



19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA

11 Número de publicación: **2 359 646**

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

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Número de solicitud europea: **06818480 .3**

96 Fecha de presentación : **10.11.2006**

97 Número de publicación de la solicitud: **2090034**

97 Fecha de publicación de la solicitud: **19.08.2009**

54 Título: **Nodo de borde para un dominio de red.**

45 Fecha de publicación de la mención BOPI:
25.05.2011

45 Fecha de la publicación del folleto de la patente:
25.05.2011

73 Titular/es: **Telefonaktiebolaget LM Ericsson (publ)**
164 83 Stockholm, SE

72 Inventor/es: **Báder, Attila;**
Takacs, Attila y
Csaszar, Andras

74 Agente: **Elzaburu Márquez, Alberto**

ES 2 359 646 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

Nodo de borde para un dominio de red

CAMPO DEL INVENTO

5 La presente aplicación se refiere a un nodo de borde de un dominio de red que comprende uno o más nodos de entrada, una pluralidad de encaminadores interiores y uno o más nodos de salida, a un encaminador interior de tal dominio de red, y a los correspondientes métodos de control de un nodo de borde y un encaminador interior.

ANTECEDENTES DEL INVENTO

10 En el campo de las comunicaciones se usa ampliamente el transporte por conmutación de paquetes. Recientemente, las soluciones de transporte con base en IP han recibido más atención. IP significa Protocolo de Internet. El interés en IP es al menos parcialmente debido a la flexibilidad y al amplio despliegue de las tecnologías IP. Sin embargo, las redes de comunicación tales como las redes de telefonía tienen características diferentes cuando se comparan con las redes IP tradicionales debido a que las redes de comunicación normalmente desean ser simples, no caras y ser capaces de proporcionar una Calidad de Servicio (QoS) definida. Debido a las exigencias de la Calidad de Servicio es importante el control de la congestión del tráfico en las redes de comunicación.

15 Uno de los conceptos para proporcionar QoS en redes IP es la reserva de recursos en encaminadores por señalización a lo largo del trayecto de los datos. Como ejemplo, la organización de normalización del Grupo de Trabajo de Ingeniería de Internet (IETF) ha especificado un protocolo de señalización llamado RSVP (Protocolo de Reserva de Recursos) para realizar reservas de recursos en encaminadores IP y para proporcionar servicios integrados para tráfico en tiempo real y en tiempo no real en Internet (véase por ejemplo R. Braden y otros: Protocolo de Reserva de Recursos (RSVP) – versión 1 “Especificación Funcional”, RFC2205, Septiembre 1997; R. Braden y otros: “Servicios Integrados en la Estructura de Internet; una visión general”, RFC 1633, 1994; J. Roclawski: “El Uso de RSVP con Servicios Integrados IETF”, RFC2210, Septiembre 1997). En esta tecnología los mensajes de señalización RSVP reservan recursos en cada encaminador a lo largo del trayecto de los datos antes de enviar un flujo de tráfico en tiempo real. Los flujos en tiempo real se admiten en la red si los recursos han sido reservados con éxito en cada encaminador a lo largo del trayecto de los datos.

20 Este método requiere almacenar estados de reserva por flujo en cada encaminador a lo largo del trayecto de los datos. Los estados de reserva son estados transitorios, lo que significa que tienen que ser renovados mediante el envío periódico de mensajes de renovación. Si un estado reservado no es renovado, el estado y los correspondientes recursos se eliminan después de un periodo de fin de temporización. Las reservas pueden también ser eliminadas mediante mensajes explícitos de anulación. Con la ayuda de los estados por flujo los mensajes RSVP siguen siempre la corriente ascendente y descendente del trayecto de los datos. Por lo tanto, es capaz de interfuncionar con protocolos de encaminamiento normales. Si el tráfico es reencaminado, por ejemplo debido a la sobrecarga de un encaminador, los mensajes de renovación realizan reservas en el nuevo trayecto de los datos. El almacenamiento de estados por flujo en cada encaminador también permite la posibilidad de reparación del trayecto local. Después del reencaminamiento, si no hay recursos suficientes para soportar todos los flujos deseados para el nuevo trayecto, algunos flujos pueden ser terminados.

30 El almacenamiento y mantenimiento de los estados por flujo en cada encaminador puede ser un problema en redes grandes, en las que el número de flujos, y por tanto el número de estados de reserva puede ser alto. En respuesta a este problema de escalabilidad de RSVP, el IETF especificó un método de agregación que permite realizar reservas para flujos agregados (véase RFC3175, Septiembre 2001). Los estados de reserva agregados no son necesariamente creados, modificados o renovados para cada solicitud de flujo.

35 El método RSVP de agregación es capaz de hacer una decisión de control de admisión solamente en un nivel de agregación, lo que significa que todos los flujos que se tratan conjuntamente como un agregado se terminan si no hay recursos suficientes en un nuevo trayecto después del reencaminamiento.

40 Como medio para proporcionar QoS en redes a gran escala, se propuso una estructura denominada Servicios Diferenciados (DiffServ), véase RFC 2475, 1998 en la estructura DiffServ la escalabilidad se consigue ofreciendo servicios en una base agregada más que por flujo, y disponiendo el sistema de forma que se conserve lo máximo posible la información de estado por flujo en los nudos de borde de la red, y solamente la información de estado agregado se conserve en los nodos interiores. La diferenciación del servicio se consigue usando un campo específico en el encabezamiento IP, el denominado campo de Servicios Diferenciados (DS). Los paquetes se clasifican en grupos Per-Hop Behaviour (PHB) en los nodos de borde DiffServ. Los paquetes son tratados en encaminadores DiffServ de acuerdo con el PHB indicado en el campo DS del encabezamiento de mensaje. La estructura DiffServ no especifica modo alguno para dispositivos de fuera del dominio para reservar dinámicamente recursos o indicaciones de recepción de disponibilidad de recursos de red. En la práctica, los proveedores de servicio dependen de los Acuerdos de Nivel de Servicio (SLAs) en un periodo de tiempo de abono que estáticamente definen los parámetros del tráfico que será aceptado de un cliente.

Las redes DiffServ proporcionan QoS para tráfico en tiempo real de una forma escalable. La DiffServ es una estructura estática en la que el tráfico está limitado en los bordes del dominio, y los nodos interiores tienen que ser diseñados con unas dimensiones apropiadas. No existen mecanismos dinámicos para tratar un sobrevolumen debido al reencaminamiento dentro del dominio, de forma que los nodos y enlaces interiores necesitan ser sobredimensionados con respecto a la carga media con el fin de proporcionar una garantía de QoS al menos hasta un cierto nivel.

El grupo de trabajo de Sigüientes Pasos en Señalización (NSIS) del IETF está trabajando en un protocolo para abordar las exigencias de nueva señalización en redes IP, véase RFC 3726, Abril 2004. El protocolo de aplicación de señalización de QoS del NSIS es similar al RSVP aunque tiene algunas características adicionales tales como el que soporta modelos de QoS diferentes. Uno de los modelos QoS es la Gestión de Recursos en DiffServ (RMD). El RMD define métodos de control de admisión escalables para redes DiffServ. También incluye una función de control de congestión que es capaz de terminar varios flujos en una situación de congestión con el fin de mantener un QoS requerido para el resto de los flujos.

Una característica importante de las redes IP y de algunas otras redes conmutadas por paquetes en comparación con otras tecnologías usadas en comunicaciones, tales como ATM o SDH es que los protocolos de encaminamiento IP se adaptan automáticamente a los cambios de topología, por ejemplo, si un enlace o nodo falla, el tráfico es automáticamente reencaminado. Sin embargo, puede darse el problema de que el nuevo trayecto no sea capaz de soportar todos los flujos reencaminados, por ejemplo debido a una falta de anchura de banda. En tal situación es conveniente una gestión de la congestión.

La RMD define funciones para notificar a nodos de borde de un dominio DiffServ sobre la congestión que ocurre en los encaminadores interiores. Estas notificaciones están basadas en el "marcado de nuevo" de los paquetes de datos en el campo DS en los encaminadores interiores en proporción a la sobrecarga. Un nodo de borde puede entonces medir el número de paquetes marcados y enviar un mensaje de terminación para algunos flujos, si es necesario para reducir la carga de tráfico.

Un algoritmo para reaccionar ante la congestión está, por ejemplo, descrito en Andràs Csàzàr, Attila Takàcs, Robert Szabo, Vlora Rexhept y Georgios Karagiannis "Tratamiento de la Congestión Grave con Gestión de Recursos en DiffServ a Petición", en las actas de Interfuncionamiento 2002. La Segunda Conferencia Internacional de Redes de Ordenadores IFIP-TC6, volumen 2345 de INCS, páginas 443-464, 2002, Pisa, Italia, o en Lars Westberg, Andràs Csàzàr, Georgios Karagiannis, Adam Marquetant, David Parlain, Octavian Pop, Vlora Rexhept, Robert Szabo y Attila Takàcs "Gestión de Recursos en DiffServ (RMD): Una Visión de Conjunto de Funcionalidad y de Comportamiento de Funcionamiento" en actas de PfSHN 2002, 7º Seminario sobre Protocolos para Redes de Alta Velocidad, volumen 2334 de INCS, páginas 17-34, 2002, Berlín, Alemania.

Un problema surgido en tales algoritmos puede ser el de la sobre-reacción frente a una situación de congestión por la terminación de más flujos que los necesarios para disminuir la congestión. Esto también se llama subimpulso, porque la carga real cae por debajo de un nivel deseado debido a la sobre-reacción.

El documento WO 2206/052174 describe las formas de tratamiento del problema de la sobre-reacción. En el sistema descrito los encaminadores interiores de un dominio marcan los paquetes en proporción a la sobrecarga, y la sobrecarga señalada se almacena. En otras palabras, los encaminadores o nodos interiores tienen una forma de "memoria" de la sobrecarga medida. Esta información memorizada es después tenida en cuenta al marcar nuevos paquetes en el futuro. Además, los encaminadores interiores marcan las unidades de datos enviadas proporcionalmente al grado de sobrecarga, y los nodos de salida cuentan las unidades de datos marcados con el fin de medir el grado de sobrecarga.

El documento US 2004/0008628 A1 describe además una red de comunicación que comprende nodos extremos y una pluralidad de nodos interiores. Además, se describe un nodo de control de flujo que determina si hay una situación de congestión. Además, se describe un nodo interior. No obstante, este método no comprende la terminación de uno o más flujos para de este modo reducir una carga de tráfico y esperar un periodo de tiempo predeterminado después de que se haya determinado si todavía llegan unidades de datos marcadas de congestión. Además, el nodo interior no está dispuesto para detectar si una carga excede de valor de umbral de carga que es más alto que un valor de umbral de carga que inicia la generación de unidades de datos marcadas de congestión.

OBJETO DEL INVENTO

El objeto del presente invento es proporcionar un modo mejorado de tratar la congestión en los dominios de la red que comprenden uno o más nudos de entrada, especialmente con objeto de simplificar los encaminadores interiores. Preferiblemente, también se desea conseguir una buena respuesta a la congestión sin provocar subimpulso, es decir sin terminar flujos más allá de lo necesario.

RESUMEN DEL INVENTO

El objeto se resuelve mediante el tema objeto de las reivindicaciones independientes. Realizaciones ventajosas se describen las reivindicaciones dependientes.

