

OFICINA ESPAÑOLA DE PATENTES Y MARCAS

ESPAÑA

① Número de publicación: **2 377 087**

② Número de solicitud: 201101048

⑤ Int. Cl.: **G06F 11/34** (2006.01)

⑫

SOLICITUD DE PATENTE

A1

② Fecha de presentación: **22.09.2011**

④ Fecha de publicación de la solicitud: **22.03.2012**

④ Fecha de publicación del folleto de la solicitud: **22.03.2012**

⑦ Solicitante/s: **Universitat Autònoma de Barcelona Edif. A - Campus Universitario, s/n 08193 Bellaterra, Barcelona, ES**

⑧ Inventor/es: **Zarza, Gonzalo; Lugones, Diego; Franco Puentes, Daniel y Luque Fadón, Emilio**

⑦ Agente/Representante: **No consta**

⑤ Título: **Método, sistema y encaminador de evitación de bloqueos en una red de interconexión.**

⑦ Resumen:

Método, sistema y encaminador de evitación de bloqueos en una red de interconexión.

El método comprende:

- a) detectar una situación propensa a un bloqueo; e
- b) identificar un ciclo de encaminamiento involucrado en la situación propensa a un bloqueo detectada, mediante la realización de las siguientes sub-etapas por parte de un encaminador ( $r^i$ ), mediante un mecanismo de búsqueda asíncrono intra-encaminador e inter-encaminadores que no requiere del uso de temporizadores:
  - b1) componer y enviar un mensaje de identificación desde un búfer de entrada ( $a^i_j$ ) del encaminador ( $r^i$ ) a un búfer de salida ( $b^h_k$ ) de otro encaminador ( $r^h$ ); y
  - b2) recibir el mensaje de identificación en el búfer de salida ( $b^i_k$ ) asociado a dicho búfer de entrada ( $a^i_j$ ) del encaminador ( $r^i$ ) que lo compuso, tras su retransmisión por parte de como mínimo otro encaminador ( $r^h$ ) desde un búfer de entrada ( $a^h_j$ ) del mismo.

El sistema y el encaminador están adaptados para implementar el método propuesto por la invención.

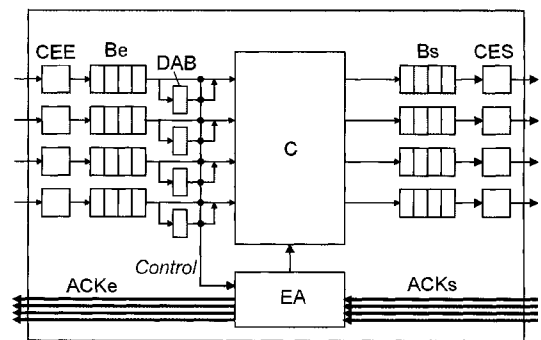


Figura 2

## DESCRIPCIÓN

Método, sistema y encaminador de evitación de bloqueos en una red de interconexión.

### 5 Sector de la técnica

La presente invención concierne en general, en un primer aspecto, a un método de evitación de bloqueos en una red de interconexión basado en la identificación de ciclos de encaminamiento involucrados en situaciones propensas a un bloqueo, y más particularmente a un método que comprende llevar a cabo dicha identificación mediante un mecanismo que no requiere del uso de temporizadores.

La invención pertenece al ámbito de las redes de interconexión de computadores paralelos y se centra en el área de tolerancia a fallos para redes de interconexión de alta velocidad.

El objeto de la presente invención es identificar ciclos de dependencias de recursos en redes de interconexión de alta velocidad de forma asíncrona intra-encaminador e inter-encaminadores. La identificación tiene como finalidad evitar que la ocurrencia de fallos en los dispositivos de red generen situaciones de bloqueo en la red de interconexión, permitiendo así la finalización de las aplicaciones que se encuentren en ejecución en el computador paralelo.

Un segundo aspecto de la invención concierne a un sistema de evitación de bloqueos en una red de interconexión adaptado para implementar el método del primer aspecto.

Un tercer aspecto de la invención concierne a un encaminador previsto para la implementación del método del primer aspecto de la invención.

### Estado de la técnica anterior

El rendimiento de los sistemas de cómputo paralelo de última generación se encuentra íntimamente relacionado con las prestaciones de la red de interconexión que comunica sus elementos de cómputo. Esta interrelación confiere vital importancia a los mecanismos de tolerancia a fallos de la red de interconexión ya que, en última instancia, es la red la que permite el funcionamiento de dichos computadores como entidades coherentes y cohesionadas.

Ante estas circunstancias, la ocurrencia de tan solo un fallo de red es capaz de generar anomalías potencialmente dañinas en el sistema de cómputo, e impedir la correcta finalización de las aplicaciones que se encuentren en ejecución en dicho sistema. De entre estas anomalías, los escenarios de bloqueos, "deadlocks", son quizás los de mayor probabilidad de ocurrencia debido a que la mayor parte de los sistemas de cómputo actuales utilizan algoritmos de encaminamiento que no fueron diseñados para tolerar fallos (situación que incrementa la probabilidad de aparición de bloqueos en dichos sistemas). Este problema se vuelve crítico en situaciones en las que el sistema de cómputo debe lidiar con fallos que aparecen aleatoriamente, es decir con fallos dinámicos. Algunos ejemplos pueden ser los sistemas de transacciones bancarias o los sistemas utilizados para estudios geológicos y de catástrofes naturales, entre otros.

En el contexto de las redes de interconexión, los bloqueos se producen cuando al menos un mensaje no puede llegar a destino, debido a que en su trayecto solicita recursos que no se encuentran disponibles, más específicamente, espacio de almacenamiento en la cola de un dispositivo de red. Esta situación tiene como resultado el bloqueo total del sistema debido a que los mensajes no liberan los recursos que necesitan otros mensajes, mientras solicitan recursos que se encuentran en posesión de otros mensajes. Esto lleva a que un grupo de mensajes se bloquee de forma permanente, provocando que todo el sistema de cómputo quede fuera de servicio. Desafortunadamente, la aparición de bloqueos es frecuente en las situaciones en las que fallan uno o más dispositivos de red. Esto se debe a que incluso un número relativamente bajo de fallos puede generar dependencias cíclicas de recursos, derivadas de los cambios en las características de la red, principalmente en la topología.

Durante las últimas décadas, varios autores han propuesto soluciones al problema de los bloqueos en las redes de interconexión, entre las que se encuentran: limitar la inyección de mensajes en la red [1], [2]; utilizar canales virtuales [3], [4]; y modificar los dispositivos de red y algoritmos de encaminamiento [5], [6], [7]. Desafortunadamente, todas las soluciones propuestas hasta el momento han sido diseñadas para evitar bloqueos en redes libres de fallos, por lo que su aplicabilidad en el ámbito de la tolerancia a fallos es muy reducida.

La presencia de fallos en la red causa que la gran mayoría -por no decir todos- los mecanismos de evitación de bloqueos actuales se vean seriamente afectados y limitados, llegando incluso a no ser efectivos. Dichas limitaciones se relacionan con aspectos importantes de la teoría general de la tolerancia a fallos, tales como el momento y la forma de aparición de los fallos, el número de fallos a tolerar, la duración de los fallos, las características de los sistemas utilizados, etc.

El modo en que se da el fallo es el aspecto más importante a tener en cuenta debido a que se encuentra íntimamente relacionado con la aplicabilidad e idoneidad de las técnicas de evitación de bloqueos. Si se asume un modo de fallos estático, es indispensable conocer de antemano la ubicación de dichos fallos en la red. En contraste, al asumir un

modo de fallos dinámico, los fallos aparecen en tiempos y localizaciones aleatorias durante la ejecución del sistema. El modelo de fallos estático permite implementar soluciones más simples pero el modelo de fallos dinámico plasma de forma fidedigna el comportamiento de los fallos en los sistemas de cómputo reales.

5 Si se utilizan los algoritmos de encaminamiento más simples, los basados en el modo de fallos estático, resulta imposible evitar fallos dinámicos debido a que dichos algoritmos no permiten (por sí mismos) la utilización de caminos alternativos y adaptativos. En contraste, los algoritmos de encaminamiento más complejos, los adaptativos, permiten utilizar caminos alternativos para la evitación de fallos. Aun así, pese a perfilarse como una solución muy prometedora, desde el punto de vista de la tolerancia a fallos, el enfoque de los algoritmos adaptativos presenta un nuevo e importante problema: la aparición de bloqueos o “deadlocks”. Este problema se vuelve crítico en situaciones en las que el sistema de cómputo debe lidiar con fallos que aparecen aleatoriamente a lo largo de la ejecución del sistema, es decir, con fallos dinámicos.

15 Actualmente, la gran mayoría de las técnicas de evitación de bloqueos capaces de tolerar fallos en redes de interconexión utilizan canales virtuales para romper las dependencias cíclicas de recursos entre los mensajes. El uso de canales virtuales limita la escalabilidad de dichas técnicas debido a que el número de canales virtuales necesarios es directamente proporcional al número de fallos a tolerar. Por este motivo, la mayor parte de las técnicas propuestas se basan en modelos de fallos estáticos, a fin de reducir el número de canales virtuales requeridos. Otro grupo de técnicas comúnmente utilizadas se basa en el concepto de limitar la inyección de paquetes, evitando el acaparamiento 20 los recursos de la red a fin de evitar las situaciones de bloqueo indefinidas. Sin embargo, estas técnicas son útiles para algoritmos de encaminamiento basados en la utilización de caminos mínimos, lo que dificulta o simplemente imposibilita tolerar un número medio o alto de fallos. Prácticamente todos estos métodos y técnicas de evitación de bloqueos se basan en diferentes tipos de mecanismos de identificación de ciclos. Desafortunadamente, la mayoría de estos métodos fueron diseñados para operar en ausencia de fallos, por lo que se vuelven inaplicables ante la ocurrencia de fallos de red.

