(19) OFICINA ESPAÑOLA DE PATENTES Y MARCAS

ESPAÑA

(11) Número de publicación: **2 237 346**

(21) Número de solicitud: 200500530

(51) Int. Cl.
H04L 12/56 (2006.01)

(12) PATENTE DE INVENCIÓN CON EXAMEN PREVIO B2

(22) Fecha de presentación: **01.03.2005**

(43) Fecha de publicación de la solicitud: **16.07.2005**

Fecha de la concesión: **14.06.2006**

(45) Fecha de anuncio de la concesión: **16.07.2006**

(45) Fecha de publicación del folleto de la patente: **16.07.2006**

(73) Titular/es: **Universidad de Cantabria**
Pabellón de Gobierno
Avda. de los Catros, s/n
39005 Santander, Cantabria, ES

(72) Inventor/es: **Puente Varona, Valentín;**
Gregorio Monasterio, José Ángel;
Valle Alonso, Fernando y
Beivide Palacio, Ramón

(74) Agente: **No consta**

(54) Título: **Mecanismo de encaminamiento tolerante a fallos altamente escalable.**

(57) Resumen:
Mecanismo de encaminamiento tolerante a fallos altamente escalable denominado S-Immunet, que se caracteriza por ser un mecanismo eficiente para tolerar fallos en redes de interconexión de computadores paralelos y distribuidos. El mecanismo está basado, por un lado, en un método de reencaminar los mensajes cuando se produce un fallo en la red y por otro lado, en una estructura hardware específica del aparato encaminador de mensajes. Las principales diferencias con los mecanismos previos de tolerancia a fallos son que la invención puede ser aplicada a redes muy grandes (miles de nodos) de tipo k-ary n-cube; no produce una sobrecarga de la red en ausencia de fallo; reconfiguración automática y transparente a la aplicación; reparados los componentes se recupera el rendimiento de la red antes del fallo y el nuevo mecanismo además es capaz de tolerar cualquier número de fallos de enlace y cualquier combinación espacial y temporal de fallos.

ES 2 237 346 B2

DESCRIPCIÓN

Mecanismo de encaminamiento tolerante a fallos altamente escalable.

Sector de la técnica

La invención está relacionada con las redes de interconexión de múltiples computadores (multicomputadores ó multiprocesadores) y más concretamente con la tolerancia a fallos que puedan producirse en las redes que los componen.

Estado de la técnica

Lograr elevadas tasas de disponibilidad es objetivo deseable en la mayoría de los computadores. Sin embargo, lograr esta meta en grandes computadores paralelos con tareas de carácter crítico deja de ser una opción para convertirse en una necesidad. Cualquier computador de este tamaño debe disponer de mecanismos eficientes para tolerar la aparición de fallos a todos sus niveles, pues la multiplicidad de componentes incrementa de forma considerable su tiempo medio entre fallos (MTBF).

El contexto en el que se centra esta invención está en el subsistema de interconexión. Los problemas de tolerancia a fallos dentro de la red de interconexión se pueden subdividir entre errores o fallos a nivel de enlace, errores a nivel de encaminador y a nivel de red. Nosotros nos centraremos sobre el nivel de red. A este nivel, el problema fundamental no esta relacionado con las técnicas de implementación, como ocurre en los otros dos niveles, siendo un problema más amplio. Cuando una red de interconexión sufre un fallo en uno de sus recursos (enlaces o nodos) inmediatamente sus propiedades topológicas se ven alteradas. Esto no resulta difícil de solventar desde el punto de vista del mecanismo de encaminamiento empleado para acercar los paquetes a destino. Sin embargo, ese cambio afecta a otros aspectos más sutiles. El más importante es que las estrategias empleadas para evitar cualquier posible anomalía en la red libre de fallo dejan de ser válidas. De acuerdo a esto, después de un fallo en un componente, incluso en redes con centenares de miles de componentes, puede quedar completamente inutilizada. El tipo de anomalía más íntimamente ligada a los aspectos topológicos es el bloqueo entre paquetes o *deadlock*. La producción de un *deadlock* en una parte de la red como consecuencia del fallo y posterior invalidación del mecanismo para su evitación puede hacer inútil el sistema completo.

