendaciones de precios de viajes en los documentos PCT/EP2013/002390 y EP 2541473 A1. El término "base de datos" pretende abarcar cualquier tipo de sistema de almacenamiento de información estructurada, tal como bases de datos autónomas estándar como SQL Server u Oracle, así como sistemas de almacenamiento complejos, distribuidos y/o propietarios, bases de datos relacionales, incluidos sistemas de gestión de bases de datos o sistemas de bases de datos orientados a objetos y similares. El término "petición de datos" se utiliza en la presente memoria como un término general para cualquier tipo de consulta a un sistema de base de datos con el fin de recuperar datos, que incluyen cualquier tipo de petición de recuperación de información, tales como consultas transaccionales, peticiones de cálculos por lotes, consultas SQL y otras formas.

35 La FIG. 1 ilustra dicho entorno 1 de base de datos en un nivel abstracto. Los datos originales, de aquí en adelante también denominados resultados originales, se proporciona mediante una primera plataforma 3. En general, la primera plataforma 3 es una fuente de datos originales en sí misma, tal como una base de datos de inventario o una base de datos que mantiene cualquier tipo de resultados originales y en general válidos, o accede a una o más fuentes de datos originales para preparar resultados originales. Ejemplos de este último tipo son un motor de búsqueda que accede a sitios web de Internet y una plataforma de cálculo que calcula recomendaciones de precios de viajes basadas en tarifas almacenadas en una base de datos de tarifas. Si la primera plataforma 3 genera/calcula/recopila los resultados originales al acceder a otras fuentes de datos originales para preparar resultados originales, la primera plataforma 3 proporciona resultados que en general reflejan con exactitud el contenido actual de los datos de respuesta originales. Puede haber más de una primera plataforma 3 presente en el entorno 1 de base de datos.

45 El entorno 1 de base de datos también incluye una segunda plataforma 4 que mantiene un grupo 5 de resultados que se han preparado por medio de la primera plataforma 3. La segunda plataforma 4 es, p. ej., una memoria caché que almacena en memoria caché los resultados que se han recuperado por medio de la primera plataforma 3, o la segunda plataforma 4 es, p. ej., un servidor que contiene un repositorio del contenido de los sitios web rastreados, o la segunda plataforma 4 es, p. ej., una plataforma de búsqueda que mantiene recomendaciones de precios de viajes calculados previamente, como se describe en el documento WO 2014/026753 A1.

55 En general, la primera plataforma 3 que proporciona datos de respuesta originales y la segunda plataforma 4 que mantiene el grupo 5 de resultados preparados difieren entre sí en términos técnicos como sigue: La segunda plataforma 4, debido a la preparación de los resultados, proporciona un acceso más barato (en términos de costes de cálculo) y/o más rápido a los resultados en comparación con la primera plataforma 3. Esta relación entre la primera plataforma 3 y la segunda plataforma 4 también se puede expresar al revés, es decir, la primera plataforma 3 en

general responde más lentamente y las respuestas de la primera plataforma 3 implican costes de cálculo más altos que las respuestas de la segunda plataforma 4 porque la primera plataforma 3 todavía tiene que preparar los resultados (p. ej., recuperar el contenido solicitado del sitio web de Internet original, o calcular las recomendaciones de precios de viaje solicitadas según las tarifas guardadas en una base de datos de tarifas). Como los recursos computacionales de la primera plataforma 3 son limitados, el número de accesos a la primera plataforma 3 debe mantenerse dentro de un determinado límite. Por ejemplo, el consumo de CPU o la carga de memoria de la primera plataforma 3 no debe superar un límite dado. Esta restricción puede traducirse en un límite del número de accesos a la primera plataforma por unidad de tiempo (p. ej., 10 accesos por segundo). Por este motivo, no se solicita cada vez la primera plataforma 3 para responder a las peticiones de datos. Más bien, las respuestas en general se forman mediante el uso de los resultados preparados almacenados en el grupo 5 de la segunda plataforma 4. De esta manera, la segunda plataforma 4 actúa como un escudo frente a la primera plataforma 3, lo que reduce la carga en la primera plataforma 3 (computacionalmente costosa).

Por otro lado, sin embargo, los resultados preparados almacenados en el grupo de la segunda plataforma 4 no están necesariamente perfectamente sincronizados con los resultados originales provistos por la primera plataforma 3. En general, los resultados preparados de la segunda plataforma 4 representan versiones anteriores del resultado original provisto por la primera plataforma 3 y la versión respectiva del resultado original provisto por la primera plataforma 3 podría haber cambiado desde la última actualización del resultado preparado correspondiente que se mantiene en la segunda plataforma 4. Además, la segunda plataforma 4 no necesariamente almacena los resultados preparados correspondientes a todos los resultados originales existentes, es decir, es posible que falten resultados preparados en la segunda plataforma 4. Los resultados preparados almacenados en el grupo 5 de la segunda plataforma 4 que reflejan correctamente su resultado correspondiente provisto por la primera plataforma se denominan de aquí en adelante resultados preparados válidos o resultados preparados exactos, mientras que los resultados preparados obsoletos almacenados en el grupo 5 de la segunda plataforma 4 se denominan resultados preparados no válidos o resultados preparados inexactos.

Por tanto, surge el problema técnico de maximizar la exactitud y validez de los resultados preparados almacenados en el grupo 5 de la segunda plataforma 4 y la exactitud y validez de los resultados preparados devueltos a los clientes que realizan consultas, respectivamente, es decir, mantener los resultados preparados en la segunda plataforma 4 que sean congruentes con los datos de respuesta originales provistos por la primera plataforma 3 tanto como sea posible, con el fin de responder a la petición de datos con una alta porción de resultados preparados válidos.

El entorno 1 de base de datos también está acoplado a por lo menos uno, pero en general a una pluralidad de clientes 6. Los clientes 6, tal como las aplicaciones en terminales de usuario, recuperan resultados del entorno 1 de base de datos al dirigir peticiones de datos al entorno 1 de base de datos por medio de una interfaz 7. Las características técnicas de la interfaz 7 dependen de la implementación particular del entorno 1 de base de datos. Por ejemplo, la interfaz 7 abarca la comunicación inalámbrica que incluye 2G/3G/4G (incluido el intercambio de datos móviles orientado a paquetes, así como SMS) y/o la comunicación wifi en el caso de que el cliente 6 esté situado en un dispositivo de comunicaciones móviles. De forma adicional o alternativa, la interfaz 7 presenta comunicación por cable utilizando protocolos de red estándar empleados en redes de área local y/o redes de área amplia, que incluyen Internet, tales como Ethernet, TCP/IP, SMTP con POP3 o IMAP, HTTP, protocolos relacionados con servicios web tales como SOAP, etc.

Una petición de datos transmitida desde un cliente 6 por medio de la interfaz 7 incluye uno o más criterios de recuperación que restringen la petición. Por ejemplo, si la petición de datos es una petición de búsqueda en Internet, la petición de datos podría llevar una cadena de búsqueda, un texto de búsqueda o una frase de búsqueda como criterio de búsqueda. Otro criterio de búsqueda puede ser el idioma de los sitios web que se van a buscar o una indicación del momento en el que estuvo disponible por primera vez la cadena de búsqueda, el texto de búsqueda o la frase de búsqueda solicitados. Según otro ejemplo, la petición de datos es una petición de base de datos para un producto o servicio ofrecido por una plataforma de proveedor de servicios, tal como una librería de Internet o un proveedor de viajes. En ese caso, la petición de datos podría incluir, p. ej., un límite de precio superior o un intervalo de precios para el servicio o producto y las características deseadas del producto/servicio, tal como el título del libro, el origen y destino del viaje, etc.

Las peticiones de datos emitidas por un cliente 6 se reciben en una entidad adicional del entorno 1 de base de datos, la unidad 2 de control (FIG. 1). En general, la unidad 2 de control procesa las peticiones de datos entrantes para determinar si los resultados de la respuesta se recuperan a partir de la segunda plataforma 4 y/o de la primera plataforma 3. Con este fin, la unidad de control funciona como una unidad intermedia que controla el flujo de datos dentro del entorno 1 de base de datos y los resultados que se devolverán al cliente 6.

Básicamente, el control del flujo de datos mediante la unidad 2 de control tiene en cuenta los dos aspectos siguientes: Por una parte, como ya se ha explicado anteriormente, la recuperación de resultados preparados que cumplen los criterios de petición incluidos en una petición de datos del grupo 5 de la segunda plataforma 4 tiene la ventaja de un tiempo de respuesta más rápido (en comparación con la generación de los resultados por la primera plataforma 3) y libera a la primera plataforma 3 de posibles sobrecargas. Por otro lado, sin embargo, atender una petición de datos al recuperar los resultados originales provistos por la primera plataforma 3 permite actualizar los resultados de datos preparados respectivos que se mantienen en el grupo 5 de la segunda plataforma 4. Por tanto, una porción de los

resultados que se devolverán en respuesta a una petición de datos se puede recuperar deliberadamente de la primera plataforma 3, aunque los resultados preparados correspondientes estén disponibles en la segunda plataforma 4. Estos resultados son así preparados nuevamente por la primera plataforma 3 y devueltos al cliente solicitante 6 y almacenados en el grupo 5 de la segunda plataforma 4, de este modo se actualizan los resultados preparados correspondientes.

El problema técnico es entonces decidir qué porciones de las peticiones de datos que se transmiten desde los clientes 6 a la unidad 2 de control deben responderse sobre la base del grupo 5 de la segunda plataforma 4 y qué porciones de las peticiones de datos deben responderse sobre la base de la primera plataforma 3 para actualizar los resultados preparados que se mantienen en el grupo 5 de la segunda plataforma 4.

Con este fin, la unidad 2 de control utiliza un indicador de actualización que está asociado a cada resultado preparado que se mantiene en el grupo 5 de la segunda plataforma 4. En general, el indicador de actualización es una medida de si el resultado preparado asociado guardado en el grupo 5 de la segunda plataforma 4 necesita o no actualizarse. El indicador de actualización se define por $(1 - \text{acc}) \cdot t$ (véase la FIG. 1), en el que acc es una probabilidad de que el resultado preparado asociado sea válido y t es una antigüedad del resultado preparado asociado. Antes de pasar a las secuencias de mensajes implicadas en el procesamiento y la respuesta a una petición de datos, primero se explica con más detalle la naturaleza del indicador de actualización.

Un resultado preparado dado guardado en el grupo 5 de la segunda plataforma 4 se puede especificar mediante los parámetros siguientes:

- La antigüedad t_i del resultado preparado i denota el tiempo que ha transcurrido desde la primera preparación o la última actualización (sea cual sea lo último que haya sucedido) del resultado preparado i por medio de la primera plataforma 3. La antigüedad t_i se puede calcular dinámicamente en un punto de tiempo dado al calcular el intervalo de tiempo desde la última preparación del resultado preparado i (tiempo actual menos el tiempo de la última preparación/actualización). Con este fin, se almacena la marca de tiempo de la preparación más reciente del resultado preparado para calcular la antigüedad t_i del resultado preparado i cuando sea necesario.

- La probabilidad acc_i que el resultado preparado i sea válido. Esta probabilidad se puede predecir mediante un modelo probabilístico que modela el comportamiento de la validez de los resultados preparados. Un ejemplo de dicho modelo se basa en una tasa de validez de un resultado preparado i que se describe a continuación.

- La tasa de validez λ_i del resultado preparado i es un indicador de la frecuencia con la que cambia el resultado i preparado por la primera plataforma 3 y, por tanto, de la rapidez con la que el resultado preparado i almacenado en el grupo 5 de la segunda plataforma 4 se vuelve inválido debido a cambios del resultado i provisto por la primera plataforma 3 (p. ej., debido a cambios de los datos guardados en la primera plataforma 3 o por cambios en los datos guardados en otras fuentes de datos a partir de los cuales la primera plataforma 3 calcula el resultado i). Esta tasa de validez λ_i de un resultado preparado i dado se deriva, por ejemplo, estadísticamente de la aparición y los resultados de preparaciones y actualizaciones pasadas y comparaciones del resultado preparado i actualizado con su estado o valores anteriores para determinar si el resultado preparado i con la antigüedad t_i todavía era válido (su estado o valor no cambia con la actualización) o si el resultado preparado i con la antigüedad t_i no era válido (su estado o valor se cambió con la actualización). Por ejemplo, se ha determinado que un resultado preparado particular i tiene una tasa de validez λ_i del 10 % por hora, lo que significa que la probabilidad de que i sea válido disminuye en un 10 % cada hora. En el momento de una actualización del resultado preparado i, i es en general el 100 % válido. Después de una hora, i es válido con una probabilidad del 90 %. Después de dos horas la validez de i es del 81 % (=90 % disminuido en otro 10 %). Después de tres horas, su validez probable es del 72,9 %, y así sucesivamente.