Desde el punto de vista del estado de la técnica, el antecedente más cercano a la invención aquí propuesta es un mecanismo de evitación de bloqueos de tres etapas, denominado “Non-blocking Adaptive Cycles” (NAC), propuesto por los autores de la presente invención en [8] y [9]. A partir de la explicación del método y su entorno de aplicación, es posible inferir de forma directa que el método NAC presenta una serie de problemas de sincronización que limitan su aplicación en redes de interconexión de alta velocidad. Más concretamente, NAC utiliza un mecanismo de identificación de ciclos de tipo “best effort”, que se aplica paso a paso de forma serial y está basado en el control de los ciclos de reloj de los encaminadores para intentar evitar las situaciones conocidas como “falsos negativos”. Es decir, el método utiliza los relojes de los encaminadores para evitar situaciones en las que no se identifica correctamente un ciclo de encaminamiento propenso a generar bloqueos, pese a la existencia del mismo. En este método, la identificación se realiza de forma serial a partir del primer encaminador que detecta las condiciones preestablecidas o de disparo. Durante dicho proceso, se aplica la identificación sobre un búfer, luego, al expirar un tiempo predefinido por el reloj, se procede a realizar la identificación sobre el siguiente búfer de entrada, y así sucesivamente. Esta situación imposibilita la aplicación del mecanismo propuesto en [8] y [9] en redes de interconexión de alta velocidad debido a que su aplicación se ve restringida a casos en los cuales no se pierda la sincronización de los relojes de los encaminadores. Esto se debe a que en NAC pueden ocurrir situaciones en las que más de un encaminador inicie el proceso de identificación sobre un mismo ciclo de dependencias. En esta situación, cada encaminador atiende las peticiones de forma serial, por lo que las respuestas a las peticiones de identificación pueden sufrir ciertos retrasos de duración variable. En los casos en que la acumulación de retrasos haga que el tiempo de identificación supere el valor umbral del reloj, al menos un encaminador llegará a la conclusión errónea de que no existe un ciclo de dependencias, por lo que no aplicará las acciones correspondientes y se generará una situación de bloqueo permanente o “deadlock” en la red. En resumen, el mecanismo de identificación propuesto en [8] y [9] puede alcanzar tanto situaciones de bloqueo como así también situaciones de “falsos negativos”. Adicionalmente, el mecanismo de evitación de bloqueos propuesto en [9] posee otras restricciones que hacen que su aplicación sea muy costosa, aunque no imposible. Estas limitaciones se basan en la disposición lógica de los elementos de evitación de bloqueos, más concretamente de los búferes de evitación de bloqueos, que se encuentran en el camino crítico de los dispositivos de red, lo que ralentiza considerablemente el funcionamiento del encaminador ya que es necesario realizar constantemente comprobaciones acerca de cuál de los dos caminos físicos conectados al búfer de entrada se debe utilizar, tanto en presencia como en ausencia de situaciones propensas a generar bloqueos en la red de interconexión.

### Explicación de la invención

60 Aparece necesario ofrecer una alternativa al estado de la técnica que cubra las lagunas halladas en el mismo, y que en particular supere las mencionadas limitaciones de las que adolecen las propuestas hechas en [8] y [9].

Con tal fin, la presente invención proporciona, en un primer aspecto, un método de evitación de bloqueos en una red de interconexión, que comprende realizar secuencialmente las siguientes etapas:

65 a) detectar al menos una situación propensa a un bloqueo; e

b) identificar al menos un ciclo de encaminamiento involucrado en dicha situación propensa a un bloqueo detectada, también conocido como ciclo de dependencia de recursos, mediante la realización de las siguientes etapas por parte de un encaminador de dicha red de interconexión:

- 5           b1) componer y enviar un mensaje de identificación desde un búfer de entrada de dicho encaminador a al menos un búfer de salida, de al menos otro encaminador, conectado al mismo; y
- b2) recibir dicho mensaje de identificación en el búfer de salida asociado a dicho búfer de entrada de dicho encaminador que lo compuso, tras su retransmisión por parte de al menos dicho otro encaminador desde  
10           un búfer de entrada del mismo.

15           A diferencia del método divulgado en [9], el cual requiere del uso de temporizadores, el método propuesto por la presente invención comprende realizar dicha etapa b) mediante un mecanismo de búsqueda asíncrono intra-encaminador e inter-encaminadores que no requiere del uso de temporizadores.

20           El término búfer debe entenderse en el presente texto como el correspondiente a un dispositivo de almacenamiento utilizado para compensar las diferencias que puedan producirse en la tasa de transmisión de datos entre dispositivos, así como también en el tiempo de ocurrencia de dichos eventos de transmisión.

25           El método propuesto por el primer aspecto de la invención es un método asíncrono y escalable, compatible con las tecnologías de red actuales, capaz de identificar y recorrer dependencias cíclicas de recursos de forma asíncrona intra-encaminador e inter-encaminadores, para asegurar la evitación de bloqueos en redes de interconexión con múltiples fallos dinámicos.

30           Según un ejemplo de realización, el método comprende, para identificar una pluralidad de ciclos de encaminamiento involucrados en situaciones propensas a un bloqueo, realizar, en paralelo, una pluralidad de etapas b) por parte de dicho encaminador, iniciadas mediante una correspondiente pluralidad de etapas b1) de composición y envío, en paralelo, de una pluralidad de dichos mensajes de identificación desde unos respectivos búferes de entrada, y finalizadas mediante una correspondiente pluralidad de etapas b2) de recepción de dichos mensajes de identificación en el búfer de salida asociado a dicho búfer de entrada de dicho encaminador. Este ejemplo de realización es posible gracias a las características de asincronicidad intra-encaminador que el mecanismo de búsqueda utilizado por el método de la invención utiliza, el cual permite que para que un búfer de entrada de un encaminador inicie un proceso de identificación no sea necesario que otro proceso de identificación iniciado por otro búfer de entrada finalice, pudiéndose  
35           realizar ambos en paralelo.

            Para un ejemplo de realización, dicho mecanismo de búsqueda asíncrono intra-encaminador e inter-encaminadores es un mecanismo de búsqueda en amplitud, o BFS (del inglés “Breadth First Search”).

40           Dicha etapa a) comprende, según una realización, evaluar localmente en cada búfer de entrada de dicho encaminador unas condiciones de funcionamiento cuyo cumplimiento establece que el encaminador se encuentra en un estado previo a una situación de bloqueo, que es interpretado como dicha detección de una situación propensa a un bloqueo.

45           El método comprende, por parte de como mínimo dicho otro encaminador, en general por parte de varios otros encaminadores, llevar a cabo la etapa a) para sus búferes de entrada y realizar dicha retransmisión del mensaje o mensajes de identificación recibidos a través del búfer o búferes de entrada donde se cumplan dichas condiciones de funcionamiento, antes de ser recibido o recibidos, en dicha sub-etapa b2), en el búfer o búferes de salida del encaminador que compuso el mensaje o mensajes de identificación enviados.

50           Asimismo, para otro ejemplo de realización, el método comprende al menos iniciar una correspondiente sub-etapa b1) por parte de dicho otro encaminador o de al menos un encaminador de dicha pluralidad de encaminadores, durante la realización de la etapa b) por parte de dicho encaminador. Es decir que es posible que un encaminador implicado en un proceso de identificación iniciado por parte de otro encaminador, inicie por su parte su propio proceso de identificación antes de que finalice aquél en el que se encuentra implicado ya que, aunque esto implique acumulación de retrasos, gracias a las características de asincronicidad inter-encaminadores que el mecanismo de búsqueda utilizado por el método de la invención utiliza, ningún encaminador llegará a la conclusión errónea de que no existe un ciclo de dependencias aunque se retrase la recepción del mensaje de la sub-etapa b2), ya que, a diferencia de lo propuesto en [9], dicha recepción no está limitada a ningún valor umbral temporal.

55           Para un ejemplo de realización, el método comprende, tras identificar un ciclo de encaminamiento en la etapa b) asociado a un búfer de entrada y uno de salida de dicho encaminador, liberar un espacio en dicho búfer de entrada moviendo un primer mensaje de dicho búfer de entrada a un búfer de evitación de bloqueos unido al mismo fuera de su camino crítico.

60           El método comprende, tras dicha liberación de un espacio en el búfer de entrada de dicho encaminador, iniciar un protocolo de recuperación gradual de movimiento de mensajes que garantice el movimiento de mensajes en dicho ciclo de encaminamiento identificado en la etapa b).

Según un ejemplo de realización, el método comprende iniciar una pluralidad de dichos protocolos de recuperación gradual de movimiento de mensajes, en paralelo, para una correspondiente pluralidad de ciclos de encaminamiento identificados en la etapa b).

5 El método propuesto por el primer aspecto de la invención constituye un método de identificación y seguimiento de ciclos de dependencias de recursos tolerante a fallos para redes de interconexión de alta velocidad, basado en una novedosa técnica que permite la realización paralela y distribuida de búsquedas de prioridad en amplitud. El método propuesto consiste en identificar de forma asíncrona los ciclos de dependencias de recursos mediante un proceso de identificación aplicado en sentido contrario al que se utiliza en el proceso de asignación de recursos para el encaminamiento de mensajes a través de la red.

15 El método tiene como finalidad dotar a las redes de interconexión de un mecanismo de identificación de ciclos asíncrono, escalable y fácil de implementar mediante soluciones hardware en los dispositivos de red actuales, para evitar situaciones de bloqueos ante la ocurrencia de uno o más fallos en los dispositivos de red. En resumen, el nuevo método que aquí se propone es un mecanismo de identificación de ciclos asíncrono no bloqueante que, a diferencia de las propuestas anteriores, prescinde de la utilización y el control de los relojes de los encaminadores y permite la aplicación en paralelo de búsquedas en amplitud.

20 Al prescindir de dichos controles, el nuevo mecanismo puede ser aplicado de forma asíncrona ante cualquier circunstancia y topología de red. Además, este nuevo mecanismo permite aplicar de forma paralela tanto los procesos de identificación como así también los de recuperación, lo que representa un novedoso avance en términos de optimización de recursos que permite mejorar el rendimiento total del sistema bajo todas las condiciones de operación de la red de interconexión.

25 El método propuesto por el primer aspecto de la invención permite realizar búsquedas:

1. *Asíncronas*: ya que al no utilizar los relojes de los encaminadores durante el proceso de identificación de ciclos de dependencias de recursos, el método no depende de ningún tipo de sincronización inter-encaminadores. A diferencia de los métodos anteriores, la invención que aquí se propone no presenta problemas en situaciones en las que más de un encaminador pueda iniciar el proceso de identificación de ciclos de forma simultánea o casi simultánea sobre un mismo ciclo de dependencias de recursos;
- 30 2. *En paralelo*: ya que la replicación intra-encaminador de los paquetes de identificación de ciclos se puede realizar de forma paralela y simultánea sobre todos los búferes de entrada que cumplan las condiciones preestablecidas. A diferencia de los métodos anteriores, que aplican el proceso de identificación de forma serial, la propuesta que aquí se presenta permite realizar la identificación en paralelo, prescindiendo de la utilización de relojes en los encaminadores;
- 35 3. *No bloqueantes*: porque diferencia las situaciones de bloqueo de las situaciones de congestión en la red de interconexión sin necesidad de utilizar los relojes de los encaminadores como límites temporales. Esta aportación de la invención evita que el proceso de identificación de ciclos pueda bloquearse bajo ninguna circunstancia (a diferencia de los métodos anteriormente publicados).

45 Gracias a la modificación estructural consistente en la reubicación de los búferes de evitación de bloqueos a un lado del camino crítico del encaminador, llevada a cabo para un ejemplo de realización, se evita que el funcionamiento del encaminador se bloquee y/o ralentice durante el proceso de identificación y tratamiento de ciclos de dependencias de recursos.

50 Un segundo aspecto de la invención concierne a un sistema de evitación de bloqueos en una red de interconexión, que comprende dos o más encaminadores previstos para la implementación del método del primer aspecto, con unos búferes de entrada, unos búferes de salida y una unidad de control apta para realizar dicha detección de dicha etapa a) y dicha identificación de dicha etapa b).

