



19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA

11 Número de publicación: **2 359 399**

51 Int. Cl.:  
**G06F 9/455** (2006.01)  
**G06F 11/14** (2006.01)  
**G06F 17/30** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Número de solicitud europea: **02804063 .2**  
96 Fecha de presentación : **25.11.2002**  
97 Número de publicación de la solicitud: **1466246**  
97 Fecha de publicación de la solicitud: **13.10.2004**

54 Título: **Sistema y procedimiento de duplicación selectiva de datos.**

30 Prioridad: **23.11.2001 US 332549 P**

45 Fecha de publicación de la mención BOPI:  
**23.05.2011**

45 Fecha de la publicación del folleto de la patente:  
**23.05.2011**

73 Titular/es: **COMMVAULT SYSTEMS, Inc.**  
**2 Crescent Place**  
**Oceanport, New Jersey 07757-0090, US**

72 Inventor/es: **Kavuri, Srinivas**

74 Agente: **Carpintero López, Mario**

ES 2 359 399 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 procedimiento de duplicación selectiva de datos

**Apunte sobre derechos de autor**

- 5 Una parte de la divulgación del presente documento contiene material que está sometido a la protección de los derechos de autor. El propietario de los derechos de autor no se opone a la reproducción en facsímil por parte de nadie respecto al documento de la patente o a las divulgaciones que la patente contempla, tal y como aparecen en los archivos o registros de la Oficina de Patentes y Marcas, pero en cualquier caso se reserva todos los derechos de autor en toda su extensión.

**Antecedentes de la invención**

- 10 La invención divulgada en la presente memoria se refiere, en general, al almacenamiento de datos en una red informática y, más concretamente, a la copia selectiva de datos en un sistema modular de gestión y almacenamiento de datos.

Los documentos EP-A-0,809,184, WO 01/06368, US-B1-6,266,679, WO 99/12098, y US 2001044910 se refieren a la extracción / recuperación de datos utilizando sistemas de salvaguardia de datos.

- 15 En el software GALAXY de gestión de almacenamiento, fabricado por CommVault Systems, Inc. de Oceanport, New Jersey, se utilizan políticas de almacenamiento para la dirección del modo en que los datos deben ser almacenados. El almacenamiento presenta al usuario unos colectores lógicos para dirigir operaciones de almacenamiento de datos, como por ejemplo la salvaguardia y extracción. Cada cliente señala una política de almacenamiento que hace posible que el usuario defina cómo, dónde, y la duración respecto de la cual los datos deben ser almacenados en un nivel más alto de abstracción sin que necesite tener un conocimiento o comprensión íntima de la arquitectura y tecnología del almacenamiento subyacente. Los detalles de gestión de las operaciones de almacenamiento son transparentes para el usuario.

- 25 Las políticas de almacenamiento constituyen, por tanto, un concepto lógico asociado con uno o más conjuntos de datos de salvaguardia siendo cada conjunto de datos de salvaguardia una unidad de información autosuficiente. Cada conjunto de datos de salvaguardia puede contener datos de aplicaciones múltiples y procedentes de múltiples clientes. Dentro de cada conjunto de datos de salvaguardia hay uno o más archivos los cuales son bloques o "blobs" discretos de datos en general relativos a una aplicación concreta. Por ejemplo un archivo podría contener ficheros de registro relacionados con un almacén de datos y otro archivo del mismo conjunto de datos de salvaguardia podría contener el almacén de datos mismo.

- 30 Los sistemas de salvaguardia presentan diversos niveles de almacenamiento. Un conjunto de datos de salvaguardia primario, por ejemplo, indica el destino por defecto de operaciones de almacenamiento para un conjunto concreto de datos al que se refiere la política de almacenamiento y está unido a un conjunto práctico de unidades. Estas unidades son consideradas con independencia de la biblioteca o del agente de medios a los cuales están unidades. El conjunto de datos de salvaguardia primario podrá, por ejemplo, contener datos a los que se acceda con frecuencia durante un periodo de una a dos semanas después de que son almacenados. Un administrador del almacenamiento podría encontrar el almacenamiento de dichos datos en un conjunto de unidades con tiempos preferentes de acceso rápido. Por otro lado, dichas unidades rápidas son costosas y una vez que ya no se accede a los datos con frecuencia, el administrador de almacenamiento podría encontrar oportuno desplazar y copiar estos datos hasta un conjunto de datos de salvaguardia secundario o auxiliar en una biblioteca de cintas menos costosa o en otro dispositivo con tiempos de acceso más lentos. Una vez que los datos procedentes del conjunto de datos de salvaguardia primario es desplazado a un conjunto de datos de salvaguardia auxiliar, los datos pueden ser podados a partir del conjunto de datos de salvaguardia primario liberando espacio de la unidad para nuevos datos.

- 45 Aunque los sistemas de almacenamiento de datos existentes proporcionan capacidad para copiar datos a partir del conjunto de datos de salvaguardia primario en el conjunto de datos de salvaguardia auxiliares, con este procedimiento de copia es una operación síncrona, lo que significa que en general todos los datos procedentes del conjunto de datos de salvaguardia primario deben ser copiados en todos los conjuntos de datos de salvaguardia auxiliares. Este proceso se llama también duplicación síncrona de datos y es ineficiente en términos de gestión de datos.

- 50 Un conjunto de datos de salvaguardia contendrá probablemente más de una salvaguardia de datos concreta relacionada con una aplicación determinada además de salvaguardias incrementales o diferenciales tomadas entre las salvaguardias completas. Por ejemplo, un administrador del almacenamiento podría especificar para un conjunto de datos de salvaguardia concreto de una política de almacenamiento que una salvaguardia completa se produce una vez por semana, produciéndose salvaguardias incrementales diariamente. Si el conjunto de datos de salvaguardia fuera retenido durante un periodo de dos semanas antes de ser podado, el conjunto de datos de salvaguardia contendría un primer conjunto de salvaguardia de datos completo, F1, con salvaguardias incrementales I1, I2, I3, I4, I5, I6, y una segunda salvaguardia completa F2. Si cada salvaguardia completa requiriera una cinta y cada salvaguardia incremental requiriera media tinta, entonces se requerirían 5 cintas para almacenar los datos de

este conjunto ejemplar de datos de salvaguardia primario. El conjunto de datos de salvaguardia auxiliar, requeriría, así mismo, 5 cintas cuando los datos fueran transferidos desde el conjunto de datos de salvaguardia primario al auxiliar.

5 De esta manera, aun cuando la duplicación de datos síncronas permite la flexibilidad de promover que cualquier conjunto de datos de salvaguardia auxiliar sea un conjunto de datos de salvaguardia primario, dado que el conjunto de datos de salvaguardia auxiliar es una copia completa del conjunto de datos de salvaguardia primario, el consumo de cintas es muy elevado. Si por alguna razón, los datos no pueden ser copiados en un conjunto de datos de salvaguardia auxiliar, las cintas procedentes del primer conjunto de datos de salvaguardia primario no serán rotadas. De esta manera, los usuarios pueden querer copiar solo salvaguardias completas a medida que su grado de granularidad requerida cambia. Un esquema destacado en el campo que ilustra este principio se llama "Abuelo, Padre, Hijo" ["Grandfather, Father, Son"] (GFS), en el cual cada uno de los tres representa un periodo de tiempo diferente. Por ejemplo, el hijo puede representar un grado semanal de granularidad, el padre puede representar un grado mensual de granularidad, y el abuelo puede representar un grado anual de granularidad.

15 Muchos usuarios no desean copiar todas las salvaguardias procedentes del conjunto de datos de salvaguardia primario en todos los conjuntos de datos de salvaguardia auxiliares. Con el tiempo, el grado de granularidad que los usuarios requiere cambia y, aunque los datos recientes necesitarían ser restaurados a partir de cualquier punto determinado en el tiempo, se requiere, en general, una menor precisión a la hora de restaurar datos más antiguos. Piénsese en un esquema de almacenamiento ejemplar en el que las salvaguardias completas se adopten semanalmente, donde las salvaguardias incrementales se adoptan diariamente, donde los datos son podados después de dos semanas, las salvaguardias completas requieran una cinta y donde las salvaguardias incrementales requieran media cinta. Un administrador del almacenamiento, en este ejemplo, podría requerir que los datos almacenados el mes pasado pudieran ser restaurados en un nivel de granularidad de un día, lo que significaría que los datos podrían ser restaurados a partir de cualquier día determinado del mes pasado. En este grado de granularidad, serían necesarias las salvaguardias incrementales para restaurar los datos. Si el conjunto de datos de salvaguardia contenido en una primera salvaguardia completa de datos, F1, con unas salvaguardias incrementales I1, I2, I3, I4, I5, I6, y una segunda salvaguardia completa, F2, entonces se requerirían F1, I1, I2, I3, I4, I5, I6. Si la salvaguardia incremental I6 se lleva a cabo al mismo tiempo que se lleva a cabo la salvaguardia completa F2, la cinta que contiene F2 sería innecesaria, dado que la salvaguardia completa F2 podría ser reproducida a partir de F1 y de las salvaguardias incrementales I1 - I6. Por otro lado, el administrador de almacenamiento, en este ejemplo, podría requerir solo un grado de granularidad de una semana para datos con un antigüedad de más de un mes, con lo que no se requerirían salvaguardias incrementales y las salvaguardias completas bastarían. En este caso, solo se requerirían las cintas que contienen las salvaguardias completas F1 y F2 y las tres cintas que contienen las salvaguardias incrementales I1, I2, I3, I4, I5, I6 serían innecesarias.

