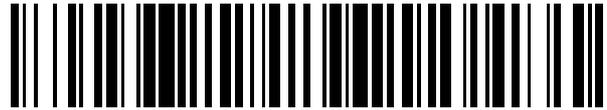


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 715 506**

51 Int. Cl.:

G06F 11/14 (2006.01)

G06F 11/10 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **29.12.2015 PCT/US2015/067761**

87 Fecha y número de publicación internacional: **21.07.2016 WO16114921**

96 Fecha de presentación y número de la solicitud europea: **29.12.2015 E 15831112 (6)**

97 Fecha y número de publicación de la concesión europea: **12.12.2018 EP 3245589**

54 Título: **Seguimiento de cambios usando redundancia en tiempo lógico**

30 Prioridad:

12.01.2015 US 201514595047

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

04.06.2019

73 Titular/es:

**MICROSOFT TECHNOLOGY LICENSING, LLC
(100.0%)
One Microsoft Way
Redmond, Washington 98052-6399 , US**

72 Inventor/es:

**STARKS, JOHN y
BEZBARUAH, ANGSHUMAN**

74 Agente/Representante:

ELZABURU, S.L.P

ES 2 715 506 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

Seguimiento de cambios usando redundancia en tiempo lógico

Antecedentes

5 Los sistemas de computación y las redes asociadas han revolucionado la manera en que los seres humanos trabajan, juegan y se comunican, preconizando lo que se conoce en la actualidad como “era de la información”. La gestión de datos es por lo tanto un campo importante en los tiempos modernos. Un aspecto de la gestión de datos es el seguimiento de cambios. Por ejemplo, con frecuencia resulta de ayuda estar capacitado para distinguir qué porciones de datos han cambiado entre dos circunstancias en el tiempo.

10 Como ejemplo, cuando se hace una copia de seguridad de un sistema de almacenaje, se escribe una copia del sistema de almacenaje en un sitio de copia de seguridad. La siguiente vez que se hace copia de seguridad del sistema de almacenaje, en vez de copiar de nuevo el sistema de almacenaje completo, solamente se hace copia de seguridad de un subconjunto del sistema de almacenaje que ha cambiado. Por consiguiente, para realizar esta copia de seguridad incremental, es un requisito previo determinar qué porciones del sistema de almacenaje han cambiado. Además, cuando se recupera un sistema de almacenaje en un tiempo lógico particular (p. ej., durante una recuperación), el seguimiento de los cambios permite que el sistema de recuperación determine qué posiciones de los datos son acordes con el estado del sistema de almacenaje en ese tiempo lógico particular.

15 El documento US 2013/0067179 se refiere a seguimiento de región sucia de medios no volátiles y presenta técnicas para rastrear escrituras en el conjunto de almacenaje dividiendo el conjunto de almacenaje en regiones y almacenando en el medio de almacenaje no volátil descriptores de regiones sucias que comprenden escrituras en curso. El proceso de recuperación post-fallo puede ser limitado entonces a las regiones identificadas como sucias.

20 El documento US 6.981.114 se refiere a reconstrucción de instantáneas en la que una instantánea borrada de un volumen de datos puede ser reconstruida a partir de una instantánea existente y de uno o más registros de modificación.

25 El documento US 2010/0274980 describe técnicas para recuperación de sistemas usando seguimiento de cambios, que comprenden identificar una región en un almacenaje electrónico asociado a una petición de escritura recibida, establecer un indicador de región que identifique la región de almacenaje electrónico como sucia, y establecer indicadores de porción que identifiquen porciones sucias de la región de almacenaje electrónico.

Breve compendio

30 La invención proporciona un método para mantener el seguimiento de cambios entre porciones unitarias de un sistema de almacenaje, junto con un producto de programa informático correspondiente, según se define en las reivindicaciones independientes 1 y 14 anexas, respectivamente.

35 Durante la copia de seguridad incremental, los identificadores de tiempo lógico se usan para determinar las porciones unitarias que han cambiado, y por lo tanto para determinar las porciones unitarias de las que se necesita hacer copia de seguridad. Puesto que los datos de redundancia inconsistentes para una porción de escritura dan como resultado que los identificadores de tiempo lógico para la porción completa de escritura reciban el tiempo lógico actual, esto provoca la apariencia en el sistema de copia de seguridad de que todas las porciones unitarias de esa porción de escritura hayan sido escritas de nuevo. Por consiguiente, el sistema de copia de seguridad hace que se haga copia de seguridad de la porción de escritura completa, incluso aunque una o más de sus porciones unitarias pudieran no haber cambiado. Mientras que esto podría quizás dar como resultado que se transfieran más datos de respaldo que los absolutamente necesarios en el caso raro de que los datos de redundancia pierdan consistencia, esto protege contra la inconsistencia de los datos cuando esos casos ocurren.

Breve descripción de los dibujos

45 Con el fin de describir la manera en que pueden ser obtenidas las ventajas mencionadas con anterioridad y otras ventajas y características, se va a realizar una descripción más particular de varias realizaciones mediante referencia a los dibujos anexos. Hay que entender que esos dibujos representan solamente realizaciones de muestra y que no deben ser considerados, por lo tanto, como limitativas del alcance de la invención, donde las realizaciones van a ser descritas y explicadas con especificidad adicional y con detalle mediante el uso de los dibujos que se acompañan, en los que:

50 La Figura 1 ilustra resumidamente un sistema de computación en el que se pueden emplear algunas realizaciones descritas en la presente memoria;

La Figura 2 ilustra una cronología de operación en conformidad con los principios descritos en la presente memoria, e incluye una fase de aprovisionamiento, una fase operativa (que incluye operaciones de escritura y verificación) y una fase de copia de seguridad;

La Figura 3 ilustra un método para mantener el seguimiento de cambios entre porciones unitarias de un sistema de almacenaje;

La Figura 4 ilustra un ejemplo de sistema de almacenaje que incluye almacenaje de datos y una estructura identificadora de tiempo lógico;

5 Las Figuras 5A a 5G ilustran varias fases de una realización de la estructura identificadora de tiempo lógico de la Figura 4 siguiendo a lo largo de un escenario particular, y en donde la estructura identificadora de tiempo lógico está dispuesta a modo de matriz lineal de identificadores de tiempo lógico;

10 Las Figuras 6A a 6F ilustran varias fases de otra realización de la estructura identificadora de tiempo lógico de la Figura 4 siguiendo a lo largo del escenario particular, y en donde la estructura identificadora de tiempo lógico está dispuesta a modo de estructura de árbol, y

La Figura 7 ilustra resumidamente varios componentes de procesamiento que pueden operar dentro del sistema de almacenaje para aprovisionar, operar y hacer copia de seguridad del sistema de almacenaje.

Descripción detallada

15 Se va a describir algo de discusión introductoria de un sistema de computación con respecto a la Figura 1. A continuación, se presentarán la estructura y la operación de las realizaciones descritas en la presente memoria con respecto a las Figuras posteriores.

20 Los sistemas de computación están ahora adoptando de manera creciente una amplia variedad de formas. Los sistemas de computación pueden ser, por ejemplo, dispositivos portátiles, aparatos, ordenadores portátiles, ordenadores de sobremesa, servidores, sistemas de computación distribuidos, centros de datos, o incluso dispositivos que no hayan sido considerados convencionalmente como un sistema de computación, tal como
25 dispositivos que se pueden llevar puestos (p. ej., gafas). En la presente descripción y en las reivindicaciones, el término "sistema de computación" se define de manera amplia de modo que incluye cualquier dispositivo o sistema (o combinación de los mismos) que incluya al menos un procesador físico y tangible, y una memoria física y tangible capacitada para tener en la misma, instrucciones ejecutables con ordenador que puedan ser ejecutadas por un procesador. La memoria puede adoptar cualquier forma y puede depender de la naturaleza y de la forma del sistema de computación. Un sistema de computación puede estar distribuido a través de un entorno de red y puede incluir múltiples sistemas de computación integrantes.

30 Según se ha ilustrado en la Figura 1, en su configuración más básica, un sistema 100 de computación incluye típicamente al menos una unidad 102 de procesamiento de hardware y memoria 104. La memoria 104 puede ser una memoria física del sistema, la cual puede ser volátil, no volátil o alguna combinación de las dos. El término "memoria" puede ser usado también en la presente memoria para referirse a un almacenaje masivo no volátil tal como medios físicos de almacenaje. Si el sistema de computación está distribuido, el procesamiento, la memoria y/o la capacidad de almacenaje pueden estar distribuidos también. Según se utiliza en la presente memoria, el término "módulo ejecutable" o "componente ejecutable" puede referirse a objetos de software, rutinas o métodos que pueden
35 ser ejecutados en el sistema de computación. Los diferentes componentes, módulos, motores y servicios descritos en la presente memoria pueden ser implementados como objetos o procesos que se ejecuten en el sistema de computación (p. eje., como aspectos independientes).

