

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 931 177**

51 Int. Cl.:

H04L 67/1097 (2012.01)

H04L 67/06 (2012.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **09.12.2016** **PCT/IB2016/001867**

87 Fecha y número de publicación internacional: **15.06.2017** **WO17098326**

96 Fecha de presentación y número de la solicitud europea: **09.12.2016** **E 16872483 (9)**

97 Fecha y número de publicación de la concesión europea: **28.09.2022** **EP 3387819**

54 Título: **Sistema y método para lanzamiento de información a través de un tapiz de red y granularidad de una marca**

30 Prioridad:

11.12.2015 US 201562266060 P

17.02.2016 US 201662296257 P

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

27.12.2022

73 Titular/es:

UMBRA TECHNOLOGIES LTD. (100.0%)
Suite 2006 20th Floor, Hua Qin International
Building, 340 Queen's Road Central
Hong Kong 100015, CN

72 Inventor/es:

RUBENSTEIN, JOSEPH, E.

74 Agente/Representante:

LINAGE GONZÁLEZ, Rafael

ES 2 931 177 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

Sistema y método para lanzamiento de información a través de un tapiz de red y granularidad de una marca

5 **Campo de la divulgación**

Esta presente divulgación se refiere de manera general a redes y, más particularmente, a acelerar la transferencia de datos a través de distancias de medias a largas modificando fundamentalmente la naturaleza de los intercambios entre dispositivos remitentes y receptores. El cambio de metodología es de transferencias de paquetes de ida y vuelta a envío de archivos de tipo lanzamiento ("*slingshot*") unidireccional, así como adoptar principios de informática de alto rendimiento para mejorar el rendimiento.

Esta divulgación describe específicamente un lanzamiento de información unidireccional que guarda archivos de datos completos de cualquier tamaño casi a velocidad de cableado a través de un tapiz de red en un almacenamiento remoto. El lanzamiento de información puede integrarse con tejidos de red existentes o puede ser un mecanismo autónomo que funciona de manera independiente, o una combinación de ambos.

La presente divulgación también se refiere de manera general a software y a redes y, más particularmente, a la granularidad de una marca de tiempo que rige las operaciones de funcionalidad de nivel superior.

20 **Antecedentes de la divulgación**

Lanzamiento de información a través de un tapiz de red

Las tecnologías convencionales y las limitaciones o inconvenientes de esas tecnologías sirven para demostrar por qué un enfoque diferente puede resultar beneficioso. Las características, métodos y utilidad de la presente divulgación descritas en el presente documento exponen cómo construir y usar esta tecnología. También ofrecen una perspectiva de cómo la presente divulgación infunde ciertas mejoras para superar las limitaciones de las tecnologías convencionales.

La presente divulgación puede ser beneficiosa para cumplir diversas necesidades de conexión en red. El uso de la presente divulgación por la industria financiera se ofrece en el presente documento como ejemplo para ilustrar características y cómo superan los inconvenientes de las tecnologías convencionales. Sin embargo, el alcance y la escala de la aplicación y el uso de la presente divulgación trascienden este estrecho enfoque y pueden aplicarse para muchos otros casos de uso.

Internet es una red de redes. Un tejido de red puede definirse o bien como una red bajo el control de un organismo administrativo o bien como un tipo de protocolo de red, tal como el protocolo de Internet (IP), o para una red dentro de una zona regional, etc. Los tejidos pueden tejerse entre sí mediante comunicación entre iguales de tejidos de red unidos entre sí en los bordes. Pueden usarse métodos tales como traducción de dirección de red (NAT) y otros.

Una red virtual global (GVN) es un tipo de red que existe como transmisión libre (OTT) de Internet y teje capas de tejidos en un tapiz de red con puntos de salida y entrada (EIP) en cada borde de tejido.

Un tipo de tejido de red puede ser una red de área amplia (WAN) que consiste en una línea dedicada tal como conmutación de etiquetas multiprotocolo (MPLS) o red de datos digital (DDN) u otra entre dos puntos de red de área local (LANS). Estos enlaces dedicados también pueden comunicarse como iguales con otras redes o bien de una manera de uno a uno definida como de ajuste individual o bien con una relación de una a muchas redes mediante comunicación entre iguales de ajuste múltiple, con muchos trayectos o rutas posibles a través del centro entre dos o más puntos de extremo.

Un tejido de red también puede definir la escala y el alcance de un tipo de tejido de red desde un extremo hasta otro. Ethernet define un tipo de red pero también puede clasificarse adicionalmente mediante protocolos tales como protocolo de Internet (IP) a través de Ethernet, tales como protocolo de Internet versión 4 (IPv4) o protocolo de Internet versión 6 (IPv6) y otros tipos de red. Basándose en IP se construyen protocolos tales como protocolo de control de transmisión (TCP) y protocolo de datagrama de usuario (UDP). TCP a través de IP (TCP/IP) es más detallado y tiene una mayor comprobación y gestión de errores incorporada para fiabilidad de datos enviados frente a UDP a través de IP (UDP/IP), que está más dirigido a la transmisión en flujo continuo de datos tales como audio o vídeo en la que un paquete perdido no tendrá un efecto drásticamente adverso.

Además de diferentes protocolos y versiones de IP construidas basándose en Ethernet, el propio Ethernet tiene Ethernet, Gigabit Ethernet (disponible a velocidades de 1 ó 10 ó 40 ó 100 gigabits) y otras versiones del mismo que se espera que se introduzcan. En el momento de la redacción, Ethernet es el protocolo popular de elección para conectar diversos dispositivos entre sí que pueden comunicarse entre sí.

Infiniband (IB) es un protocolo alternativo a Ethernet, usando IB puertos de NIC físicos, enchufes y cableado diferentes, así como funcionando IB de una manera diferente. Para conectar muchos nodos en un entorno de computación/informático de alto rendimiento (HPC), IB es la opción preferida. IB permite el acceso directo a memoria remoto (RDMA) nativo entre nodos evitando la red y pilas de sistema operativo (O/S) de los dispositivos anfitriones para el acceso en el que están conectados los dispositivos de almacenamiento de RDMA (u otros). También hay alternativas tales como canal de fibra y otras tecnologías de red que pueden ofrecer la capacidad de mantener transferencias de RDMA nativas rápidas y guardar archivos a distancia en un sistema de archivos en paralelo (PFS) o en RAM o guardar en otro dispositivo de almacenamiento remoto.

Aunque el canal de fibra es una alternativa a IB, para los fines de la divulgación, esta divulgación se centra en IB como el tipo de red de elección para conectar nodos de HPC, especialmente a través de grandes distancias. IB se usa como ejemplo con fines de demostración pero cualquier tipo de red alternativo que pueda proporcionar conectividad de base similar puede usarse como sustituto para soportar un lanzamiento de información.

Cuando se comparan tipos de red, la fiabilidad tiene una importancia fundamental. Otro factor impulsor principal que afecta a la elección del tipo de red y protocolo de red también está relacionado con el tiempo para que las comunicaciones se desplacen a través de la distancia del trayecto físico. La latencia es una medida de tiempo para que los datos se desplacen en un sentido o también puede expresarse como un tiempo de ida y vuelta (RTT) a lo largo de una distancia especificada entre dos puntos. El ancho de banda es una medida de la tasa de transferencia de datos de un segmento de red, ruta de comunicaciones, tejido de red u otro trayecto de red.

Hay reglas de la física que se mencionan en el presente documento como base. Dado que el tiempo y la distancia son importantes, esta divulgación usa el siguiente nivel de referencia:

- las distancias en el presente documento se miden en millas según el sistema imperial
- las medidas de distancia en el presente documento pueden ser un número con o sin puntos, y/o decimales, o pueden expresarse como un número entero
- una excepción para la distancia usada es en el índice de refracción de cables de fibra óptica indicado a continuación, en el que la distancia se expresa en una medida de metros según el sistema métrico
- el tiempo se mide en segundos, expresado como número enteros, fracciones y/o decimales de segundos
- las unidades de tiempo mencionadas en el presente documento son segundos, milisegundos (ms) y microsegundos (μ s)
- cualquier granularidad más fina que microsegundos, tal como nanosegundos (ns), puede ser importante en determinadas aplicaciones prácticas de esta divulgación pero, con fines de demostración, la granularidad más fina en el presente documento es μ s.

En informática, las medidas de tiempo principales o convencionales actuales para la conexión en red son milisegundos (ms) y para el procesamiento son microsegundos (μ s). La granularidad de una marca de tiempo puede medirse, por ejemplo, o bien como fracción (cada 1/20 ó 1/10 ó 1/100) o bien como decimales (0,05, 0,1, 0,01) de un a milisegundo. La siguiente tabla es una referencia de algunos valores posibles y su conversión equivalente correspondiente.

TABLA 1 – medidas de tiempo

N.º	Descripción	Segundos	Milisegundos (ms)	Microsegundos (μ s)
1	1/10 de un segundo	0,10000	100	100.000
2	1/20 de un segundo	0,05000	50	50.000
3	1/100 de un segundo	0,01000	10	10.000
4	10 microsegundos	0,00001	0,010	10
5	100 microsegundos	0,00010	0,100	100
6	1.000 microsegundos	0,00100	1,000	1.000

La siguiente tabla compara la velocidad de la luz en el vacío frente a la velocidad de la luz dentro del núcleo de vidrio de fibra óptica. Esto ilustra la limitación física de la eficiencia de fibra para establecer un nivel de referencia para la mejor velocidad teórica que puede esperarse lograr para que se desplace la luz a través de fibra.

TABLA 2 – velocidad de línea de fibra teniendo en cuenta la resistencia en la latencia de fibra óptica debida a la refracción: basándose en datos encontrados en el sitio: <http://www.m2optics.com/blog/bid/70587/Calculating-Optical-Fiber-Latency>

	millas/segundo en el vacío	millas/segundo a través de fibra	eficiencia de fibra
Velocidad de la luz	186.282,34	126.759,88	68,05%

Aunque el índice de refracción de cables de fibra óptica puede variar ligeramente, en el presente documento se supone que el promedio es de la siguiente manera: promedio de aproximadamente 203 m a 204 m/μs frente a la velocidad de la luz de 299,792 m/μs para una eficiencia del 68,05%.

5

Existen industrias especializadas cuyas aplicaciones tienen un requisito absoluto para la mejor calidad de servicio (QoS) con la menor latencia posible o pérdida de paquetes nula a través de la distancia. Estas industrias especializadas necesitan intercambiar información para aplicaciones tales como información financiera, ejecución de transacciones para confirmación, cálculos científicos y más. En muchos casos, se usará una línea de fibra óptica dedicada para el transporte para ofrecer más fiabilidad y un rendimiento superior a lo que puede esperarse para el transporte a través de Internet abierto.

10

Una alternativa a las comunicaciones por línea de fibra óptica terrestre es la retransmisión mediante transmisión por satélite(s). El lanzamiento puede funcionar a través de enlaces de satélite y puede ser posible que las velocidades de comunicación se aproximen a la velocidad de la luz usando tal medio de transporte de base. Los enlaces de satélite son propensos a limitaciones de ancho de banda, riesgos meteorológicos, interferencia cósmica y otros desafíos. Para esta divulgación, el enfoque para el lanzamiento de información es que su demostración se despliegue a través de redes de línea de fibra óptica terrestre.

15

Ethernet y su combinación de tecnologías de conexión en red es el tipo de red más ampliamente usado desplegado desde las redes de área local dentro de oficinas, centros de datos y otras agrupaciones de dispositivos hasta las estructuras centrales globales que conectan redes gestionadas por proveedores de servicios de Internet (ISP) a través de Internet global. El protocolo de Internet IPv4, aunque está ampliamente desplegado, es víctima de su propio éxito ya que se enfrenta a una situación en la que el número de direcciones IP disponibles se ha agotado casi por completo. Las tecnologías tales como NAT para dispositivos en una LAN han ayudado a mitigar esto, pero todavía no han resuelto el problema y, por consiguiente, hay una gran escasez de direcciones de IPv4. En esencia, IPv6 tiene lo que parece ser un amplio suministro de direcciones IP disponibles, sin embargo IPv6 no se ha desplegado universalmente debido a varios factores, siendo uno de ellos un bloqueo de equipos de legado instalados que sólo gestionan IPv4 y no tanto IPv4 como IPv6. Ethernet pasó a ser el tipo de red dominante y su amplio uso es prevalente tanto en LAN como a través de Internet más amplio porque era una norma relativamente fácil de implementar y desplegar globalmente. Los efectos de red también intervienen a medida que cada vez más dispositivos usan un protocolo o tipo de red, se vuelve más fácil que otros decidan adoptar el mismo, con un incentivo adicional de altos volúmenes que conducen a la popularización que tiene el efecto de reducir los precios.

35

En el centro de datos en el que operaciones concentradas de cálculo, almacenamiento, procesamiento y otras están extendidas a través de diversos servidores montados en bastidores, se requirió un transporte más rápido que Ethernet para conectar mediante canal secundario estos servidores entre sí para que compartan datos y se comuniquen. El canal de fibra e IB son dos de tales tecnologías que ofrecen ancho de banda de latencia ultrabaja y alta capacidad. Las transferencias sin pérdidas y en paralelo de IB ofrecen fuertes ventajas, particularmente junto con el despliegue de acceso directo a memoria remoto (RDMA) y también la oportunidad de desplegar y usar un sistema de archivos en paralelo (PFS) dispersado de manera global. La limitación de IB era que sólo podía desplegarse a una distancia relativamente corta medida en metros. Después se extendió a tan sólo unos pocos kilómetros. Hasta hace poco, los enlaces de IB de "larga distancia" estaban limitados al interior de una red de área metropolitana (MAN) o entre dos áreas metropolitanas cercanas que conectan centros de datos entre sí mediante IB superrápida a través de líneas dedicadas entre sí. Tecnologías recientes permiten extender IB hasta 20.000 kilómetros entre dos dispositivos de pasarela de IB a través de una línea de fibra oscura.

40

45

Se conocen inconvenientes en las técnicas convencionales del protocolo IP de Ethernet, especialmente a través de larga distancia. Los problemas de IP están provocados de manera inherente por cómo se envían los paquetes. Las comunicaciones IP que usan su protocolo más popular TCP/IP conllevan desplazamientos de ida y vuelta y descripciones dando como resultado un RTT superior al necesario.

50

IP transmite datos en paquetes. Cada paquete incluye cabeceras, carga útil de datos y posiblemente otra información. El tamaño de un paquete permitido por un tejido de red de IP se define como unidad de transmisión máxima (MTU) y puede estar limitado a cientos de bytes a 4 KB o más, pero si la carga útil es mayor de lo que permite el límite de paquete establecido, para transportar esa carga útil de datos desde el punto A hasta el punto B, se necesitará dividir la carga útil en múltiples partes con dos o más paquetes y posteriores desplazamientos de ida y vuelta requeridos por cada transmisión y confirmación. El tamaño de paquete también tiene un impacto sobre si un paquete experimenta fragmentación o no. Resulta habitual que un paquete fragmentado se corrompa si se fuerza la fragmentación por un dispositivo intermedio que tiene una MTU inferior a los otros dispositivos entre segmentos de red en un trayecto de transmisión a través de una cadena de dispositivos. Este inconveniente de IP puede conducir a abandono de paquetes que tienen que volver a enviarse.

55

60

- Aunque lleguen todos los paquetes, la fragmentación aumenta la congestión en el centro y también puede inundar los dispositivos receptores con más paquetes de lo esperado. Por ejemplo, si un dispositivo de envío usa un ajuste común de 1500 bytes de MTU, y descompone una carga útil en un flujo de paquetes cada uno de los cuales tiene 1500 bytes, el dispositivo de envío los transmitirá en flujo continuo a la velocidad que puede gestionar la pila y el NIC hasta el ancho de banda disponible de la conexión de la "última milla". Y si un dispositivo en el centro en Internet tiene un ajuste de MTU de 1492, entonces cada paquete que fluye a través de este dispositivo se fragmentará o dividirá en múltiples paquetes. El resultado es que el dispositivo de recepción está recibiendo el doble del número de paquetes que transmitió en flujo continuo el dispositivo de envío. Esto puede saturar e inundar el dispositivo de recepción conduciendo a congestión y posible pérdida, o al menos una ralentización. Además, los dispositivos de red están clasificados para enviar una cierta cantidad de paquetes por segundo, independientemente de lo "completo" que esté un paquete. Normalmente los paquetes fragmentados no están completos hasta su capacidad, reduciendo adicionalmente el posible ancho de banda con el efecto en cadena de tener una reducción en el rendimiento de red.
- La fiabilidad y verificación de paquetes de TCP/IP ofrecen tranquilidad pero con un alto coste para la velocidad y una rigidez forzada porque algunas desventajas con respecto a técnicas convencionales de IP a distancia son que hay límites debidos a limitaciones físicas y la naturaleza del/de los protocolo(s) de red.
- Específicamente, el transporte en paquetes tal como transferencia de paquetes en múltiples partes que divide una carga útil más grande corre el riesgo de fragmentarse por un dispositivo en el centro que puede no ser capaz de o estar configurado para gestionar una gran carga útil.
- Se sabe que el transporte de ida y vuelta de Ethernet/IP tiene degradación de QoS a lo largo de la distancia con retardos, problemas conocidos tales como congestión que conduce a retardos, especialmente a lo largo de la distancia. TCP/IP puede atrapar un paquete perdido y lo retransmitirá. Aunque esto es bueno desde un punto de vista de la fiabilidad, aumenta la latencia y congestión para las redes.
- También es importante observar cómo se gestiona un paquete de datos mediante Ethernet, porque tiene un modelo de almacenamiento y reenvío en el que se recibe un paquete, se examina y después se reenvía únicamente después de haberse recibido y examinado completamente todo el paquete que incluye la carga útil completa. La latencia para esta operación dentro de un ordenador/enrutador/dispositivo de red para gestionar un paquete de tráfico de Ethernet es de aproximadamente 100 microsegundos (μs) por dispositivo. Este arrastre de retardo de 100 μs por cada dispositivo intermedio para que un paquete se reciba y se pase puede parecer minúsculo, pero a lo largo de una ruta de larga distancia, este retardo puede multiplicarse al pasar a través de muchos dispositivos. NAT también añade otro punto de resistencia de fricción pequeño pero todavía medible. El retardo debido a la metodología de cliente-servidor (preparar un paquete, enviarlo, transportarlo, recibirlo, combinar paquetes, usarlo) aumenta etapas sobredimensionadas e innecesarias a un flujo de comunicación dentro de la pila de un O/S de ese dispositivo.
- Existe posibilidad de mejora en los protocolos IP actuales, incluyendo la necesidad de reducir la latencia, eliminar la pérdida de paquetes y resolver otros problemas tales como ineficiencias en la formación de paquetes y transferencias bidireccionales, abordar el producto de ancho de banda y retardo (BDP) y algoritmos relacionados que rigen las transferencias de archivos, y más.
- Aunque las metodologías de mercado actuales se han diseñado por ingeniería para ser lo más eficientes posible con información de mercado emitida por radiodifusión a través de multidifusión de UDP/IP y pedidos y confirmaciones de negociación enviados en los paquetes más pequeños posible a través de TCP/IP con la latencia más baja, todavía existen problemas. UDP/IP no tiene el mismo mecanismo de corrección de errores que TCP/IP y, como tal, ni el receptor ni el remitente saben si se pierde un paquete. Y es durante momentos de mucho tráfico cuando la pérdida de paquetes es más probable, y precisamente este es el momento en el que resulta imperativo que los comerciantes tengan la información más completa posible.
- También hay cambios para calcular qué protocolos de red relacionados van a tener problemas para abordar la necesidad de mover grandes cantidades de datos.
- Se requiere un nuevo paradigma debido a tecnologías emergentes tales como cadena de bloques (y otras tecnologías de registro distribuidas), Internet de las cosas industrial (HOT), medicina remota y otras en las que existe un requisito de transmisión rápida y fiable de cada vez más cargas útiles de datos. El lanzamiento es un mecanismo para enviar información en un sentido como un único archivo de prácticamente cualquier tamaño.
- Un lanzamiento de información usa RDMA a través de IB o un tipo de red de base similar para guardar un archivo de datos de tamaño ilimitado en un dispositivo de almacenamiento remoto tal como un PFS, y para que un gestor de cola de archivos (FQM) o mecanismo similar en un dispositivo en esa región remota objetivo extraiga el archivo del PFS. La corrección de errores está incorporada en el dispositivo de almacenamiento y posteriormente rige el procedimiento de escritura del archivo, de la misma manera que interacciona un dispositivo de almacenamiento conectado localmente con un O/S dentro de un dispositivo anfitrión. En un

sentido, el lanzamiento consiste en distribuir lo que era un procedimiento interno en uno globalmente distribuido. El transporte unidireccional es otra característica del lanzamiento que produce la eficiencia de UDP/IP al tiempo que se logra la fiabilidad de TCP/IP abordando la necesidad de confirmación de paquetes enviados. El lanzamiento usa guardados de archivos de flujo en paralelo que tienen el efecto de que el último byte llega al mismo tiempo o casi al mismo tiempo que la llegada del primer byte transmitido.

El lanzamiento puede usar cualquier red de base que puede soportar RDMA y que tiene atributos similares a IB. Los principios de informática de alto rendimiento (HPC) tal como RDMA, IB y otros atributos proporcionan las bases sobre las que producir un enfoque novedoso. Se indica que el almacenamiento en un dispositivo de PFS puede ser físico o virtual, o un HDD, SAS, SSD, disco de RAM, o cualquier otro dispositivo que puede proporcionar un medio de almacenamiento rápido y permite RDMA o un tipo equivalente de acceso directo remoto.

Para la demostración de esta divulgación, se usa IB como ejemplo de red sobre la cual realizar el lanzamiento. El ancho de banda de IB con carga puede alcanzar del 90 al 95 por ciento del máximo de BW teórico. IB presenta conmutación de corte, en la que los dispositivos de IB reciben las cabeceras de un paquete, usan lógica para la decisión de reenvío y transmiten la carga útil hacia delante, incluso antes de recibirse la carga útil completa. Por tanto, IB tiene una latencia extremadamente baja en comparación con Ethernet. También es mucho menos detallada que TCP/IP o UDP/IP. Se ejecuta en conexiones de fibra oscura. En comparación con Ethernet a través de fibra oscura, todavía es relativamente más rápido y, si se usa RDMA nativo, puede medirse la latencia como desplazamiento unidireccional para una transmisión eficaz en vez de ida y vuelta como RTT para TCP/IP de Ethernet.

En el contexto de la transmisión de datos financieros, esta divulgación es una infraestructura para aprovechar el tapiz y usar enfoques de conectividad avanzada para abordar la necesidad de acelerar la ejecución de transacciones globales y otras necesidades de comunicaciones exigentes.

Una distinción importante es que el lanzamiento no realiza una transferencia de archivos tradicional sino que, en vez de eso, ejecuta un archivo guardado en un almacenamiento remoto. Como transporte para una agrupación de diversos "archivos" de carga útil de datos, este método de lanzamiento mueve un bloque más grande de datos más rápido desde una región hasta otra región de una manera muy eficiente.

Ganancias significativas a través de largas distancias. Véanse los ejemplos de realizaciones, figuras y sus explicaciones adjuntas para más características.

El lanzamiento de información puede integrarse en un tapiz en un tipo de red tal como una GVN a través de puntos de salida y entrada (EIP) en los bordes entre tejidos tales como EIP 2970 en la figura 29 en el presente documento. Además, usando un mecanismo de lanzamiento de información como base, pueden soportarse otras características tales como salto de lanzamiento ("*Sling Hop*") (solicitud de patente provisional estadounidense n.º 62/327.907), baliza de datos de información (solicitud de patente provisional estadounidense n.º 62/327.846), ruta de lanzamiento ("*Sling Route*") (solicitud de patente provisional estadounidense n.º 62/327.911) y otras.

Agrupación de HPC con todos los nodos que pueden actuar como nodo de control que pone archivos en ubicaciones de almacenamiento de PFS para que los recuperen otros nodos. Aunque es posible que un nodo extraiga archivos de un PFS en una región remota, la latencia añadida para extraer y marcar y mover archivos no tiene sentido. Lo mejor es que un remitente guarde el archivo en un PFS lo más cerca posible del objetivo. Una característica de esta divulgación es que los nodos sean conscientes unos de otros, de instancias de almacenamiento de PFS, así como de la proximidad de la instancia de almacenamiento de PFS al dispositivo objetivo que el remitente desea que lea el archivo guardado.

Cualquier archivo presentado al mecanismo de lanzamiento de información, desde cualquier origen, se guardará de manera remota en una región objetivo como archivo completo. El gestor de cola de archivos recuperará el archivo de la cola en el PFS o de RAM o cualquier otro medio de almacenamiento sustituto. Procesará el archivo, comprobando que está completo en el momento del procesamiento. Los archivos completos estarán disponibles para la siguiente cadena de cálculo.

Un lanzamiento de información supera los inconvenientes de técnicas convencionales ya que proporciona una infraestructura eficiente, rápida y viable para compartir grandes volúmenes de información entre dispositivos a través de grandes distancias.

La latencia de transacción se reduce significativamente. La transmisión a larga distancia se realiza casi a velocidad de cableado, sin necesidad de dividir el archivo en múltiples partes ni formar paquetes al enviarlo o hacer que vuelvan a ensamblarse paquetes. Esta divulgación también elimina el sobredimensionamiento de paquetes de cabeceras y los límites de delimitación de múltiples partes. La transferencia de corte es mucho más eficiente que el almacenamiento y reenvío de la gestión de paquetes de IP. La detección y corrección de errores avanzadas al escribir el archivo proporcionan adicionalmente una eficiencia superior.

La naturaleza del intercambio de información se transforma de transferencia de paquetes secuencial en transferencia de archivos de múltiples hilos en paralelo a través de distancia y el guardado de archivos en paralelo en el almacenamiento en la región remota es más eficiente.

5

También hay un beneficio añadido de disponibilidad superior de archivos accesibles a través de una estructura que contiene RDMA y carga útil extendida que facilita transferencias de datos mucho mayores. RDMA tiene una lectura y escritura superrápidas en/de dispositivos de almacenamiento sin tener que pasar por el anfitrión ya que evita la pila tanto en el origen como en el destino.

10

Mediante esta divulgación también se evitan cuellos de botella y puntos de resistencia debidos a la naturaleza de protocolos de IP. Ya no hay necesidad de formación de paquetes al guardar archivos completos en un PFS remoto para que otro dispositivo los extraiga y los use.

15

Desacoplar la naturaleza de ida y vuelta RTT de la comunicación en paquetes para dar el envío de archivos unidireccional reduce el tiempo de transmisión. Otra ventaja sutil es que los datos de cuerpo pueden ser variables en cuanto al tamaño de archivo.

20

Para tecnologías convencionales, si los datos que van a transportarse son más grandes de lo especificado por el tamaño de carga útil de un paquete, se dividirán en múltiples partes y, por tanto, requerirán múltiples desplazamientos de ida y vuelta para la transmisión. Esto añade un tiempo significativo para completarse la transmisión y más procesamiento para crear y después validar y volver a ensamblar un mensaje de múltiples partes. El lanzamiento de información ofrece un tiempo de transmisión conocido y sistemático entre puntos sin necesidad de que el dispositivo de envío divida el archivo ni de que el dispositivo de recepción tenga que volver a ensamblarlo.

25

Supera los retardos debidos al total acumulativo de diversos tiempos de transmisión y tiempos de procesamiento para una cadena de paquetes. Cuando se forman paquetes, el tiempo de transmisión de paquetes de IP de Ethernet se calcula como tiempo de RTT * número total de paquetes, el lanzamiento de información proporciona un mecanismo de sincronismo más preciso.

30

El lanzamiento de información puede enviar un archivo entero sin necesidad de descomponer datos para dar cargas útiles de múltiples partes enviadas mediante un flujo de paquetes tal como se realiza mediante IP de Ethernet. Esto también tiene el efecto positivo de mantener el movimiento de datos a un nivel muy bajo, sin tener que encargar a la CPU, RAM u otros elementos que realicen el apilamiento de un dispositivo.

35

El lanzamiento de información evita además la necesidad de recibir un flujo de paquetes. Ethernet consume tiempo y recursos valiosos cuando tiene que recibir un flujo de paquetes y después valida cada uno, se asegura de que están en el orden correcto y después vuelve a ensamblar el archivo de datos.

40

Las ventajas principales del lanzamiento de información (RDMA OTT IB) son flujos paralelos cuando se transfiere, baja latencia, transferencia sin pérdidas, conexión en red sin copias, carga reducida sobre el hardware mediante conmutación de corte, RDMA que evita la pila de los dispositivos anfitriones en ambos extremos, uso nativo de sistema de archivos en paralelo (PFS), transferencia unidireccional sin necesidad de paquetes de confirmación, sin RTT, simplemente transferencia unidireccional, sin "límite de tamaño de paquete".

45

La naturaleza de una transferencia mediante lanzamiento puede cambiar cómo se enfoca la composición de la transmisión. La cabecera de archivo de datos y los datos de cuerpo pueden ser más detallados, y también pueden incluir una cabecera de archivo y marcador de fin de archivo (EOF) distintivo. Los datos de cuerpo, incluyendo parámetros, matriz de información, infraestructura lógica algorítmica y más, pueden permitir la transmisión de datos más ricos en cuanto a características. También puede incluirse un pie de archivo de datos, con sumas de comprobación, seguridad y otros elementos relacionados con la integridad, dentro del archivo transmitido.

50

La presente divulgación puede ofrecer ventajas de comunicaciones. A larga distancia, puede lograrse casi tiempo real, casi con una velocidad de línea cableada. A corta distancia, puede ahorrar μ s (microsegundos) con respecto a métodos anteriores y el resultado es una rápida disponibilidad de datos en una región remota. El lanzamiento de información cambia cómo se ejecutan las transacciones acelerando la transmisión, simplificando el direccionamiento, permitiendo cargas útiles de archivo más grandes y ofreciendo otras ventajas.

60

Las figuras adjuntas y los ejemplos de realizaciones describen más sobre las técnicas convencionales, los inconvenientes de las técnicas convencionales y las características de esta divulgación así como cómo superan estas características los inconvenientes de las técnicas convencionales.

65

El lanzamiento de red y las tecnologías derivadas pueden proporcionar beneficios, tales como, pero sin limitarse a: a) el lanzamiento de red es una entrega rápida y unidireccional de datos de una vez; b) puede no haber

necesidad de formar paquetes de tráfico, tamaño de archivo ilimitado, llegada del primer byte transmitido casi a velocidad de cableado, con la llegada del último bit $\Delta tiempo$ lo más rápido posible mediante flujos de transmisión paralelos dinámicos ajustados a escala basándose en el margen de BW disponible; c) la conexión en red de alto rendimiento (HPN) basándose en lanzamiento ofrece numerosas ventajas; y d) puede comentarse otra tecnología relacionada con el lanzamiento durante una reunión en una divulgación más formal.

Las tecnologías de red financieras y otras basadas en lanzamiento de información pueden proporcionar beneficios, tales como, pero sin limitarse a:

1. Transmisión de información mediante generador de pulsos de baliza de datos: aborda cuestiones con la alimentación de datos a partir de mercados, entre otras ventajas, no hay ninguna pérdida debida a congestión o abandono de paquetes, se envían datos como archivos con tamaño de archivo ilimitado. Hay una fiabilidad del 100%, de punto a punto casi a velocidad de cableado, independientemente de la distancia entre el origen y el objetivo.

2. Del pedido de negociación a la ejecución: se produce casi a velocidad de cableado, permitiendo tipos de mensajes más grandes que pueden permitir el envío de algoritmos complejos desde el anfitrión de origen hasta un anfitrión objetivo en el borde de un mercado en el que puede llevarse a cabo una serie de instrucciones con la mayor parte de la lógica enviada de una vez como bloque de datos integral frente a la necesidad de paquetes de datos individuales.

3. Cadena de bloques: aunque se ha prestado mucho interés y atención a esta tecnología, una cuestión no abordada de manera suficiente es que la altura de los datos relacionados da como resultado un tamaño de archivo bastante grande y que los protocolos IP actuales no son el transporte más eficiente. El lanzamiento y la tecnología correspondiente permiten el transporte de todo el archivo de cadena de bloques casi a velocidad de cableado, con una fiabilidad del 100%.

4. La integración de potenciaciones también debe ajustarse en el flujo de trabajo de redes de legado existentes al menos localmente dentro de regiones de los dispositivos cliente/servidor hacia y desde el mecanismo de lanzamiento. Los componentes de red de lanzamiento de información están diseñados para potenciar las largas distancias entre diversas redes en regiones o zonas y acelerar el flujo de datos entre las mismas.

5. El salto de lanzamiento también puede integrarse como segmento sin interrupciones y transparente en infraestructuras de IP existentes, de tal manera que: a) se permite la integración en flujos de trabajo de trayecto de IP de legado existentes; y b) la optimización en la parte media aprovecha las ventajas de lanzamiento sin interrupciones pero de manera transparente para el tráfico de cliente a través del salto de lanzamiento.

TABLA 3 – Comparación del tiempo de ida y vuelta (RTT) de IP de Ethernet frente a la latencia de estructura central de fibra (FBB):

N.º	Ubicaciones		RTT de IP de Internet de Ethernet, min./prom (ms)	Distancia de punto a punto (millas)	FBB, latencia unidireccional	
	Desde	Hasta			Milisegundos (ms)	Microsegundos (µs)
1	Nueva York	Londres	65/73	3.465	27,3	27.335
2	Hong Kong	Londres	174/217	5.969	47,1	47.089
3	Nueva York	Singapur	209,7 /241	9.538	75,2	75.245
4	Nueva York	Los Ángeles	67/69	2.448	19,3	19.312
5	Nueva York	Tokio	142/172	6.737	53,1	53.148
6	Nueva York	Frankfurt	73,4/87	3.858	30,4	30.435
7	Nueva York	Hong Kong	191,6 /253	8.058	63,6	63.569
8	Nueva York	París	82/85	3.631	28,6	28.645
9	Nueva York	Sídney	233 /261	9.946	78,5	78.463
10	Los Ángeles	Londres	131/144	5.447	43,0	42.971
11	Los Ángeles	Hong Kong	168/168	7.245	57,2	57.155
12	Los	Singapur	193/215	8.788	69,3	69.328

	Ángeles					
13	Los Ángeles	Sídney	163/166	7.497	59,1	59.143
14	Londres	Sídney	294/298	10.571	83,4	83.394
15	Londres	Singapur	172/188	6.748	53,2	53.235
16	Londres	Frankfurt	9/20	396	3,1	3.124
17	Tokio	Hong Kong	43/55	1.788	14,1	14.105
18	Hong Kong	Singapur	30/31	1.609	12,7	12.693
19	Punto A	A más de 100 millas	- ms	100	0,8	789
20	Punto A	A más de 1000 millas	- ms	1.000	7,9	7.889
21	Punto A	A más de 12000 millas	- ms	12.000	94,7	94.667

Fuentes para los datos:

https://www.sprint.net/lg/lg_start.php – pruebas de ping entre puntos

<http://www.verizonenterprise.com/about/network/latency/#latency> – tablas estáticas para referencia

<https://wondernetwork.com/pings/Hong+Kong> – tablas dinámicas, puede establecerse la ciudad de base y recuperarse datos estadísticos

https://ipnetwork.bgtmo.ip.att.net/pws/network_delay.html – tablas * (EE.UU.)

https://ipnetwork.bgtmo.ip.att.net/pws/global_network_avgs.html – tablas, promedios globales

Fuentes para cifras de latencia de FBB:

Latencia unidireccional de FBB calculada específicamente para esta divulgación.

La granularidad de una marca

Esta divulgación describe la metodología para usar una granularidad de grano muy fino para marcas de tiempo para regir ciclos de procesamiento de datos o bien fijos o bien variables. El tiempo asignado para un ciclo puede calcularse dinámicamente basándose en varios factores incluyendo el tiempo requerido para procesar un elemento durante un ciclo. Esta información puede usarse para calcular de manera algorítmica el límite del número máximo de elementos que pueden procesarse teniendo en cuenta el procesamiento posterior durante un ciclo de tiempo. Hay información disponible tanto hacia como desde la capa de aplicación.

El software en forma de secuencias de comandos o programas o rutinas u otros tipos de bloques de código (SW) puede ejecutarse con un factor desencadenante de un uso. En algunos casos, esto es suficiente. Sin embargo, hay tipos de SW que se necesita ejecutar a intervalos de tiempo regulares y repetitivos.

En el sistema de UNIX y LINUX está incorporado un planificador denominado ejecución de comandos (CRON) que ejecutará el SW en un punto de tiempo específico. También puede establecer tiempos de intervalo. Una alternativa a una tarea de CRON será tener un centinela que se ejecuta como proceso de fondo que tiene un reloj interno con un tiempo establecido u otros factores desencadenantes, momento en el cual pueden lanzarse otras tareas de SW. En teoría, esto funcionará como un CRON pero puede establecerse de una manera más personalizable.