De acuerdo con una realización básica del presente invento se dispone un nodo de borde de un dominio de red que comprende un nodo de entrada para recibir y encaminar unidades de datos que pertenecen a una pluralidad de flujos al interior del dominio de red, una pluralidad de encaminadores interiores para encaminar las unidades de datos a través del dominio de red, y al menos un nodo de salida para encaminar unidades de datos fuera del dominio de red. Al menos uno de los encaminadores interiores está dispuesto para detectar si está sometido a una situación de congestión y para marcar de congestión las unidades de datos encaminadas si está sometido a la situación de congestión. Obsérvese que no es necesario que cada encaminador interior tenga esta funcionalidad, pero es preferible que cada encaminador interior pueda actuar de este modo. El marcado de congestión de unidades de datos encaminadas puede, por ejemplo, conseguirse fijando un cierto valor en un campo previamente determinado del encabezamiento de una unidad de datos encaminados. El nodo de borde del invento está dispuesto para actuar como un nodo de salida y comprende un elemento de control para realizar una función de tratamiento de la congestión que comprende una parte para detectar la presencia de unidades de datos marcadas de congestión y una parte para reaccionar frente a la detección de la presencia de unidades de datos marcadas de congestión llamando a un proceso de control de congestión. El proceso de control de congestión comprende enviar al nodo de entrada una instrucción para terminar uno o más flujos para de este modo reducir una carga de tráfico, esperar un determinado periodo de tiempo y después determinar si las unidades de datos marcadas de congestión están todavía llegando al nodo de borde, y si las unidades de datos marcadas de congestión están todavía llegando se repetirán los pasos de envío y de espera.

Esperando un periodo de tiempo predeterminado antes de posiblemente enviar una instrucción adicional para reducir la carga de tráfico, se puede tener en cuenta la demora entre el envío de esa instrucción y la instrucción real de la misma que está siendo aplicada y su efecto. En otras palabras, el concepto antes descrito reconoce que hay una cierta demora entre el envío de una instrucción de reducción de carga por el nodo de salida, y el punto en el tiempo en el que el nodo de salida es capaz de medir una reacción a esa instrucción. La demora comprende el tiempo que emplea la instrucción para alcanzar el nodo de salida, el tiempo que el nodo de entrada emplea para seguir la instrucción, y el tiempo que emplea para que las unidades de datos inyectadas en el dominio por el nodo de entrada después de seguir la instrucción nuevamente alcancen el nodo de salida. No enviando instrucciones de reducción adicionales dentro del periodo de demora se puede evitar de una forma muy simple un subimpulso debido a reaccionar excesivamente ante la recepción de las unidades marcadas de congestión. Es decir, no es necesario adaptar específicamente los encaminadores interiores del dominio, es decir éstos pueden ser mantenidos simples.

De acuerdo con una realización preferida, uno o más encaminadores interiores están dispuestos de tal forma que puede detectarse si una carga excede de un primer umbral de carga o de un segundo umbral de carga que es mayor que el primer umbral de carga, y para generar unidades de datos marcadas de congestión que comprenden un primer marcador si la carga de tráfico excede del primer umbral de carga, y un segundo marcador que es diferente del primer marcador si la carga de tráfico excede del segundo umbral de carga. A su vez, el nodo de salida puede entonces estar dispuesto de tal modo que su proceso de control de congestión comprende la fijación de un primer parámetro de carga en un primer valor de reducción de carga en respuesta a la detección del primer marcador en las unidades de datos recibidas, y en un segundo valor de reducción de carga en respuesta a la detección del segundo marcador en las unidades de datos recibidas, en donde el segundo valor de reducción de carga corresponde a una reducción de carga mayor que la del primer valor de reducción de carga.

La ventaja de este concepto es que se puede conseguir un control de congestión efectivo con unos medios muy simples. Es decir, los encaminadores interiores solamente necesitan ser capaces de distinguir dos diferentes umbrales de carga diferentes en el tráfico de entrada, y en consecuencia marcar las unidades de datos encaminados. No es necesario que las unidades de datos estén marcadas en proporción con la sobrecarga o congestión, ya que es posible que el marcado se produzca de un modo escalonado, es decir una vez que se hayan excedido los respectivos umbrales, todas las unidades de datos encaminados posteriormente se marcan en consecuencia. De este modo los encaminadores interiores pueden ser puestos en práctica de una forma sencilla. Naturalmente, sin embargo es posible combinar el concepto con el que un marcado proporcional, es decir que el nodo interior marque solamente una fracción de las unidades de datos enviadas, proporcional al grado de sobrecarga o congestión, una vez que se ha excedido un umbral dado.

En cualquier caso, proporcionando dos umbrales de carga diferentes y disponiendo el nodo de salida para en consecuencia dar instrucciones sobre una mayor o menor reducción de carga se puede proporcionar una respuesta efectiva a la congestión sin la necesidad de tener que poner en práctica mecanismos complicados en los encaminadores interiores o en los nodos de salida. Por ejemplo, no es necesario realizar un marcado proporcional en los encaminadores interiores, ni es por tanto necesario realizar medidas en el grado de congestión sobre la base de tales unidades de datos marcadas proporcionalmente en el nodo de salida.

Se ha observado que la provisión de dos umbrales y de dos valores de reducción de carga correspondientes es solamente un ejemplo, ya que este concepto puede naturalmente ser aplicado también con un número mayor de umbrales y por tanto con los valores de reducción de carga asociados. Por ejemplo, además de los dos umbrales que se acaba de mencionar, se podría proporcionar un tercer umbral que indicara una carga mayor que el segundo umbral, y por lo tanto el nodo de salida podría ser dispuesto para responder a unos terceros marcadores que indicaran el tercer umbral mediante el envío de la correspondiente instrucción de un tercer valor de reducción de carga correspondiente a un valor de reducción de carga mayor que el del segundo valor de reducción de carga.

BREVE DESCRIPCIÓN DE LAS FIGURAS

- Figura 1: es una representación esquemática de un dominio de red que comprende un nodo de entrada, un nodo de salida y una pluralidad de encaminadores interiores;
- 5 Figura 2: es un diagrama de bloques que muestra esquemáticamente un nodo de borde de acuerdo con una realización del invento;
- Figura 3: es un cuadro de flujos que muestra una realización del método básico del presente invento;
- Figura 4: es un diagrama de bloques que muestra esquemáticamente una realización posterior de un nodo de borde de la presente aplicación;
- Figura 5: es un cuadro de flujos de una realización adicional de un método del presente invento;
- 10 Figura 6: es un cuadro de flujos de otra realización del método del presente invento;
- Figura 7: es un diagrama de bloques que muestra esquemáticamente una realización de un encaminador interior de acuerdo con una realización del presente invento; y
- Figura 8: es un cuadro de flujos de un método para controlar un encaminador interior.

DESCRIPCIÓN DETALLADA DE LAS REALIZACIONES

15 En lo que sigue, las realizaciones específicas del presente invento se describirán haciendo referencia a las figuras. Las realizaciones solamente pretenden ser vistas como ejemplos de los conceptos generales del invento resaltados en las reivindicaciones, de forma que no tienen efecto limitativo alguno sobre la exposición de estos conceptos del invento, que solamente sirven para proporcionar a las personas expertas una buena comprensión del mismo.

20 La Figura 1 muestra de forma esquemática un dominio de red 10 que comprende un nodo de entrada 11 para recibir y encaminar unidades de datos que pertenecen a una pluralidad de flujos al interior del dominio de red 10. Comprende además una pluralidad de encaminadores interiores 13, 14, 15 para encaminar las unidades de datos a través del dominio de red 10, y un nodo de entrada 12 para encaminar las unidades de datos hacia fuera del dominio de red 10.

25 Se ha observado que en la presente especificación y en las reivindicaciones, el término “dominio de red” se usa genéricamente para describir cualquier disposición de nodos que tienen la estructura que se acaba de describir, es decir uno o más nodos de entrada, una pluralidad de encaminadores interiores, y uno o más nodos de salida. Como tal, un dominio de red puede ser una red que tenga sus propios encaminadores, en la que los nodos de entrada y salida son proporcionados por servidores de pasarela apropiados, o un dominio de red puede ser igualmente una subdivisión de una red mayor. Además, en conexión con diferentes protocolos y diferentes tecnologías las subdivisiones de datos enviadas por una red pueden tener nombres diferentes tales como paquetes, celdas, segmentos, unidades de datos de protocolo, etc, y el término “unidad de datos”, tal como se usa en la presente especificación y reivindicaciones, es genérico y se refiere a cualquier subdivisión de una cantidad de datos mayor.

30 En el ejemplo de la Figura 1 se ha supuesto que un fallo de enlace 19 ha ocurrido entre los encaminadores interiores 14 y 15, de forma que todo el tráfico procedente del nodo de entrada 11 al nodo de salida 12 que normalmente pasaba a través del encaminador interior 14 tiene que ser reencaminado a través de los trayectos 16, 17 y 18 por los encaminadores interiores 13 y 15. Como consecuencia de este reencaminamiento podría ocurrir que se diera una situación de sobrecarga o de congestión en el encaminador interior 13.

35 En la realización de la Figura 1 se ha supuesto que uno o más encaminadores interiores están dispuestos para detectar si están sometidos a una situación de congestión, y para marcar de congestión en consecuencia las unidades de datos encaminadas a través de ellos si se ha detectado una situación de congestión. La situación de congestión puede definirse de cualquier forma adecuada o deseable, e igualmente el marcado de congestión puede realizarse de cualquier forma adecuada o deseable. Por ejemplo, la situación de congestión puede determinarse mientras se mide la velocidad de datos del tráfico entrante antes de la formación de la cola, y se compara este valor con un valor umbral. La situación de congestión puede igualmente ser definida por la medición de la longitud de una cola de colocación en la memoria intermedia y la comparación de este valor con un valor umbral correspondiente. En cada caso, si se excede el valor umbral, entonces se determina que existe una situación de congestión. El marcado de las unidades de datos encaminadas para indicar congestión puede por ejemplo ser hecho fijando una información apropiada en el encabezamiento de las unidades de datos encaminadas. Como ejemplo, cuando se usan unidades de datos IP el Punto de Código DiffServ (DSCP) o campo de Notificación Explícita de Congestión (ECN) en el encabezador IP puede fijarse apropiadamente para indicar congestión.

40 El procedimiento de marcado de unidades de datos de congestión puede realizarse de cualquier forma apropiada o deseable, por ejemplo puede hacerse proporcionalmente a la carga o puede hacerse en una forma escalonada. Marcado proporcional significa que se ha medido un grado de sobrecarga, y que solamente una fracción de las unidades de datos enviadas se ha marcado de acuerdo con ese grado. Un marcado escalonado significa que una

vez que se ha excedido el umbral todas las unidades de datos están marcadas. Una tercera opción puede ser la de que el nodo interior marque los paquetes que están por encima de un umbral.

La Figura 2 muestra un diagrama de bloques de un ejemplo esquemático de un nodo de borde de acuerdo con una realización del presente invento. En la Figura 1 se muestra el nodo de borde 12 que está dispuesto para actuar como un nodo de salida. El nodo de borde 12 comprende una memoria intermedia 121 de entrada para recibir unidades de datos 124 del dominio de red 10, y una memoria intermedia 122 de salida para mantener las unidades de datos que han de ser enviadas fuera del dominio de red 10. El nodo de borde 12 comprende además un elemento de control 123 que está dispuesto para realizar una función de tratamiento de la congestión que comprende una parte para detectar la presencia de las unidades de datos marcadas de congestión que llegan al nodo de borde 12, y una parte para reaccionar ante la detección de la presencia de unidades de datos marcadas de congestión solicitando un proceso de control de la congestión apropiado. El proceso de control de congestión comprende enviar al nodo de entrada 11 una instrucción 125 para terminar uno o más flujos para de este modo reducir una carga de tráfico, esperando un periodo de tiempo predeterminado T_{relax} y después determinar si las unidades de datos marcadas de congestión están todavía llegando al nodo de borde 12. Si las unidades de datos marcadas de congestión están todavía llegando, entonces se repiten los pasos de enviar una instrucción 125 y esperar el periodo de tiempo T_{relax} .