35 Otro ejemplo es una política de almacenamiento con tres conjuntos de salvaguardias llamados Wkly Mnthly, y Yrly con criterios de retención diferentes. El conjunto de datos de salvaguardia Wkly tiene un periodo de retención de 15 días, el conjunto de datos de salvaguardia Mnthly tiene un periodo de retención de 6 meses, y el conjunto de datos de salvaguardia Yrly tiene un periodo de retención de 7 años. Las salvaguardias en este ejemplo se llevan a cabo cada día con una salvaguardia completa cada Viernes sobre el conjunto de datos de salvaguardia Wkly. Así mismo, una salvaguardia completa se lleva a cabo al final de cada mes en el conjunto de datos de salvaguardia Wkly. Solo la salvaguardia completa situada al final de la semana será copiada en el conjunto de datos de salvaguardia Mnthly y solo el final de la salvaguardia completa mensual será copiada en el conjunto de datos de salvaguardia Yrly. Partiendo de la base de que cada salvaguardia completa utiliza una cinta y que las salvaguardias incrementales requieren ¼ de cinta, el conjunto de datos de salvaguardia Wkly ocupa hasta seis cintas con, como máximo, tres salvaguardias completas y 12 salvaguardias incrementales. Estas 6 cintas se reciclan todo el tiempo. El conjunto de datos de salvaguardia Mnthly ocupa 26 cintas que son constantemente recicladas y el conjunto de datos de salvaguardia Yrly ocupa una cinta por mes durante 7 años. De esta manera se requiere un total de 84 cintas y son recicladas durante un largo periodo de tiempo.

50 Así mismo, algunas veces surgen problemas con cintas estropeadas o lagunas en datos debido a problemas de software o de hardware. En estos casos, los datos procedentes del conjunto de datos de salvaguardia primario no pueden ser podados a menos que todos los datos sean copiados en todos los conjuntos de datos de salvaguardia auxiliares, lo que representa un proceso temporal sumamente intenso y, así mismo, requiere un gran número de cintas.

55 Existe, por tanto, la necesidad de un sistema que permita la copia selectiva de datos a partir del conjunto de datos de salvaguardia primario en el conjunto de datos de salvaguardia auxiliares, promueva la eficiente rotación de las cintas, proporcione la capacidad para configurar cualquier variante del esquema GFS, y que permita, así mismo, la poda selectiva de datos a partir del conjunto de salvaguardia del conjunto de datos de salvaguardia primario.

60 El documento EP0809184 divulga un sistema de procesamiento de datos en el cual las copias de datos son almacenadas en baterías de almacenamiento primarias y secundarias. Una base de datos mantiene un repertorio de ficheros de datos y las localizaciones de sus emplazamientos mientras un gestor de los almacenamientos dirige y coordina la transferencia de ficheros entre baterías de almacenamientos. Cabrera et al. en "ADSM: Un Sistema de Almacenamiento Masivo de Salvaguardias y Archivos, de Multiplataforma Escalable" ["ADSM: A Multi Platform

5 Scalable Backup and Archive Mass Storage System”] [Boletín de Documentos de la Conferencia Informática de Sociedades Informáticas (Primavera) CompCon; Tecnologías para la autopista de la información; San Francisco, mar. 5-9 (1995) Los Alamitos; IEEE Comp. Soc. Press, EE.UU., vol. Conf. 40, 5 de marzo de 1995 (03-05-1995), páginas 420 - 427] [Digest of Papers of the Computer Society Computer Conference (Spring) CompCon; Technologies for the Information Superhighway; San Francisco,, Mar. 5-9, 1995, Los Alamitos; IEEE Comp. Soc. Press, US, vol. Conf. 40, 5 March 1995 (1995-03-05), pages 420-427] describen la forma en que los datos son almacenados y salvaguardados utilizando clases de gestión, las cuales definen políticas de administración de datos. Estas clases se fijan por elementos de datos mediante usuarios y administradores de sistemas.

### **Sumario de la invención**

10 Formas de realización de la presente invención dan respuesta, entre otras cosas, a los problemas analizados con anterioridad en relación con el almacenamiento de datos de salvaguardia dentro de una red informática.

15 De acuerdo con un primer aspecto de la presente invención se proporciona un procedimiento computerizado para copiar datos de acuerdo con lo definido en la reivindicación 1. En formas de realización, se proporcionan procedimientos computerizados para copiar datos electrónicos dentro de un primer conjunto de datos de salvaguardia, comprendiendo los procedimientos la identificación, dentro del primer conjunto de datos de salvaguardia, de un elemento de datos que satisface un criterio de selección; y la copia en un segundo conjunto de datos de salvaguardia de al menos una porción del elemento de datos. En algunas formas de realización, el elemento de datos puede comprender una salvaguardia completa dentro de un conjunto de datos de salvaguardia primario de datos de aplicación, una salvaguardia completa dentro de un conjunto de datos de salvaguardia auxiliar de datos de aplicación, un elemento de datos asociado con una ID de datos específicos, u otros elementos de datos.

20 Los criterios de selección es una propiedad o característica del primer elemento de datos utilizado por la invención para seleccionar el primer elemento de datos para copiar y para otras finalidades. En algunas formas de realización, el criterio de selección comprende un criterio temporal y la identificación del elemento de datos comprende la comparación de un tiempo en el que el elemento de datos fue almacenado con el criterio temporal. En algunas formas de realización, el criterio temporal comprende un día de una semana, y la identificación del elemento de datos comprende la comparación de un día de la semana en el que el elemento de datos fue almacenado con el día de la semana. En algunas formas de realización, el criterio temporal comprende un día de un mes y la identificación del elemento de datos comprende la comparación de un día del mes en que el elemento de datos fue almacenado con el día del mes. En algunas formas de realización, el criterio de selección comprende un criterio de ciclos y la identificación del elemento de datos comprende la comparación de una pluralidad de ciclos transcurridos desde que el elemento de datos fue almacenado con el criterio de ciclos. En algunas formas de realización, el criterio de ciclos comprende una pluralidad de salvaguardias completas llevadas a cabo y la identificación del elemento de datos comprende la comparación de una pluralidad de salvaguardias completas llevadas a cabo desde el que el elemento de datos fue almacenado con la pluralidad de salvaguardias completas.

35 En algunas formas de realización, el elemento de datos satisface los criterios de selección que son indicados o de cualquier otra forma marcados o señalizados. En algunas forma de realización, la indicación de que el elemento de datos satisface el criterio de selección comprende la asociación, en una estructura de datos, de la información con el criterio de datos que indica que el elemento de datos satisface el criterio de selección. En algunas formas de realización, que indican que el elemento de datos satisface el criterio de selección comprende la asociación, en una matriz, de la información, con el elemento de datos que indica que el elemento de datos satisface el elemento de selección. En algunas formas de realización, el elemento de datos indicado es desindicado después de que el elemento de datos es copiado en el segundo conjunto de datos de salvaguardia. En algunas formas de realización, el elemento de datos es desindicado mediante la retirada, en una estructura de datos, de la información asociada con el elemento de datos que indica que el elemento de datos satisface el criterio de selección. En algunas formas de realización, el elemento de datos es desindicado mediante la retirada, en una matriz, de la información asociada con el elemento de datos que indica que el elemento de datos satisface el criterio de selección. En algunas formas de realización, el elemento de datos indicado no será podado por un programa de poda a menos que elemento de datos sea previamente desindicado.

50 En algunas formas de realización, en las que el elemento de datos comprende una salvaguardia completa de datos de aplicación.

55 En algunas formas de realización, la copia de al menos una porción del elemento de datos comprende la ejecución de una copia auxiliar de al menos una porción del elemento de datos. En algunas formas de realización, la copia de al menos una parte del elemento de datos comprende a ejecución de una copia auxiliar de al menos una parte del elemento de datos. En algunas formas de realización, la copia de al menos una parte del elemento de datos es una operación reinicialable.

De acuerdo con un segundo aspecto de la presente invención, se proporciona un sistema para copiar datos electrónicos de acuerdo con lo definido en la reivindicación 19.

5 En una forma de realización, un sistema para copiar datos electrónicos comprende un primer conjunto de datos de salvaguardia que contiene uno o más elementos de datos; un segundo conjunto de datos de salvaguardia; y una computadora, conectable al primer conjunto de datos de salvaguardia y al segundo conjunto de datos de salvaguardia; en la que, la computadora está programada para identificar, dentro del primer conjunto de datos de salvaguardia, un primer elemento de datos que satisface unos criterios de selección; y para copiar al menos una parte del primer elemento de datos a partir del primer conjunto de datos de salvaguardia en el segundo conjunto de datos de salvaguardia.

De acuerdo con un tercer aspecto de la presente invención, se proporciona un código de programa de almacenamiento de un medio utilizable por computadora de acuerdo con lo definido en la reivindicación 20.

10 En una forma de realización, un código de programa almacena un medio utilizable por computadora, el cual, cuando se ejecuta en el dispositivo computerizado, provoca que el dispositivo computerizado ejecute un procedimiento computerizado para copiar los datos electrónicos almacenados en un primer conjunto de datos de salvaguardia, comprendiendo el procedimiento la identificación, dentro del primer conjunto de datos de salvaguardia, de un elemento de datos que satisface un criterio de selección; y la copia en un segundo conjunto de datos de salvaguardia de al menos una parte del elemento de datos.