También puede prepararse una secuencia de comandos de tal manera que, una vez ejecutada, siga realizando un bucle basándose en una condición que nunca se vuelve verdadera para forzar una salida, o basándose en un número muy grande de veces que se permite que realice un bucle con un número decreciente cada vez que realiza un bucle.

Puede desplegarse una combinación de lanzamiento de CRON y bucle interno en un caso en el que la granularidad del tiempo de bucle es menor que el minuto mínimo requerido por una configuración de crontab. Por ejemplo, si un determinado acontecimiento debe producirse cada cinco segundos, puede desencadenarse un CRON de un minuto y, dentro de esa secuencia de comandos de SW, un bucle que se repite doce veces (60 segundos/12 bucles = intervalos de 5 segundos).

También puede ejecutarse una secuencia de comandos y después, al final de su ejecución, puede llamarse a una nueva instancia de la misma para que vuelva a ejecutarse. Esto puede tener una función de suspensión al

final para realizar una pausa entre su siguiente tiempo de ejecución o puede establecerse con una suspensión de cero segundos que implica una nueva llamada a sí misma para que empiece inmediatamente al final.

Un gestor relacionado puede actuar como guardián haciendo algo como monitorizar un directorio y, cuando su lista de archivos se rellena con un determinado número de archivos, esto puede desencadenar el lanzamiento de otra secuencia de comandos para el procesamiento de los archivos. Por tanto, el lanzamiento condicional de una secuencia de comandos de SW se basa en que se alcance un factor no basado en el tiempo. Una variación de esto puede basarse en el momento en el que una cola está llena con una cantidad de elementos en la cola, en el tamaño de memoria total en RAM o en el almacenamiento de la cola, o basarse en el espacio libre restante u otros factores.

Con respecto a los bucles o la ejecución de una secuencia de comandos de SW, puede haber condiciones tanto internas como externas establecidas, tales como un tiempo de ejecución máximo o un mecanismo para intentar asegurarse de que una secuencia de comandos o función u otra rutina sólo se ejecuta una vez. Esta comprobación para asegurarse de que no se ejecutan instancias por duplicado se realiza para evitar que dos operaciones solapantes se ejecuten al mismo tiempo que pueden interferir entre sí.

Una operación de HW clave es el reloj de sistema interno en el que se basan muchas funciones. Los relojes internos están sujetos a deriva a lo largo del tiempo, otros funcionan más lento o más rápido de lo que deberían y, por tanto, se actualizan mediante ejecución regular de secuencia de comandos de SW.

El tiempo y los intervalos de ejecución se establecen o bien se codifican de forma rígida en una secuencia de comandos o bien como variables dinámicas de modo que secuencias de comandos/rutinas/managers/módulos u otro software dependientes unos de otros pueden usar los mismos parámetros de marcas. Estas variables dinámicas para marcas pueden establecerse en una base de datos, en un archivo plano incluido, pueden calcularse o generarse de otro modo.

Los factores desencadenantes de un uso pueden resultar útiles pero, algunas veces, se necesita que sigan ejecutándose, volviendo a ejecutar código, secuencias de comandos o rutinas y, por tanto, se necesita automatizar los factores desencadenantes. Con el fin de conservar recursos de sistema y también tener más control sobre estos factores desencadenantes, se requiere una infraestructura relacionada con el tiempo para permitir la planificación.

Los CRON pueden ejecutarse en un momento específico en un día específico o cada X unidades de tiempo o mediante otros diversos periodos de tiempo relacionados con el calendario. La unidad de tiempo más pequeña especificada por un CRON es un minuto. Esto es aceptable para ciertas tareas relacionadas con el sistema tales como rotación de registros, copias de seguridad planificadas u otros procedimientos de rutina que pueden y deben automatizarse. Sin embargo, algunas tareas requieren una granularidad que es mucho más fina que un minuto así como requieren más monitorización o control en la capa de aplicación de software de terreno de usuario o espacio de usuario.

En los casos de secuencias de comandos de recuento o de bucle condicional o secuencias de comandos que se llaman a sí mismas al terminar el ciclo actual, existen algunos problemas, pero una de las cuestiones principales es que la siguiente llamada se realiza en un tiempo indeterminado. Hay maneras para mantener un seguimiento del tiempo de ejecución y añadir una función de suspensión para intentar iniciar el siguiente bucle o iteración de llamada a un intervalo regular. Sin embargo, esto sólo es aplicable para un ciclo que se ejecuta en menos tiempo que el tiempo de intervalo. En el caso en el que el ciclo tarda más que el tiempo de intervalo previsto, la siguiente marca llegará tarde y, por tanto, puede descolocar el sincronismo de los siguientes ciclos. La gestión de este procedimiento puede complicarse muy rápidamente. Por tanto, esto crea un problema con la regularidad y la fiabilidad.

Un centinela que ejecuta un proceso de fondo es muy similar a un CRON, pero requiere su propio planificador y gestor para regir los tiempos de ejecución.

La ejecución condicional basada en factores tales como una longitud de lista, cantidad de elementos de cola, espacio libre en el disco u otros factores puede resultar eficiente en algunos casos, resulta ineficiente y no sistemática en otros casos. Algunos elementos pueden atascarse en una cola durante una cantidad de tiempo excesiva. Además, esto no tiene en cuenta otros factores de modo que algunas veces los recursos de servidor se desaprovechan y después, otras veces, se sobreutilizan provocando posiblemente otros problemas de efectos en cadena negativos.

Los bucles dentro de una secuencia de comandos desencadenados por minuto mediante un CRON son secuenciales y, por tanto, si se indican 12 bucles en 60 segundos entonces se asignan cinco segundos por cada bucle. Tal como se indicó anteriormente, si un bucle sólo tarda 3 segundos, puede emitirse una suspensión de 2 segundos al final del bucle para mantenerlo de manera regular a 5 segundos. Los verdaderos problemas surgen cuando uno o más bucles tardan más de 5 segundos, lo cual puede empujar el tiempo de secuencia de

comandos a más de 60 segundos. Al siguiente minuto, el planificador de CRON llamará a la siguiente instancia de la secuencia de comandos que comenzará sus 12 ciclos de bucle. Las cuestiones negativas de ejecutar simultáneamente secuencias de comandos pueden provocar problemas significativos debido a la gestión solapante de archivos o, en el caso de procesamiento de transacción, si dos secuencias de comandos están ejecutando una transacción, puede haber duplicados no pretendidos, o etapas huérfanas pertenecientes a procesos independientes y otras cuestiones.

Hay otros aspectos negativos con respecto a las técnicas convencionales y se refieren principalmente a la granularidad en milisegundos (ms) o incluso la granularidad más fina de tiempo tal como microsegundos (μs).

En informática, las medidas de tiempo convencionales o principales actuales para la conexión en red son milisegundos (ms) y, para algunos procesamientos, son microsegundos (μs). La granularidad de una marca de tiempo puede medirse, por ejemplo, o bien como fracción (cada 1/20 ó 1/10 ó 1/100) o bien como decimales (0,05, 0,1, 0,01) de un milisegundo.

En la planificación de la ejecución de secuencias de comandos, rutinas, etc., de software, existe una necesidad de un mayor control sobre el momento en el que se desencadena su ejecución. Para una granularidad de tiempo muy fina tal como en el caso en el que los tiempos de intervalo se miden en ms o μs, se requiere un gestor inteligente para regular la granularidad de marcas de tiempo.

Este gestor requiere un reloj de sistema preciso y actualizado. En el gestor están incorporados actualizadores de tiempo para mantener el reloj de sistema lo más sincronizado posible con el tiempo real. Pueden usarse algoritmos de acuerdo del mejor de la clase tales como algoritmo de intersección o algoritmo de Marzullo.

Este gestor de marcas también se basa en monitores de recursos de sistema, en entender tareas de procesamiento de sistema y otras capacidades (y restricciones) físicas, así como otros factores para basar su estimación en cuánto tiempo debe tardar un ciclo de una llamada de secuencia de comandos o rutina u otro software en completar su ejecución.

Teniendo también en cuenta los elementos que van a procesarse antes del procesamiento, puede realizarse una estimación de tiempo más razonada. En el cálculo algorítmico de cuánto debe durar una marca hay tres factores principales usados en el cálculo de la delta de t o la cantidad de tiempo desde el inicio de la marca hasta el final.

La granularidad de una marca

$$\Delta t = P + Q + R$$

t = Delta de tiempo desde el inicio de la marca hasta el final de la marca

P = Tiempo para el procesamiento por lotes de elementos para esta marca en el ciclo.

Q = Tiempo para cálculos de procesamiento de lotes posterior.

R = Tiempo para el retardo para garantizar que no hay ningún solapamiento entre elementos de lote y/o para establecer una desviación entre marcas.

Esta granularidad de una marca se basa en varios factores. El tiempo de inicio de una marca o bien se basa en completarse el último ciclo llamado por una marca o bien es según un intervalo de tiempo fijo planificado.

Puede haber dos tipos de ciclos:

Un ciclo de tiempo fijo se basa en una estimación del tiempo por cada elemento que va a procesarse (P) y después la gestión de procesamiento posterior (Q), se establece una cantidad limitada de elementos. Este límite garantiza que ningún ciclo se ejecuta más allá de su tiempo.

Un ciclo de tiempo variable permite que todos los elementos se procesen en ese lote. El tiempo de retardo R garantiza que el siguiente ciclo extrae un lote nuevo y que todos los elementos en el último lote procesado se han procesado (P) y que se ha completado el procesamiento posterior (Q). Por ejemplo, para archivos en una cola, puede procesarse una lista en la etapa P y después pueden marcarse, moverse, eliminarse o modificarse de otro modo los archivos de modo que no se extraen en el siguiente lote. El retardo de tiempo (R) garantiza que hay suficiente retardo entre el último elemento procesado y sometido a procesamiento posterior en el último lote y el primer elemento que va a procesarse en el siguiente lote.

Una parte componente de este gestor de marcas es mantener registros históricos de datos de tiempo de ejecución para cada ciclo. Estos registros registrados pueden usarse para analizar periodos de uso insuficiente y de uso pico. También puede detectar retardos debidos a elementos no procesados cortados debido al tiempo limitado para abordar las cantidades por cada ciclo. Si esto se produce con demasiada frecuencia, entonces es un indicador de que se requiere más potencia de procesamiento, de que el software no es eficiente o puede indicar otras cuestiones tales como ralentizaciones de base de datos u otras necesidades para refactorización de software u otras soluciones para resolver cuestiones.

La marca puede tener una granularidad tan fina como se requiera por la aplicación, puede establecerse para ejecutarse a un intervalo fijo, a intervalos variables basándose en longitud de ciclo u otros factores. La granularidad de una marca permite ráfagas de baliza de red sistemáticos, funcionalidad relacionada con lanzamiento y otra funcionalidad dependiente del tiempo.

Para más características, véanse ejemplos de realizaciones tal como se presentan en las figuras adjuntas y sus descripciones. La granularidad de una marca se basa en que temporizadores, circuitos de tiempo, software de mantenimiento del tiempo, lógica, actualizadores del tiempo y otras tecnologías de mantenimiento del tiempo y sincronización, funcionen de manera apropiada con el fin de que proporcionen tiempo de sistema preciso para relojes internos y otros mecanismos. La granularidad de una marca se construye basándose en estos mecanismos de tiempo.

La granularidad de marca de ciclo de tiempo fijo se establece mediante variables dinámicas y basándose en condiciones siempre variables. El número de elementos en una cola también puede establecerse mediante variable dinámica u otros factores. Las pruebas constantes, monitorización de recursos, registro de rendimiento actual, comparación con rendimientos anteriores, determinación de requisitos de carga de sistema y otros enfoques inteligentes ofrecen una potencia significativa para determinar cuánto debe tardar cada marca para el procesamiento (P), procesamiento posterior (Q) y retardo (R). Y saber cuánto debe tardar cada elemento que va a procesarse durante un ciclo llamado por una marca permite el establecimiento de un límite de elementos que van a procesarse durante ciclos de marca posteriores.

El gestor de marcas sobre la marcha ajustable y las métricas pueden indicar salud de sistema, identificar la necesidad de mejoras y también pueden incorporarse en una configuración de procesamiento distribuido para extraer recursos a demanda según se requiera y estén disponibles. El análisis algorítmico sobre la marca rige cuántos elementos pueden procesarse equilibrando recursos disponibles con cuánto pueden procesar físicamente esos recursos. El resultado es la predictibilidad y fiabilidad.

Los intervalos entre marcas también pueden ajustarse en consecuencia basándose tanto en requisitos de usuario dinámicos como en la realidad contextual de la situación.

Este gestor de marcas altamente controlado a granularidades de tiempo muy finas tales como ms o μ s resulta ventajoso por varios motivos, tales como;

- Impide solapamientos en el procesamiento por lotes durante llamadas de marca fijas.
- Pueden detectarse y controlarse procesos fugados.
- Pueden identificarse ineficiencias de código, proporcionando a los desarrolladores una infraestructura para usar como base a partir de la cual investigar y diagnosticar y, lo que es más importante, resolver cuestiones.
- Esto ofrece un método para usar hardware (HW) de manera más eficiente.
- Y más ventajas resultan evidentes para los expertos en la técnica.

La granularidad de una marca forma la base de intervalos de tiempo fiables basándose en los cuales pueden construirse otras aplicaciones de software, tales como para una baliza de información de red establecida para emitir pulsos con una regularidad de X ms, o un salto de lanzamiento, u otro tipo de aplicación sensible al tiempo.

Los datos de monitorización también pueden ponerse a disposición en la capa de software de modo que puede adaptarse sus operaciones a las condiciones actuales y disponer de información relacionada con el tiempo para usarse a un nivel muy alto.

La opción de ciclos de tiempo fijos o ciclos de tiempo ajustables puede resultar beneficiosa en diferentes situaciones. Al tener un gestor que ejecuta y monitoriza ciclos, durante el uso pico en el que los ciclos se ejecutan más tiempo de lo previsto, durante tiempos menos exigentes, los ciclos posteriores pueden ejecutarse menos tiempo llevando los ciclos de vuelta a un intervalo regular para alcanzar la planificación cuando se retrasó anteriormente.

Este gestor y sus datos son fáciles de integrar como gestor de flujo en una arquitectura existente para obtener beneficios. También existen otras aplicaciones y beneficios que pueden obtenerse al tener más control sobre la granularidad de una marca.

“Open Radio equipment Interface (ORI); ORI Interface Specification; Part 2: Control and Management (Release 4)”, GROUP SPECIFICATION, EUROPEAN TELECOMMUNICATIONS STANDARDS INSTITUTE (ETSI), 650, ROUTE DES LUCIOLES; F-06921 SOPHIA-ANTIPOLIS; FRANCIA, (01/10/2014), vol. ORI, n.º V4.1.1,

XP014223146, da a conocer una técnica mediante la cual puede subirse un archivo de registro desde un equipo de radio (RE) hasta un controlador de equipo de radio (REC).

El documento EP 1498809 A da a conocer que, tras recibir una petición de acceso a partir de un dispositivo cliente mediante una red interna, un controlador emite una instrucción de transferencia de datos a un dispositivo de almacenamiento en caché o un dispositivo de almacenamiento. El dispositivo de almacenamiento en caché o el dispositivo de almacenamiento, que ha recibido la instrucción de transferencia de datos, lleva a cabo una negociación directa requerida para la transferencia de datos con el dispositivo cliente según la instrucción, transfiriendo de ese modo directamente datos al dispositivo cliente.

Robert Russell, "Introduction to RDMA Programming", (17/04/2014), URL: web.archive.org/web/20140417205540/http://www.cs.unh.edu/~rdr/rdma-intro-module.ppt, (02/12/2015), XP055232895, proporciona algo de información referente a la técnica de acceso directo a memoria remoto (RDMA).

Sumario de la divulgación

La presente invención se refiere a un sistema para transferencia y procesamiento de archivos en un entorno de red tal como se expone en las reivindicaciones adjuntas.

Breve descripción de los dibujos

Con el fin de facilitar una mayor comprensión de la presente divulgación, ahora se hace referencia a los dibujos adjuntos, en los que se hace referencia a elementos similares con números o referencias similares. Estos dibujos no deben interpretarse como limitativos de la presente divulgación, sino que se pretende que sólo sean ilustrativos.

La figura 1 muestra una comparación de distancias para segmentos de un trayecto de datos a través de larga distancia.

La figura 2 ilustra un problema a modo de ejemplo con transacciones a larga distancia.

La figura 3 ilustra una ejecución de técnica convencional a modo de ejemplo en el mundo real.

La figura 4 ilustra ineficiencias de la técnica actual.

La figura 5 ilustra una topología de GVN, que incluye un segmento de estructura central a través de Internet o fibra oscura.

La figura 6 ilustra técnicas convencionales, archivo enviado como transferencia en múltiples partes mediante transporte por paquetes a través de TCP/IP.

La figura 7 ilustra una topología, potencia de HPC a través de un tapiz.

La figura 8 ilustra un flujo de trabajo a modo de ejemplo para archivo guardado en PFS a través de una larga distancia.

La figura 9 ilustra una replicación de REQ-RESP de cliente-servidor y DB de extremo posterior tradicional (múltiples maestros).

La figura 10 ilustra el lanzamiento de información mediante RDMA, directo a Db.

La figura 11 ilustra el lanzamiento de información mediante RDMA, archivos en cola.

La figura 12 ilustra el lanzamiento de información mediante RDMA, independencia de archivos en cola con verificación de DB cruzada mediante SRV ACC.

La figura 13 ilustra un caso de uso a modo de ejemplo de lanzamiento de información, retransmisión de información de mercado desde el mercado hasta la mesa de negociación.

La figura 14 ilustra un caso de uso a modo de ejemplo de lanzamiento de información, petición de ejecución de negociación desde la fuente de cliente hasta el mercado.

La figura 15 ilustra un caso de uso a modo de ejemplo de lanzamiento de información, resultado de negociación desde el mercado de vuelta al cliente.

- La figura 16 ilustra un flujo de trabajo a modo de ejemplo de información para escribir archivo y después leer, procesar, etc.
- 5 La figura 17 ilustra un flujo de trabajo a modo de ejemplo de información para escribir archivo.
- La figura 18 ilustra un flujo de trabajo a modo de ejemplo de leer archivo y procesar.
- 10 La figura 19 ilustra una realización a modo de ejemplo de lanzamiento de información, con respecto a cabeceras de archivo/secciones de archivo: cabecera, datos de cuerpo, pie.
- La figura 20 ilustra una realización a modo de ejemplo de lanzamiento de información, con respecto al disparo, almacenamiento en PFS y procesamiento de archivos a partir del almacén.
- 15 La figura 21 ilustra una realización a modo de ejemplo de lanzamiento de información, con respecto a la recogida de cola de archivos para recuperar datos de cuerpo para su análisis y uso.
- La figura 22 ilustra una realización a modo de ejemplo de lanzamiento de información, con respecto a datos de cuerpo de archivo y EOF.
- 20 La figura 23 ilustra una realización a modo de ejemplo de lanzamiento de información, con respecto a flujo de procesamiento de archivo.
- La figura 24 ilustra posibles problemas con la cola de extracción de archivos tales como procesamiento por duplicado, etc.
- 25 La figura 25 ilustra una realización a modo de ejemplo de sincronización de lotes extrayendo únicamente archivos no marcados como leídos durante una marca.
- La figura 26 ilustra un flujo de trabajo a modo de ejemplo de procedimientos de extremo posterior postcríticos (para escribir en remoto).
- 30 La figura 27 ilustra un flujo de trabajo a modo de ejemplo de procedimientos de extremo posterior postcríticos (para leer a partir de PFS y procesar).
- 35 La figura 28 ilustra un diagrama de sistema a modo de ejemplo para lanzamiento de información.
- La figura 29 ilustra una combinación estructurada de tejidos de red en capas para dar un tapiz de red.
- La figura 30 ilustra capas ordenadas a modo de ejemplo de diversos tejidos de red para dar un tapiz.
- 40 La figura 31 ilustra técnicas convencionales, CRONS.
- La figura 32 ilustra técnicas convencionales, bucles.
- 45 La figura 33 ilustra técnicas convencionales, recursión, secuencia de comandos que se llama a sí misma.
- La figura 34 ilustra problemas con bucles.
- La figura 35 ilustra $\Delta \text{ tiempo} = P + Q + R$.
- 50 La figura 36 ilustra el sincronismo de diversos ciclos en los que $X \neq Y$.
- La figura 37 ilustra el sincronismo de diversos ciclos en los que X es sistemática .
- 55 La figura 38 ilustra un umbral de distancia para obtener una ganancia a partir de salto de lanzamiento.
- La figura 39 ilustra la sincronización de lotes, con respecto a marcar archivo(s).
- La figura 40 ilustra la granularidad de una marca.
- 60 La figura 41 ilustra la granularidad de una marca, con respecto a un intervalo de tiempo variable.
- La figura 42 ilustra un algoritmo a modo de ejemplo, con respecto a ajuste de marca.
- 65 La figura 43 ilustra un análisis de inteligencia artificial, granularidad y periodos.

La figura 44 ilustra componentes a modo de ejemplo de marca, componentes.

La figura 45 ilustra un uso de marca a modo de ejemplo en ráfagas de baliza.

5 La figura 46 ilustra un uso de marca a modo de ejemplo en múltiples pulsos de baliza.

La figura 47 ilustra un uso de marca a modo de ejemplo en baliza en puntos de intersección.

La figura 48 ilustra un diagrama de sistema a modo de ejemplo de granularidad de una marca.

10

Descripción detallada

La figura 1 muestra una comparación de distancias para segmentos de un trayecto de datos a través de larga distancia. Este ejemplo de realización describe una línea dedicada entre dos puntos de extremo que conectan
15 redes en cada extremo. Es un error común considerar que las comunicaciones a través de larga distancia a través de fibra óptica se desplazan a la velocidad de la luz, sin embargo, en el mejor de los casos la velocidad de cableado auténtica tiene una eficiencia del 68% (véase la tabla 2 en el presente documento) en comparación con la velocidad de la luz. Para la transferencia de datos de comunicaciones a través de la distancia, por ejemplo, al pasar información financiera, realizar una petición de negociación, recibir una confirmación o algún tipo
20 equivalente de intercambio de información, se ha dedicado mucho esfuerzo a hacer que el trayecto sea lo más eficiente posible. Aunque pueden ser posibles mejoras mediante innovación continuada en zonas locales regionales, la cantidad de tiempo relativa (delta de tiempo (Δt)) ahorrada en una o ambas de estas zonas locales regionales es mucho menor en comparación con el tiempo a través de la WAN 0154 a larga distancia.

25 Este ejemplo de realización demuestra un enlace de una red en la región A 0100 con una red en la región B 0120, conectadas mediante un trayecto dedicado a través de una conexión de estructura central de fibra, FBB 0130. La FBB 0130 está conectada en ambos extremos mediante dispositivos de punto de extremo (EP) 0102 y EP 0122 que actúan como anclajes en cada extremo 0132 y 0138 del puente a través de la FBB 0130.

30 Las distancias más cortas están dentro de cada región, región A 0150 local y región B 0152 local. La distancia más larga es a través del segmento 0154 de WAN a larga distancia y es a través de este segmento donde pueden realizarse las mayores mejoras (por ejemplo, método de transporte, distancia más significativa a través del segmento más largo, FBB 0130).

35 Según el sincronismo de extremo a extremo, la transferencia de datos desde un dispositivo en la red en la región A 0100 hasta dispositivos en una red en la región B 0120 consiste en tres bloques de segmentos de red. La región A 0100 local desde los dispositivos en 0100 hasta el punto 0102 de extremo a través del trayecto 0106 de comunicaciones.

40 La larga distancia entre regiones es desde el EP 0102 hasta la FBB 0130 a través del trayecto 0132 de comunicaciones y en el otro extremo a través de 0138 hasta el EP 0122. Finalmente, el segmento 0126 de red local desde el EP 0122 hasta los dispositivos en la región B 0120.

45 En una situación tal como transmisión de información o instrucción de negociación financiera, el segmento de WAN 0154 a larga distancia puede representar una porción de tiempo significativa y principal requerida para la transmisión de extremo a extremo, de 0100 a 0120. Por tanto, resulta esencial mejorar la conectividad del segmento de WAN 0154 a larga distancia.

50 La figura 2 ilustra un problema a modo de ejemplo con transacciones a larga distancia. Esta figura es un ejemplo de realización basado en la figura n.º 1, que extiende el trayecto en ambos extremos de la red a partir del EP 02100 y el EP 02200 a través de los trayectos 02302 y 02306.

55 En cada extremo en la región A donde está ubicado el EP 02100 y la región B donde está ubicado el EP 02200, conecta el EP a una LAN 02102/02202 y también ilustra un punto 02106/02206 de salida y entrada (EIP) en cada LAN con Internet 02108/02208 abierto en el EIP 02106 y el EIP 02206 a través de trayectos de conexión tales como 02116 y 02216, siguiendo una metodología convencional de pasarela al punto de presencia (POP) a Internet ampliamente usada en la práctica actual.

60 Por ejemplo, cada LAN 02102/02202 y zona 02108/02208 de Internet ilustrada tiene diversos dispositivos conectados a esas redes. En el caso de cada LAN, por ejemplo la LAN 02102, hay dispositivos conectados a la misma tales como un servidor SRV_A 02122, un cliente C_A 02126 y un nodo de sistema de archivos en paralelo (PFS) PFS 02128 a través de trayectos 02152, 02156 y 02158 respectivamente. Internet, por ejemplo Internet 02108, tiene dos servidores SRV_A 02132 y SRV_A 02136 conectados al mismo a través de los trayectos 02162 y 02166 de comunicaciones.

65

En un caso de uso real, puede haber la misma cantidad de, o más o menos, dispositivos conectados a cada zona

de red LAN 02102, Internet 02108, LAN 02202, Internet 02208, con respecto a lo que se describe en el presente documento. La cantidad y tipos de estos dispositivos conectados sólo son con fines ilustrativos.

Pueden usarse otros dispositivos tales como enrutadores, cortafuegos, conmutadores gestionados, servidores y otros tipos de dispositivos que pueden regir cómo puede fluir el tráfico dentro de, y entre, diversas redes.

La figura 3 ilustra una ejecución de técnica convencional a modo de ejemplo en el mundo real. Este ejemplo de realización describe la colocación de, y conexiones entre, nodos de lanzamiento ubicados en algunos de los principales mercados financieros del mundo, en tres regiones: Norteamérica, Europa y Asia.

Tal como se describe en el recuadro 03000 de leyenda en la parte inferior derecha, cada mercado financiero indicado en el presente documento es desde una perspectiva de conexión en red descrita como nodo 03008 global. Alrededor del nodo global, dos anillos designan el tipo de zona de calidad de conectividad, por ejemplo con fines de un radio desde el centro en el que está ubicado el mercado o intercambio. Esto es únicamente por simplificación, ya que muchos factores determinan el tamaño y la forma de estas zonas. Sin embargo, las dos zonas pueden distinguirse una de otra, siendo la más cercana una zona 03006 de alto rendimiento y siendo la otra una zona 03002 de servicio óptimo.

Los mercados están conectados entre sí mediante enlaces de red de alto rendimiento para compartir información de precios y también facilitar el procesamiento de pedidos de negociación y proporcionar confirmación de transacción de negociaciones, y otra información desde ubicaciones remotas hasta un mercado en otra ubicación.

Cuanto más lejos está el mercado desde un cliente o servidor u otro tipo de dispositivo que realiza una consulta, más tarda la información en fluir, incluyendo precios, volumen u otra información de mercado en tiempo real, peticiones de transacción, confirmaciones de cumplimiento tanto si se cumple el pedido como si no, u otros tipos de intercambio de datos.

Las zonas indicadas en el presente documento son:

- NYC 03100 para Nueva York, NY, EE.UU.
- LDN 03110 para Londres, Inglaterra, R.U.
- TOK 03120 para Tokio, Japón
- HKG 03130 para Hong Kong, SAR, China

Hay muchos otros mercados en todo el mundo que son significativos, pero los mencionados se indican con fines ilustrativos. También hay trayectos indicados entre cada mercado tales como 03200 entre NYC 03100 y LDN 03110. En realidad, hay una multitud de trayectos que representan cables submarinos entre los dos puntos. Se pretendía que esto simplificara el ejemplo ilustrado en el presente documento.

Cuanto menor es la latencia entre los puntos, más rápidamente puede conocerse la información de mercado y más rápidamente puede ejecutarse y procesarse la información relacionada con la transacción.

La figura 4 ilustra ineficiencias de la técnica actual. Esta figura describe las técnicas convencionales de formación de paquetes de datos transmitidos a través de una red de IP.

El trayecto de comunicaciones descrito en el presente documento es desde el inicio 04110 hasta el final 04210.

La LAN_A 04100 está conectada a la WAN 04000 y a la LAN_B 04200 como un flujo de procesamiento que ilustra el flujo de información en paquetes. La información se prepara en un dispositivo en la etapa 04112 y se envía a un servidor o a través de un servidor 04120. Esta figura demuestra el envío de un flujo de paquetes en un procedimiento de envío de IP/recepción de IP.

Los trayectos 04322, 04332, 04352, 04522, 04532, 04552, 04622, 04632, 04652 son el envío de uno de un número múltiple de paquetes de transmisión en múltiples paquetes. Los trayectos 04326, 04336, 04356, 04526, 04536, 04556, 04626, 04636, 04656 son la confirmación de paquetes que indica que se recibió. Un paquete perdido no llegará al dispositivo de destino y, por tanto, esta confirmación no se enviará de vuelta al dispositivo de origen. También es posible que un paquete llegue al destino pero que la confirmación se pierda en tránsito.

Se toma el tamaño de paquete total permitido menos el tamaño de cabecera(s), menos el tamaño de pie(s), y eso es igual al tamaño más grande permitido para la carga útil de un único paquete.

Para calcular el tamaño de carga útil total:

tamaño de carga útil = tamaño de carga útil permitido – tamaño de cabecera – tamaño de pie

El número total de paquetes necesarios para enviar un archivo de datos con un tamaño X es:

$$\text{Número de paquetes} = \text{INT.rounded.up} \left(\frac{\text{tamaño de datos que van a enviarse}}{\text{tamaño de carga útil calculado}} \right)$$

Por tanto, un tamaño de datos más grande se necesita dividir en paquetes para dar un flujo de paquetes.

Esta figura ilustra un ejemplo en el que el paquete puede retransmitirse entre dispositivos en vez de simplemente pasar los paquetes donde, en más de una etapa, se descomponen los datos para dar diversos paquetes y después vuelven a ensamblarse para dar un archivo de datos una vez que se han recibido todos los paquetes en el procedimiento ilustrado en envío de IP de paquete 04116 a recepción de IP de paquete 04118.

Por ejemplo, el SRV_A 04120 puede consultar un mercado u otra fuente de datos para obtener información y la guardará localmente en memoria intermedia antes de replicar esa información mediante de envío de IP de paquete 04006 a recepción de IP de paquete 04008 al SRV_B 04220.

Y posteriormente, un cliente puede procesar información 04212 consultando al SRV_B 04220 mediante de envío de IP de paquete 04216 a recepción de IP de paquete 04218.

El punto clave es que la división en paquetes de los datos para dar cargas útiles de tamaño de paquete permitido y después la recombinación en el destino o el origen pueden producirse múltiples veces a lo largo de un trayecto de comunicaciones.

La figura 5 ilustra una topología de GVN, que incluye un segmento de estructura central a través de Internet o fibra oscura. En la solicitud de patente internacional n.º PCT/US15/64242, "SYSTEM AND METHOD FOR CONTENT RETRIEVAL FROM REMOTE NETWORK REGIONS", hay una característica en la que se agrupan más de un archivo entre sí para dar un archivo más grande para enviarse mediante una transferencia de archivo a través de "caché en cadena" desde una región geográfica hasta otra región geográfica. Para que esta característica sea ventajosa, se necesita que la transferencia de archivo sea lo más rápida posible. Como transporte para una agrupación de diversos "archivos" de carga útil de datos, el método de lanzamiento de información de esta divulgación mueve un bloque de datos más grande desde un extremo del mundo hasta el otro más rápidamente que métodos de técnicas convencionales.

Esta divulgación puede usarse dentro de la topología descrita en la figura 5 de la siguiente manera:

El SRV_BBX 05508 en la región o zona ZD2 05500 puede conectarse al SRV_BBX 05608 en otra región o zona ZD3 05600 a través de una conexión 05818 de fibra oscura a través del tejido 05808 de fibra oscura.

El SRV_BBX 05508 usa esta divulgación para escribir directamente un archivo en el almacenamiento de archivos en paralelo, PFS 05620, a través de acceso directo a memoria remoto (RDMA) a través de 05818 evitando la pila del SRV_BBX 05608 a través del trayecto 05650.

El SRV_IBX 05608 usa esta divulgación para escribir directamente un archivo en el almacenamiento de archivos en paralelo, PFS 05520, a través de acceso directo a memoria remoto (RDMA) a través de 05818 evitando la pila del SRV_IBX 05508 a través del trayecto 05550.

El trayecto 05812 puede ser IPv4 o alguna clase de protocolo de Internet normalizado a través del cual fluye tráfico desde el SRV_AP 05510 hasta y/o desde el SRV_AP 05610 a través del trayecto 05812 en transmisión libre de la GVN 05802 a través de un túnel u otro tipo de trayecto de comunicación.

Esta realización ilustra que pueden combinarse diversos tipos de tejidos de red para dar un tapiz de red más grande. Estos tejidos pueden entretorse sin interrupciones entre sí tal como se describe en la solicitud de patente provisional estadounidense n.º 62/174.394. El lanzamiento de información puede o bien ser un método autónomo o bien integrarse como un segmento de red dentro de un trayecto de red más grande compuesto por diversos segmentos de red.

Este ejemplo de realización ilustra la topología de una red virtual global (GVN), sus diversos dispositivos, trayectos de comunicaciones y otras realizaciones. Muestra cómo diversas regiones geográficas o zonas o territorios están conectados entre sí a través de diversos tipos de trayectos.

Leyenda:

ZL0 – Zona de LAN cero	ZL1 – Zona de LAN uno	ZI0 – Zona de Internet cero
ZI1 – Zona de Internet uno	ZI2 – Zona de Internet dos	ZI3 – Zona de Internet tres
ZD2 – Zona de centro de datos de Internet dos	ZD3 – Zona de centro de datos de Internet tres	

La figura 6 ilustra técnicas convencionales, archivo enviado como transferencia de múltiples partes mediante transporte por paquetes a través de TCP/IP. Como extensión de la figura 4, cuando una carga útil de datos contiene un archivo de datos estructurado que es más grande que el tamaño permitido de la carga útil, puede enviarse mediante transferencia de archivo en múltiples partes de IP.

En este caso, se requiere cabecera adicional y posiblemente más información de pie para indicar qué fragmento del archivo de múltiples partes está contenido en la sección de carga útil del paquete. Cuando se “guarda” un archivo, se hace en un sistema de archivos jerárquico (HFS) 06120, 06320 y el envío se realiza a partir de la lectura del archivo local guardado a partir de este HFS y la formación de paquetes de una transferencia de archivo en múltiples partes a través de tantos paquetes como se requiera a través de enviar archivo 06200. A medida que se recibe el archivo, el procedimiento de rec. archivo 06220 leerá cada paquete, se asegurará de que están en orden y después tomará cada parte de la transferencia de múltiples partes y volverá a juntar el archivo parte por parte.

El archivo nuevamente ensamblado se guarda en el HFS 06320 en el dispositivo 06310 anfitrión.

La figura 7 ilustra una topología, potencia de HPC a través de un tapiz. Este ejemplo de realización describe la conexión de dos EPD 07100 y 07200 con un SRV 07300 dentro de una topología de red tal como una red 07000 virtual global (GVN).

Es una simplificación de esta red, pero está diseñada para ilustrar cómo un PFS puede distribuir dispositivos de almacenamiento a través de cuatro diversas regiones.

Cada trayecto de comunicación en el tapiz puede ejecutar diversos protocolos de tejido. Puede accederse a n almacenamiento de archivos en un PFS tal como el PFS 07810, el PFS 07812, dentro de una LAN 07120 o LAN 07812, o el PFS 07800 o el PFS 07802 en la nube, a través de acceso directo a memoria remoto (RDMA) mediante un protocolo tal como InfiniBand (IB) u otro protocolo que permitirá el acceso de RDMA nativo a nodos distribuidos en un sistema de archivos en paralelo (PFS).