Este problema no es nuevo y las soluciones propuestas por otros autores pasan desde eliminar aquellos recursos sanos que como consecuencia del fallo pueden inducir el *deadlock* [1], incluir nuevos recursos físicos o virtuales para solventar la situación [2]-[10] o reconfigurar adecuadamente la información topológica de la red empleando un mecanismo de evitación del *deadlock* lo suficientemente versátil. En [11] los autores de esta invención propusieron un nuevo mecanismo en este sentido, llamado Immunet. Sus principales virtudes eran que toleraba cualquier combinación espacio-temporal en los fallos, siempre y cuando la red permaneciera conexa.

Sin embargo Immunet presenta una serie de deficiencias que la hacen casi impracticable cuando se trata de redes con decenas de miles de nodos. La pérdida de rendimiento observada después de un número de fallos reducido tiende a ser elevada. Esto limitaba la aplicabilidad de la idea en los grandes supercomputadores paralelos, que es donde la tolerancia a fallos tiene mas sentido. La invención que aquí se presenta, S- Immunet, es una evolución del mecanismo que permite lograr con total éxito una muy baja pérdida de rendimiento a muy bajo coste.

La invención se apoya en la explotación de las características típicas de las redes de interconexión empleadas en estos sistemas. Así como la aplicabilidad de Immunet era completamente genérica, S- Immunet se especializa en redes *k-ary n-cube*. De esta manera, el nuevo mecanismo, además de tolerar de forma dinámica la aparición de fallos en la red, logra pérdidas de rendimiento por debajo del 1% cuando lo empleamos en una red con 1024 nodos con un solo fallo de enlace (los mejores resultados los obtenía la antigua versión que degeneraba hasta un 15% con el mismo fallo). Además, la invención permite la recuperación dinámica del rendimiento inicial de la red de interconexión cuando se lleva a cabo la reparación. Es decir, una vez que todos los fallos han ido siendo eliminados, el mecanismo es capaz de detectar la nueva situación y sin detener la ejecución de las aplicaciones que estén ejecutándose en la máquina paralela, puede llevar a cabo la mencionada recuperación.

[1] NR **Adiga**, GS **Almasi**, Y **Aridor**, M **Bae**, **Rajkishore Barik**, *et al*., "An Overview of the BlueGene/L Supercomputer", *Supercomputing* 2002.

[2] R.V. **Boppana** and S. **Chalasani**. "Fault-tolerant wormhole routing algorithms for mesh networks". IERF, *Trans. on Computers*, vol. 44, no.7, pp. 848-864, July 1995.

[3] J. **Bruck**, R. **Cypher**, and C. **Ho**. "Fault-tolerant meshes with small degree". *SIAM Journal of Computing*, vol. 26, no. 6, pp. 1764-1784, December 1997.

[4] R. **Casado**, A. **Bermudez**, F. J. **Quiles**, J. L. **Sanches**, and J. **Duato**, "Performance evaluation of dynamic reconfiguration in high-speed local area networks". *International Symposium on High-Performance Computer Architecture* (HPCA), January 2000.

[5] S. **Chalasani** and R V. **Boppana**, "Communication in Multicomputers with Nonconvex Faulty". *IEEE Trans. on Computers*, vol. 46, no.5, pp. 616-622, 1997.

[6] M.S. **Chen** and K.G. **Shin**, "Adaptive Fault-Tolerant Routing in Hypercube Multicomputers", *IEEE Trans. on Computers*, vol. 39, no.12, pp. 1406-1416, December 1990.

[7] D. **Avresky**, "Embedding and Reconfiguration of Spanning Trees in Faulty Hypercubes", *IEEE Transactions on Parallel and Distributed Systems* vol. 10 no.3, pp. 211-222, March 1999.

[8] M. **Galles**, "Spider: a high-speed network interconnect", *IEEE Micro*, vol. 17, no.1, pp. 34-39, Jan-Feb. 1997.