En la Figura 3 se muestra este procedimiento en donde el paso S31 determina si las unidades de datos marcadas de congestión están llegando, y si éste es el caso, entonces se ejecutan los pasos S32 y S33, en los que se envía una instrucción de reducción de carga al nodo de entrada, y después se espera un periodo de tiempo T_{relax} . Después se determina de nuevo en el paso S31 si están llegando las unidades de datos marcadas de congestión, y si éste es el caso, entonces se repiten los pasos S32 y S33. Por el contrario, el procedimiento básicamente hace un bucle de retroceso para continuar determinando si están llegando las unidades de datos marcadas de congestión, en donde el paso S34 indica una posibilidad de finalizar el procedimiento, por ejemplo mediante una interrupción externa.

Se ha observado que el elemento de control 123 mostrado en la Figura 2 puede ser proporcionado de cualquier forma adecuada o deseable. Por ejemplo, puede ser un microprocesador de un tipo conocido para nodos de red, el cual sea programable de tal forma que ponga en práctica los conceptos novedosos del presente invento. Como tal, el presente invento puede también ser realizado como un programa informático que comprende partes de códigos de programa dispuestas para realizar el método del invento cuando se ejecutan en un nodo de red programable. El invento puede además ser realizado como un producto de programa informático, tal como una portadora de datos que comprende un programa informático que está dispuesto para poner en práctica el método del invento.

Por otra parte, el elemento de control 123 puede también estar compuesto por elementos de soporte físico especializados para poner en práctica la función de tratamiento de la congestión, o pueden comprender una combinación apropiada de soporte físico y de soporte lógico.

Se ha observado también que el nodo de borde 12 mostrado en la Figura 2 solamente indica las características pertinentes para la explicación del presente invento, de forma que otros elementos conocidos de nodos de red no se muestran por motivos de simplicidad.

La instrucción de reducción de nodos enviada por el nodo de salida al nodo de entrada puede estar dispuesta de cualquier forma adecuada o deseable. La instrucción puede, por ejemplo, ser enviada mediante mensajes de señalización especializados entre los nodos de salida y de entrada, en donde tales mensajes de señalización pueden ser enviados por los mismos trayectos que las unidades de datos, posiblemente usando los mismos protocolos, tales como el IP, o tales mensajes pueden también ser enviados usando trayectos de señalización y/o usando protocolos de señalización especializados.

La información de instrucción propiamente dicha puede ser elegida de diversas formas, como sea adecuado o deseable. Por ejemplo, puede ser una única señal binaria que pueda aceptar uno o dos valores, en la que un valor indica una reducción y el otro indica no reducción. Cuando el nodo de entrada recibe la señal que indica reducción, reduce el tráfico en un valor previamente determinado, por ejemplo la reducción corresponde a una cantidad de velocidad de datos previamente fijada, y el nodo de entrada realiza entonces una selección entre todos los flujos que están siendo tratados con el fin de conseguir la reducción de carga deseada. La operación de selección puede realizarse de cualquier modo adecuado o deseable, por ejemplo por la elección aleatoria de los flujos hasta que se hayan terminado tantos flujos que se consiga la reducción deseada.

En el contexto de las presentes especificación y reivindicaciones el término "flujo" se entiende de forma genérica como unidades de datos relativos a unidades de datos que están identificadas por un cierto identificador de la fuente y del destino, por ejemplo una dirección de la fuente y una dirección de destino. Los identificadores de fuente y de destino pueden también tener elementos adicionales tal como números de puerto. Como consecuencia, los flujos de datos en tiempo real o TCP son un ejemplo de los flujos genéricos a los que se hace referencia en la presente aplicación.

Como se ha indicado antes, la instrucción enviada por el nodo de salida al nodo de entrada puede ser simplemente una simple información binaria que puede por ejemplo ser codificada por un único bit. Sin embargo, es preferible que

la instrucción comprenda más información. De acuerdo con una realización preferida el elemento de control 123 está dispuesto de forma que el proceso de control de la congestión comprende además la fijación de un valor de un parámetro de reducción de carga en respuesta a la detección de unidades marcadas de congestión, y la instrucción es tal que da orden al nodo de entrada para terminar uno o más flujos de acuerdo con el valor del parámetro de reducción de carga. En otras palabras, el parámetro de reducción de carga diferencia más exactamente cuánta carga reducir. Esto puede hacerse de varias formas. Por ejemplo, el parámetro de reducción de carga podría expresarse en términos de un factor de ajuste previamente dispuesto de reducción de carga. Por ejemplo, puede ser una escala de valores de reducción, en la que un primer valor corresponde a un pequeño ajuste previamente determinado A, un segundo valor puede corresponder a un valor de ajuste $2xA$, y un tercer valor puede corresponder a un ajuste mayor $3xA$, en donde la instrucción contiene el respectivo valor primero, segundo o tercero, y el nodo de entrada está dispuesto para reducir en consecuencia la carga mediante la terminación de un número de flujos apropiado.

Más preferiblemente, el elemento de control y el proceso de control de congestión están dispuestos de forma que el parámetro de reducción de carga está expresado en términos de una carga absoluta o relativa. En otras palabras, la carga puede por ejemplo medirse en un valor absoluto tal como una velocidad de datos, y la instrucción puede por lo tanto contener una información de reducción en términos de una carga absoluta, por ejemplo como un valor absoluto de la velocidad de datos. La carga puede también ser expresada en términos relativos, por ejemplo como un porcentaje de la carga total. Como por ejemplo, una carga relativa puede ser indicada como una fracción de la carga causada por varios flujos. El nodo de entrada puede entonces terminar los flujos procedentes de entre ese número hasta la carga total (por ejemplo, la velocidad de datos total proporcionada por los flujos) ha sido reducida en la proporción indicada.

De acuerdo con una realización preferida el elemento de control y el proceso de control de congestión en el nodo de salida son además tales que se proporciona una rutina de detección del flujo para detectar qué flujos están afectados por la congestión, y la instrucción al nodo de entrada comprende una indicación de los flujos de los que uno o más tienen que ser terminados de acuerdo con el valor del parámetro de reducción de carga. En otras palabras, cuando se considera el ejemplo de la Figura 3 es preferible que haya un paso adicional entre S31 y S32, lo que comprende detectar los flujos que están afectados por la congestión. Esto significa que después de haber detectado las unidades de datos marcadas de congestión el elemento de control y el proceso de control de congestión en el nodo de salida identifican qué flujos contienen marcas de congestión, esto es qué flujos llevan una marca que indica congestión. Los flujos están identificados por sus indicaciones de fuente y de destino. Por ejemplo, esta rutina de detección puede comprender el examen de las unidades de datos que son recibidas dentro de un cierto periodo de tiempo, y notificar a todos los identificadores de flujo de las unidades de datos que llevan una marca de congestión. Los identificadores de flujo así determinados son enviados a continuación al nodo de entrada con la instrucción, de forma que el nodo de entrada pueda entonces terminar los flujos del grupo indicado. La decisión con respecto a qué flujos terminar puede ser tomada en el nodo de entrada o en el nodo de salida. Si la decisión está tomada en el nodo de entrada entonces no es necesario proporcionar más información que los identificadores de los flujos afectados por la congestión. No obstante, de acuerdo con una realización posterior el nodo de salida está dispuesto de tal manera que su elemento de control y el proceso de control de congestión comprenden también una rutina de selección de flujo para seleccionar qué flujos terminar y en la que la instrucción al nodo de entrada comprende entonces una indicación de los flujos seleccionados. Esta realización es ventajosa por cuanto el procedimiento de control del nodo de entrada puede ser mantenido simple, ya que el proceso de selección se realiza en el nodo de salida.

Como ya se ha especificado anteriormente el proceso de selección con respecto a qué flujos terminar puede ser realizado de cualquier forma adecuada o deseable. Por ejemplo, puede ser un proceso de selección aleatoria de un flujo afectado, de medir la carga de ese flujo (por ejemplo, en términos de velocidad de datos), de determinar la reducción de carga cuando se termina el flujo seleccionado, y después juzgar si la terminación lleva a la reducción deseada en el valor fijado del parámetro de reducción de carga. Si no se consigue la reducción de carga deseada entonces se repite el proceso de selección aleatoria hasta que se hayan considerado tantos flujos para la terminación que se consiga la reducción de carga deseada. Los flujos así seleccionados son después indicados al nodo de entrada para la terminación real.

Naturalmente que son posibles otras formas de selección, por ejemplo es posible que se mida la carga individual de los flujos y que se clasifiquen de acuerdo con su carga, desde la carga más alta a la más baja. Entonces, el proceso de selección comienza seleccionando los flujos con la carga más alta y examinando el efecto de terminación en la reducción de carga y seleccionando los flujos desde el de carga más alta al de carga más baja, hasta que se consiga la reducción de carga deseada. Este procedimiento de selección tiene la ventaja de que resulta en el menor número de flujos terminados para conseguir la reducción de carga deseada.

Un criterio adicional para seleccionar los flujos que hay que terminar puede depender del tipo de flujo. En otras palabras, el proceso de selección puede además tener en cuenta un identificador de tipo de flujo (por ejemplo, flujo en tiempo real, flujo en tiempo no real) y después realizar el proceso de selección según el tipo de flujo. Por ejemplo, el proceso de selección puede comenzar terminando solamente los flujos en tiempo no real. La selección basada en el tipo de flujos puede también ser combinada con la selección aleatoria antes descrita o la selección basada en la contribución de la carga.

Se ha observado que la información de flujo puede también usarse como un medio de hacer saber a los nodos de entrada qué instrucciones tienen que procesar. Esto se explicará con más detalle en lo que sigue. En una situación simple un dominio de red solamente tiene un nodo de entrada. En este caso cada uno de uno o más nodos de salida envían su instrucción de reducción de carga a ese nodo de entrada. Por otra parte, si un dominio de red comprende una pluralidad de nodos de entrada, entonces existen posibilidades diferentes de asegurar una comunicación fiable entre los nodos de salida que envían instrucciones y los nodos de entrada que reciben las instrucciones. Una posibilidad es que las unidades de datos que fluyen a través del dominio de red comprendan una indicación del nodo de entrada que ha inyectado la unidad de datos en el dominio de red. En este caso un nodo de salida que recibe una unidad de datos marcada de congestión puede no solamente identificar a qué flujo pertenece esa unidad de datos sino también de qué nodo de entrada procede. Por lo tanto, las instrucciones de reducción de carga pueden ser enviadas al nodo de entrada apropiado basándose en esa información. Como alternativa, en un sistema en el que las unidades de datos no llevan indicaciones que identifican los nodos de entrada, o si tales informaciones no tienen que usarse, cada nodo de salida puede enviar sus instrucciones a todos los nodos de entrada juntamente con al menos una indicación de los flujos de los que uno o más tienen que ser terminados, o con una indicación de los flujos específicos que tienen que ser terminados. Cada nodo de entrada puede entonces comparar los flujos identificados con los flujos que están siendo tratados por él mismo, con el fin de determinar si es el receptor correcto de la instrucción. Si lo es, termina los flujos apropiados, y si no lo es simplemente ignora la instrucción.