40 En la descripción que sigue, se describen realizaciones con referencia a acciones que son realizadas por uno o más sistemas de computación. Si tales acciones están implementadas en software, uno o más procesadores (del sistema de computación asociado que lleva a cabo la acción) dirigen la operación del sistema de computación en respuesta a haber ejecutado instrucciones ejecutables con ordenador. Por ejemplo, tales instrucciones ejecutables con ordenador pueden estar materializadas sobre uno o más medios legibles con ordenador que forman un producto de programa informático. Un ejemplo de operación de ese tipo incluye la manipulación de datos. Las instrucciones ejecutables con ordenador (y los datos manipulados) pueden ser almacenados en la memoria 104 del sistema 100 de computación. El sistema 100 de computación puede contener también canales 108 de comunicación que permitan que el sistema 100 de computación comunique con otros sistemas de computación a través de, por ejemplo, la red 110.

45 Las realizaciones descritas en la presente memoria pueden comprender o utilizar un sistema de computación de propósito general o de propósito especial que incluya hardware de ordenador, tal como, por ejemplo, uno o más procesadores y memoria de sistema, según se discute con mayor detalle más adelante. Las realizaciones descritas en la presente memoria incluyen también medios físicos y otros medios legibles con ordenador para llevar a cabo o almacenar instrucciones ejecutables con ordenador y/o estructuras de datos. Tales medios legibles con ordenador pueden ser cualquier medio disponible al que se pueda acceder por medio de un sistema de computación de propósito general o de propósito especial. Los medios legibles con ordenador que almacenan instrucciones ejecutables con ordenador son medios físicos de almacenaje. Los medios legibles con ordenador que portan instrucciones ejecutables con ordenador son medios de transmisión. De ese modo, a título de ejemplo y sin limitación, las realizaciones de la invención pueden comprender al menos dos clases completamente diferentes de medios legibles con ordenador: medios de almacenaje y medios de transmisión.

Los medios de almacenaje legibles con ordenador incluyen RAM, ROM, EEPROM, CD-ROM u otro almacenaje en disco óptico, almacenaje en disco magnético u otros dispositivos de almacenaje magnético, o cualquier otro medio de almacenaje físico y tangible que pueda ser usado para almacenar medios de código de programa informático en forma de instrucciones ejecutables con ordenador o de estructuras de datos, y a los que se pueda acceder por medio de un sistema de computación de propósito general o de propósito especial.

Una "red" se define como uno o más enlaces de datos que permiten el transporte de datos electrónicos entre sistemas de computación y/o módulos y/u otros dispositivos electrónicos. Cuando se transfiere o se proporciona información a través de una red o de otra conexión de comunicaciones (ya sea cableada, inalámbrica o una combinación de cableada o inalámbrica) a un sistema de computación, el sistema de computación ve apropiadamente la conexión como un medio de transmisión. Los medios de transmisión pueden incluir una red y/o enlaces de datos que pueden ser usados para portar medios de código de programa informático deseado en forma de instrucciones ejecutables con ordenador o de estructuras de datos, y a los que se puede acceder por medio de un sistema de computación de propósito general o de propósito especial. Combinaciones de los anteriores pueden estar también incluidas dentro del ámbito de los medios legibles con ordenador.

Además, tras alcanzar varios componentes del sistema de computación, los medios de código de programa informático en forma de instrucciones ejecutables con ordenador o de estructuras de datos, pueden ser transferidos automáticamente desde medios de transmisión a medios de almacenaje (o viceversa). Por ejemplo, las instrucciones ejecutables con ordenador o las estructuras de datos recibidas a través de una red o de un enlace de datos pueden ser almacenadas en memoria RAM dentro de un módulo de interfaz de red (p. ej., un "NIC"), y a continuación transferidas eventualmente a una RAM del sistema de computación y/o a medios de almacenaje menos volátiles en un sistema de computación. De ese modo, se debe entender que los medios de almacenaje pueden estar incluidos en componentes del sistema de computación que utilicen también (o al menos mayormente) medios de transmisión.

Las instrucciones ejecutables con ordenador comprenden, por ejemplo, instrucciones y datos que, cuando se ejecutan en un procesador, provocan que un sistema de computación de propósito general, un sistema de computación de propósito especial, o un dispositivo de procesamiento de propósito especial realicen una función determinada o un grupo de funciones. Las instrucciones ejecutables con ordenador pueden ser, por ejemplo, instrucciones binarias o uniformes que se sometan a alguna traducción (tal como compilación) con anterioridad a su ejecución directa por los procesadores, tal como instrucciones de formato intermedio tal como lenguaje ensamblador, o incluso código fuente. Aunque el presente objeto ha sido descrito en lenguaje específico para características estructurales y/o acciones metodológicas, debe entenderse que el presente objeto que definido en las reivindicaciones anexas no se limita necesariamente a las características descritas o a las acciones descritas con anterioridad. En cambio, las características y acciones descritas se divulgan como ejemplos de formas de implementación en las reivindicaciones.

Los expertos en la materia podrán apreciar que la invención puede ser puesta en práctica en entornos de computación de red con muchos tipos de configuraciones de sistemas de computación, incluyendo ordenadores personales, ordenadores de sobremesa, ordenadores portátiles, procesadores de mensajes, dispositivos portátiles, sistemas de multiprocesador, electrónica de consumo basada en microprocesador o programable, PCs de red, miniordenadores, ordenadores centrales, teléfonos móviles, PDAs, localizadores, enrutadores, conmutadores, centros de datos, dispositivos que se pueden llevar puestos (tal como gafas) y similares. La invención puede ser también puesta en práctica en entornos de sistema distribuido donde sistemas de computación locales y remotos, que están enlazados (ya sea mediante enlaces de datos cableados, enlaces de datos inalámbricos, o mediante una combinación de enlaces de datos cableados e inalámbricos) a través de una red, realizan las tareas. En un entorno de sistema distribuido, módulos de programa informático pueden estar situados en dispositivos de almacenaje en memoria tanto locales como remotos.

La Figura 2 ilustra un cronograma 200 de operación acorde con los principios descritos en la presente memoria. La primera, es una fase 210 de aprovisionamiento en la que un sistema de computación se prepara para la operación de mantener el seguimiento de cambios en un sistema de almacenaje estableciendo una estructura identificadora de tiempo lógico. La segunda, es una fase 220 de operación en la que el sistema de computación participa en operación normal y como tal, realiza operaciones 221 de escritura en el sistema de almacenaje y escrituras equivalentes en la estructura identificadora de tiempo lógico, y a continuación verifica 222 más tarde la consistencia de las operaciones de escritura usando la estructura identificadora de tiempo lógico en conformidad con los principios descritos en la presente memoria. Finalmente, existe una fase 230 de copia de seguridad potencial en la que se usa la estructura identificadora de tiempo lógico para la copia de seguridad del sistema de almacenaje en un estado coherente.

Más específicamente, la Figura 3 ilustra un diagrama de flujo de un método 300 para mantener el seguimiento de cambios entre porciones de un sistema de almacenaje. En la fase 210 de aprovisionamiento, un identificador de tiempo lógico está asociado a cada porción unitaria en una estructura identificadora de tiempo lógico (acción 310). La Figura 4 ilustra un ejemplo de sistema 400 de almacenaje que incluye una estructura identificadora de tiempo lógico. Ahora se va a describir el ejemplo de la Figura 4 antes de volver al diagrama de flujo del método 300 de la Figura 3.

Como ejemplo, la Figura 4 ilustra un ejemplo de sistema 400 de almacenaje que incluye un almacenaje 400A de datos y una estructura 400B identificadora de tiempo lógico. El almacenaje 400A de datos incluye porciones 401A a 409A unitarias. En una realización, el sistema 400A de almacenaje es un sistema de archivo en el que cada porción 401A a 409A unitaria es un archivo. Sin embargo, el sistema 400A de almacenaje podría también ser un sistema basado en bloques en el que cada porción 401A a 409A unitaria es un bloque. Por supuesto, pueden existir muchas más porciones unitarias en un sistema de almacenaje típico, pero el ejemplo de la Figura 4 se mantiene simple por razones de claridad.

El sistema 400A de almacenaje incluye también porciones 411A a 413A de escritura. En conformidad con el aspecto más amplio descrito en la presente memoria, las porciones de escritura pueden incluir una o más porciones unitarias, y representan la manera en la que las porciones unitarias pueden ser escritas en el sistema de almacenaje. Por ejemplo, si el sistema 400A de almacenaje es un sistema de archivo, entonces típicamente, se escribe un archivo cada vez. Por consiguiente, cada una de las porciones de escritura podrían contener solamente una porción unitaria (es decir, un archivo). Sin embargo, si el sistema 400A de almacenaje es un sistema de almacenaje a base de bloques, entonces típicamente los bloques se escriben en una página cada vez, donde una página abarca múltiples bloques. En el ejemplo ilustrado, las porciones 411A a 413A de escritura están representadas, cada una de ellas, de modo que cada una abarca tres porciones unitarias. Esto es, de nuevo, por razones de simplicidad y claridad en cuanto a la descripción del ejemplo. En un sistema de almacenaje típico, pueden existir millones o miles de millones de porciones unitarias y porciones de escritura, y cada porción de escritura puede abarcar potencialmente números diferentes de porciones unitarias.