- Como se ha mencionado anteriormente, la tasa de validez λ_i se puede emplear para proporcionar una estimación de la probabilidad de que un resultado preparado siga siendo válido después de un tiempo determinado $P(\text{sin cambios después de } t) = e^{-\lambda_i t}$ lo cual se considera la probabilidad de que un resultado preparado sea válido o, en otras palabras, no esté desactualizado $\text{acc}_i = e^{-\lambda_i t_i}$. En la FIG. 3 se representan dos funciones ejemplares de esta probable exactitud que disminuye con el tiempo. La función 16 representa un resultado preparado que potencialmente sigue siendo más exacto (o, más correctamente, tiene una mayor probabilidad de ser válido a lo largo del tiempo) que otro resultado preparado asociado con la función 17. Por ejemplo, el resultado preparado representado por la función 16 tiene un 70 % de probabilidad de seguir siendo válido 35 horas después de su último recálculo, mientras que el otro resultado preparado caracterizado por la función 17 solo es válido hasta aproximadamente el 50 % 35 horas después de su último recálculo. Las funciones 16 y 17 también pueden representar conjuntos completos de resultados preparados y luego indicar las proporciones de los conjuntos de los resultados preparados que probablemente sean válidos en un tiempo que ha transcurrido desde la última actualización del conjunto.

- La popularidad p_i del resultado preparado i es una frecuencia de acceso promedio a este resultado preparado por parte de los clientes 7 por medio de la interfaz 7. Algunas realizaciones tienen como objetivo lograr una mayor exactitud para estos resultados preparados que se solicitan más a menudo por parte de los clientes 6 que otros resultados preparados.

- El coste de actualización c_i se refiere a los recursos computacionales (tales como la cantidad de tiempo de CPU, pero que también abarca, p. ej., los recursos de red requeridos, p. ej., para recopilar datos de respuesta originales de otras fuentes como se ha mencionado anteriormente) de la primera fuente para actualizar el resultado preparado i . En algunas realizaciones, están disponibles indicaciones de C_i para diferentes resultados preparados (p. ej., de actualizaciones anteriores). En este caso, el parámetro de coste de actualización se puede utilizar para favorecer la actualización de resultados preparados que requieren menos recursos computacionales que otros resultados preparados para actualizar más resultados preparados. De lo contrario, se puede suponer que todos los resultados preparados tienen los mismos costes de actualización. En cualquier caso, en algunas realizaciones, el parámetro de coste de actualización se utiliza para limitar el número de actualizaciones por unidad de tiempo a la cantidad de recursos informáticos disponibles en la primera plataforma 3, es decir, $\sum_i \text{actualizados } c_i \leq \text{recursos por cada unidad de tiempo}$.

Con estos parámetros, la probabilidad de validez de los resultados preparados y almacenados en el grupo 5 de la segunda plataforma 4 se puede definir como sigue:

La probabilidad de validez de todos los resultados preparados guardados en el grupo 5 de la segunda plataforma 4 se puede considerar como la validez media, también denominada "exactitud global", que se define por

$$\text{Exactitud global} = \frac{1}{N} \sum_{i \in \text{grupo}} \text{acc}_i$$

siendo N el número de resultados preparados almacenados en el grupo 5 de la segunda plataforma 4.

A continuación, se presenta el modelo probabilístico ejemplar mencionado anteriormente basado en la tasa de validez λ_i , la exactitud global viene dada por:

$$\text{Exactitud global} = \frac{1}{N} \sum_{i \in \text{grupo}} e^{-\lambda_i t_i}$$

La probabilidad de validez de los resultados preparados en el grupo 5 de la segunda plataforma 4 desde la perspectiva de los clientes 6 (también denominada "exactitud del usuario") se define de manera que cada valor de probabilidad de validez se pondera en función de la popularidad del resultado preparado respectivo. Por tanto, la proporción de accesos de clientes probablemente exactos a los resultados preparados en contraposición a la proporción previstas de resultados preparados exactos es:

$$\text{Exactitud del usuario} = \sum_{i \in \text{grupo}} \frac{p_i}{p_{\text{tot}}} \text{acc}_i = \sum_{i \in \text{grupo}} \frac{p_i}{p_{\text{tot}}} e^{-\lambda_i t_i}$$

en el que P_{tot} se define como la suma de todos los p_i .

El indicador de actualización utilizado por la unidad 2 de control para decidir si se deben devolver los resultados preparados que se mantienen en el grupo 5 de la segunda plataforma 4 o si se deben recuperar los resultados preparados actualizados de la primera plataforma 3 tanto para devolver los resultados actualizados al cliente solicitante 6 como para actualizar los resultados respectivos preparados en el grupo 5 de la segunda plataforma 4 con lo que aumenta la validez de los resultados preparados en la segunda plataforma 4. En la forma más básica, el indicador de actualización descrito en la presente memoria se da por $(1 - \text{acc}) \cdot t_i$ para un resultado preparado particular i por $(1 - \text{acc}_i) \cdot t_i$. En algunas realizaciones, se utilizan definiciones perfeccionadas basadas en esta forma básica, tal como $\frac{(1 - \text{acc}) \cdot t}{c}$, o $\frac{p \cdot (1 - \text{acc}) \cdot t}{c}$. Estas variaciones de los indicadores de actualización y su importancia técnica se explican con más detalle a continuación.

Independientemente de cuál de las definiciones mencionadas anteriormente del indicador de actualización se utilice, la unidad 2 de control emplea el indicador de actualización como sigue (véase la FIG. 2A y la FIG. 2B).

La unidad 2 de control recibe una petición 30 de datos de un cliente 6. A continuación, la unidad 2 de control determina por lo menos un resultado preparado correspondiente a la petición 30 de datos (actividad 31). Por tanto, la unidad 2 de control procesa la petición 30 de datos, determina los criterios de recuperación de datos incluidos en la petición 30 de datos (ya se han dado ejemplos anteriormente) e identifica un número de resultados preparados que cumplen los criterios de recuperación de datos. La forma específica en que la unidad 2 de control identifica estos resultados preparados que cumplen los criterios de recuperación de datos depende de la arquitectura del entorno 1 de base de datos (un ejemplo particular se da más adelante con referencia a las FIG. 4 y 5). Por ejemplo, en el caso de que la petición 30 de datos sea una consulta SQL, la unidad 2 de control determina los parámetros de consulta indicados en

la consulta SQL y recupera la información de identificación que identifica los resultados preparados (p. ej., valores de clave primaria) que cumplen los parámetros de consulta de la consulta SQL desde la segunda plataforma 4 u otro módulo existente dentro del entorno 1 de base de datos para ese propósito.

5 Esta determinación de que los resultados preparados cumplen los criterios de recuperación de datos de la petición 30 de datos no incluye necesariamente todavía una recuperación del contenido de los resultados preparados. Más bien, es suficiente identificar estos resultados preparados, p. ej., mediante identificadores únicos tal como sus valores de clave primaria. En algunas realizaciones, sin embargo, la actividad 31 ya abarca la recuperación del contenido de los resultados preparados que cumplen los criterios de recuperación de datos de la petición 30 de datos desde la segunda plataforma 4.

10 A continuación, la unidad 2 de control compara los indicadores de actualización de los resultados preparados determinados con un valor umbral (actividad 32 en la FIG. 2A y FIG. 2B). A un nivel más detallado, esta actividad 32 implica una determinación de los valores del indicador de actualización de cada uno de los resultados preparados determinados. Con este fin, la unidad 2 de control, p. ej., recupera datos de control respectivos tales como las marcas de tiempo de la última actualización de los resultados preparados determinados, p. ej., de una base de datos de control, calcula las antigüedades t de los resultados preparados determinados, recupera las tasas de validez que están asociadas con los resultados preparados determinados, calcula la probabilidad de validez acc de los resultados preparados determinados y luego calcula $(1 - \text{acc}) \cdot t$ para cada uno de los resultados preparados determinados. Los valores calculados para el indicador de actualización de los resultados determinados se comparan luego con el valor umbral. El valor umbral controla la cantidad de resultados preparados que se actualizan por medio de la primera plataforma 3 en respuesta a las peticiones 30 de recepción de datos y la cantidad complementaria de resultados preparados que no se actualizan en respuesta a las peticiones 30 de recepción de datos, pero que se devuelven al cliente 6 desde el grupo 5 de la segunda plataforma 4.

25 La comparación indica la necesidad de actualizar el resultado preparado respectivo si el valor del indicador de actualización sobrepasa el umbral o la comparación indica que no es necesario actualizar el resultado preparado respectivo si el valor del indicador de actualización no sobrepasa el umbral (por motivos de sencillez, asumimos aquí que la comparación produce las mismas indicaciones para todos los resultados preparados determinados - en la práctica, la comparación también puede indicar que una porción de los resultados preparados determinados se debe actualizar y la porción restante de los resultados preparados determinados no necesita actualizarse; esta situación se explica más adelante con referencia a las FIG. 6A y 6B y también es aplicable a cualquier otro ejemplo descrito en la presente memoria).

30 En el caso de que la comparación indique el requisito de actualizar los resultados preparados (recuadro 33 en la FIG. 2A), la unidad 2 de control recupera versiones actualizadas de los resultados preparados determinados desde la primera plataforma 3 (p. ej., al retransmitir la petición de datos a la primera plataforma 3) (mensaje 34 en la FIG. 2A). La primera plataforma 3 prepara las versiones actualizadas de los resultados preparados y devuelve estas versiones actualizadas de los resultados preparados a la unidad 2 de control. La unidad 2 de control recibe las versiones actualizadas de los resultados preparados desde la primera plataforma 3 mediante el mensaje 35. A continuación, la unidad 2 de control actualiza el resultado de búsqueda preparado determinado en el grupo 5 de la segunda plataforma 4 (actividad 36 en la FIG. 2A). Además, la unidad 2 de control también actualiza los valores del indicador de actualización asociados con los resultados preparados determinados (también actividad 36 en la FIG. 2A) al almacenar la nueva marca de tiempo de esta actualización de los resultados preparados determinados. Finalmente, la unidad 2 de control devuelve la versión actualizada de los resultados preparados determinados al cliente 6 mediante el mensaje 37. Obsérvese que las actividades 36 y 37 también pueden producirse en un orden diferente, es decir, la unidad 2 de control primero devuelve los resultados preparados al cliente 6 y luego actualiza los resultados preparados en la segunda plataforma y los indicadores de actualización asociados.

45 Por otro lado, en el caso de que la comparación 32 no indique el requisito de actualizar los resultados preparados determinados (recuadro 38 en la FIG. 2B), la unidad 2 de control devuelve los resultados preparados determinados al cliente 6 mediante el mensaje 39, sin actualizar los resultados preparados consultando a la primera plataforma 3.

La definición básica anterior del indicador de actualización, dada por $(1 - \text{acc}) \cdot t$, abarca el efecto técnico de proporcionar una validez mejorada de los resultados preparados en el largo plazo, como se explicará a continuación.

50 Los presentes inventores han reconocido que actualizar repetidamente dichos resultados preparados con una tasa de validez λ_i más alta los cuales quedan obsoletos con frecuencia por un cambio de sus datos de respuesta originales subyacentes correspondientes o actualizar dichos resultados preparados que fomentan la exactitud del usuario según se define anteriormente no es óptimo. Una estrategia de este tipo se describe, por ejemplo, en el documento WO 99/22315, que propone actualizar aquellos objetos en una memoria caché con el producto $P_i = \text{Psi}(t) \times \text{Pri}(h)$ más alto, tal como se ha explicado anteriormente al principio. Asimismo, centrarse en actualizar los resultados preparados con

$$\frac{p_i(1 - \text{acc}_i)}{c_i}$$

la mejor relación ganancia-coste, p. ej., definida como $\frac{p_i(1 - \text{acc}_i)}{c_i}$ tampoco es óptimo. Los autores de la presente invención se han dado cuenta de que estas estrategias solo producen aumentos a corto plazo de la exactitud de los resultados preparados, pero podrían provocar una degradación de la exactitud del grupo 5 a largo plazo.

Por otro lado, el indicador de actualización como se propone en la presente memoria proporciona un aumento de la exactitud a largo plazo de los resultados preparados almacenados en la segunda plataforma 4 al establecer una métrica de decisión para responder a una petición de datos ya sea con resultados preparados actualizados por la primera plataforma 3 (que al mismo tiempo actualizan los resultados preparados correspondientes en la segunda plataforma 4) o para responder a una petición de datos ya sea con resultados preparados no actualizados que se mantienen en la segunda plataforma 4 sin consultar a la primera plataforma 3 y actualizar estos resultados preparados en la segunda plataforma 4. El efecto del aumento de la exactitud a largo plazo del indicador de actualización propuesto en la presente memoria será evidente para el experto en la materia a partir de las explicaciones siguientes.