Con respecto a la fijación de un valor para el parámetro de reducción de la carga, esto puede ser en principio realizado de cualquier forma adecuada o deseable. Por ejemplo, si los encaminadores interiores del dominio de red están dispuestos para marcar proporcionalmente las unidades de datos cuando se excede un umbral de congestión, en las formas descritas previamente, entonces el elemento de control y la función de tratamiento de la congestión del nodo de salida pueden estar dispuestos para determinar el grado de congestión o sobrecarga contando el número de unidades de datos marcadas y no marcadas de cada flujo afectado para de este modo recibir una indicación del grado de sobrecarga sobre la base de la relación entre las unidades marcadas y no marcadas. El valor del parámetro de reducción de carga puede entonces fijarse en consecuencia, por ejemplo proporcional al grado de sobrecarga o congestión medido. Como se ha especificado en la introducción a la presente aplicación, los conceptos para disponer encaminadores interiores de los dominios de red de tal forma que puedan marcar las unidades de datos proporcionalmente al grado de sobrecarga son conocidos en la técnica anterior, de forma que no es necesaria una explicación adicional aquí.

De acuerdo con una realización preferida del presente invento, con el fin de ser capaz de simplificar los encaminadores interiores y también el procedimiento de control en los nodos de salida dispuestos para detectar situaciones de sobrecarga, se exponen los siguientes nuevos conceptos. Uno puede ordenar los nodos interiores, como por ejemplo los mostrados en el ejemplo de la Figura 7, con un procedimiento de control que tiene que ser explicado con referencia a la Figura 8.

La Figura 7 muestra un encaminador interior que comprende una memoria intermedia de entrada 710 para recibir las unidades de datos 720 que tienen que ser encaminados, una memoria intermedia de salida 711 para generar unidades de datos 721 que tienen que ser enviadas, y un elemento de control 712 para controlar la operación de recibir y enviar unidades de datos. Como en el ejemplo de la Figura 2, no se muestran los elementos adicionales conocidos de tales encaminadores por motivos de simplicidad.

El elemento de control 712 comprende una parte 7121 dispuesta para detectar si una carga en el encaminador interior excede un primer umbral de carga th_1 o un segundo umbral de carga th_2 , en donde el segundo umbral de carga th_2 es mayor que el primer umbral de carga th_1 . La carga puede, por ejemplo, medirse en términos de velocidad de datos de las unidades de datos que llegan 720 antes de la formación de la cola. Los umbrales pueden ser en consecuencia expresados en términos absolutos, es decir como un valor de carga, o en términos relativos, por ejemplo como una fracción de un valor de carga máximo previamente determinado. El encaminador interior de la Figura 7 comprende además una parte 7122 en el elemento de control 712 para generar unidades de datos marcadas de congestión que comprenden un primer marcador L1 si la carga de tráfico excede el primer umbral de carga th_1 , y un segundo marcador L2 diferente del primer marcador si la carga de tráfico excede el segundo umbral de carga th_2 . En el ejemplo de la Figura 7 se ha supuesto que se ha excedido el primer umbral de carga th_1 pero no el segundo umbral de carga th_2 , y por lo tanto las unidades de datos salientes 721 están marcadas con L1.

En la Figura 8 se muestra un ejemplo de un proceso de control. En un primer paso S81 se ha determinado si la carga excede el segundo umbral th_2 . Si lo excede, las unidades de datos salientes están marcadas con L2, véase el paso S82, y el proceso hace un bucle de retroceso al comienzo. Si no se excede el umbral de carga th_2 el procedimiento va al paso S83 en el que se determina si se ha excedido el umbral th_1 . Si éste es el caso, entonces las unidades de datos salientes son marcadas con L1, véase el paso S84, y el procedimiento hace un bucle de retroceso al comienzo. Si no se ha excedido th_1 , entonces el paso S85 determina si ha de finalizarse el procedimiento, por ejemplo basado en una interrupción externa, y si el procedimiento tiene que continuar, hace un bucle de retroceso al comienzo.

Se ha observado que el método de control de la Figura 8 es convenientemente sencillo por cuanto el marcado de las unidades de datos se realiza en un modo escalonado, es decir una vez que se ha excedido un umbral dado, entonces todas las unidades de datos salientes son marcadas en consecuencia. En este ejemplo no hay marcado

proporcional. El método de control de la Figura 8 es un ejemplo de cómo controlar el encaminador interior mostrado en la Figura 7. Sin embargo, se ha observado que el encaminador interior de la Figura 7 puede también ser operado de una forma que incluya un marcado proporcional, por ejemplo el elemento 7121 no solamente determina lo que excede del umbral de th_1 o th_2 , sino que al mismo tiempo mide también el grado en que se ha excedido el umbral, y el elemento 7122 marca una proporción de las unidades de datos en correspondencia con el grado en que se han excedido los umbrales respectivos.

En ambos casos el nodo de borde que actúa como un nodo de salida está preferiblemente dispuesto de la siguiente forma. El nodo de salida puede estar dispuesto de modo que el proceso de control comprenda la fijación del parámetro de reducción de carga en un primer valor LR_1 de reducción de carga en respuesta a la detección del primer marcador L1, y en un segundo valor LR_2 de reducción de carga en respuesta a la detección del segundo marcador L2. El segundo valor LR_2 de reducción de carga corresponde a una reducción de carga mayor que la del primer valor LR_1 de reducción de carga.

Un ejemplo de este concepto se muestra en la realización del método de la Figura 5. En un primer paso S500 se ha determinado si están presentes las unidades de datos marcadas de congestión. Si no lo están, el sistema hace un bucle de retroceso al comienzo a través del paso 510 en el que se determina si el procedimiento debe terminar, por ejemplo sobre la base de una interrupción externa. Si las unidades de datos marcadas de congestión están presentes, el paso S501 determina si el marcado es una marca L1. Si lo es, el procedimiento avanza al paso 506, en el que el parámetro de reducción de carga se fija en un primer valor LR_1 . Entonces, en el paso S507 el procedimiento identifica los flujos afectados (esto es, qué flujos llevan marcas L1) y selecciona qué flujos terminar con el fin de conseguir una reducción de carga de acuerdo con LR_1 . Además, después se envía una correspondiente instrucción al nodo de entrada. Después de esto, en el paso S508 se fija un periodo T_{relax} durante el cual se espera, es decir no se emprende reacción alguna más contra las unidades de datos marcadas de congestión que se reciben.

Después de haber expirado el periodo de espera T_{relax} el paso S509 determina si están llegando todavía las unidades de datos marcadas de congestión que llevan L1. Si están llegando, entonces el procedimiento hace un bucle de retorno al paso S506 con el fin de repetir el proceso de fijación del parámetro de reducción de carga y para enviar una instrucción al nodo de entrada para reducir la carga. Si no hay más unidades de datos que transportan L1, entonces el procedimiento hace un bucle de retroceso al paso S500.

Si en el paso S501 por otra parte se ha determinado que las unidades de datos marcadas de congestión llevan marcas L2, entonces el procedimiento avanza a los pasos S502 a S504, los cuales son básicamente idénticos a los pasos S506 a S508, excepto que en el paso S502 el parámetro de reducción de carga se fija en el segundo valor LR_2 , el cual indica una reducción de carga mayor que LR_1 . Después del paso S504, el paso S505 determina si las unidades de datos que llevan marcas L2 están todavía llegando, y si es éste el caso, el sistema hace un bucle de retroceso al paso S502 con el fin de repetir los pasos de establecimiento del parámetro de reducción de carga en LR_2 y después enviar una instrucción apropiada al nodo de entrada. Por otra parte, si no están llegando más unidades de datos que llevan L2, el procedimiento va al paso S509 para determinar si están llegando posiblemente unidades de datos que llevan L1. La importancia de esto reside en el hecho de que LR_2 está asociado con L2, el cual a su vez está asociado con el umbral th_2 que es mayor que el umbral th_1 , el cual a su vez está asociado con L1 y LR_1 . Es decir, si la reducción en LR_2 dada en el paso S503 lleva a una reducción de carga en un encaminador interior sobrecargado por debajo de th_2 es posible que la carga todavía esté por encima de th_1 , es decir que se generan las unidades de datos marcadas con L1, de forma que se debería seguir un procedimiento para reducir adicionalmente la carga de acuerdo con los pasos S506 a S508.

La realización actualmente descrita tiene las siguientes ventajas. Cuando el nodo de salida recibe las unidades de datos marcadas de congestión que llevan el marcador L2 se puede concluir que está ocurriendo una sobrecarga que excede th_2 . Si están llegando las unidades de datos que llevan L1 se puede concluir que está ocurriendo una sobrecarga que excede th_1 . Basado en el conocimiento de que th_2 corresponde a una carga mayor que th_1 se puede iniciar una correspondiente reducción de carga mayor (es decir, basada en LR_2) sin tener que realizar medidas que indiquen el grado de sobrecarga. Como consecuencia, ambos encaminadores interiores y nodos de salida pueden ser simplificados en comparación con la técnica anterior.

En lo que sigue se presentarán ejemplos para elegir los valores primero y segundo del valor de reducción de carga.

Al usar el concepto de marcado de unidades de datos en dos niveles, el marcado proporciona información sobre el grado de carga en un encaminador anterior sin tener que realizar medidas específicas. A saber, cuando las unidades de datos están marcadas L1 se sabe que la carga está comprendida entre th_1 y th_2 . Además, cuando se ha detectado la marca L2, la carga excede th_2 .

th_1 y th_2 se eligen preferiblemente de tal modo que th_1 esté asociado con una ligera sobrecarga, en tanto que th_2 esté asociado con una fuerte sobrecarga que requiera una radical disminución de carga. Por ejemplo, th_1 puede fijarse en el 5% por encima de un valor de carga óptimo, en tanto que th_2 puede fijarse en un 10% por encima del valor de carga óptimo. Naturalmente, éstos son solamente ejemplos, y son posibles otras selecciones.

Como ya se ha indicado previamente, th_1 y th_2 pueden expresarse de cualquier forma o modo deseados. Por ejemplo, pueden ser expresados como valores de carga absolutos, por ejemplo como una velocidad de datos, o como un valor relativo, por ejemplo como una fracción de una carga límite superior absoluta. Se ha observado que de acuerdo con el presente invento es preferible que todos los encaminadores interiores del dominio de red en el que se aplica la presente aplicación usen los mismos valores de th_1 y th_2 . Debido a esto, si los diferentes encaminadores interiores del dominio de red tienen capacidades diferentes es preferible expresar th_1 y th_2 como valores relativos con respecto a una capacidad máxima absoluta dada.

Como ya se ha mencionado, si el nodo de salida recibe unidades de datos marcadas L2, entonces se puede concluir que la carga en un encaminador interior excede th_2 . Cuando se reduce la carga se desea disminuir bruscamente la carga, aunque preferiblemente no por debajo de th_1 , ya que esto constituiría un subimpulso. Como consecuencia, se prefiere elegir el segundo valor de reducción de carga que se fija si las unidades de datos L2 son recibidas en un valor que corresponde a una reducción de carga que es igual a la diferencia entre el segundo umbral de carga th_2 y el primer umbral de carga th_1 . Es decir, reduciendo la carga en este factor, se asegura que la carga no caiga por debajo de th_1 .