[9] P.T. **Gaughan** and S. **Yalamanchili**, "A Family of Fault-Tolerant Routing Protocols for Direct Multiprocessor Networks", *IEEE Trans. on Parallel and Distributed Systems*, vol. 6, no.5, pp. 482-497, May 1995.

[10] **Rpang**, T. **Pinkston**, "The Double Scheme: *deadlock*-free reconfiguration of Cut-through Networks", *International Conference of Parallel Processing* (ICPP) August 2000.

[11] V. **Puente**, J.A. **Gregorio**, F. **Vallejo** and R. **Beivide**, "Immunet: A Cheap and Robust Fault-Tolerant Packet Routing Mechanism". The 31st Annual International Symposium on Computer Architecture (ISCA2004), pp. 198-209, Munchen, Germany, June 2004.

[12] V. **Puente**, C. **Izu**, R. **Beivide**, J.A. **Gregorio**, F. **Vallejo** and J.M. **Prellezo**, "The Adaptive Bubble Router", *Journal of Parallel and Distributed Computing*. vol 61, no. 9, September 2001.

Descripción de la invención

Breve descripción de la invención

La invención, denominada S-Immunet, es un mecanismo eficiente para tolerar fallos en redes de interconexión de computadores paralelos y distribuidos. El mecanismo está basado, por una parte, en un método de reencaminar los mensajes cuando se produce un fallo en la red y por otra parte, en una estructura hardware específica del aparato encaminador de mensajes. Las principales diferencias con los mecanismos previos de tolerancia a fallos son que la invención puede ser aplicada a redes muy grandes (miles de nodos) de tipo *k-ary n-cube* y además la reparación de los componentes es tenida en cuenta y es recuperado el rendimiento de la red antes del fallo. El mecanismo es capaz de tolerar cualquier número de fallos de enlace y cualquier combinación espacial y temporal de fallos; después de cualquier fallo, permite llevar a cabo una reconfiguración automática y además es transparente a la aplicación que esta ejecutándose en el sistema multiprocesador. El mecanismo además no produce una sobrecarga de la red en ausencia de fallos.

Breve descripción del contenido de las figuras

Figura 1. Ejemplo de una rotura de un enlace en una red toroidal (3-ary 2-cubo).

Figura 2. Nodo padre y nodos hijos del nodo 5 tras la comunicación del fallo por el nodo 4.

Figura 3. Camino o "tour" a lo largo del árbol y que formará el Anillo o Camino Seguro.

Figura 4. Ejemplos de uso de los canales de escape de primer y segundo orden sobre una red toroidal de dos dimensiones.

Figura 5. Estructura del encaminador de mensajes que puede ser empleado en la invención.

Descripción detallada de la invención

Ante la aparición en la red de interconexión de uno o varios fallos de enlace o de nodo, S-Immunet permite la obtención automática de un "Anillo" que abarca todos los nodos de la red y aplicando sobre él la técnica conocida como *Buffer Flow Control, BFC* [12], dicho anillo pasa a ser un "Anillo o Camino Seguro" para que cualquier paquete alcance su destino en presencia de fallos. El proceso de obtención del Anillo se lleva a cabo sin tirar ningún paquete en tránsito.

El Anillo Seguro se obtiene a partir de la generación de un árbol (*spanning tree*), que siempre existe embebido en cualquier topología, y cuya raíz es el nodo que detectó el fallo. Llevando a cabo un "recorrido" (*tour*) a lo largo del mencionado árbol y aplicando sobre él la técnica BFC mencionada se garantiza la existencia de un Anillo Seguro para el intercambio de paquetes entre todos los nodos supervivientes de la red.

La generación del árbol se lleva a cabo mediante un proceso trivial de difusión (*broadcast*) a sus inmediatos vecinos de la situación de fallo. Con la respuesta, o no, de sus vecinos cada nodo crea una tabla provisional de rutas que permite la actualización de las tablas de ruta necesarias para soportar el cambio topológico que ha forzado el fallo.