La estructura 400B identificadora de tiempo lógico incluye identificadores 401B a 409B de tiempo lógico, asociados a cada una de las porciones 401A a 409A unitarias, respectivamente. Las Figuras 5A a 5G ilustran estados progresivos de la estructura 400B identificadora de tiempo lógico de modo que el estado estará en un ejemplo particular que se describe en lo que sigue. En el ejemplo de las Figuras 5A a 5G, la estructura identificadora de tiempo lógico incluye una matriz lineal de entradas, siendo cada entrada un identificador de tiempo lógico para una porción unitaria asociada del sistema de almacenaje. Las Figuras 6A a 6F ilustran estados progresivos de la estructura 400B identificadora de tiempo lógico, en los que se usa una estructura de árbol para un seguimiento de datos de redundancia. Las Figuras 5A a 5D corresponden a las Figuras 6A a 6D, respectivamente. Las Figuras 5E y 5F corresponden a la Figura 6E. Las Figuras 5G y 5H corresponden a las Figuras 6E y 6F, respectivamente. Por consiguiente, los ejemplos de las Figuras 5A a 5G, así como de las Figuras 6A a 6F, van a ser mencionadas de manera frecuente en la descripción del escenario que sigue.

En el ejemplo particular de las Figuras 5A a 5G y de las Figuras 6A a 6F, los identificadores de tiempo lógico son números de secuencia, y el tiempo lógico está representado por un número de secuencia que se incrementa en uno cada vez que el tiempo lógico avanza. Por ejemplo, en el estado 500A de la Figura 5A, y en el estado 600A de la Figura 6A, cada uno de los identificadores 401B a 409B de tiempo lógico se ponen a cero. En algunos casos, los identificadores de tiempo lógico podrían estar asociados a un tiempo real, de tal modo que se pueda introducir un tiempo real, y el identificador de tiempo lógico correspondiente pueda ser recuperado.

Haciendo de nuevo referencia a la Figura 4, el grupo de identificadores de tiempo lógico para cada una de las porciones 411A a 413A de escritura, tiene asociados datos 411B a 413B de redundancia, respectivamente. Los datos 411B a 413B de redundancia pueden ser usados para confirmar la exactitud de los identificadores de tiempo lógico dentro de las porciones 411A a 413A de escritura correspondientes, respectivamente. Por ejemplo, la porción 411A de escritura está asociada a las porciones 401A a 403A unitarias, y las porciones 401A a 403A unitarias tienen asociados identificadores 401B a 403B de tiempo lógico. Por consiguiente, los datos 411B de redundancia representan datos de redundancia (p. ej., una suma de control) que pueden ser usados para verificar la precisión de los identificadores 401B a 403B de tiempo lógico. También, la porción 412A de escritura está asociada a porciones 404A a 406A unitarias, y las porciones 404A a 406A unitarias tienen asociados identificadores 404B a 406B de tiempo lógico. Por consiguiente, los datos 412B de redundancia representan datos de redundancia (p. ej., una suma de control) que pueden ser usados para verificar la precisión de los identificadores 404B a 406B de tiempo lógico. Finalmente, la porción 413A de escritura está asociada a porciones 407A a 409A unitarias, y las porciones 407A a 409A unitarias tienen asociados identificadores 407B a 409B de tiempo lógico. Por consiguiente, los datos 413B de redundancia representan datos de redundancia (p. ej., una suma de control) que pueden ser usados para verificar la precisión de los identificadores 407B a 409B de tiempo lógico.

Con los identificadores 401A a 409A de tiempo lógico puestos a cero en la Figura 5A, los datos 411B, 411B y 411C de redundancia representan, cada uno de ellos, los datos de redundancia para tres ceros. En la nomenclatura de las Figuras 5A a 5G ilustradas, los datos de redundancia para una secuencia de números se simboliza al tener la secuencia de números ilustrada con un subrayado dentado. Por ejemplo, en la Figura 5A, los datos 411B de redundancia se han representado como que tienen un valor 000.

Las Figuras 6A a 6F ilustran varias fases de otra realización de la estructura identificadora de tiempo lógico de la Figura 4 que sigue a lo largo del mismo escenario particular que se describió con respecto a las Figuras 5A a 5G. Sin embargo, en esta segunda realización, la estructura identificadora de tiempo lógico puede adoptar la forma de una estructura de árbol.

En esta estructura de árbol, si una porción de escritura no tiene asociado un grupo de identificadores de tiempo lógico, entonces se supone que los identificadores de tiempo lógico son un valor por defecto representado por un nodo raíz en el árbol. Si cualquiera del grupo de estructuras identificadoras de tiempo lógico para una porción de escritura es diferente del valor por defecto representado en el nodo raíz, entonces ese grupo de estructuras identificadoras de tiempo lógico tendrá su propio nodo hijo. Si el grupo de identificadores de tiempo real son los mismos para una porción de escritura dada, entonces el nodo hijo simplemente mantendrá ese valor. Si cada uno del grupo de identificadores de tiempo lógico no son iguales para la porción de escritura, entonces el nodo hijo contendrá los datos de redundancia para la serie de identificadores de tiempo lógico.

Por ejemplo, con referencia al ejemplo de la Figura 6A, ninguno de los identificadores de tiempo lógico es diferente de un valor por defecto de cero. Por consiguiente, puesto que cada uno de los valores de los identificadores 401B a 409B de tiempo lógico es 0, existe un único nodo 610 raíz que representa el valor 0.

Pasando a la operación 221 de escritura de la fase 220, cuando se escribe en una porción de escritura particular que incluye porción(es) unitaria(s) del sistema de almacenaje, el sistema actualiza los identificadores de tiempo lógico para cualquiera de la(s) porción(es) unitaria(s) de la porción de escritura que haya(n) cambiado (acción 320), y calcula los datos de redundancia (p. ej., una suma de control) asociados al (a los) identificador(es) de tiempo lógico asociado(s) a la(s) porción(es) unitaria(s) de la porción de escritura (acción 321). El sistema escribe la porción de escritura (acción 322) y también escribe los datos de redundancia asociados (acción 323). En algunas realizaciones, los datos de redundancia se escriben (acción 323) en paralelo con la actualización de los identificadores de tiempo lógico (acción 320). Por consiguiente, si se encuentra después que los datos de redundancia son incompatibles con los identificadores de tiempo lógico (acción 324), entonces ha ocurrido probablemente un fallo de potencia, y los identificadores de tiempo lógico pueden ser tratados como dañados. Ésta es la razón de que los identificadores de tiempo lógico sean marcados, en ese caso, con el tiempo lógico actual según se mida mediante la aplicación que usa los datos.

Por ejemplo, con referencia al estado 500B de la Figura 5B, supóngase que se realiza una escritura en porciones 402A y 403A unitarias, y que el tiempo lógico ha avanzado hasta el tiempo 1. Por consiguiente, los identificadores 401B a 403B de tiempo lógico de las porciones 401A a 403A unitarias correspondientes para la porción 411A de escritura cambian de 000 a 011 (acción 320). Los datos 411B de redundancia se recalculan a continuación (acción 321) como 011. El sistema escribe a continuación la nueva porción 411A de escritura en las porciones 401A a 403A unitarias (acción 322) y también escribe los nuevos datos 411B de redundancia que tienen ahora el valor 011 (acción 323). Con referencia al estado 600B de la Figura 6B, dado que la porción 411A de escritura ya no está más representada de manera precisa por el nodo 610 raíz, se crea un nodo 611 hoja, que contiene los datos de redundancia (p. ej., una suma de control) para los identificadores 401B a 403B de tiempo lógico.

A continuación, con referencia al estado 500C de la Figura 5C, supóngase que se realiza una escritura en las porciones 403A y 404A unitarias, y que el tiempo lógico ha avanzado hasta tiempo 2. Por consiguiente, los identificadores 401B a 403B de tiempo lógico de las porciones 401A a 403A unitarias correspondientes para la porción 411A de escritura cambian desde 011 a 012 (acción 320). Además, los identificadores 404B a 406B de tiempo lógico de las porciones 404A a 406A unitarias correspondientes para la porción 412B de escritura cambian desde 000 a 200 (también acción 320). Los datos 411B de redundancia se recalculan a continuación (acción 321) como 012. Adicionalmente, los datos 412B de redundancia se recalculan a continuación (también acción 321) como 200. El sistema escribe entonces la nueva porción 411A de escritura en las porciones 401A a 403A unitarias (acción 322), la nueva porción 412A de escritura de las porciones 404A a 406A unitarias (también acción 322), los nuevos datos 411B de redundancia que tienen ahora el valor 012 (acción 323), y los nuevos datos 412B de redundancia que tienen ahora el valor 200 (también acción 323). Con referencia al estado 600C de la Figura 6C, puesto que la porción 412A de escritura ya no está más representada de manera precisa por el nodo 610 raíz, se crea un nodo 612 hoja, que contiene los datos de redundancia (p. ej., una suma de control) para los identificadores 404B a 406B de tiempo lógico. Además, el nodo 611 hoja ha sido alterado de la misma manera que los datos 411B de redundancia.