Como se ha explicado previamente, las estrategias de actualización conocidas anteriormente se centran en actualizar los resultados preparados que probablemente no sean válidos, es decir, emplean un indicador de actualización de $(1 - acc_i)$, o versiones perfeccionadas de este tal como $\frac{p_i(1 - acc_i)}{c_i}$. Además, como también se ha explicado anteriormente, los costes de cálculo para actualizar un resultado preparado i por la primera plataforma se denotan como c_i . Estos costes C_i se generan con cada actualización del resultado preparado i por la primera plataforma 3. Por tanto, si, p. ej., el resultado preparado i se actualiza con el doble de frecuencia que otro resultado preparado durante un período de tiempo más largo, la cantidad acumulada de c_i durante este período de tiempo más largo es el doble de los costes acumulados para actualizar el otro resultado preparado durante el mismo período de tiempo. Por lo tanto, los costes de actualización a largo plazo de un resultado preparado i se pueden definir teniendo en cuenta la frecuencia de actualización a largo plazo f_i del resultado preparado i (p. ej., $f_i = 20$ actualizaciones de i en 120 horas = $1/6$ o $f_i = 6$ actualizaciones de i en 120 horas = $1/20$)

$$\frac{p_i(1 - acc_i)}{c_i}$$

Además, como también se ha explicado anteriormente, los costes de cálculo para actualizar un resultado preparado i por la primera plataforma se denotan como c_i . Estos costes C_i se generan con cada actualización del resultado preparado i por la primera plataforma 3. Por tanto, si, p. ej., el resultado preparado i se actualiza con el doble de frecuencia que otro resultado preparado durante un período de tiempo más largo, la cantidad acumulada de c_i durante este período de tiempo más largo es el doble de los costes acumulados para actualizar el otro resultado preparado durante el mismo período de tiempo. Por lo tanto, los costes de actualización a largo plazo de un resultado preparado i se pueden definir teniendo en cuenta la frecuencia de actualización a largo plazo f_i del resultado preparado i (p. ej., $f_i = 20$ actualizaciones de i en 120 horas = $1/6$ o $f_i = 6$ actualizaciones de i en 120 horas = $1/20$)

$$\text{Costes actualiz. a largo plazo}_i = c_i \times f_i$$

Si se tienen en cuenta estos costes de actualización a largo plazo, se llega a la conclusión de que los resultados preparados deben actualizarse de una manera que proporcione la mayor ganancia de exactitud con los menores

$$\frac{1 - acc_i}{c_i f_i} = \frac{1 - acc_j}{c_j f_j}$$

costes de actualización a largo plazo, es decir, con el máximo $\frac{1 - acc_i}{c_i f_i}$. La frecuencia real de actualización a largo plazo f_i es en general desconocida, pero se puede aproximar mediante la consideración siguiente: Si, en respuesta a una petición de datos, se decide actualizar un resultado preparado i dado que no se ha actualizado durante un período de tiempo t_i , se supone que la frecuencia de actualización a largo plazo de este

resultado preparado i es $\frac{1}{t_i}$. Esto en realidad será verdad si se toma efectivamente la decisión de actualizar el resultado preparado. En otras palabras, se considera como una frecuencia potencial de actualización a largo plazo del resultado

$$\frac{c_i}{t_i}$$

preparado i . Por este motivo la proporción $\frac{c_i}{t_i}$ puede utilizarse como una estimación de los costes de actualización a largo plazo de un resultado preparado dado. En general, una optimización de la exactitud a largo plazo del resultado preparado almacenado en el grupo 5 de la segunda plataforma 4 debería, por lo tanto, actualizar los resultados preparados con el máximo

$$\frac{(1 - acc_i)}{c_i} t_i$$

Esta definición del indicador de actualización todavía se puede generalizar al suponer que el coste de actualización C_i es igual y constante para todos los resultados preparados y la restricción de costes de actualización se emplea, p. ej., solo para limitar el número de actualizaciones de resultados preparados por unidad de tiempo por medio de la primera plataforma 3. En este caso, el factor C_i se puede eliminar del indicador de actualización, lo que da como resultado la definición del indicador de actualización como se ha presentado anteriormente, es decir,

$$(1 - acc_i) \cdot t_i$$

Como también se ha indicado brevemente más arriba, son posibles refinamientos adicionales de este indicador de actualización, p. ej., si se tiene en cuenta asimismo la popularidad de los resultados preparados, es decir, la exactitud del usuario (con fines ilustrativos):

$$\frac{p_i(1 - acc_i)}{c_i} t_i$$

$$0$$

$$p_i(1 - acc_i) \cdot t_i$$

Además, acc_i puede sustituirse por el modelo probabilístico específicamente empleado que modela la validez de los resultados preparados que disminuyen con el tiempo, tal como $acc_i = e^{-\lambda_i t_i}$ como se deduce más arriba. Esto da como resultado una versión más específica de las definiciones del indicador de actualización, concretamente

$$(1 - e^{-\lambda_i t_i}) t_i$$

5 o (con fines ilustrativos)

$$\frac{(1 - e^{-\lambda_i t_i})}{c_i} t_i$$

o (con fines ilustrativos)

$$p_i (1 - e^{-\lambda_i t_i}) t_i$$

o (con fines ilustrativos)

$$\frac{p_i (1 - e^{-\lambda_i t_i})}{c_i} t_i$$

10 La FIG. 4 presenta un ejemplo más específico de un entorno 1 de base de datos. Este ejemplo difiere de la arquitectura más general que se muestra en la FIG. 1 en que incluye determinados componentes adicionales, concretamente, el almacén 10 de datos de control y el gestor 12 de modelos predictivos. El almacén 10 de datos de control está conectado a la unidad 2 de control por medio de la interfaz 11 y al gestor de modelos predictivos por medio de la interfaz 15. El gestor de modelos predictivos está conectado a la primera plataforma 3 por medio de la interfaz 13, a la unidad 2 de control por medio de la interfaz 14 y al almacén de datos de control por medio de la interfaz 15.

El almacén 10 de datos de control almacena datos para calcular el indicador de actualización de unos resultados preparados dados almacenados en el grupo 5 de la segunda plataforma 4. En particular, el almacén 10 de datos de control almacena marcas de tiempo que indican la última hora de actualización de los resultados preparados para calcular la antigüedad t_i en un momento dado del tiempo. Opcionalmente, el almacén 10 de datos de control almacena datos para modelar la probabilidad de que los resultados preparados que se mantienen en el grupo 5 de la segunda plataforma 4 sean exactos, por ejemplo, las tasas de validez λ_i de los resultados preparados para calcular $acc_i = e^{-\lambda_i t_i}$ en un momento dado del tiempo. Opcionalmente, el almacén 10 de datos de control almacena también datos de control adicionales utilizados para determinar los valores de los diversos ejemplos del indicador de actualización como se describe anteriormente, p. ej., la tasa de petición (popularidad) p_i de los resultados preparados almacenados en el grupo 5 de la segunda plataforma 4 y/o los costes de cálculo c_i para actualizar los resultados preparados. Los datos de control guardados en el almacén 10 de datos de control están asociados a los resultados preparados almacenados y que se mantienen en el grupo 5 de la segunda plataforma 4. Por tanto, los datos de control almacenados 10 también almacenan información de identificación de los resultados preparados para asociar los datos de control con los resultados preparados. Por ejemplo, un registro de base de datos del almacén 10 de datos de control se define como sigue:

(identificación del resultado preparado i , hora de la última actualización de i , λ_i , p_i , c_i)

La identificación de los resultados preparados utilizados en el almacén 10 de datos de control son, por ejemplo, los valores de clave primaria de los resultados preparados tal como están almacenados en el grupo 5 de la segunda plataforma 4.

En algunas realizaciones, el almacén 10 de datos de control es una parte integrada de la unidad 2 de control. En otras realizaciones, el almacén 10 de datos de control es una parte integrada de la segunda plataforma 4. En este caso, los datos de control mantenidos por el almacén 10 de datos de control se almacenan, p. ej., en una tabla de base de datos que está asociada con la(s) tabla(s) de base de datos de los resultados preparados en el grupo 5. En otras realizaciones más, el almacén de datos de control es un componente autónomo (p. ej., realizado por un sistema de servidor dedicado).

El gestor 12 de modelos predictivos está dispuesto para actualizar determinados datos de control almacenados mediante el almacén de datos de control, tal como las tasas de validez λ_i , la popularidad p_i , y los costes de actualización c_i . Para determinar los valores actuales de estos parámetros, el gestor 12 de modelos predictivos recibe datos de entrada de la primera plataforma 3 y de la unidad 2 de control. Más específicamente, la primera plataforma 3 transmite información sobre los resultados preparados actualizados (las actualizaciones que se producen en respuesta a los mensajes 34 (FIG. 2A)) al gestor 12 de modelos predictivos, es decir, qué resultados se han actualizado en qué momento y si las actualizaciones han dado lugar o no a un contenido diferente de los resultados actualizados. La primera plataforma 3 también transmite información sobre los costes de cálculo para actualizar los resultados

preparados al gestor 12 de modelos (esta información es determinada por la primera plataforma misma al monitorizar las actualizaciones de los diversos resultados preparados y los costes de actualización implicados y/o a partir de información de registro recibida de una fuente original de los resultados preparados tales como los servidores web). La unidad 2 de control transmite información sobre las peticiones 30 de datos entrantes y los resultados preparados solicitados por las peticiones 30 de datos al gestor 12 de modelos predictivos, p. ej., qué resultados preparados se devolvieron a los clientes que realizaron consultas y en qué momentos. En algunas realizaciones, esta comunicación de entrada desde la primera plataforma 3 y/o desde la unidad 2 de control al gestor 12 de modelos predictivos se produce de forma asíncrona a partir del procesamiento y servicio de las peticiones 30 de datos tal como se visualiza en las FIG. 2A, 2B, 5A, 5B, 6A y 6B. En otras realizaciones, la primera plataforma 3 transmite la información sobre un resultado preparado actualizado al gestor 12 de modelos predictivos de forma sincronizada con la actualización, p. ej., inmediatamente después de que la primera plataforma 3 haya devuelto el resultado actualizado a la unidad 2 de control. En dichas realizaciones, la unidad 2 de control también puede transmitir la información sobre las peticiones 30 de datos entrantes y los resultados respectivos devueltos a los clientes 6 al gestor 12 de modelos predictivos de forma sincronizada con el procesamiento de las peticiones 30 de datos, p. ej., inmediatamente después de que la unidad de control 3 haya devuelto los resultados correspondientes al cliente 6 que realizó la consulta.

El gestor 12 de modelos predictivos procesa estas entradas procedentes de la primera plataforma 3 y de la unidad 2 de control y proporciona información de control actual resultante del procesamiento al almacén 10 de datos de control. Más específicamente, el gestor 12 de modelos predictivos utiliza la información sobre los resultados preparados actualizados recibidos de la primera plataforma 3, que abarca por lo menos la información sobre si una actualización de un resultado preparado por la primera plataforma da lugar o no a un cambio del contenido del resultado preparado y una marca de tiempo de la actualización, para mantener las tasas de validez λ_i para cada uno de los resultados preparados almacenados en el grupo 5 de la segunda plataforma 4. El gestor 12 de modelos predictivos utiliza la información sobre los costes computacionales recibida desde la primera plataforma 3 para mantener los parámetros C_i para cada resultado preparado i almacenado en el grupo 5 de la segunda plataforma 4. El gestor 12 de modelos predictivos utiliza la información sobre las peticiones 30 de datos entrantes recibidas desde la unidad 2 de control y los resultados preparados solicitados en consecuencia para mantener los valores de popularidad p_i para cada uno de los resultados preparados almacenados en el grupo 5 de la segunda plataforma 4. Los valores actuales resultantes de λ_i , p_i y/o c_i se transmiten desde el gestor de modelos predictivos al almacén 10 de datos de control y, de este modo, se ponen a disposición de la unidad 2 de control para determinar los valores del indicador de actualización en el curso del procesamiento de las peticiones 30 de datos entrantes (actividad 31). Las actualizaciones de los valores actuales de λ_i , p_i y/o c_i se envían mediante el gestor 12 de modelos predictivos al almacén de datos de control en intervalos regulares y/o bajo demanda, es decir, cuando se han modificado los valores de los datos de control.

En algunas realizaciones, el gestor 12 de modelos predictivos y/o el almacén 10 de datos de control emplean un marco de lotes distribuido que utiliza, p. ej., Apache Hadoop®, lo que permite que ambos componentes ejecuten procesos informáticos intensivos con grandes cantidades de datos (p. ej., del orden de petabytes) en clústeres de servidores.