Otro modo de conseguir el mismo efecto es elegir el segundo valor de reducción de carga de tal manera que corresponda a una reducción de carga que sea igual al cociente de la diferencia entre el segundo umbral de carga th_2 y el primer umbral de carga th_1 dividido por el segundo umbral de carga th_2 , es decir igual a $(th_2 - th_1) / th_2$. Eligiendo el segundo umbral de carga de este modo es preferible si el invento se aplica en un dominio de red que tiene una pluralidad de nodos de salida. Para dar un ejemplo, uno puede aceptar que th_1 y th_2 se expresen como valores relativos de una carga máxima absoluta, por ejemplo como el 59% y el 60%, respectivamente. En este caso el segundo valor de reducción de carga se puede fijar en $(60-59)/60 = 1/60$. Como consecuencia, el nodo de borde que selecciona los flujos con el fin de reducir la carga de acuerdo con el valor de reducción de carga determinará la carga total de todos los flujos afectados (por ejemplo, la velocidad de datos total) y después reducirá el número de flujos de forma que la carga total se reduzca en $1/60$.

Cuando el nodo de salida recibe unidades de datos marcadas L1 se puede concluir que la carga en un encaminador interior está comprendida entre th_1 y th_2 . Como consecuencia, el primer valor de reducción debería corresponder a una reducción de carga menor que el segundo valor de reducción de carga, debido al uso de los anteriores ejemplos, una reducción de carga que corresponda a $th_2 - th_1$ reduciría ciertamente la carga por debajo de th_1 , por lo que se produciría un subimpulso. El valor de la primera reducción de carga puede elegirse de cualquier modo adecuado o deseable con respecto a un segundo valor de reducción de carga, por ejemplo como una cierta fracción tal como $1/2$, $1/3$, $1/4$, etc.

De acuerdo con una realización posterior el elemento de control y el proceso de control de congestión del nodo de salida están dispuestos de tal forma que si el paso de fijar el parámetro de reducción de carga en el primer valor de reducción de carga LR_1 se repite más de una vez, es decir que después del periodo de tiempo T_{relax} las unidades de datos que llevan L1 son todavía recibidas, el valor de LR_1 se reduce consecutivamente en cada repetición. En otras palabras, el proceso de control de congestión es tal que comprende un procedimiento para enviar al nodo de entrada la instrucción de terminación que corresponde a un primer valor transitorio de reducción de carga $LR_1(i)$ y después esperar durante un periodo de tiempo T_{relax} previamente determinado, subsiguientemente determinar si las unidades de datos marcadas de congestión que contienen el primer marcador L1 están todavía llegando, y si todavía se detecta el primer marcador, fijar el parámetro de reducción de carga en un nuevo valor de reducción de carga $LR_1(i+1)$ que sea menor que el valor transitorio de reducción de carga previo $LR_1(i)$, y después repetir los pasos de enviar la instrucción y esperar un periodo de tiempo T_{relax} .

Esto puede, por ejemplo, ser visto a partir del ejemplo de la Figura 6, el cual es una variante del método mostrado en la Figura 5. Los pasos de la Figura 6 que son idénticos a los pasos correspondientes de la Figura 5 llevan el mismo número de referencia, de forma que no es necesaria una nueva descripción. La diferencia entre el ejemplo de la Figura 6 y el de la Figura 5 es que después de determinar en el paso S501 que las unidades de datos marcadas de congestión llevan los marcadores L1, se fija un valor i en cero en el paso S511. Posteriormente, en el paso S512, que corresponde al paso S506 de la Figura 5, el parámetro de reducción de carga se fija en un primer valor $LR_1(i) = LR_1(0)$. Entonces, en el paso S507 se seleccionan los flujos que han de ser terminados y se envía una instrucción correspondiente. Después de esto, en el paso S513 se incrementa i en 1. A continuación, en el paso S508, se inicia el periodo de tiempo de espera T_{relax} . Si después del paso S509 el método hace un bucle de retroceso al paso S512 se fija un nuevo primer parámetro de reducción de carga $LR_1(i) = LR_1(1)$ en el que en general $LR_1(i+1) < LR_1(i)$. Se ha observado que en el método de la Figura 6 existe también un paso S514 para reposicionar el valor de i subsiguiente al paso S505 que determina que no se están recibiendo unidades de datos marcadas L2.

La importancia de elegir valores siempre decrecientes de LR_1 consiste en que uno conoce que la carga está comprendida entre th_1 y th_2 , entonces cada paso de reducción de la carga hace más probable que uno caiga por debajo de th_1 . Como consecuencia, con el fin de evitar el subimpulso es preferible que el valor de reducción de carga LR_1 sea reducido en cada iteración sucesiva.

La definición de reducir los subsiguientes valores de $LR_1(i)$ puede elegirse de cualquier modo adecuado o deseable. Por ejemplo, $LR_1(i)$ puede elegirse como $LR_1(0) / (i+1)$. De acuerdo con una realización preferida, la fijación del

nuevo primer valor transitorio de reducción de carga $LR_1(i+1)$ menor que el valor provisional previo $LR_1(i)$ está dispuesto de forma que la suma de la secuencia general

$$\sum_{i=1 \dots n} LR_1(i)$$

5 del primer valor transitorio de reducción de carga fijado consecutivamente debería ser menor que o igual al segundo valor de reducción de carga LR_2 . Por ejemplo, si LR_2 se elige para corresponder a $th1-th1$, entonces la suma mayor de $LR_1(i)$ debería preferiblemente ser menor que o igual a $th2 - th1$.

10 En el concepto general del invento y de las realizaciones anteriormente descritas se dispone un periodo de espera T_{relax} con el fin de tener en cuenta la demora entre el envío de una instrucción de reducción de carga por el nodo de salida, siendo el nodo de salida capaz de determinar realmente un efecto de haber enviado esa instrucción. El periodo de espera T_{relax} puede elegirse de cualquier modo adecuado o deseable. Por ejemplo, puede ser fijado en un valor, es decir cuando se instala el nodo de borde que actúa como un nodo de salida, entonces T_{relax} se establece en un valor fijado, por ejemplo elegido sobre la base de los tiempos de demora medidos entre nodos de borde en el dominio de red que se considera. Como ejemplo, si el periodo medio de propagación entre nodos de borde en el dominio de red es T , entonces T_{relax} puede fijarse en un valor algo superior a $2T$, por ejemplo $2,5T$.

15 De acuerdo con una realización preferida, el periodo de tiempo previamente determinado T_{relax} es un valor adaptable, y el elemento de control y el procedimiento de respuesta de congestión del nodo de salida están dispuestos de forma que se proporciona una rutina de adaptación del tiempo T_{relax} . Esta rutina de adaptación preferiblemente comprende una rutina para determinar una indicación de un tiempo de propagación de una unidad de datos desde el nodo de entrada al nodo de salida, y para adaptar T_{relax} sobre la base de esa indicación del tiempo de propagación de la unidad de datos. La rutina para determinar una indicación puede comprender la monitorización de mensajes de renovación y/o de mensajes de petición/respuesta y/o mensajes de señalización de mantenimiento intercambiados entre el nodo de entrada y el nodo de salida. Esto se explicará con respecto a los ejemplos que vienen a continuación.

20 Como se ha mencionado antes, el periodo de tiempo T_{relax} sirve para tener en cuenta la demora entre el envío de una instrucción y ser capaz de avisar de un efecto de esa instrucción. Como consecuencia, el valor debe ajustarse sobre la base de la demora de propagación o tiempo del viaje de ida y vuelta (RTT) asociado con los mensajes de intercambio entre el nodo de salida y el nodo de entrada. Se debe tener en cuenta que el efecto de la terminación de flujos en el nodo de entrada se propagará con las unidades de datos de los flujos que no están terminados, y que el tiempo empleado por estas unidades de datos en alcanzar el nodo de salida puede en realidad ser mayor que el tiempo de propagación medio debido a las demoras de encaminamiento y de formación de cola a lo largo del trayecto a través del dominio de red. Como consecuencia, el periodo de tiempo T_{relax} puede fijarse para que sea significativamente mayor que la demora de propagación, por ejemplo de $3xRTT$ a $5xRTT$.

35 Se pueden usar diferentes enfoques para determinar la demora de propagación o RTT, dependiendo de la naturaleza del dominio de red y de los protocolos y rutinas que se empleen. Por ejemplo, si se utiliza un protocolo que usa estados transitorios (como el RSVP o el protocolo de señalización NSIS) entonces se usan los mensajes de renovación para mantener estados de reserva. Estos mensajes normalmente se envían de forma periódica y proporcionan una buena oportunidad para ser usados para medidas de RTT. Similarmente, se pueden usar mensajes de petición/respuesta especializados para obtener un valor de RTT. En NSIS se define un mensaje de petición, el cual puede ser enviado desde un nodo de salida a un nodo de entrada. Como respuesta el nodo de entrada envía un correspondiente mensaje de respuesta de vuelta al nodo de salida, en el que el nodo de salida puede medir la demora entre el envío de la petición y la recepción de la respuesta, y usar esta medida como una estimación de RTT.

40 Otra posibilidad para evaluar RTT puede ser el uso de mensajes de señalización de mantenimiento enviados entre los nodos de borde de un dominio de red. Sobre la base de tales mensajes de señalización de mantenimiento cada nodo de salida puede mantener una tabla de demoras de propagación o de valores de RTT entre él mismo y cualquier otro nodo de borde.

45 Otro enfoque para actualizar el valor de T_{relax} usa las estimaciones del peor caso de RTT. En vez de medir RTT señalizando mensajes uno puede también usar bases de datos de rutinas de los protocolos de encaminamiento intradominio (tales como USPF, IS-IS) si tales protocolos están siendo usados. Con tal base de datos se puede obtener el trayecto real entre cada par de nodos de borde, y mediante el acceso a las propiedades de longitud de cola del enlace se puede calcular el peor caso de RTT.

50 La Figura 4 muestra una realización preferida de un elemento de control 123. El número de referencia 1231 se refiere a una parte que determina si se están recibiendo unidades de datos marcadas de congestión, pero solamente si T_{relax} no está actualmente ejecutándose, es decir si el sistema no se encuentra en un periodo de espera. El número de referencia 1232 se refiere a una parte para calcular la cantidad de tráfico que terminar, es decir fijar el valor del parámetro de reducción de carga si se detectan unidades de datos marcadas de congestión. El número de referencia 1233 se refiere a una parte para identificar los flujos afectados y para seleccionar qué flujos terminar, con

5 el fin de proporcionar la reducción de carga de acuerdo con el valor de reducción de carga. La parte 1233 puede por ejemplo estar asociada con una base de información en los flujos candidatos para terminación, flujos que pueden ser agrupados de acuerdo con su nodo de entrada. Después de que se haya fijado el valor de reducción de carga la parte 1233 selecciona el número de flujos necesario para conseguir la disminución de carga deseada. El número de referencia 1235 se refiere a una parte para estimar la demora entre el envío de una instrucción y ser capaz de medir el efecto de la instrucción. Por ejemplo, puede ser un estimador de RTT que usa uno o más de los conceptos antes descritos. La parte 1235 puede por ejemplo mantener una base de información de estimaciones de RTT como una función de los nodos de entrada. El número de referencia 1234 se refiere a una parte para calcular o fijar el periodo de espera T_{relax} , preferiblemente sobre la base de la estimación hecha por la parte 1235 y sobre la base de los flujos seleccionados en la parte 1233. El tiempo T_{relax} puede por ejemplo fijarse en el RTT máximo de entre los flujos seleccionados, es decir se ha determinado qué nodo de entrada entre los nodos de entrada asociados con los flujos seleccionados es el más distante, es decir que tiene el RTT mayor, y este RTT es después fijado como T_{relax} . El valor fijado de T_{relax} puede por ejemplo ser contado en sentido descendente en un temporizador, en el que la operación de la parte 1231 se deshabilite mientras que el temporizador esté contando en sentido descendente. Se ha observado que la operación de selección de la parte 1233 puede también ser hecha dependiente de las estimaciones de RTT contenidas en la base de información del elemento 1235 de tal manera que la parte 1233 seleccione los flujos para terminar que tengan el RTT almacenado menor.