55 Para un ejemplo de realización, como mínimo uno de dichos dos encaminadores comprende un búfer de evitación de bloqueos unido a un búfer de entrada del mismo fuera de su camino crítico, estando su unidad de control prevista para implementar el método del primer aspecto según el ejemplo de realización explicado anteriormente con referencia a la liberación de un espacio en el búfer de entrada del encaminador.

60 Asimismo, según un ejemplo de realización del sistema del segundo aspecto, dicha unidad de control está prevista para ejecutar como mínimo un protocolo de recuperación gradual de movimiento de mensajes mediante la implementación del método del primer aspecto.

65 Un tercer aspecto de la invención concierne a un encaminador de evitación de bloqueos en una red de interconexión, previsto para la implementación del método del primer aspecto, comprendiendo unos búferes de entrada, unos búferes de salida y una unidad de control apta para realizar la detección de la etapa a) y la identificación de la etapa b).

Según un ejemplo de realización, el encaminador propuesto por el tercer aspecto de la invención comprende un búfer de evitación de bloqueos unido a un búfer de entrada del mismo fuera de su camino crítico, estando dicha unidad de control prevista para implementar el método del primer aspecto según el ejemplo de realización explicado anteriormente con referencia a la liberación de un espacio en el búfer de entrada del encaminador.

Dicha unidad de control está prevista, según un ejemplo de realización, para ejecutar como mínimo un protocolo de recuperación gradual de movimiento de mensajes mediante la implementación del método del primer aspecto.

### Breve descripción de los dibujos

Las anteriores y otras ventajas y características se comprenderán más plenamente a partir de la siguiente descripción detallada de unos ejemplos de realización con referencia a los dibujos adjuntos, que deben tomarse a título ilustrativo y no limitativo, en los que:

la Figura 1 es un cuadro de resumen de la notación y los operadores utilizados en la descripción de la presente invención;

la Figura 2 es un diagrama de la arquitectura simplificada del encaminador o dispositivo de red propuesto por el tercer aspecto de la invención;

la Figura 3 ilustra el detalle del proceso de activación y desactivación de los identificadores de ciclos en un encaminador, según un ejemplo de realización del método propuesto por el primer aspecto de la invención;

la Figura 4 es una representación de grafos de dos identificaciones de ciclos de dependencias de recursos en la red de interconexión en base al dispositivo de red de la Figura 2, para un ejemplo de realización del método propuesto por el primer aspecto de la invención;

la Figura 5 es una representación gráfica de las tres etapas que componen el método propuesto por el primer aspecto de la invención, para un ejemplo de realización;

la Figura 6 ilustra a dos encaminadores interconectados que incluyen la notación propuesta para la descripción de unos ejemplos de realización que se realizará en el siguiente apartado con respecto al método propuesto por el primer aspecto de la invención;

la Figura 7A ilustra a la red de interconexión de la Figura 6, donde un encaminador tiene tres búferes de entrada llenos y uno de salida también lleno y conectado a un búfer de entrada lleno de un segundo encaminador, y la Figura 7B es una representación de grafos de dicha red de interconexión;

la Figura 8 es una representación gráfica de las condiciones de detección de situaciones propensas a generar bloqueos a detectar según la etapa a) el método propuesto por el primer aspecto de la invención, ilustradas sobre el esquema de la Figura 6; y

la Figura 9 es una representación gráfica, sobre el esquema de la Figura 6, de las condiciones de recuperación en las que, aplicando el método propuesto por el primer aspecto de la invención para un ejemplo de realización, se ha liberado un espacio en el búfer de entrada de uno de los encaminadores ilustrados como paso previo al inicio de un protocolo de recuperación gradual de movimiento de mensajes.

### Descripción detallada de unos ejemplos de realización

La notación utilizada en el presente apartado para la descripción de la invención se detalla en el cuadro de la Figura 1, así como en la red de interconexión ilustrada en la Figura 6.

Asimismo, en la Figura 2 se muestra un diagrama simplificado de la arquitectura de un dispositivo de red o encaminador basado en los postulados de la invención aquí presentada, incluyendo las líneas de Acknowledge de entrada (ACKe) y de salida (ACKs), que representa, para un ejemplo de realización, tanto al encaminador del tercer aspecto de la invención, como a los incluidos en el sistema del segundo aspecto, como al utilizado por el método del primer aspecto.

En dicha Figura 2 se ha referenciado de manera genérica los búferes de entrada como Be y los de salida como Bs, un conmutador como C, así como una serie de controladores de enlace de entrada como CEE y una serie de controladores de enlace de salida como CES, todos ellos elementos convencionales en la clase de dispositivo de red ilustrado.

Asimismo, siguiendo con la Figura 2, en ella se han indicado los búferes de evitación de bloqueos como DAB y la unidad de encaminamiento y arbitraje como EA. La función de estos elementos se describirá posteriormente.

La novedad de la propuesta hecha por la presente invención se basa en la utilización de un mecanismo de búsqueda asíncrona intra-encaminador e inter-encaminadores que permite identificar en tiempo real la aparición de dependencias

cíclicas en la asignación de recursos de encaminamiento, así como realizar dicha identificación de manera paralela y distribuida.

5 El mecanismo de identificación de ciclos propuesto en la invención es, para un ejemplo de realización, una extensión original de los procesos de búsqueda de prioridad en amplitud o “Breadth-First Search” (BFS), basado en la aplicación paralela de búsquedas distribuidas en cada uno de los dispositivos de red que formen parte del ciclo de dependencias de recursos. Durante la búsqueda, se aplican una serie de acciones de forma paralela y asíncrona partiendo de un dispositivo de red, o encaminador, encargado de iniciar el proceso de identificación de ciclos.

10 Dicho dispositivo de red, indicado en las Figuras 6 a 9 como nodo  $r^i$ , constituye el punto inicial del proceso de identificación y es el dispositivo que envía los mensajes de identificación a través de al menos un búfer de entrada  $a_j^i$  que no posea espacio de almacenamiento disponible y que, a su vez, necesite acceder a un búfer de salida  $b_k^i$  del mismo dispositivo que tampoco posea espacio disponible, es decir que cumpla con las anteriormente referidas como condiciones de funcionamiento cuyo cumplimiento establece que el encaminador  $r^i$  se encuentra en un estado previo a una situación de bloqueo. Estas acciones forman parte del proceso de identificación intra-encaminador.

15 Los dispositivos de red  $r^h$  (ver Figura 9) que se encuentran físicamente conectados a los búferes de entrada  $a_j^i$  del dispositivo de red  $r^i$ , por los que se envíen los mensajes de identificación, repiten el mismo procedimiento aplicado en el dispositivo  $r^i$ , es decir, replican el mensaje de identificación a los dispositivos de red conectados a sus búferes de entrada  $a_j^h$  que cumplan las condiciones detalladas anteriormente, ampliando el proceso de identificación al espacio de búsqueda inter-encaminadores. En caso contrario, los dispositivos de red  $r^h$  que no posean búferes de entrada  $a_j^h$  que cumplan dichas condiciones drenan -eliminan- los mensajes de identificación en lugar de replicarlos. En otras palabras, los dispositivos de red que replican los mensajes de identificación activan de forma distribuida y asíncrona los indicadores de existencia de ciclos, mientras que los dispositivos de red que drenan los mensajes los desactivan.

20 Este proceso se ejemplifica en la Figura 3, donde los círculos grises corresponden a los indicadores activados y los círculos blancos a los indicadores desactivados.

25 En dicha Figura 3 se ilustran cinco nodos  $r^y$ ,  $r^h$ ,  $r^z$ ,  $r^i$ ,  $r^m$ , y puede verse que el nodo central  $r^i$  dispone de cuatro búferes de entrada  $a_1^i$ ,  $a_2^i$ ,  $a_3^i$ ,  $a_4^i$  y cuatro de salida  $b_1^i$ ,  $b_2^i$ ,  $b_3^i$ ,  $b_4^i$  y que ha iniciado tres ciclos de identificación en paralelo a través de tres respectivos búferes de entrada  $a_1^i$ ,  $a_3^i$ ,  $a_4^i$  en los cuales se cumplían las tres condiciones de funcionamiento anteriormente referidas e indicativas de un estado previo a una situación de bloqueo, y que se encuentran conectados respectivamente a un búfer de salida  $b_4^y$  del nodo  $r^y$ , a un búfer de salida  $b_3^h$  del nodo  $r^h$  y a un búfer de salida  $b_1^z$  del nodo  $r^z$  para la transmisión de unos respectivos mensajes de identificación de ciclos.

30 Los ciclos se identifican cuando el dispositivo de red  $r^i$  que inició el proceso de identificación y seguimiento recibe el correspondiente mensaje de identificación de ciclos por el búfer de salida  $b_k^i$  correcto, que en el caso ilustrado por la Figura 3 es el búfer  $b_3^i$ , que ha recibido dicho mensaje de identificación a través del búfer de entrada  $a_m^i$  del nodo  $r^m$ .

35 Este último proceso se ejemplifica en el diagrama de la Figura 4, donde se muestran dos procesos de identificación de ciclos aplicados de manera asíncrona sobre el mismo ciclo de dependencias de recursos, a partir de los dispositivos de red  $r^i$  y  $r^h$ , respectivamente. Explicado de forma coloquial, el método de identificación y seguimiento de ciclos consiste en iluminar paso a paso el camino (ciclos de dependencias), encendiendo (activando) bombillos (indicadores de ciclos) a lo largo del camino. Al ser los propios dispositivos de red -tanto los que replican como los que drenan los mensajes- los encargados de activar y desactivar los indicadores de ciclos, la invención aquí propuesta logra que el método de identificación de ciclos se aplique de forma distribuida y asíncrona a lo largo del ciclo de dependencias de recursos.

40 Este enfoque asíncrono distribuido de búsquedas en amplitud no ha sido propuesto ni aplicado con anterioridad por ningún método de evitación de bloqueos.

50 El nuevo método de identificación que propone la presente invención en su primer aspecto está compuesto, para un ejemplo de realización, por tres etapas, representadas gráficamente en la Figura 5:

- 55 1. La primera identificación del ciclo de dependencias de recursos, iniciada por el primer encaminador  $r^i$  (correspondiente al nodo 1 de la Figura 5 como 1) que detecta las condiciones de ocupación de búferes explicadas anteriormente. Esta etapa finaliza cuando el encaminador  $r^i$  recibe el correspondiente mensaje de identificación de ciclos a través del búfer de salida  $b_k^i$  correcto. Estas acciones corresponden a la etapa 1 de la Figura 5;
- 60 2. Las identificaciones secundarias, que se realizan sobre el mismo ciclo de dependencias de recursos pero son iniciadas por parte de al menos otro encaminador  $r^h$  desde un búfer de entrada  $a_j^h$  del mismo. Estas acciones corresponden a la etapa 2 de la Figura 5, ilustradas para cuatro de dichas identificaciones secundarias iniciadas desde cuatro respectivos encaminadores (o nodos 2, 3, 4 y 5 en la Figura 5);
- 65 3. La propagación de la señal de “recuperación”, iniciada por el encaminador  $r^i$  inmediatamente después de la finalización de la etapa 1. La utilización de la señal de recuperación se explica en detalle más adelante, para un ejemplo de aplicación de la invención. Estas acciones (R1 a R5) corresponden a la etapa 3 de la Figura 5.