El punto clave es que puede lograrse un pilar de informática de alto rendimiento (HPC) mediante una topología proporcionada por una GVN

La figura 8 ilustra un flujo de trabajo a modo de ejemplo para archivo guardado en PFS a través de una larga distancia. Esta figura describe el guardado de un archivo 08100 desde un dispositivo 08000 hasta un PFS 08200 en otra ubicación mediante RDMA y la posterior recuperación del archivo a partir del PFS 08200 mediante otro dispositivo 08300.

Las etapas en el flujo desde el inicio 08000 que puede ser un dispositivo que genera un archivo, que pasa un archivo, que extrae un archivo o que gestiona de otro modo un archivo tal como el archivo 08100. El trayecto 08120 describe el guardado de un archivo directamente en un dispositivo de PFS mediante RDMA con una escritura directa en el dispositivo de nodo de almacenamiento desde una región hasta otra región.

La etapa de recuperar archivos de la cola 08300 en el PFS 08200 a través del trayecto 08220 puede realizarse de manera local o también remota con respecto a la región en la que está ubicado el dispositivo de PFS.

El fin 08600 implica el acceso al archivo tal como leer, cargar o usar de otro modo el archivo.

La figura 9 ilustra una replicación de REQ-RESP de cliente-servidor y DB de extremo posterior tradicional (múltiples maestros). Este ejemplo de realización describe la infraestructura de petición/respuesta de IP actual a través de los trayectos de REQ1 09500 y RESP2 09502.

Los archivos se guardan en un HFS local tal como el HFS 09040 por el SRV_APP_1 09020 y en el HFS 09240 por el SRV_APP_2 09220. El SRV_APP puede representar un servidor de aplicación u otro tipo de dispositivo y se marca con fines ilustrativos. La WAN 09100 puede ser LAN/Internet/línea dedicada u otro tipo de conexión de red.

El n.º de MMDBR se refiere a la replicación de base de datos de múltiples maestros. En un despliegue de este tipo, la definición de la tabla de base de datos y estructura de campos son equivalentes e idénticas. Cada anfitrión tiene un número inicial en filas con un valor de desviación de incremento automático igual al número de anfitriones. Por ejemplo, en el caso de tan sólo dos bases de datos conectadas a los anfitriones maestros, DB_1 tendrá los números de filas 1, 3, 5 y así sucesivamente. Y DB_2 tendrá la numeración de ID de fila de 0, 2, 4, 6, 8, 10, etc. También puede implementarse otra metodología frente a colisiones.

La replicación de base de datos puede ser un procedimiento automatizado o manual u otro en el que el contenido de cada una de la DB_1 09050 y la DB_2 09250 se replica en la otra. Para proteger datos frente a la sobrescritura, los desplazamientos de fila logran esta protección garantizando que no hay ID de fila idénticos entre los dos lados.

La figura 10 ilustra el lanzamiento de información mediante RDMA, directo a Db. Este ejemplo de realización explora un enfoque para intentar acelerar las comunicaciones. Esto puede consistir en enviar una tabla de base de datos entera y contenido mediante RDMA directamente a los medios de almacenamiento para una base de datos en una región remota. Esto es bueno para la estructura de tabla y todas las filas de datos de una base de datos estática, no en uso o bloqueada de otro modo. Esto puede usarse cuando se realiza la replicación en un esclavo fuera de línea antes de (volver a) ponerse en línea.

Sin embargo, este enfoque no es eficiente para una única o un número pequeño de filas y también puede conducir a problemas con la sincronización.

Por tanto, antes de usar este método, se requiere un bloqueo de la base de datos durante la transferencia en la MMDBR1 10510 y la MMDBR2 10512 para enviar mensajes entre las Db 10050 y 10250 para asegurarse de que están sincronizadas. Esto añade resistencia no deseada al sistema.

Este ejemplo puede no ser un uso ideal de lanzamiento.

La figura 11 ilustra el lanzamiento de información mediante RDMA, archivos en cola. Un servidor de intercambio troncal (SRV_BBX) infunde atributos de muy buena conectividad en un gran conducto y tiene una fuerte potencia de desplazamiento.

Esta figura ilustra un SRV_BBX_1 11080 que escribe directamente un archivo mediante RDMA_1 11600 en el PFS 11802, y un SRV_IBX_2 12280 que envía de manera cruzada un archivo diferente mediante RDMA_2 11610 en el PFS 11800.

Las cargas útiles de datos se transfieren de manera eficiente. Se mantiene un registro de la transferencia en cada lado indicando respectivamente qué archivos se recibieron. La replicación de DB de múltiples maestros se realiza con la DB_1 11050 usando filas impares y la DB_2 11250 usando filas pares. Este ejemplo se basa en el mecanismo de replicación de base de datos para mantener un seguimiento de acontecimientos de replicación de múltiples maestros. Por tanto, pueden producirse problemas de sincronización. También hay posibilidad de un retraso significativo entre el tiempo en el que puede ejecutarse una escritura remota por un dispositivo y un largo retardo para obtener confirmación, lo cual puede ser en sí mismo significativamente perjudicial. Por tanto, existe un diferencial de velocidad así como un lapso de fiabilidad que deben superarse entre el acoplamiento de escrituras de RDMA superrápidas y la replicación de información de base de datos tradicional.

Este ejemplo de realización no ilustra la implementación ideal del lanzamiento. El guardado en dispositivos de PFS remotos es eficiente, pero hacer que el SRV_APP en la región objetivo tenga conocimiento del archivo no es lo suficientemente rápido ni fiable.

La figura 12 ilustra el lanzamiento de información mediante RDMA, independencia de archivos en cola con verificación de DB cruzada mediante SRV_ACC. Este ejemplo de figura ilustra el desacoplamiento completo de una necesidad de realizar una ida y vuelta de los registros de base de datos en un sentido activo. En este ejemplo, un SRV_BBX tal como el SRV_BBX_1 12080 enviará el archivo al PFS 12802 y el registro se almacenará en el SRV_DB_1 12030 en la base de datos local Db_1 12050. Y no se realizará nada más.

Debido a la fiabilidad de RDMA, una vez guardado el archivo en un nodo de almacenamiento de PFS, hay confianza de que se encuentra ahí. Los servidores de contabilidad (SRV ACC) tales como el SRV_ACC_1 12060 y el SRV_ACC_2 12260 se comunican entre sí para catalogar transferencias recibidas.

De nuevo, esto no es un enfoque ideal para el lanzamiento y apunta a la necesidad de desacoplar completamente el enfoque de lanzamiento con respecto a métodos actuales desplegados a través de IP y abarcar la potencia del lanzamiento.

La figura 13 ilustra un caso de uso a modo de ejemplo de lanzamiento de información, retransmisión de información de mercado desde el mercado hasta la mesa de negociación. Como ejemplo de lanzamiento, en la industria financiera, puede pedirse y recuperarse información de mercado, tal como precios y volumen, a través de TCP/IP, o puede enviarse mediante una multidifusión de UDP/IP desde un origen hasta un destino. Se requiere un tiempo de ida y vuelta, RTT, para la petición y recuperación de TCP/IP. Un envío de UDP/IP sólo es unidireccional para que la información sea útil (la confirmación para el remitente no es tan importante), sin embargo, todavía existe una resistencia significativa debido a ineficiencias de IP y transferencia en paquetes.

Usando tecnologías de IP en paquetes actuales, ni siquiera un enfoque de multidifusión de UDP/IP es el más eficiente.

- 5 Esta figura contiene ejemplos de realizaciones que describen la recopilación de información por un dispositivo DO 13100 en una región, la región A 13000, y después el flujo hacia delante de transmitir esa información a un dispositivo en la región B 13200.

- 10 La etapa de recopilar información de mercado 13120 a preparar contenido del archivo 13130 al lanzamiento del archivo al PFS remoto 13180 se realizan en el DO 13100. El lanzamiento mediante 13882 al nodo de sistema de archivos en paralelo (PFS), PFS 13880, en el dispositivo de almacenamiento S2 13800. Hay numerosas ventajas de usar RDMA a través de IB 13882 o tipo de red equivalente frente a transferir archivos mediante TCP/IP o UPD/IP a través de Ethernet, principalmente debido al hecho de que no se necesita dividir el archivo en paquetes, pero también debido a cómo se gestionan los paquetes, y la eficiencia de RDMA a través de IB.

- 15 El archivo guardado en el PFS 13880 se recupera en el dispositivo D2 13300 mediante el procedimiento de recuperar archivos de uno en uno a partir del PFS 13380 a través del trayecto 13382. Puede recuperarse simultáneamente un único o un lote de múltiples archivos.

- 20 Las siguientes etapas 13332 a 13325 a 13320 describen el uso del archivo en la región remota con respecto al origen del archivo. No hay necesidad adicional de que los dispositivos D0 13100 y D2 13300 se comuniquen entre sí con respecto a la planificación o la transferencia o el uso del propio archivo.

- 25 La figura 14 ilustra un caso de uso a modo de ejemplo de lanzamiento de información, petición de ejecución de negociación desde la fuente de cliente hasta el mercado. Este ejemplo de figura ilustra adicionalmente el lanzamiento en el contexto de un caso de uso de la industria financiera que demuestra la simple eficiencia de poner un pedido de negociación a partir del dispositivo D2 14300 para usarse en última instancia por el dispositivo D1 14100. Empieza desde 14310 y termina en 14100. Es la inversa del trayecto descrito en la figura 13.

- 30 El punto clave es que ambos dispositivos funcionan independientemente uno de otro. Sólo cuando el D2 14300 desea enviar una instrucción para que se ejecute una negociación en el D1 14100 entonces hay un acontecimiento de comunicación, pero esto tampoco es directamente desde el D2 14300 hasta el D1 14100 sino para que el D2 14300 guarde uno o más archivos de datos en el sistema de archivos en paralelo (PFS) 14000. Una vez escrito el archivo, el D2 14300 puede pasar a otras cosas.

- 35 El D1 14100 tiene un procedimiento automatizado que se ejecuta según un intervalo de tiempo. Véanse las figuras 24 y 25 referentes a este intervalo. Este procedimiento automatizado buscará en el PFS 14880 para comprobar 14180 si hay nuevos archivos no marcados (no marcados como leídos) para recuperarlos y procesarlos.

- 40 Este enfoque es eficiente para pedidos de negociación tales como una petición de ejecución de negociación. El dispositivo D1 puede estar lo más cerca posible del mercado. El dispositivo D2 puede ser el terminal de cliente o un servidor que está en proximidad muy estrecha a los clientes que acceden al mismo. El punto clave es que la carga útil de negociación puede enviarse lo más cerca posible de la velocidad de cableado.

- 45 Esto resulta ventajoso y, dado que puede enviar un gran tamaño de archivo, este mecanismo puede usarse para enviar/recibir cadenas de bloques de altura cada vez mayor (tamaño de datos), algo-bots (algoritmos inteligentes que van a ejecutarse de manera remota) y otras cargas útiles más grandes que el tamaño de paquete.

- 50 La figura 15 ilustra un caso de uso a modo de ejemplo de lanzamiento de información, resultado de negociación desde el mercado de vuelta al cliente. Este ejemplo es muy similar a las figuras 13 y 14 ya que ambos dispositivos D1 15100 y D2 15300 están funcionando de manera independiente. Las comunicaciones entre los mismos se realizan mediante un archivo 15882 enviado en un PFS 15880 por un dispositivo 15100, y para recuperar este archivo por un procedimiento 15380 que se ejecuta en el otro dispositivo 15300. Están en comunicación pero con todo el sobredimensionamiento eliminado del intercambio.

La diferencia en esta figura son ejemplos de realizaciones de confirmación a partir del mercado 15120 y notificación de ejecución de negociación 15320.

- 60 El D1 15100 y el D2 15300 pueden producirse ambos dentro del mismo dispositivo o mediante operaciones distribuidas entre diferentes dispositivos en una zona representada por el D1 15100 y el D2 15300.

- 65 La figura 16 ilustra un flujo de trabajo a modo de ejemplo de información para escribir y después leer, procesar, etc., un archivo. Esta figura simplifica el lanzamiento en un flujo de procesamiento lineal y es genérica sin una referencia específica a la industria financiera ni ninguna otra.

Esta figura empieza el procedimiento a partir del inicio 16000 a través del trayecto 16010 hasta la etapa de información 16100 para evaluar y preparar la información 16200 a través de 16110 o bien como copiado de un archivo o bien para estructurar la información para dar un archivo específicamente diseñado para guardarse de manera remota mediante RDMA en un PFS 16300 a través de 16210. Después, este archivo guardado se recuperará mediante el procedimiento de recuperar archivo a partir de la cola en el PFS 16500 a través del trayecto 16310. El contenido del archivo se analizará sintácticamente mediante la etapa de analizar sintácticamente el contenido del archivo 16600 a través del trayecto 16510. La etapa de fin 16900 a la que se llega a través del trayecto 16610 implica que puede usarse el archivo. El punto clave es que este procedimiento puede simplificarse en unas pocas etapas clave.

La figura 17 ilustra un flujo de trabajo a modo de ejemplo de información para escribir un archivo. Esta figura se basa en la figura 16. Simplifica adicionalmente el flujo mostrando únicamente las etapas desde la información 17100 hasta guardarse el archivo en el PFS 17300.

La figura 18 ilustra un flujo de trabajo a modo de ejemplo de leer y procesar un archivo. Esta figura también se basa en la figura 16. También simplifica el flujo, pero esta vez sólo muestra las etapas desde el archivo en el PFS 18300 que se recupera mediante el procedimiento de recuperar archivo a partir de la cola en el PFS 18500. El contenido del archivo se analizará sintácticamente mediante la etapa de analizar sintácticamente el contenido del archivo 18600 y terminará en el fin 18900.

La figura 19 ilustra una realización a modo de ejemplo de lanzamiento de información, con respecto a cabeceras de archivo/secciones de archivo: cabecera, datos de cuerpo, pie. Este ejemplo de realización de lanzamiento de información describe un archivo de datos organizado en tres secciones definidas; información 19100 de cabecera, carga 19200 útil que contiene datos de cuerpo, y un pie 19300.

Este archivo puede almacenarse en RAM, memoria, guardarse en un disco o almacenarse en otro tipo de memoria o almacenamiento en un dispositivo o de otro modo.

La cabecera 19100 puede contener información sobre el origen de anfitrión, destino de anfitrión, sello de tiempo y otra información.

Puede almacenarse información de seguridad en campos tanto en la cabecera 19100 como en la sección 19300 de pie. Esta información de seguridad puede contener referencias a claves usadas para el descifrado de la carga 19200 útil, así como otra información. La carga útil (datos de cuerpo) puede cifrarse en su totalidad o en parte o enviarse sin cifrarse.

Se usa(n) suma(s) de comprobación de carga útil en el pie para validar la integridad de los datos de cuerpo.

Otras características pueden aprovechar la flexibilidad ofrecida.

Una ventaja significativa es que puede haber un algoritmo en la carga útil que se envía como parte de la carga útil.

La indicación de EOF en el pie 19300 indicará que el archivo ha llegado, está completo y listo para validarse/verificarse para determinar la precisión y después usarse en última instancia.

En el alcance de la industria financiera, un algoritmo ejecutado de manera remota en la carga 19200 útil, por ejemplo, puede contener una condición de salida que se indica para realizar una acción indicada si cambian las condiciones de mercado con instrucciones algorítmicas para evaluar la dirección del cambio de mercado y después modificar las instrucciones en consecuencia, para:

- cancelar

- revertir

- intensificar

- o cambiar de otro modo las instrucciones.

La figura 20 ilustra una realización a modo de ejemplo de lanzamiento de información, con respecto al disparo, almacenamiento en PFS y procesamiento de archivos a partir del almacén. Esta figura combina los principios establecidos e ilustrados en la figura 8 y las figuras 12 a 19. La adición de un gestor de cola de archivos (FQM) 20810 gestiona los intervalos para extracciones de lotes de archivos, comprobaciones de archivos para determinar su completitud y realiza otras acciones relacionadas con archivo.

La conectividad del FQM con el archivo 20888 en el PFS 20880 se realiza a través del trayecto FQM 20386 y la

coordinación entre el FQM y la recuperación de archivo a partir de la cola en el PFS 20380 se realiza a través del trayecto de FQM 20388.

5 Esta figura también examina cada sección, la cabecera 20852, la carga 20856 útil y el pie 20858 de un archivo 20850 para demostrar ejemplos de campos. En la cabecera 20862, la carga 20866 útil, el pie 20868.

Estos ejemplos de campos se demuestran como pares de información de clave = valor y se realizan únicamente con fines ilustrativos. Los datos pueden almacenarse en otra estructura de datos y los campos pueden ser diferentes de los descritos en esta figura.

10

La figura 21 ilustra una realización a modo de ejemplo de lanzamiento de información, con respecto a la recogida de cola de archivos para recuperar datos de cuerpo para su análisis y uso. Esta figura empieza en 21000 y demuestra la extracción de un lote de archivos a partir del PFS 21800 por el gestor 21100 de cola de archivos en la RAM 21900. Los archivos extraídos mediante el procedimiento de recogida 21120 se comprueban para determinar su completitud 21300. Los archivos 21802, 21804, 21806, 21808 completos que han llegado se ponen en una lista en 21380 y los archivos 21814 incompletos se ponen en una lista en 21340. Y aquí es donde el flujo de procesamiento de esta figura termina en 21990. La implicación es que los archivos 21340 incompletos, tales como el archivo 21814, se recogerán por un procedimiento posterior cuando estén completos.

15

20

La figura 22 ilustra una realización a modo de ejemplo de lanzamiento de información, con respecto a datos de cuerpo de archivo y EOF.

Esta figura es una representación algorítmica de lo que se describió en las figuras 20 y 21.

25

La puerta de decisión de comprobar para ver si cada archivo está completo 22300 implica que analizará cada uno de los archivos extraídos por el gestor 22100 de cola de archivos a partir del PFS 22800. Si un archivo 22802, 22804, 22806, 22806 está completo, se procesará a través del trayecto de sí BFP 22510 y se usará en última instancia. Si no está completo 22410, entonces el archivo se ignorará 22440 y su procesamiento terminará a través de BFP 22448 en el fin 22990.

30

El fin 22990 implica el final de las operaciones del gestor 22100 de cola de archivos.

35

La figura 23 ilustra una realización a modo de ejemplo de lanzamiento de información, con respecto a flujo de procesamiento de archivo. Esta figura es una continuación de la descripción de lo que sucede en la figura 22, con elementos añadidos tales como la etapa de extraer y marcar cada archivo en la lista de archivos 23590 a procedimientos simultáneos de procesar archivo 23610 mediante la ruta de archivo FR 23620, procesar archivo 23612 mediante la ruta de archivo FR 23622, procesar archivo 23616 mediante la ruta de archivo FR 23626, procesar archivo 23618 mediante la ruta de archivo FR 23628. También hay dos puntos de fin diferentes: el fin 23990 es donde terminan los archivos completos después de procesarse mediante los trayectos de archivo FP 23630, 23632, 23636, 23638, con un "indicador" de leído que indica que se ha leído el archivo, que estaba completo y se ha usado. Estos archivos no se extraerán mediante extracciones de archivos de lotes posteriores 23120.

40

45

El fin 23940 implica que se ha ignorado un archivo porque está incompleto y 23400 la implicación es que volverá a leerse durante extracciones de lotes de archivos posteriores a partir del PFS 23800.

50

La figura 24 ilustra posibles problemas con la cola de extracción de archivos tales como procesamiento por duplicado, etc. Esta figura describe la extracción de lotes de archivos a partir del PFS 24800 a lo largo de un periodo de tiempo desde el inicio 24000 hasta el fin 24990. El tiempo se describe en ambos ejes con intervalos a lo largo de la parte superior y lotes a lo largo del lado izquierdo.

55

El intervalo A 24500 y la extracción de lotes A 24600 se producen al mismo tiempo. El uso 24602 se produce al final de este intervalo. El intervalo B 24510 y la extracción de lotes B 24610 se producen al mismo tiempo. El uso 24612 se produce al final de este intervalo. El intervalo C 24520 y la extracción de lotes C 24620 se producen al mismo tiempo. El uso 24622 se produce al final de este intervalo. Existe un problema por tener un intervalo que empieza justo después de haber terminado el otro porque puede no haber suficiente tiempo para marcar un archivo como leído, aunque ya se haya leído mediante un procedimiento por lotes anterior.

60

Por ejemplo, el archivo 25802 se extrae tanto por la extracción de lotes A 24600 como por la extracción de lotes B 24610. Esto también se produce de nuevo cuando el archivo 25808, que se leyó por la extracción de lotes B 24610, también se extrae por la extracción de lotes C 24620.

65

Esto es un defecto muy peligroso ya que, por ejemplo, una petición de negociación que se ejecuta dos veces puede provocar daño financiero u otro significativo debido a consecuencias no deseadas.

Marcar como leído puede implicar establecer un indicador en el registro de archivo en el dispositivo de

almacenamiento, o mover el archivo a una carpeta diferente, o eliminarlo del dispositivo de almacenamiento después de haberse movido, u otra manera de “marcarlo” de modo que los lotes posteriores sepan que el archivo se ha leído y “usado” por una extracción de lotes anterior.

5 La figura 25 ilustra una realización a modo de ejemplo de sincronización de lotes extrayendo únicamente archivos no marcados como leídos durante una marca. Esta figura es idéntica a la figura 24 en todos los aspectos excepto porque la figura 25 incluye un retardo A 25505 entre el intervalo A 25500 y el intervalo B 25510 (y un retardo A 25605 entre la extracción de lotes A 25600 y la extracción de lotes B 25610).

10 Hay un retardo B 25515 adicional entre el intervalo B 25510 y el intervalo C 25520 (y 25615 entre la extracción de lotes B 25610 y la extracción de lotes C 25620).

El punto clave es que este retardo permite que el lote actual evalúe y procese todos los archivos que ha extraído y, cuando ha usado archivos completos, marque esos archivos como leídos.

15 Este retardo añadido al mecanismo es totalmente dinámico y puede prolongarse si se requiere más tiempo de procesamiento para un lote o acortarse a un procesamiento de lote que se ha completado. Los tiempos de intervalo también pueden ajustarse dinámicamente basándose en varios factores. La granularidad de una marca describe $\Delta t = P + Q + R$. t = Delta de tiempo desde el inicio de la marca hasta el final de la marca, P = tiempo para el procesamiento por lotes de elementos para esta marca en el ciclo. Q = tiempo para cálculos de procesamiento de lotes posterior. R = tiempo para el retardo para garantizar que no hay ningún solapamiento entre elementos de lote y/o para establecer una desviación entre marcas.

25 La figura 26 ilustra un flujo de trabajo a modo de ejemplo de procedimientos de extremo posterior postcríticos (para escribir en remoto). Esta figura describe lógica de algoritmo del procedimiento de escritura en un PFS 26860 en una región remota. Comienza en el inicio 26000 para prepararse para ejecutar 26100 a través del trayecto CP 26010. Después el flujo pasa al punto 26120 de unión a través del trayecto CP 26110. En el punto 26120 de unión, el flujo se divide en dos corrientes, mediante los trayectos CP 26130 y CP 26150. Estas corrientes discurren de manera simultánea entre sí y son independientes.

30 Trayecto CP 26150 desde el punto 26120 de unión hasta un procedimiento de uso tal como escribir archivo en el PFS remoto 26860. La escritura se realiza mediante escribir archivo en PFS remoto 26300 a través del trayecto CP 26810.

35 El otro trayecto CP 26130 es para guardado y registro en registro de base de datos. El registro de Db se guarda localmente en la Db 26250 mediante el procedimiento de guardar en DB 26200. El registro registra las copias de seguridad u otros archivos relacionados que se almacenan en un HFS 26240 local mediante el procedimiento de guardar en registro 26220. La escritura en remoto termina en el fin 26998. El flujo administrativo de guardados de base de datos y registro termina en el fin 26990.

40 La figura 27 ilustra un flujo de trabajo a modo de ejemplo de procedimientos de extremo posterior postcríticos (para leer a partir de PFS y procesar). Este ejemplo de realización describe un algoritmo que rige la extracción de lotes 27100 y lectura de archivos a partir de PFS 27180 que van a procesarse de manera local. Cuando se lee un archivo, se comprueba para determinar la completitud 27200. Si está realmente completo, se marca como que se extrae 27188 de modo que un siguiente lote mediante el procedimiento 27100 no repite la lectura. Véase la figura 25 para ver cómo aborda el problema de que se extraen archivos duplicados tal como se describió en la figura 24.

50 En paralelo y de manera simultánea al procesamiento, el registro mediante la indicación de escritura de guardar en Db 27320 en la Db_1 27350 y guardar en registro 27600 registra los registros en el HFS 27650.

Sólo si la suma de comprobación calculada para la carga útil de un archivo en está el archivo completo 27200 es igual a la suma de comprobación incorporada en el pie, entonces puede marcarse el archivo como leído y correcto 27188.

55 El uso de este archivo 27220 implica que se hará algo con la información o instrucciones u otros datos enviados dentro de la estructura de archivo. Si es información de precios de mercado, se visualiza o procesa en un algoritmo para procesamiento. Si se pide una instrucción de negociación, entonces en 27220 se alimentará al mercado para su ejecución. Entonces puede escribirse de manera remota la respuesta mediante la misma metodología. Esto puede usarse en el sector financiero o para otras aplicaciones.

60 Un punto clave es que cada lote de archivos en una lista de ciclo de procesamiento de archivos extraída como lista en el inicio 27100. Y posteriormente se extrae cada archivo 27110 y se comprueba para determinar la completitud 27200. Si un archivo está completo, “sí” 27222, se usa 27220. La corriente se divide después de esta etapa. Marcar el archivo CP 27188 marca el archivo como leído en el PFS 27180. El trayecto CP 27302 conduce al punto 27300 de unión. El trayecto se divide aquí hacia CP27322 que conduce a guardar en una base de datos

27350 y un registro en el HFS 27650. El otro trayecto CP 27312 es para usar el archivo.

Por motivos de simplificación de esta demostración, la extracción de archivos 27110 se describe como una extracción en bucle. La etapa CP 27102 entre extraer la lista de archivos a partir del PFS 27100 y extraer un
 5 archivo a partir del PFS 27110 también puede dar lugar a extracciones de archivos paralelas y simultáneas que se ejecutan, cada una, al mismo tiempo, siguiendo este procedimiento.

Existen otras variaciones y posibilidades y los expertos en este campo podrán realizar ligeras variaciones a este algoritmo para lograr los resultados de guardado.

10 El punto clave es que pueden extraerse lotes de archivos, comprobarse individualmente cada archivo para determinar si está completo, si no está completo 27242, el archivo no se usa 27240 pero todavía se guarda 27320 en la Db 27350 y se registra 27600 en el HFS 27650. Si un archivo está completo 27222, se usa 27220, se marca 27188, se procesa 27512 y se guarda un registro de su extracción, comprobación, uso y otra
 15 información en la base 27350 de datos y también se registran los archivos en el HFS 27650.

La figura 28 ilustra un diagrama de sistema a modo de ejemplo para lanzamiento de información. Esta figura describe los diversos módulos que se ejecutan en diversos dispositivos que pueden usar el lanzamiento de información descrito en el presente documento. El punto 28880 de unión entre los dispositivos 28000 de EPD, el
 20 SRV_AP 28100, el SRV_CNTRL 28200 y el SRV_BBX 28300 y los dispositivos 28800, 28810, 28820 y 28830 de PFS.

Los servidores SRV_BBX también pueden conectarse entre sí, tal como el SRV_BBX 28300 y el SRV_BBX 28732 a través del trayecto 28738. Esta figura también incluye un módulo 28388 de nodo de lanzamiento en un
 25 SRV_BBX 28300. Este nodo de lanzamiento puede o bien residir en el SRV_BBX 28300 o bien un nodo de lanzamiento puede ser un dispositivo/servidor independiente que funciona en unísono con el SRV_BBX 28300.

La figura 29 ilustra una combinación estructurada de tejidos de red en capas para dar un tapiz de red. Esta figura representa diversos tejidos de red ilustrados como nubes independientes en capas en diferentes niveles. OTT¹
 30 representa una capa de primer grado en transmisión libre de una LAN 29002 o Internet 29000 o mecanismo 29020 de lanzamiento. Un tipo de red tal como GVN 29008 existe en OTT¹.

OTT² representa una transmisión libre de segundo grado y estos módulos se construyen basándose en una GVN 29008 tal como tejido 29010 de extremo a extremo de GVN de OTT² o mecanismo 29012 de cortafuegos de
 35 perímetro múltiple de MPFWM de OTT².

UTI¹ es la capa 29020 lanzamiento "bajo Internet" o en paralelo con Internet 29000. UTI² es donde existen los módulos 29030 de lanzamiento tales como salto de lanzamiento, ruta de lanzamiento, baliza de información y
 40 otra técnica relacionada.

En cada límite entre un tejido de red y otro hay puntos de salida y entrada (EIP). Los EIP, tales como el EIP 29052, implican un puente entre una LAN y una GVN, normalmente a través de un dispositivo de punto de
 45 extremo (EPD) en el borde de la LAN 29002 que construye un túnel seguro con un servidor de punto de acceso (SRV_AP) en la GVN 29008. Un EIP 29058 entre la GVN 29008 e Internet 29000 implica un EIP en un SRV_AP. Los EIP entre otras capas pueden representar otras topologías e implicar varios otros dispositivos.

Estos diversos tejidos de red, cuando se conectan mediante EIP, constituyen juntos un tapiz de red.

La figura 30 ilustra capas ordenadas de diversos tejidos de red para dar un tapiz. Esta figura está relacionada con la figura 29 presentada como bloques en capas. Los elementos añadidos son ruta + rendimiento 30115, construcción 30225, lanzamiento 30715 y lógica 30815 de lanzamiento.

Los nuevos elementos también son lógica 30050 de nivel 2 de GVN, lógica 30150 de nivel 4 de GVN, lógica 30650 de subnivel 1 de GVN y lógica 30750 de subnivel 2 de GVN. Estas capas lógicas son donde se usan
 55 determinados ajustes, operaciones, rutinas y otros métodos para alimentar los EIP entre capas.

La figura 31 ilustra técnicas convencionales, CRONS. Este ejemplo de realización describe técnicas convencionales de una tarea planificada por CRON. Las secuencias de comandos/programas llamados 31200 se controlan mediante el procedimiento de gestor de CRON, gestor 31100 de CRON.

60 Pueden establecerse tareas de CRON o bien manualmente añadiendo elementos a la lista 31108 de CRON o bien puede gestionarse esta lista mediante una capa de gestor de software. Existen algunos problemas relacionados con saber si este procedimiento está ejecutándose de manera apropiada o no. Y si sólo está ejecutándose un procedimiento tal como se pretende o si están ejecutándose muchas instancias paralelas
 65 generadas y no intencionadas. Se necesita que los elementos 31318 de entorno estén actualizados y funcionando dentro de los parámetros previstos. Se necesita que los elementos 31680 de base estén

ejecutándose de manera óptima y dentro de los parámetros previstos.

El tiempo 31600 de sistema/reloj interno tiene que estar actualizado 31360 y ser preciso hasta una granularidad de tiempo muy fina. Se necesita que esta regularidad de actualización sea más frecuente que la deriva a la que es propenso el sistema. El hardware y los controladores 31620 asociados tienen cierta dependencia de recursos disponibles.

La figura 32 ilustra técnicas convencionales, bucles. Este ejemplo de realización describe técnicas convencionales de bucles dentro de una secuencia de comandos que están regidos de manera lógica por factores 32150 condicionales o de recuento. Presenta la estructura de código para un bucle 32200 condicional y/o de recuento.

Cuando se llama al archivo mediante un factor desencadenante en el inicio 32000. En este ejemplo, el código de preparación se ejecuta entre los ejemplos de líneas 004 y 009. En el caso de un bucle de contador, el bucle condicional se ejecutará mientras la condición 32200 de bucle evalúa que el contador de ciclo de bucle actual \$ii es menor que el número de veces que debe ejecutarse este bucle. En la línea 27 del ejemplo de secuencia de comandos, el contador o la condición 32150 (externa) añadirá uno a \$ii mediante \$ii++ o algún equivalente. Este bucle continuará ejecutándose mientras la condición siga siendo verdadera 32220.

Si el contador \$ii es igual a o mayor que el número de veces que el bucle debe realizar un ciclo, el bucle terminará y el flujo de procesamiento continuará. El código posterior al bucle que va a ejecutarse está ubicado entre las líneas 033 y 037.

La secuencia de comandos/rutina/software termina en la línea 038 y su ejecución termina en el fin 32990.

Aunque se llama de manera externa a la secuencia de comandos una vez y realiza un bucle tantas veces como se indique o bien mediante un contador o bien mediante una condición, también pueden surgir problemas tales como:

- ¿Qué sucede si/cuando falla una secuencia de comandos? ¿Cómo puede recuperarse?
- ¿Qué sucede si una secuencia de comandos tarda más en ejecutarse por cada ciclo?
- ¿Qué sucede si el reloj de base del sistema presenta una deriva que es demasiado rápida o demasiado lenta?
- ¿Qué sucede si la secuencia de comandos se divide y se produce un número mayor que el previsto de instancias simultáneas de sí misma?
- ¿Cómo se realiza un ajuste fino de la granularidad de una marca hasta ms o incluso µs?

La figura 33 ilustra técnicas convencionales, recursión, secuencia de comandos que se llama a sí misma. Este ejemplo de realización es similar a la figura 32 y la principal diferencia es que es más sencillo. No contiene un bucle dentro de su código sino que más bien tiene un comando que se llama a sí mismo en un bucle de tiempo de ejecución constante.

Comienza en el inicio 33000. El código 33100 se indica de las líneas 001 a 020. Para prepararse para la ejecución, el código está ubicado entre las líneas 002 y 004. El código que va a ejecutarse durante esa primera instancia está entre las líneas 005 y 014. Tras la ejecución, el código que va a ejecutarse está entre 015 y 018. En la línea 020 de la figura, la secuencia de comandos se llama a sí misma y, por ejemplo, abre un intérprete de CLI> para volver a lanzar la secuencia de comandos 33200. Si hay una condición para terminar la ejecución del código (lógica no ilustrada), entonces, en la línea 016, puede guardarse 33800 información de código. La secuencia de comandos termina en el fin 33990.

Si no existe ninguna condición o lógica para terminar el bucle de la secuencia de comandos que está ejecutándose, entonces el bucle se ejecutará de manera indefinida.

La figura 34 ilustra problemas con los bucles. La figura 34 se basa en la figura 33 y el principal factor diferenciador son los problemas indicados como 344###. Estos problemas son los siguientes:

34400 No puede lanzarse

Debido a permisos del usuario del sistema que llama u otras cuestiones, una secuencia de comandos 4-100 no puede ejecutarse inicialmente. Esto puede resultar inmediatamente evidente o no. Un cambio de software o actualización de software, falta de archivos de inclusión u otros factores pueden evitar la ejecución.

34402 Demasiadas instancias

Si, por algún motivo, se llama la secuencia de comandos 4-100 para que se ejecute más de una vez,

entonces estará ejecutándose más de una instancia de la secuencia de comandos. Esto resulta problemático porque cada instancia de una secuencia de comandos se llamará a sí misma y se ejecutará de manera perpetua consumiendo recursos de sistema. También suponen riesgos de conflicto debido a que múltiples secuencias de comandos intentan procesar o gestionar el mismo objeto con resultados no intencionados y posiblemente destructivos o disruptivos. Más no siempre es mejor.

- 34404 Excepción/fallo no detectado
Si una secuencia de comandos se ejecuta y presenta un fallo, a menos que se detecte esta excepción, el siguiente bucle no se desencadenará y no se realizará la funcionalidad que logra la secuencia de comandos.
- 34406 Tiempo de ejecución > tiempo permitido de marca
Si una secuencia de comandos tarda más que el tiempo asignado para una marca, la siguiente marca se retardará de promedio descolocando la planificación.
- 34408 Información de estado corrupta
Si el estado en el que se basa una secuencia de comandos para la realización de bucles condicional se corrompe, entonces pueden producirse como resultado consecuencias no intencionadas.
- 34410 No puede volver a lanzarse
Si se ejecuta una secuencia de comandos pero no puede volver a lanzarse a sí misma porque el procedimiento en ejecución tiene derechos diferentes del usuario o procedimiento que llamó inicialmente a la secuencia de comandos, entonces sólo se ejecutará una vez y no seguirá ejecutándose.
- 34412 Archivo de secuencia de comandos corrupto
Si el archivo de secuencia de comandos se corrompe en los medios de almacenamiento (ya sea de manera física o virtual en almacenamiento en memoria), la siguiente vez que se llame por el CLI> de volver a lanzar la secuencia de comandos 4-200 o bien presentará un fallo o bien puede ejecutarse y producir consecuencias no intencionadas.
- 34414 Problemas desconocidos
Puede haber otros problemas que se producen que son desconocidos y no pueden detectarse. El lanzador que llamó a la secuencia de comandos puede pensar que está ejecutándose pero en realidad no es así.