Por ejemplo, se supone que en la red de la figura 1 se ha roto el enlace entre los nodos #4 y #1 y que el nodo #5 ha recibido una notificación de fallo del nodo #4 (y por consiguiente se ha convertido en su nodo padre) y sus inmediatos vecinos #3 y #8 han respondido a su reemisión (y se han convertido en sus hijos ordenados local y arbitrariamente). A partir de ese instante, el nodo #5 únicamente necesita recordar cuál es puerto correspondiente al "padre" en el árbol y cuál el de cada uno de los "hijos", como se muestra en la figura 2.

Es obvio que el recorrido a lo largo del árbol (figura 3) se genera sin más que cada nodo reenvíe los paquetes aplicando las reglas de la tabla 1. Además, cualquier paquete inyectado en este camino alcanzará su destino (aunque en el peor caso tenga que recorrer todo el anillo).

TABLA 1

Reglas para reenviar mensajes tras la aparición de fallos

Paquetes provenientes de	Deben ser enviados a
hijo i	hijo i+1
padre	Hijo 1 (primer hijo)
último hijo	padre (1)
puerto de inyección	padre (1)
canal adaptativo	padre (1)

(1) Una excepción es la raíz del árbol (no tiene padre). Los paquetes deben ser enviados al primer hijo.

Una vez que cada nodo dispone de la información necesaria para enviar paquetes a cualquier otro nodo de la red, se actualizan las tablas de rutas definitivas resultantes de la nueva situación de fallo y se lleva a cabo el envío, tanto de los nuevos paquetes que se generen por las aplicaciones, como los que se encontraban en las colas de tránsito (*buffers*) esperando la reconfiguración.

La capacidad de soportar cualquier combinación espacio-temporal de fallos se consigue mediante del empleo de un identificador único de fallo, EPL (*Emergency Priority Level*), generado con cada nuevo fallo y dependiente del propio identificador (ID) del nodo que lo detecta y del número de fallos previos que haya soportado dicho nodo hasta ese instante. Por ejemplo, en una red de N nodos, si un encaminador cuyo identificador es *ID=x* detecta un nuevo fallo, generará un EPL igual a *tN+x*, siendo t el número de reconfiguraciones previas experimentadas por el nodo. Así, el valor EPL generado tras la detección de un fallo siempre es distinto y por tanto se genera un solo árbol.

Aunque la metodología es completamente general, cuando disponemos de una red compuesta por miles de nodos, el fallo en un solo componente, o de unos pocos, (son las situaciones más probables) puede producir una degradación en el rendimiento muy acusada, como consecuencia de la elevada longitud del Anillo Seguro resultante. Además, si el componente se repara, el mecanismo no es capaz de recuperar el funcionamiento original libre de fallos. Mediante S-Immunet se resuelven estos dos problemas para redes de cualquier tamaño con una estructura del tipo *k-ary n-cube*.

La invención asocia a la red de interconexión tres redes virtuales diferentes. La primera red es completamente adaptativa y con encaminamiento mínimo. En la segunda red virtual, denominada "red virtual de escape de primer orden", se supondrá que la red se encuentra libre de fallo, aunque esto no sea cierto. Por lo tanto, los paquetes en esta red se encaminarán en orden de dimensión (X, Y, Z,..). En una situación de fallo es imposible lograr que todo el tráfico alcance su destino porque el camino en alguna dimensión estará cortado. Cuando un paquete que se encuentre empleando esta red alcance un recurso en fallo que impida seguir aplicando encaminamiento en orden de dimensión, estará habilitado para emplear la tercera red virtual, denominada "red virtual de escape de segundo orden". En esta red aplicaremos el Anillo Seguro descrito anteriormente como mecanismo de evitación de bloqueo. Puesto que sabemos que la red virtual de escape de segundo orden es libre de bloqueo, independientemente de la configuración de los fallos, toda la red será libre de fallo.

Bajo las condiciones previamente expuestas, se dispone de dos redes virtuales con posibilidad de bloqueo y una tercera libre de bloqueo bajo cualquier circunstancia. Los tipos de tráfico que pueden emplear cada una de las tres redes virtuales y que el encaminador se encargará de determinar son los siguientes:

- La red virtual adaptativa está disponible para cualquiera que sea el tipo de tráfico y su recorrido previo.