La demostración de la aplicabilidad e idoneidad del mecanismo de identificación de ciclos perteneciente a la invención aquí presentada se basa en el hecho de que una de las finalidades de las búsquedas de tipo BFS es “encontrar ciclos en un grafo o probar que dichos ciclos no existen”. Dicho enfoque de búsquedas de tipo BFS se basa en los conceptos matemáticos a partir de los cuales la arquitectura de una red puede ser representada como un grafo (representación abstracta de un conjunto vértices o nodos y aristas) donde los vértices representan los dispositivos de red o procesamiento, según corresponda, y las aristas representan los enlaces que conectan los dispositivos [11]. Si se consideran los búferes de entrada ( $a_j^i$ ) y de salida ( $b_k^i$ ) que componen los búferes lógicos ( $c_{k,j}^{i,m}$ ) de la Figura 6 como nodos, y las conexiones del sistema de control de flujo y señales ACK -con origen y finalización en la unidad de encaminamiento y arbitraje EA- del dispositivo de red como las aristas, se puede obtener una representación equivalente de la red como un grafo, tal y como se ejemplifica en las Figuras 7a y 7b, respectivamente. A partir de dicho grafo se pueden realizar búsquedas de tipo BFS asíncronas y en paralelo intra-encaminador y asíncronas y distribuidas inter-encaminadores, según el método propuesto en la invención aquí presentada.

A fin de mejorar la comprensión de la invención, se describe a continuación un ejemplo de aplicación del método de identificación y seguimiento de ciclos aquí propuesto. El ejemplo de aplicación incluye la detección de situaciones propensas a generar bloqueos, y la identificación y seguimiento de los ciclos de encaminamiento involucrados en dichas situaciones, que es el objeto de la presente invención.

Conceptualmente, la propuesta se fundamenta en evitar que se cumplan dos de las cuatro principales condiciones que causan los bloqueos: la condición de “retención y espera” (“hold-and-wait”), y la condición de “espera circular” (“circular wait”) [10]. Para cumplir este objetivo y evitar así la aparición de bloqueos se incluye un búfer de evitación de bloqueos DAB (por la denominación inglesa “Deadlock Avoidance Buffer”) unido a cada búfer de entrada como se ejemplifica en la Figura 2, y la aplicación de un grupo de acciones simples al momento de acceder y utilizar los búferes de salida que no posean espacio libre disponible. Es importante aclarar que estas acciones solamente se aplican bajo circunstancias específicas directamente relacionadas con el espacio disponible en los búferes de ciertos dispositivos de red, tal y como se explicó anteriormente.

Con la finalidad de poder detectar cuáles son las situaciones que tienen la capacidad de generar bloqueos en la red, se proponen tres condiciones de funcionamiento (ilustradas en la Figura 8, e indicadas como (1), (2) y (3)) que actúan como disparadores de los mecanismos de evitación de bloqueos. La situación descrita por estas condiciones es la de un búfer de entrada de un nodo cualquiera que no posee espacio de almacenamiento disponible (Condición (3)) e intenta escoger como próximo paso un búfer de salida (dentro del mismo nodo de red) que tampoco posee espacio disponible (Condición (2)) y se encuentra físicamente enlazado con otro nodo de red cuyo correspondiente búfer de entrada tampoco tiene espacio disponible (Condición (1)).

Estas tres condiciones se evalúan localmente en cada dispositivo de red  $r^i$  antes de cada ciclo de encaminamiento de mensajes. Si las tres condiciones anteriores se cumplen, el dispositivo de red se encuentra en un estado previo a una situación de bloqueo (ya que aún quedan recursos disponibles en los búferes DAB). Ante dichas circunstancias, el dispositivo de red en el que se plantean estas condiciones (que es el nodo  $r^i$  en la Figura 8) aplica las siguientes acciones:

1. Detiene la inyección de mensajes en el búfer lógico ( $c_{k,j}^{i,m}$ );
2. Detiene la inyección de nuevos mensajes desde el búfer de entrada conectado al nodo de procesamiento local (si fuese necesario), es decir desde el búfer de entrada del nodo  $r^i$  marcado en gris.

Luego de aplicar estas dos acciones, el dispositivo de red  $r^i$  debe identificar y seguir los ciclos de encaminamiento entre pares origen-destino propensos a generar bloqueos. Este proceso es el responsable de identificar el inicio y final de los ciclos de encaminamiento entre pares origen-destino en cada dispositivo de red involucrado en una posible situación de bloqueo. Nótese que dicho proceso corresponden a las etapas 1 y 2 de la Figura 5. La identificación de los ciclos de dependencias de recursos se puede realizar de forma asíncrona y paralela intra-encaminador y asíncrona y distribuida inter-encaminadores, realizando la identificación en sentido contrario al que se utiliza en el proceso de asignación de recursos para el encaminamiento de mensajes a través de la red. Este proceso tiene como objeto asegurar el acceso a los últimos recursos libres en un ciclo de encaminamiento por parte del búfer de entrada correcto.

Los pasos necesarios para la identificación y seguimiento de los ciclos de encaminamiento aplicados por un dispositivo de red son:

1. Componer un nuevo mensaje de identificación que incluya como información: el identificador del dispositivo de red  $r^i$  que lo creó; la identificación del búfer de entrada  $a_j^i$  donde se cumplieron las condiciones de disparo; y el identificador del búfer de salida  $b_k^i$  por el que se debería recibir el mensaje si es parte del ciclo de encaminamiento;
2. Enviar el mensaje de identificación en la dirección opuesta a la dirección de encaminamiento de mensajes, mediante el mecanismo de control de flujo punto a punto del dispositivo de red  $r^i$ . El mensaje se envía a los búferes de salida ( $b_k^h \in c_{k,j}^{h,i}$ ) de los dispositivos de red  $r^h$  (ver Figura 9) físicamente conectados con los búferes de entrada  $a_j^i$  que cumplen con las condiciones de disparo.



Al recibir uno de estos mensajes de identificación, cada dispositivo de red  $r^h$  debe verificar las condiciones y, si al menos uno de los búferes de entrada  $a^h_j$  cumple con dichas condiciones, el dispositivo de red replica (retransmite) este mensaje de identificación a los búferes apropiados. En caso contrario, el dispositivo de red drena (elimina) el mensaje. La primer etapa de identificación del ciclo de la Figura 5 concluye cuando el mensaje de identificación llega al dispositivo de red que lo compuso  $r^i$  a través del búfer de salida  $b^i_k$  correcto, según ilustra la flecha de la línea gris de la Figura 4 para el nodo  $r^i$  o la de la negra para el nodo  $r^h$ .

Una vez identificados los búferes de entrada y salida  $a^i_j, b^i_k$  que forman parte del ciclo, se procede a liberar un espacio en el búfer de entrada  $a^i_j$  mediante la operación  $\text{move}(a^i_j, d^i_j)$ . Esta operación es la que finalmente introduce un nuevo espacio libre en el ciclo de encaminamiento al mover el primer mensaje del búfer de entrada  $a^i_j$  al búfer DAB  $d^i_j$  (ver Figura 9), destrabando la situación de semi bloqueo del búfer de entrada  $a^i_j$  y permitiendo que los mensajes del ciclo avancen hacia sus respectivos destinos. Los casos en los que la identificación no concluye se deben a que el mensaje de identificación ha sido enviado a través del búfer de entrada  $a^i_j$  correcto pero existe al menos un dispositivo de red  $r^s$  a lo largo del ciclo en el que las condiciones no se cumplen por lo que el mensaje de identificación se elimina. En este caso, el ciclo de encaminamiento no se encuentra bloqueado sino congestionado.

Como primera medida, el dispositivo de red  $r^i$  en el que se identificó la situación de riesgo debe componer y enviar (en dirección opuesta al encaminamiento de mensajes de aplicaciones) la señal de recuperación correspondiente a la etapa 3 de la Figura 5, a fin de que el dispositivo de red  $r^h$  que le precede en el ciclo de encaminamiento inicie el protocolo de recuperación gradual de movimiento. Dicha señal es utilizada para alertar a los dispositivos de red precedentes de que el nuevo espacio generado en su búfer de salida  $b^h_k$  es en realidad un espacio generado a partir del uso del búfer DAB. Esta señal permite al dispositivo de red  $r^h$  distinguir entre las dos siguientes situaciones:

- El nuevo espacio fue generado por la desaparición de la situación propensa a generar un bloqueo;
- El nuevo espacio fue generado por la utilización del búfer DAB en el dispositivo de red conectado a su búfer de salida  $b^h_k$ .

Luego de recibir la señal de recuperación, el dispositivo de red  $r^h$  utiliza la información obtenida a partir de su propio proceso de identificación (una de las identificaciones secundarias de la etapa 2 de la Figura 5) para identificar localmente cuál de los búferes de entrada ( $a^h_j$ ) pertenece al ciclo de encaminamiento. Esta acción es crítica ya que cada dispositivo de red que forme parte del ciclo de encaminamiento debe aplicar la operación  $\text{move}(a^h_j, d^h_j)$  sobre el búfer de entrada  $a^h_j$  correcto. Una vez identificados los ciclos de encaminamiento involucrados en situaciones propensas a generar bloqueos, se debe aplicar un conjunto de acciones que garanticen el movimiento de los mensajes en dichos ciclos sin que se lleguen a producir las situaciones de bloqueo. Este conjunto de acciones tiene como punto de partida la etapa de identificaciones secundarias de la Figura 5, en la que al menos un encaminador  $r^h$  inicia su propia identificación del ciclo de dependencias para saber cuáles de sus búferes de entrada  $a^h_j$  y de salida  $b^h_k$  están involucrados en el ciclo y poder luego asegurar la recuperación gradual de movimiento de mensajes hasta que se alcanzan las condiciones de funcionamiento normal de la red, es decir, que no se cumpla al menos una de las condiciones de la Figura 9 (indicadas como (4), (5), (6) y (7)). Si alguna de estas cuatro condiciones no se cumple, el protocolo de recuperación gradual de movimiento de mensajes debe ser finalizado ya que esto implica la red ha vuelto a alcanzar sus condiciones de funcionamiento normal. Como primer medida, se reinicia la inyección de mensajes en el búfer lógico ( $c^{h_i}_{k_j}$ ) mediante la acción  $\text{start\_st}(b^h_k)$  (dando prioridad a los mensajes almacenados en los búferes DAB). Finalmente se vuelve a permitir la inyección de mensajes desde el nodo de procesamiento local mediante la acción  $\text{start\_fw}(a^h_j)$ .