10 Los elementos 1231-1235 del elemento de control 1233 mostrados en la Figura 4 pueden ser proporcionados como elementos de soporte físico, como elementos de soporte lógico ejecutados en un procesador programable o como combinaciones de soporte lógico y de soporte físico.

15 Los conceptos anteriormente descritos pueden ser aplicados en cualquier contexto apropiado de dominios de red que tengan nodos de entrada, encaminadores interiores y nodos de salida. Se ha considerado específicamente aplicar el invento en el contexto de los dominios de red que hacen uso de los conceptos descritos en la introducción, tales como los servicios diferenciado (DiffServ) o más específicamente la Gestión de Recursos en DiffServ (RMD). Los conceptos pueden también aplicarse en conexión con el RSVP o el protocolo de señalización NSIS.

20 Aunque el presente invento ha sido descrito con referencia a realizaciones específicas, éstas no pretenden tener efecto limitativo alguno, sirviendo solamente para proporcionar a la persona experta una mejor comprensión, y el alcance de protección del invento viene definido por los términos de las reivindicaciones independientes. Los números de referencia sirven para hacer las reivindicaciones más fáciles de leer, pero sin efecto limitativo alguno.

30

REIVINDICACIONES

1. Un nodo de borde de un dominio de red (10) que comprende al menos un nodo de entrada (11) para recibir y encaminar unidades de datos que pertenecen a una pluralidad de flujos en dicho dominio de red (10), una pluralidad de encaminadores interiores (13, 14, 15) para encaminar dichas unidades de datos a través de dicho dominio de red (10), y al menos un nodo de salida (12) para encaminar dichas unidades de datos fuera de dicho dominio de red (10), estando dispuesto al menos uno de dichos encaminadores interiores (13, 14, 15) para detectar (7121) si está sometido a una situación de congestión, y para marcar de congestión las unidades de datos encaminadas (7122) si está sujeto a dicha situación de congestión, estando dicho nodo dispuesto para actuar como dicho nodo de salida (12) y que comprende:
- 5 un elemento de control (123, 712) dispuesto para realizar una función de tratamiento de la congestión que comprende:
- una parte (S31) para detectar la presencia de unidades de datos marcadas de congestión que llegan a dicho nodo de borde,
 - una parte (S32) para reaccionar ante la detección de la presencia de unidades de datos marcadas de congestión solicitando un proceso de control de la congestión,
- 15 **caracterizado porque**
- el proceso de control de congestión está dispuesto para:
- 1a) enviar (S32) a dicho nodo de entrada (11) una instrucción para terminar uno o más flujos para de este modo reducir una carga de tráfico,
 - 20 1b) esperar (S33) un periodo de tiempo determinado previamente (T_{relax}) y después determinar si las unidades de datos marcadas de congestión están todavía llegando a dicho nodo de borde, y
 - 1c) si las unidades de datos marcadas de congestión están todavía llegando, se repiten los pasos 1a) a 1b).
2. El nodo de borde de la reivindicación 1, en el que dicho elemento de control (123, 712) está dispuesto de tal modo que dicho proceso de control de congestión comprende además la fijación (S502, S506) de un valor de un parámetro de reducción de carga en respuesta a la detección de unidades de datos marcadas de congestión, y disponer dicha instrucción de forma que instruya a dicho nodo de entrada (11) para terminar uno o más flujos para así reducir la carga de tráfico de acuerdo con dicho valor de dicho parámetro de reducción de carga.
- 25 3. El nodo de borde de la reivindicación 2, en el que dicho elemento de control (123, 712) está dispuesto de forma que dicho parámetro de reducción de carga se exprese en términos de una carga absoluta y de una carga relativa.
- 30 4. El nodo de borde de la reivindicación 2 ó 3, en el que dicho elemento de control (123, 712) está dispuesto de forma que dicho proceso de control de congestión comprenda además una rutina de detección y de selección de flujo para detectar qué flujos están afectados por la congestión, y dicha instrucción comprenda una indicación de flujos de los que uno o más han de ser terminados de acuerdo con dicho valor de dicho parámetro de reducción de carga.
- 35 5. El nodo de borde de las reivindicaciones 2 a 4, en el que dichas unidades de datos marcadas de congestión pueden comprender un primer marcador o un segundo marcador, indicando dicho primer marcador que una carga de tráfico en un nodo interior (13, 14, 15) excede un primer umbral de carga, e indicando dicho segundo marcador que una carga de tráfico en un nodo interior (13, 14, 15) excede un segundo umbral de carga que corresponde a una carga más alta que dicho primer umbral de carga, y dicho elemento de control (123, 712) está dispuesto de forma que dicho proceso de control de congestión comprende además la fijación de dicho parámetro de reducción de carga en respuesta a la detección de dicho primer marcador y a un segundo valor de reducción de carga en respuesta a la detección de dicho segundo marcador, correspondiendo dicho segundo valor de reducción de carga a una reducción de carga mayor que dicho primer valor de reducción de carga.
- 40 6. El nodo de borde de la reivindicación 5, en el que dicho elemento de control (123, 712) está dispuesto de forma que dicho segundo valor de reducción de carga corresponde a una reducción de carga que es igual a la diferencia entre el segundo umbral de carga y el primer umbral de carga.
- 45 7. El nodo de borde de la reivindicación 5, en el que dicho elemento de control (123, 712) está dispuesto de forma que dicho segundo valor de reducción de carga corresponde a una reducción de carga que es igual al cociente de la diferencia entre el segundo umbral de carga y el primer umbral de carga, dividido por el segundo umbral de carga.
- 50 8. El nudo de borde de las reivindicaciones 5 a 7, en el que dicho elemento de control (123, 712) está dispuesto de forma que dicho proceso de control de congestión comprende además un procedimiento para:

5a) enviar a dicho nodo de entrada (11) dicha instrucción para terminar uno o más flujos para de este modo reducir una carga de tráfico que corresponde a un primer valor transitorio de reducción de carga,

5b) esperar dicho periodo de tiempo previamente determinado (T_{relax}) y después determinar si las unidades de datos marcadas de congestión que contienen dicho primer marcador están todavía llegando a dicho nodo de borde,

5c) si todavía se detecta dicho primer marcador se fija dicho parámetro de reducción de carga en un nuevo primer valor transitorio de reducción de carga menor que el valor previo provisional de reducción de carga y se repiten los pasos 5a) y 5b).

5
10
9. El nodo de borde de la reivindicación 8, en el que dicho elemento de control (123, 712) está dispuesto de forma que dicha fijación de dicho nuevo primer valor transitorio de reducción de carga menor que el valor transitorio previo está dispuesto de forma que una suma de una secuencia total de primeros valores transitorios de reducción de carga fijados es menor o igual que el segundo valor de reducción de carga.

10. El nodo de borde de una de las reivindicaciones 1 a 9, en el que dicho elemento de control (123, 712) está dispuesto de forma que dicho periodo de tiempo previamente determinado (T_{relax}) es un valor fijado.

15
11. El nodo de borde de las reivindicaciones 1 a 9, en el que dicho periodo de tiempo previamente determinado (T_{relax}) es un valor adaptable, y dicho elemento de control (123, 712) está dispuesto de forma que se proporciona una rutina de adaptación para dicho periodo de tiempo previamente determinado (T_{relax}).

20
12. El nodo de borde de la reivindicación 11, en el que dicho elemento de control (123, 712) está dispuesto de forma que dicha función de tratamiento de congestión comprende una rutina para determinar una indicación de un tiempo de propagación de una unidad de datos desde dicho nodo de entrada (11) a dicho nodo de salida (12), y dicha rutina de adaptación está dispuesta para adaptar dicho periodo de tiempo previamente determinado (T_{relax}) sobre la base de dicha identificación de un tiempo de propagación de una unidad de datos.

25
13. El nodo de borde de la reivindicación 12, en el que dicha rutina para la determinación de una indicación de un tiempo de propagación de una unidad de datos comprende la monitorización de uno o más mensajes de renovación, mensajes de petición/respuesta y mensajes de señalización de mantenimiento intercambiados entre dicho nodo de entrada (11) y dicho nodo de salida (12).

30
14. Un encaminador interior para un dominio de red (10) que comprende al menos un nodo de entrada (11) para recibir y encaminar unidades de datos que pertenecen a una pluralidad de flujos a dicho dominio de red (10), una pluralidad de encaminadores interiores (13, 14, 15) para encaminar dichas unidades de datos a través de dicho dominio de red (10), y al menos un nodo de salida (12) para encaminar dichas unidades de datos fuera de dicho dominio de red (10), comprendiendo dicho encaminador interior un elemento de control (123, 712),

caracterizado porque

35
el elemento de control (123) está dispuesto para detectar (7121) si una carga en dicho encaminador interior excede un primer umbral de carga o un segundo umbral de carga más alto que dicho umbral de carga, y para generar (7122) unidades de datos marcadas de congestión que comprenden un primer marcador si la carga de tráfico excede el primer umbral de carga y un segundo marcador si la carga de tráfico excede el segundo umbral de carga, en el que las unidades de datos marcadas de congestión son enviadas al nodo de salida (12) de forma que puedan ser detectadas y que provoquen una reacción en dicho nodo de salida solicitando un proceso de control de congestión (S31, S32, S33).

40
45
15. Un método para controlar un nodo de salida (12) de un dominio de red (10) que comprende al menos un nodo de entrada (11) para recibir y encaminar unidades de datos que pertenecen a una pluralidad de flujos al interior de dicho dominio de red (10), una pluralidad de encaminadores interiores (13, 14, 15) para encaminar dichas unidades de datos a través de dicho dominio de red (10), y al menos un nodo de salida (12) para encaminar dichas unidades de datos hacia fuera de dicho dominio de red (10), estando al menos uno de dichos encaminadores (13, 14, 15) dispuesto para detectar (7121) si está sometido a una situación de congestión y para marcar de congestión las unidades de datos encaminadas (7122) si está sujeto a dicha situación de congestión, que comprende:

realizar una función de tratamiento de congestión que comprende:

- una parte (S31) para detectar la presencia de las unidades de datos marcadas de congestión que llegan a dicho nodo de borde,

50
- una parte (S32) para reaccionar frente a la detección de la presencia de unidades de datos marcadas de congestión solicitando un proceso de control de congestión,

caracterizado porque

el proceso de control de congestión comprende:

- 1a) enviar (S32) a dicho nodo de entrada (11) una instrucción para terminar uno o más flujos para así reducir una carga de tráfico,
- 1b) esperar (S33) un periodo de tiempo determinado previamente (T_{relax}) y después determinar si las unidades de datos marcadas de congestión están todavía llegando a dicho nodo de borde, y
- 5 1c) si las unidades de datos marcadas de congestión están todavía llegando, se repiten los pasos 1a) a 1b).
16. El método de la reivindicación 15, en el que dicho método está adaptado para controlar un nodo de borde de acuerdo con cualquiera de las reivindicaciones 2 a 13.
- 10 17. Un método para controlar un encaminador interior de un dominio de red (10) que comprende al menos un nodo de entrada (11) para recibir y encaminar unidades de datos que pertenecen a una pluralidad de flujos al interior de dicho dominio de red (10), una pluralidad de encaminadores interiores (13, 14, 15) para encaminar dichas unidades de datos a través de dicho dominio de red (10), y al menos un nodo de salida (12) para encaminar dichas unidades de datos hacia fuera de dicho dominio de red, que comprende:
- 15 detectar si una carga en dicho encaminador interior excede un primer umbral de carga (S83) o un segundo umbral de carga (S81) que es más alto que dicho primer umbral de carga, y generar unidades de datos marcadas de congestión que comprenden un primer marcador si la carga de tráfico excede el primer umbral de carga (S83) y un segundo marcador si la carga de tráfico excede el segundo umbral de carga (S82), en el que las unidades de datos marcadas de congestión se envían al nodo de salida (12) de forma que puedan ser detectadas y que provoquen una reacción en dicho nodo de salida solicitando un proceso de control de congestión (S31, S32, S33).
- 20 18. Un producto de programa informático que comprende partes de código de programa dispuestas para ejecutar todos los pasos del método de una de las reivindicaciones 15 a 17 cuando se ejecutan en un nodo de red programable.
19. Un producto de programa informático que comprende el programa informático de la reivindicación 18.