### **Breve descripción de los dibujos**

La invención se ilustra en las figuras de los dibujos que se acompañan, los cuales pretenden ser ejemplares y no limitativos, en los cuales las mismas referencias pretenden referirse a las mismas o correspondientes partes, y en los que:

20 La Fig. 1 es un diagrama de bloques de un sistema de almacenamiento ejemplar de acuerdo con una forma de realización de la presente invención; y

la Fig. 2 es un diagrama de bloques de otro sistema de almacenamiento ejemplar de acuerdo con una forma de realización de la invención; y

25 la Fig. 3 es un diagrama de bloques de unas políticas de almacenamiento ejemplares de acuerdo con una forma de realización de la presente invención;

la Fig. 4 es un diagrama de bloques de una lógica de almacenamiento ejemplar de acuerdo con una forma de realización de la invención;

la Fig. 5 es un diagrama de bloques de unas secuencias de almacenamiento ejemplares de acuerdo con una forma de realización de la invención;

30 la Fig. 6 es un diagrama de bloques de un conjunto ejemplar de datos de salvaguardia primarios de acuerdo con una forma de realización de la invención;

la Fig. 7 es un diagrama de flujo de un proceso de duplicación selectiva de datos de acuerdo con una forma de realización de la invención; y

35 la Fig. 8 es un diagrama de flujo de otro proceso de duplicación selectiva de datos de acuerdo con una forma de realización de la invención.

### **Descripción detallada**

A continuación se describen formas de realización preferentes de la invención con referencia a los dibujos. Tal y como se describe con más detalle más adelante, las preferencias asociadas con políticas específicas de almacenamiento de datos son utilizadas para llevar a cabo una duplicación selectiva de datos. La Fig. 1 presenta un diagrama de bloques de un sistema de almacenamiento ejemplar 100 de acuerdo con una forma de realización de la presente invención. El sistema de almacenamiento 100 incluye unas aplicaciones de software 102 que almacenan datos en diversos medios de almacenamiento 108. Las aplicaciones de software 102 comunican con un control de almacenamiento 104 que determina dónde almacenar los datos de las aplicaciones de software 102 de acuerdo con las políticas de almacenamiento 106 como se describe con mayor detalle en la presente memoria. El control de almacenamiento 104 típicamente recibe los datos que deben ser almacenados a partir de las aplicaciones de software 102, sin embargo, los expertos en la materia advertirán que los datos no se limitan a ser recibidos a partir de las aplicaciones de software y los datos podrían, así mismo, comprender datos en bruto y otros tipos de datos conocidos en la técnica. Tras la recepción de los datos, el control de almacenamiento 104 determina si almacenar los datos en los medios de almacenamiento 108, como por ejemplo en un primer medio de almacenamiento 110, un segundo medio de almacenamiento 112, o cualquier otra serie de medios de almacenamiento hasta un medio de almacenamiento Noveno 114. El control de almacenamiento 104 controla en qué medios de almacenamiento 108 los datos serán almacenados, así como hacia dónde serán desplazados los datos que han sido almacenados con anterioridad.

Por ejemplo, el control de almacenamiento 104 puede desplazar los datos que han sido almacenados en el primer medio de almacenamiento 110, hasta el segundo medio de almacenamiento 112 en base a determinadas políticas de almacenamiento 106. Así mismo, el control de almacenamiento 104 podría desplazar datos del segundo medio de almacenamiento 112 hasta cualquier otro de los medios de almacenamiento 108, incluyendo el medio de almacenamiento Noveno 114. Así mismo, el control de almacenamiento 104 podría desplazar datos desde el primer medio de almacenamiento 110 hasta el medio de almacenamiento Noveno 114. Por supuesto, los datos podrían ser desplazados en cualquier dirección, y el control de almacenamiento 104 es capaz de desplazar datos genéricamente entre los medios de almacenamiento 108.

La Fig. 2 es un diagrama de bloques de otro sistema de almacenamiento ejemplar 116 de acuerdo con una forma de realización de la invención. El sistema de almacenamiento 116 incluye un dispositivo informático 118 para llevar a cabo las operaciones del sistema de almacenamiento 116. El dispositivo informático 118 incluye una aplicación de software 120 que interactúa con un sistema de ficheros instalados 122 cuando los datos son transferidos desde la aplicación de software 120 hasta los medios de almacenamiento 108. Por ejemplo, cuando un usuario ahorra datos desde la aplicación de software 120, el usuario solo es consciente de que los datos están siendo ahorrados en el sistema de ficheros instalado 122. El sistema de ficheros instalado 122 interactúa con un módulo gestor 124, el cual, a su vez, interactúa con un módulo de medios 128 para almacenar datos en los medios de almacenamiento 108, sin embargo, este proceso es en general transparente para el usuario y se lleva a cabo de manera automática tal y como se describe con mayor detenimiento en la presente memoria.

El módulo gestor 124 es un control de almacenamiento ejemplar 104 e incluye unas políticas de almacenamiento 106 las cuales son utilizadas para determinar la forma en que el módulo gestor de datos 124 recibe desde el sistema de ficheros instalado 122 serán almacenados en los medios de almacenamiento. El módulo gestor 124 incluye, así mismo, una estructura de datos denominada mapa maestro 126 para ayudar a las decisiones de almacenamiento iniciales en los medios de almacenamiento. En algunas formas de realización, cuando la aplicación de software tiene por objeto los datos almacenados, los datos son enviados al sistema de ficheros instalado y, a continuación, el módulo gestor 124 accede a las políticas de almacenamiento 106 para determinar el emplazamiento apropiado del almacenamiento de los datos. El mapa maestro 126 incluye información adicional para dirigir los datos que deben ser enviados a otros servicios, por ejemplo a otros dispositivos informáticos para su procesamiento adicional o a los diversos dispositivos de almacenamiento 108.

Un módulo de medios 128 es un módulo de hardware o software que incluye un índice de datos 130 que proporciona detalles adicionales acerca de dónde deben ser almacenados los datos en el sistema de almacenamiento 116. El índice de datos 130 incluye detalles tales como la localización de los medios de almacenamiento 108, como por ejemplo el medio de disco magnético 132 y el medio de cinta magnética 134. El índice de datos 130 es actualizado con la información de localización de ficheros cuando cualquier dato es desplazado desde un medio de almacenamiento 108 a otro, por ejemplo desde el medio de disco magnético 132 hasta la unidad de cinta magnética 134. Por supuesto, podrían ser incorporados en el sistema de almacenamiento 116 tipos adicionales y más de un tipo de medios de almacenamiento 108.

La Fig. 3 es un diagrama de bloques de unas políticas 106 de almacenamiento de datos específicos ejemplares, descritas con mayor detalle en el documento US 2 001 034 812. Las políticas de almacenamiento 106 incluyen unas secuencias de almacenamiento 136 y una lógica de almacenamiento 138. Las secuencias de almacenamiento 136 representan patrones de almacenamiento potenciales los cuales pueden ser utilizados para almacenar tipos concretos de datos. La lógica de almacenamiento 138 incluye una selección de secuencias de almacenamiento inicial 140, la cual representa las configuraciones de almacenamiento potenciales de los datos cuando llegan al control de almacenamiento 104. En línea de puntos se representa la selección de secuencias de almacenamiento 142 y la adaptación de secuencias de almacenamiento 144. La reelección de secuencias de almacenamiento 142 representa la lógica que permitiría que fuera utilizada una lógica alternativa de selección de secuencias en lugar de la selección de secuencias de almacenamiento inicial 140. Aunque no se requiere, se produciría una selección de secuencias de almacenamiento 142, por ejemplo, cuando la selección de secuencias de almacenamiento inicial 140 fuera inadecuada para los datos que están siendo transmitidos desde la aplicación de software 102, como por ejemplo cuando el conjunto de datos de salvaguardia primarios no está disponible o presenta un espacio de almacenamiento insuficiente para los datos. La adaptación de secuencias de almacenamiento 144 representa la lógica que permite que la selección de secuencias de almacenamiento inicial 140 sea modificada de acuerdo con los cambios introducidos en los datos que están siendo ahorrados en los mecanismos de almacenamiento, por ejemplo, si se producen cambios en el tipo de ficheros u otros cambios. Como se describe con mayor detenimiento en la presente memoria, las preferencias e informaciones de datos específicos pueden ser asociadas con las secuencias de almacenamiento 136, con la lógica de almacenamiento 138, con las selecciones de secuencias de almacenamiento iniciales 140, con las reelecciones de almacenamiento 142 y con las adaptaciones de secuencias de almacenamiento 144 para procesar diferentes tipos de datos asociados con las políticas de almacenamiento 106. Por ejemplo, los datos de Microsoft Exchange pueden, de esta manera, ser distinguidos de la base de datos Oracle y procesados de manera diferente.