A continuación, haciendo referencia al estado 500D de la Figura 5D, supóngase que se realiza una escritura en las porciones 405A y 406A unitarias, y que el tiempo lógico ha avanzado hasta tiempo 3. En consecuencia, los identificadores 404B a 406B de tiempo lógico de las porciones 404A a 406A unitarias correspondientes para la porción 412A de escritura cambian desde 200 a 233 (acción 320). Los datos 412B de redundancia se recalculan a continuación (acción 321) como 233. El sistema escribe a continuación la nueva porción 412A de escritura en las porciones 404A a 406A unitarias (acción 322), y también escribe los nuevos datos 412B de redundancia que ahora tienen el valor 233 (acción 323). De forma similar, con referencia al estado 600D de la Figura 6D, el valor del nodo 612 hoja también cambia correspondientemente.

Haciendo de nuevo referencia a la Figura 2, durante la fase 220 operativa, se puede llevar a cabo ocasionalmente una operación 222 de verificación. Esta operación 222 de verificación puede ser realizada en respuesta a determinados eventos, y/o podría ser llevada a cabo de manera periódica. Como ejemplo, la operación 222 de verificación podría ser llevada a cabo después de que se haya detectado que ha ocurrido un ciclo de potencia en el sistema.

Durante una operación 222 de verificación, se podría verificar tras la operación 221 de escritura, que los datos de redundancia no son coherentes con el grupo de uno o más identificadores de tiempo lógico (acción 324). Si los datos de redundancia no son coherentes, entonces se lleva a cabo una operación 221 de escritura con respecto al tiempo lógico 3 actual para todas las porciones unitarias en la porción de escritura correspondiente.

5 Por ejemplo, con referencia al estado 500E de la Figura 5E, supóngase mientras está aún en el tiempo lógico 3, que los datos 412B de redundancia tienen un valor dañado (representado por “****”) que ya ha dejado de ser consistente con el valor 233 de los identificadores 404B a 406B de tiempo lógico de las porciones 404A a 406A unitarias asociadas a la porción 412A de escritura correspondiente (acción 324). En ese caso, según se ha representado en el estado de la Figura 5F, los identificadores 404B a 406B de tiempo lógico de las porciones 404A a 406A unitarias correspondientes de la porción 412A de escritura son reescritos, cada uno de ellos, con el identificador de tiempo lógico 3 actual (acción 325). Los nuevos datos 412B de redundancia son calculados a continuación como 333 (acción 321), y se escriben los datos de redundancia (acción 323). Con referencia al estado 600E de la Figura 6E, puesto que los identificadores 404B a 406B de tiempo lógico de la porción de escritura correspondiente al nodo 611 hoja son ahora todos iguales (tiempo lógico 3), el nodo 612 hoja puede transformarse opcionalmente de modo que tenga un valor único de 3. Alternativamente, el nodo 612 hoja podría haber tenido el valor de redundancia de 333.

En otras palabras, si los datos de redundancia (p.ej., la suma de control) son incorrectos, la porción de escritura asociada completa se trata como si ésta tuviera el último número de secuencia. En el momento siguiente se actualiza la porción de escritura, el tiempo lógico actual se almacena explícitamente para cada identificador de tiempo lógico para esa porción de escritura. Esto significa que si la página resulta dañada de alguna manera, cuando el usuario requiere el conjunto de cambios desde el tiempo t (que es menor que el tiempo actual), el usuario verá que todo lo representado por esta parte de la estructura identificadora de tiempo lógico, ha cambiado, incluso aunque solamente haya cambiado alguna o incluso ninguna de las porciones unitarias de esta porción de escritura. Esto es mucho más seguro que suponer que nada ha cambiado (puesto que se desconoce lo que ha cambiado).

A continuación, con referencia al estado 500G de la Figura 5G, supóngase que se realiza una escritura en porciones 406A y 407A unitarias, y que el tiempo lógico ha avanzado hasta tiempo 4. En consecuencia, los identificadores 404B a 406B de tiempo lógico de las porciones 404A a 406A unitarias correspondientes para la porción 412A de escritura, cambian de 333 a 334 (acción 320). Además, los identificadores 407B a 409B de tiempo lógico de las porciones 407A a 409A unitarias correspondientes para la porción 413A de escritura, cambian de 000 a 400 (también acción 320). Los datos 412B de redundancia se recalculan a continuación (acción 321) como 334. Adicionalmente, los datos 413B de redundancia son recalculados a continuación (acción 321) como 400. El sistema escribe a continuación la nueva porción 412A de escritura en las porciones 404A a 406A unitarias (acción 322), la nueva porción 413A de escritura de las porciones 407A a 409A unitarias (también acción 322), los nuevos datos 412B de redundancia que ahora tienen un valor 334 (acción 323), y los nuevos datos 413B de redundancia que ahora tienen el valor 400 (también acción 323). Con referencia al estado 600F de la Figura 6F, puesto que la porción 413A de escritura ha dejado ya de ser representada de manera precisa por el nodo 610 raíz, se crea un nodo 613 hoja, el cual contiene los datos de redundancia (p. ej., una suma de control) para los identificadores 407B a 409B de tiempo lógico. Además, puesto que el valor en el nodo 612 hoja (de 3) ya no representa más a todos los identificadores de tiempo lógico de la porción 412A de escritura, el nodo 612 hoja se modificada de modo que incluya los datos 334 de redundancia.

Haciendo de nuevo referencia a la Figura 4, supóngase ahora que se necesita hacer una copia de seguridad del sistema 400 de almacenaje. Por ejemplo, en la Figura 2, el cronograma 200 avanza desde la fase 220 de operación normal a la fase 230 de copia de seguridad. Los identificadores de tiempo lógico son usados por el sistema de copia de seguridad para determinar las porciones unitarias que han cambiado. Por ejemplo, con referencia a las Figuras 5G y 6F, si la última copia de seguridad había sido tomada después del tiempo lógico 0 pero antes que el tiempo lógico 1, entonces no se hace copia de seguridad de las porciones 402A a 407A unitarias debido a que sus respectivos identificadores 402B a 407B de tiempo lógico están, cada uno de ellos, en tiempo lógico 1 o posterior. Si la última copia de seguridad se tomó después del tiempo lógico 1 pero antes del tiempo lógico 2, entonces se hace copia de seguridad de las porciones 403A a 407A puesto que sus respectivos identificadores 403B a 407B de tiempo lógico están, cada uno de ellos, en tiempo lógico 2 o posterior. Si la última copia de seguridad fue tomada después del tiempo lógico 2 pero antes del tiempo lógico 3, entonces se hace copia de seguridad de las porciones 404A a 407A unitarias puesto que sus respectivos identificadores 404B a 407B de tiempo lógico están, cada uno de ellos, en tiempo lógico 3 o posterior. Si la última copia de seguridad se tomó después del tiempo lógico 3 pero antes del tiempo lógico 4, se debe hacer copia de seguridad de las porciones 406A y 407A unitarias puesto que sus respectivos identificadores 406B y 407B de tiempo lógico están, cada uno de ellos, en tiempo lógico 4 o posterior.

Puesto que los datos de redundancia inconsistentes para una porción de escritura dan como resultado los identificadores de tiempo lógico para la totalidad de la porción de escritura que recibe el tiempo lógico actual, esto provoca la apariencia en el sistema de copia de seguridad de que todas las porciones unitarias de esa porción de escritura hayan sido escritas de nuevo. Por consiguiente, el sistema de copia de seguridad provoca que se haga copia de seguridad de la totalidad de la porción de escritura, incluso aunque una o más de las porciones unitarias pudieran no haber cambiado. Mientras que esto podría quizás dar como resultado que fueran transferidos más datos de copia de seguridad de los absolutamente necesarios en el caso raro de que los datos de redundancia pierdan consistencia, esto protege la consistencia de los datos reales cuando ocurren tales casos.