- La red virtual de escape de primer orden está disponible para todos aquellos paquetes que no han encontrado un fallo en su camino a destino, provengan de inyección, de la red virtual adaptativa o estén avanzando por ella misma.

- La red de escape de segundo orden podrá ser empleada exclusivamente por los paquetes que habiendo circulado previamente por la red de escape de primer orden encontraron un recurso en fallo en su camino.

En la Figura 4 se muestran varios ejemplos de rutas para diferentes paquetes. Esta política resulta extremadamente simple de aplicar e implica únicamente un bit de sobrecarga por paquete. Este tipo de política de asignación permite limitar la aplicación del Anillo Seguro solamente a aquel tráfico que se encuentre frontalmente con un recurso en fallo. Cuando el número de recursos en fallo es relativamente bajo, la mayor parte del tráfico fluirá como si la red estuviera libre de fallos. Así, la pérdida de rendimiento debido al Anillo Seguro se ve eliminada en su mayor parte, pues el número de paquetes que alcanzan los fallos es reducido, siendo, en consecuencia, el nivel de ocupación de la red de escape de segundo orden muy reducido y logrando que el efecto de la longitud del Anillo Seguro sea prácticamente despreciable.

En una situación libre de fallo no es necesario aplicar ninguna clase de mecanismo adicional para evitar el bloqueo originalmente propuesto en [12]. Aplicando encaminamiento en orden de dimensión en una de las tres redes virtuales de que disponemos es suficiente (la red de escape de primer orden). La red de escape de segundo orden es empleada como un segundo canal adaptativo adicional.

Por último, la invención permite la recuperación completa del rendimiento tras la reparación completa de la red. Así, si suponemos que se desencadena un proceso de reconfiguración en la red después de cualquier clase de reconfiguración, resulta sencillo determinar si todos los nodos de la red tienen todos sus enlaces funcionando. Para una red *k-ary n-cube*, cualquiera de los encaminadores libres de fallos debe contar con 2*n enlaces bi-direccionales operativos. Globalmente, la red ha de disponer de $2*n*k^2$ enlaces en funcionamiento. Por tanto, conocida la cantidad, denominada w, de enlaces operativos, la red estará libre de fallo si y solo si $w= 2*n*k^2$.

W puede conocerse fácilmente durante la actualización de las tablas en el proceso de reconfiguración. Cada nodo envía un paquete de actualización hacia su nodo padre con el identificador del nodo emisor, el nivel de prioridad EPL y el número de enlaces disponibles en ese nodo particular. Así, el nodo raíz determina el número de enlaces correctos w. Si $w= 2*n*k^2$, el nodo raíz envía un paquete de control específico junto con el nivel EPL de la reconfiguración a sus inmediatos hijos y estos lo retransmitían a los suyos. Cada nodo que reciba este paquete desechará el uso del Anillo Seguro y a partir de ese momento empleará los recursos previamente dedicados a la red de escape de segundo orden (Anillo Seguro) a encaminar los paquetes de forma adaptativa. Repitiendo este proceso en todos los nodos de la red se dispondrá otra vez de dos canales adaptativos y se restaurará el rendimiento original de la red.

Ejemplo de realización de la invención

Para mejor comprensión del invento, a continuación se describe un ejemplo de aplicación. Suponer la existencia de una red de interconexión de computadores del tipo *k-ary n-cube* como la de la figura 1 pero que puede estar compuesta por miles de nodos o elementos de cómputo. Cada uno de los encaminadores que forman la red tiene una estructura como la de la figura 5. Tres canales virtuales (sin fallos, dos serán adaptativos y uno de escape), un elemento de conmutación o *crossbar* que permite la interconexión de las entradas y las salidas, así como con los elementos de cómputo local. La selección de los canales de entrada que tendrán acceso a los de salida la lleva a cabo un Árbitro. La Unidad de Encaminamiento (RU) emplearía dos tablas. Una pequeña, mediante la que se indica los puertos que constituyen el camino de escape en cada momento (Tabla Escape) y la otra tabla (Tabla Adaptativa) tendrá una entrada por cada nodo de la red y será empleada para encaminar los mensajes dependiendo del nodo destino.

