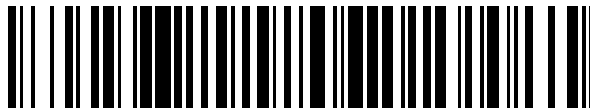


19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 424 859**

51 Int. Cl.:

**H04W 40/10** (2009.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **09.04.2010 E 10716762 (9)**

97 Fecha y número de publicación de la concesión europea: **19.06.2013 EP 2430801**

54 Título: **Encaminamiento teniendo en cuenta la eficiencia energética**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:  
**09.10.2013**

73 Titular/es:

**NEC EUROPE LTD. (100.0%)  
Kurfürsten-Anlage 36  
69115 Heidelberg, DE**

72 Inventor/es:

**DUDKOWSKI, DOMINIQUE y  
HASSELMAYER, PEER**

74 Agente/Representante:

**ROEB DÍAZ-ÁLVAREZ, María**

**ES 2 424 859 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

Encaminamiento teniendo en cuenta la eficiencia energética

5 La presente invención se refiere a una red de comunicación, que incluye una pluralidad de elementos de red - nodos de comunicación - que están habilitados para solicitar el establecimiento de conexiones - flujos - entre sí, una pluralidad de elementos de red - conmutadores - que soportan el encaminamiento de dichos flujos, unos medios para interceptar solicitudes de establecimiento de flujo, por lo menos un módulo de cálculo que está configurado para seleccionar previamente, para cada flujo, un conjunto de trayectos de red posibles que son compatibles con los requisitos de rendimiento y/o las restricciones de calidad de dicho flujo, en la que cada trayecto de red incluye un conjunto específico de conmutadores que ha de recorrer dicho flujo, y por lo menos un módulo de selección que está configurado para aplicar un algoritmo de selección que analiza la carga de los conmutadores que están contenidos en cualquier trayecto de red de dicho conjunto previamente seleccionado de trayectos de red para determinar un único trayecto de red a partir de dicho conjunto previamente seleccionado de trayectos de red que es el más adecuado en términos de eficiencia energética.

Además, la presente invención se refiere a un método para accionar una red de comunicación, incluyendo dicha red una pluralidad de elementos de red - nodos de comunicación - que están habilitados para solicitar el establecimiento de conexiones - flujos - entre sí, y una pluralidad de elementos de red - conmutadores - que soportan el encaminamiento de dichos flujos, comprendiendo el método las etapas de interceptar solicitudes de establecimiento de los flujos, para cada flujo, seleccionar previamente un conjunto de trayectos de red posibles que son compatibles con los requisitos de rendimiento y/o las restricciones de calidad de dicho flujo, en el que cada trayecto de red incluye un conjunto específico de conmutadores que ha de recorrer dicho flujo, y aplicar un algoritmo de selección que analiza la carga de los conmutadores que están contenidos en cualquier trayecto de red de dicho conjunto previamente seleccionado de trayectos de red para determinar un único trayecto de red a partir de dicho conjunto previamente seleccionado de trayectos de red que es el más adecuado en términos de eficiencia energética.

Las redes de comunicación de hoy en día se caracterizan, a menudo, por una carga de red que se encuentra significativamente por debajo de la capacidad de red disponible total. Por ejemplo, en las redes de centro de datos, en determinados momentos del día, por ejemplo durante la noche, la carga de red es pequeña debido a que los servicios que se proporcionan mediante el centro de datos se solicitan con una frecuencia mucho más baja que durante el día y la red está diseñada para manejar picos de tráfico sin congestión. En consecuencia, los conmutadores de red se utilizan, a menudo, muy por debajo del 100 %. En otras redes, tal como las redes de banda ancha, otra razón para la pequeña carga es que el equipo de red, cuando acaba de actualizarse con el fin de proporcionar recursos de reserva que tengan en cuenta el futuro aumento de tráfico hasta que se realice la siguiente actualización, está provisto en exceso. Este hecho implica que, inicialmente, el equipo está significativamente subutilizado.