- 5 La Figura 7 ilustra de manera resumida una arquitectura 700 que incluye varios componentes de procesamiento que pueden operar dentro del sistema de almacenaje para aprovisionar, operar y hacer copia de seguridad del sistema de almacenaje. En particular, un componente 710 de aprovisionamiento puede estar estructurado y configurado de modo que realice las operaciones descritas con anterioridad para la fase 210 de aprovisionamiento, un componente 721 de escritura puede estar estructurado y configurado para realizar las operaciones 221 de escritura, un componente 722 de verificación puede estar estructurado y configurado para realizar las operaciones 222 de verificación, y un componente 730 de copia de seguridad puede estar estructurado y configurado para llevar a cabo las operaciones descritas con anterioridad para la fase 230 de copia de seguridad.
- 10 En consecuencia, los principios descritos en la presente memoria proporcionan un mecanismo eficaz para mantener el seguimiento de cambios entre versiones arbitrarias de un sistema de almacenaje. Las realizaciones descritas han de ser consideradas en todos los aspectos solamente como ilustrativas y no restrictivas. El alcance de la invención está indicado, por lo tanto, por las reivindicaciones anexas en vez de por la descripción que antecede.

REIVINDICACIONES

- 1.- Un método para mantener el seguimiento de cambios entre porciones unitarias de un sistema de almacenaje, comprendiendo el método:
- 5 para cada porción unitaria en el sistema de almacenaje, asociar (310) un identificador de tiempo lógico a la porción en una estructura identificadora de tiempo lógico;
- cuando se escriba en una porción de escritura particular que incluye una o más de las porciones unitarias del sistema de almacenaje, llevar a cabo lo siguiente:
- una acción de calcular (321) datos de redundancia asociados a un grupo de uno o más identificadores de tiempo lógico asociados a las una o más porciones de la porción de escritura;
- 10 una acción de escribir (322) en la porción de escritura del sistema de almacenaje, y
- una acción de escribir (323) los datos de redundancia correspondientes a la porción de escritura en la estructura identificadora de tiempo lógico que corresponda a la porción de escritura;
- una acción de verificar (324) a continuación que los datos de redundancia no son consistentes con el grupo de uno o más identificadores de tiempo lógico, y
- 15 si los datos de redundancia no son consistentes, asignar (325) un identificador de tiempo lógico actual a cada uno de los identificadores de tiempo lógico en el grupo de uno o más identificadores de tiempo lógico en la estructura identificadora de tiempo lógico, de tal modo que las porciones de escritura inconsistentes sean tratadas como recientemente escritas.
- 2.- El método según la reivindicación 1, siendo el sistema de almacenaje un disco duro virtual.
- 20 3.- El método según la reivindicación 1, siendo el sistema de almacenaje un sistema de archivo, y siendo las porciones unitarias del sistema de almacenaje archivos.
- 4.- El método según la reivindicación 1, siendo el sistema de almacenaje un sistema de almacenaje a base de bloques, y siendo las porciones unitarias del sistema de almacenaje bloques.
- 5.- El método según la reivindicación 4, siendo la porción de escritura una página de múltiples bloques.
- 25 6.- El método según la reivindicación 1, siendo los datos de redundancia una suma de control.
- 7.- El método según la reivindicación 1, incluyendo la estructura identificadora de tiempo lógico una matriz de entradas, incluyendo una entrada que incluye un identificador de tiempo lógico para cada una de al menos algunas de las porciones unitarias del sistema de almacenaje.
- 8.- El método según la reivindicación 1, incluyendo la estructura identificadora de tiempo lógico una estructura de árbol en la que los nodos hoja son cada uno de los grupos de identificadores de tiempo lógico asociados a porciones de escritura, incluyendo cada uno de al menos algunos de los nodos hoja datos de redundancia correspondientes para el grupo asociado de identificadores de tiempo lógico.
- 30 9.- El método según la reivindicación 8, en donde, si un grupo de identificadores de tiempo lógico asociados a una porción de escritura son los mismos, entonces la porción de escritura no está representada por un nodo hoja, sino que más bien el valor del identificador de tiempo lógico está representado en un nodo padre.
- 35 10.- El método según la reivindicación 1, comprendiendo los identificadores de tiempo lógico números de secuencia.
- 11.- El método según la reivindicación 1, con cada uno de al menos algunos identificadores de tiempo lógico asociado a un tiempo real.
- 12.- El método según la reivindicación 1, en donde la acción de verificar (324) a continuación que los datos de redundancia no son consistentes con el grupo de uno o más identificadores de tiempo lógico ocurre después de que se haya detectado que ha ocurrido un ciclo de potencia, y después de que la porción de escritura haya sido escrita.
- 40 13.- El método según la reivindicación 1, que comprende además:
- una acción de restablecer el sistema de almacenaje a un tiempo lógico particular usando la estructura identificadora de tiempo lógico.
- 45 14.- Un producto de programa informático que comprende uno o más medios de almacenaje legibles con ordenador que incluyen instrucciones ejecutables con ordenador que están estructuradas de tal modo que, cuando se ejecutan mediante uno o más procesadores del sistema de computación, provocan que el sistema de computación lleve a cabo un método para mantener el seguimiento de cambios entre porciones de un sistema de almacenaje,

comprendiendo el método:

para cada porción unitaria en el sistema de almacenaje, asociar (310) un identificador de tiempo lógico a la porción en una estructura identificadora de tiempo lógico;

5 cuando se escriba en una porción de escritura particular que incluye una o más de las porciones unitarias del sistema de almacenaje, llevar a cabo lo siguiente:

una acción de calcular (321) datos de redundancia asociados a un grupo de uno o más identificadores de tiempo lógico asociados a las una o más porciones de la porción de escritura;

una acción de escribir (322) en la porción de escritura del sistema de almacenaje, y

10 una acción de escribir (323) los datos de redundancia correspondientes a la porción de escritura en la estructura identificadora de tiempo lógico que corresponda a la porción de escritura;

una acción de verificar (324) a continuación que los datos de redundancia no son consistentes con el grupo de uno o más identificadores de tiempo lógico, y

15 si los datos de redundancia no son consistentes, asignar (325) un identificador de tiempo lógico actual a cada uno de los identificadores de tiempo lógico en el grupo de uno o más identificadores de tiempo lógico en la estructura identificadora de tiempo lógico, de tal modo que las porciones de escritura inconsistentes sean tratadas como recientemente escritas.

15.- El producto de programa informático según la reivindicación 14, siendo el sistema de almacenaje un sistema de almacenaje a base de bloques, y siendo las porciones unitarias del sistema de almacenaje bloques.