La Fig. 4 es un diagrama de bloques de una lógica de almacenamiento ejemplar 138 de acuerdo con una forma de realización de la presente invención. Tal y como se muestra, la lógica de almacenamiento representa una estructura de datos que muestra unas filas y una columnas para indicar las IDs de datos específicos 146 que están asociadas

con las preferencias de secuencias de almacenaje 148. Las IDs de datos específicos 146 están asociados con tipos concretos de datos, como por ejemplo aplicaciones, tipos de ficheros, y otros tipos de datos conocidos en la técnica. En algunas formas de realización, las IDs de datos específicos 146 son utilizados como IDs de aplicaciones para distinguir diferentes tipos de datos de aplicación. La columna izquierda representa una pluralidad de IDs de datos específicos ejemplares 146 que se corresponde con programas de aplicaciones, como por ejemplo Microsoft Exchange, la base de datos Oracle de Oracle Corporation, Microsoft SQL Server, y otras aplicaciones conocidas en la técnica. La columna de la derecha asocia una preferencia de secuencias de almacenamiento correspondientes 148 con cada ID de datos específicos 146. Por ejemplo, las primera y segunda IDs de datos específicos 146 están asociadas con una primera secuencia de almacenamiento. De esta manera, el modulo gestor 124 y otros módulos que consultan la lógica de almacenamiento 138 son capaces de determinar que los datos de aplicación generados por las aplicaciones correspondientes a las primera y segunda IDs de datos específicos 146 deben ser almacenados de acuerdo con los patrones de almacenamiento especificados en la secuencia de almacenamiento primaria tal y como se describe con mayor detalle en la presente memoria. De modo similar, los datos de aplicación generados por la aplicación correspondiente a la tercera ID de datos específicos 146 deben ser almacenados de acuerdo con los patrones de datos especificados en la secuencia de almacenamiento secundaria. Otras IDs de datos específicos 146 hacen referencia a patrones de almacenamiento especificados en una secuencia de almacenamiento duplicada y a una secuencia de almacenamiento a largo plazo.

La Fig. 5 es un diagrama de bloques de unas secuencias de almacenamiento ejemplares 136 de acuerdo con una forma de realización de la invención. La secuencia de almacenamiento 136 incluye una secuencia de almacenamiento primaria 150, una secuencia de almacenamiento secundaria 152, una secuencia de almacenamiento duplicada 154 y una secuencia de almacenamiento a largo plazo 156. Los expertos en la materia advertirán que estas son únicamente secuencias de almacenamiento ejemplares y que podrían ser añadidas a las secuencias de almacenamiento 136 otras secuencias de almacenamiento. Las secuencias de almacenamiento 136 están dispuestas en filas con columnas para ilustrar diferentes características de secuencias de almacenamiento.

Por ejemplo, la columna izquierda última ilustra el nombre de la secuencia de almacenamiento seguido por una primera ID de almacenamiento en la segunda columna. Cuando las aplicaciones 102 dirigen el almacenamiento de datos, el control de almacenamiento 104 consulta las políticas de almacenamiento 106 para determinar una secuencia de almacenamiento apropiada 136 para su uso para el almacenamiento de los datos de acuerdo con la ID de datos específicos 146 de la aplicación que genera los datos. Una política de almacenamiento por defecto puede, por ejemplo, incluir el almacenamiento de los datos de acuerdo con la secuencia de almacenamiento primaria 150 donde los datos entraría en el primer almacenamiento con una ID de 001. Los datos serían almacenados en la primera ID de almacenamiento durante 13 semanas tal y como se indica en la siguiente columna de las secuencias de almacenamiento 136. En este punto los datos serían desplazados hasta una segunda ID de almacenamiento 005 donde serían almacenados durante un periodo de 26 semanas. Después de 26 semanas los datos serían desplazados hasta una tercera ID de almacenamiento 002 durante un periodo de 52 semanas. Este proceso continuaría hasta que los datos fueran almacenados en una ID de almacenamiento Noveno 004. La secuencia de almacenamiento 150, por supuesto, es una secuencia de almacenamiento ejemplar, pero se explica aquí como medio para comprender el funcionamiento de la secuencias de almacenamiento 136. Tal y como puede apreciarse en la Fig. 5, las demás secuencias de almacenamiento funcionan de manera similar.

La Fig. 6 muestra un ejemplo de un conjunto de datos de salvaguardia primario 158 de a cuerdo con una forma de realización de la presente invención. Tal y como se muestra, el conjunto de datos de salvaguardia primario 158 es un bloque de memoria u otro espacio de almacenamiento e incluye unas salvaguardias completas A<sub>1</sub> 160 A<sub>2</sub> 168, B<sub>1</sub> 172, B<sub>2</sub> 178 y A<sub>3</sub> 180, y unas salvaguardias incrementales A'<sub>1</sub> 162, A'<sub>2</sub> 164, A'<sub>3</sub> 166, A'<sub>4</sub> 170, B'<sub>1</sub> 174 y A'<sub>5</sub> 176. Los expertos en la materia advertirán que los conjuntos de datos de salvaguardia, como por ejemplo el conjunto de datos de salvaguardia primario 158 no se limitan al almacenamiento solo de este tipo de datos de salvaguardia y pueden almacenar muchos tipos diferentes de datos incluyendo salvaguardias diferenciales, salvaguardias sintéticas, instantáneas, volúmenes de recuperación rápida y otros tipos de datos conocidos en la técnica.

De acuerdo con lo descrito con anterioridad, los datos de aplicación son almacenados según las secuencias de almacenamiento 136 y asociados con las IDs de datos específicos 146 especificadas en la lógica de almacenamiento 138. De esta manera, por ejemplo, la secuencia de almacenamiento 136 podría requerir que se llevara a cabo una salvaguardia completa semanalmente con unas salvaguardias incrementales efectuadas a diario. Así mismo, las IDs de datos específicos múltiples 146 podrían estar asociadas con la misma secuencia de almacenamiento y, de esta manera, el conjunto de datos de salvaguardia podría contener numerosas salvaguardias completas y salvaguardias incrementales a partir de una o más aplicaciones diferentes. Por ejemplo, el conjunto de datos de salvaguardia primario 158 contiene cuatro salvaguardias completas A<sub>1</sub> 160, A<sub>2</sub> 168, B<sub>1</sub> 172, B<sub>2</sub> 178 y A<sub>3</sub> 180 y 6 salvaguardias incrementales A'<sub>1</sub> 162, A'<sub>2</sub> 164, A'<sub>3</sub> 166, A'<sub>4</sub> 170, B'<sub>1</sub> 174 y A'<sub>5</sub> 176.

Cada salvaguardia completa, cada salvaguardia incremental, y otros bloques de datos almacenados en un conjunto de datos de salvaguardia tiene una pluralidad de características asociadas con ellas, como por ejemplo una ID de datos específicos 146 para la aplicación que genera los datos, una fecha en la que los datos fueron almacenados, la cantidad de datos almacenados, y otras características conocidas en la materia útiles en la identificación de datos. Entre otras cosas, estas características pueden ser utilizadas, tal y como se describe con mayor detenimiento en la presente memoria, para identificar bloques de datos individuales discretos dentro de un conjunto de datos de

salvaguardia y llevar a cabo una duplicación de datos selectiva mediante la copia de datos individuales a partir del conjunto de datos de salvaguardia en otro conjunto de datos de salvaguardia utilizando la copia auxiliar y otros procedimientos conocidos en la técnica.

5 La Fig. 7 es un diagrama de flujo que representa un proceso de duplicación de datos selectiva de acuerdo con una forma de realización de la invención. El módulo gestor 124 u otro programa controlador busca en primer término los datos asociados con un primer conjunto de datos de salvaguardia 182. Los datos contenidos en el primer conjunto de datos de salvaguardia pueden ser buscados directamente o la búsqueda puede llevarse a cabo sobre unos datos indizados y sobre otros datos asociados con los datos del primer conjunto de datos de salvaguardia. El módulo gestor 124 identifica unos elementos de datos que satisfacen los criterios de selección especificados por un usuario o de acuerdo con las preferencias asociadas con un segundo conjunto de datos de salvaguardia 184. Por ejemplo, las preferencias asociadas con un segundo conjunto de datos de salvaguardia en los cuales los elementos de datos del primer conjunto de datos de salvaguardia serán duplicados de manera selectiva, podrían ser utilizadas para identificar los elementos de datos del primer conjunto de datos de salvaguardia. Aquellos elementos de datos que son identificados como que satisfacen los criterios de selección son duplicados de manera selectiva en el segundo conjunto de datos de salvaguardia, etapa 186.

La Fig. 8 es un diagrama de flujo que representa otro proceso de duplicación de datos selectiva de acuerdo con una forma de realización de la invención. El módulo gestor 124 u otro programa controlador, consulta en primer término el mapa maestro 126 y determina las IDs de datos específicos 146 señalando la política de almacenamiento 106 para la cual debe llevarse a cabo la duplicación de datos selectiva, etapa 188. En algunas formas de realización, las IDs de datos específicos pueden ser almacenadas en el índice de datos 130 del módulo de medios 128 o en las políticas de almacenamiento 106. Por ejemplo, si un usuario desea llevar a cabo una duplicación selectiva de datos para un conjunto de datos de salvaguardia primario, el módulo gestor 124 llevaría a cabo una consulta en el mapa maestro 126 para determinar qué IDs de datos específicos 146 están asociadas con la política de almacenamiento 106 para el conjunto de datos de salvaguardia primario.