Las secuencias de mensajes ejemplares que se producen en la arquitectura ejemplar de la FIG. 4 se muestran en las FIG. 5A y 5B. Parecido a las FIG. 2A and 2B, la FIG. 5A muestra la variante en la que los valores indicadores de actualización de los resultados preparados que se devolverán al cliente 6 indican una necesidad de actualizar los resultados preparados y la FIG. 5B muestra la variante en la que los valores del indicador de actualización de los resultados preparados que se devolverán al cliente 6 no indican la necesidad de actualizar los resultados preparados.

En ambas variantes, el proceso comienza con una transmisión de una petición 30 de datos por parte del cliente 6 a la unidad 2 de control. En respuesta a la recepción de la petición 30 de datos, la unidad 2 de control determina información de identificación (p. ej., valores de clave primaria) de los resultados preparados que cumplen los criterios incluidos en la petición 30 de datos y los valores indicadores de actualización asociados (actividad 31A). A modo de ejemplo sencillo, la petición 30 de datos solicita el número de asientos disponibles en el vuelo LH123 del 1 de julio de 2015. En respuesta a la recepción de esta petición 30 de datos, la unidad 2 de control utiliza la clave "LH123-1JUL15" ya incluida en la petición para recuperar el último tiempo de actualización del resultado preparado LH123-1JUL15-42 (es decir, el resultado preparado indica 42 asientos libres, en el que el número 42 de asientos disponibles no es conocido por la unidad 2 de control en esa etapa) y la tasa de validez $\lambda_{LH123-1JUL15}$ a partir del almacén 10 de datos de control. El almacén 10 de datos de control responde con los valores respectivos de la marca de tiempo de actualización y la tasa de validez $\lambda_{LH123-1JUL15}$ (actividad 31B en la FIG. 5A y 5B). A continuación, la unidad 2 de control calcula el indicador de actualización del resultado preparado LH123-1JUL15-42 (actividad 31C en la FIG. 5A y FIG. 5B) al calcular la antigüedad $t_{LH123-1JUL15}$ (marca de tiempo de la última actualización de LH123-1JUL15-42 menos la hora actual) y luego calcular $(1 - e^{-\lambda_i t_i}) t_i$.

La unidad 2 de control realiza entonces la comparación del valor del indicador de actualización calculado con el valor umbral (actividad 32 en la FIG. 5A y la FIG. 5B). Si la comparación indica la necesidad de actualizar el LH123-1JUL15-42 (recuadro 33 en la FIG. 5A), la unidad 2 de control recupera una versión actualizada del resultado preparado desde la primera plataforma 3 (actividad 34 en la FIG. 5A). En una realización, la primera plataforma en sí misma es un inventario original de asientos de vuelo disponibles y devuelve el valor de actualización (p. ej., LH123-1JUL15-23, es decir, 23 asientos libres) a la unidad 2 de control sobre la base de sus propios datos de inventario. En otra realización, la primera plataforma 3 recupera la información de asientos disponibles actuales solicitados a partir de otras fuentes y devuelve el resultado preparado actualizado LH123-1JUL15-23 a la unidad 2 de control posteriormente (actividad

35 en la FIG. 5A). La unidad 2 de control, después de haber recibido el resultado preparado actualizado desde la primera plataforma 3, devuelve el resultado preparado actualizado al cliente 6 (actividad 37 en la FIG. 5A), actualiza el resultado preparado en el grupo 5 de la segunda plataforma 4 (por ejemplo al sustituir el valor obsoleto anterior 42 con el valor actual actualizado 23, actividad 36A) y también actualiza la marca de tiempo de la última actualización de LH123-1JUL15 en el almacén 10 de datos de control (actividad 36B). Obsérvese que las actividades 36A, 36B y 37 pueden producirse en cualquier orden.

De forma alternativa, la comparación entre el valor del indicador de actualización actual de LH123-1JUL15 y el umbral, realizada por la unidad 2 de control (actividad 32), indica que no es necesaria una actualización de LH123-1JUL15 (recuadro 38 en la FIG. 5B). En este caso, la unidad 2 de control recupera el valor de LH123-1JUL15 (p. ej., 42 asientos disponibles) del grupo 5 de la segunda plataforma 4 (actividad 40A en la FIG. 5B) y devuelve este valor como respuesta a la petición 30 de datos al cliente 6 (actividad 39).

En otras realizaciones, dependiendo del contenido y la definición de los resultados preparados y los criterios incluidos en la petición 30 de datos, la actividad 31 es más compleja (FIG. 6A y 6B), p. ej., porque los valores clave de los resultados preparados en los que está interesado el cliente 6 no están incluidos en la petición 30 de datos y, por tanto, puede requerirse una búsqueda para determinar qué resultados preparados cumplen con los criterios de búsqueda incluidos en la petición 30 de datos. Por ejemplo, los resultados preparados y almacenados por el grupo 5 especifican recomendaciones de precios de viajes calculados previamente y la petición 30 de datos, p. ej., solicita los viajes más baratos entre Frankfurt y Boston con una salida entre el 1 de julio de 2015 y el 10 de julio de 2015 y una duración de estadía entre 11 y 15 días. En escenarios de este tipo, en algunas realizaciones, la unidad 2 de control identifica primero resultados preparados almacenados en el grupo 5 de la segunda plataforma 4 que cumplen los criterios de búsqueda de la petición 30 de datos. Por tanto, en estas realizaciones, la actividad 31A (FIG. 6A) es una consulta de la unidad 2 de control a la segunda plataforma 4 para recuperar información de identificación de resultados preparados que cumplen con los criterios de búsqueda incluidos en la petición 30 de datos. En referencia al último ejemplo, la segunda plataforma 4, p. ej., devuelve todas las recomendaciones de viaje almacenadas en el grupo 5 para los viajes entre Frankfurt y Boston con una salida entre el 1 de julio de 2015 y el 10 de julio de 2015 y una duración de estancia entre 11 y 15 días (actividad 31B en la FIG. 6A).

La unidad 2 de control determina entonces los valores del indicador de actualización para todos los resultados preparados identificados de la manera descrita anteriormente (recupera los datos de control respectivos del almacén 10 de datos de control mediante las actividades 31C y 31D en la FIG. 6A y calcula los valores del indicador de actualización mediante la actividad 31E en la FIG. 6A) y realiza la comparación entre el valor del indicador de actualización y el umbral para cada resultado preparado identificado (actividad 32 en la FIG. 6A). En todos los resultados preparados para los cuales la comparación indica una necesidad de actualizar el resultado preparado respectivo (recuadro 33 en la FIG. 6A), la unidad 2 de control recupera una versión actualizada del resultado preparado respectivo de la primera plataforma 3 (actividades 34 y 35 en la FIG. 6B). Opcionalmente, en lo que respecta a los resultados preparados restantes para los cuales la comparación entre el indicador de actualización y el valor umbral no indica una necesidad de actualización, la unidad 2 de control recupera el contenido de estos resultados preparados que no se van a actualizar de la segunda plataforma 4 (actividades 40A y 40B en la FIG. 6B; esto se aplica si la unidad 2 de control no ha recuperado ya de antemano el contenido de todos los resultados preparados que cumplen los criterios de búsqueda con las actividades 31A y 31B). Después de haber recibido las versiones actualizadas de los resultados preparados en respuesta a la petición 34 (actividad 35), la unidad 2 de control almacena los resultados preparados actualizados en el grupo 5 de la segunda plataforma 4 (actividad 36A en la FIG. 6B), actualiza las marcas de tiempo de actualización de los resultados preparados actualizados en el almacén 10 de datos de control (actividades 36B en la FIG. 6B), opcionalmente posprocesa los resultados preparados recuperados a partir de la segunda plataforma 4 y las versiones actualizadas de los resultados preparados recuperados a partir de la primera plataforma 3 (actividad 41 en la FIG. 6B), que es, p. ej., una determinación de un subgrupo de resultados preparados de los resultados preparados actualizados y los resultados preparados no actualizados determinados antes de devolverse al cliente 6 (en el ejemplo anterior, seleccionar la recomendación de viaje más económica), y devuelve estos resultados al cliente 6 (actividad 37 en la FIG. 6B).

En algunas realizaciones, se emplean flujos de mensajes aún diferentes, dependiendo de las características particulares de los modelos de base de datos empleados por el entorno 1 de base de datos y la estructura y los criterios incluidos en las peticiones 30 de datos. Por ejemplo, en algunas realizaciones, no se emplea el posprocesamiento de los resultados preparados (actividad 41 en la FIG. 6A), sino que todos los resultados preparados recuperados a partir de la segunda plataforma 4 y todas las versiones actualizadas de los resultados preparados recuperados a partir de la primera plataforma 3 se devuelven al cliente 6. En dichas realizaciones, los resultados preparados recuperados a partir de la segunda plataforma 4 ya se devuelven al cliente 6 tan pronto como están disponibles (actividad 39 en la FIG. 6B), lo que en general es antes de que concluya la recuperación de las versiones actualizadas de los resultados preparados de la primera plataforma 3 (porque la actualización de los resultados preparados por la primera plataforma en general requiere sustancialmente más tiempo que recuperar los resultados preparados de la segunda plataforma 4, como se ha explicado en detalle anteriormente). Así pues, en esta situación, la actividad 37 de la FIG. 6B entonces solo abarca la devolución de las versiones actualizadas de los resultados preparados recuperados a partir de la primera plataforma 3.

En algunas realizaciones, el flujo de mensajes que se muestra en las FIG. 6A y 6B también está adaptado porque la actividad 41 de posprocesamiento se realiza ya en una etapa anterior, por ejemplo, mediante la unidad 2 de control después de haber recibido el contenido de los resultados preparados con la actividad 31B o mediante la segunda plataforma 4 antes de devolver los resultados preparados (o por lo menos sus identidades) con la actividad 31B. De nuevo en referencia al ejemplo anterior de devolución de la recomendación de viaje más barata para los viajes entre Frankfurt y Boston con una salida entre el 1 de julio de 2015 y el 10 de julio de 2015 y una duración de estadía entre 11 y 15 días, en algunas realizaciones, la segunda plataforma 4 está dispuesta para determinar un número limitado de las recomendaciones de viaje más económicas para los viajes entre Frankfurt y Boston con una salida entre el 1 de julio de 2015 y el 10 de julio de 2015 y una duración de estadía entre 11 y 15 días (a diferencia de devolver todas las recomendaciones de precios de viajes para los viajes entre Frankfurt y Boston con una salida entre el 1 de julio de 2015 y el 10 de julio de 2015 y una duración de estadía entre 11 y 15 días almacenadas en el grupo 5). Por ejemplo, con la actividad 31B en la FIG. 6B, la segunda plataforma 4 devuelve solo las cinco recomendaciones de viaje más baratas. Las actividades posteriores de la unidad 2 de control (actividades 31C, 31D, 31E, 32-38) se limitan entonces únicamente a estos cinco resultados preparados recuperados y se realizan de la misma forma que se describe anteriormente.

En algunas realizaciones, el valor umbral se adapta dinámicamente dependiendo de la carga de la primera plataforma 3. De esta manera, mediante el uso del valor umbral adaptable, se controla la cantidad de resultados preparados que se actualizan por medio de la primera plataforma 3 y, por lo tanto, la carga de la primera plataforma. Si la carga de la primera plataforma 3 es demasiado alta (se actualizan demasiados resultados preparados por medio de la primera plataforma 3 en una unidad de tiempo dada) y se debe reducir la carga, se aumenta el valor umbral con el efecto de que se reduce la porción de los resultados preparados que se actualizan por medio de la primera plataforma 3 y se aumenta la porción de los resultados preparados que se recuperan a partir de la segunda plataforma 4 y se devuelven al cliente 6 sin actualizar. Si, por otro lado, la primera plataforma 3 tiene recursos libres que pueden utilizarse para actualizar más resultados preparados en una unidad de tiempo dada que los actualizados actualmente, el valor umbral se reduce con el efecto de que se incrementa la porción de los resultados preparados que se actualizan por medio de la primera plataforma 3 y se reduce la porción de los resultados preparados que se recuperan a partir de la segunda plataforma 4 y se devuelven al cliente 6 sin actualizar.

En algunas realizaciones, la adaptación dinámica del valor umbral se logra manteniendo un bucle 20 de control de realimentación entre la primera plataforma 3 y la unidad 2 de control (FIG. 7). El bucle 20 de control de realimentación está definido por tres variables, concretamente, el valor 21 de control, el punto de ajuste 18 y la variable 19 de accionamiento. El valor 21 de control es la variable a controlar (a mantener constante en un nivel objetivo determinado) y, por lo tanto, está definido por un factor de carga real actual de la primera plataforma 3. El punto de ajuste 18 del bucle 20 de control de realimentación viene dado por un factor de carga objetivo deseado de la primera plataforma 3. La variable de accionamiento del bucle 20 de control de realimentación es el valor umbral que incide en el factor de carga real de la primera plataforma, es decir, controla el valor 21 de control.