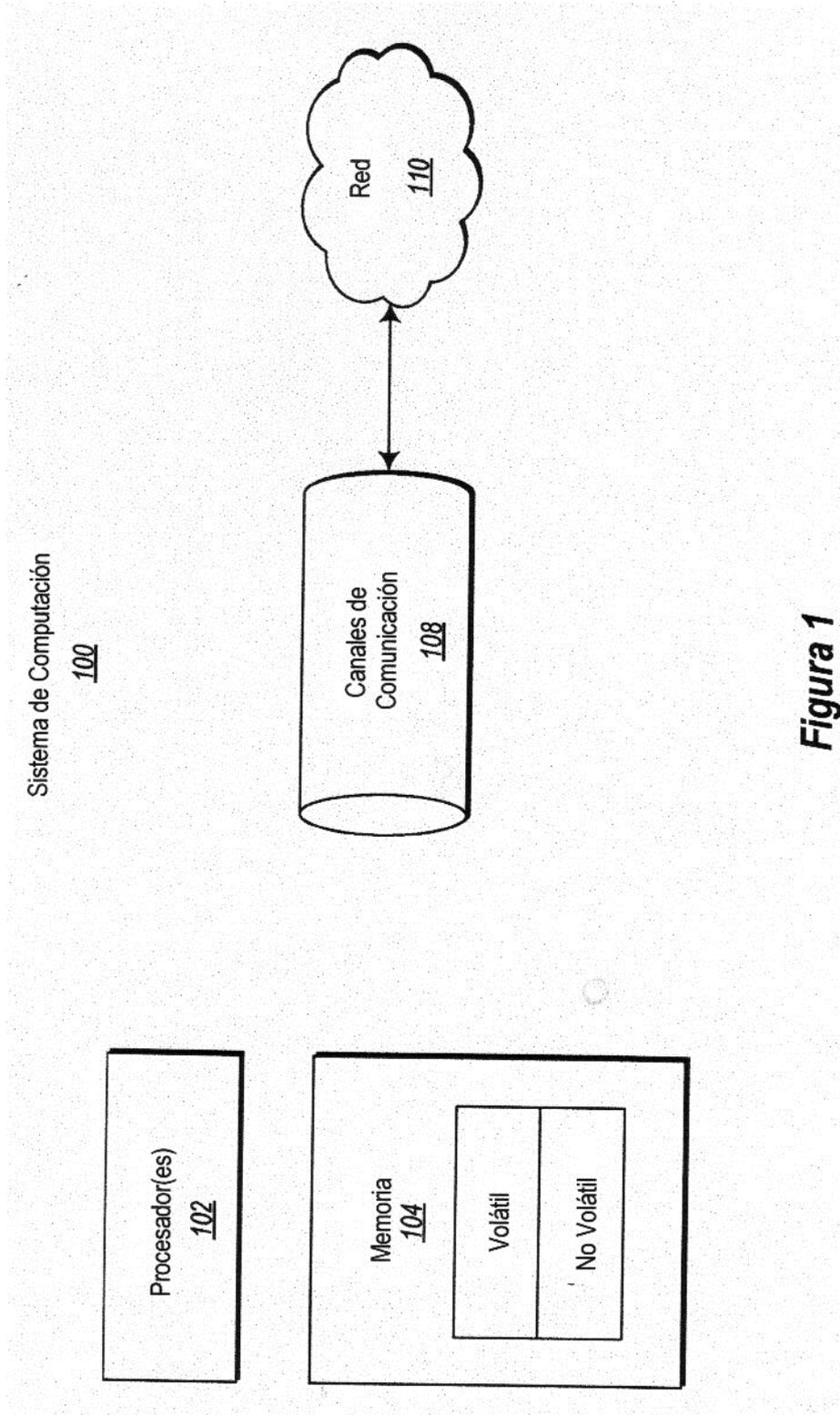


Figura 1

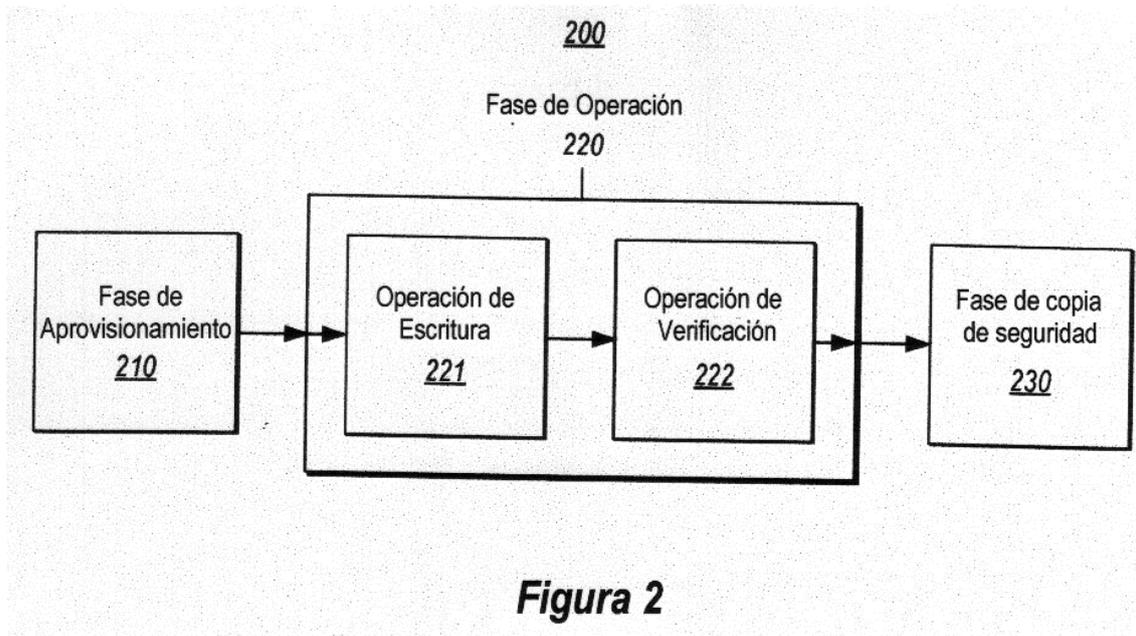


Figura 2

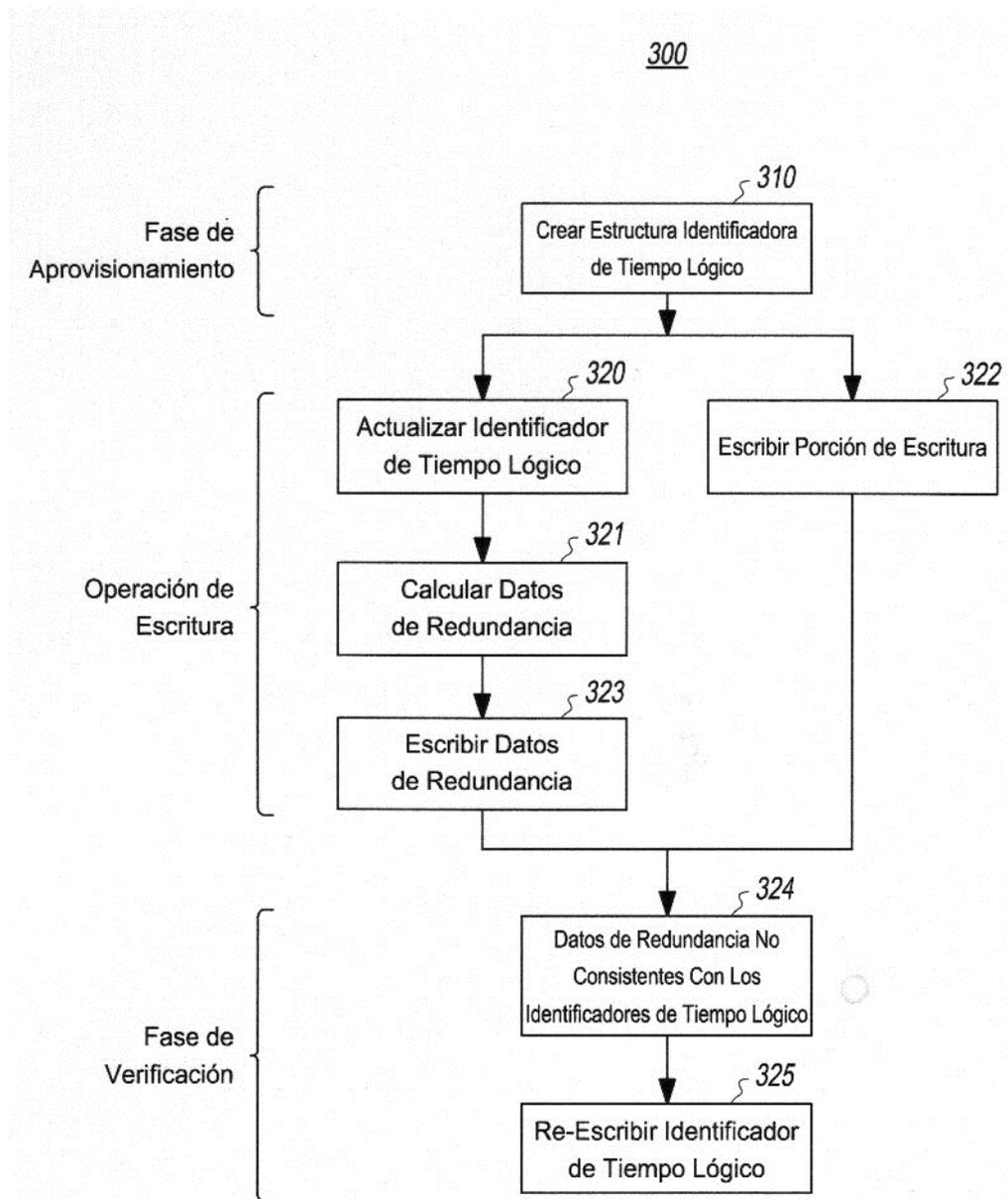


Figura 3

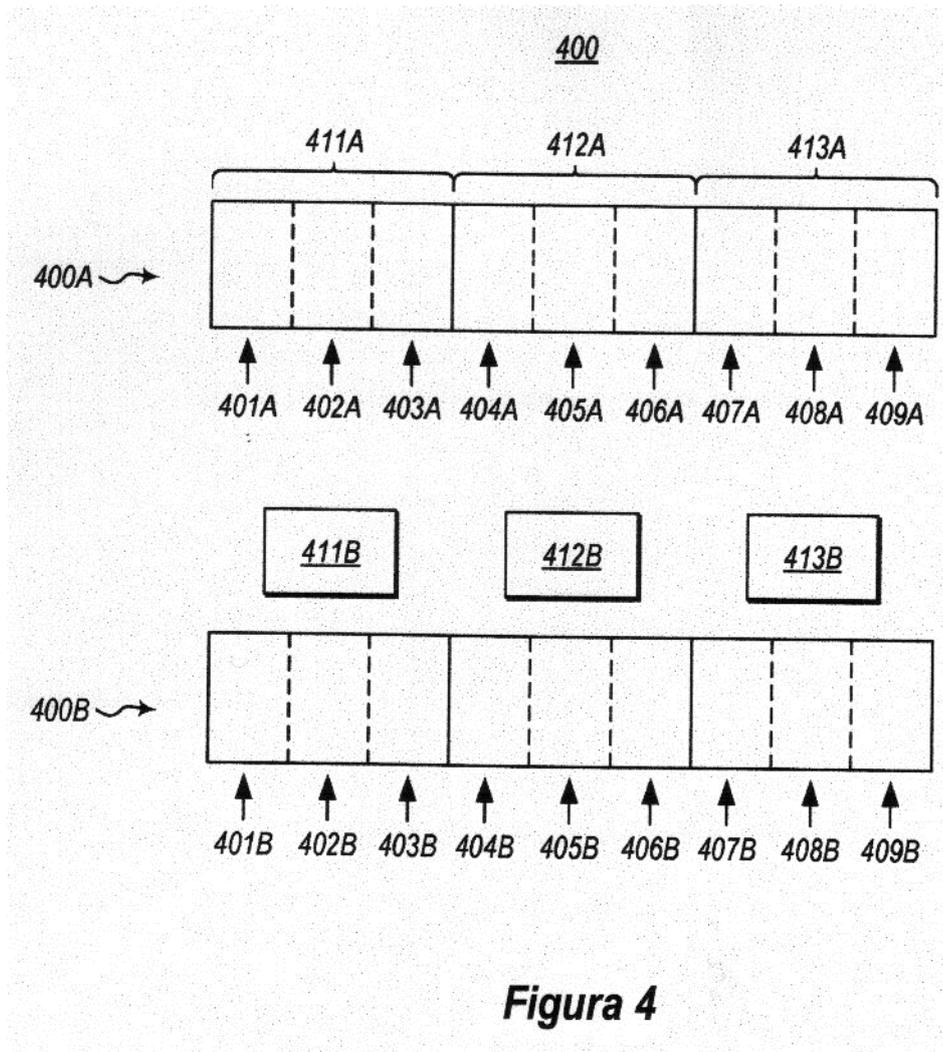


Figura 4

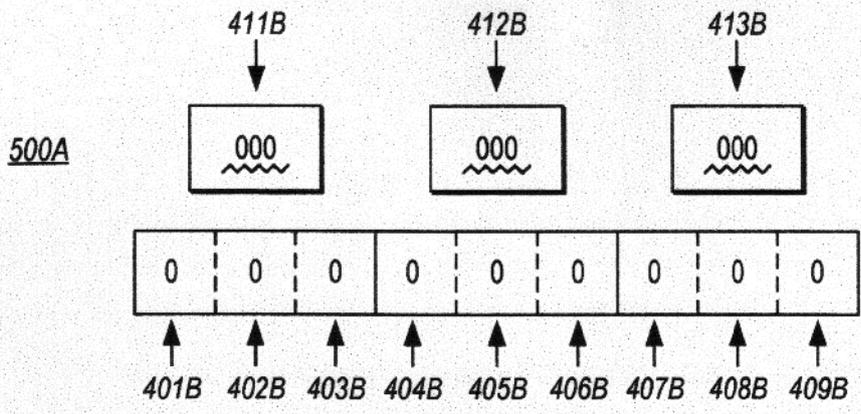


Figura 5A

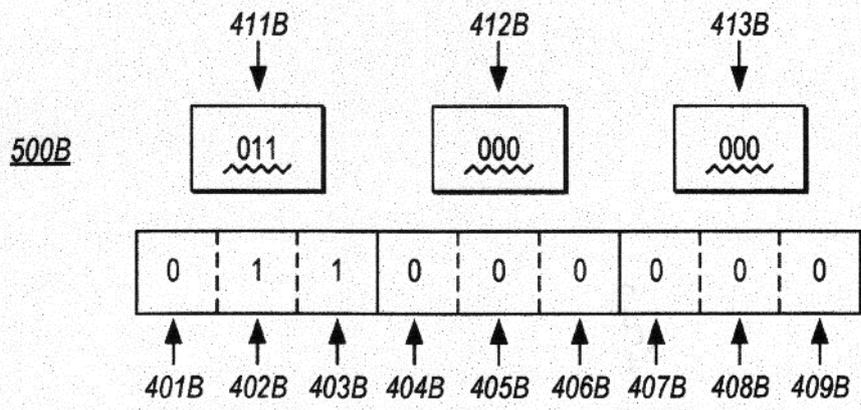