Los problemas anteriores se incluyen para destacar algunas de las posibles cuestiones. No todos, o sólo algunos de los mismos, pueden estar presentes en cualquier momento. Pueden existir más problemas no expuestos.

- 5 La figura 35 ilustra $\Delta t = P + Q + R$. Esta figura describe el ciclo de vida de una marca. La granularidad de tiempo de una marca puede definirse de la siguiente manera.

La granularidad de una marca

$$10 \quad \Delta t = P + Q + R$$

Δt = delta de tiempo desde el inicio de la marca hasta el final de la marca 35100

- 15 • Inicio basándose en completarse la última marca

o

- según un tiempo planificado

- 20 P = tiempo para el procesamiento por lotes de elementos para esta marca en el ciclo.

Q = tiempo para cálculos de procesamiento de lotes posterior.

- 25 R = tiempo para el retardo para garantizar que no hay ningún solapamiento entre elementos de lote y/o para establecer una desviación entre marcas.

El periodo 35000 es el delta de tiempo o Δt durante la duración de una marca.

- 30 La figura 36 ilustra el sincronismo de diversos ciclos en los que $X \neq Y$. Esta figura se basa en la figura 35 y describe dos marcas con tiempos variables y diferentes.

$\Delta t = x \ 36100$ y $\Delta t = y \ 36200$ no son iguales entre sí, tal como se describe por $x \neq y \ 36010$.

- 35 En este ejemplo, los retardos $R \ 36130$ y $R \ 36230$ son iguales entre sí. Los tiempos de procesamiento $P \ 36110$ y $P \ 36210$ no son iguales, reflejando la cantidad diferente de tareas que se requiere que ejecuten cada uno. Los

tiempos de procesamiento posterior Q 36120 y Q 36220 también son respectivamente diferente basándose en la cantidad de tareas que tienen que completar cada uno.

Por tanto, la duración de tiempo para cada ciclo cambia basándose en cuántos elementos tienen que procesarse.

La figura 37 ilustra el sincronismo de diversos ciclos en los que X es sistemático. Esta figura se basa en la figura 35 y, aunque describe dos ciclos de marca, a diferencia de la figura 36, estos ciclos de marca tienen la misma duración de tiempo $x_0 = x_2$ 37010. Aunque $\Delta t = x_0$ 37100 y $\Delta t = x_2$ 37200 son iguales, los tiempos de procesamiento por cada marca P 37110 y P 37210 pueden o bien ser iguales entre sí o bien pueden ser uno más largo que el otro. Cada uno se basa todavía en la cantidad de tareas que tiene que procesar. Y respectivamente los tiempos de procesamiento posterior Q 37120 y Q 37220 también pueden diferir uno de otro o ser iguales.

Las marcas se mantienen en una duración Δt igual 37010 al menos de dos maneras:

- Limitando la cantidad de elementos que van a procesarse por P de una marca (P 37110 o P 37210) nunca se forzará al Δt a pasar más allá de su presupuesto de tiempo asignado.

- Los retardos R 37130 y R 37230 se ajustan de manera dinámica para garantizar que el retardo es lo suficientemente largo antes de ejecutarse el siguiente ciclo. El retardo puede ser un temporizador para suspensión o puede ser una acción nula pero tiene que haber alguna clase de hueco entre el final de un Q del ciclo de marca actual y el comienzo del P del siguiente ciclo de marca.

Este tiempo fijo es posible si hay un límite sobre la cantidad del número de elementos procesados basándose en cuánto tiempo se requiere para cada P y Q.

La figura 38 ilustra un umbral de distancia para obtener una ganancia a partir de salto de lanzamiento. Esta figura destaca la lógica de decisión algorítmica sobre si es o no más rápido y/o más eficiente enviar paquetes a través de un trayecto de Internet IP 38562 a 38588 o a través de un trayecto de internet potenciado por lanzamiento 38662 a 38688 con un segmento transparente de salto de lanzamiento 38800 y 38880 entre un par (o más de un par) de direcciones de IP 38668 y 36682.

No tiene en cuenta la fiabilidad o pérdida o congestión relacionada u otras cuestiones además de la latencia que puede tener un efecto sobre el trayecto de IP.

Hay ineficiencias de usar un salto de lanzamiento de Δt en cada cabeza de puente 38610 y 38620 del segmento de salto transparente de salto de lanzamiento 38600 frente a la ganancia lograda 38370 mediante el uso de salto de lanzamiento.

Se necesita una distancia mínima para compensar el tiempo de procesamiento requerido del mecanismo de salto de lanzamiento en 38615 y 38625.

Medir la latencia de IP 38100 usa la siguiente ecuación para la medición

$$\text{Tiempo de trayecto de Internet total } \Delta t \text{ ms} = TX_{38510} + TW_{38500} + TY_{34590}$$

Medir la latencia de SL 38200 usa la siguiente ecuación para la medición:

$$\text{Tiempo de trayecto híbrido de Internet-salto de lanzamiento total } \Delta t = TX_{38630} + TV_{38610} + TZ_{38600} + TV_{38620} + TY_{38640}$$

La comparación de latencia de IP frente a SL usa la ecuación

$$\Delta t\{TV + TZ + TV\} < \Delta t\{TW \text{ ms}\}$$

Si es verdadero que el trayecto potenciado por salto de lanzamiento ($\Delta t\{TV + TZ + TV\}$) tiene una latencia inferior, SL es inferior 38370, al trayecto de Internet, entonces "SL es el uso de transporte más eficiente" 38380 es el trayecto más óptimo.

Si es falso e Internet ($\Delta t\{TW \text{ ms}\}$) es más rápido que salto de lanzamiento, "IP es inferior" 38320, entonces "IP es el uso de transporte más eficiente" 38360 es el trayecto más óptimo.

La importancia del sincronismo exacto no es sólo que los archivos se reciben y procesan sin duplicados, sino que además hay un retardo limitado, el archivo se procesa con un retardo mínimo.

La figura 39 ilustra la sincronización de lotes, con respecto a marcar archivo(s). Esta figura se basa en la figura

25 anterior. El uso de los archivos ilustrado para el uso 39602 de archivos en la extracción de lote A 39600, para el uso 39612 de archivos en la extracción de lote B 39610, y para el uso 39622 de archivos en la extracción de lote C 39620. El uso de archivos (lote A) 39702 envía los archivos, archivo 39800 y archivo 39802, en paralelo y termina en el fin del lote 39902. El uso de archivos (lote B) 39712 envía los archivos, archivo 39804 y archivo 39806 y archivo 39808, en paralelo y termina en el fin del lote 39912. El uso de archivos (lote C) 39722 envía los archivos, archivo 39810 y archivo 39812 y archivo 39814, en paralelo y termina en el fin del lote 39922. El punto clave es que, aunque los archivos se extraen en lotes, la extracción de archivos que están disponibles se realiza en paralelo y, por tanto, el uso también se realiza en paralelo. Los archivos no listos para la extracción de lote A 39600, tales como el archivo 39804 y el archivo 39806, no se extraerán ni se usarán hasta que estén disponibles. Los archivos 39804 y 39806 se extraen con el archivo 39808 en la extracción de lote B 39610.

La capacidad de extraer por fragmentos y procesar en paralelo permite ajustar a escala de manera masiva el mecanismo de lanzamiento.

La figura 40 ilustra la granularidad de una marca. Esta figura ilustra dos marcas 40100 y 40200. Este ejemplo demuestra dos extracciones de archivos de lote en un servidor de intercambio troncal (SRV_BBX) con leer cola + procesar 40510 y leer cola + procesar 40610. Ambos extraen archivos de los mismos medios de almacenamiento, sistema de archivos en paralelo, PFS de archivos entrantes 40800. Los archivos extraídos en 40510 a través del trayecto 40512 se procesan y después, en el procesamiento posterior, P posterior 40520, se marcan los archivos a través del trayecto 40522.

Esto es un punto críticamente importante porque la siguiente extracción de archivos de lote, leer cola + procesar 40610 a partir de PFS de archivos entrantes 40800 a través del trayecto 40612 sólo debe incluir archivos no marcados o archivos no llenos por lotes anteriores. Entonces, en P posterior 40620, los archivos extraídos y usados se marcan a través del trayecto 40622 de modo que no se extraerán de manera inadvertida por una extracción de archivos de lote posterior.

La figura 41 ilustra la granularidad de una marca, con respecto a un intervalo de tiempo variable. Esta figura es similar a la figura 40, sin embargo, ilustra el procesamiento de cantidades variables de archivos y demuestra cómo afecta al tiempo tanto de procesamiento como de procesamiento posterior así como indica un retardo mínimo antes de que deba ejecutarse el siguiente ciclo.

Se indican tres ciclos en el presente documento; 41100, 41200 y 41300, y cada ciclo contiene diferentes cantidades respectivas de cantidad 58 en 41810, de cantidad 22 en 41830, y de cantidad 8 en 48860.

La figura 42 ilustra un algoritmo a modo de ejemplo, con respecto al ajuste de marca. Esta figura describe el algoritmo usado para diferentes tipos de marcas. Comienza en el inicio 42000 y termina en el fin 42990.

En la unión 42020, se ejecutan dos procedimientos paralelos para recopilar parámetros de marca 42030 obtenidos a partir de los ajustes de marca 42050 y comprobar monitores de sistema 42100 obtenidos a partir de datos de monitorización de sistema 42150 acumulados a partir de la fuente de datos de monitorización de sistema 42160.

Si todo es correcto para que la marca siga ejecutándose en ¿Todo correcto? 42140, entonces el flujo continuará a lo largo del trayecto de sí 42142. Si no es correcto, el flujo pasará a la unión de error 42200 a través del trayecto de no 42144 hasta la gestión de errores 42400.

Los dos tipos de marca son ¿marca variable o fija? 42170. Una marca de tiempo variable 42176 comienza en el inicio de secuencia de comandos 42300. Una marca de tiempo fija 42172 comienza en el inicio de secuencia de comandos 42500 usando información a partir de usar parámetros de marca 42175. También hay una puerta lógica adicional de ¿tiempo de marca restante? 42550 para garantizar que no hay rebasamiento del tiempo. Los procedimientos de compleción de ejecución de marca habituales, usados por procedimientos tanto fijos como variables, incluyen fin de preparación de secuencia de comandos 42600 y notificación 42680.

La figura 43 ilustra un análisis de inteligencia artificial, granularidad y periodos. Esta figura ilustra el análisis y la interpretación de periodos de ciclos de marca. La inteligencia artificial puede ayudar en cálculos futuros para estimar cuánto tardarán los tiempos de procesamiento y de procesamiento posterior de una marca basándose en cantidades de elementos que van a procesarse teniendo en cuenta recursos de sistema y otros factores.

Por cada periodo calculado, pueden representarse gráficamente métricas con datos a partir del periodo actual 43100, datos históricos a corto plazo 43200 y a largo plazo 43300 en una curva de desviación estándar basándose en las diversas tendencias cíclicas. La indicación de puntos bajos, puntos altos, promedios y otros análisis puede indicar si el performance actual está en línea con las expectativas o si es mejor, peor o diferente de otro modo de la experiencia anterior.

Esta recopilación de datos y análisis contextual ayudan a algoritmos de inteligencia artificial (AI) en la toma de

decisiones. El periodo actual 43100 puede compararse con datos de rendimiento a corto plazo 43200 y a largo plazo 43300.

La figura 44 ilustra componentes a modo de ejemplo de la marca. Hay tres tipos de dispositivos ilustrados en esta figura. Un servidor 44200 de control central (SRV_CNTRL), un dispositivo 44100 genérico y un servidor 44600 de tiempo (SRV_TIME). El dispositivo puede ser cualquier tipo de dispositivo de punto de extremo (EPD), concentrador de punto de extremo (EPH), servidor de punto de acceso (SRV_AP) u otro tipo de dispositivo.

En el servidor 44200 de control central (SRV_CNTRL) se almacena información sistémica en la base 44295 de datos (Db) del repositorio 44295 de información central. Esta información puede usarse por gestores relacionados con marca, monitores y otros componentes de sistema. Véase la figura 48 en la que se indican algunos elementos relacionados con marca.

Para actualizar el tiempo de su reloj interno, un dispositivo 44100 puede consultar el escuchador 44650 de servidor de NTP de un SRV_TIME 44600 mediante una llamada de API u otro tipo de infraestructura de petición-respuesta a través de los trayectos REQ 44660 para la petición y RESP 44680 para la respuesta. Las bases de datos en el dispositivo 44100, Db 44155, y en el SRV_TIME 44600, base 44655 de datos Db, pueden almacenar información sobre peticiones de tiempo, deriva de tiempo del dispositivo, otra marca, información relacionada con el tiempo y otra información.

El/los dispositivo(s) 44100 guarda(n) información sobre el tiempo en sus bases 44155 de datos locales Db y también notifican su información relacionada con el tiempo a partir de su gestor 44120 de sistema y tiempo al SRV_CNTRL 44200 mediante un envío de REQ-RESP al SRV_CNTRL 44200 a través de REQ 44600 con confirmación de vuelta a través de RESP 44620.

El módulo 44250 de tiempo de NTP en el SRV_CNTRL 44200 analiza información relacionada con tiempo sistémica.

La figura 45 ilustra un uso de marca a modo de ejemplo en ráfagas de baliza. Esta figura ilustra la importancia del uso de marcas en una aplicación en el mundo real, una baliza de datos. Todos los pulsos de información comienzan en el mismo origen 45200 que está en un servidor 45000 de estructura central (SRV_BB) y radian hacia fuera. Esta figura demuestra 10 pulsos desde el origen hasta el punto más alejado en el que se escribe un archivo de información en un dispositivo de almacenamiento de sistema de archivos en paralelo (PFS) remoto, PFS 45800, para su recuperación y uso por el SRV_BB 45900 en la región remota. Un SRV_BB también puede ser un nodo de lanzamiento u otro dispositivo.

Los pulsos son de 45300 a 45318 y la duración de tiempo entre cada pulso se indica como ts00 45100 a ts18 45118. Estos pulsos deben producirse a intervalos definidos de modo que es posible fiarse de cada ráfaga. Estos pueden transportar datos simples a complejos. El punto clave es que la baliza es fiable.

El sincronismo de combinación de múltiples ráfagas de baliza se denomina tasa de transmisión de pulsos. Y la frecuencia es Δt entre dos ráfagas. Con respecto a la tolerancia, esto tiene que ver con la latencia y la demanda de actualidad de información. Lo actual que es la información cuando se recibe depende de la latencia. Lo reciente que es la información depende de la latencia más el retardo entre pulsos.

Por ejemplo, si la latencia es de 50 ms desde 45000 hasta 45800 y la frecuencia de pulsos y la granularidad de marca es cada 1 ms, entonces la información siempre es reciente. Sin embargo, si la frecuencia es de 3 ms entre pulsos y la granularidad de marca es de 1 ms, entonces hay una posibilidad de que la información pueda tener una antigüedad de 2 ms.

Lo reciente de la información supone el envío de un pulso a la mejor velocidad de cableado posible y con la mayor eficiencia de envío y recepción posible.

La figura 46 ilustra un uso de marca a modo de ejemplo en múltiples pulsos de baliza. Esta figura ilustra el uso de marca en una aplicación en el mundo real, tal como una baliza de información de múltiples pulsos para diversas regiones. El punto de origen para la información es SRV_BB 46000 en el perímetro 46100. Hay tres pulsos de baliza que se envían exactamente en la misma marca de tiempo a tres ubicaciones diferentes que están a distancias variables desde el origen. Las tres ubicaciones de destino son el PFS 46280 objetivo en 46200 a través del trayecto 46282, al PFS 46580 a través del trayecto 46582 y al PFS 46380 a través del trayecto 46382.

La suposición es que hay un transporte unidireccional de información como un archivo enviado lo más cerca posible de la velocidad de cableado desde el SRV_BB 46000 para escribirse a través de acceso directo a memoria remoto (RDMA) en cada PFS. La baliza garantiza que la información se transmite de manera global y se combina con múltiples pulsos y garantiza que la información siempre es lo más reciente posible mediante combinación de latencia conocida entre dos puntos con granularidad sistemática de marcas para el envío del

pulso/ráfaga de información de baliza y la recepción y después la granularidad de una marca para el procesamiento y después el uso posterior de archivos recibidos.

Hay más aplicaciones para esto además de la que se ilustra.

5

La figura 47 ilustra un uso de marca a modo de ejemplo en la baliza en puntos de intersección. Esta figura ilustra la granularidad de una marca usada en una aplicación en el mundo real, tal como el envío de múltiples pulsos de baliza desde dos ubicaciones y el punto de intersección de los pulsos en una tercera ubicación.

10

El primer origen de información transmitida mediante una baliza de información por pulsos está en NYC 47100. Se envían pulsos a LDN 47110 a través del trayecto 74200. El segundo origen de información transmitida mediante una baliza de información por pulsos está en TOK 47120. Se envían pulsos a LDN 47110 a través del trayecto 47208.

15

En este ejemplo, los pulsos de ambos orígenes se han sincronizado para que se solapen. Esto puede realizarse mediante ajuste de tiempos de inicio o alineando marcas y enviando tasas de transmisión.

Para este ejemplo, NYC representa Nueva York, NY, EE.UU., LDN representa Londres, Inglaterra, R.U., y TOK representa Tokio, Japón.

20

Se indican los segmentos entre las ubicaciones así como la mejor velocidad de cableado en una dirección para la información.

25

La distancia entre Nueva York y Londres es de 3.456 millas y la mejor velocidad de cableado unidireccional 47000 es igual a 27,3 ms. La distancia entre Tokio y Londres es de 5.936 millas y la mejor velocidad de cableado unidireccional es 47020 es igual a 46,8 ms. No se ilustra en la parte superior del diagrama, pero debajo se compara que la distancia entre Nueva York y Tokio es de 6.737 millas y la mejor velocidad de cableado unidireccional entre las mismas es 47080 es de 53,1 ms.

30

Por ejemplo, en mercados financieros en los que resultan relevantes datos de precios, otra información y noticias relacionadas de diferentes regiones mundiales y pueden influir en la negociación en un mercado local, tener la información más reciente y actualizada resulta crítico para la misión. Un supernodo informático (SCN) en Londres puede recibir información sobre Nueva York y Tokio y puede reaccionar en consecuencia. Si, por ejemplo, el precio en Tokio empieza a fluctuar, Londres puede reaccionar más rápido que Nueva York mediante una ventaja de distancia de ubicación de 801 millas 47082, lo cual es igual a 6,3 ms.

35

Por tanto, la granularidad de una marca y su impacto sobre el sincronismo y la secuenciación tiene una importancia crítica para la misión para agentes de negocios que buscan una ventaja.

40

Hay muchos otros ejemplos de cómo múltiples pulsos a partir de diversas ubicaciones pueden beneficiarse de una granularidad de marcas fiable.

45

La figura 48 ilustra un diagrama de sistema a modo de ejemplo de granularidad de una marca. Esta figura describe los diversos módulos que se ejecutan en diversos dispositivos que pueden usar la granularidad de una marca descrita en el presente documento. El dispositivo genérico, todos los dispositivos 48900 con marca, puede aplicarse a cualquier dispositivo para el que la granularidad de una marca es importante. El monitor 48916 de HW + gestor evalúa los recursos disponibles, la carga de sistema, el uso de HW potencial y actual y otros factores que pueden influir en el sincronismo.

50

El monitor 48918 de O/S + gestor evalúa los procedimientos y otros elementos relacionados con el sistema que están ejecutándose que pueden tener una influencia sobre el sincronismo. El actualizador 48970 de tiempo interacciona con el servidor 4831800 de tiempo (SRV TIME) a través del escuchador 4831850 de NTP para pedir y recibir el tiempo más actual.

55

El software 48950 de marca en todos los dispositivos 48900 con marca funciona para asegurarse de que la granularidad de una marca es lo más eficiente posible. Interacciona con monitores y gestores. Se pueden usar software incorporado y complementos y otros métodos para extender esto. El monitor 48980 de recursos y el mon. 48985 + gestor de procedimientos pueden proporcionar información sobre el gestor 48988 de marca.

60

Los módulos, gestores, componentes de HW y otras partes constituyentes mostrados en el presente documento son únicamente como ejemplo y pueden diferir en el uso en el mundo real. Otros dispositivos también pueden usar granularidad de una marca para diversas clases de funcionalidad.

65

Un hilo unificador clave entre todos ellos es la importancia del tiempo. Y, por tanto, el tiempo de sistema debe ser preciso y la granularidad de una marca fiable de modo que los acontecimientos, y especialmente acontecimientos secuenciales, puedan procesarse y ejecutarse de una manera predecible.

REIVINDICACIONES

1. Sistema para transferencia y procesamiento de archivos en un entorno de red, que comprende:

5 uno o más procesadores, acoplados a un primer dispositivo, configurados para:

recuperar un lote de múltiples archivos (22802, 22804, 22806, 22808, 22810) a partir de una cola de archivos, en el que la cola de archivos está almacenada en un almacén (22800) local del primer dispositivo, en el que el lote de archivos se transfiere desde uno o más dispositivos remotos a través de acceso directo a memoria remoto;

procesar cada archivo en el lote, en el que el procesamiento comprende:

determinar (22300) si el archivo está completo; y retirar el archivo a partir de la cola de archivos, si se determina que el archivo está completo;

en el que el uno o más procesadores están configurados además para retardar la recuperación de un siguiente lote de múltiples archivos a partir de la cola de archivos en una cantidad de tiempo de tal manera que se completa el procesamiento del lote de archivos antes de comenzar la recuperación del siguiente lote de archivos.

2. Sistema según la reivindicación 1, en el que la cantidad de tiempo se establece basándose en un tiempo de procesamiento del lote de archivos (22802, 22804, 22806, 22808, 22810).

3. Sistema según la reivindicación 1, en el que la cantidad de tiempo es igual a la suma de: un tiempo de procesamiento del lote de archivos (22802, 22804, 22806, 22808, 22810), un tiempo de procesamiento para que el uno o más procesadores realicen cálculo de procesamiento de lotes posterior; y un retardo establecido a un valor suficiente para garantizar que no hay ningún solapamiento entre el procesamiento del lote de archivos y el procesamiento del siguiente lote de archivos.

4. Sistema según cualquier reivindicación anterior, en el que el uno o más procesadores están configurados además para: para cualquier archivo dentro del lote de archivos (22802, 22804, 22806, 22808, 22810) que se determina que está incompleto (22810), ignorar dicho archivo (22810) de tal manera que no se procesa con el lote de archivos y se procesa con un lote de archivos posterior.

5. Sistema según cualquier reivindicación anterior, en el que el uno o más procesadores están configurados además para marcar un archivo como leído en el caso en el que se determina que el archivo está completo.

6. Sistema según cualquier reivindicación anterior, en el que, como parte de determinar si el archivo está completo, el uno o más procesadores están configurados para:

calcular una suma de comprobación para una carga (22526) útil del archivo;

comparar la suma de comprobación calculada con una suma de comprobación incorporada en un pie (22528) del archivo; y

determinar que el archivo está completo en el caso en el que la suma de comprobación calculada es igual a la suma de comprobación incorporada en el pie del archivo.

7. Sistema según la reivindicación 1, en el que al menos uno del uno o más dispositivos remotos es un servidor de intercambio troncal (SRV BBX).

8. Sistema según la reivindicación 1, que comprende además un enlace dedicado entre el primer dispositivo y al menos uno del uno o más dispositivos remotos.

9. Sistema según la reivindicación 1, en el que cada archivo (22802, 22804, 22806, 22808, 22810) comprende una cabecera (22522), un cuerpo (22526) y un pie (22528).

10. Sistema según la reivindicación 9, en el que el pie (22528) comprende un marcador que indica un final del archivo respectivo.

11. Sistema según la reivindicación 9, en el que cada archivo (22802, 22804, 22806, 22808, 22810) comprende un par de valores clave.

12. Sistema según la reivindicación 9, en el que el pie (22528) comprende una suma de comprobación para validar la integridad del cuerpo (22526).

13. Sistema según la reivindicación 1, en el que el acceso directo a memoria remoto se implementa basándose en al menos uno de Infiniband o canal de fibra.

5 14. Sistema según la reivindicación 1, en el que cada archivo (22802, 22804, 22806, 22808, 22810) comprende código ejecutable.

10 15. Sistema según la reivindicación 1, en el que el uno o más procesadores están configurados además para, como parte del procesamiento de cada archivo en el lote (22802, 22804, 22806, 22808, 22810): si se determina que el archivo está completo, al menos uno de establecer un indicador en el archivo, mover el archivo a una carpeta o eliminar el archivo del almacén local.