25 Empezando con la primera ID de datos específicos 146, etapa 190, el módulo gestor 124 pasa revista a los datos almacenados en cada conjunto de datos de salvaguardia para determinar si algo debe ser copiado con destino a esa ID de datos específicos 146 en base a los criterios de selección del conjunto de datos de salvaguardia auxiliar en el cual los datos deben ser copiados, etapa 192. Por ejemplo, un conjunto de datos de salvaguardia auxiliar podría especificar que los datos con destino a una aplicación concreta, como por ejemplo los datos de Microsoft Exchange, van a ser copiados de manera selectiva a partir del conjunto de datos de salvaguardia primario de acuerdo con determinados criterios de selección.

Cada conjunto de datos de salvaguardia auxiliar presenta unos criterios de selección utilizados para decidir cuándo copiar cuál salvaguardia completa en él. Los criterios de selección utilizados en la duplicación selectiva de datos pueden ser definidos o bien por tiempo o bien por ciclos. Los criterios de tiempo, por ejemplo, pueden ser especificados como un día determinado del mes y cada n meses y el mes inicial o, así mismo, un día determinado de la semana indicada n semanas y un día inicial de la semana. El día del mes del ejemplo anterior podría adoptar la forma del último día del mes. En el caso de los criterios por ciclos, un ciclo representa los datos almacenados entre salvaguardias completas. Por ejemplo, un ciclo podría incluir una primera salvaguardia completa F1, unas salvaguardias incrementales I1, I2, I3, I4, I5, I6 y una segunda salvaguardia completa F2. Por ejemplo, en algunas formas de realización, un gestor de aplicaciones mantiene el seguimiento del número de ciclos para las salvaguardias completas sobre una base ID de datos específicos 146. Por ejemplo, esto permite que las salvaguardias sean podadas con menor retención en el conjunto de datos de salvaguardia primario el cual no deja ningún rastro para determinar el número de ciclos entre las salvaguardias completas existente y la salvaguardia completa copiada en el conjunto de datos de salvaguardia auxiliar. Los expertos en la materia advertirán que pueden ser utilizados otros muchos criterios para llevar a cabo la duplicación selectiva de datos.

El módulo gestor 124 a continuación marca o de cualquier otra forma señala en el mapa maestro 126 aquellos trabajos que satisfacen los criterios de selección como trabajos que deben ser duplicados, etapa 194. El marcado de estos trabajos, en cuanto tal, asegura que no serán podados antes de que la duplicación pueda ser completada. A menudo, los conjuntos de datos de salvaguardia son podados para promover en general un uso de las cintas y un almacenamiento de datos más eficiente. Por ejemplo, un administrador de almacenamiento o un programa de podado podría podar todas las salvaguardias de un conjunto de datos de salvaguardia más antiguo que una fecha determinada o de acuerdo con otros criterios de selección de poda útiles conocidos en la técnica. Cuando un programa de poda busca datos para podar en un conjunto de datos de salvaguardia, primero comprueba para ver si los elementos de datos que satisfacen los criterios de selección de poda están marcados para que sean duplicados de manera selectiva. Si un elemento de datos es marcado de esta forma, entonces un programa de poda no copiará el elemento de datos hasta que el elemento de datos haya sido duplicado de manera selectiva y no marcado de acuerdo con lo descrito con mayor detalle en la presente memoria.

A diferencia de la duplicación de datos síncrona en la que los datos son duplicados fichero de archivo por fichero de archivo y, de esta manera, todas las salvaguardias, las salvaguardias incrementales, las salvaguardias diferenciales y otras salvaguardias son copiadas en el conjunto de datos de salvaguardia auxiliar, con una duplicación selectiva de datos, el módulo gestor 124 inicia la operación de copia sobre una base de trabajo por trabajo en todos los



conjuntos de salvaguardia necesarios, etapa 196, y copia solo las salvaguardias completas que satisfacen los criterios de selección. En algunas formas de realización, esta operación de copia auxiliar es reinicializable dado que, en otro caso, el usuario puede no saber reiniciar la operación sobre la política de almacenamiento 106, y puede todavía por ejemplo, tener pérdida de datos o cintas que no han sido liberadas debido a los fallos de las copias auxiliares.

5 Una vez que la duplicación selectiva de datos de un trabajo concreto ha sido completada, el módulo gestor 124 diseña ese trabajo indicando que los datos han sido copiados de manera satisfactoria, etapa 198, y que el trabajo puede ahora ser podado o de cualquier otra forma manipulado. En algunas formas de realización, el módulo gestor 124 compensa el mismo trabajo que está siendo duplicado en múltiples conjuntos de datos de salvaguardia reflejando este estado utilizando una estructura de datos matriz u otra técnica apropiada para efectuar el  
10 seguimiento de múltiples elementos y operaciones con el fin de que los trabajos no puedan ser podados antes de que se haya completado la duplicación en todos los conjuntos de datos de salvaguardia. El módulo gestor 124 comprueba si hay IDs de datos específicos 146 restantes, etapa 200, y o bien retorna a la etapa 192 o sale de la subrutina etapa 202, si no hay.

15 Un ejemplo basado en el tiempo que ilustra el proceso descrito en la Fig. 7, es el siguiente. En base a los criterios de selección utilizados en la etapa 192, el módulo gestor 124 determina si es el momento de localizar una salvaguardia completa en el pasado reciente para que sea copiada. Si se encuentra una salvaguardia que vaya a ser guardada, el módulo gestor 124 marca el trabajo que va a ser duplicado y el trabajo no puede ser podado. El módulo gestor 124 añade la ID del trabajo a la vista de trabajos que necesitan ser duplicados en el conjunto de datos de salvaguardia completo. Esta puede ser una matriz dado que el mismo trabajo puede tener que ser copiado en múltiples conjuntos  
20 de datos de salvaguardia.

Por ejemplo, si se considera que un criterio de selección debe ser un día determinado del mes cada mes n, el módulo gestor 124 copiará la salvaguardia completa satisfactoria más reciente desde que la copia selectiva por tiempos fue configurada, que tuvo lugar en el primer conjunto de datos de salvaguardia primario. El módulo gestor 124 copia la primera salvaguardia completa en todos los conjuntos de datos de salvaguardia con independencia del  
25 criterio. Después de que se ha copiado la primera salvaguardia, el módulo gestor 124 tratará de encontrar la salvaguardia completa con éxito más reciente como la del día del mes a partir del tiempo presente buscando hacia atrás en el tiempo y copiará esa salvaguardia completa. En algunas formas de realización, si la salvaguardia completa encontrada es la misma que la salvaguardia completa que ya fue copiada, el módulo gestor 124 emitirá una señal de episodio crítico y una alarma.

30 Un ejemplo basado en ciclos que ilustra el proceso descrito en la Fig. 7 es el que sigue. En base a los criterios de selección utilizados en la etapa 192, el módulo gestor 124 determina el número de ciclos de salvaguardia completa previos para la ID de datos específicos 146 en cuestión. Si la diferencia entre el ciclo duplicado y la ID de datos específicos 146 en cuestión es igual al número de ciclos del criterio, el módulo gestor 124 marca el trabajo que va a ser duplicado y el trabajo no puede ser podado. De acuerdo con lo descrito con anterioridad, este sistema puede  
35 emplear una matriz u otra técnica, dado que el mismo trabajo puede tener que ser copiado en múltiples conjuntos de datos de salvaguardia.

Por ejemplo, si el criterio de selección se da en ciclos en forma de 4 ciclos, el módulo gestor 124 copiará la primera salvaguardia completa que se encuentra en el conjunto de datos de salvaguardia primario. El módulo gestor 124 registra el número de ciclos para esta salvaguardia completa. El módulo gestor 124 tratará entonces de encontrar la  
40 cuarta salvaguardia completa con éxito yendo atrás en el tiempo. Si existe una, entonces esa salvaguardia completa es también copiada.

Los sistemas y módulos descritos en la presente memoria pueden comprender software, firmware, hardware, o cualquier combinación de software, firmware o hardware apropiada para los fines descritos en la presente memoria. El software u otros módulos pueden residir en servidores, estaciones de trabajo, computadoras personales, tabletas  
45 computerizadas PDAs y otros dispositivos apropiados para los fines descritos en la presente memoria. El software y otros módulos pueden ser accesibles por medio de una memoria local, por medio de una red, por medio de un buscador o de otra aplicación en un contexto ASP, o a través de otros medios apropiados para los fines descritos en la presente memoria. Las estructuras de datos descritas en la presente memoria pueden comprender ficheros informáticos, variables, matrices de programación, estructuras de programación, o cualquier esquema o  
50 procedimiento de almacenamiento de información electrónico, o cualquier combinación de éstos, apropiados para los fines descritos en la presente memoria. Los elementos de interfaz descritos en la presente memoria pueden comprender elementos a partir de interfaces de usuario gráficas, interfaces de línea de comando, y otras interfaces apropiadas para los fines descritos en la presente memoria. Las capturas de pantalla presentadas y descritas en la presente memoria pueden ser representadas de manera diferente tal y como es conocido en la técnica para entrar,  
55 acceder, cambiar, manipular, modificar, alterar y trabajar con información.

Aunque la invención ha sido descrita e ilustrada en conexión con formas de realización preferentes, pueden llevarse a cabo muchas variantes y modificaciones como resultará evidente para los expertos en la materia, y la invención, por tanto, no queda limitada a los precisos detalles de metodología y estructura definidos con anterioridad, en cuanto tales variantes y modificaciones pretenden quedar incluidas dentro del alcance de la invención.