Figura 5B

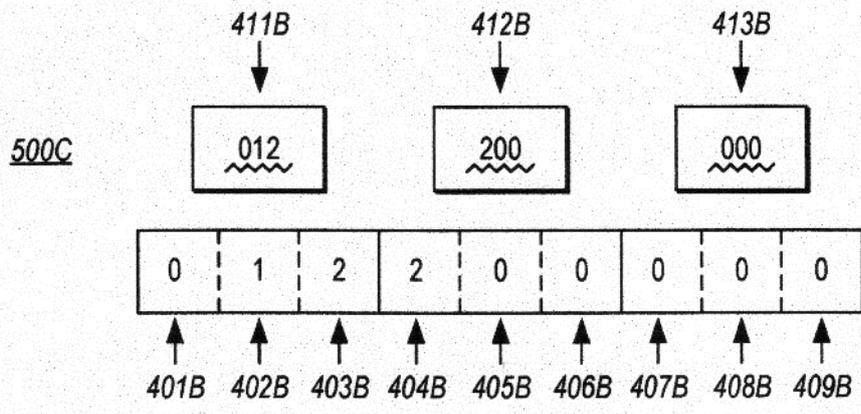


Figura 5C

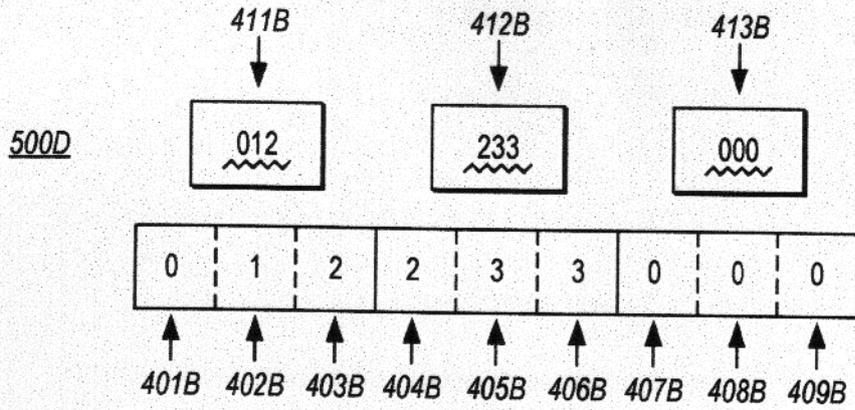


Figura 5D

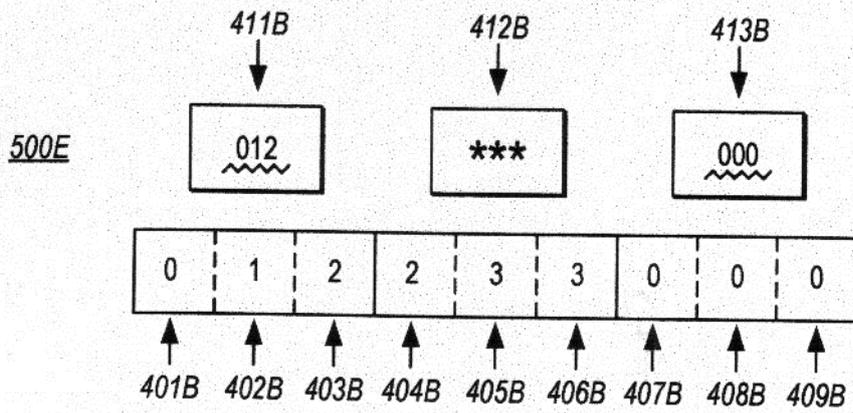


Figura 5E

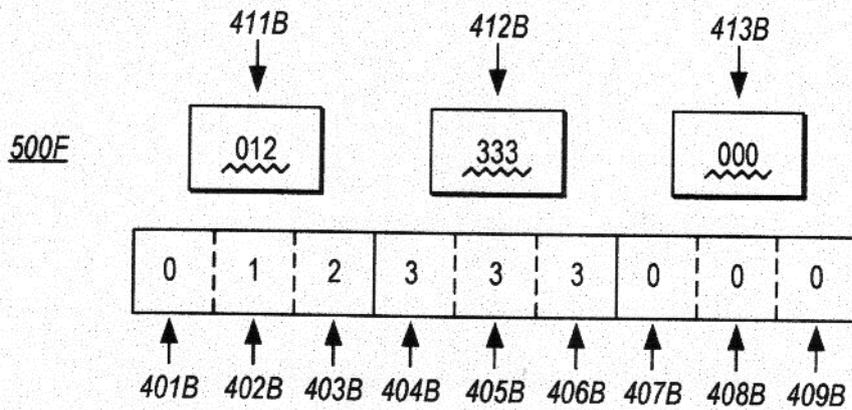
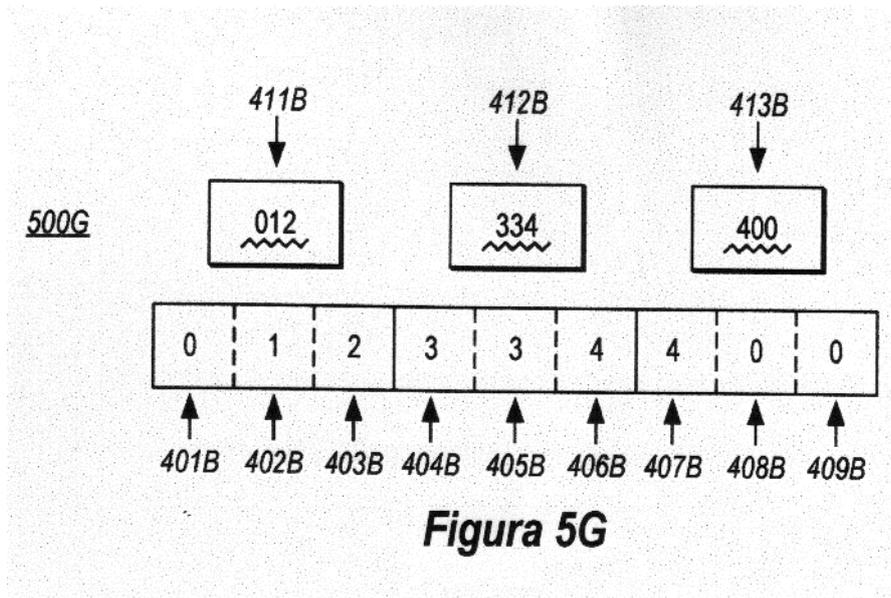


Figura 5F



600A

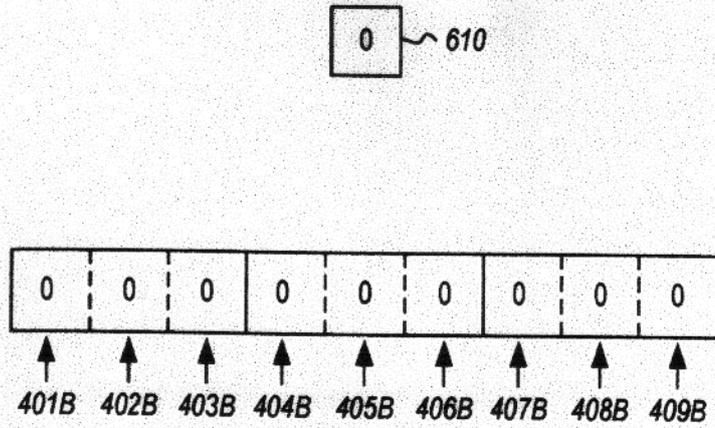