En tales situaciones de subutilización de red, en principio es posible desactivar algunos elementos de red en la red para ahorrar energía. No obstante, el tráfico se encamina, en general, de forma dinámica de una forma tal que todos se usan los elementos de red. Además, los requisitos de QoS (*Quality of Service*, calidad de servicio) y otras restricciones no permiten determinar con facilidad si un elemento de red es esencial o no y puede desactivarse sin afectar al rendimiento de la red o vulnerar las restricciones de QoS. Como resultado, a pesar de la ligera carga de red, virtualmente todo el equipo de red se mantiene en funcionamiento, y una gran cantidad de energía se malgasta cuando la utilización se encuentra significativamente por debajo del 100 %, en particular cuando se tiene en cuenta una gran cantidad de elementos de red.

La dificultad al inducir una concentración de carga en una red de comunicación es lograr la concentración de una forma tal que se cumple cualquier tipo de restricción, incluyendo la topología de red, la dinámica del tráfico de red a lo largo del tiempo, y diferentes tipos de QoS. Existen unos pocos enfoques que intentan abordar estos problemas. Por ejemplo, tal como se describe en el documento de L. Chiaraviglio y col., "*Reducing Power Consumption in Backbone Networks*", en la *IEEE International Conference on Communication*, 2009, algunos métodos se basan en un conocimiento completo global que captura todos los requisitos, que es extensivo y puede ser sumamente dinámico. Por lo tanto, tales enfoques no pueden proporcionar, debido a la complejidad de cálculo, unos medios eficientes y escalables para concentrar la carga con el fin de ahorrar energía a la vez que se cumplen la totalidad de las restricciones que se mencionan anteriormente. Estos no son capaces de determinar, de una manera oportuna, qué decisiones de encaminamiento han de realizarse con el fin de lograr una red más energéticamente eficiente.

Además, con el fin de mejorar la eficiencia de cálculo, algunas soluciones siguen un enfoque diferente intentando conseguir una concentración de carga sin depender de un conocimiento completo, tal como se describe, por ejemplo, en el documento de M. Gupta y col., "*A Feasibility Study for Power Management in LAN Switches*", en la *12<sup>th</sup> IEEE International Conference on Network Protocols*, págs. 361-371, 2004. No obstante, tales soluciones o bien no pueden evitar por completo la situación de comprometer el rendimiento de la red debido a que estas no pueden cumplir con algunos parámetros relevantes para la QoS en su cálculo, o bien no pueden aprovechar el pleno potencial de ahorro de energía.

El documento US 200910010189 A1 divulga una red con una métrica de encaminamiento de duración de batería restante. De forma más específica, la red incluye una pluralidad de nodos de sensor acoplados de forma comunicativa que incluyen un nodo de destino. Un primer trayecto seleccionado de un nodo de origen al nodo de destino se determina a partir de cualquier pluralidad de trayectos posibles basándose en un coste del primer trayecto seleccionado. El coste de cualquier trayecto se basa, por lo menos en parte, en una función acumulativa de la duración de batería restante para cada nodo de sensor en ese trayecto.

Un objeto de la presente invención es mejorar y desarrollar adicionalmente una red de comunicación y un método para accionar una red de comunicación del tipo que se describe inicialmente de una forma tal que, mediante el empleo de unos mecanismos que pueden implementarse fácilmente, la eficiencia energética de la red se optimiza de una forma fiable y eficiente sin afectar al rendimiento global de la red.

De acuerdo con la invención, el objeto que se menciona anteriormente se logra mediante una red de comunicación que comprende las características de la reivindicación 1. De acuerdo con esta reivindicación, una red de este tipo se caracteriza por que dicho módulo de selección está configurado para realizar dicho algoritmo de selección tal como sigue:

- 1) Crear el conjunto de conmutadores que contiene todos los conmutadores que son parte de cualquier trayecto de red de dicho conjunto previamente seleccionado de trayectos de red y apostillar cada uno de los conmutadores con una métrica de utilización,
- 2) Desmarcar todos los conmutadores en todos los trayectos de red,
- 3) Seleccionar, a partir del conjunto creado en 1), el conmutador con el valor más alto de su métrica de utilización,
- 4) Eliminar, con respecto al conjunto creado en 1), el conmutador seleccionado,
- 5) Marcar el conmutador seleccionado en todos los trayectos de red que contienen el mismo, y
- 6) Si todos los conmutadores en un trayecto de red están marcados, emitir como salida ese trayecto de red, en caso contrario, continuar con 3).