## REIVINDICACIONES

5 1. Un procedimiento computerizado para copiar datos electrónicos almacenados en un primer conjunto de datos de salvaguardia (158), comprendiendo el procedimiento: la identificación, dentro del primer conjunto de datos de salvaguardia, de un elemento de datos que satisfaga un elemento de selección y la copia en un segundo conjunto de datos de salvaguardia de al menos una parte del elemento de datos

**caracterizado porque:**

10 - el criterio de selección se basa, al menos en parte, en una ID de datos específicos (146), y en el que la ID de datos específicos está asociada con una aplicación de software para el elemento de datos o está asociada con un tipo de ficheros para el elemento de datos,

y en el que la copia (196) se basa en uno de al menos dos diferentes preferencias de secuencias de almacenamiento (148), y en el que la una preferencia de secuencias de almacenamiento se basa en la ID de datos específicos.

15 2. El procedimiento de la reivindicación 1, en el que la identificación del elemento de datos comprende la identificación de una salvaguardia completa dentro de un conjunto de datos de salvaguardia primario de datos de aplicación.

3. El procedimiento de la reivindicación 1, en el que la identificación del elemento de datos comprende la identificación de una salvaguardia completa dentro de un conjunto de datos de salvaguardia auxiliar de datos de aplicación.

20 4. El procedimiento de la reivindicación 1, en el que el criterio de selección comprende un criterio de tiempo y en el que la identificación del elemento de datos que satisface el criterio de selección comprende la comparación de un tiempo en el que el elemento de datos fue almacenado con arreglo al criterio de tiempo.

25 5. El procedimiento de la reivindicación 4, en el que el criterio de tiempo comprende un día de una semana y en el que la identificación del elemento de datos que satisface el criterio de selección comprende la comparación de un día de una semana en el que el elemento de datos queda almacenado en el día de la semana.

6. El procedimiento de la reivindicación 4, en el que el criterio de tiempo comprende el día de un mes en el que la identificación del elemento de datos que satisface el criterio de selección comprende la comparación de un día del mes en el que el elemento de datos fue almacenado con el día del mes.

30 7. El procedimiento de la reivindicación 1, en el que el criterio de selección comprende un criterio de ciclos y en el que la identificación del elemento de datos que satisface el criterio de selección comprende la comparación de un número de ciclos que se producen desde que el elemento de datos fue almacenado con el criterio de ciclos.

35 8. El procedimiento de la reivindicación 7, en el que el criterio de ciclos comprende un número de salvaguardias completas llevadas a cabo y en el que la identificación del elemento de datos que satisface el criterio de selección comprende la comparación de una pluralidad de salvaguardias completas llevadas a cabo desde que el elemento de datos fue almacenado con la pluralidad de salvaguardias completas.

9. El procedimiento de la reivindicación 1, que comprende la indicación de que el elemento de datos satisface el criterio de selección.

40 10. El procedimiento de la reivindicación 9, en el que la indicación de que el elemento de datos satisface el criterio de selección comprende la asociación, dentro de una estructura de datos, de la información con el elemento de datos que indica que el elemento de datos satisface el elemento de selección.

11. El procedimiento de la reivindicación 9, en el que la indicación de que el elemento de datos satisface el criterio de selección comprende la asociación, dentro de una matriz, de la información con el elemento de datos que indica que el elemento de datos satisface el criterio de selección.

45 12. El procedimiento de la reivindicación 9, en el que el elemento de datos indicado es desindicado después de que el elemento de datos es copiado en el segundo conjunto de salvaguardia.

13. El procedimiento de la reivindicación 12, en el que la desindicación del elemento de datos comprende la retirada, dentro de una estructura de datos, de la desinformación asociada con el elemento de datos que indica que el elemento de datos satisface el criterio de selección.

50 14. El procedimiento de la reivindicación 12, en el que la desindicación de que el elemento de datos comprende la retirada, dentro de una matriz, de la relación asociada con el elemento de datos que indica que el elemento de datos satisface el elemento de selección.

15. El procedimiento de la reivindicación 12, en el que el elemento de datos indicado no será podado por un programa de poda a menos que el elemento de datos sea primeramente desindicado.

16. El procedimiento de la reivindicación 1, en el que el elemento de datos comprende una salvaguardia completa de datos de aplicación.

5 17. El procedimiento de la reivindicación 1, en el que la copia de al menos una parte del elemento de datos comprende la realización de una copia auxiliar de al menos una parte del elemento de datos.

18. El procedimiento de la reivindicación 1, en el que la copia de al menos una parte del elemento de datos es una operación reinicializable.

10 19. Un sistema para copiar datos electrónicos, comprendiendo el procedimiento: un primer conjunto de datos de salvaguardia (158) que contiene uno o más elementos de datos; un segundo conjunto de datos de salvaguardia y una computadora, que puede ser conectada al primer conjunto de datos de salvaguardia y al segundo conjunto de datos de salvaguardia, en el que la computadora está programada para identificar, dentro del primer conjunto de datos de salvaguardia, un primer elemento de datos que satisfaga un criterio de selección y para copiar (196) al menos una parte del primer elemento de datos a partir del primer conjunto de datos de salvaguardia en el segundo conjunto de datos de salvaguardia,  
15

**caracterizado porque:**

- el criterio selección se basa, al menos en parte, en una ID de datos específicos (146), y en el que la ID de datos específicos está asociada con una aplicación de software para el elemento de datos o está asociada con un tipo de fichero para el elemento de datos;

20 y en el que la copia se basa en una de al menos dos diferentes preferencias de secuencias de almacenamiento (148), y en el que una preferencia de secuencias de almacenamiento se basa en una ID de datos específicos.

20. Un medio utilizable por computadora que almacena un código de programa el cual, cuando es ejecutado en un dispositivo computerizado, provoca que el dispositivo computerizado ejecute un procedimiento computerizado para copiar datos electrónicos almacenados en un primer conjunto de datos de salvaguardia, comprendiendo el procedimiento: la identificación, dentro del primer conjunto de datos de salvaguardia (158), de un elemento de datos que satisfaga un criterio de selección y la copia (196) en un segundo conjunto de datos de salvaguardia de al menos una parte del elemento de datos  
25

**caracterizado porque:**

30 - el criterio de selección se basa, al menos en parte, en una ID de datos específicos (146), y en el que la ID de datos específicos está asociada con una aplicación de software para el elemento de datos o está asociada con un tipo de fichero para el elemento de datos;

y en el que la copia se basa en una de al menos dos preferencias de secuencias de almacenamiento diferentes (148) y en el que una preferencia de secuencias de almacenamiento se basa en una ID de datos específicos.