Figura 6A

600B

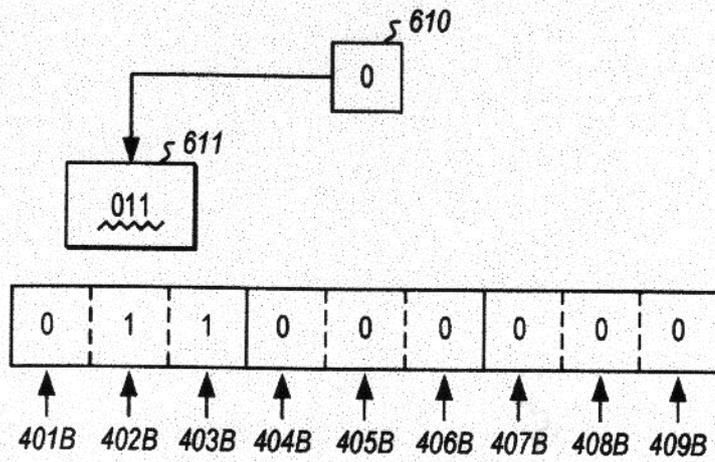


Figura 6B

600C

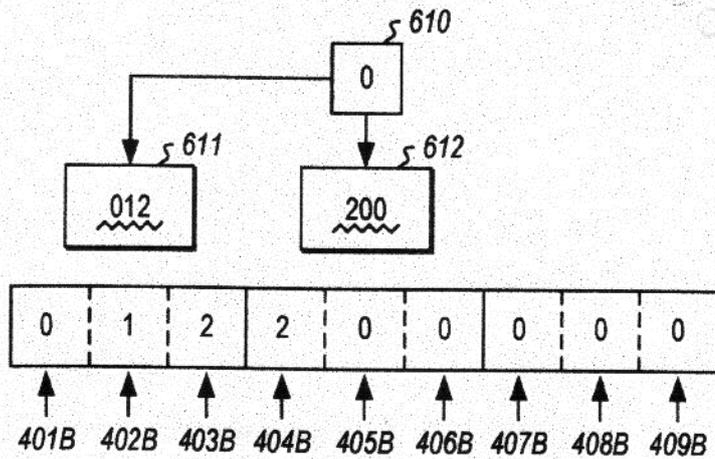


Figura 6C

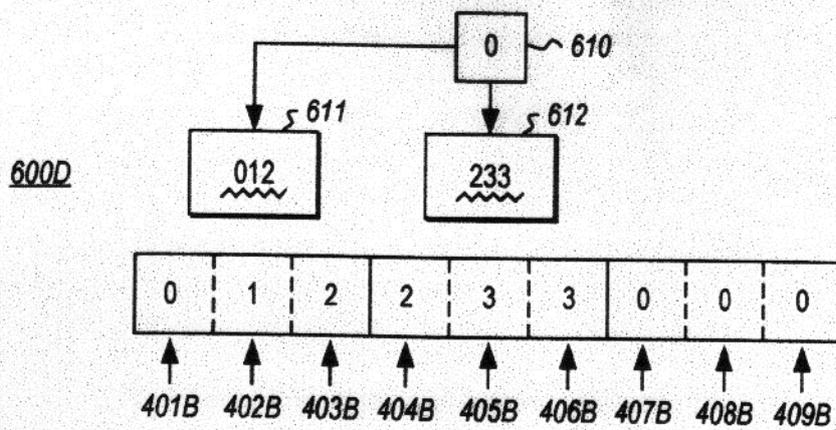


Figura 6D

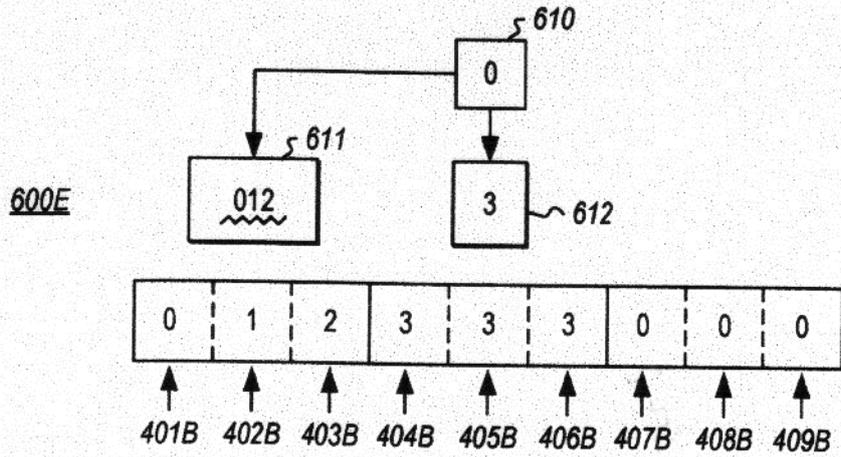


Figura 6E

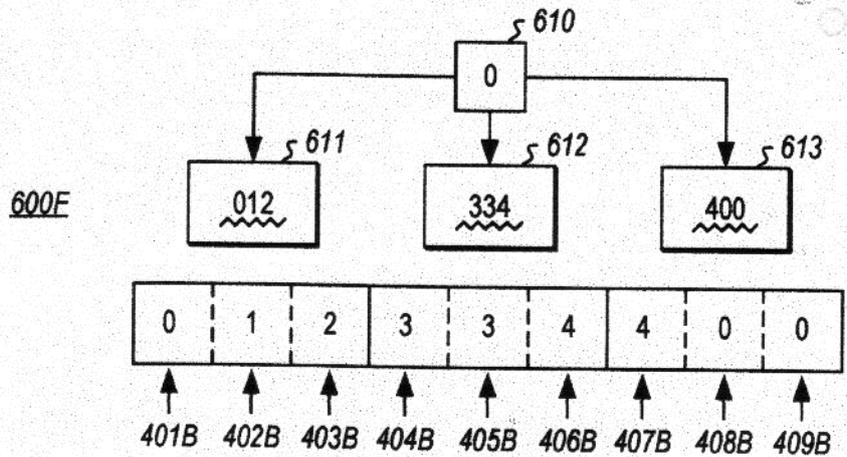


Figura 6F