Además, el objeto que se menciona anteriormente se logra mediante un método que comprende las características de la reivindicación independiente 12. De acuerdo con esta reivindicación, tal método se caracteriza por que dicho algoritmo de selección incluye las siguientes etapas:

apostillar cada uno de dichos conmutadores que están contenidos en cualquier trayecto de red de dicho conjunto previamente seleccionado de trayectos de red con un valor escalar que está relacionado con su métrica de utilización,

a partir del conjunto de conmutadores que están contenidos en cualquier trayecto de red de dicho conjunto previamente seleccionado de trayectos de red, seleccionar el conmutador con la métrica de utilización más alta, eliminar dicho conmutador seleccionado con respecto a dicho conjunto de conmutadores, marcar el conmutador seleccionado en todos los trayectos de red de dicho conjunto previamente seleccionado de trayectos de red que contienen dicho conmutador seleccionado,

en el caso en el que todos los conmutadores de un trayecto de red preseleccionado están marcados, seleccionar ese trayecto de red preseleccionado como trayecto de red para el flujo respectivo.

De acuerdo con la invención, en primer lugar se ha reconocido que la eficiencia energética de una red de comunicación puede mejorarse concentrando la carga de red sobre solo aquellos elementos de red que son realmente necesarios, a la vez que se mantienen todas las otras restricciones de red, tal como la topología de red, la dinámica del tráfico de red a lo largo del tiempo, y muchos tipos diferentes de calidad de servicio (QoS). Este problema es extremadamente difícil debido a que una red puede ser muy grande e implicar muchas alternativas de encaminamiento, los patrones de tráfico pueden ser sumamente dinámicos en una escala temporal corta y difíciles de anticipar, y los requisitos de QoS pueden cambiar en cualquier momento.

La presente invención propone una red y un método de acuerdo con los cuales la concentración de carga se induce sobre un subconjunto de los elementos de red que soportan el encaminamiento de los flujos en la red. En lo sucesivo, se hace brevemente referencia a los elementos de red como conmutadores. Tal concentración de carga permite la desactivación de elementos de red sin usar y de ahorrar, potencialmente, una gran cantidad de energía. La ventaja clave de la invención es que no es necesario que esta sepa acerca de cualquier restricción que pudiera dictarse, por ejemplo, por, los sistemas de gestión y de accionamiento de un centro de datos.

La presente invención logra esto mediante la aplicación de un enfoque que tiene influencia sobre la selección de trayecto de red al recomendar esos trayectos que se prefieren en términos de ahorro de energía a partir de un conjunto de múltiples alternativas de trayectos que se proporcionan mediante un módulo de cálculo (por ejemplo, un módulo de QoS). Con el tiempo, tales decisiones se agregan de una forma tal que, finalmente los elementos de red que no se requieren se liberan de cualquier tráfico de red y pueden desactivarse. A este respecto, la presente invención puede considerarse como un tipo de tecnología de ahorro de energía predictiva para las redes de comunicación. En lugar de intentar determinar qué cambios se requieren para liberar algunos elementos de red de forma instantánea, tales cambios no se hacen, sino que el comportamiento de colocar el tráfico sobre los conmutadores se ve influido de una forma que ocasiona un cambio de este tipo de manera gradual. El algoritmo de

selección se basa en unas métricas que expresan el potencial de ahorro de energía y ocasiona el cambio gradual a lo largo del tiempo.

5 De acuerdo con la invención, el algoritmo de selección funciona de la siguiente forma: En una primera etapa cada uno de los conmutadores que están contenidos en cualquier trayecto de red del conjunto previamente seleccionado de trayectos de red se apostillan con un valor escalar que está relacionado con su métrica de utilización. La métrica de utilización puede expresar la utilización (aproximada) del conmutador, por ejemplo en términos del tráfico de red y/o el número de flujos que el conmutador está procesando. El cálculo de la métrica de utilización puede desencadenarse con cualquier cambio en el estado de un conmutador. Este también puede realizarse solo cuando  
10 un nuevo trayecto de red para un flujo se determina mediante el módulo de selección.