En algunas realizaciones, el bucle 20 de control de realimentación se realiza mediante la primera plataforma 3 y la unidad 2 de control que se comunican entre sí a intervalos regulares como sigue. La primera plataforma compara periódicamente el factor 21 de carga real actual con el factor 18 de carga objetivo deseado. Si el factor 21 de carga real actual es el N % mayor que el factor 18 de carga objetivo deseado (p. ej., 10 %), el valor umbral se incrementa en una relación N/B % (p. ej., 10 % dividido por 2 = 5%), en el que B es un suavizado que evita reacciones exageradas de la adaptación del valor umbral dinámico y permite una convergencia suave entre la carga 21 actual y la carga 18 objetivo. Como resultado de la disminución del valor umbral, el factor 21 de carga real actual disminuirá. Si, por otro lado, el factor 21 de carga real actual es inferior al factor 18 de carga objetivo deseado en el N %, el valor umbral se reduce en el N/B %, lo que da como resultado una mayor cantidad de resultados preparados que se actualizan con la primera plataforma 3 y, de este modo, aumenta el factor 21 de carga real actual.

Finalmente, la FIG. 8 es una representación esquemática de un sistema informático que proporciona la funcionalidad de la unidad 2 de control. Dentro de la unidad 2 de control, se ejecutan un conjunto de instrucciones para hacer que el sistema informático realice cualquiera de los procedimientos analizados en la presente memoria. La unidad 2 de control incluye un procesador 81, una memoria 82 principal y un dispositivo 83 de interfaz de red, que se comunican entre sí por medio de un bus 84. Opcionalmente, la unidad 2 de control incluye además una memoria 85 estática y una unidad 86 de disco. Un visualizador 87 de vídeo, un dispositivo 88 de entrada alfanumérica y un dispositivo 89 de control del cursor pueden formar una interfaz de usuario para que, p. ej., un gestor controle la unidad 2 de control.

El dispositivo 83 de interfaz de red conecta la unidad 2 de control a la primera plataforma 3 por medio de la interfaz 9 y a la segunda plataforma 4 por medio de la interfaz 8. El dispositivo 83 de interfaz de red también conecta la unidad 2 de control a los clientes 6 por medio de la interfaz 7. Un conjunto de instrucciones (es decir, software) 90 que incorpora cualquiera o todos los procedimientos descritos anteriormente reside completamente, o por lo menos parcialmente, en o sobre un medio legible por máquina, p. ej., la memoria 82 principal y/o el procesador 81. Un medio legible por máquina en el que reside el software 90 también puede ser un soporte de datos no volátil (p. ej., un disco duro magnético no extraíble o un disco extraíble óptico o magnético) que forma parte de la unidad 86 de disco. El software 90 puede además transmitirse o recibirse como una señal propagada, p. ej., por medio de Internet o cualquier otra red a través del dispositivo 83 de interfaz de red. El funcionamiento básico de la unidad 2 de control, incluida la interfaz de usuario y la comunicación de red, está controlado por el sistema 91 operativo.

- 5 El presente planteamiento para controlar el encaminamiento de las peticiones de datos entrantes mediante la unidad de control a la primera plataforma y/o a la segunda plataforma a partir del indicador de actualización $(1 - \text{acc}) \cdot t$ que se compara con el valor umbral proporciona una validez mejorada a largo plazo de los resultados preparados que se mantienen en el grupo de la segunda plataforma. Este planteamiento permite determinar de forma óptima qué resultados preparados que se devolverán en respuesta a las peticiones de datos se recuperarán de la primera plataforma y, por lo tanto, se actualizarán en la segunda plataforma. De esta manera, se evita centrarse ineficazmente en actualizar repetidamente resultados preparados muy volátiles, con lo que mejora la exactitud y validez de los resultados preparados a largo plazo.
- 10 Aunque en la presente memoria se han descrito determinados productos y procedimientos contruidos según las enseñanzas de la invención, el alcance de cobertura de esta patente no se limita a ellos. Por el contrario, esta patente cubre todas las realizaciones de las enseñanzas de la invención que se encuentran razonablemente dentro del alcance de las reivindicaciones adjuntas, ya sea literalmente o según la doctrina de los equivalentes.

REIVINDICACIONES

1. Un procedimiento para manejar peticiones de datos dirigidas a un entorno de base de datos, comprendiendo el entorno (1) de base de datos por lo menos una primera plataforma (3) que proporciona resultados originales para almacenarse en una segunda plataforma (4) como resultados preparados, manteniendo la segunda plataforma (4) un grupo de los resultados preparados que se devolverán a las peticiones de datos, y una unidad (2) de control para procesar las peticiones de datos dirigidas al entorno (1) de base de datos, en el que cada resultado preparado que se mantiene en el grupo de la segunda plataforma (4) está asociado con un indicador de actualización que es una medida de que el resultado preparado asociado que se mantiene en el grupo de la segunda plataforma (4) debe actualizarse, comprendiendo el procedimiento:
- 5 - recibir, mediante la unidad (2) de control, una petición de datos;
- determinar, mediante la unidad (2) de control, por lo menos un resultado preparado correspondiente a la petición de datos;
- comparar, mediante la unidad (2) de control, el indicador de actualización del resultado preparado determinado correspondiente a la petición de datos con un valor umbral;
- 15 - en respuesta a la determinación mediante la unidad (2) de control de que la comparación indica el requisito de actualizar el resultado preparado,
- recuperar, mediante la unidad (2) de control de la primera plataforma (3), una versión actualizada del por lo menos un resultado preparado;
- 20 → actualizar, mediante la unidad (2) de control, el resultado preparado en el grupo de la segunda plataforma (4) y el indicador de actualización asociado en función de la versión actualizada del por lo menos un resultado preparado al almacenar una nueva marca de tiempo de la actualización del por lo menos un resultado preparado; y
- devolver, mediante la unidad (2) de control, la versión actualizada del por lo menos un resultado,
- en respuesta a la determinación mediante la unidad (2) de control de que la comparación no indica el requisito de actualizar el por lo menos un resultado preparado,
- 25 → la unidad (2) de control que devuelve el por lo menos un resultado preparado determinado;
- en el que el indicador de actualización está definido por $(1 - acc) \cdot t$, en el que acc es una probabilidad de que el resultado preparado asociado sea válido y t es una antigüedad del resultado preparado asociado, en el que $acc = e^{-\lambda t}$, en el que la tasa de validez λ es un indicador de la frecuencia con la que cambia el resultado preparado.
2. El procedimiento de la reivindicación 1, que además comprende
- 30 → consultar, mediante la unidad (2) de control, un almacén de datos de control según la información clave que identifica el por lo menos un resultado preparado;
- recibir, mediante la unidad (2) de control del almacén de datos de control, datos de control que especifican el indicador de actualización asociado con el por lo menos un resultado preparado.
3. El procedimiento de la reivindicación 2, en el que la determinación del por lo menos un resultado preparado comprende
- 35 → consultar, mediante la unidad (2) de control, a la segunda plataforma el por lo menos un resultado preparado determinado que cumple los criterios especificados por la petición de datos;
- recibir, mediante la unidad (2) de control, el por lo menos un resultado preparado a partir del grupo de resultados preparados de la segunda plataforma (4).
4. El procedimiento de la reivindicación 1, en el que determinar el por lo menos un resultado preparado correspondiente a la petición de datos comprende:
- 40 → retransmitir, mediante la unidad (2) de control, la petición de datos a la segunda plataforma (4); y
- recibir, mediante la unidad (2) de control de la segunda plataforma (4), el por lo menos un resultado preparado que cumple los criterios especificados por la petición de datos;
- 45 en el que el procedimiento además comprende:
- recuperar, mediante la unidad (2) de control, el indicador de actualización asociado al resultado preparado recibido.

5. El procedimiento de cualquiera de las reivindicaciones 1 a 4, en el que el por lo menos un resultado preparado correspondiente a la petición de datos comprende por lo menos dos resultados preparados, un primer resultado preparado y un segundo resultado preparado, comprendiendo además el procedimiento:

→ determinar, mediante la unidad (2) de control, que la comparación indica el requisito de actualizar el primer resultado preparado, pero no indica el requisito de actualizar el segundo resultado preparado,

→ la unidad (2) de control recupera una versión actualizada del primer resultado preparado de la primera plataforma (3) y actualiza el primer resultado preparado en el grupo de la segunda plataforma (4) y el indicador de actualización asociado basado en la versión actualizada recuperada del primer resultado preparado; y

→ devolver la versión actualizada del primer resultado preparado y del segundo resultado preparado.

6. Una unidad (2) de control o de manejo de peticiones de datos dirigidas a un entorno (1) de base de datos, estando acoplada la unidad (2) de control a una primera plataforma (3) que proporciona resultados originales que se almacenarán en una segunda plataforma (4) como resultados preparados, manteniendo la segunda plataforma (4) un grupo de los resultados preparados que se devolverán a las peticiones de datos, en el que cada resultado preparado mantenido por la segunda plataforma (4) está asociado con un indicador de actualización que es una medida de que el resultado preparado asociado que se mantiene en la segunda plataforma (4) se actualizará, estando dispuesta la unidad (2) de control para:

→ recibir una petición de datos;

→ determinar por lo menos un resultado preparado correspondiente a la petición de datos;

→ comparar el indicador de actualización del resultado preparado determinado correspondiente a la petición de datos con un valor umbral;

→ en respuesta a la determinación de que la comparación indica el requisito de actualizar el resultado preparado,

→ recuperar una versión actualizada del por lo menos un resultado preparado a partir de los datos de respuesta originales de la primera plataforma (3);

→ actualizar el resultado preparado en el grupo de la segunda plataforma (4) y el indicador de actualización asociado en función de la versión actualizada del por lo menos un resultado preparado al almacenar una nueva marca de tiempo de la actualización del por lo menos un resultado preparado; y

→ devolver la versión actualizada de por lo menos un resultado;

- en respuesta a la determinación de que la comparación no indica el requisito de actualizar por lo menos un resultado preparado,

→ devolver por lo menos un resultado preparado determinado;

en el que el indicador de actualización se define por $(1 - \text{acc}) \cdot t$, en el que acc es una probabilidad de que el resultado preparado asociado sea válido y t es una antigüedad del resultado preparado asociado, en el que $\text{acc} = e^{-\lambda t}$, en el que la tasa de validez λ es un indicador de la frecuencia con la que cambia el resultado preparado.

7. La unidad (2) de control de la reivindicación 6 está dispuesta para realizar el procedimiento de cualquiera de las reivindicaciones 2 a 5.

8. Un programa informático para su ejecución por una unidad (2) de control para manejar peticiones de datos dirigidas a un entorno (1) de base de datos, estando acoplada la unidad (2) de control a una primera plataforma (3) que proporciona resultados originales que se van a almacenar en una segunda plataforma (4) como resultados preparados, manteniendo la segunda plataforma (4) un grupo de los resultados preparados que se devolverán a las peticiones de datos, en el que cada resultado preparado mantenido por la segunda plataforma (4) está asociado a un indicador de actualización que es una medida de que el resultado preparado asociado que se mantiene en la segunda plataforma (4) se va a actualizar, estando configurado el programa informático para -cuando se ejecuta mediante la unidad (2) de control- hacer que la unidad de control:

- reciba una petición de datos;

- determine por lo menos un resultado preparado correspondiente a la petición de datos;

- compare el indicador de actualización del resultado preparado determinado correspondiente a la petición de datos con un valor umbral;

- en respuesta a la determinación de que la comparación indica el requisito de actualizar el resultado preparado,

→ recuperar una versión actualizada del por lo menos un resultado preparado a partir de los datos de respuesta originales de la primera plataforma (3);

5 → actualizar el resultado preparado en el grupo de la segunda plataforma (4) y el indicador de actualización asociado en función de la versión actualizada del por lo menos un resultado preparado al almacenar una nueva marca de tiempo de la actualización del por lo menos un resultado preparado; y

→ devolver la versión actualizada de por lo menos un resultado;

- en respuesta a la determinación de que la comparación no indica el requisito de actualizar por lo menos un resultado preparado,

→ devolver por lo menos un resultado preparado determinado;

10 en el que el indicador de actualización se define por $(1 - acc) \cdot t$, en el que acc es una probabilidad de que el resultado preparado asociado sea válido y t es una antigüedad del resultado preparado asociado, en el que $acc = e^{-\lambda t}$, en el que la tasa de validez λ es un indicador de la frecuencia con la que cambia el resultado preparado.

9. El programa informático de la reivindicación 8 está configurado para -cuando se ejecuta mediante la unidad (2) de control- hacer que la unidad de control realice el procedimiento de cualquiera de las reivindicaciones 2 a 5.