Posible aplicación industrial y puesta en producción de la invención:

La meta de la invención es lograr su incorporación, utilización y puesta en práctica -a nivel industrial y de producción- en redes de interconexión de altas prestaciones. Es por este motivo que la invención busca proveer un método de evitación de bloqueos que permita a una red de interconexión de alta velocidad tolerar uno o más fallos en sus componentes. El esquema de la Figura 2 representa un posible punto de partida para el diseño de dispositivos compatibles con la invención, aunque solamente es ilustrativo y no limitativo.

Un experto en la materia podría introducir cambios y modificaciones en los ejemplos de realización descritos sin salirse del alcance de la invención según está definido en las reivindicaciones adjuntas.

## Referencias

[1] A. W. Roscoe, "Routing messages through networks: an exercise in deadlock avoidance", in Programming of Transputer Based Machines: Proceedings of 7th occam User Group Technical Meeting, 1987.

[2] V. Puente, C. Izu, R. Bevide, J. A. Gregorio, F. Vallejo, and J. M. Pallezo, "The adaptive bubble router", *J. Parallel Distrib. Comput.*, vol. 61, no. 9, pp. 1180-1208, 2001.

[3] J. Duato, "A theory of deadlock-free adaptive multicast routing in wormhole networks", *Parallel and Distributed Systems, IEEE Transactions on*, vol. 6, pp. 976-987, Sep 1995.

## ES 2 377 087 A1

[4] J. Duato, "A necessary and sufficient condition for deadlock-free routing in cut-through and store-and-forward networks", *IEEE Trans. Parallel Distrib. Syst.*, vol. 7, no. 8, pp. 841-854, 1996.

5 [5] S. Konstantinidou and L. Snyder, "The chaos router: a practical application of randomization in network routing", in SPAA '90: Proc. of the 2nd Annual ACM Symposium on Parallel Algorithms and Architectures, (New York, USA), pp. 21-30, 1990.

10 [6] S. Konstantinidou and L. Snyder, "Chaos router: architecture and performance", *SIGARCH Comput. Archit. News*, vol. 19, no. 3, pp. 212-221, 1991.

[7] S. Konstantinidou and L. Snyder, "The chaos router", *Computers, IEEE Transactions on*, vol. 43, pp. 1386-1397, Dec 1994.

15 [8] G. Zarza, D. Lugones, D. Franco, and E. Luque, "Deadlock avoidance for interconnection networks with multiple dynamic faults", in 18th Euromicro International Conference on Parallel, *Distributed and Network-Based Computing*, pp. 276-280, Feb 2010.

20 [9] G. Zarza, D. Lugones, D. Franco, and E. Luque, "Non-blocking adaptive cycles: Deadlock avoidance for fault-tolerant interconnection networks", in Cluster Computing Workshops and Posters (CLUSTER WORKSHOPS), 2010 IEEE International Conference on, pp. 1-4, Sep 20th 2010.

[10] E. G. Coffman, M. Elphick, and A. Shoshani, "System deadlocks", *ACM Comput. Surv.*, vol. 3, no. 2, pp. 67-78, 1971.

25 [11] L.-H. Hsu and C.-K. Lin, *Graph Theory and Interconnection Networks*. *CRC Press*, 2008.

30

35

40

45

50

55

60

65

## REIVINDICACIONES

1. Método de evitación de bloqueos en una red de interconexión, que comprende realizar secuencialmente las siguientes etapas:

a) detectar al menos una situación propensa a un bloqueo; e

b) identificar al menos un ciclo de encaminamiento involucrado en dicha situación propensa a un bloqueo detectada, mediante la realización de las siguientes sub-etapas por parte de un encaminador ( $r^i$ ) de dicha red de interconexión:

b1) componer y enviar un mensaje de identificación desde un búfer de entrada ( $a^i_j$ ) de dicho encaminador ( $r^i$ ) a al menos un búfer de salida ( $b^h_k$ ), de al menos otro encaminador ( $r^h$ ), conectado al mismo; y

b2) recibir dicho mensaje de identificación en el búfer de salida ( $b^i_k$ ) asociado a dicho búfer de entrada ( $a^i_j$ ) de dicho encaminador ( $r^i$ ) que lo compuso, tras su retransmisión por parte de al menos dicho otro encaminador ( $r^h$ ) desde un búfer de entrada ( $a^h_j$ ) del mismo;

estando el método **caracterizado** porque comprende realizar dicha etapa b) mediante un mecanismo de búsqueda asíncrono intra-encaminador e inter-encaminadores que no requiere del uso de temporizadores.

2. Método según la reivindicación 1, **caracterizado** porque, para identificar una pluralidad de ciclos de encaminamiento involucrados en situaciones propensas a un bloqueo, el método comprende realizar, en paralelo, una pluralidad de etapas b) por parte de dicho encaminador ( $r^i$ ), iniciadas mediante una correspondiente pluralidad de sub-etapas b1) de composición y envío, en paralelo, de una pluralidad de dichos mensajes de identificación desde unos respectivos búferes de entrada ( $a^i_1, a^i_2, a^i_3, a^i_4$ ), y finalizadas mediante una correspondiente pluralidad de etapas b2) de recepción de dichos mensajes de identificación en el búfer de salida ( $b^i_1, b^i_2, b^i_3, b^i_4$ ) asociado a dicho búfer de entrada ( $a^i_1, a^i_2, a^i_3, a^i_4$ ) de dicho encaminador ( $r^i$ ).

3. Método según una cualquiera de las reivindicaciones anteriores, **caracterizado** porque dicho mecanismo de búsqueda asíncrono intra-encaminador e inter-encaminadores es un mecanismo de búsqueda en amplitud.

4. Método según una cualquiera de las reivindicaciones anteriores, **caracterizado** porque dicho mensaje de identificación incluye por lo menos información referente al identificador del encaminador ( $r^i$ ) que lo creó, a la identificación del búfer de entrada ( $a^i_j$ ) por el que se envió y al identificador del búfer de salida ( $b^i_k$ ) por el que se debería recibir el mensaje de identificación si es parte del ciclo de encaminamiento.

5. Método según una cualquiera de las reivindicaciones anteriores, **caracterizado** porque dicha etapa a) comprende evaluar localmente en cada búfer de entrada ( $a^i_1, a^i_2, a^i_3, a^i_4$ ) de dicho encaminador ( $r^i$ ) unas condiciones de funcionamiento cuyo cumplimiento establece que el encaminador ( $r^i$ ) se encuentra en un estado previo a una situación de bloqueo, que es interpretado como dicha detección de una situación propensa a un bloqueo.

6. Método según la reivindicación 5, **caracterizado** porque comprende, por parte de al menos dicho otro encaminador ( $r^h$ ), llevar a cabo dicha etapa a) para sus búferes de entrada ( $a^h_1, a^h_2, a^h_3, a^h_4$ ) y realizar dicha retransmisión del mensaje o mensajes de identificación recibidos a través del búfer o búferes de entrada ( $a^h_1, a^h_2, a^h_3, a^h_4$ ) donde se cumplan dichas condiciones de funcionamiento.

7. Método según la reivindicación 6, **caracterizado** porque comprende realizar dicha retransmisión del mensaje o mensajes de identificación a través de una pluralidad de encaminadores ( $r^h, r^y, r^z, r^m$ ), por sus búferes de entrada que cumplan con dichas condiciones de funcionamiento, antes de ser recibido o recibidos, en dicha sub-etapa b2), en el búfer o búferes de salida ( $b^i_k$ ) de dicho encaminador ( $r^i$ ) que compuso el mensaje o mensajes de identificación enviados.

8. Método según la reivindicación 5, **caracterizado** porque dicha condiciones de funcionamiento describen una situación en la que un búfer de entrada ( $a^i_j$ ) de un encaminador ( $r^i$ ) no posee espacio de almacenamiento disponible e intenta escoger como próximo paso un búfer de salida ( $b^i_k$ ), del mismo encaminador ( $r^i$ ), que tampoco posee espacio disponible y se encuentra físicamente enlazado con el búfer de entrada ( $a^m_j$ ) de otro encaminador ( $r^m$ ) que tampoco tiene espacio disponible.

9. Método según una cualquiera de las reivindicaciones anteriores, **caracterizado** porque comprende al menos iniciar una correspondiente sub-etapa b1) por parte de al menos dicho otro encaminador ( $r^h$ ) o de al menos un encaminador de dicha pluralidad de encaminadores ( $r^h, r^m$ ), durante la realización de la etapa b) por parte de dicho encaminador ( $r^i$ ).

10. Método según una cualquiera de las reivindicaciones anteriores, **caracterizado** porque comprende, tras identificar un ciclo de encaminamiento en dicha etapa b) asociado a un búfer de entrada ( $a^i_j$ ) y uno de salida ( $b^i_k$ ) de dicho encaminador ( $r^i$ ), liberar un espacio en dicho búfer de entrada ( $a^i_j$ ) moviendo un primer mensaje de dicho búfer de entrada ( $a^i_j$ ) a un búfer de evitación de bloqueos ( $d^i_j$ ) unido al mismo fuera de su camino crítico.

## ES 2 377 087 A1

11. Método según la reivindicación 10, **caracterizado** porque comprende, tras dicha liberación de un espacio en el búfer de entrada ( $a_j^i$ ) de dicho encaminador ( $r^i$ ), iniciar un protocolo de recuperación gradual de movimiento de mensajes que garantice el movimiento de mensajes en dicho ciclo de encaminamiento identificado en la etapa b).

5 12. Método según la reivindicación 11, **caracterizado** porque comprende iniciar una pluralidad de dichos protocolos de recuperación gradual de movimiento de mensajes, en paralelo, para una correspondiente pluralidad de ciclos de encaminamiento identificados en la etapa b).

10 13. Sistema de evitación de bloqueos en una red de interconexión, **caracterizado** porque comprende al menos dos encaminadores ( $r^i, r^h$ ) previstos para la implementación del método según una cualquiera de las reivindicaciones anteriores, con unos búferes de entrada ( $a_j^i, a_j^h$ ), unos búferes de salida ( $b_k^i, b_k^h$ ) y una unidad de control apta para realizar dicha detección de dicha etapa a) y dicha identificación de dicha etapa b).

15 14. Sistema según la reivindicación 13, **caracterizado** porque al menos uno ( $r^i$ ) de dichos dos encaminadores ( $r^i, r^h$ ) comprende un búfer de evitación de bloqueos ( $d_j^i$ ) unido a un búfer de entrada ( $a_j^i$ ) del mismo fuera de su camino crítico, estando su unidad de control prevista para implementar el método según la reivindicación 9 para realizar dicha liberación de un espacio en dicho búfer de entrada ( $a_j^i$ ).

20 15. Sistema según la reivindicación 14, **caracterizado** porque dicha unidad de control está prevista para ejecutar al menos un protocolo de recuperación gradual de movimiento de mensajes mediante la implementación del método según la reivindicación 11 ó 12.