Si un nodo desea enviar un mensaje a otro que se encuentra a una distancia (Δx, Δy) 5 empleará prioritariamente cualquiera de los dos canales virtuales adaptativos (Tabla Adaptativa). En caso de que no pueda hacerlo, intentará entrar en el canal de escape que siempre seguirá una ruta en orden de dimensión (DOR). Es decir, recorrerá primero el anillo de las Xs hasta agotar esta componente de su dirección (la distancia Δx sea *0*), luego continuaría por el anillo correspondiente al eje de las Ys hasta su destino (la distancia Δy sea *0*). No obstante, en su avance volverá a los canales adaptativos siempre que pueda.

Si algún encaminador detecta un fallo en alguno de sus enlaces con algún vecino, entra en un estado de fallo, detiene la comunicación con todos los demás nodos, se convierte en el nodo raíz y, empleando las líneas hardware, se lo comunica a sus vecinos. Estos lo interpretarán como el comienzo de un proceso de reconfiguración y tras responder al requerimiento se comunicarán a su vez con sus vecinos y así hasta alcanzar el último nodo de la red. Cada uno de los encaminadores apuntará en su pequeña tabla (Tabla Escape) cuál fue el puerto por el que recibió el requerimiento de paso a modo de fallo (nodo padre a partir de entonces) y quienes de sus vecinos le han ido respondieron al suyo (nodos hijos en el orden de la respuesta). Tras un corto periodo de tiempo, todos los encaminadores tienen su pequeña tabla, de forma que juntos forman un árbol (como el de la figura 3) cuya raíz es el nodo que detectó el fallo. Por tanto, cualquier encaminador que desee enviar un paquete a cualquier otro, dispone de un camino para hacerlo. Si además para introducir un nuevo paquete en este camino se añade la condición de mantener un hueco adicional, este camino es "seguro", es decir, no podrán ocurrir bloqueos mutuos entre paquetes. A partir de este instante, en un sencillo proceso de difusión, vuelven a actualizarse las tablas adaptativas para tener en cuenta el cambio topológico producido por el enlace roto.

A partir de ese momento, si un nodo desea enviar un mensaje a otro que se encuentra a una distancia (Δx, Δy), igual que cuando se encuentra libre de fallos empleará prioritariamente el único canal adaptativo de que ahora dispone

el encaminador (Tabla Adaptativa actualizada). En caso de que no pueda hacerlo intentará entrar en el canal de escape de primer orden, que siempre seguirá una ruta DOR como si no hubiese ningún enlace roto. Si en el camino DOR el paquete se encuentra con un enlace roto y por lo tanto no puede avanzar, en este instante pasará al camino de escape de segundo orden o Anillo Seguro que se acaba de crear con la reconfiguración a costa de uno de los canales adaptativos.

Así, con un número pequeño de enlaces rotos, la degradación del rendimiento será muy pequeña porque lo más probable es que el paquete no se encuentre con el enlace en mal estado, sobremanera en las redes con un gran número de nodos que es además donde más degradación se produce cuando no se aplica el mecanismo en el que se basa la invención que se esta describiendo.

Si el enlace roto es reparado, se producirá un proceso de reconfiguración y durante la actualización de las tablas se lleva a cabo también la comunicación a los nodos padre de los enlaces en correcto estado, hasta llegar al nodo raíz que puede determinar si todos los enlaces de la red están libres de fallos. Si es así, emitirá un paquete específico para indicar a todos los nodos de la red que a partir de ese momento pueden volver a usar el canal de escape de segundo orden como un segundo canal adaptativo. Es decir, se restaura el rendimiento original de la red.

Si durante éste o en cualquier otro instante ocurriese un nuevo fallo, la existencia de un único identificador EPL garantiza una nueva y correcta actualización de las tablas de encaminamiento.