15

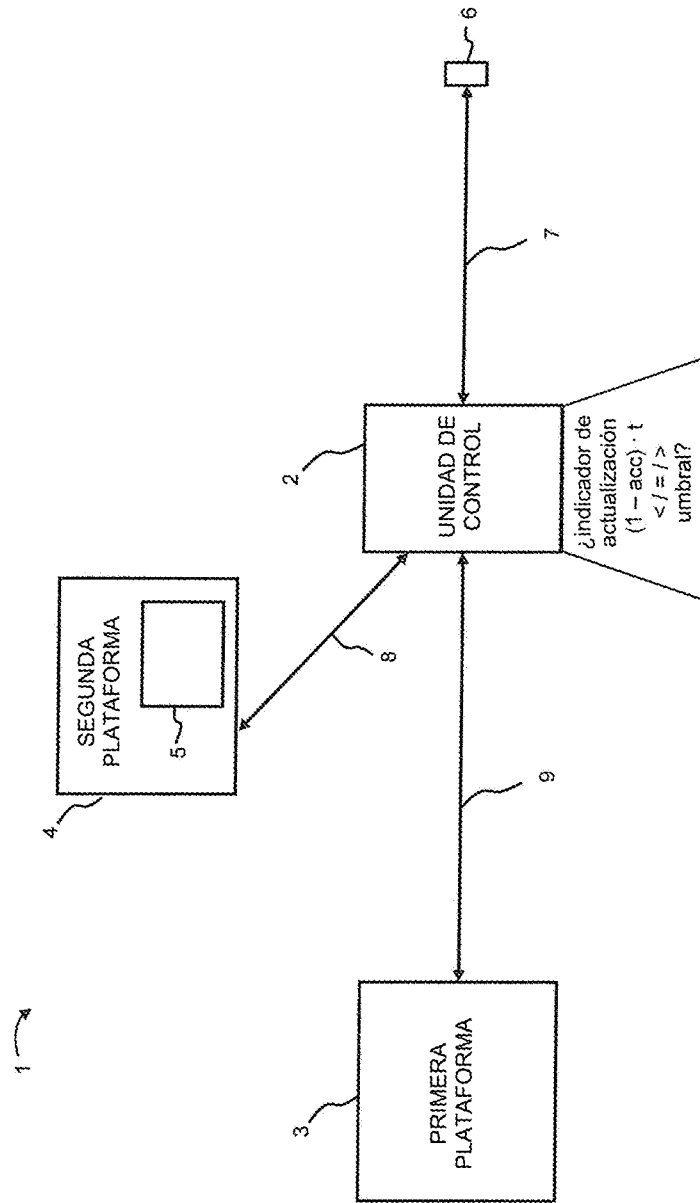


FIG. 1

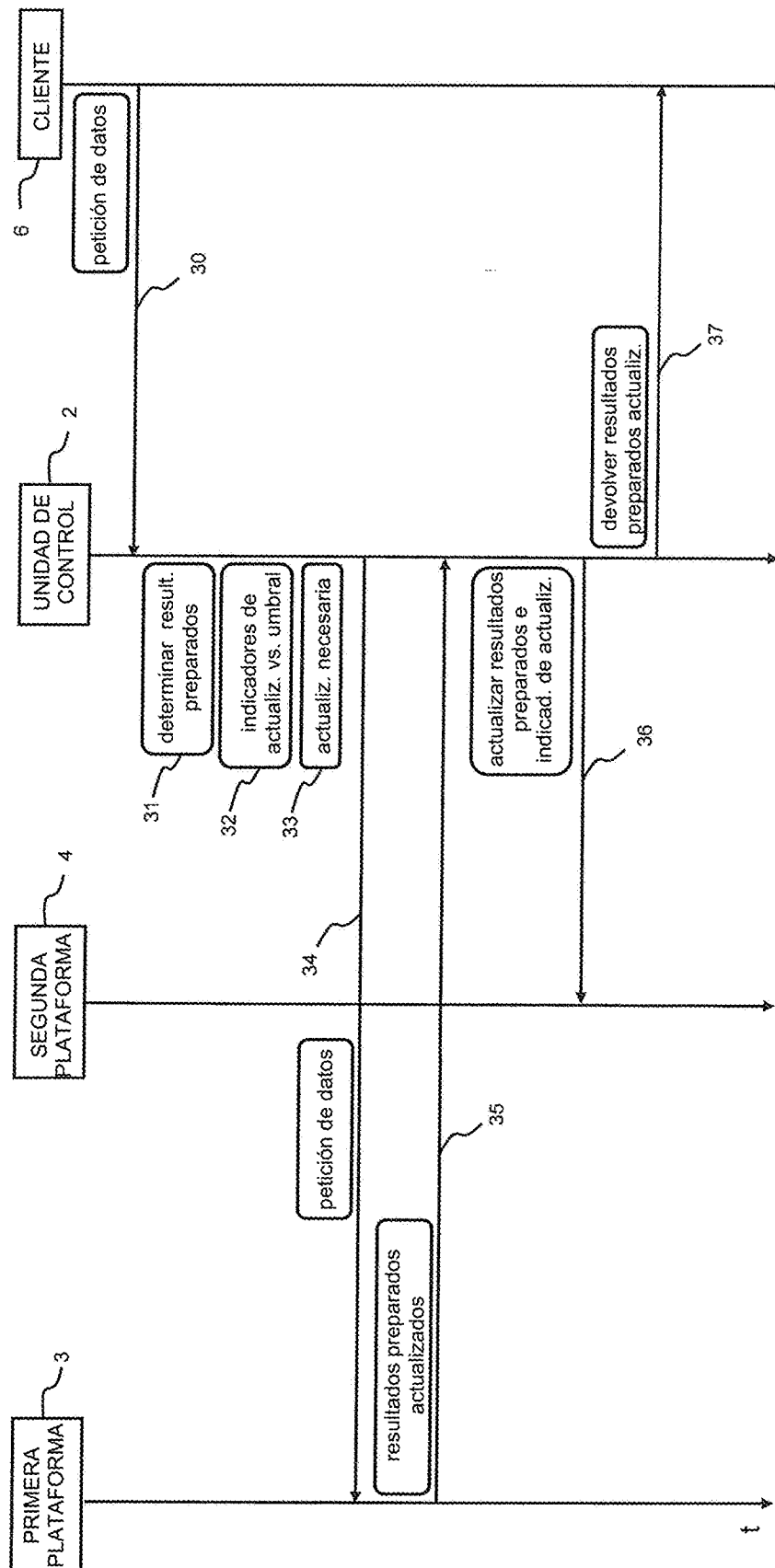


FIG. 2A

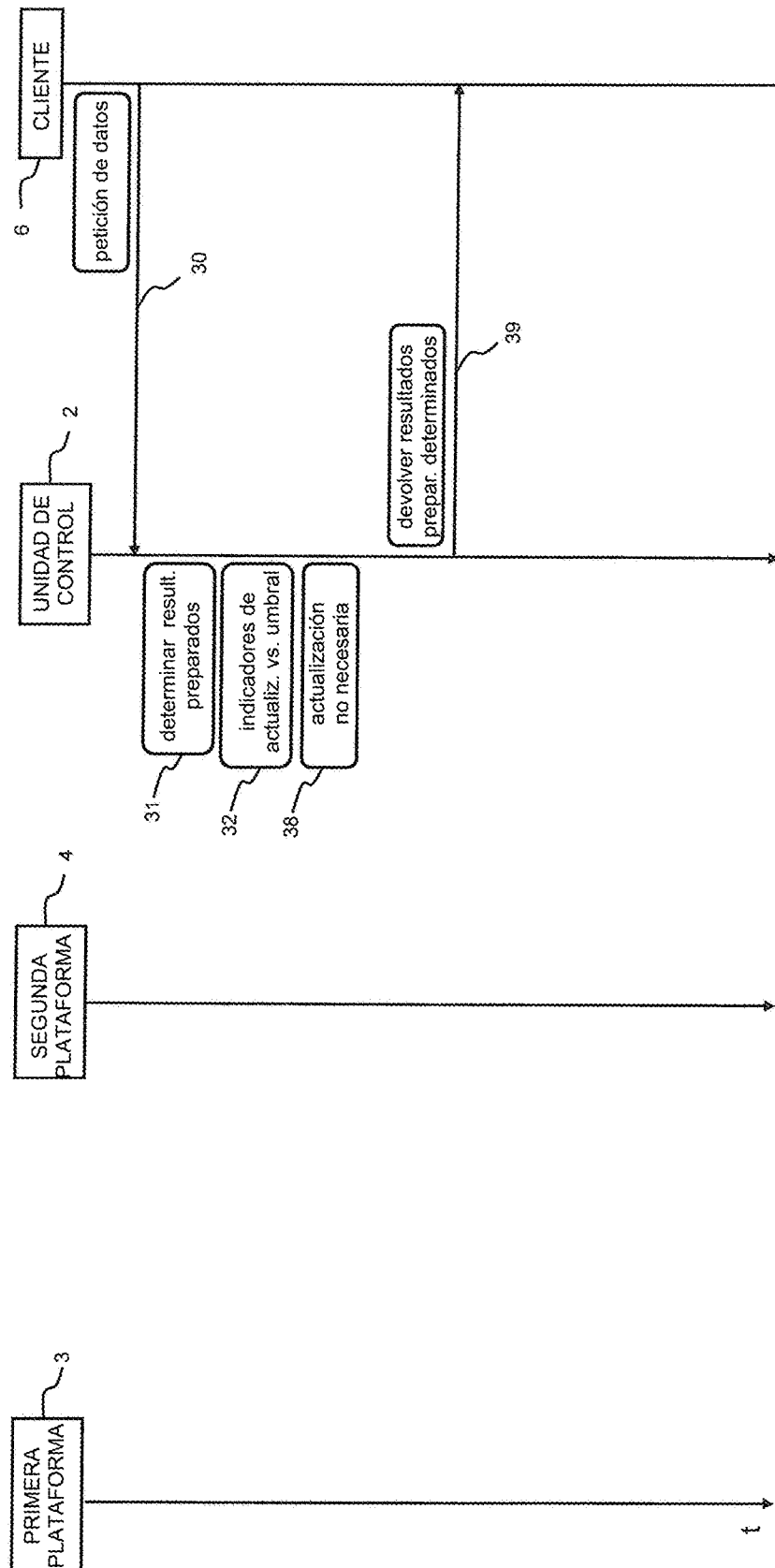


FIG. 2B