De forma más específica, el algoritmo de selección funciona de una forma iterativa en múltiples etapas, seleccionando en primer lugar a partir del conjunto de conmutadores que están contenidos en cualquier trayecto de red del conjunto previamente seleccionado de trayectos de red, el conmutador con la métrica de utilización más alta.  
15 En una siguiente etapa, este conmutador seleccionado con la métrica de utilización más alta se elimina del conjunto de conmutadores, y este se marca en todos los trayectos de red del conjunto previamente seleccionado de trayectos de red que contienen el conmutador seleccionado. A continuación, existen posibilidades: En el caso en el que todos los conmutadores de un trayecto de red preseleccionado están marcados, el algoritmo de selección selecciona ese trayecto de red como trayecto de red para el flujo respectivo. En caso contrario, el algoritmo de selección vuelve a la primera etapa, es decir, seleccionar a partir del conjunto de los conmutadores restantes el conmutador con la métrica de utilización más alta y continuar tal como se describe anteriormente. El presente enfoque garantiza que el conmutador con el valor más pequeño de su métrica de utilización se elige solo cuando no existe otra opción.

25 En otras palabras, siempre se prefieren los conmutadores de valor más grande, lo que quiere decir que los conmutadores de uso más intenso (de acuerdo con la métrica de utilización) atraen incluso más tráfico de red y los de uso más ligero están finalmente desprovistos de tráfico. De nuevo, es importante entender que, mientras que los conmutadores de uso más intenso se toman de acuerdo con las métricas, a estas alturas los requisitos de rendimiento de red (como las restricciones de QoS, etc.) ya se han analizado y no se vulnerarán.

30 Es importante observar que el módulo de selección puede decidir con libertad, sin consultar a ningún otro subsistema de gestión en el centro de datos, cual de las rutas elegir, debido a que estas propiedades se verificaron de antemano y dieron como resultado la selección previa de las rutas por el módulo de cálculo en primer lugar. Esta es la clave para garantizar que el rendimiento de la red no se ve comprometido en momento alguno y que las alternativas de flujo pueden recomendarse de una forma sumamente eficiente con una complejidad temporal lineal.  
35 Además, en el caso en el que determinadas restricciones (por ejemplo, parámetros de QoS) están sujetas a cambios, de acuerdo con la presente invención esto es completamente transparente y no se requieren adaptaciones del algoritmo.

40 Como resultado, la presente invención efectúa una minimización del consumo de energía por la red bajo cualquier tipo de restricciones de red (topología de red, tráfico de red y requisitos de QoS). Esta se adapta a cualquier cambio en las restricciones de red y no es invasiva ya que no compromete el rendimiento (QoS) de la red. Además, la presente invención es sumamente eficiente en términos de tiempo de cálculo, debido a que no requiere un conocimiento completo global que capture todos los requisitos.

45 De acuerdo con una realización preferida, el algoritmo de selección puede desencadenarse mediante un cambio de la configuración de flujo en dicha red de comunicación. Por ejemplo, un desencadenante podría ser la creación de un nuevo flujo, por ejemplo, debido a la creación de una conexión de TCP entre dos servidores unidos a la red. Un caso más complejo podría ser el reencaminamiento de múltiples flujos, lo que es una situación típica cuando una máquina virtual va a reubicarse en otra máquina física. En cualquier caso, uno o más flujos requieren el reencaminamiento a través de los conmutadores de red de una forma tal que cumpla las restricciones de rendimiento y de eficiencia energética.  
50

Con respecto a una alta eficiencia, puede proporcionarse que la red de comunicación incluya por lo menos un nodo de gestión que funciona como entidad de coordinación para un dominio de gestión que incluye por lo menos un subconjunto de dichos conmutadores. En un escenario de aplicación típico, la red de comunicación está dividida en diversos dominios de gestión, en los que cada conmutador de la red pertenece a un dominio específico, y en los que cada dominio tiene un nodo de gestión asociado.  
55