REIVINDICACIONES

1. Mecanismo de encaminamiento tolerante a fallos altamente escalable que se **caracteriza** por ser de aplicación, sobre todo, en redes directas regulares del tipo k-ary n-cube de tamaños elevados y en el que se distinguen un dispositivo encaminador de mensajes en cada nodo de la red con encaminamiento basado en tablas; un método para reencaminar mensajes basado en la creación automática de un camino de escape sobre un anillo; y un método para recuperar el rendimiento de la red tras la reparación de los fallos.

2. Mecanismo de encaminamiento tolerante a fallos altamente escalable, según reivindicación 1ª, **caracterizado** porque el dispositivo encaminador mensajes de cada nodo consta de, al menos, tres canales virtuales por cada canal físico. En ausencia de fallos dos de los canales se emplean como adaptativos y el tercero como determinista.

3. Mecanismo de encaminamiento tolerante a fallos altamente escalable, según reivindicación 1ª, **caracterizado** porque el dispositivo encaminador mensajes de cada nodo ante la presencia de uno o varios fallos, automáticamente reasigna uno de los canales adaptativos para formar parte de un gran anillo que una a todos los nodos de la red y que pueda ser empleado como un canal de escape de segundo orden.

4. Mecanismo de encaminamiento tolerante a fallos altamente escalable, según reivindicación 1ª, **caracterizado** porque el método para reencaminar mensajes, basado en la creación automática de un camino de escape sobre un anillo, se lleva a cabo creando un árbol (spanning tree) compuesto por todos los nodos de la red y cuya raíz es aquel que detectó el fallo.

5. Mecanismo de encaminamiento tolerante a fallos altamente escalable, según reivindicación 1ª, **caracterizado** porque el método para recuperar el rendimiento de la red tras la reparación de los fallos esta basado en la comunicación automática al nodo raíz, por parte de cualquier nodo de la red, de cualquier reparación detectada.

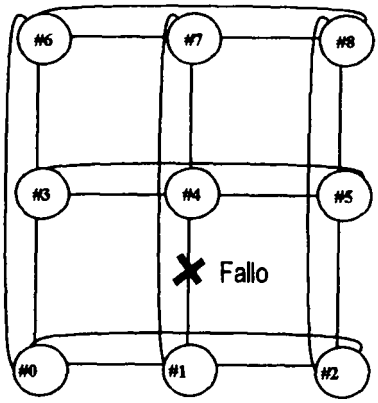

Figura 1.

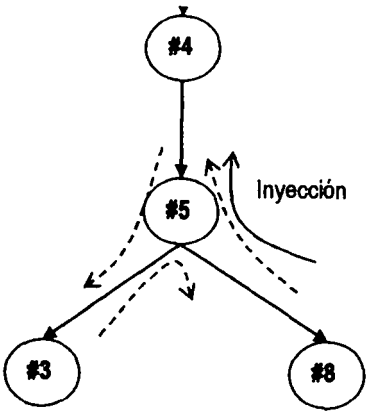

Figura 2

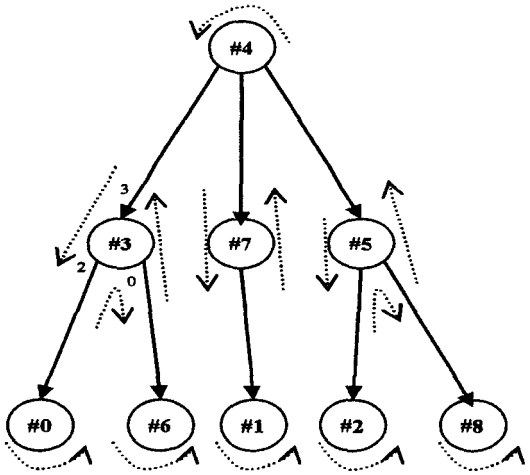

Figura 3.