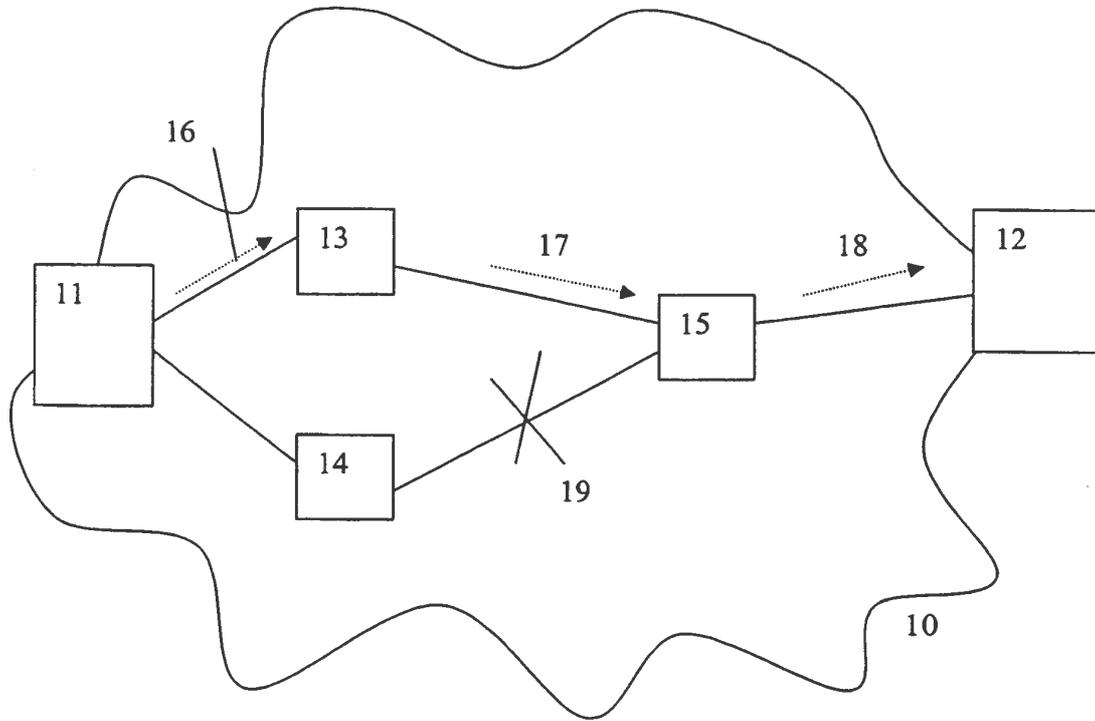


Fig. 1

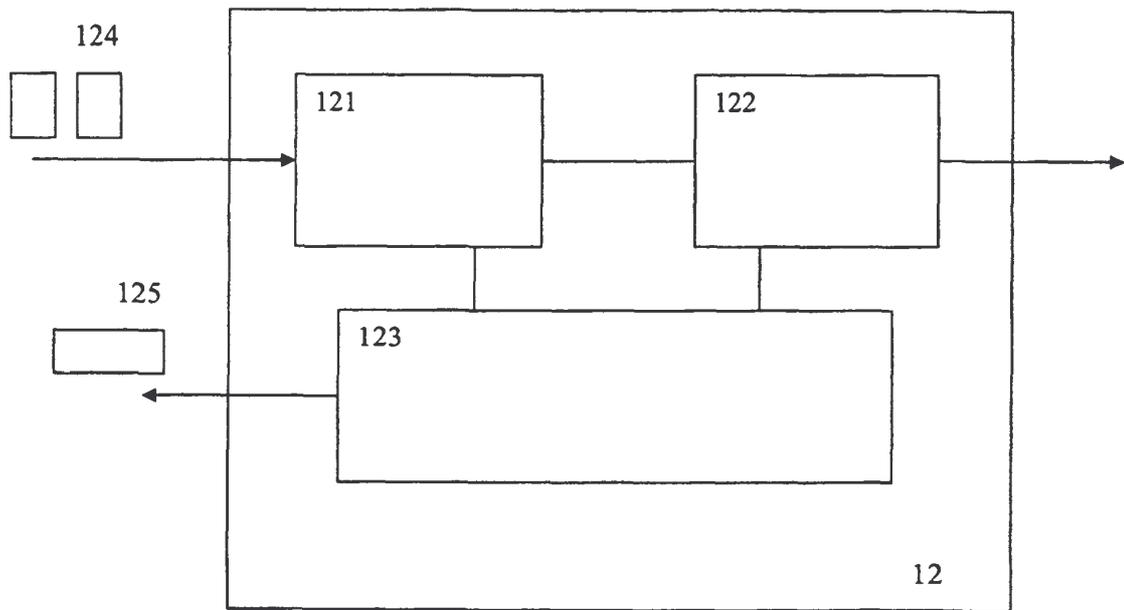


Fig. 2

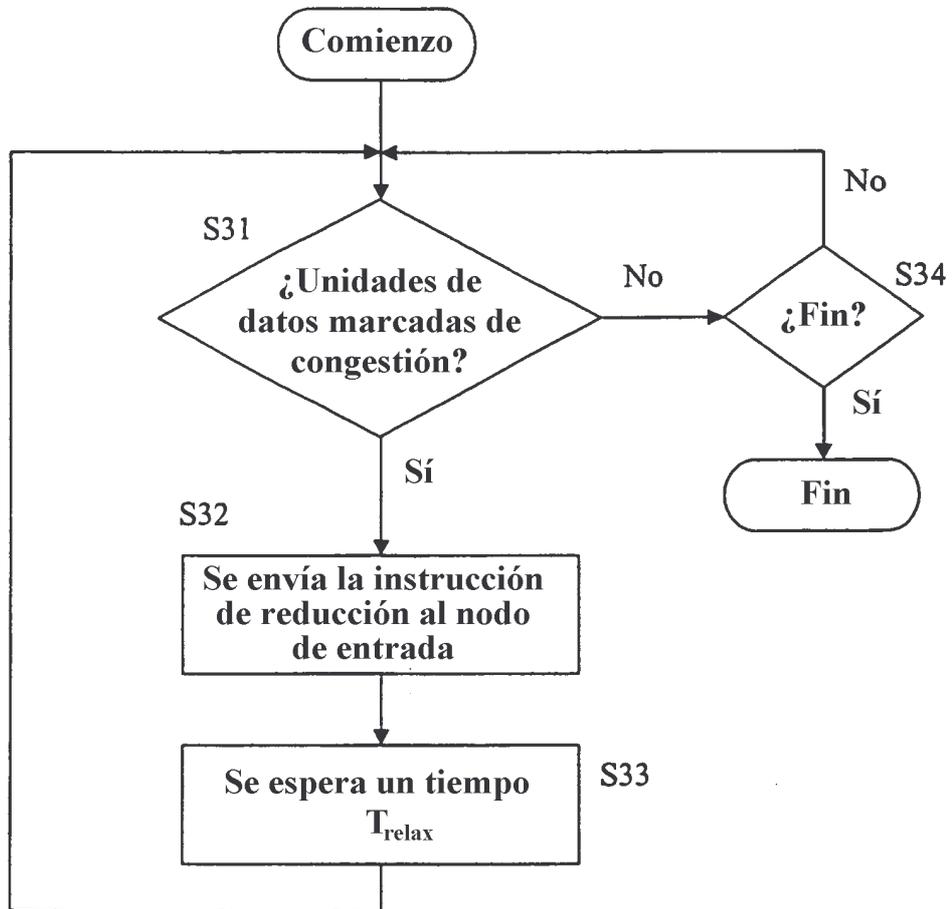


Fig. 3

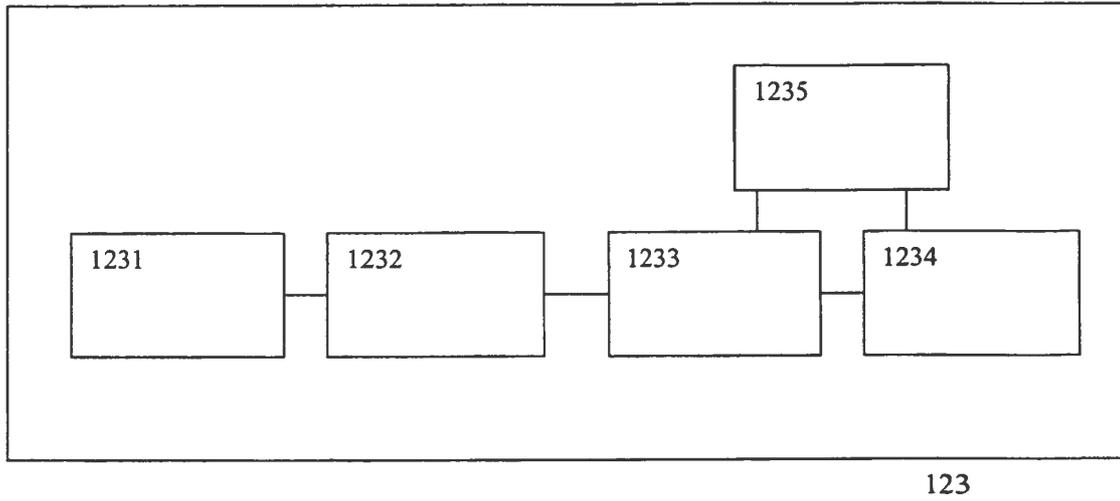


Fig. 4