60 En lo que concierne a la interceptación de solicitudes de establecimiento de flujo, se demuestra que es ventajoso que el conmutador que se encuentra lo más cerca, es decir, a una distancia de un único salto con respecto al nodo de comunicación solicitante, realice la interceptación. Con este fin, los conmutadores pueden configurarse para realizar la interceptación por medio de un reencaminamiento de las solicitudes de establecimiento de flujo hacia un nodo de gestión, preferiblemente hacia el nodo de gestión asociado con el dominio de gestión al que estos pertenecen. En una realización específica, los nodos de gestión pueden estar equipados con un módulo de interceptación que está configurado para recibir las solicitudes interceptadas de establecimiento de flujo.  
65

De acuerdo con una realización específica, el algoritmo de selección puede ejecutarse sobre un conjunto de trayectos de red alternativos que tienen diferentes puntos de extremo de trayecto. Por ejemplo, en un escenario de aplicación concreto en el que los puntos de extremo de comunicación pueden ser servidores de un centro de datos, múltiples trayectos de red para múltiples servidores pueden someterse al algoritmo de selección. En el presente caso, el algoritmo de selección trata todos los trayectos sin distinguir los puntos de extremo. Entonces, el trayecto finalmente seleccionado que se emite como salida por el algoritmo de selección también determina de forma implícita el punto de extremo (es decir, un servidor en el caso de un centro de datos) al que se impondrá un flujo a través del trayecto seleccionado.

Para liberar de manera gradual los conmutadores con poca actividad de cualquier carga, el algoritmo de selección puede configurarse para seleccionar, a partir del conjunto previamente seleccionado de trayectos de red un trayecto de red que incluye solo los conmutadores con unas métricas de utilización por encima de un umbral predefinido. Por este medio, los conmutadores con unas métricas de utilización baja se eligen con menos frecuencia como trayecto de red para un flujo que los conmutadores con unas métricas de utilización alta, de tal modo que poco a poco se sustraerá el tráfico de los mismos.

Con respecto a una realización técnica efectiva que pueda implementarse fácilmente, puede proporcionarse que el por lo menos un módulo de cálculo y/o el por lo menos un módulo de selección estén incluidos en el por lo menos un nodo de gestión. De forma más específica, puede proporcionarse que cada uno de los nodos de gestión de las redes esté equipado tanto con un módulo de cálculo como con un módulo de selección. En una realización específica, los módulos de selección pueden configurarse para recibir como entrada el conjunto previamente seleccionado de trayectos de red a partir del módulo de cálculo correspondiente del nodo de gestión respectivo. Además, los módulos de selección pueden configurarse para emitir como salida el único trayecto de red seleccionado a un módulo de control del nodo de gestión respectivo. El módulo de control puede configurarse para imponer el flujo respectivo a través del único trayecto de red seleccionado.

En un escenario de aplicación concreto, los nodos de comunicación pueden ser servidores de un centro de datos. Se considera que los centros de datos y su soporte lógico de gestión son la tecnología objetivo más apropiada en este momento. No obstante, la presente aplicación puede aplicarse, en principio, en todos los tipos de redes de comunicación en los que los flujos entre los nodos de comunicación se encaminan mediante una pluralidad de elementos de red.

Una aplicación prometedora de la presente invención se encuentra en conexión con la tecnología de OpenFlow tal como ha sido desarrollada por el consorcio OpenFlow y tal como se describe de forma ejemplar, por ejemplo, en el Libro Blanco de Nick McKeown y col. "*OpenFlow: Enabling Innovation in Campus Networks*", marzo de 2008 (disponible a través de <http://www.OpenFlowSwitch.org>). Por ejemplo, los conmutadores pueden realizarse como conmutadores de OpenFlow. Además, los módulos de intercepción de los nodos de gestión pueden realizarse como controladores de OpenFlow.