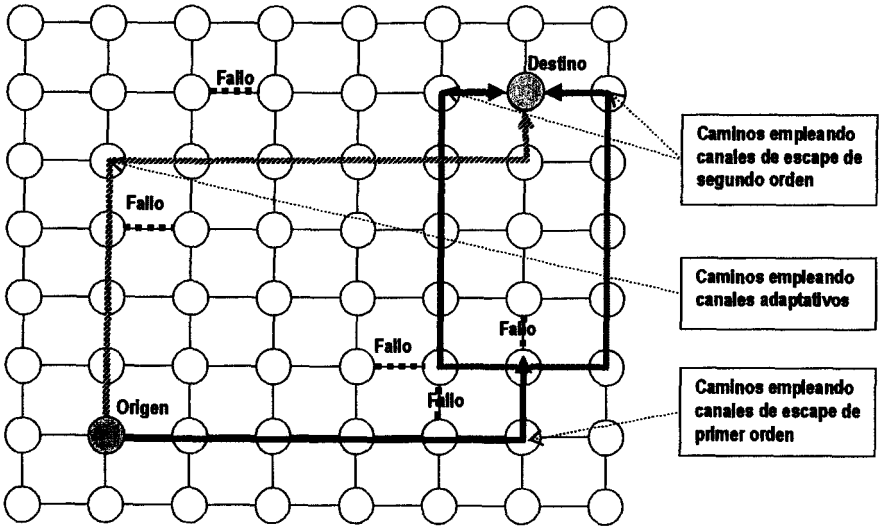

Figura 4.

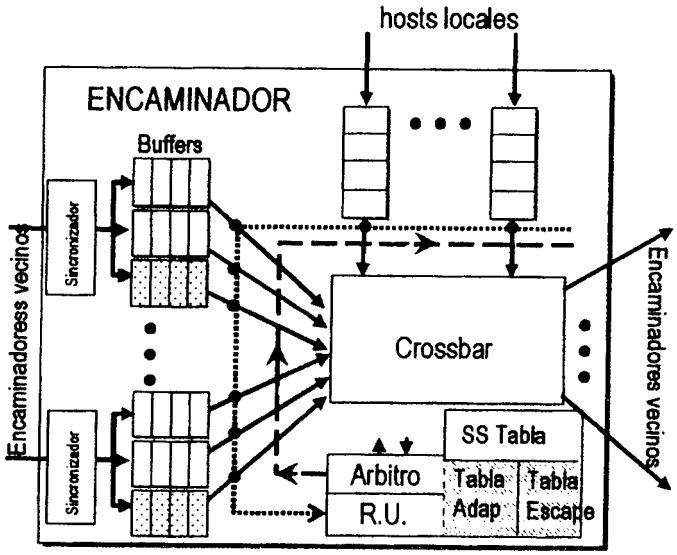
hosts locales
ENCAMINADOR
Buffers
Sincronizador
Sincronizador
Encaminadores vecinos
Crossbar
Encaminadores vecinos
SS Tabla
Arbitro
R.U.
Tabla Adap
Tabla Escape

Figura 5.

OFICINA ESPAÑOLA DE PATENTES Y MARCAS

ESPAÑA

(11) ES 2 237 346

(21) Nº de solicitud: 200500530

(22) Fecha de presentación de la solicitud: **01.03.2005**

(32) Fecha de prioridad:

INFORME SOBRE EL ESTADO DE LA TÉCNICA

(51) **Int. Cl.[7]:** H04L 12/56

DOCUMENTOS RELEVANTES

Categoría	Documentos citados	Reivindicaciones afectadas
A	Puente V. Et al. "Immunet: a cheap and robust fault-tolerant packet routing mechanism". En: Proceedings. 31st Annual International Symposium on Computer Architecture. 2004. IEEE Comput. Soc. Los Alamitos, CA, USA. Páginas 198-209. ISBN 0-7695-2143-6.	1
A	US 2004042418 A1 (HAMADA TAKEO; CZEZOWSKI PETER J; SU CHING-FONG) 04.03.2004, todo el documento.	
A	US 2004078625 A1 (RAMPURIA ASHOKE; DHARA PRADIP) 22.04.2004, todo el documento.	

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	Examinador	Página
29.06.2005	M. Alvarez Moreno	1/1