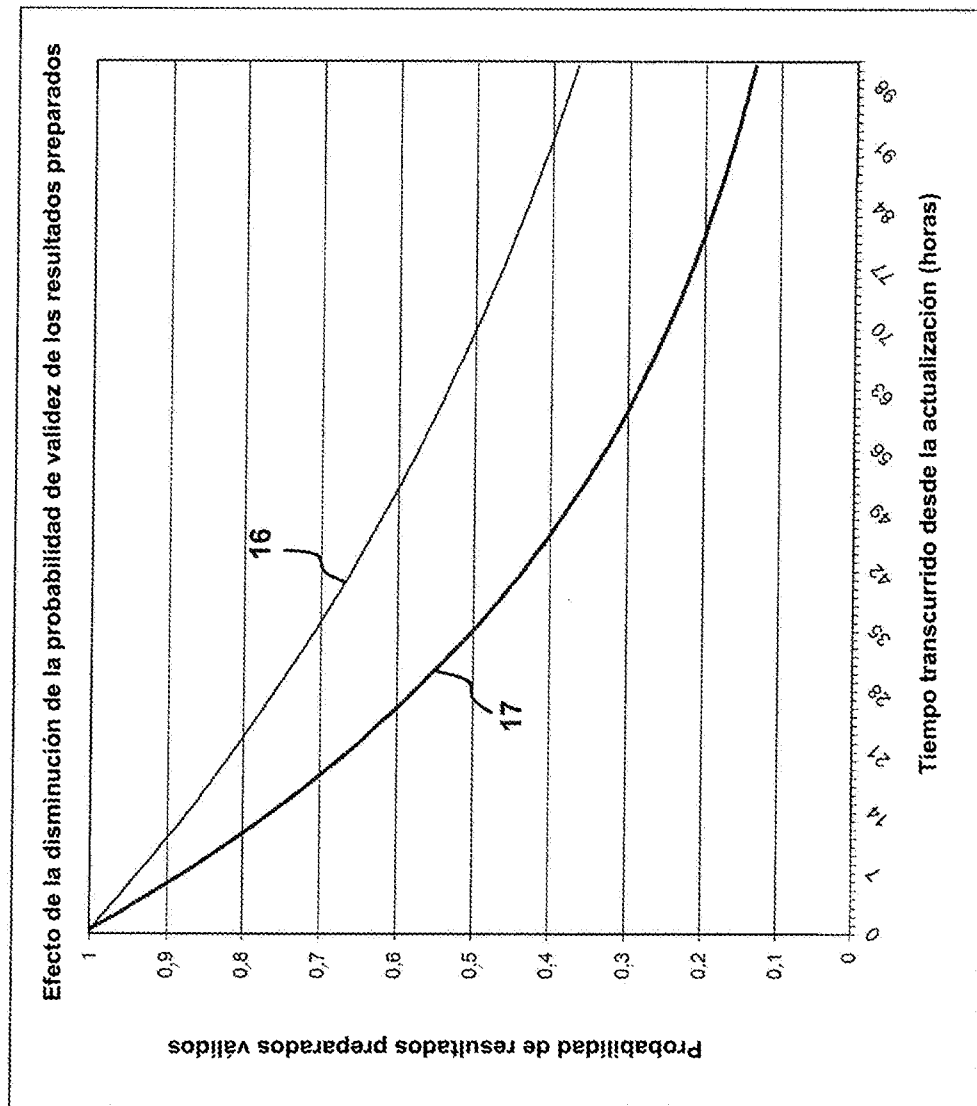


Fig. 3

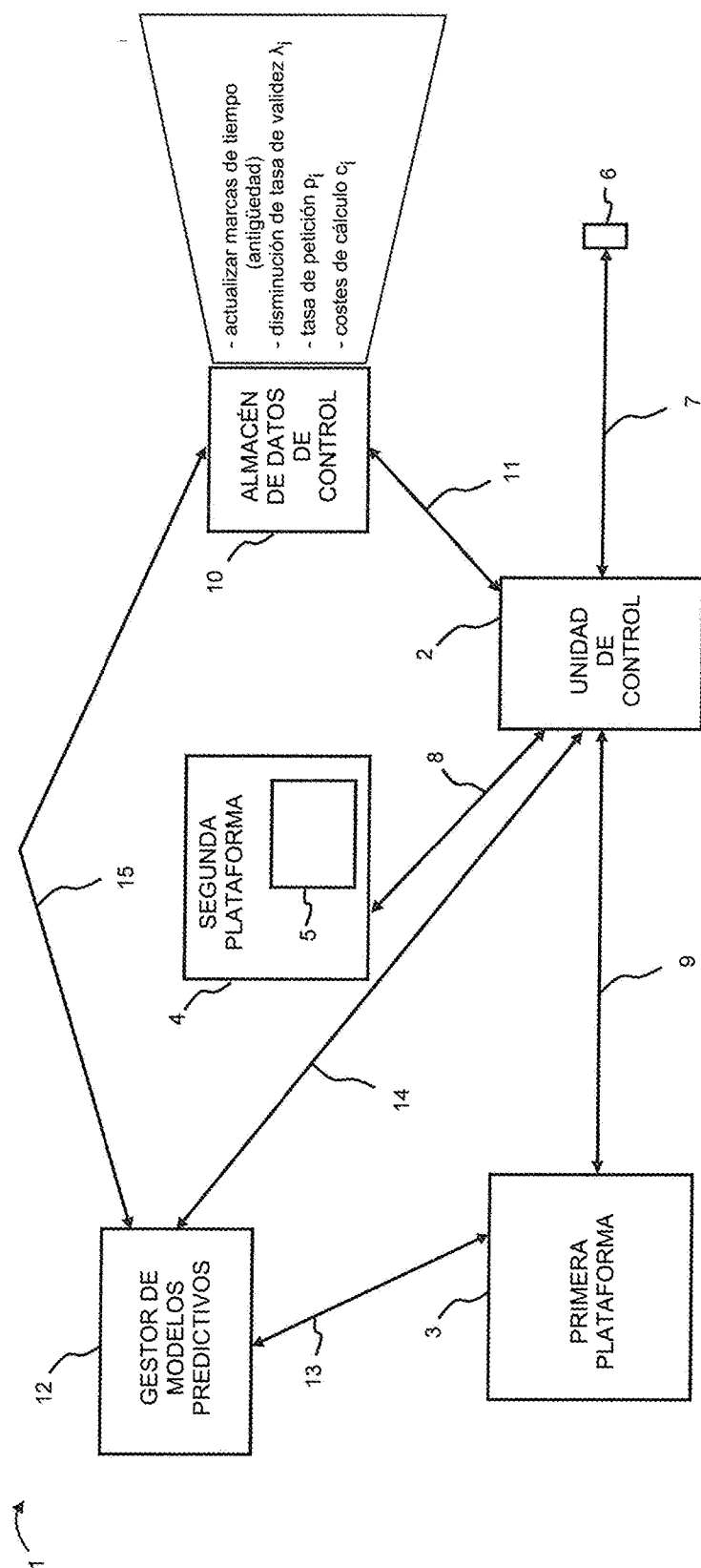


FIG. 4

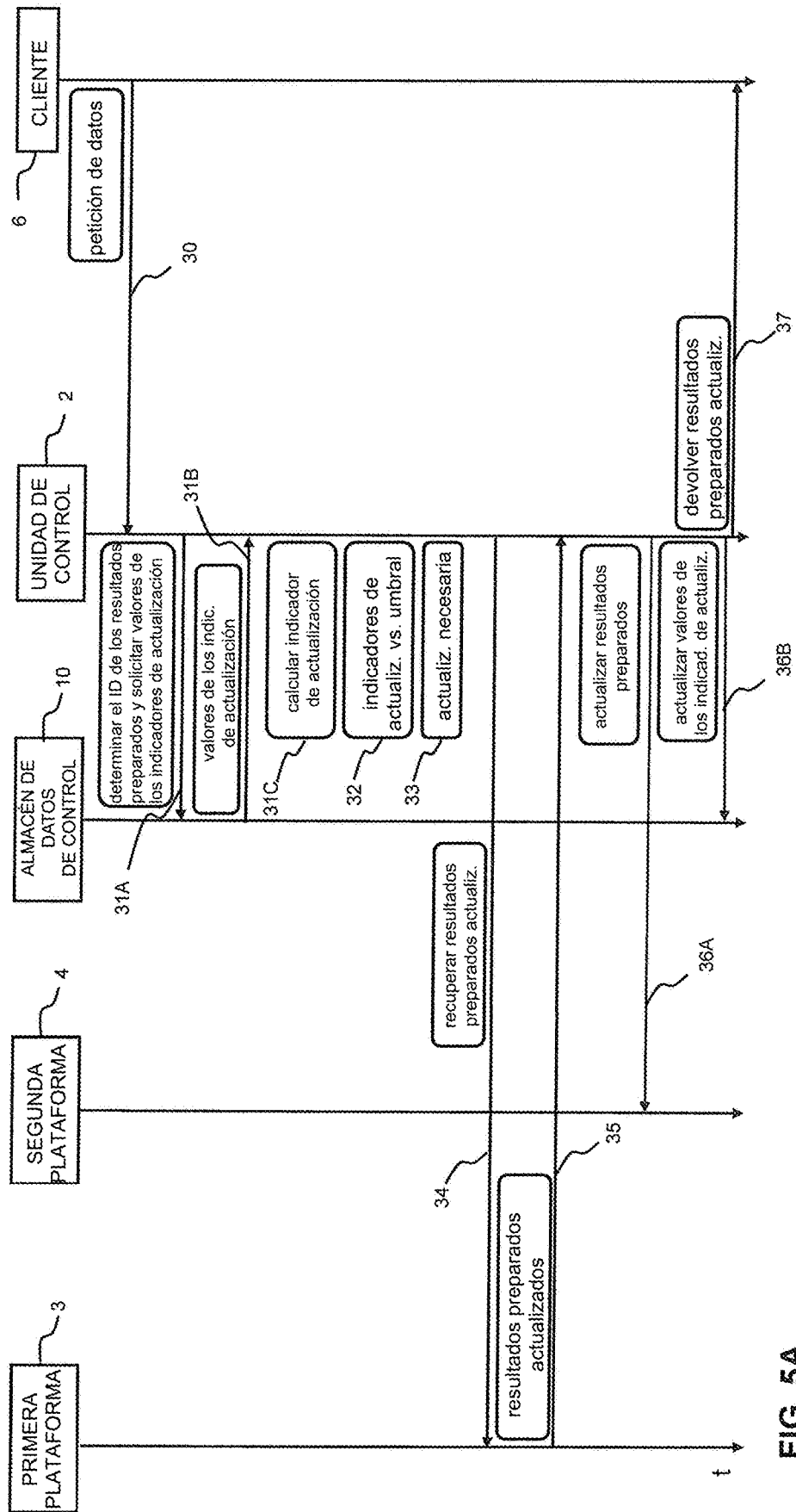
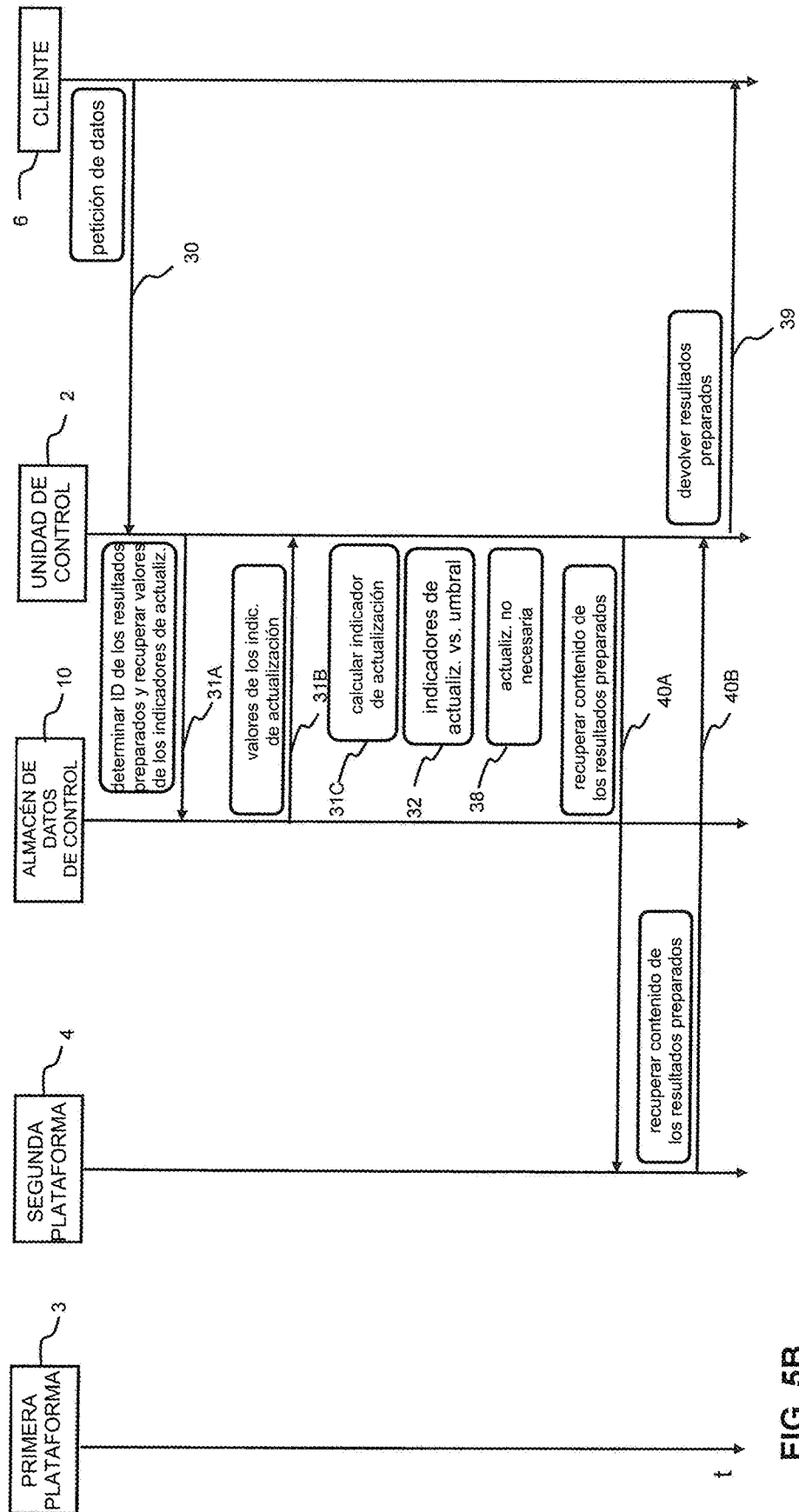