Otra realización es la implementación del método de acuerdo con la presente invención con la tecnología de MPLS (*Multiprotocol Label Switching*, conmutación de etiquetas multiprotocolo), que también permite gestionar los flujos entre nodos. En la presente tecnología, los paquetes se etiquetan con un trayecto que estos recorrerán, y este trayecto puede verse influido, adicionalmente, por el algoritmo de selección. Entonces, puede usarse el protocolo RSVP-TE (*Resource Reservation Protocol - Traffic Engineering*, protocolo de reserva de recursos – ingeniería de tráfico) para reservar recursos a través de una red de IP, es decir, para señalar el trayecto seleccionado. Otra realización es el uso de la tecnología de SNMP (*Simple Network Management Protocol*, protocolo simple de gestión de redes). Por ejemplo, pueden usarse trampas de SNMP para interceptar el envío de paquetes iniciales, y entonces puede usarse el SNMP para configurar los conmutadores con unas entradas de encaminamiento apropiadas.

Existen diversas formas acerca de cómo diseñar y desarrollar adicionalmente la enseñanza de la presente invención de una forma ventajosa. Con este fin, ha de hacerse referencia, por un lado, a las reivindicaciones de patente subordinadas a las reivindicaciones de patente 1 y 12 y, por el otro lado, a la siguiente explicación de un ejemplo preferido de una realización de la invención que se ilustra mediante el dibujo. En conexión con la explicación del ejemplo preferido de una realización de la invención mediante la ayuda del dibujo, se explicarán realizaciones generalmente preferidas y desarrollos adicionales de la enseñanza. En los dibujos

la figura 1 ilustra de forma esquemática una red de comunicación de acuerdo con una realización de la presente invención con una concentración de carga en un único segmento de red,  
 la figura 2 ilustra de forma esquemática una red de comunicación de acuerdo con otra realización de la presente invención, y  
 la figura 3 es una vista esquemática que ilustra la estructura de un nodo de gestión de la red de comunicación de la figura 2.

Una idea esencial de la presente invención es describir una red con un tipo de algoritmo de ahorro de energía predictivo (que se abrevia en lo sucesivo PES). La invención es capaz de anticipar qué equipo de red (que se denomina en lo sucesivo, en general, conmutadores) es el que se desactivará lo más probablemente en el futuro

(inmediato) con el fin de conseguir una red energéticamente eficiente.

La figura 1 muestra un escenario típico que la invención tiene como objetivo. Se muestra el ejemplo de un centro de datos 1 con un total de cuatro servidores 2<sub>1</sub>-2<sub>4</sub> y una pluralidad de elementos de red - conmutadores 3 - que soportan el encaminamiento de flujos, por ejemplo conexiones de TCP, entre los servidores 2<sub>1</sub>-2<sub>4</sub> de la red. Los puntos oscuros indican el número de flujos activos de los servidores 2<sub>1</sub>-2<sub>4</sub>. En la figura 1, mediante la aplicación de un método de acuerdo con la presente invención, la carga de red en la red completa se desplaza hacia y, de ese modo, se concentra en, uno de los segmentos de red en el centro de datos 1. Antes de la concentración de carga, todos los conmutadores 3 están activos. Después de concentrar todo el tráfico de red en el segmento de red derecho, pueden desactivarse todos los conmutadores 3 en el segmento de red izquierdo.

A pesar de que el PES puede aplicarse a diferentes granularidades, incluyendo pero sin limitarse a, puertos de conmutación, conmutadores de red, y la totalidad de los segmentos de red (tal como se muestra en la figura 1), la invención se ilustrará basándose en los conmutadores de red. La idea puede adaptarse de una forma sencilla a puertos de conmutación y segmentos de red.

La idea del PES es predecir aquellos conmutadores de red en la red de comunicación que pasarán lo más probablemente a ser inactivos en el futuro inmediato de tal modo que estos pueden desactivarse. "Inactivo" quiere decir que un conmutador de red no procesa flujo alguno, es decir, no retransmite paquete alguno. De forma intuitiva, cuanto más pequeña sea la cantidad de flujos en un conmutador de red, más sencillo será desactivar el conmutador o bien mediante el reencaminamiento de los flujos restantes o bien, si se soportan tiempos de espera, esperando que expiren los flujos restantes.