25 16. Encaminador de evitación de bloqueos en una red de interconexión, **caracterizado** porque está previsto para la implementación del método según una cualquiera de las reivindicaciones 1 a 12, comprendiendo unos búferes de entrada ( $a_j^i$ ), unos búferes de salida ( $b_k^i$ ) y una unidad de control apta para realizar dicha detección de dicha etapa a) y dicha identificación de dicha etapa b).

30 17. Encaminador según la reivindicación 16, **caracterizado** porque comprende un búfer de evitación de bloqueos ( $d_j^i$ ) unido a un búfer de entrada ( $a_j^i$ ) del mismo fuera de su camino crítico, estando dicha unidad de control prevista para implementar el método según la reivindicación 10 para realizar dicha liberación de un espacio en dicho búfer de entrada ( $a_j^i$ ).

35 18. Encaminador según la reivindicación 17, **caracterizado** porque dicha unidad de control está prevista para ejecutar al menos un protocolo de recuperación gradual de movimiento de mensajes mediante la implementación del método según la reivindicación 11 ó 12.

40

45

50

55

60

65

Notación	
$r^i$	un nodo de red (encaminador),
$a_j^i$	un buffer de entrada del nodo $r^i$ ,
$b_k^i$	un buffer de salida del nodo $r^i$ ,
$d_j^i$	un buffer de evitación de bloqueos del nodo $r^i$ ,
$cap(arg)$	la capacidad del buffer $arg$ ,
$size(arg)$	el número de mensajes almacenados en el buffer $arg$ ,
$output(arg)$	el buffer de salida $b_k^i$ asignado al primer mensaje en el buffer $arg$ ,
$c_{k,j}^{i,m}$	un buffer lógico compuesto por el buffer de salida $b_k^i$ y el buffer de entrada $a_j^m$ enlazado al buffer de salida $b_k^i$ .
Operadores	
$move(arg1, arg2)$	Quita el primer mensaje del buffer $arg1$ y lo encola en el buffer $arg2$ ,
$stop\_st(arg)$	Detiene la inyección de mensajes en el buffer $arg$ ,
$start\_st(arg)$	Permite la inyección de mensajes en el buffer $arg$ ,
$stop\_fw(arg)$	Detiene el envío de mensajes desde el buffer $arg$ ,
$start\_fw(arg)$	Permite el envío de mensajes desde el buffer $arg$ .

Figura 1

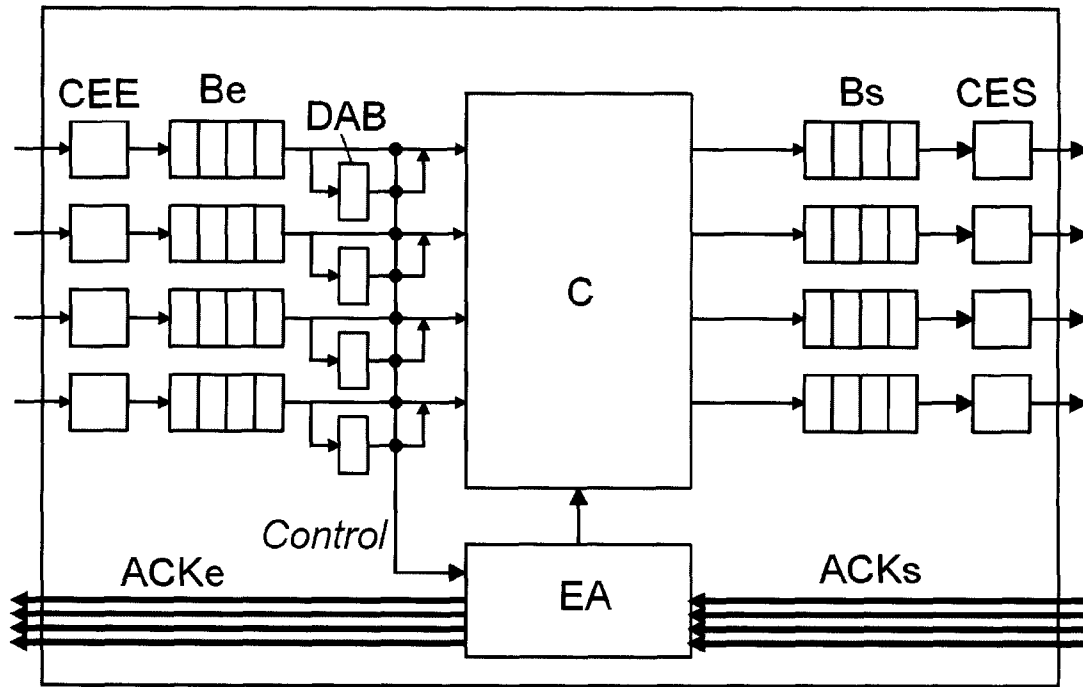


Figura 2

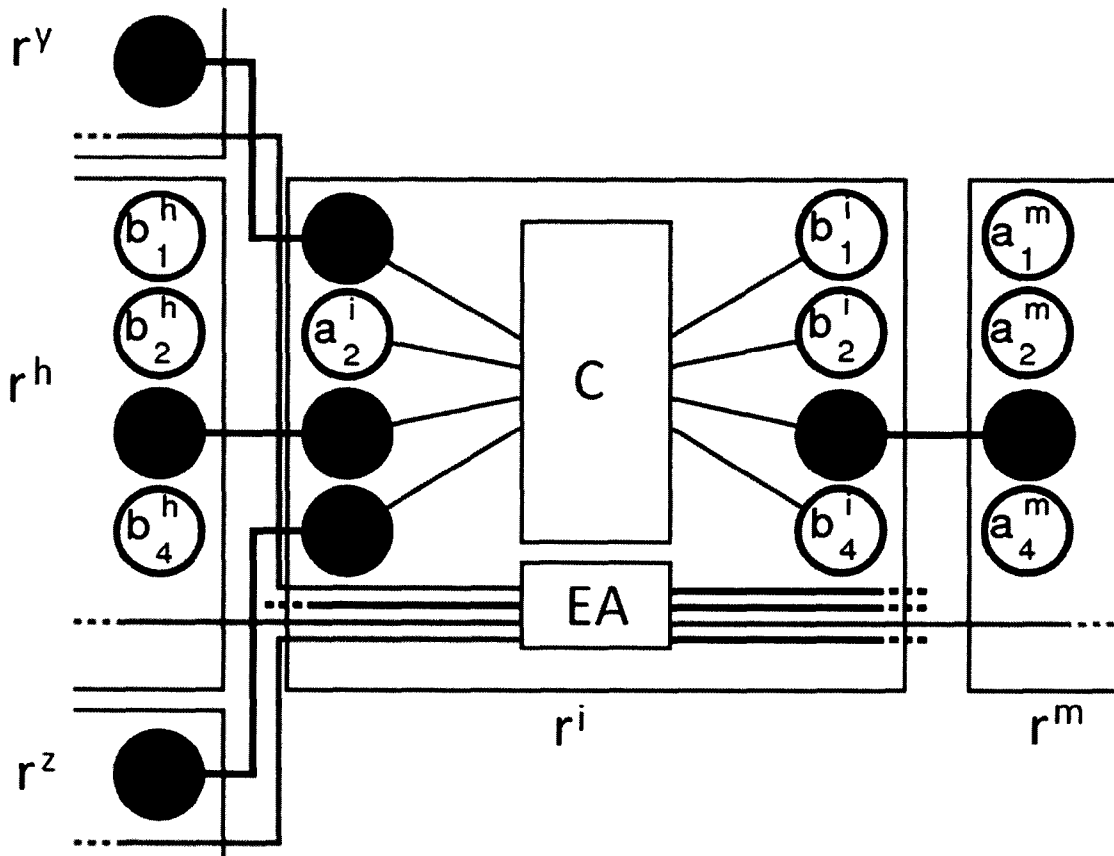
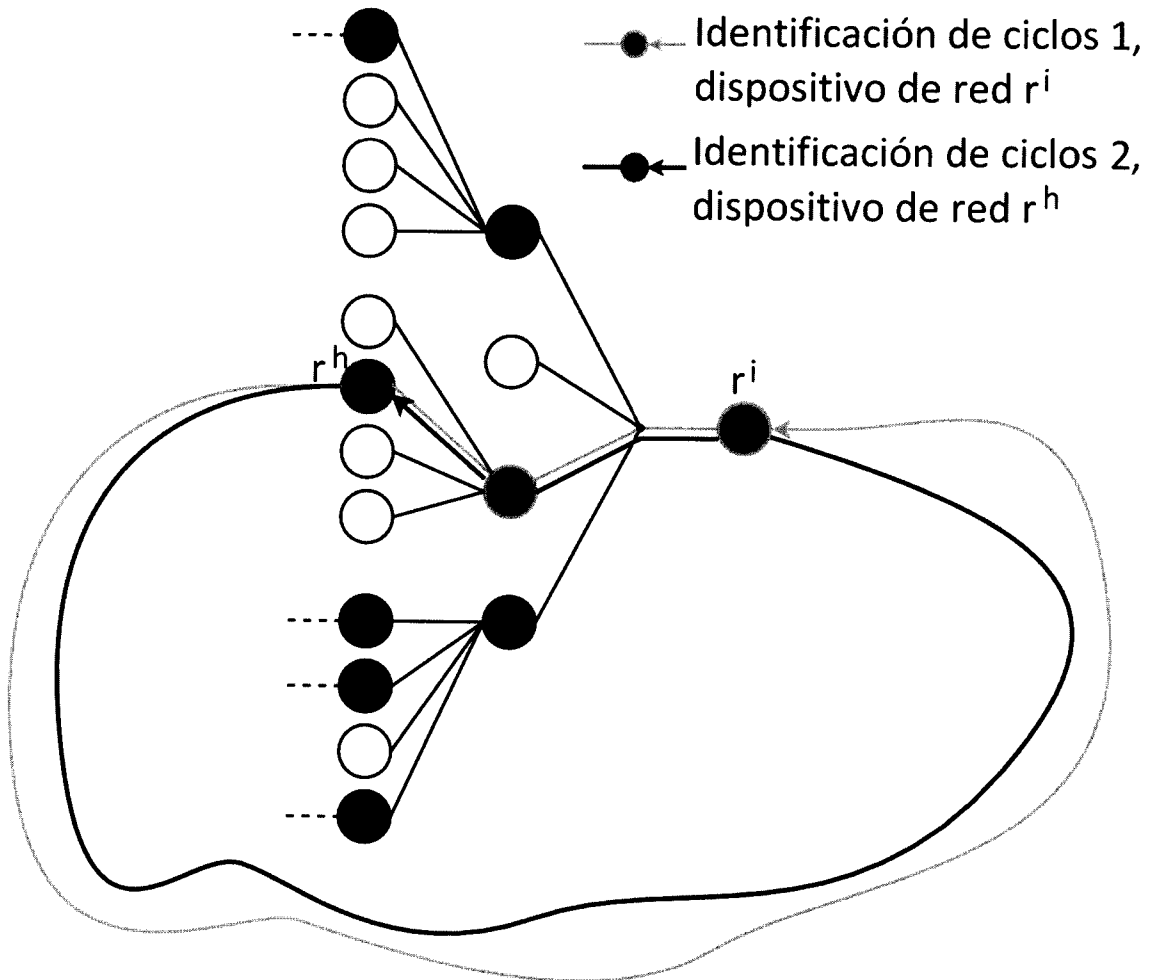
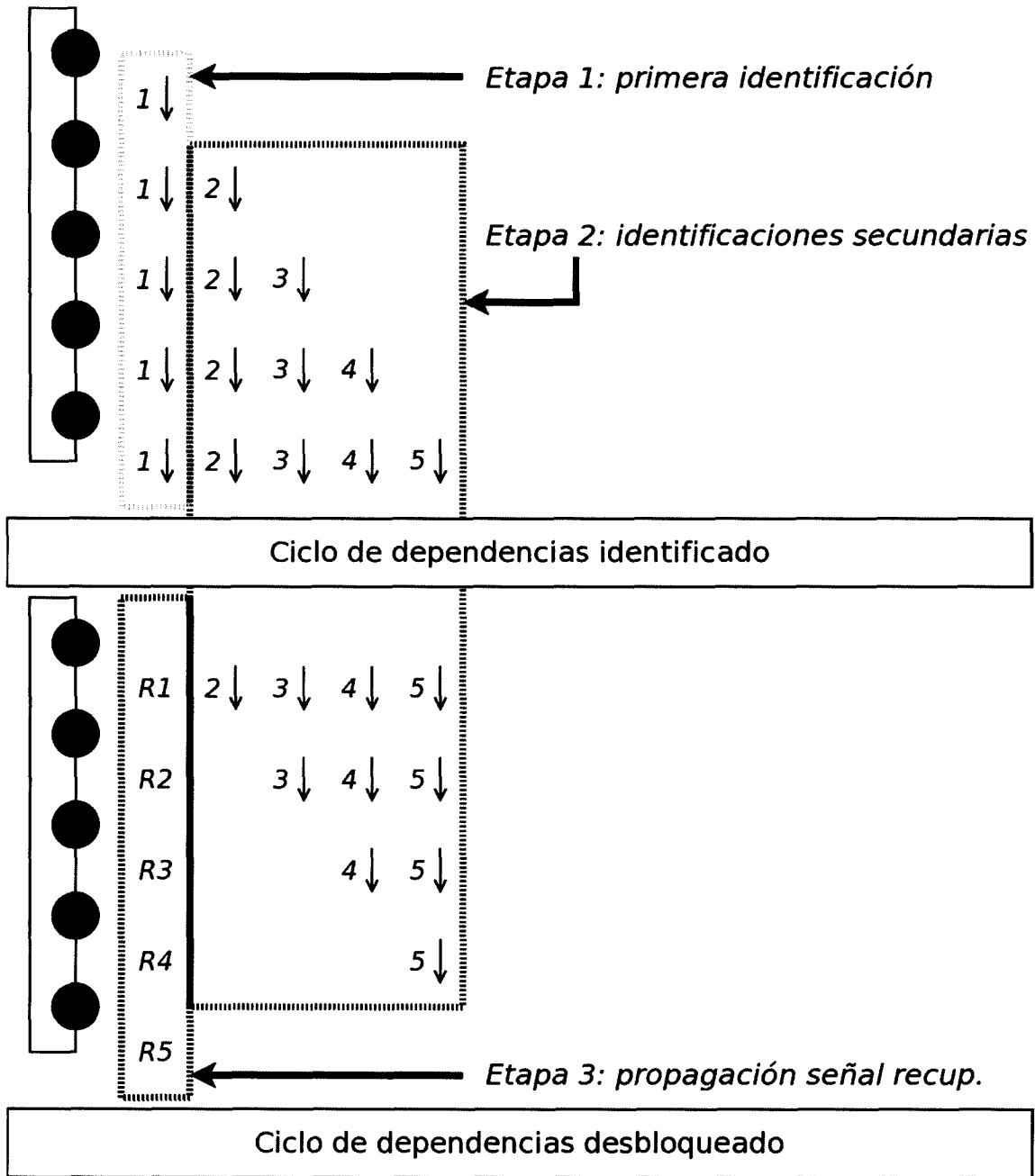