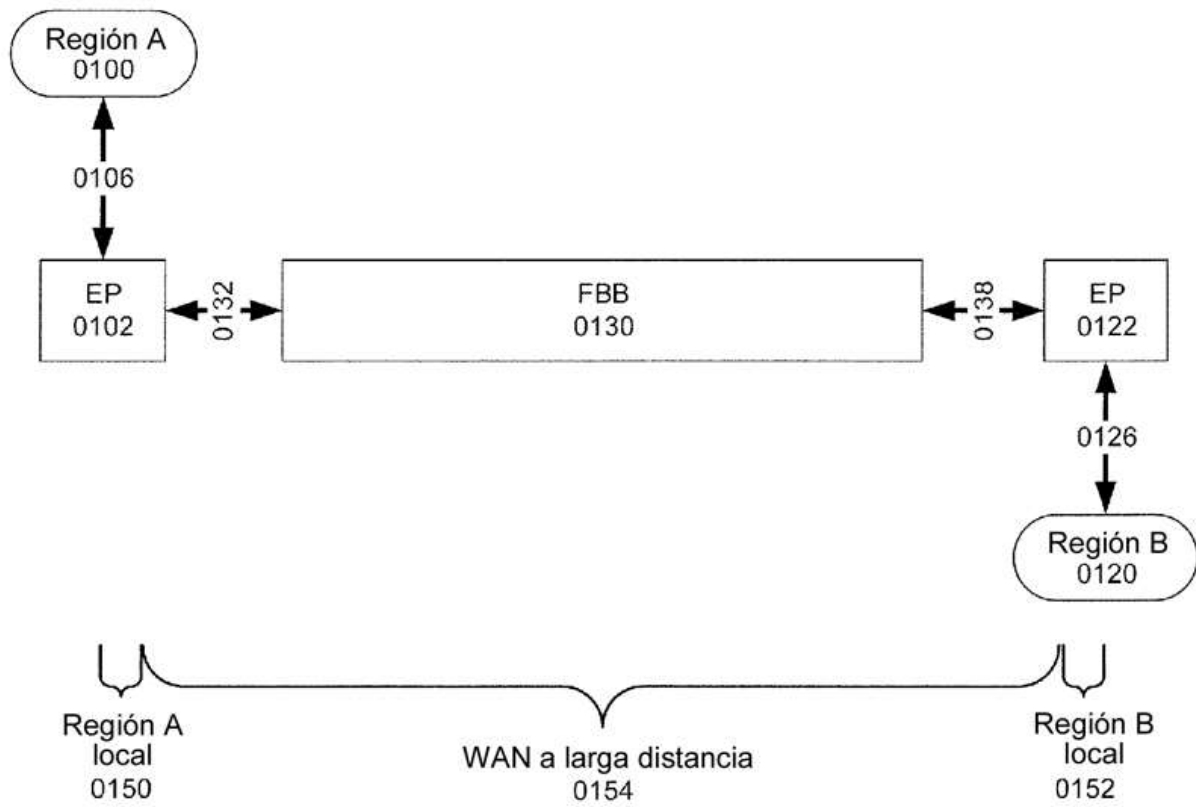


FIG. 1

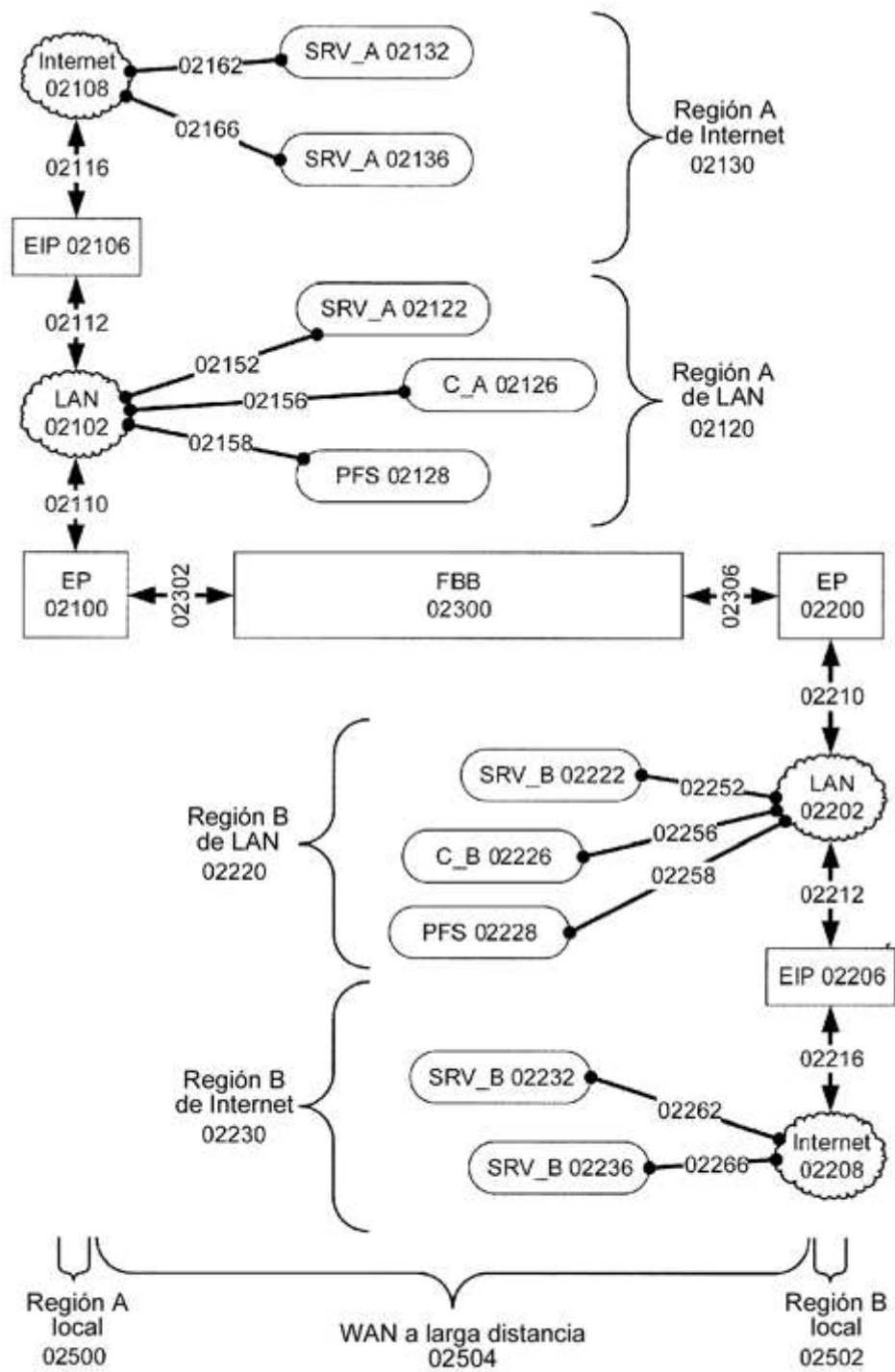


FIG. 2

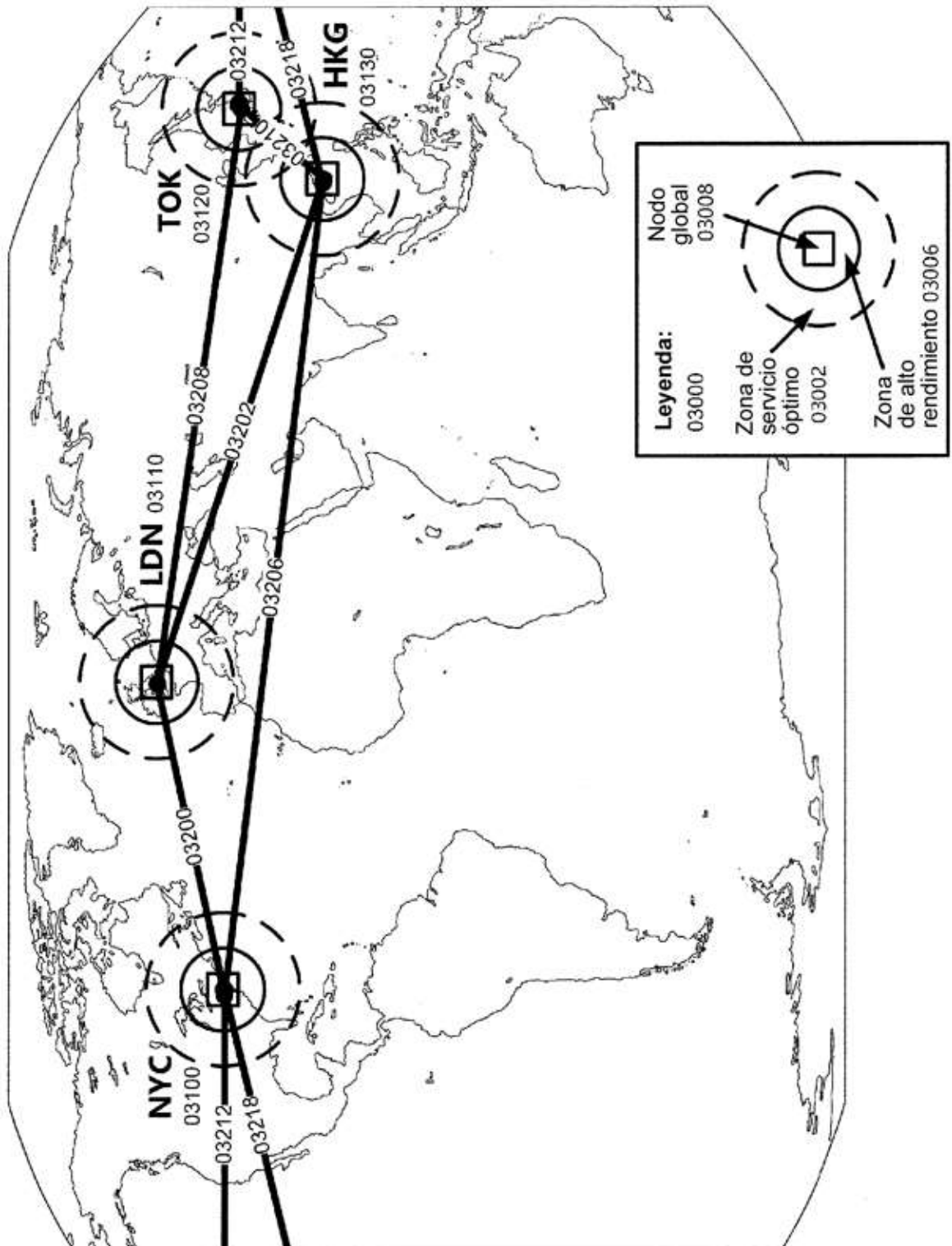


FIG. 3

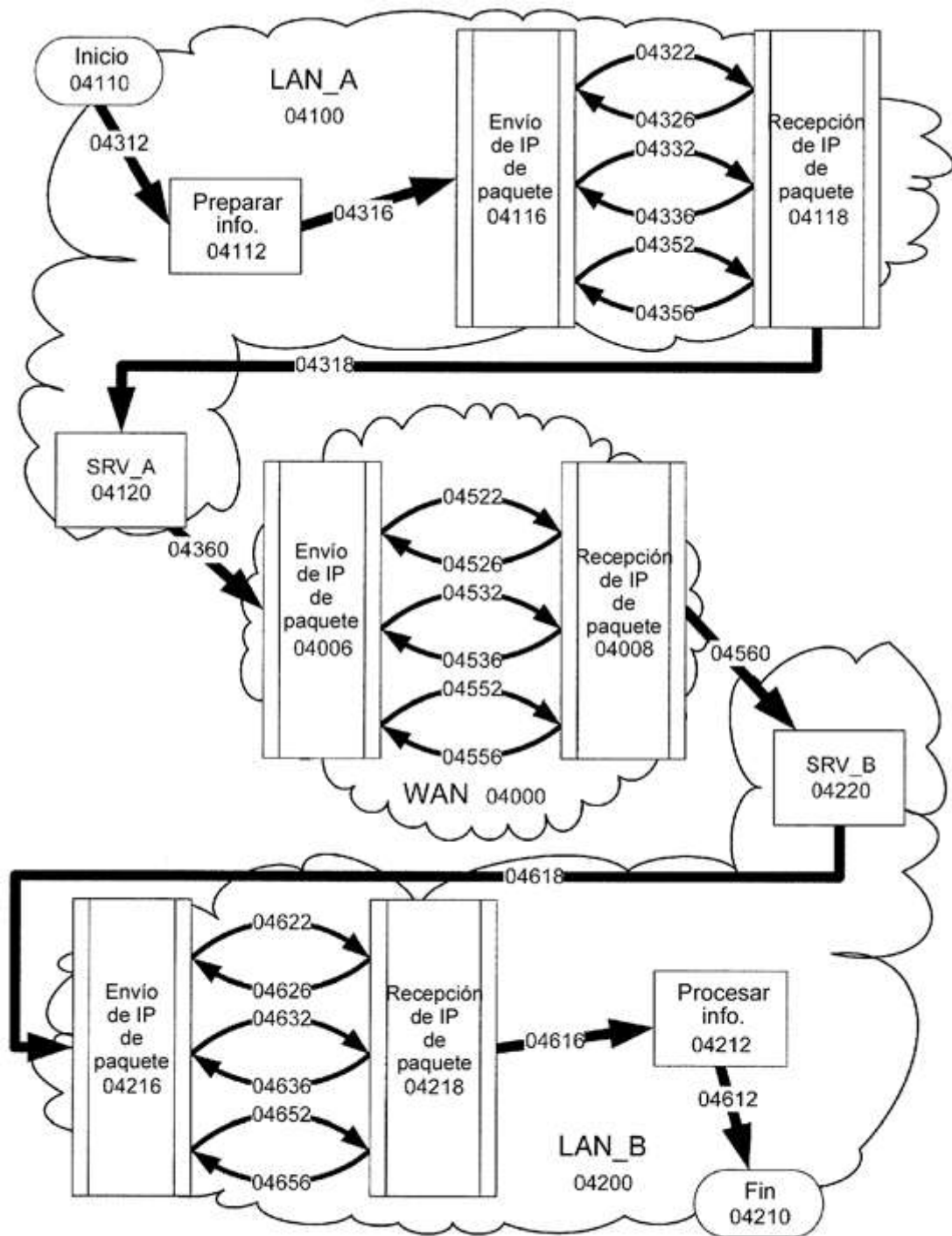


FIG. 4

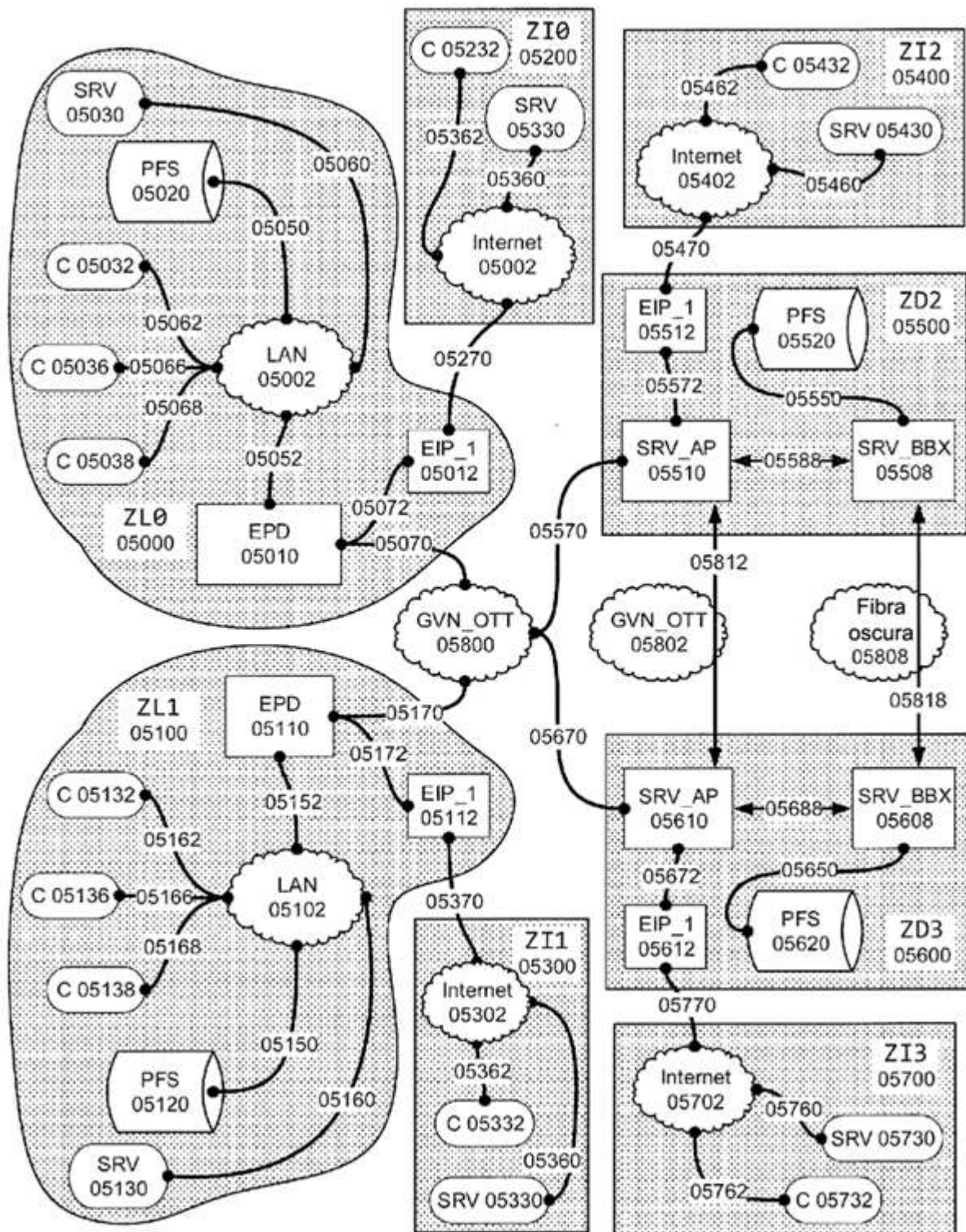


FIG. 5

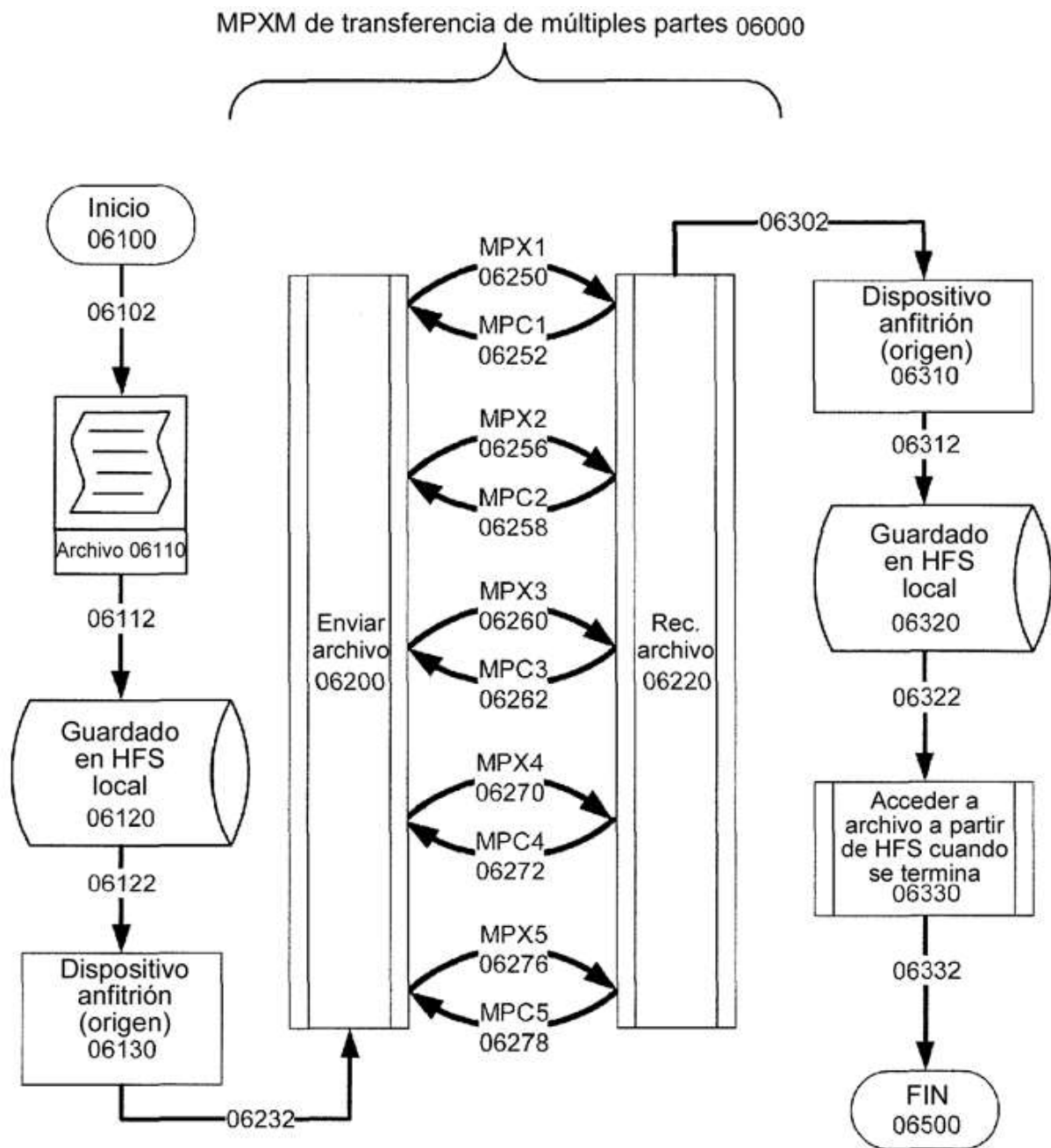


FIG. 6

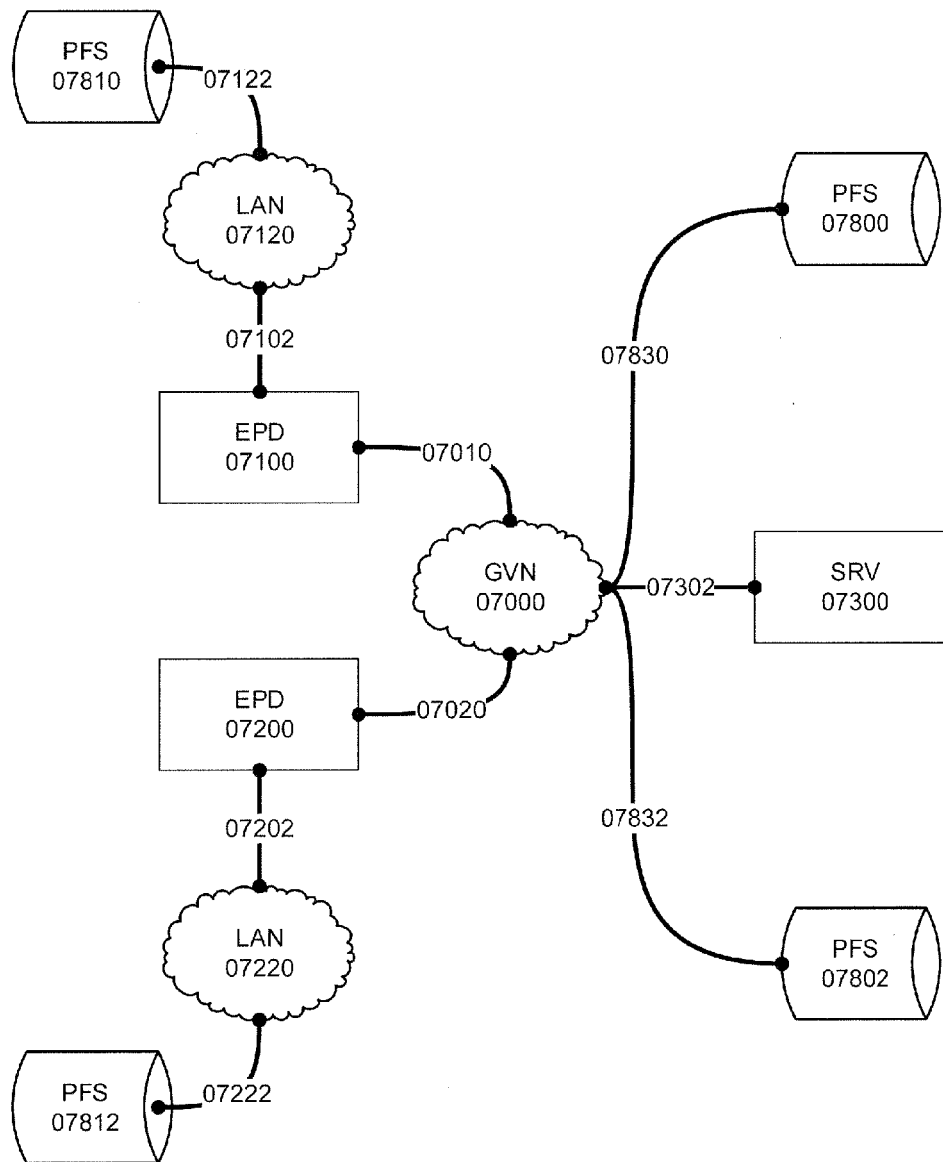


FIG. 7

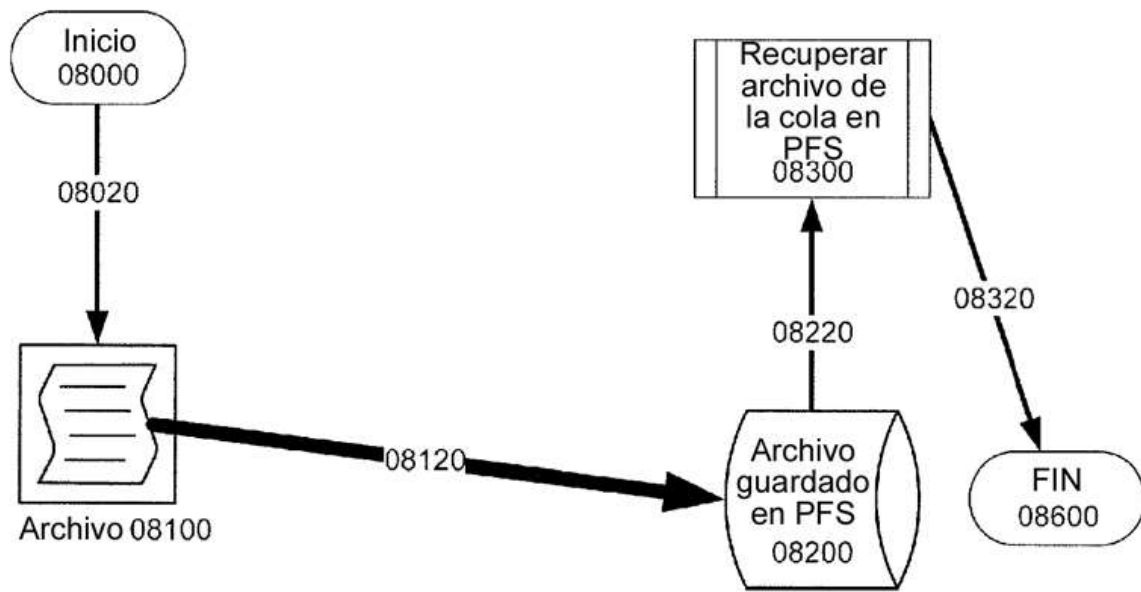


FIG. 8

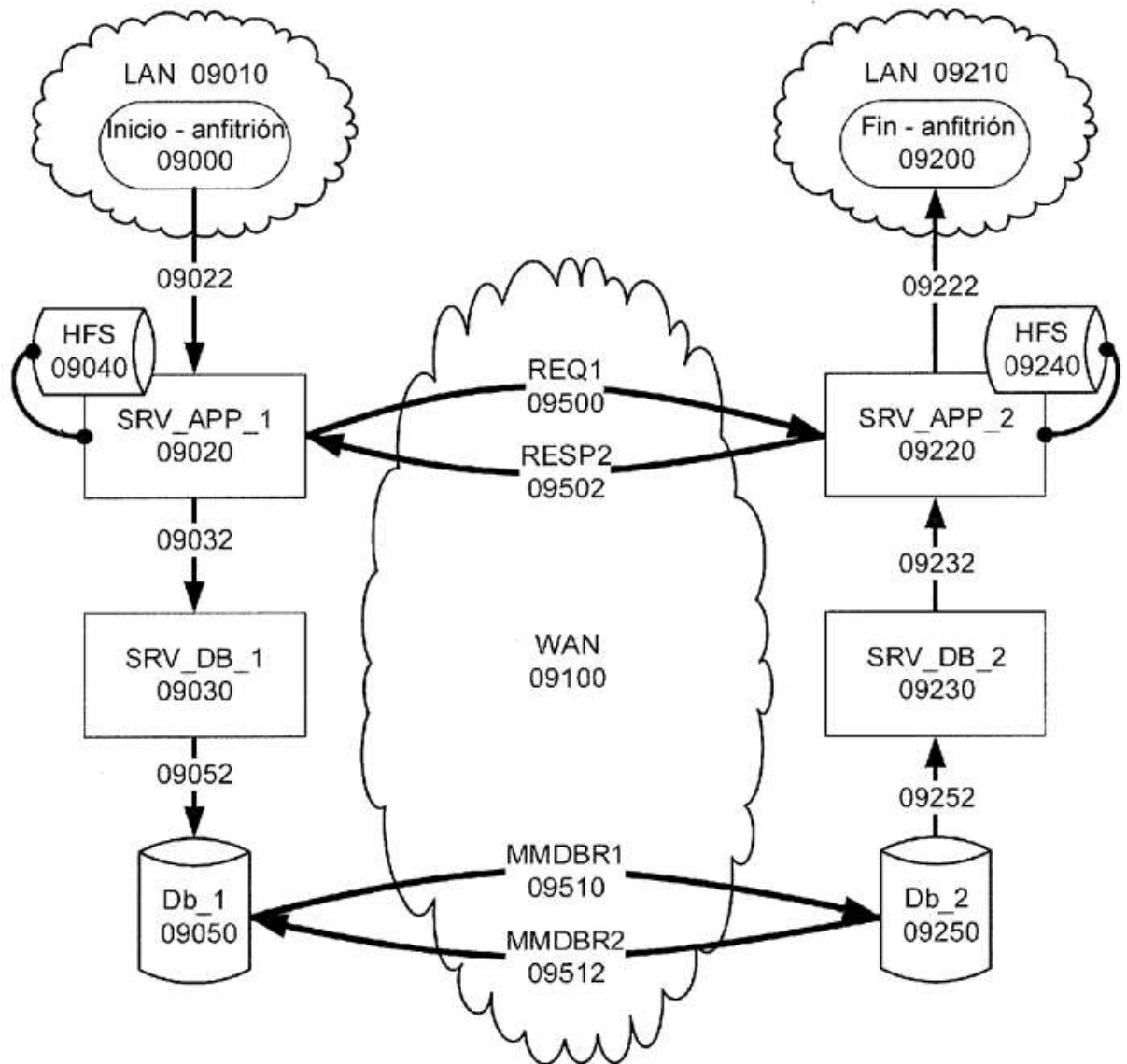


FIG. 9

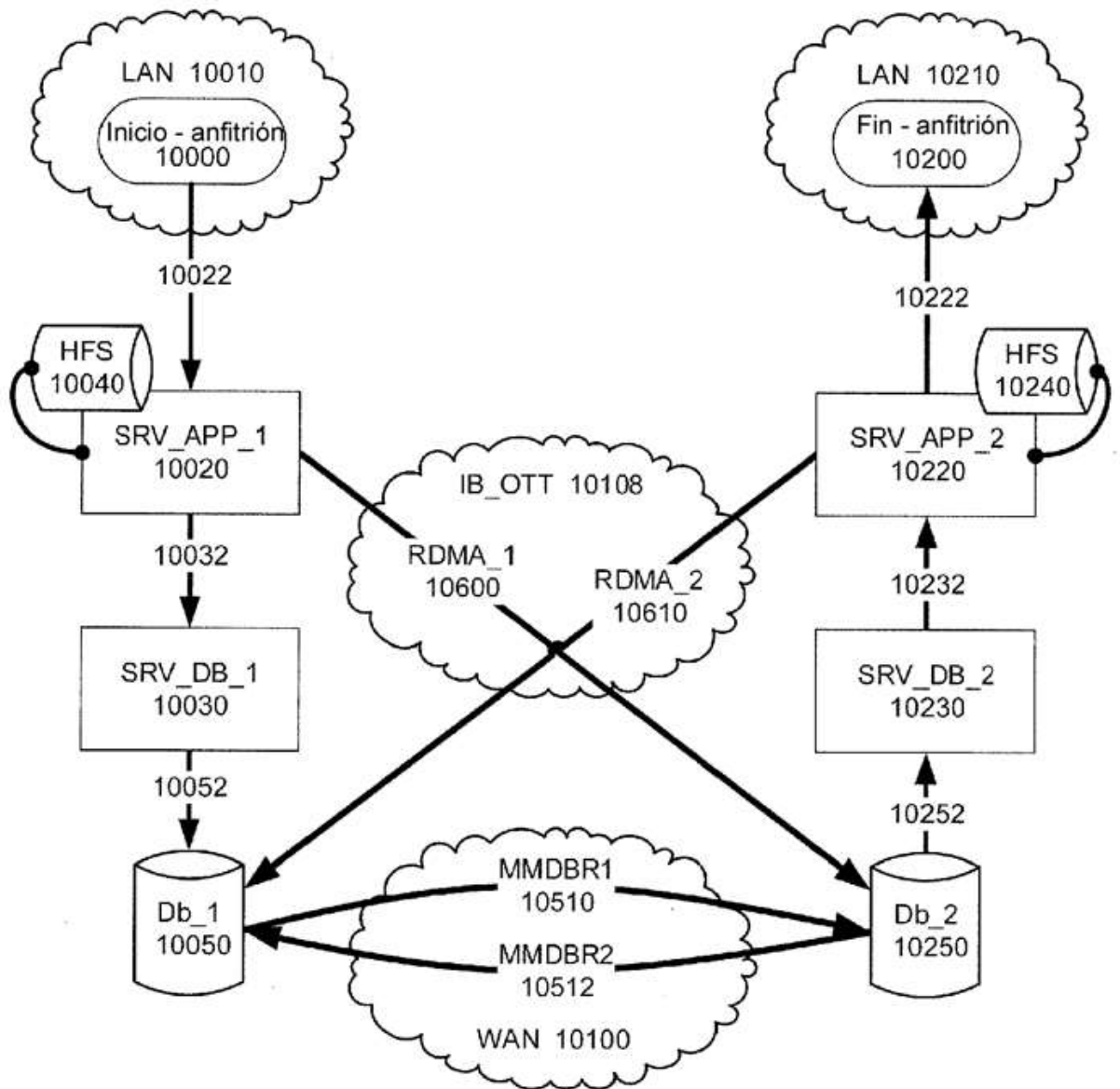


FIG. 10