Una realización de PES funciona basándose en una secuencia de operaciones que se desencadenan mediante un cambio en la configuración de flujo en la red de comunicación. Un desencadenante es la creación de un nuevo flujo, por ejemplo, debido a la creación de una conexión de TCP entre dos servidores unidos a la red.

El PES no reubica los flujos en la red de forma proactiva, pero puede estar perfectamente integrado con tales métodos. Por ejemplo, cuando el sistema de gestión de un centro de datos decide que debería reencaminarse un número de flujos de forma inmediata, esto puede hacerse sin afectar al funcionamiento del PES. En tales situaciones, el PES puede adaptarse recalculando sus métricas de utilización (véase la sección 6. a continuación) y continúa de forma consistente a partir de este estado. Esto permite, en particular, tratar los flujos persistentes, es decir, los flujos que no cambian a largo plazo (por ejemplo, en la situación de un tráfico periódico en segundo plano). En un caso de este tipo, una reubicación proactiva puede desencadenarse para liberar los conmutadores con respecto al tráfico que está bloqueando su desactivación. A este respecto, es importante observar que la reubicación proactiva es siempre significativamente más costosa.

En lo sucesivo, con referencia a las figuras 2 y 3 se describirá el funcionamiento general de PES basándose en el escenario de crear un nuevo flujo debido al establecimiento de una conexión de TCP entre dos máquinas físicas. La figura 2 ilustra de forma esquemática una distribución de red típica de un centro de datos 1 con un total de ocho servidores 2<sub>1</sub>-2<sub>8</sub> y unos conmutadores de red S<sub>1</sub>-S<sub>14</sub> que soportan el encaminamiento de los flujos. Los conmutadores de red S<sub>1</sub>-S<sub>14</sub> pertenecen a tres dominios de gestión 4 diferentes, estando gestionado cada dominio mediante un nodo de gestión 5 asociado. La figura 3 ilustra la estructura de un nodo de gestión 5 de la red que se muestra en la figura 2.

### 1. Solicitud inicial

En la primera etapa, se solicita la creación de un flujo (múltiples flujos en el caso general) cada vez. En la figura 2, el servidor 2<sub>1</sub> solicita el establecimiento de una conexión de TCP con el servidor 2<sub>8</sub>. El servidor 2<sub>1</sub> está conectado con el conmutador de red S<sub>1</sub>, a través del cual se envía el primer mensaje del TCP del establecimiento de conexión.

### 2. Interceptación de solicitudes

A continuación, la solicitud inicial se intercepta. En el ejemplo de la figura 2, la interceptación tiene lugar en el conmutador de red S<sub>7</sub>. Para ello, el conmutador soporta la interceptación por medio de un reencaminamiento de un paquete hacia el nodo de gestión 5 respectivo. Por ejemplo, los conmutadores de OpenFlow soportan tales características reconociendo que un paquete de datos no se corresponde con entrada de conmutación (denominada entrada de tabla de flujo en OpenFlow) alguna y la solicitud inicial en relación con el TCP se enviará al nodo de gestión 5 (controlador de OpenFlow) que es responsable del dominio de gestión 4 correspondiente (que se indica mediante líneas discontinuas) que contiene el conmutador de red S<sub>7</sub>. En esta etapa, el módulo de interceptación 6 en el nodo de gestión 5 que se muestra en la figura 3 recibe el paquete interceptado y lo entrega a los módulos de cálculo 7.

### 3. Selección de trayecto

Los módulos de cálculo 7 (por ejemplo, un módulo relacionado con la QoS) determinan qué conjunto de conmutadores de red  $S_1$ - $S_{14}$  puede usarse para crear un nuevo flujo. Esta etapa es completamente independiente de los algoritmos y métodos de PES y permite que el controlador de OpenFlow se centre en las restricciones de rendimiento dictadas por cualquier subsistema de gestión de centro de datos. Por lo tanto, esta etapa puede incluir cualquier tipo de negociación, tal como entre múltiples nodos de gestión 5 que controlan diferentes dominios de gestión 4 (es decir, subconjuntos de los conmutadores de red  $S_1$ - $S_{14}$ ), y coordinación con otros subsistemas de gestión del centro de datos 1 que puede que tengan que consultarse con el fin de determinar qué configuraciones de flujo son válidas.