Fig. 5

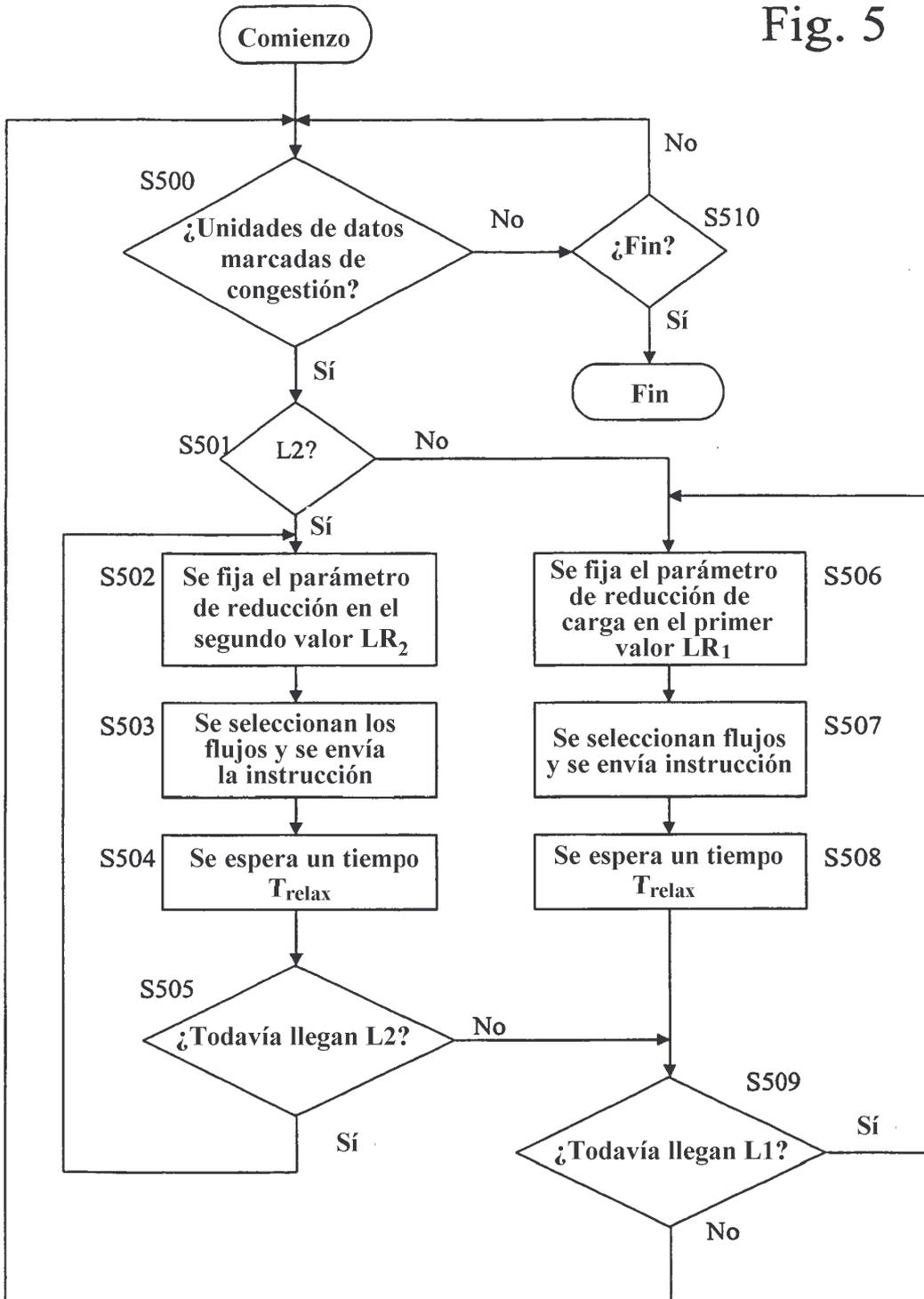
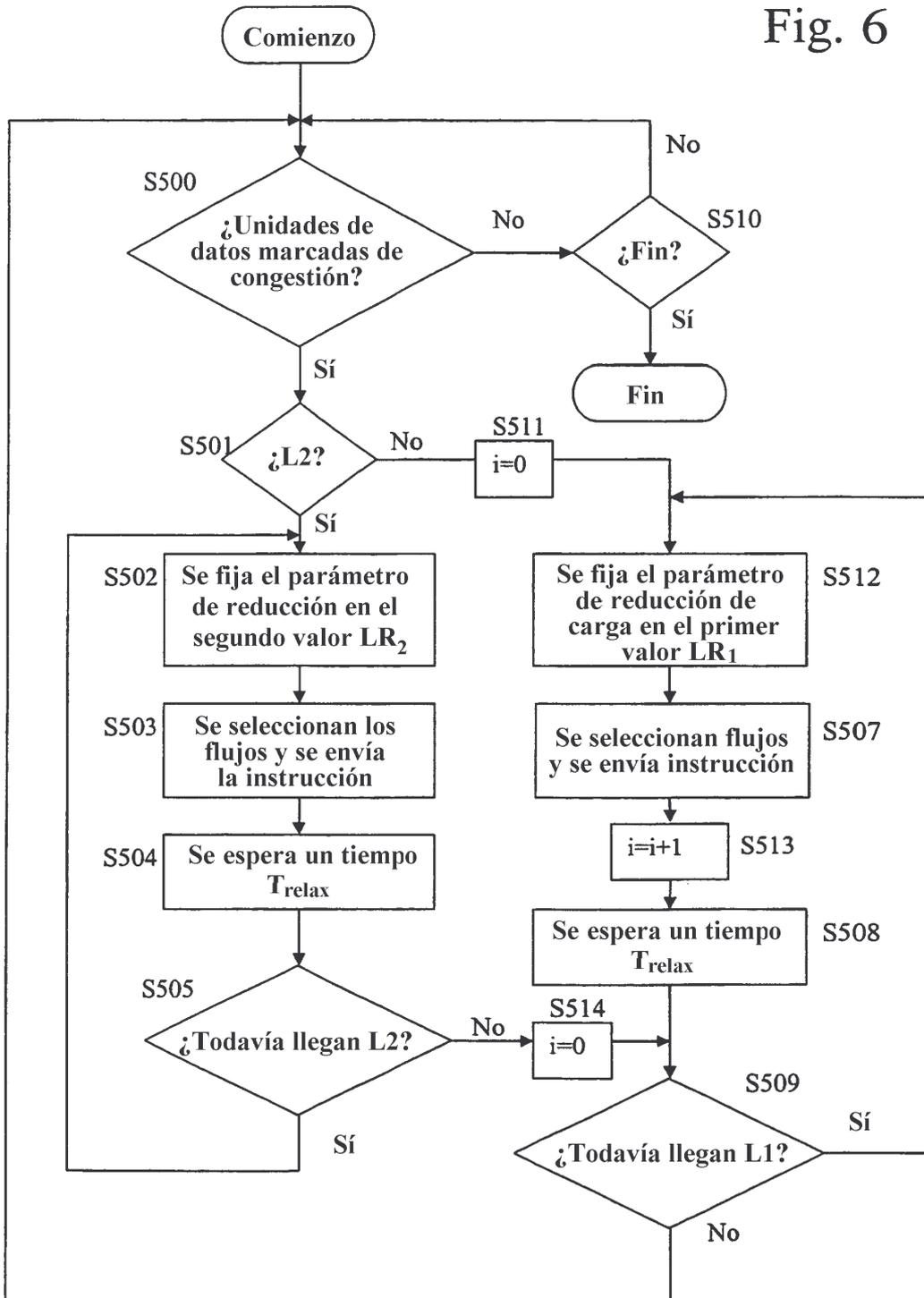


Fig. 6



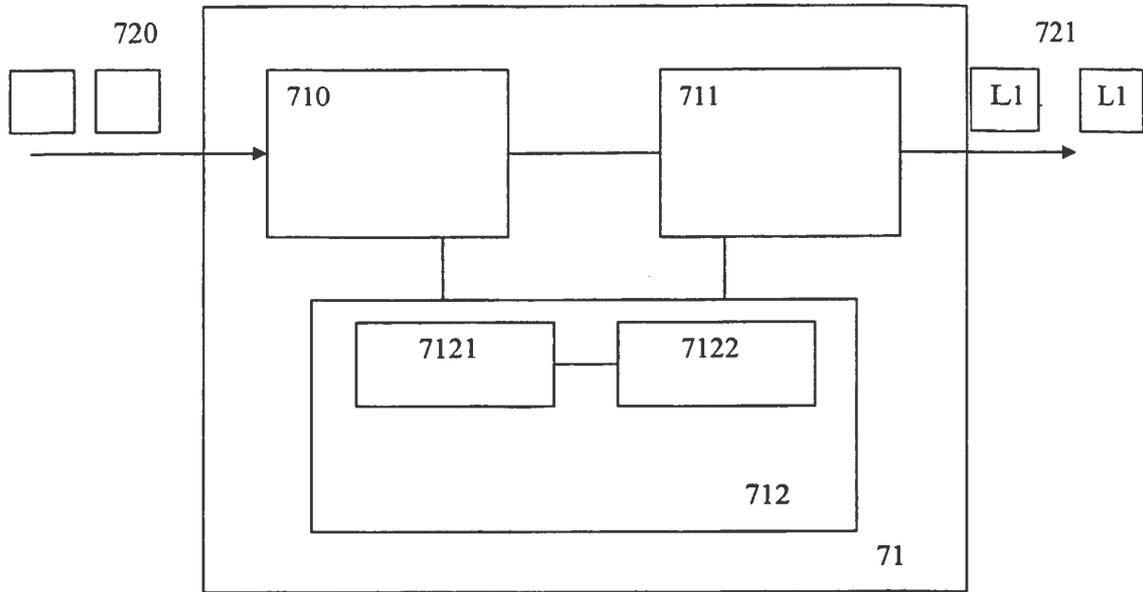


Fig. 7

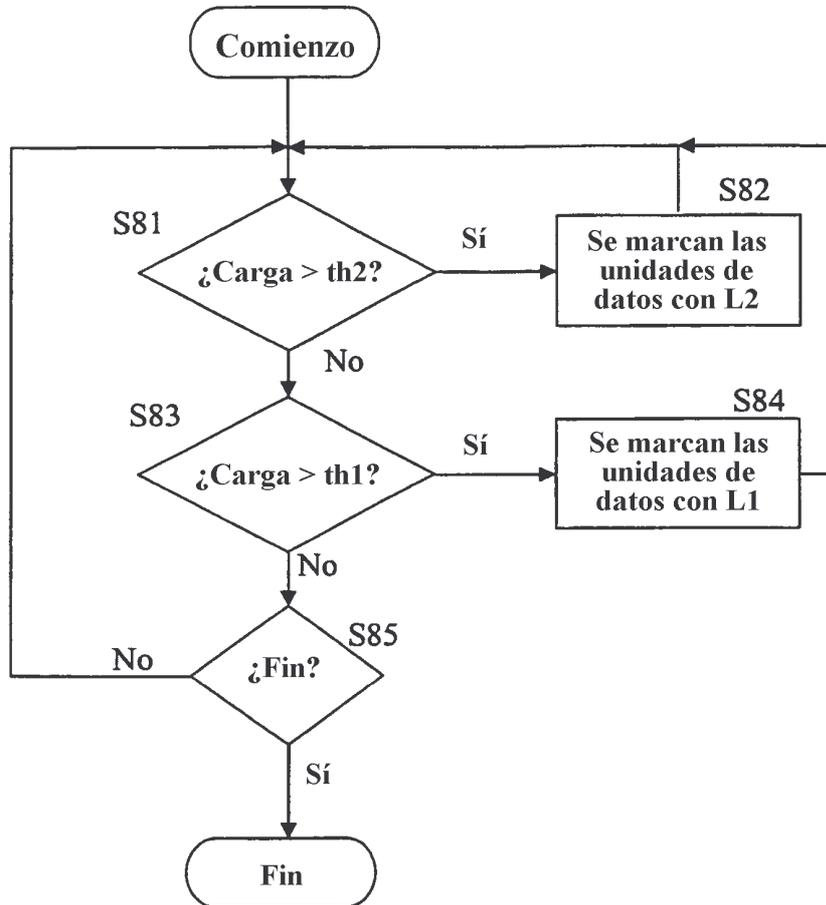


Fig. 8