Figura 3



**Figura 4**



**Figura 5**



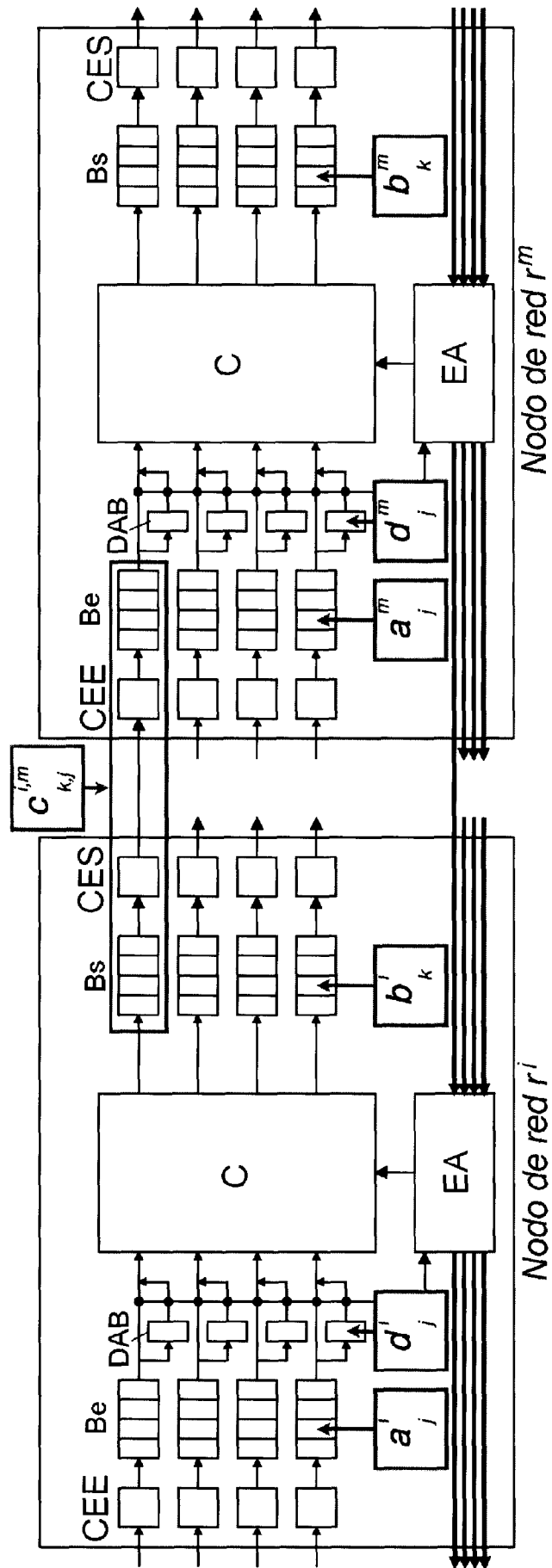


Figura 6

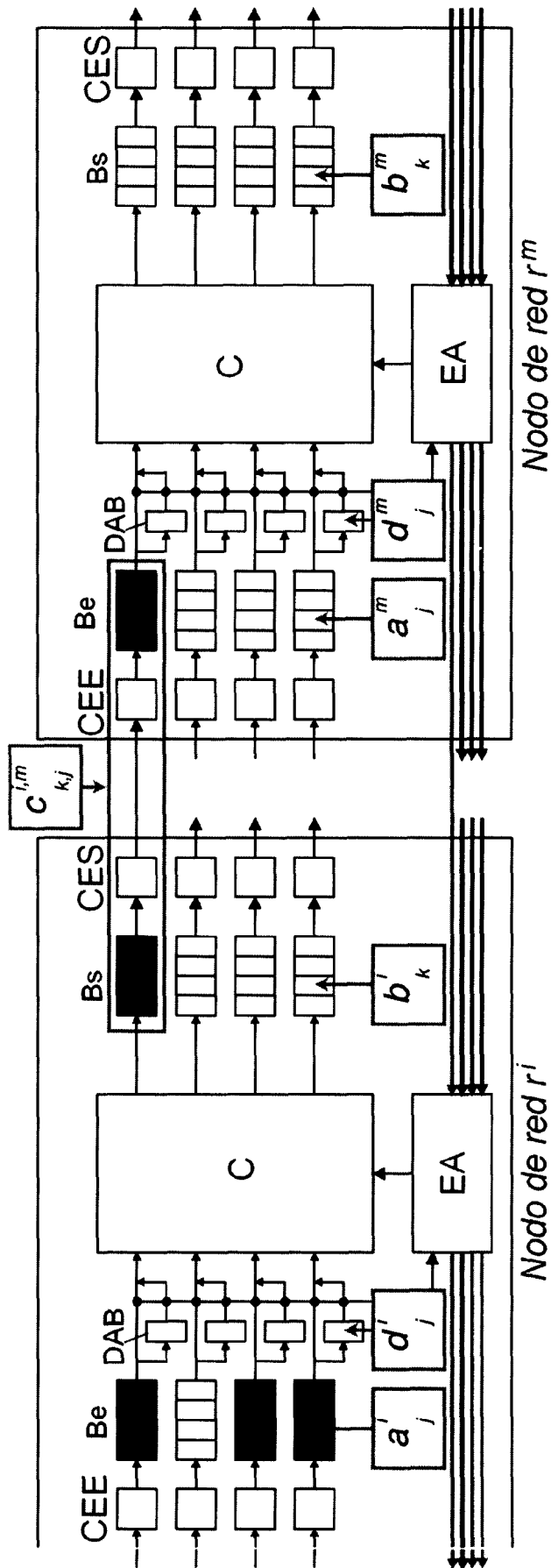


Figura 7A

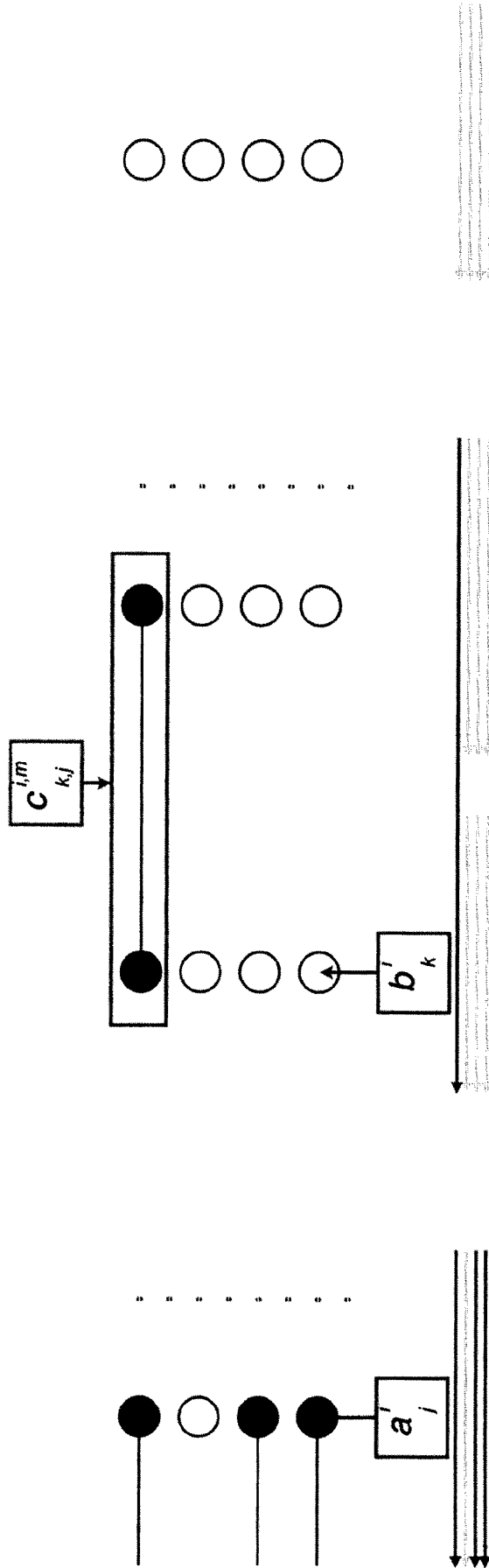


Figura 7B

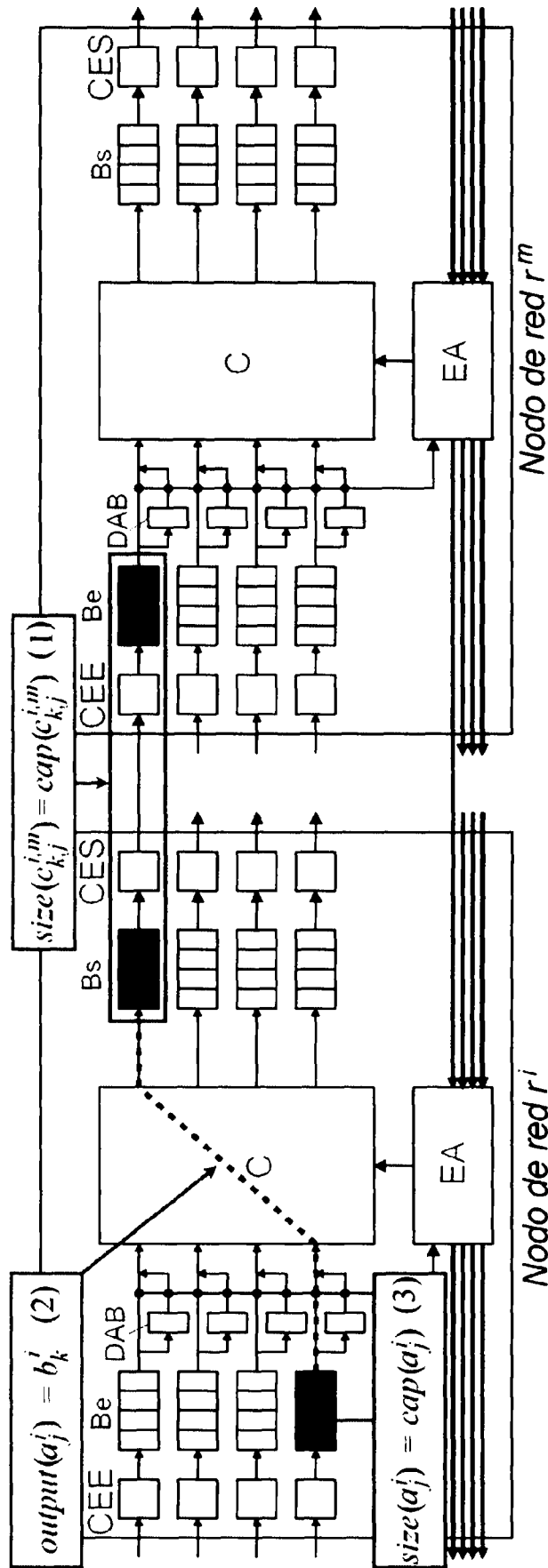


Figura 8

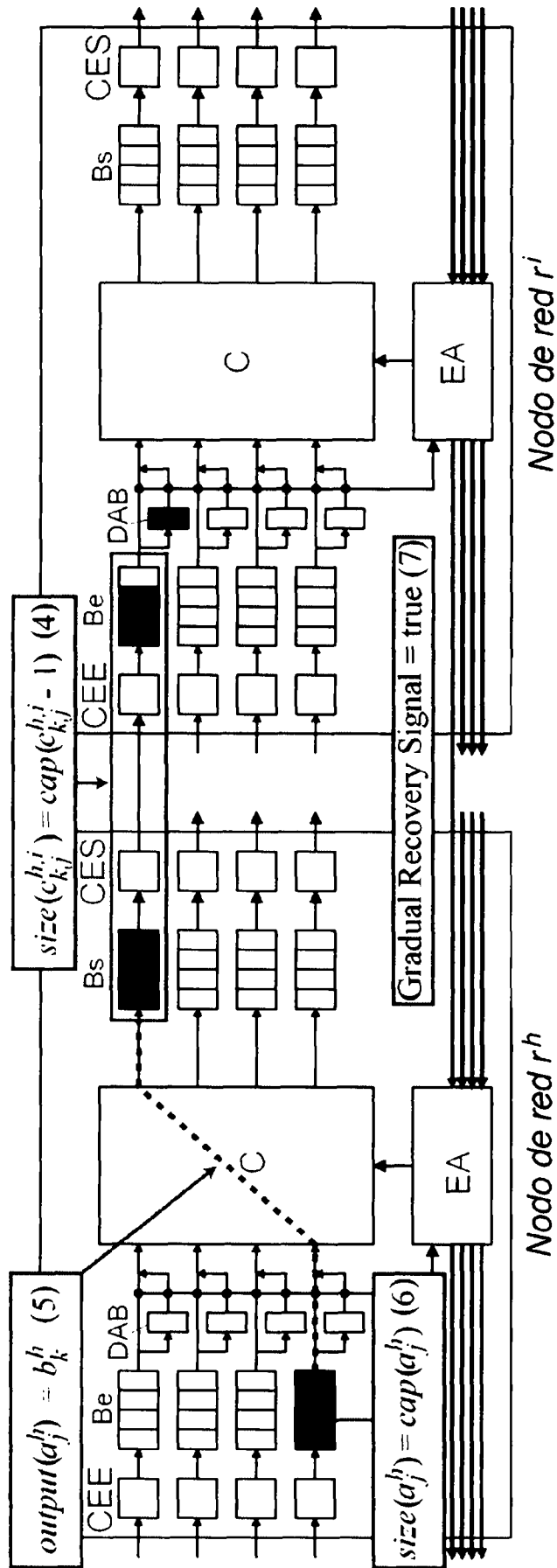


Figura 9



OFICINA ESPAÑOLA  
DE PATENTES Y MARCAS

ESPAÑA

②<sup>1</sup> N.º solicitud: 201101048

②<sup>2</sup> Fecha de presentación de la solicitud: 22.09.2011

③<sup>2</sup> Fecha de prioridad:

## INFORME SOBRE EL ESTADO DE LA TÉCNICA

⑤<sup>1</sup> Int. Cl.: **G06F11/34** (2006.01)

### DOCUMENTOS RELEVANTES

Categoría	⑥ <sup>1</sup> Documentos citados	Reivindicaciones afectadas
X	GONZALO ALBERTO ZARZA, "Multipath Fault-tolerant Routing Policies to deal with Dynamic Link Failures in High Speed Interconnection Networks" - Universidad Autónoma de Barcelona, Escuela de Ingeniería - Dpto. de Arquitectura de computadores y Sistemas Operativos. Capítulo 4, páginas 53-66. URL:// <a href="http://www.thesisenred.net/bitstream/handle/10803/51494/gaz1de1.pdf?sequence=1">http://www.thesisenred.net/bitstream/handle/10803/51494/gaz1de1.pdf?sequence=1</a>	1-18

Categoría de los documentos citados

X: de particular relevancia

Y: de particular relevancia combinado con otro/s de la misma categoría

A: refleja el estado de la técnica

O: referido a divulgación no escrita

P: publicado entre la fecha de prioridad y la de presentación de la solicitud

E: documento anterior, pero publicado después de la fecha de presentación de la solicitud

**El presente informe ha sido realizado**

para todas las reivindicaciones

para las reivindicaciones nº:

Fecha de realización del informe  
08.03.2012

Examinador  
B. Pérez García

Página  
1/4

Documentación mínima buscada (sistema de clasificación seguido de los símbolos de clasificación)

G06F, H04L

Bases de datos electrónicas consultadas durante la búsqueda (nombre de la base de datos y, si es posible, términos de búsqueda utilizados)