35

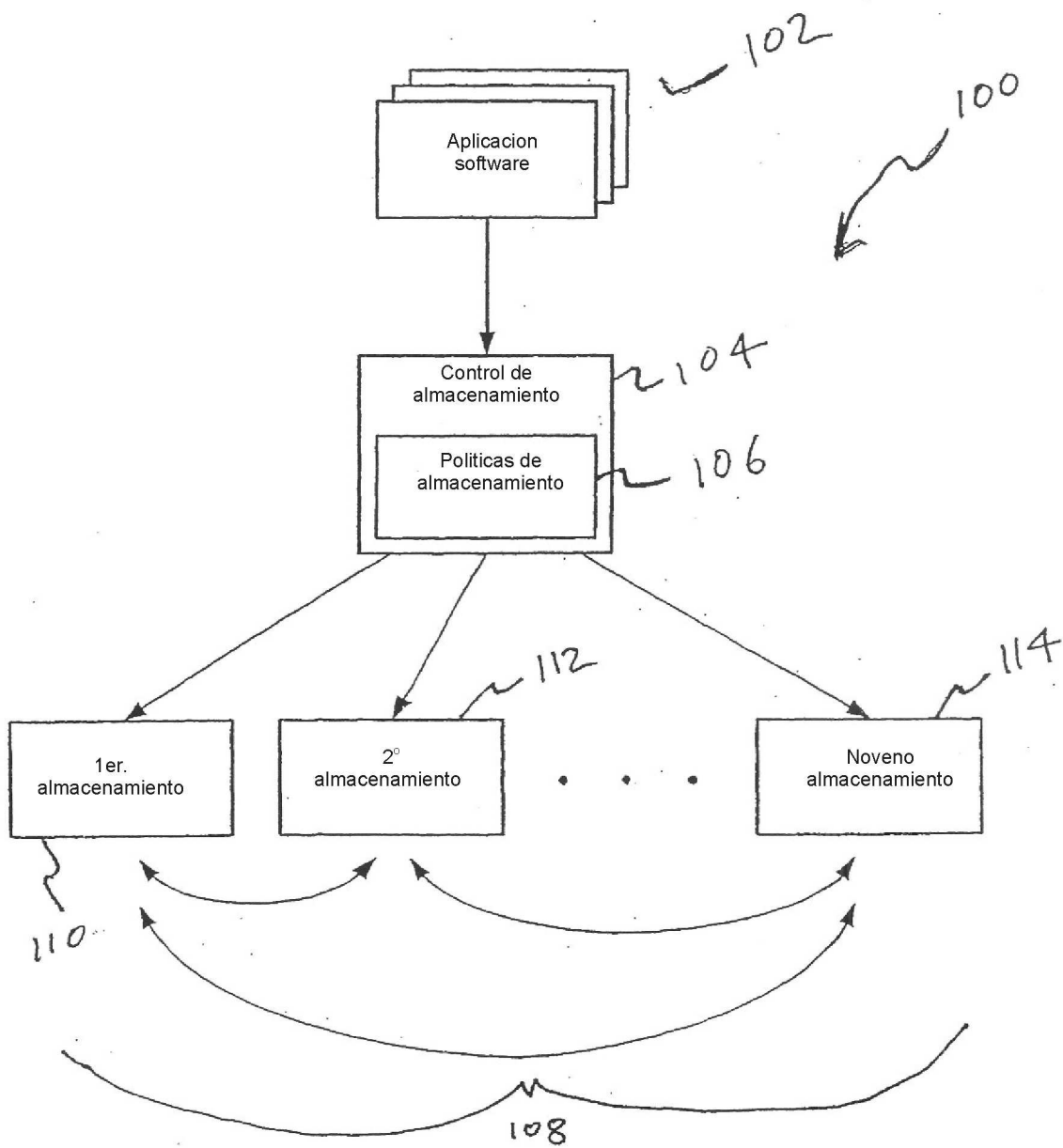


Fig. 1

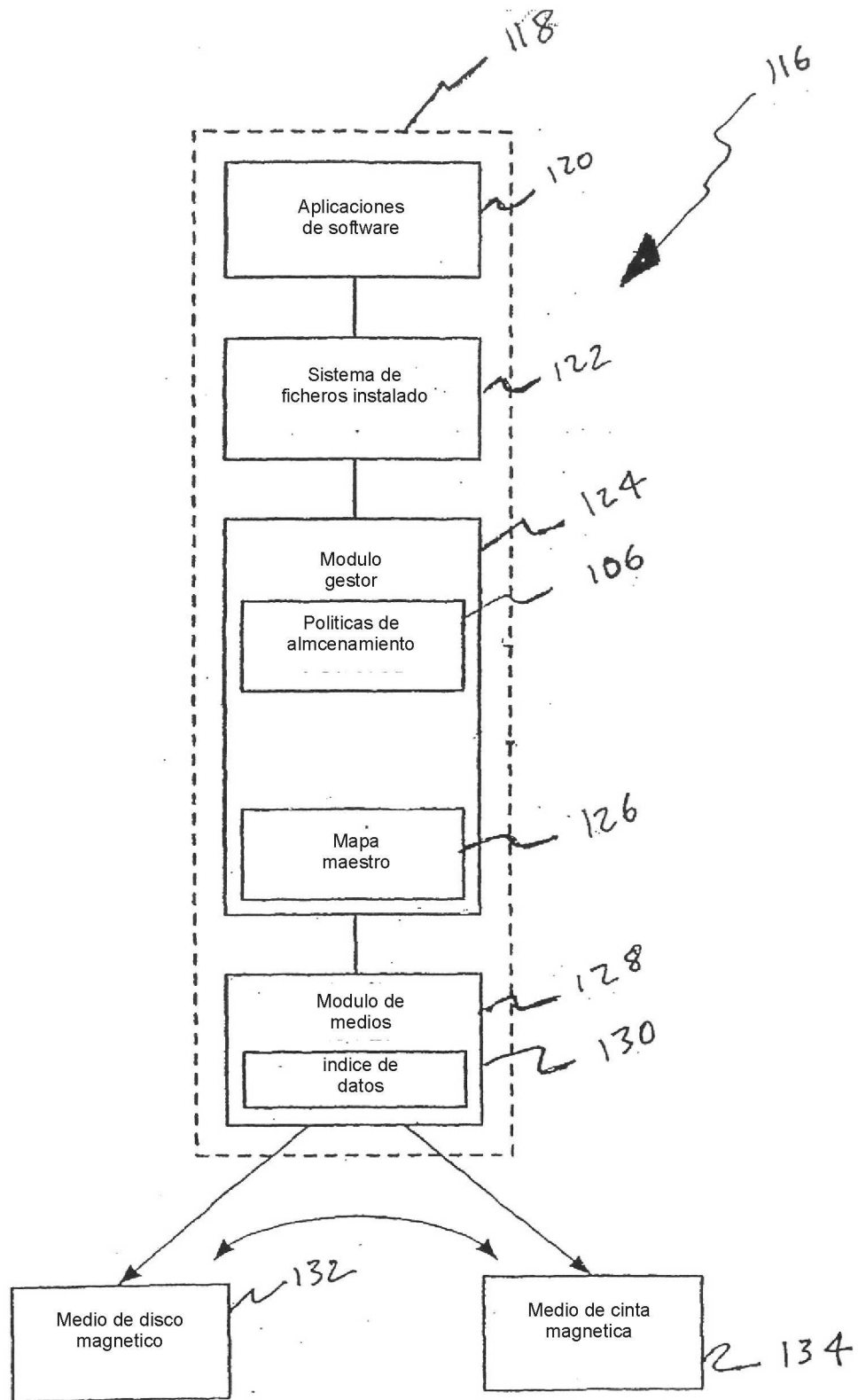


Fig. 2

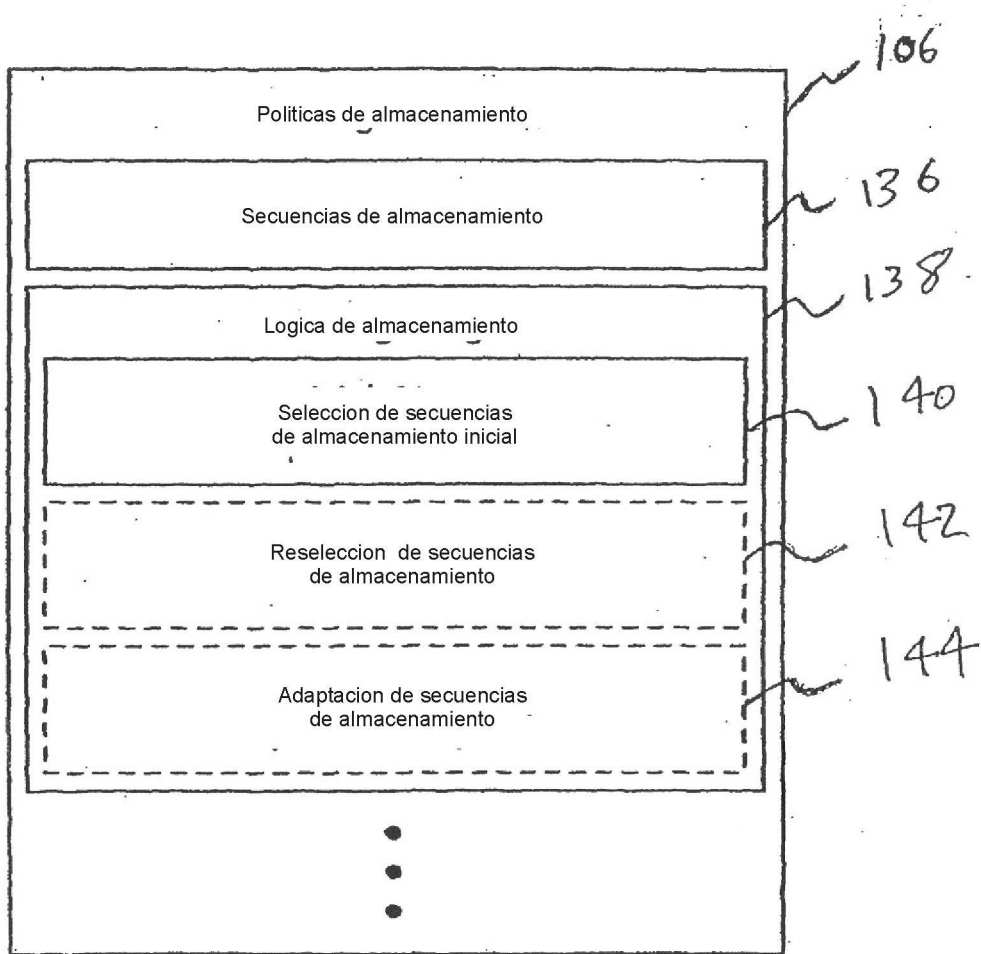
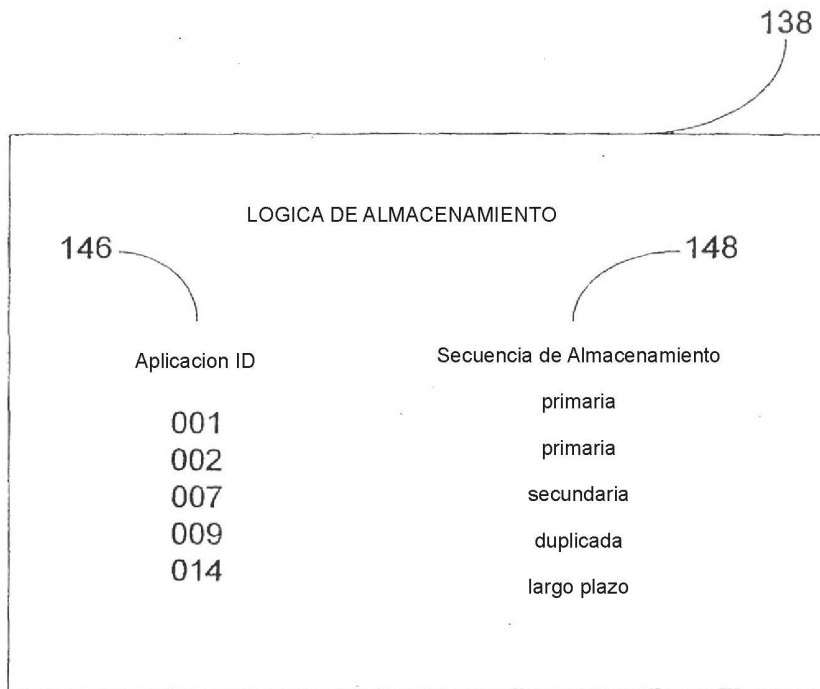