FIG. 5A



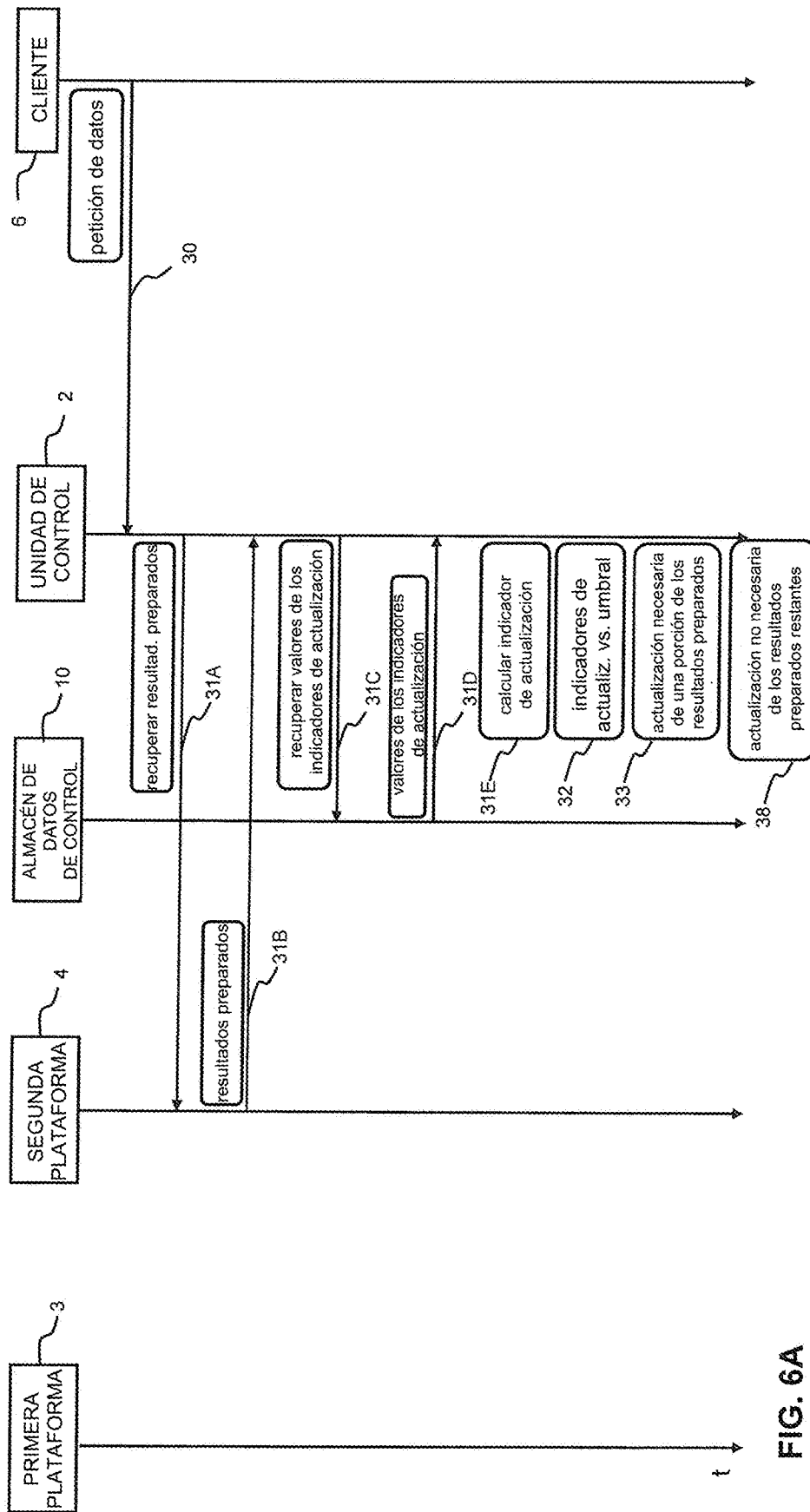


FIG. 6A

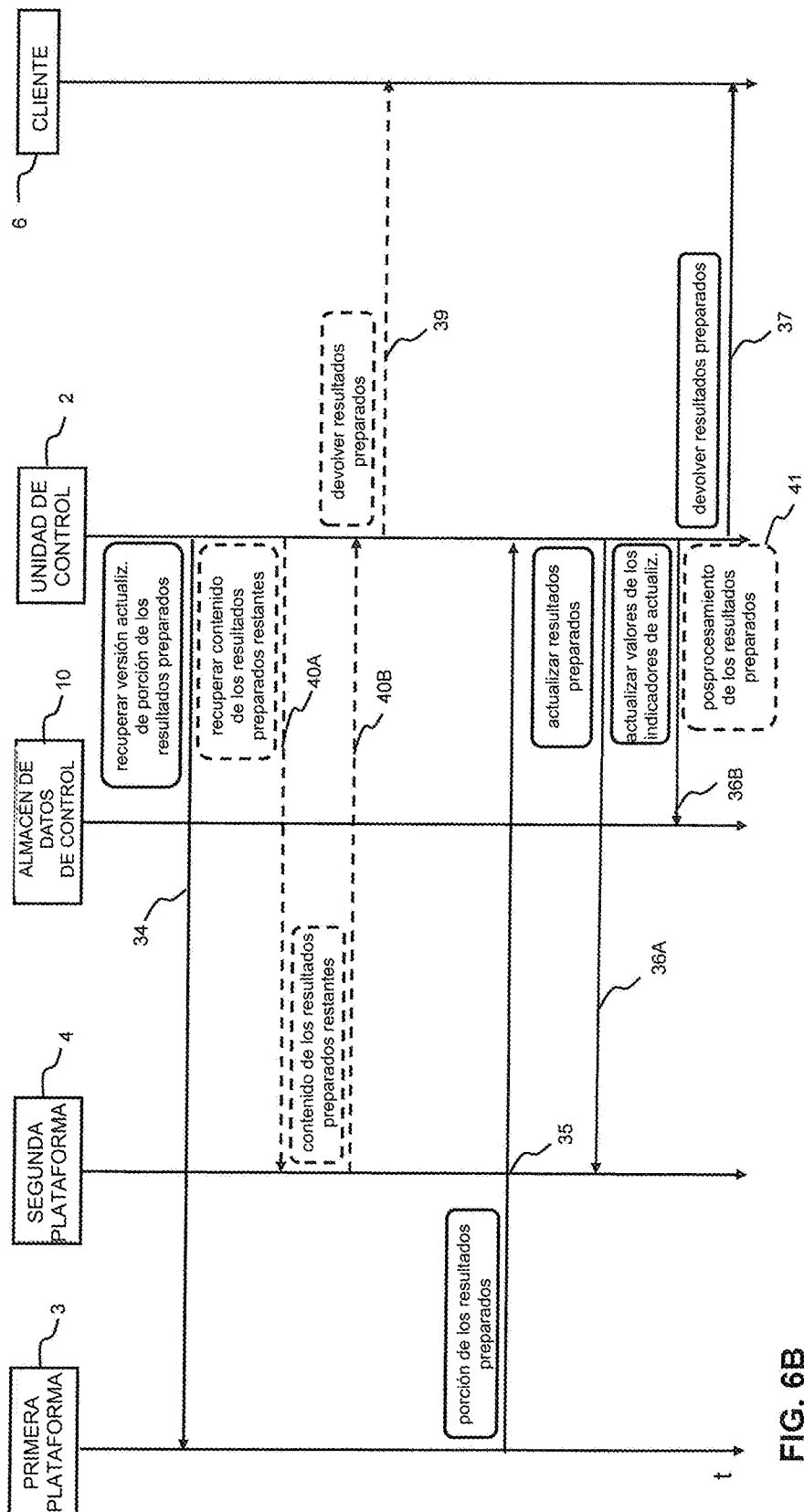


FIG. 6B

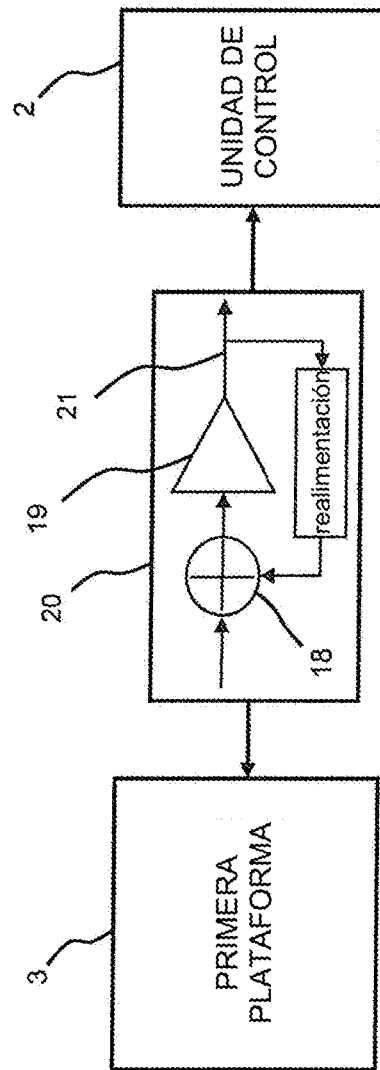


FIG. 7

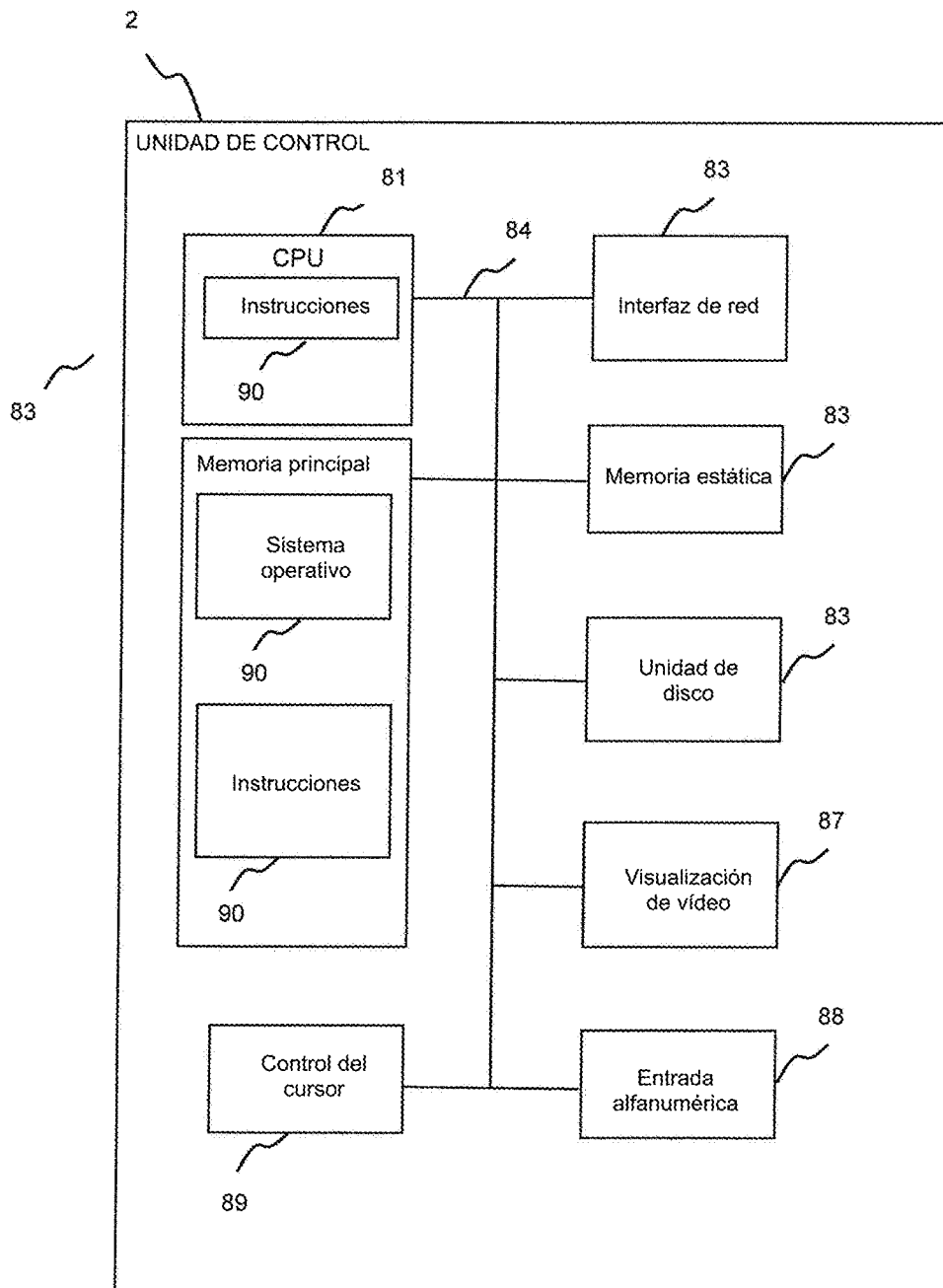


FIG. 8