INVENES, EPODOC, WPI, INSPEC

Fecha de Realización de la Opinión Escrita: 08.03.2012

**Declaración**

<b>Novedad (Art. 6.1 LP 11/1986)</b>	Reivindicaciones	<b>SI</b>
	Reivindicaciones 1-18	<b>NO</b>
<b>Actividad inventiva (Art. 8.1 LP11/1986)</b>	Reivindicaciones	<b>SI</b>
	Reivindicaciones	<b>NO</b>

Se considera que la solicitud cumple con el requisito de aplicación industrial. Este requisito fue evaluado durante la fase de examen formal y técnico de la solicitud (Artículo 31.2 Ley 11/1986).

**Base de la Opinión.-**

La presente opinión se ha realizado sobre la base de la solicitud de patente tal y como se publica.



**1. Documentos considerados.-**

A continuación se relacionan los documentos pertenecientes al estado de la técnica tomados en consideración para la realización de esta opinión.

Documento	Numero Publicación o Identificación	Fecha Publicación
D01	GONZALO ALBERTO ZARZA, "Multipath Fault-tolerant Routing Policies to deal with Dynamic Link Failures in High Speed Interconnection Networks" - Universidad Autónoma de Barcelona, Escuela de Ingeniería - Dpto. de Arquitectura de computadores y Sistemas Operativos. Capítulo 4, páginas 53-66.	15.07.2011

**2. Declaración motivada según los artículos 29.6 y 29.7 del Reglamento de ejecución de la Ley 11/1986, de 20 de marzo, de Patentes sobre la novedad y la actividad inventiva; citas y explicaciones en apoyo de esta declaración**

La solicitud presentada aparece descrita de forma completa y detallada en el documento D01, tesis doctoral del solicitante y que fue divulgado anteriormente a la fecha de presentación de la solicitud.

Este documento D01 contiene todas las características técnicas definidas en las reivindicaciones y no se ha encontrado ninguna diferencia.

Por tanto, la divulgación de D01 anula la novedad de la solicitud presentada, según el Art. 6 de la Ley Española de Patentes.