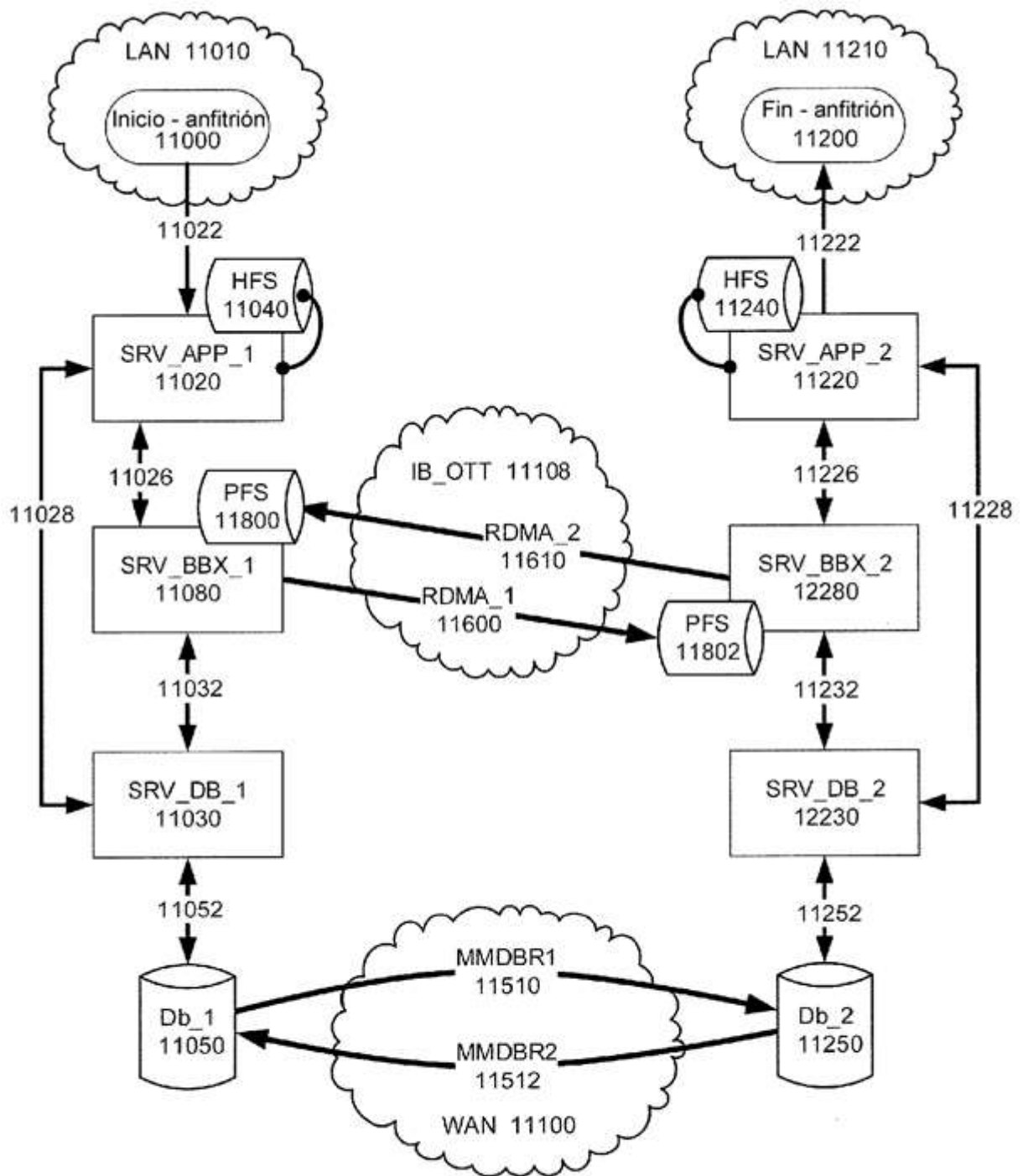


FIG. 11

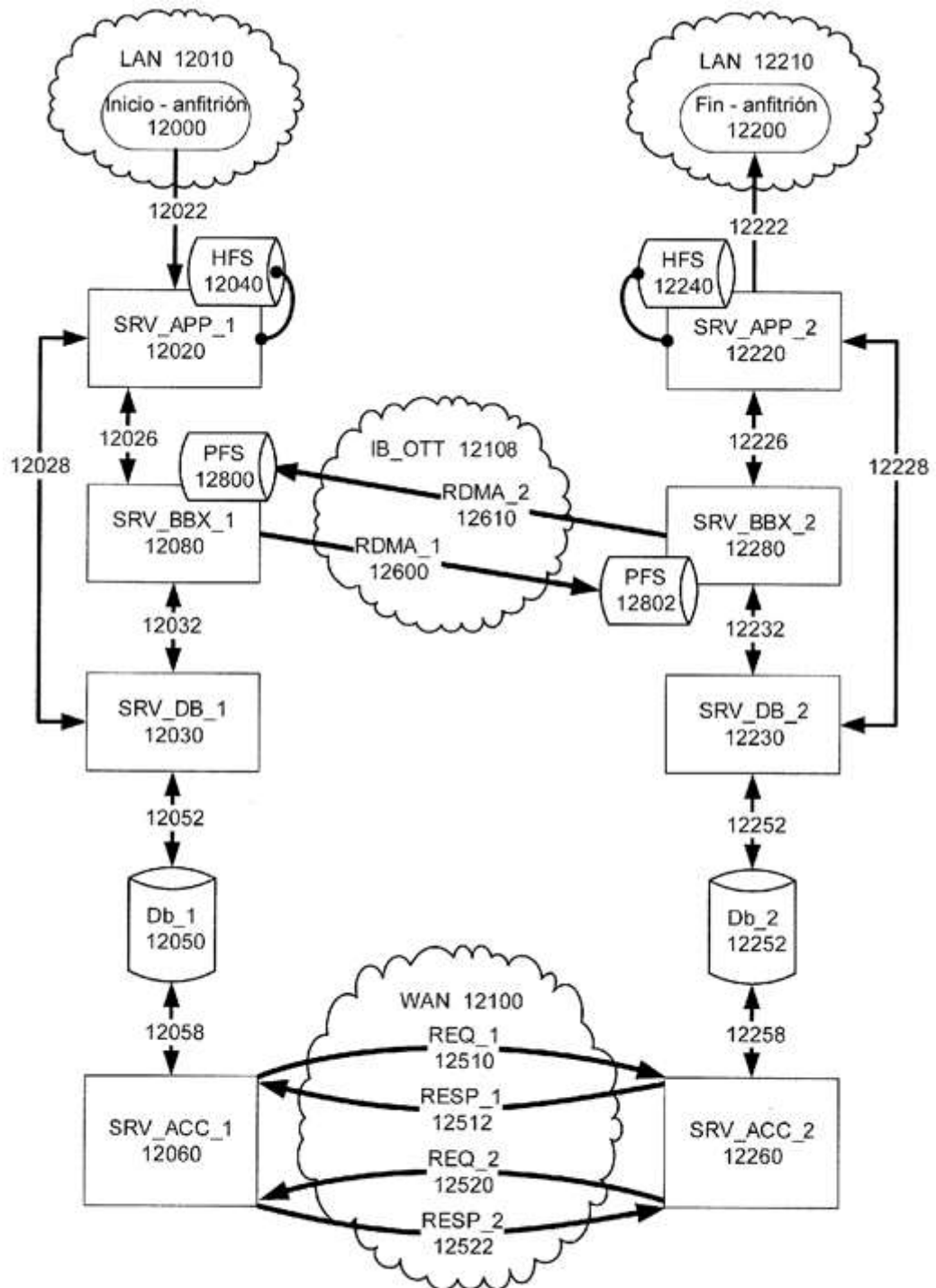


FIG. 12

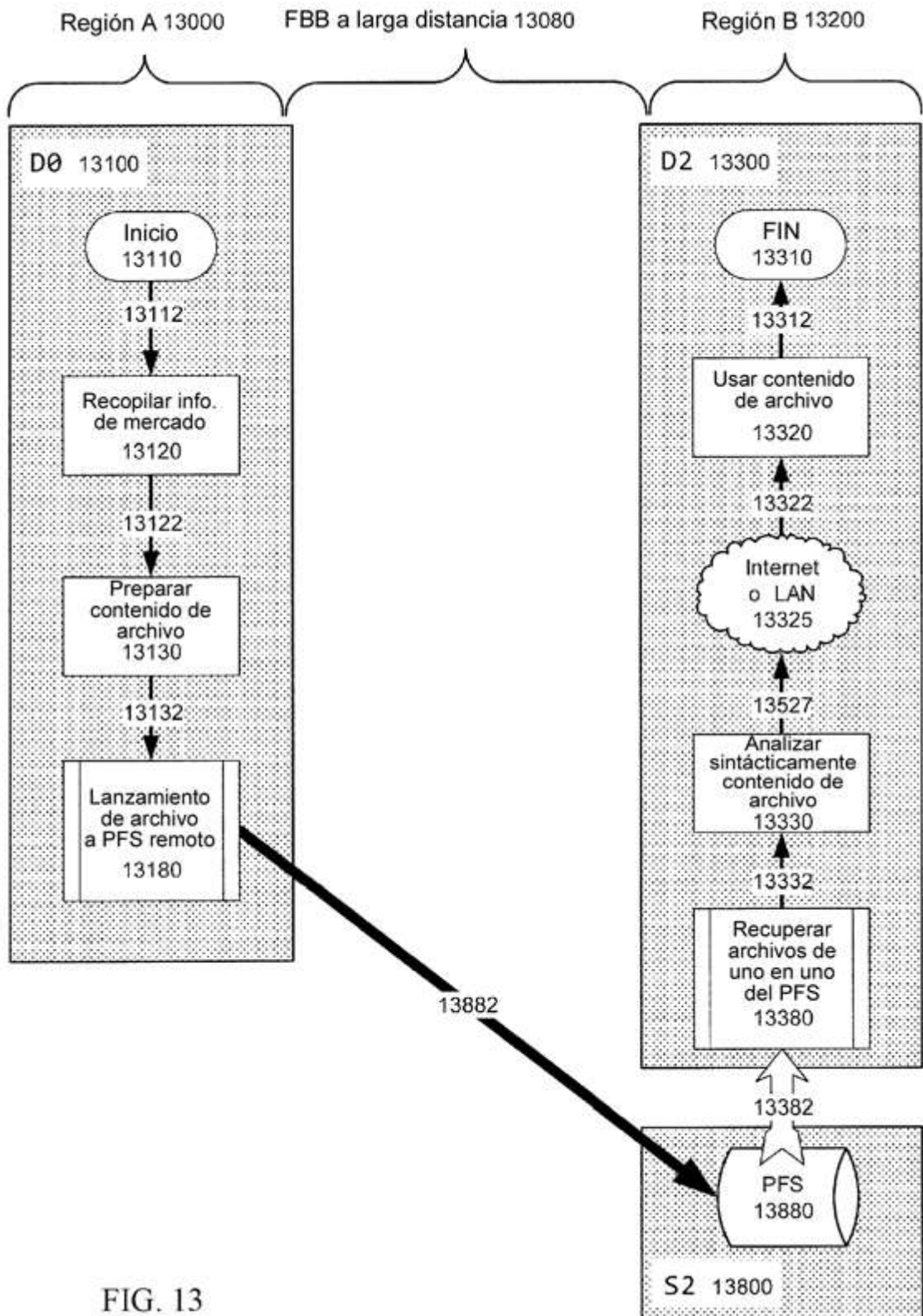


FIG. 13

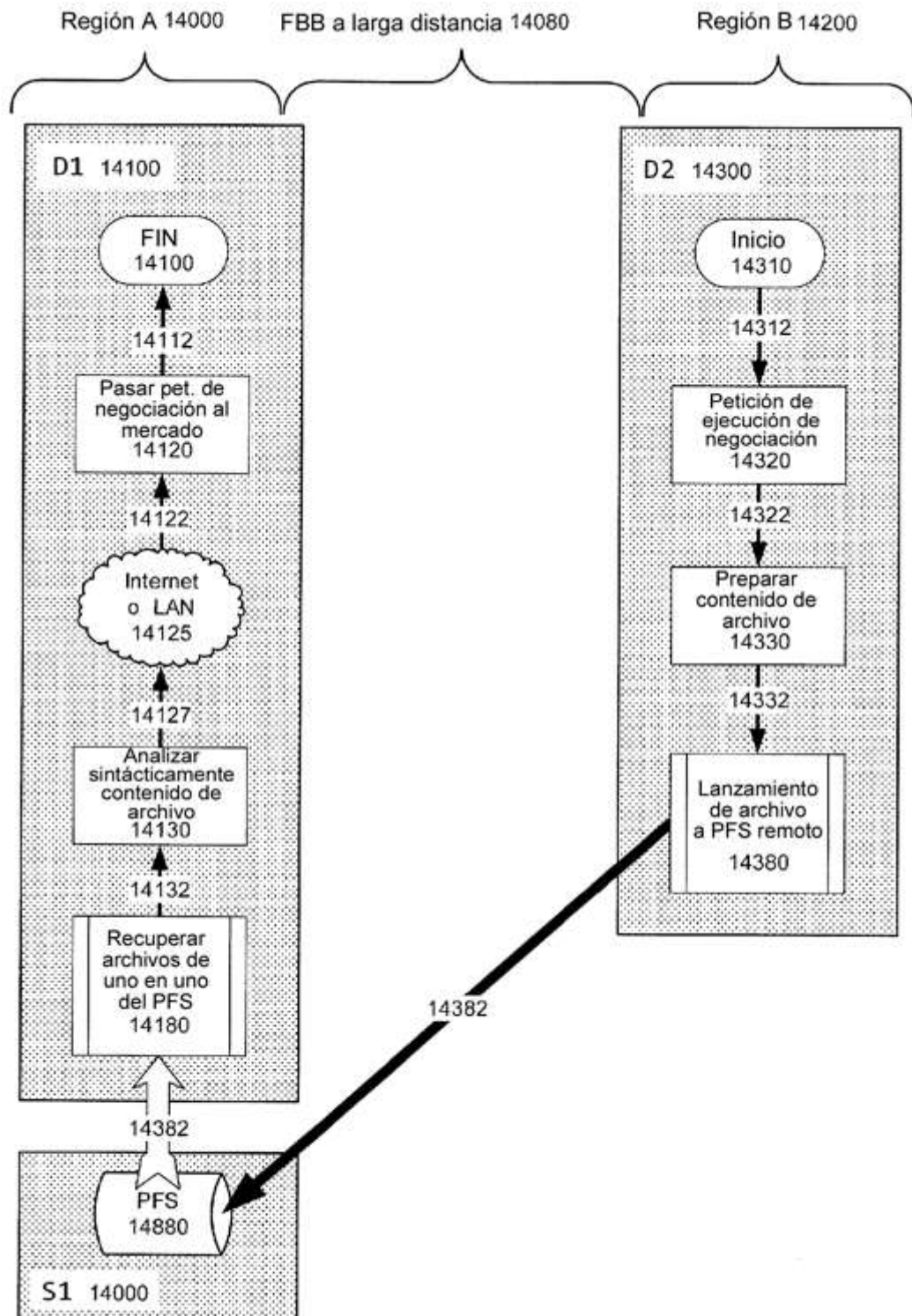


FIG. 14

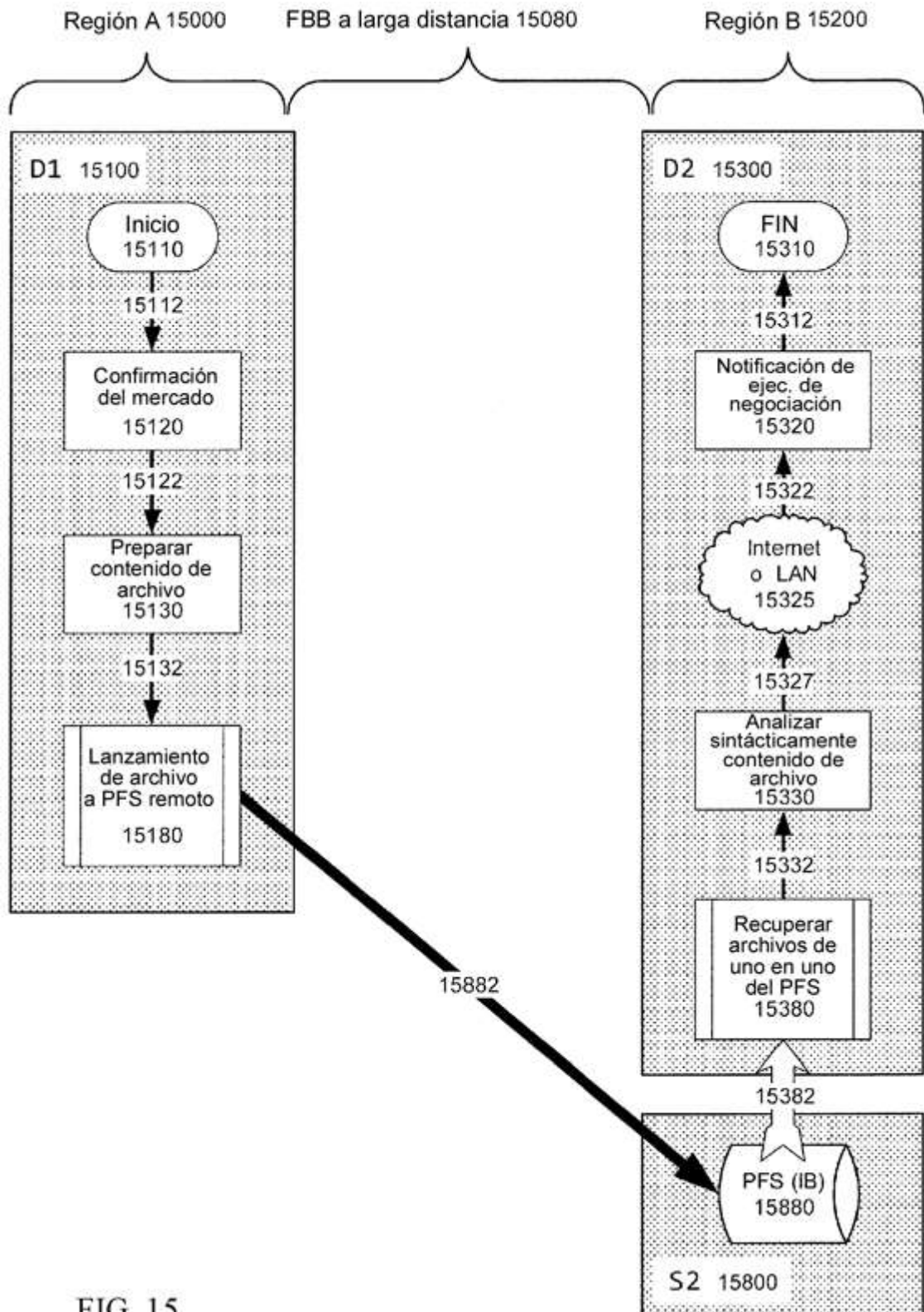


FIG. 15



FIG. 16

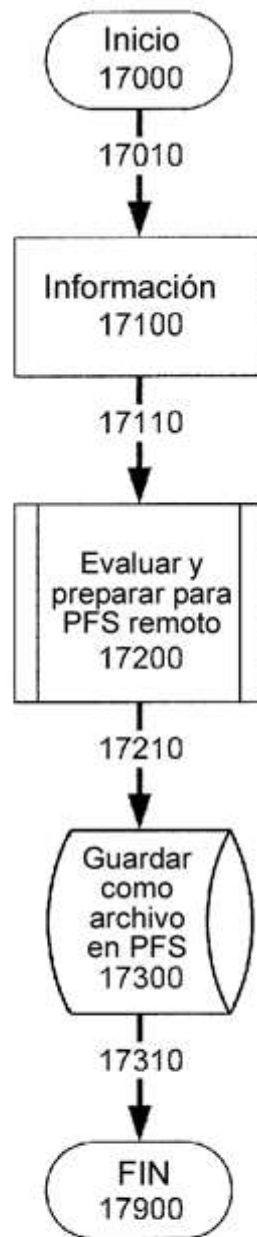


FIG. 17



FIG. 18

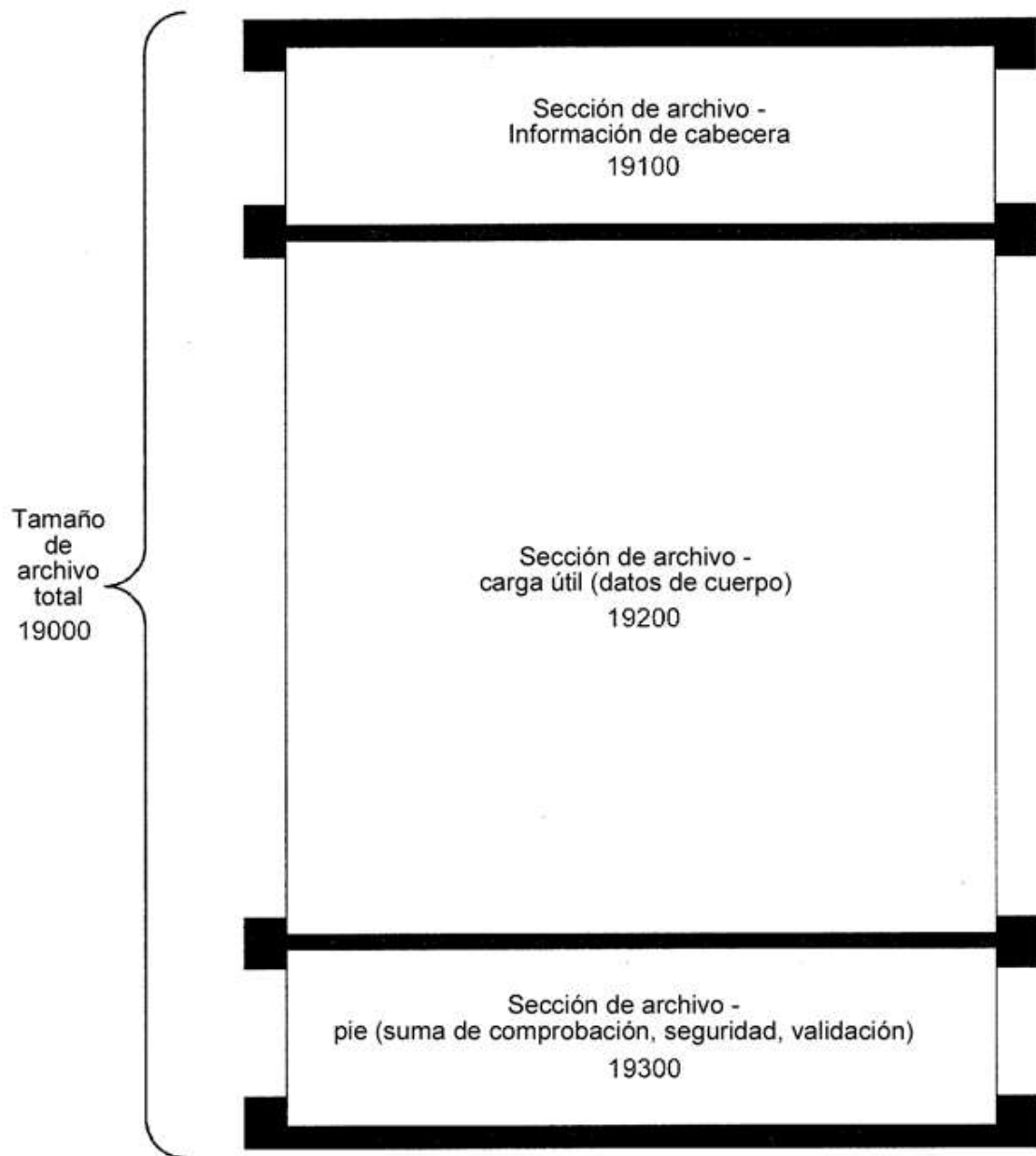


FIG. 19

Nota: Δt (delta de tiempo) es la duración de tiempo (intervalo unitario) expresada en milisegundos ms o en microsegundos μs .