Fig. 3



**Fig. 4**

Secuencias de Almacenamiento

| Nombre de almacenamiento | 1er almacenamiento |                    | 2º almacenamiento |                    | 3er almacenamiento |                    | Noveno almacenamiento ID |
|--------------------------|--------------------|--------------------|-------------------|--------------------|--------------------|--------------------|--------------------------|
|                          | ID                 | Duracion (semanas) | ID                | Duracion (semanas) | ID                 | Duracion (semanas) |                          |
| primaria                 | 001                | 13                 | 005               | 26                 | 002                | 52                 | 004                      |
| secundaria               | 005                | 13                 | 003               | 13                 | 002                | 65                 | 004                      |
| duplicada                | 001                | 13                 |                   |                    |                    |                    |                          |
|                          | 005                | 39                 | 002               | 26                 |                    |                    | 004                      |
|                          |                    |                    |                   |                    |                    |                    |                          |
|                          |                    |                    |                   |                    |                    |                    |                          |
|                          |                    |                    |                   |                    |                    |                    |                          |
| largo plazo              | 005                | 2                  | 002               | 12                 |                    |                    | 004                      |

136

150

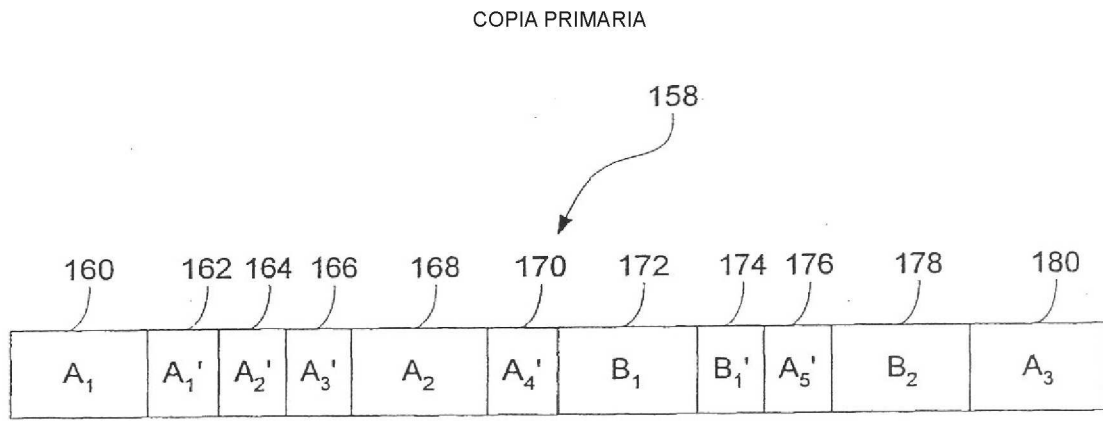
152

154

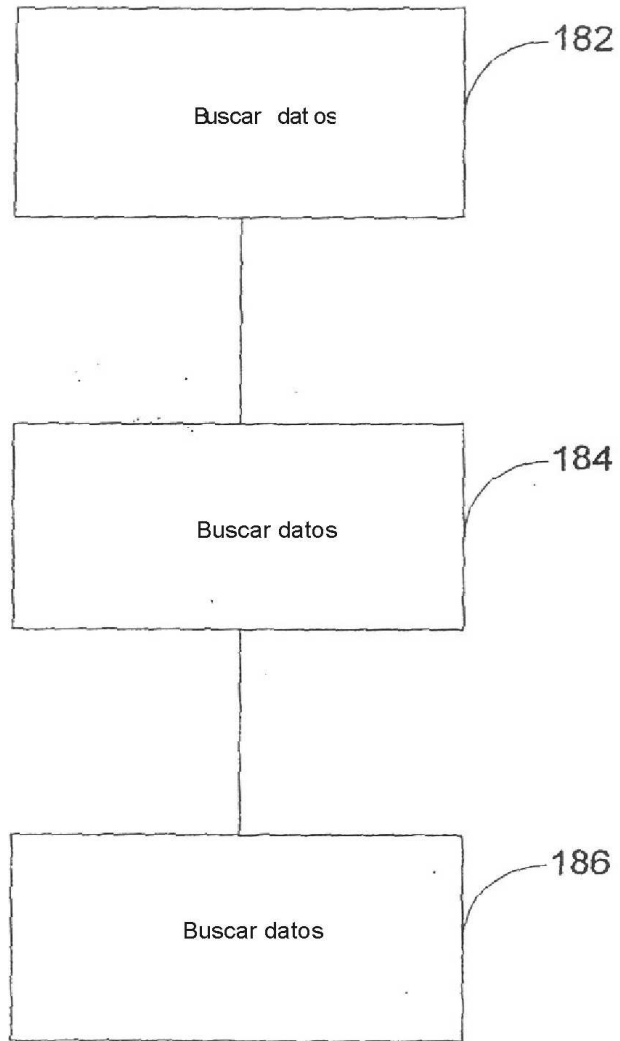
156

Fig. 5





**Fig. 6**



**Fig. 7**

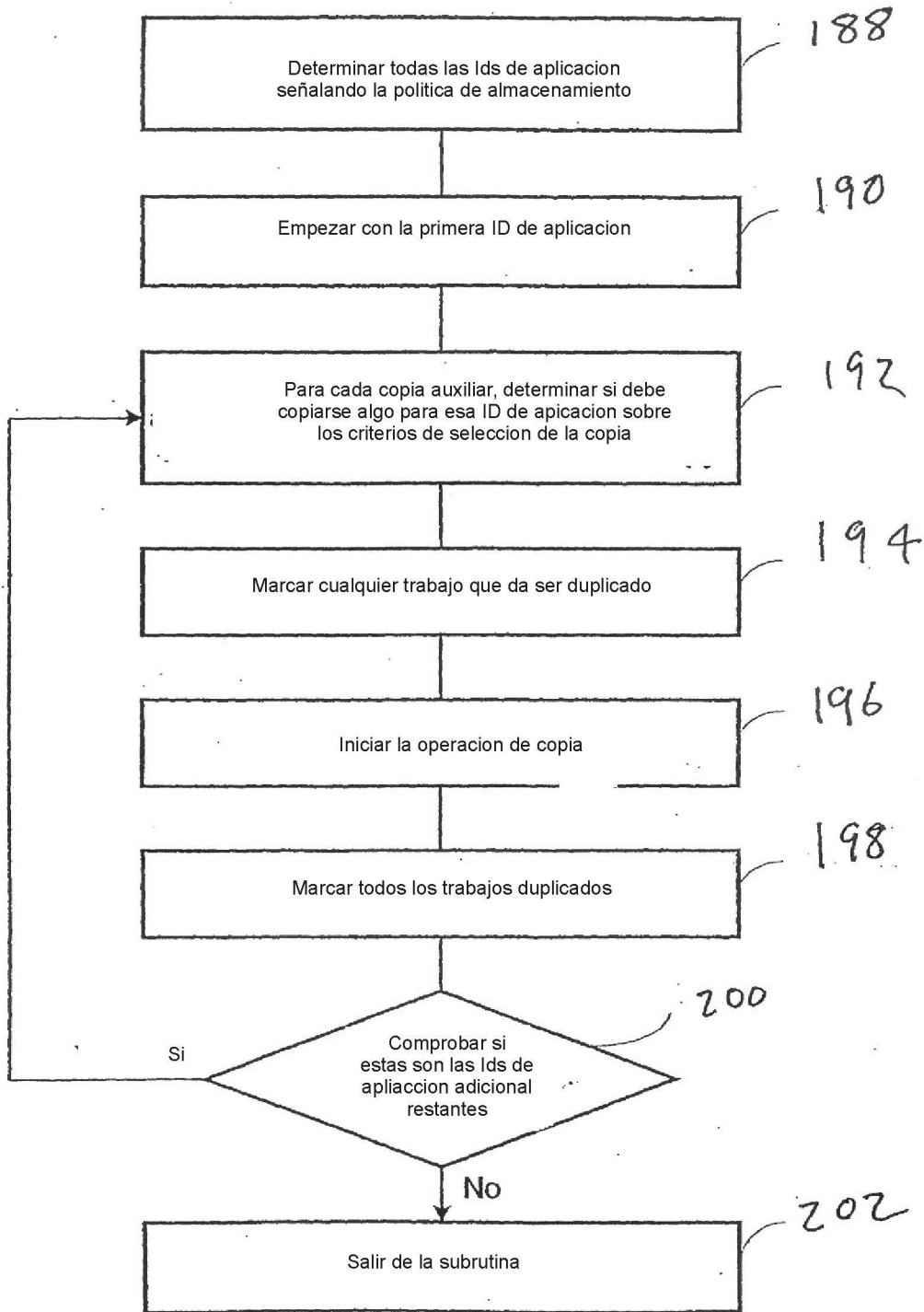


FIG. 8