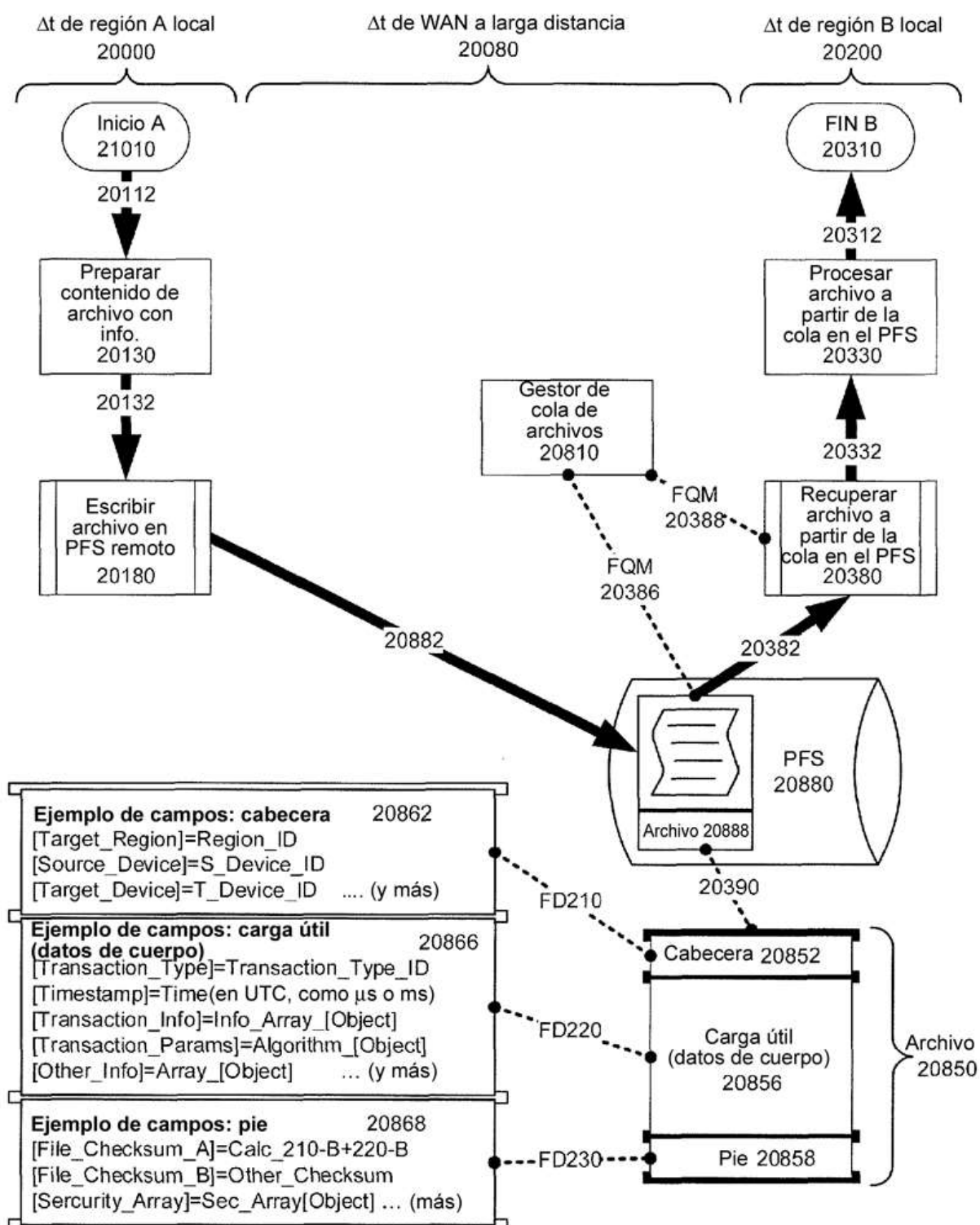


FIG. 20

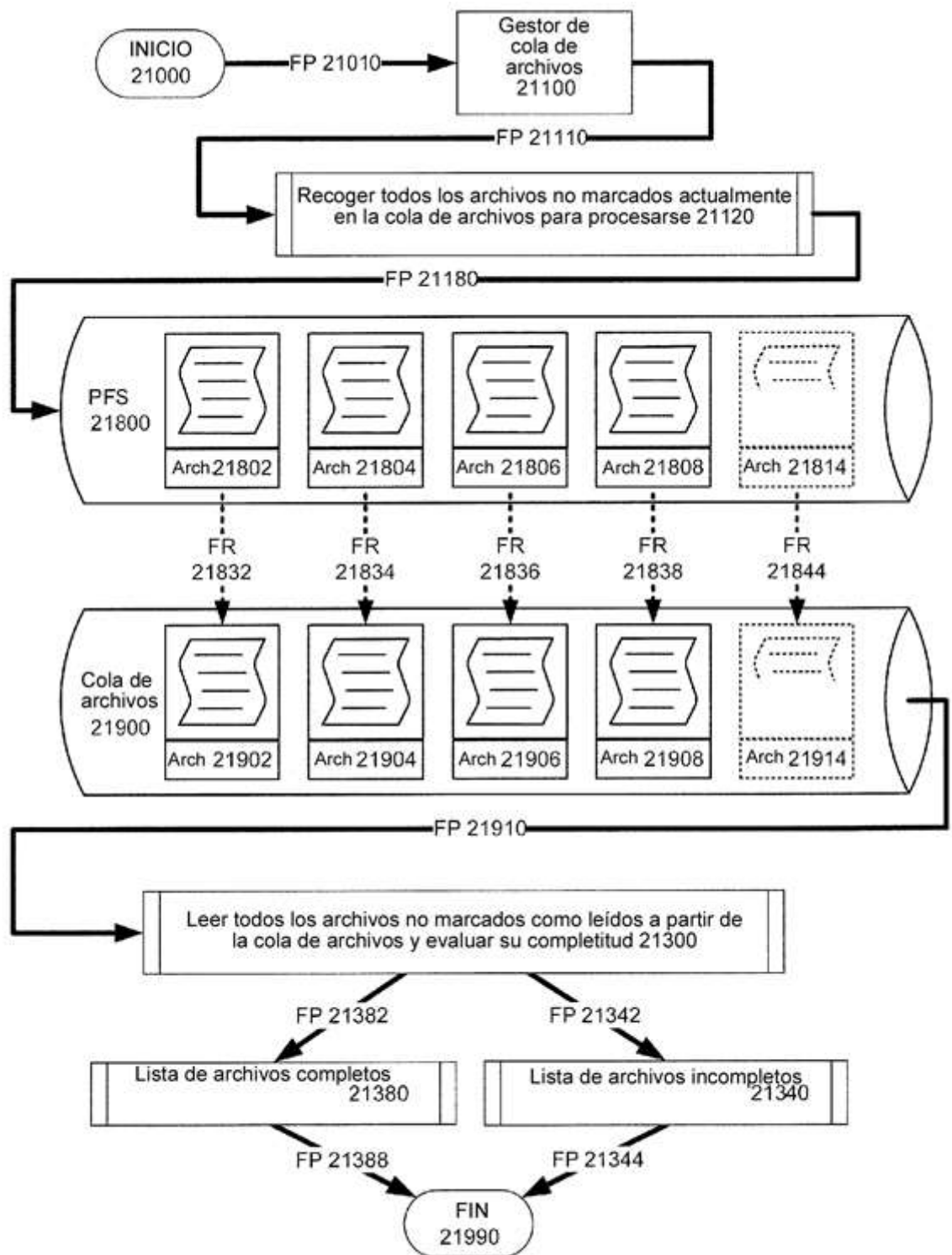


FIG. 21

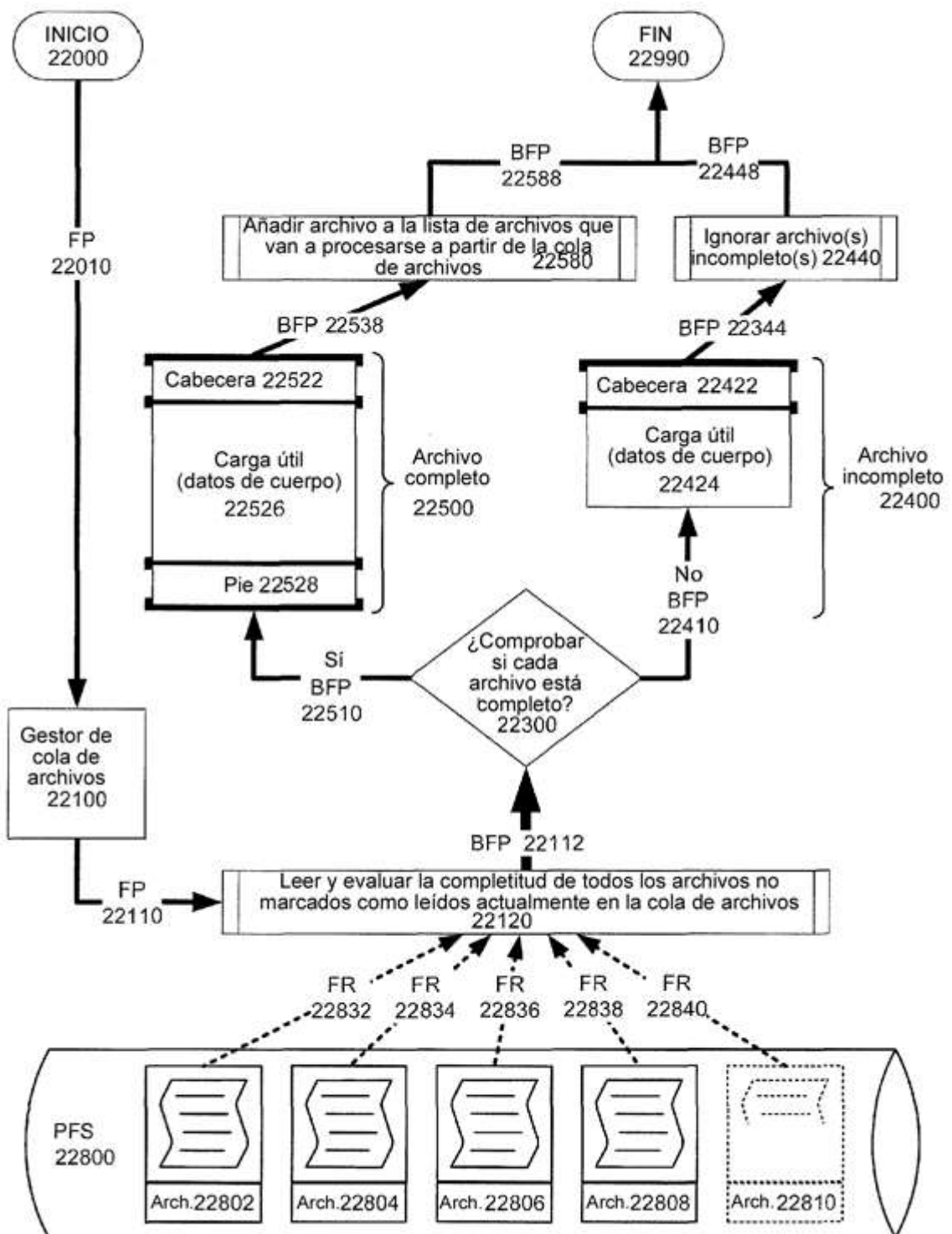


FIG. 22

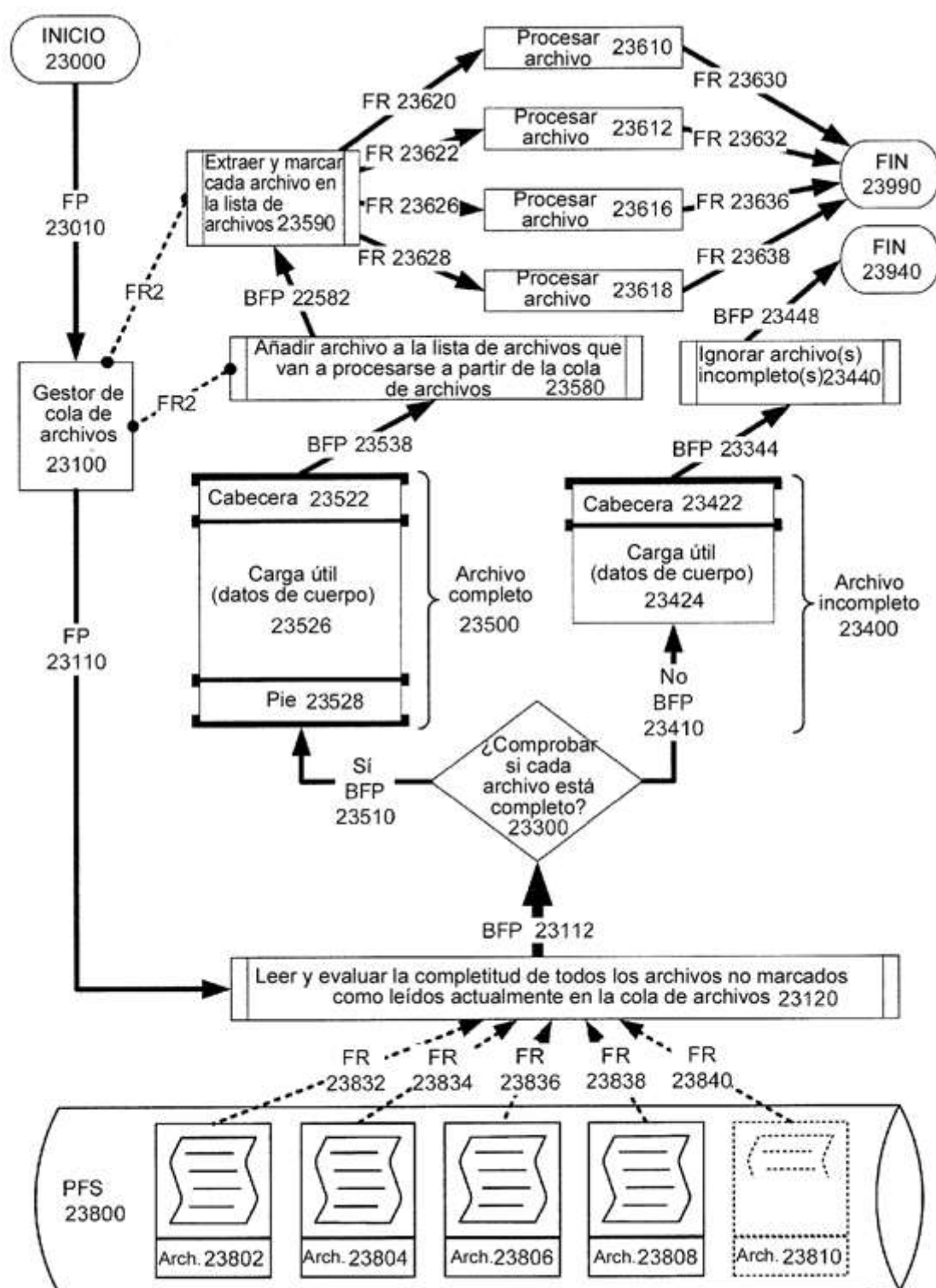


FIG. 23

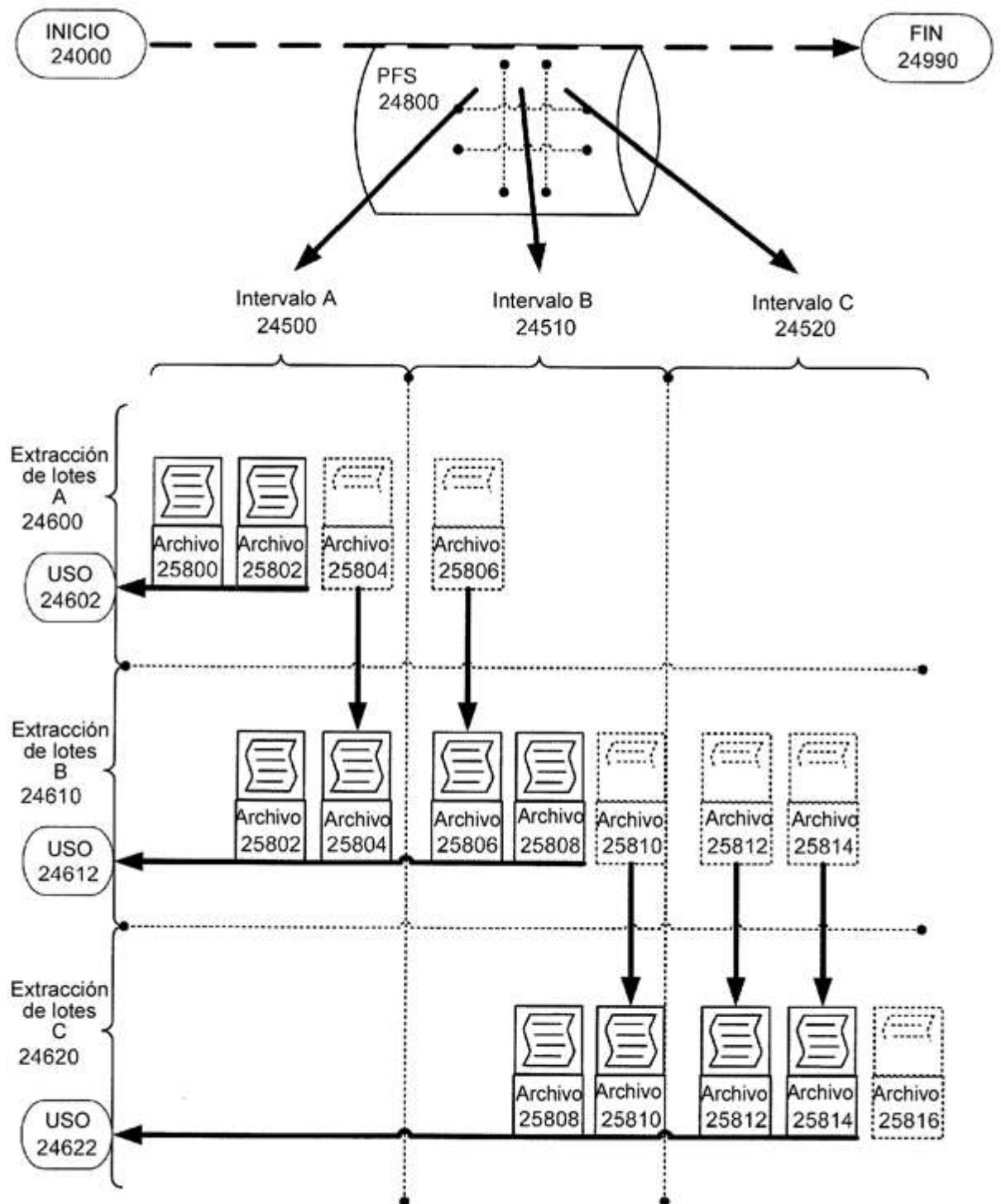


FIG. 24

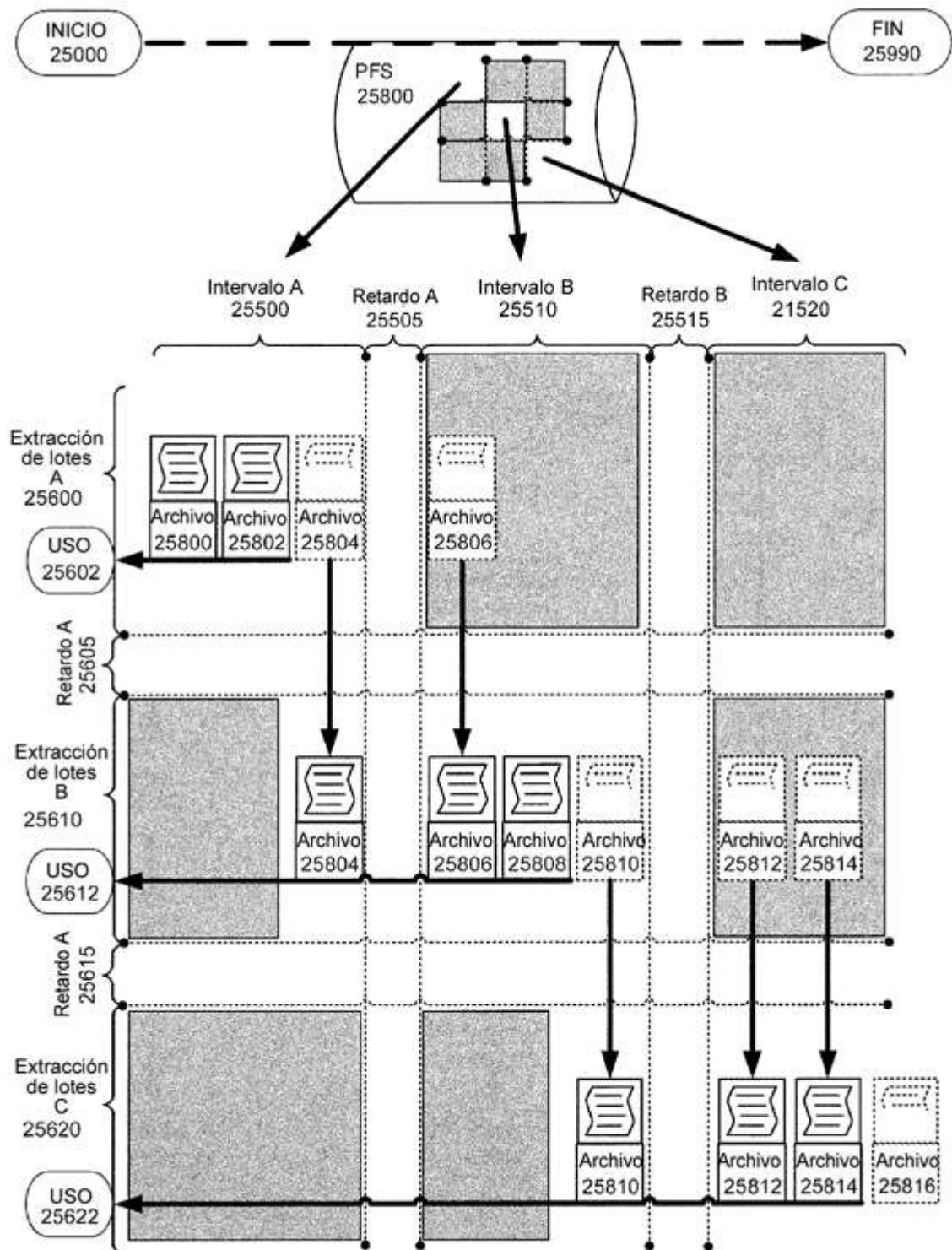


FIG. 25

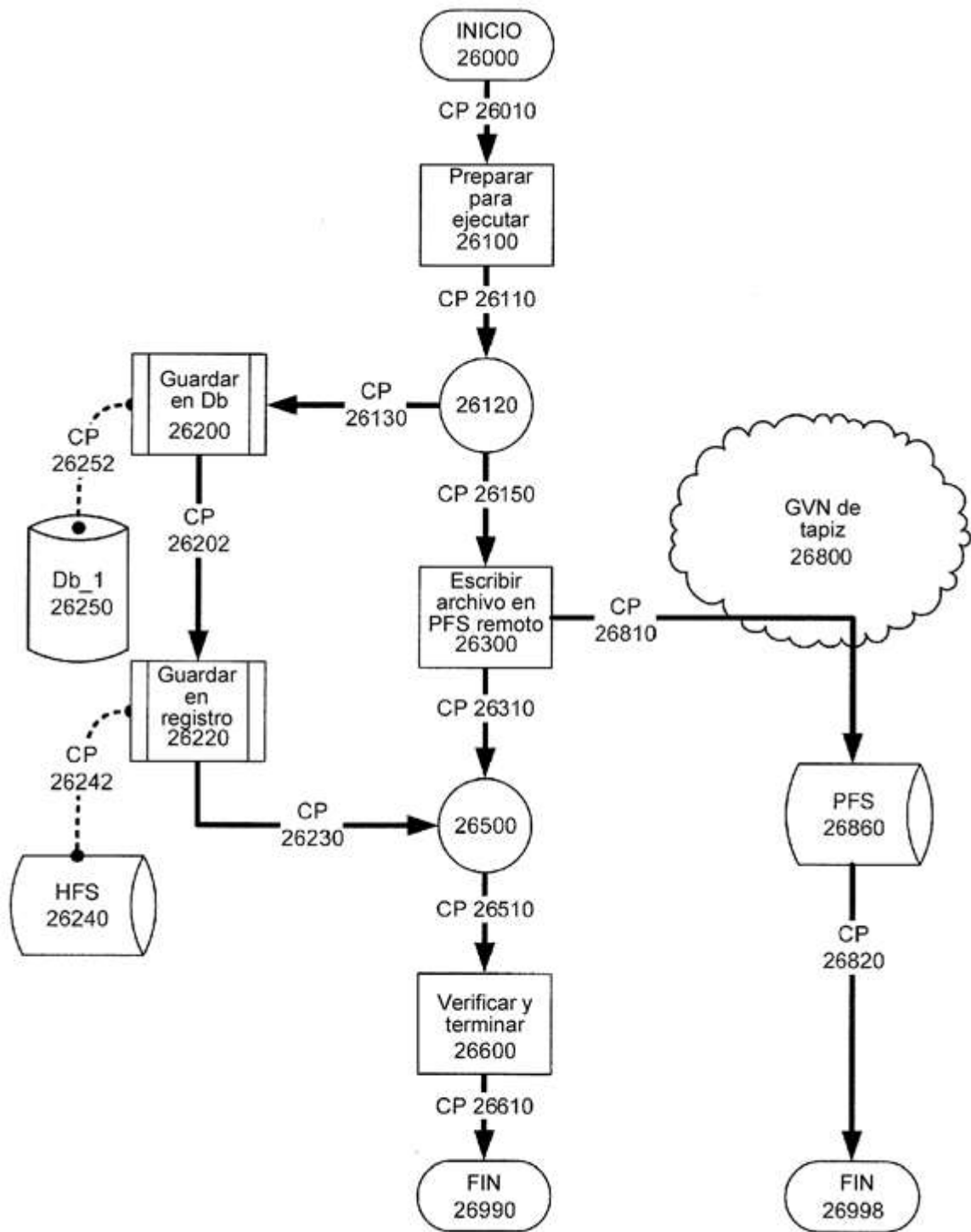


FIG. 26

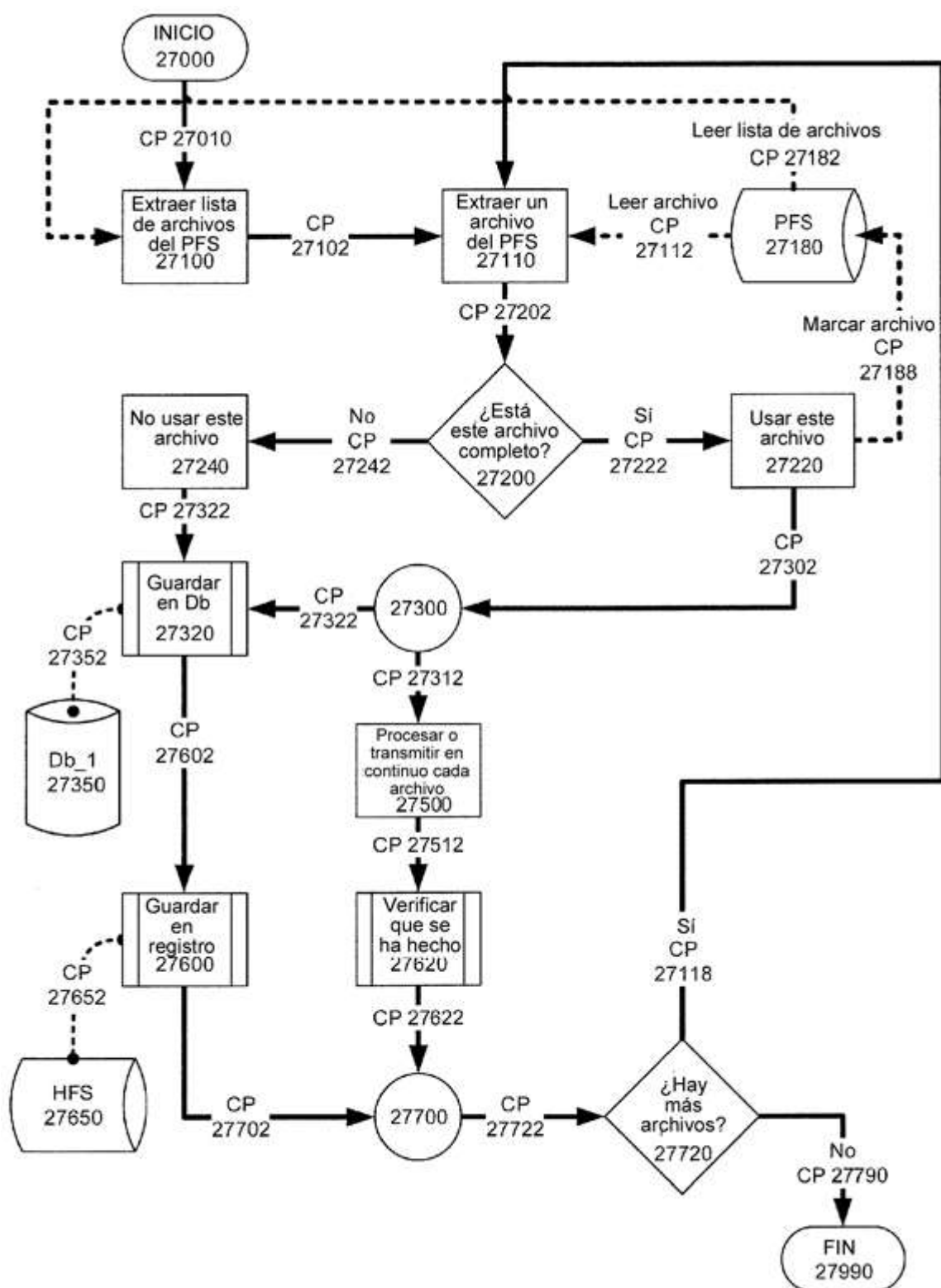


FIG. 27

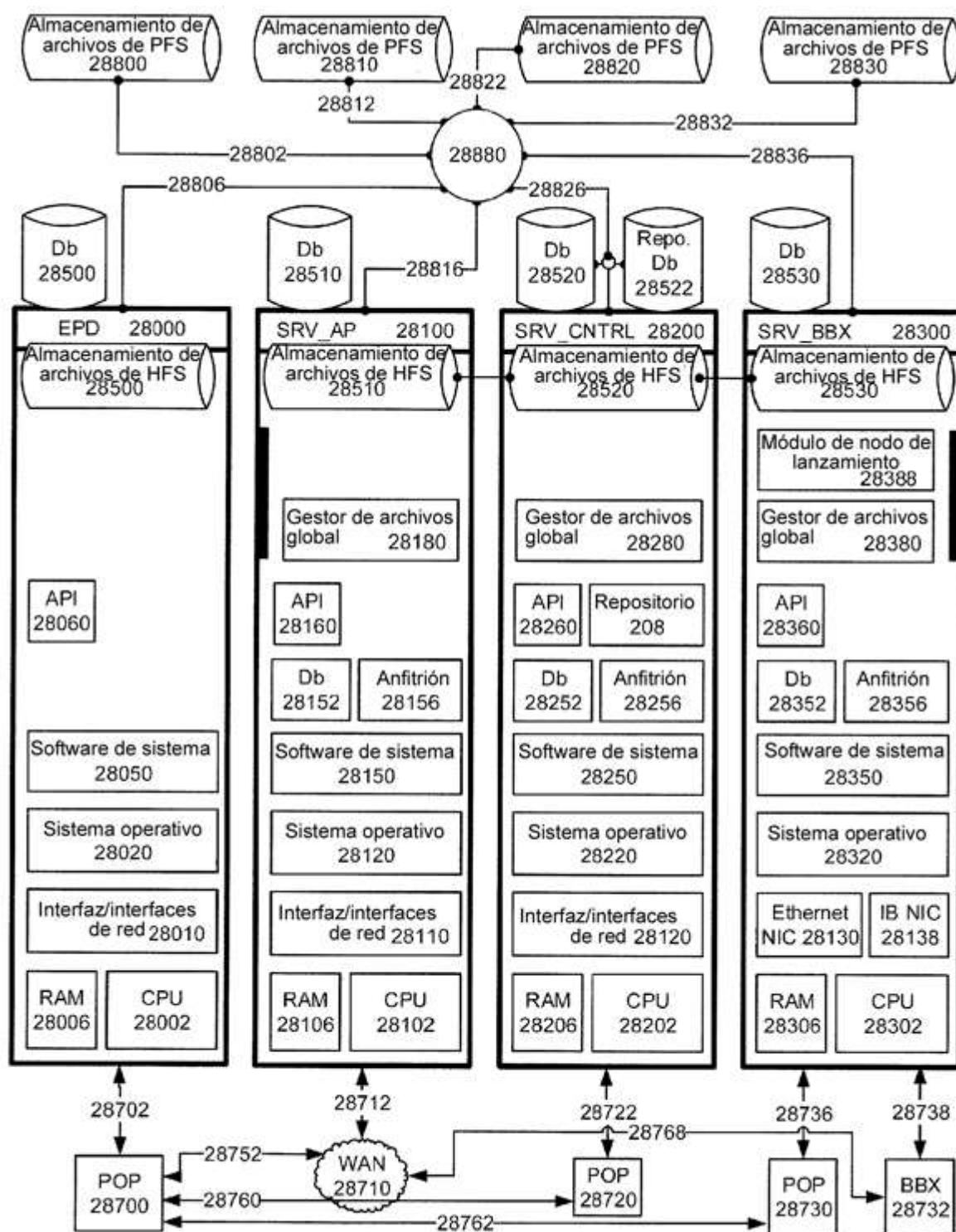


FIG. 28

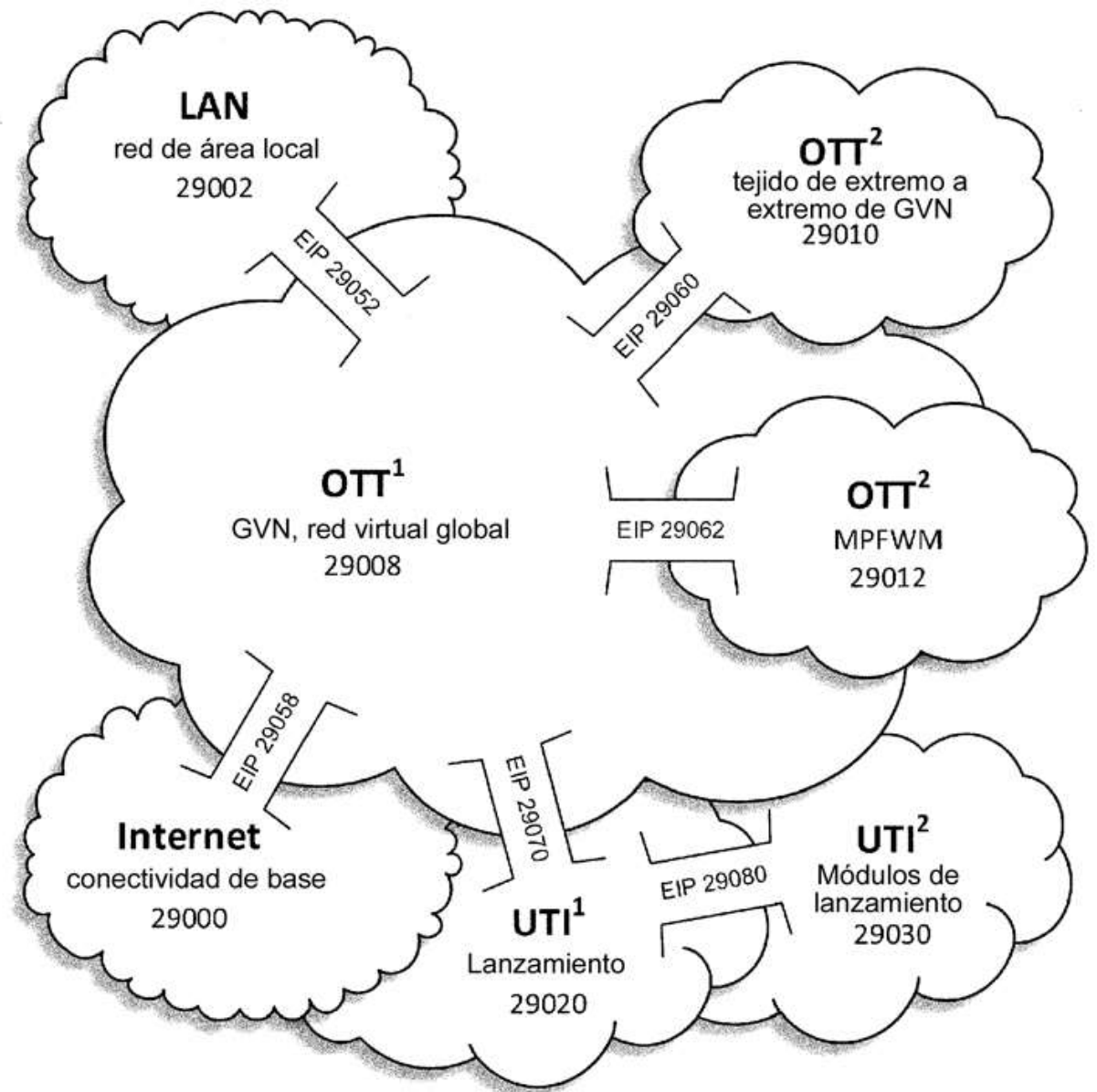


FIG. 29

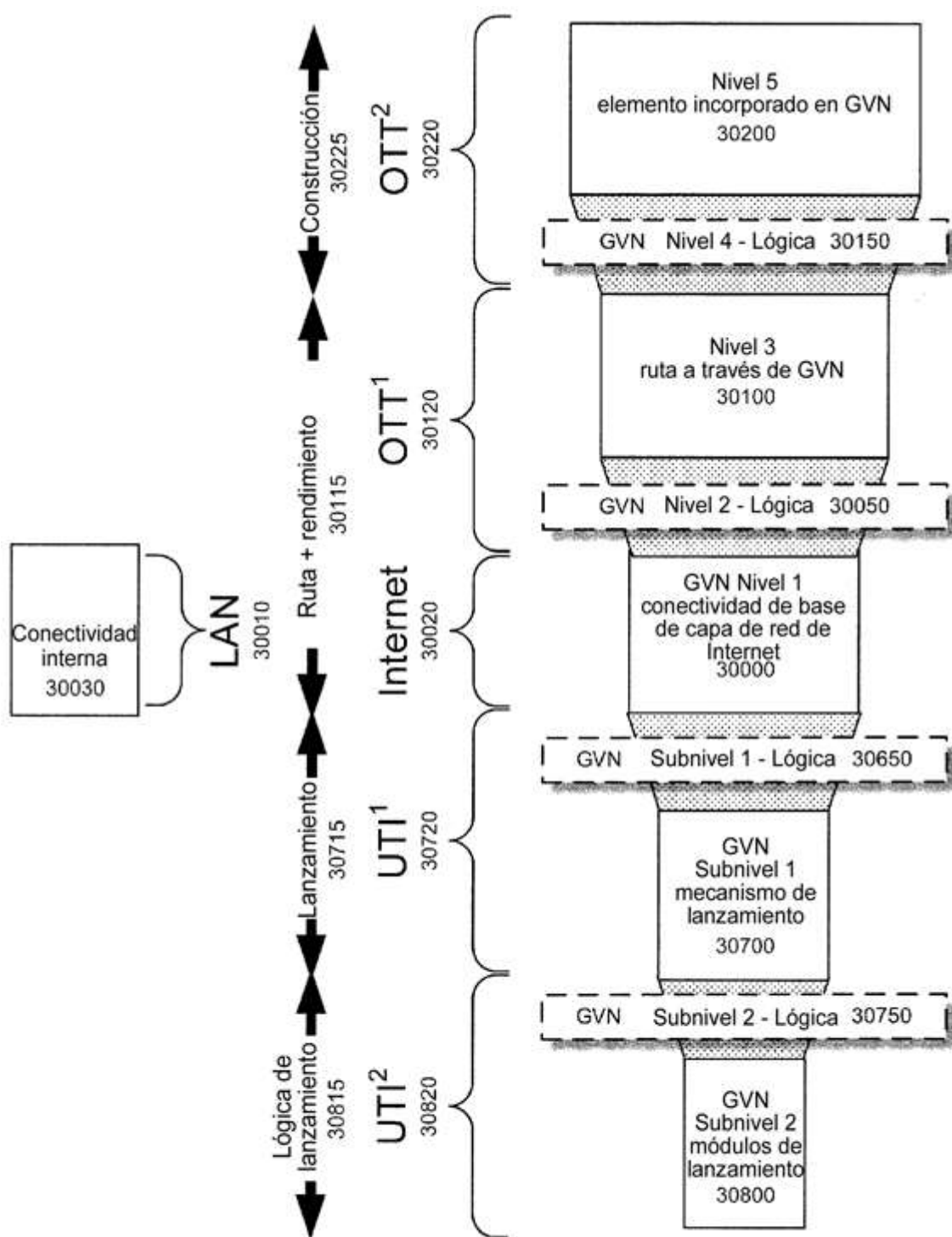


FIG. 30

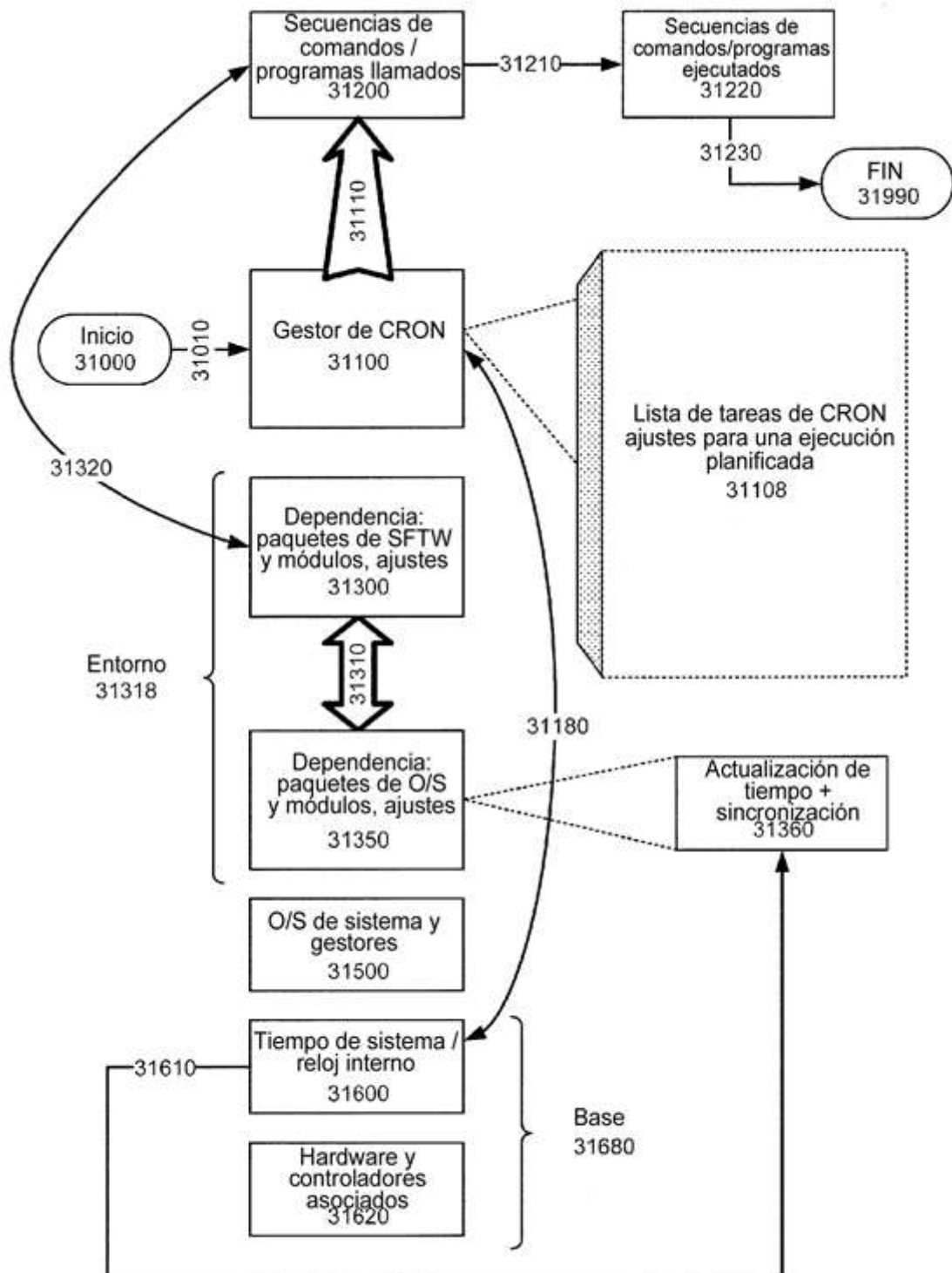


FIG. 31

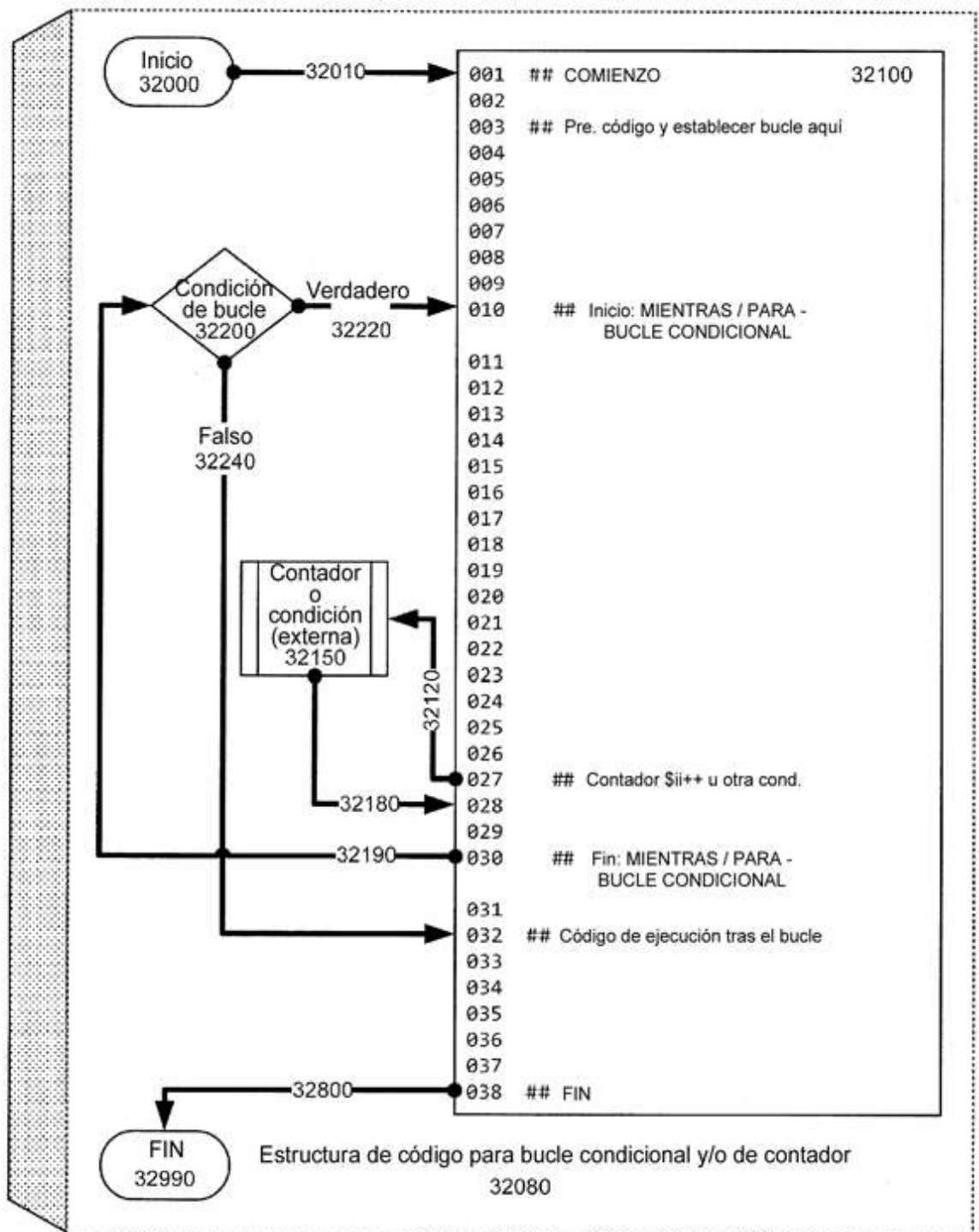


FIG. 32

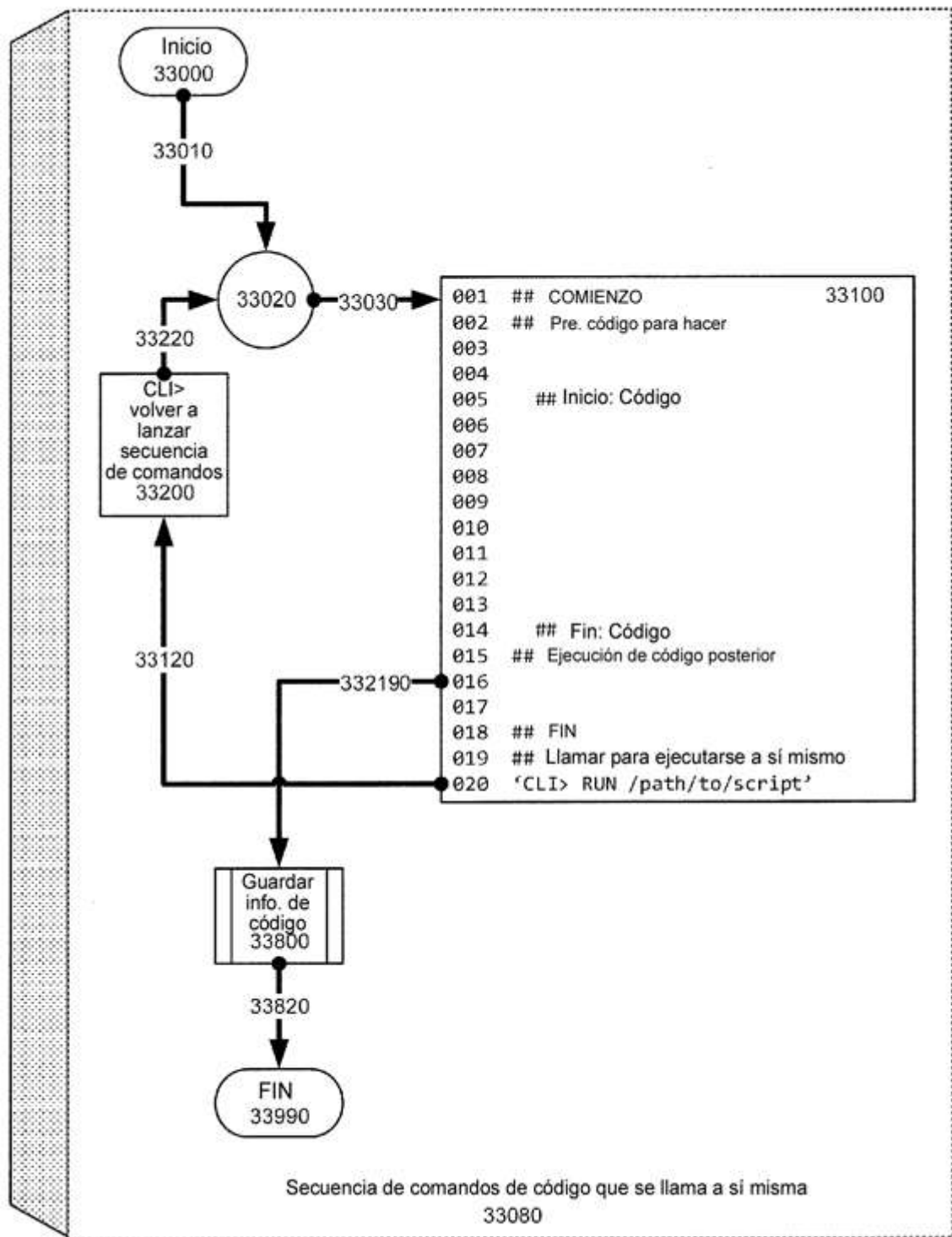


FIG. 33

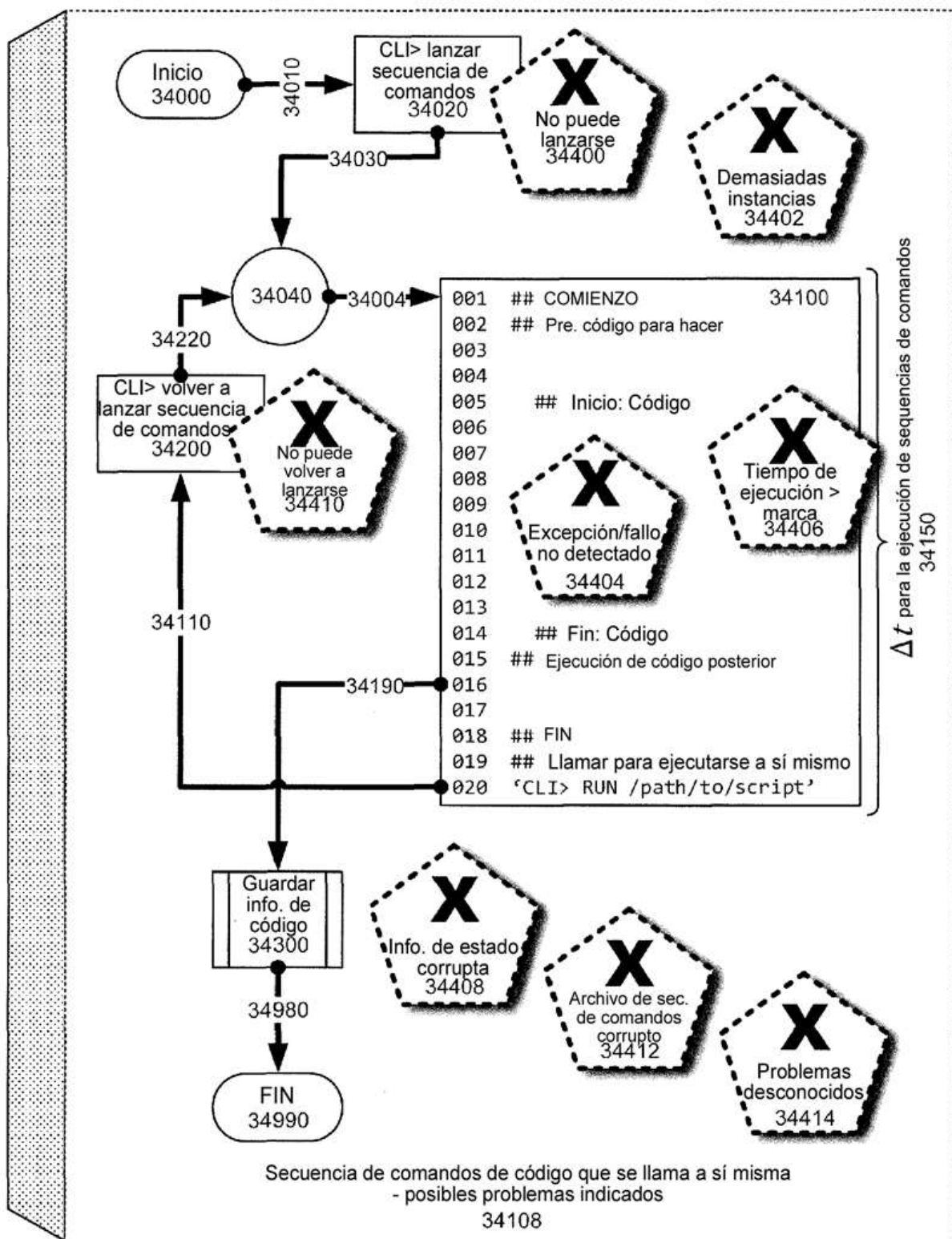
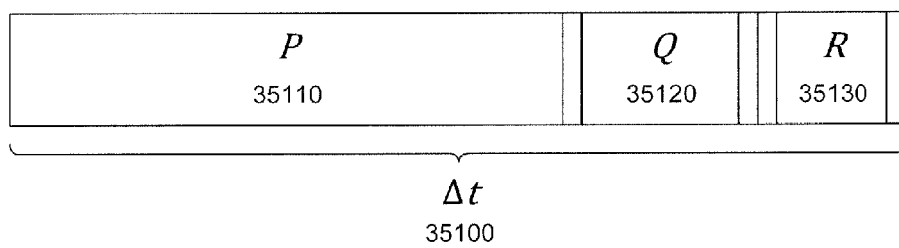


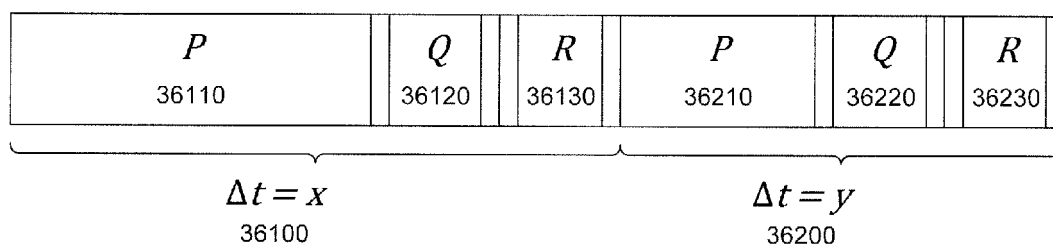
FIG. 34



$$\Delta t = \Delta P + \Delta Q + \Delta R$$

35000

FIG. 35



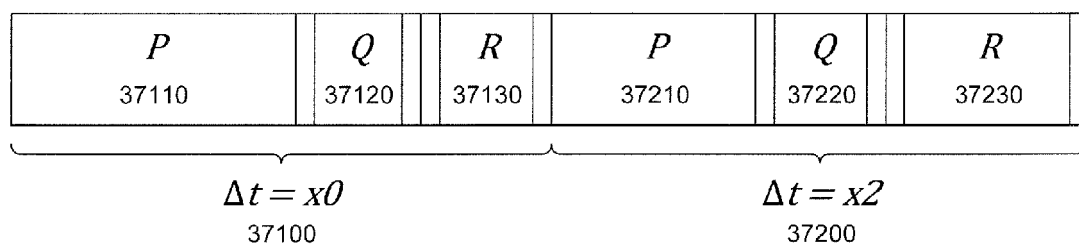
$$\Delta t = \Delta P + \Delta Q + \Delta R$$

36000

$$x \neq y$$

36010

FIG. 36



$$\Delta t = \Delta P + \Delta Q + \Delta R$$

37000

$$x0 = x2$$

37010

FIG. 37

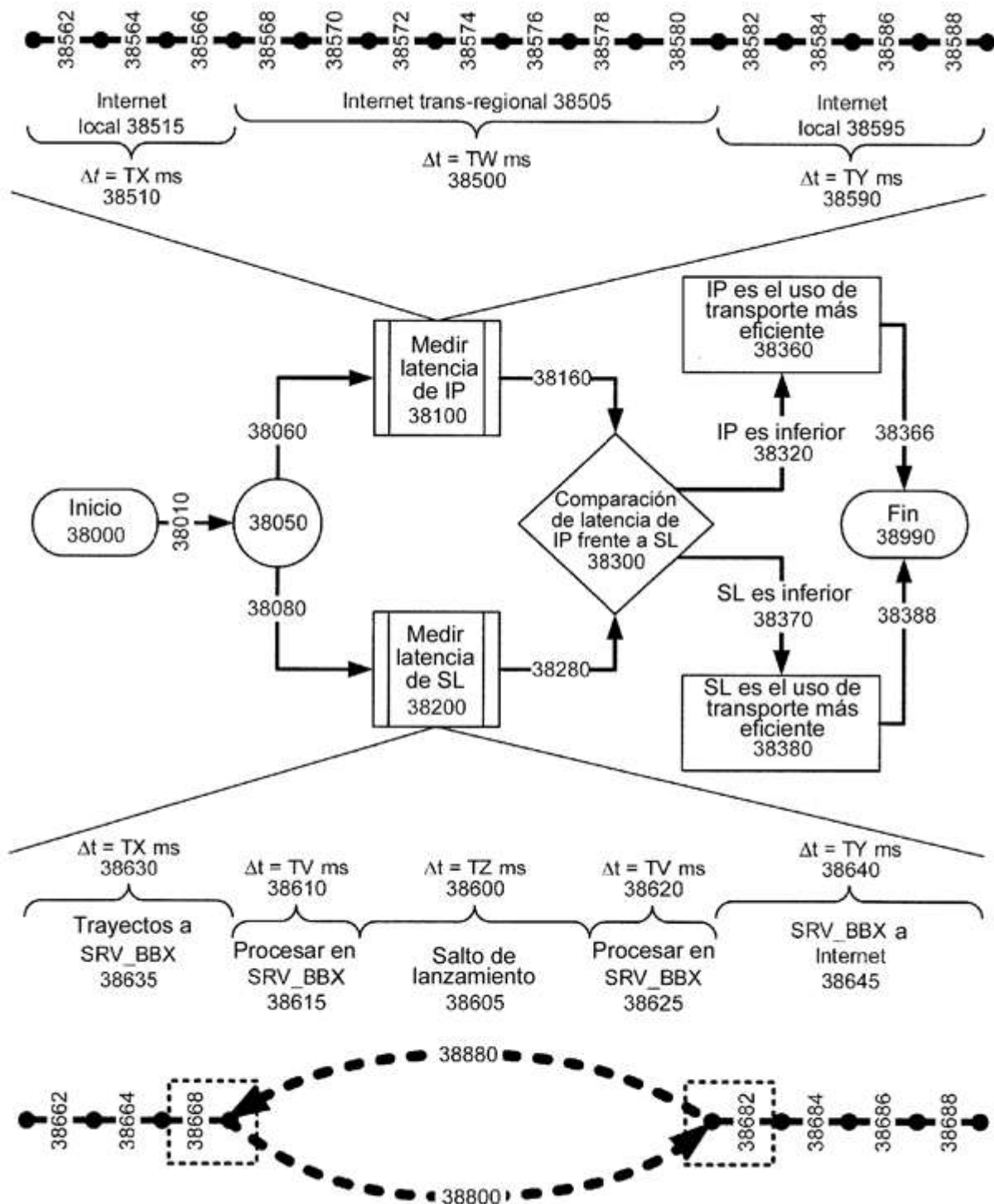


FIG. 38

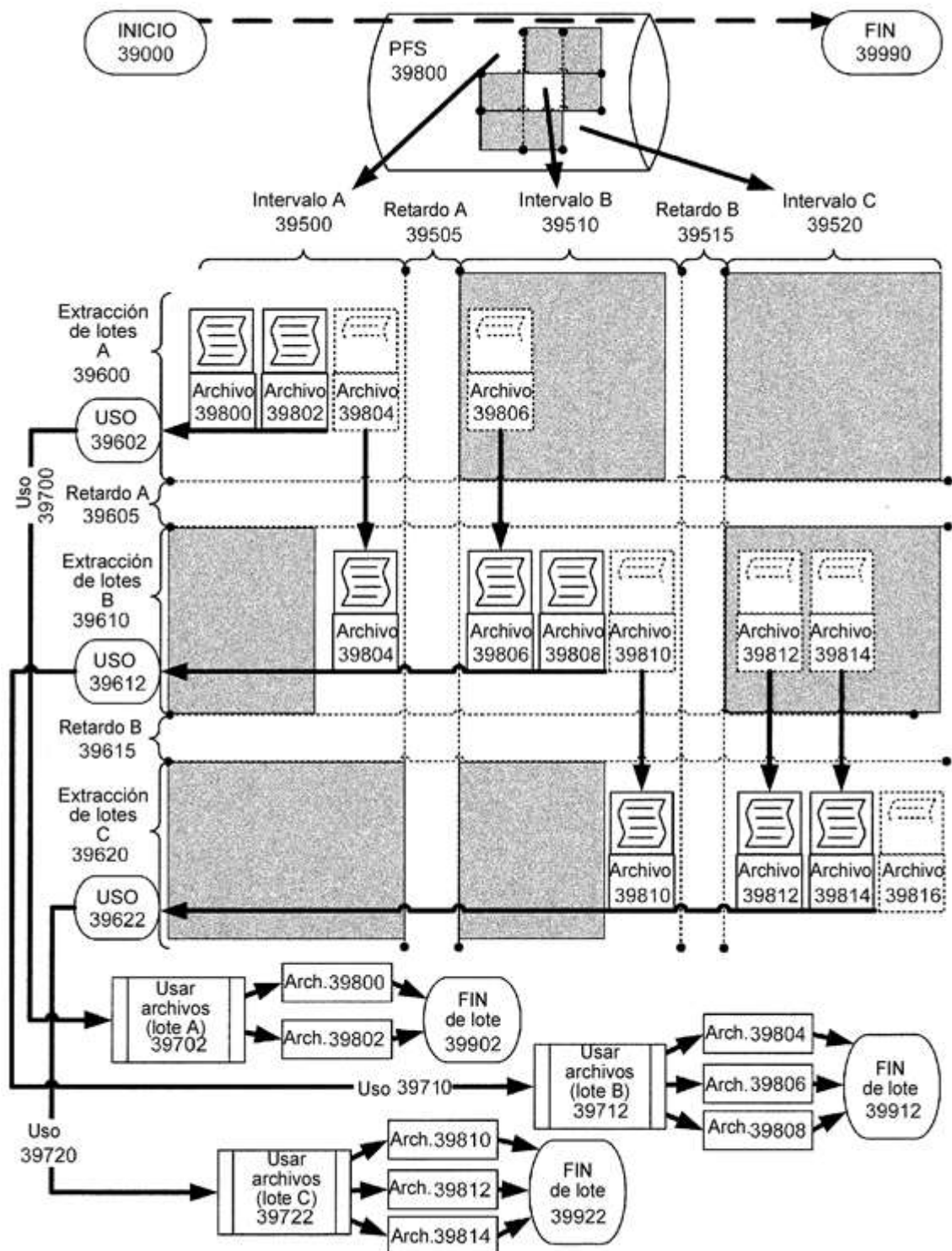


FIG. 39

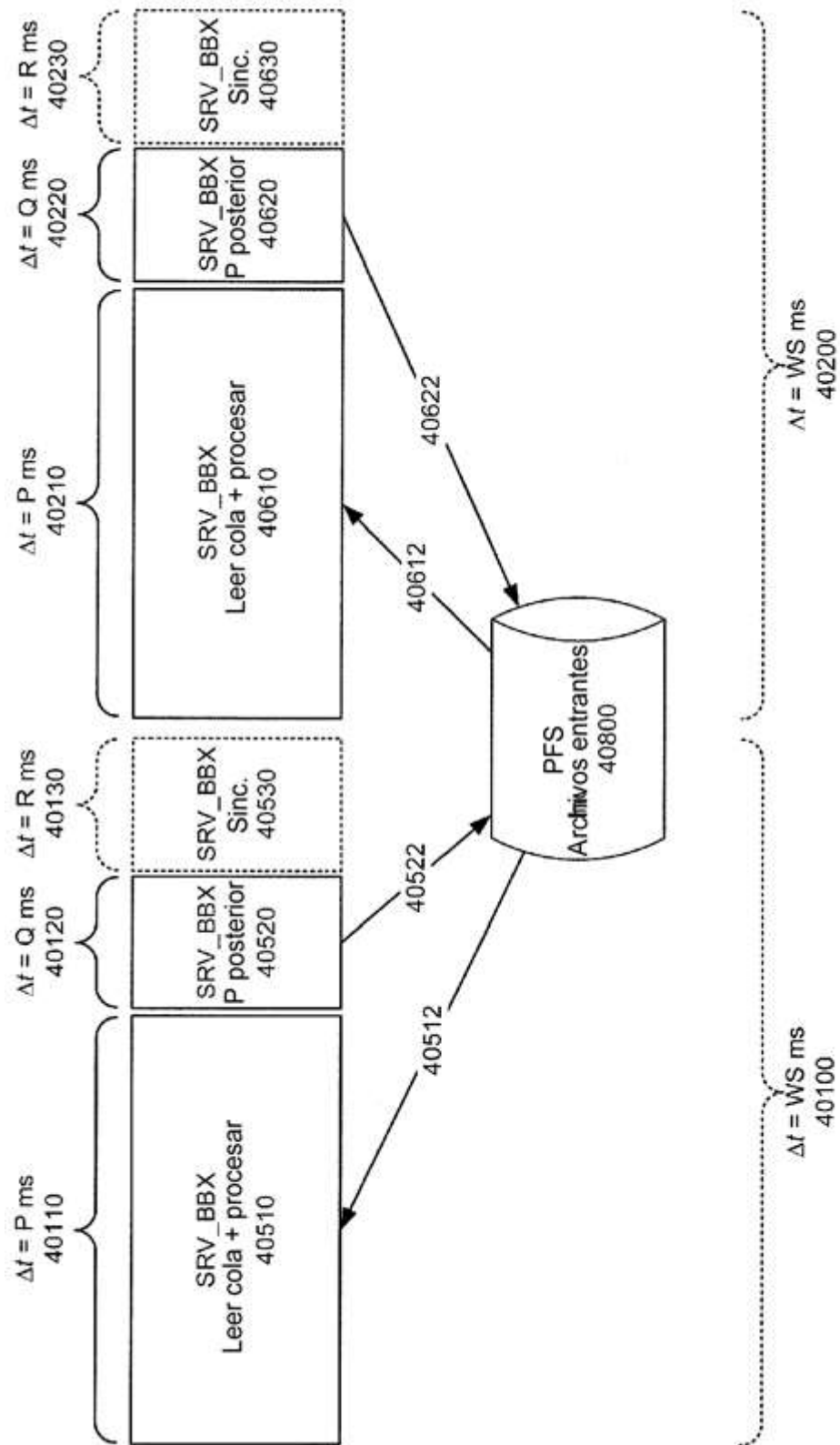


FIG. 40

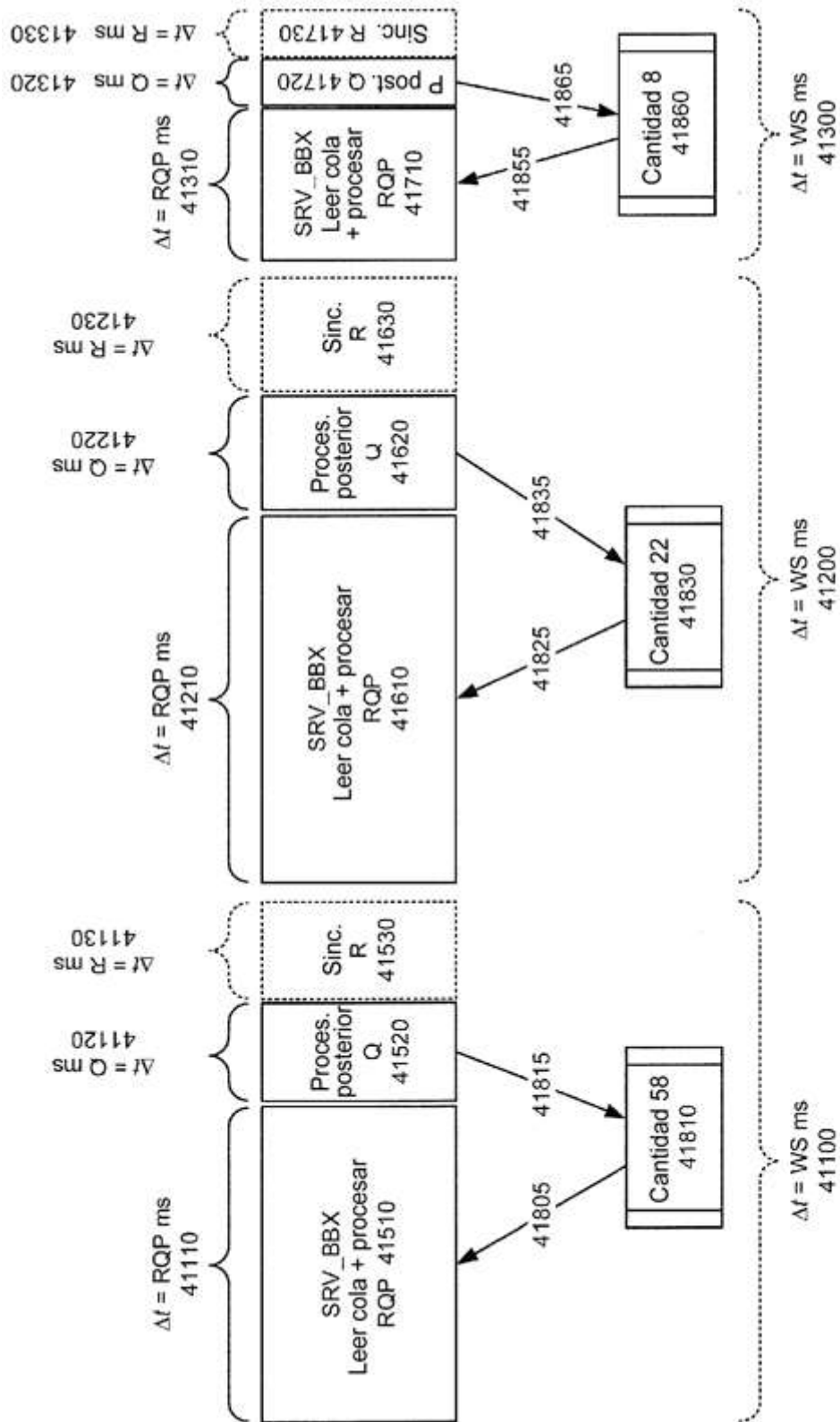


FIG. 41

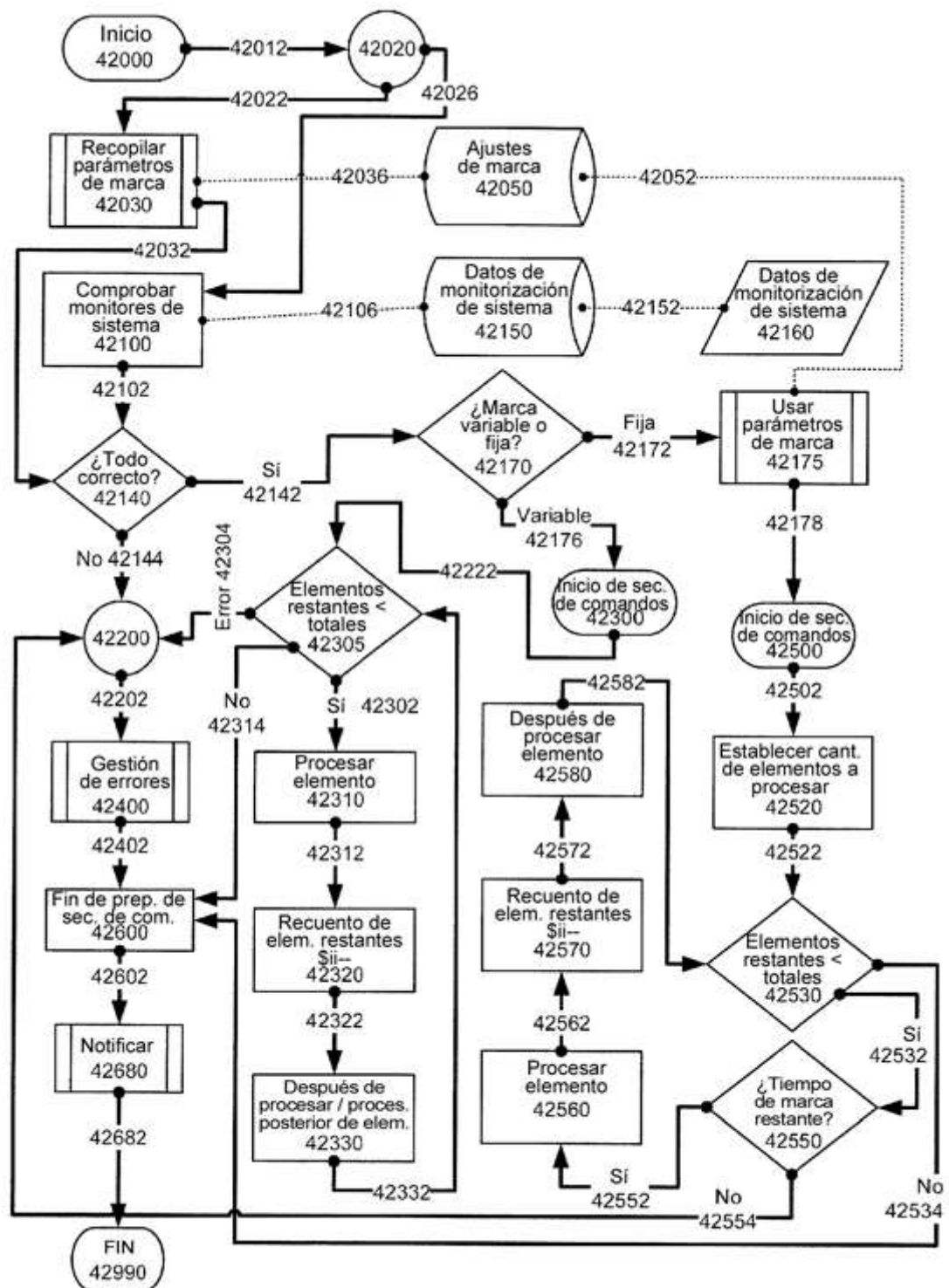


FIG. 42

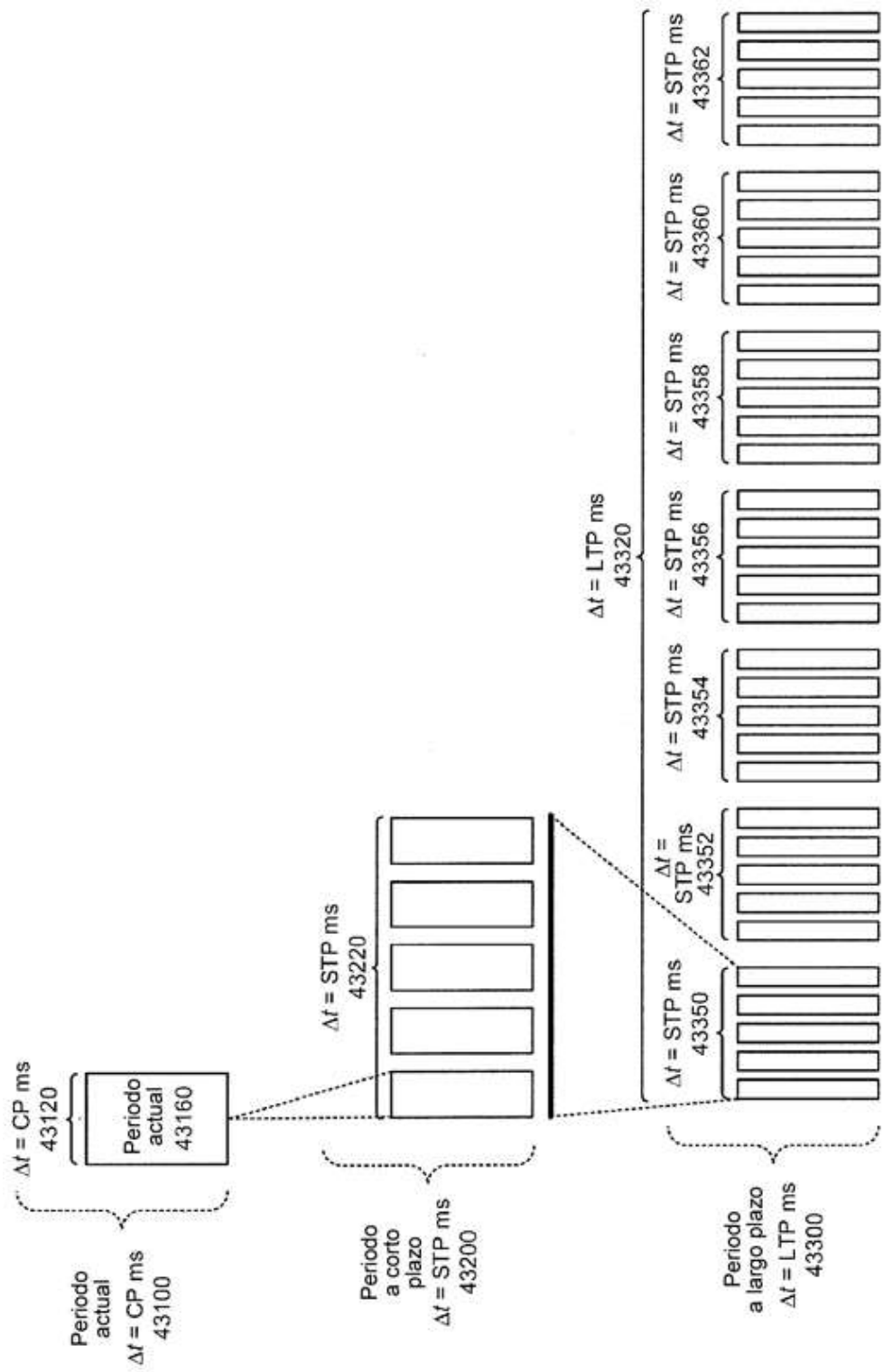


FIG. 43

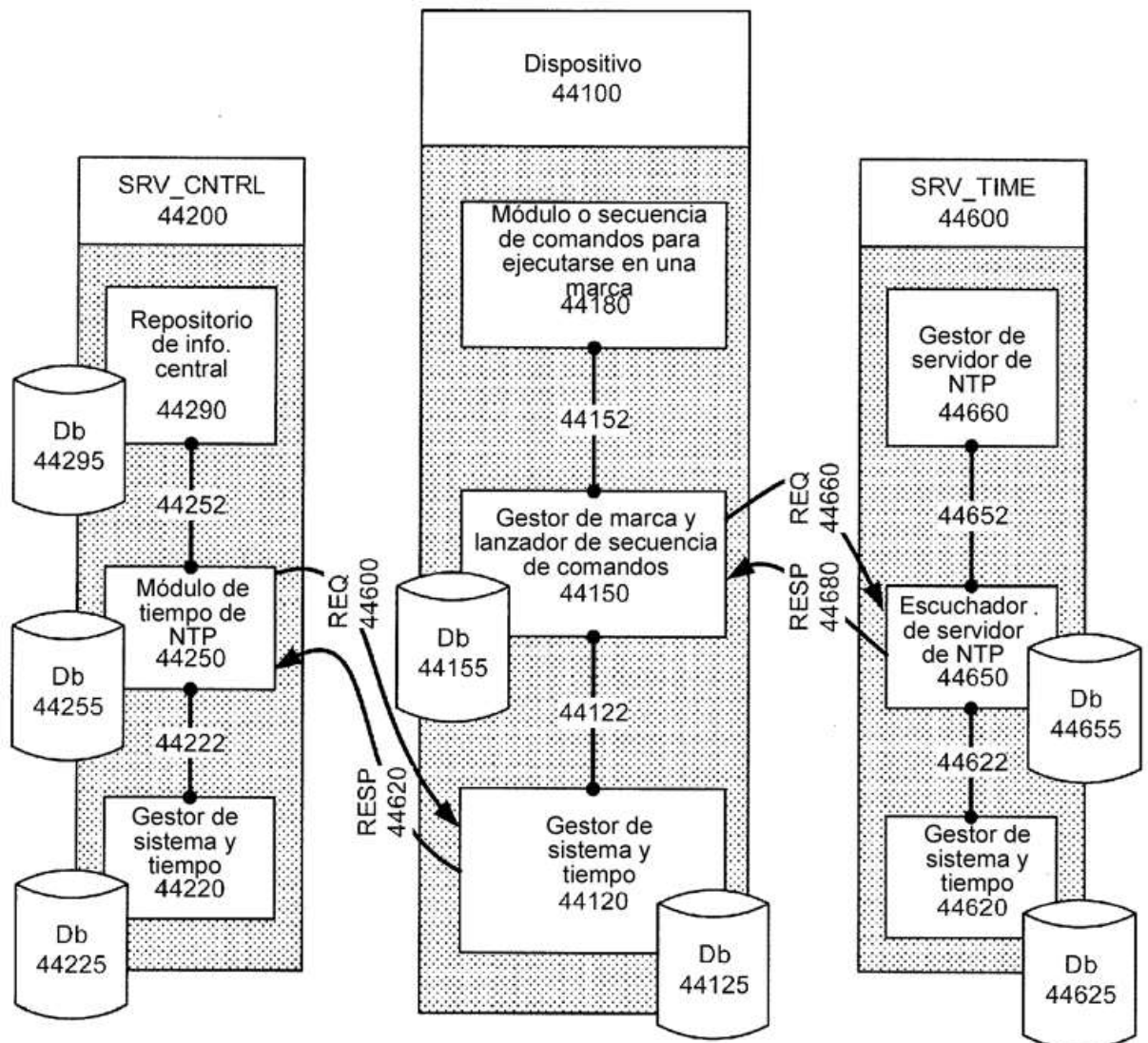


FIG. 44

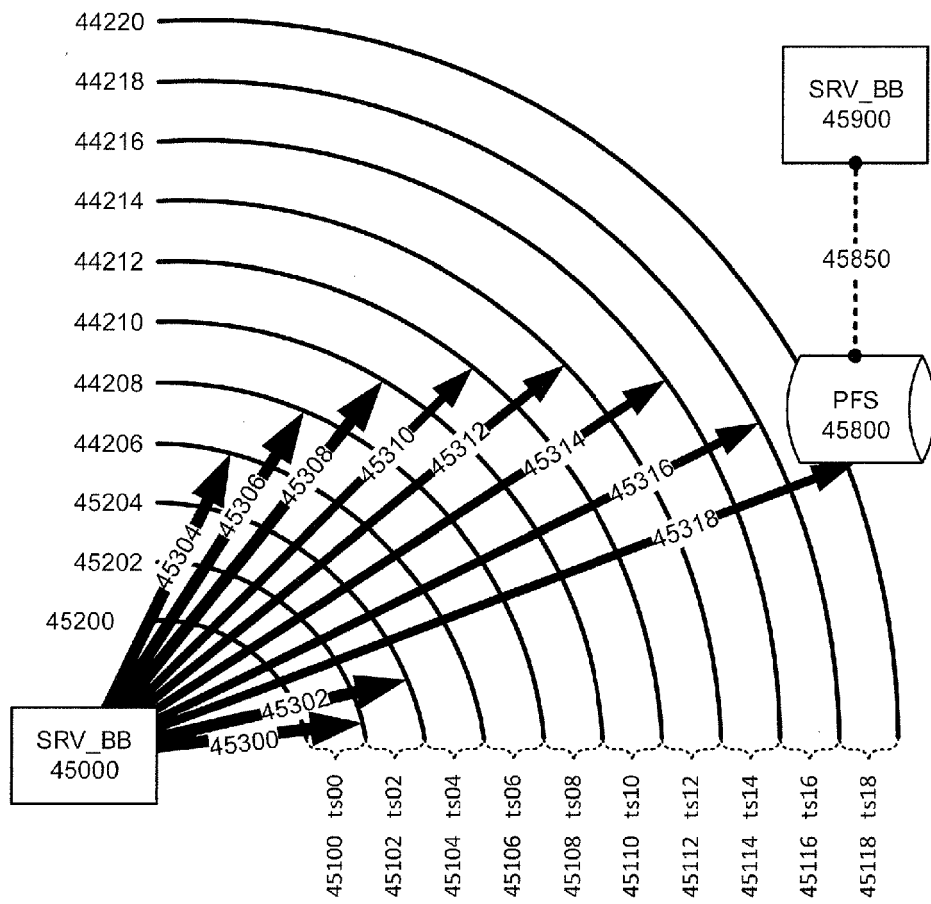


FIG. 45

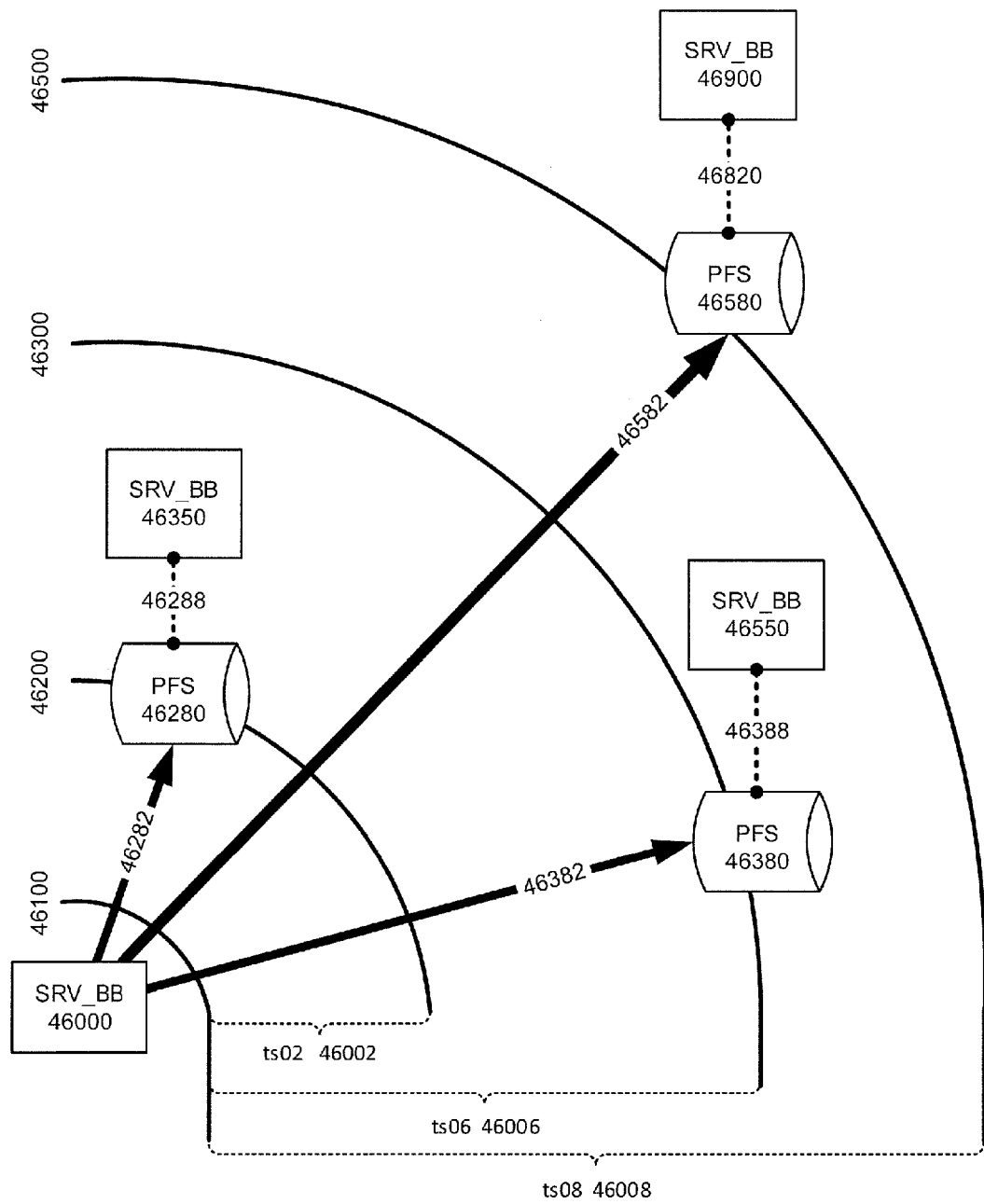


FIG. 46

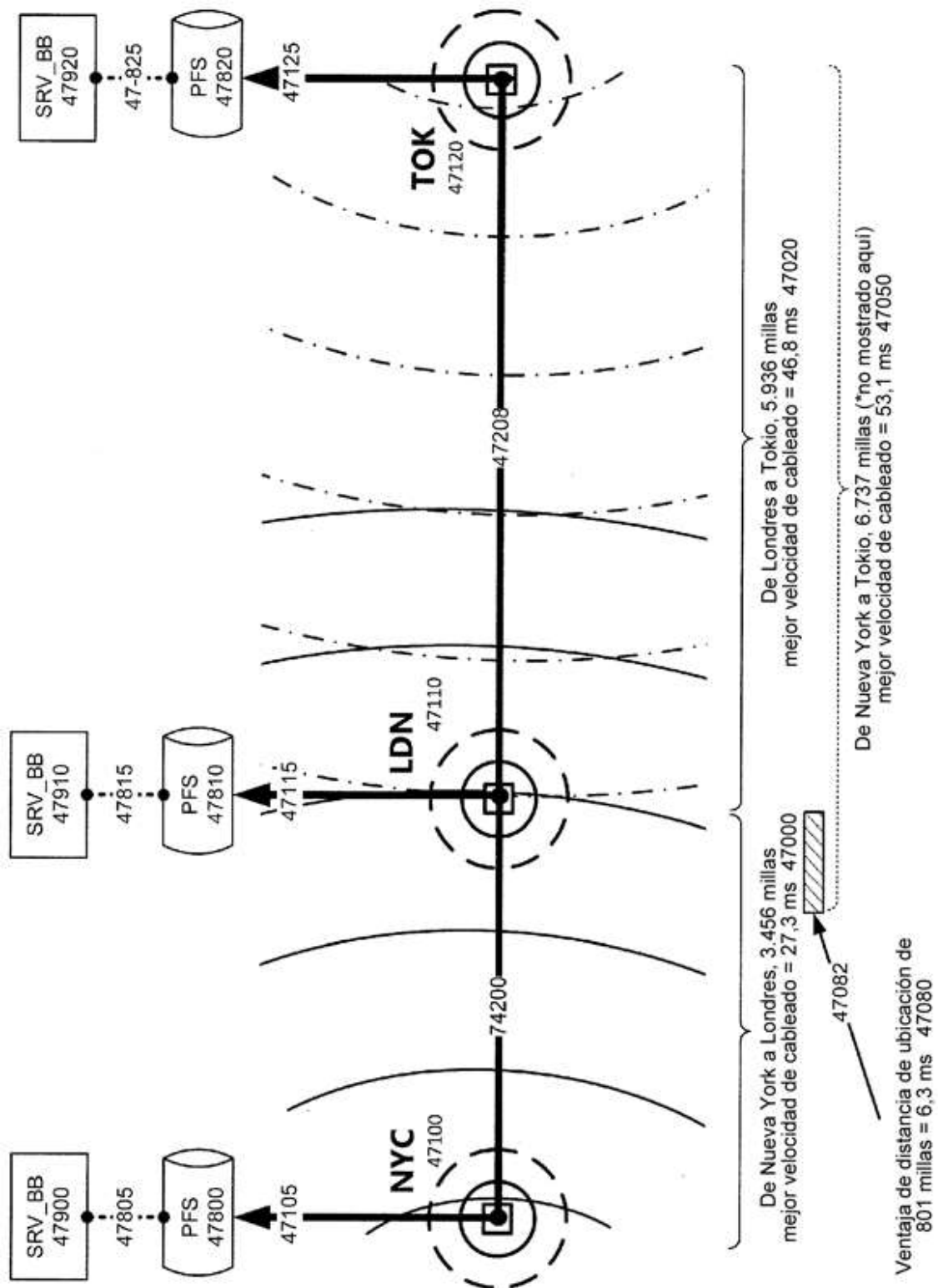


FIG. 47

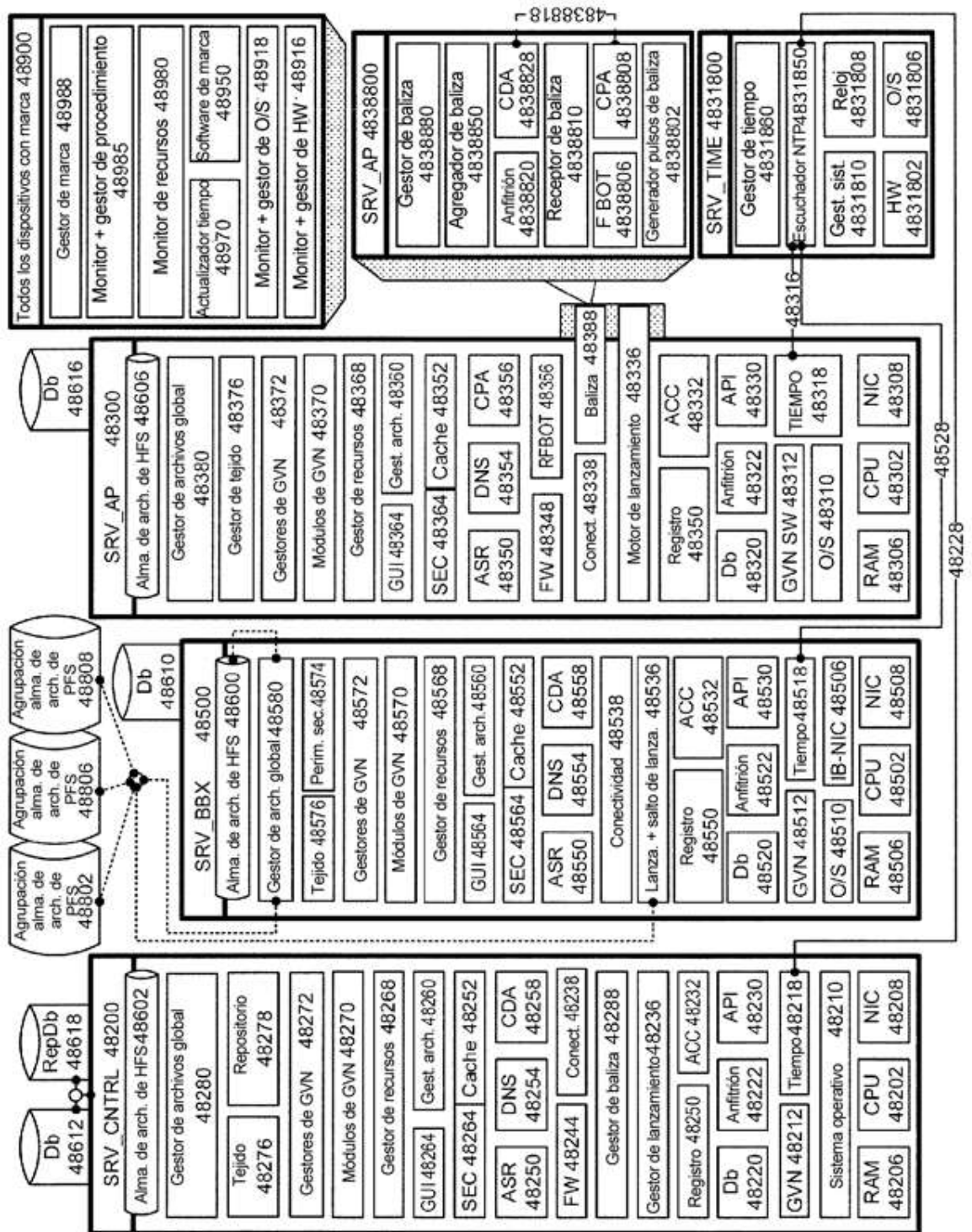


FIG. 48