El resultado de esta etapa es un conjunto de trayectos de red (flujos) compatibles con los requisitos de los módulos de cálculo 7. Cada trayecto contiene una lista ordenada de los conmutadores de red  $S$  que han de recorrer los paquetes. En el ejemplo de la figura 2, pueden determinarse dos trayectos diferentes, uno a través de la ruta  $S_7 \rightarrow S_3 \rightarrow S_2 \rightarrow S_6 \rightarrow S_{14}$  y el otro a través de  $S_7 \rightarrow S_4 \rightarrow S_2 \rightarrow S_5 \rightarrow S_{13}$ . Ambos trayectos de red (flujos) cumplen los requisitos de rendimiento (por ejemplo, en relación con el ancho de banda) dictados por los módulos de cálculo 7.

#### 4. Cálculo de PES

Una vez que se calculan los trayectos posibles, se consulta un módulo de selección - al que se hace referencia en lo sucesivo como módulo de PES 8 - con el fin de determinar una recomendación para uno de los trayectos que es el más adecuado en términos de eficiencia energética.

El módulo de PES 8 aplica un algoritmo específico, uno de los cuales se describe adicionalmente en lo sucesivo, para determinar el trayecto a partir de los trayectos de red preseleccionados que es el más eficiente. En el ejemplo de la figura 2, se supone que el primer trayecto,  $S_7 \rightarrow S_3 \rightarrow S_2 \rightarrow S_6 \rightarrow S_{14}$ , es el más adecuado en términos de eficiencia energética en el sentido de que, mediante la selección de este trayecto, los conmutadores no afectados ( $S_1$ ,  $S_3$ , etc. en el ejemplo) siguen siendo los que lo más probablemente queden libres de todos los flujos en el futuro. A estas alturas, la propiedad importante conseguida por la invención es que el módulo de PES 8 puede decidir con libertad, sin consultar a ningún otro subsistema de gestión en el centro de datos 1, cual de las rutas elegir, debido a que estas propiedades se verificaron de antemano por los módulos de cálculo 7 y dieron como resultado la selección previa de las rutas por los nodos de gestión 4 en primer lugar. Esta es una clave para garantizar que el rendimiento de la red no se ve comprometido en momento alguno.

El algoritmo de selección básico, que se ejecuta por el módulo de PES 8, funciona tal como sigue. La suposición básica es que cada uno de los conmutadores de red  $S_1$ - $S_{14}$  se apostilla con un valor escalar, que se denomina métrica de utilización. Este valor se determina mediante un proceso adicional que se describe posteriormente.

El algoritmo funciona entonces de la siguiente forma, con un conjunto de trayectos como entrada:

- 1) Crear el conjunto de conmutadores de red que contiene todos aquellos conmutadores de red que son parte de cualquier trayecto de entrada.
- 2) Desmarcar todos los conmutadores de red en todos los trayectos de entrada.
- 3) Seleccionar, a partir del conjunto creado en 1), el conmutador de red con el valor más alto de su métrica de utilización.
- 4) Eliminar, con respecto al conjunto creado en 1), el conmutador de red seleccionado.
- 5) Marcar el conmutador de red seleccionado en todas las rutas que contienen el mismo.
- 6) Si todos los conmutadores de red en un trayecto están marcados, emitir como salida ese trayecto, en caso contrario, continuar con 3).

El presente enfoque garantiza que el conmutador de red con el valor más pequeño de su métrica de utilización se elige solo cuando no existe otra opción. En otras palabras, siempre se prefieren los conmutadores de red de valor más grande (lo que quiere decir que los conmutadores de red de uso más intenso (de acuerdo con la métrica de utilización) atraen incluso más tráfico de red y los de uso más ligero están finalmente desprovistos de tráfico).

El trayecto determinado mediante el algoritmo que se menciona anteriormente se recomienda y se introduce como entrada entonces en el módulo de control 9 (véase la figura 3), que impone el flujo en la red de acuerdo con la siguiente sección 5.

#### 5. Imposición de flujo

En la etapa final, el flujo recomendado para un único trayecto por el módulo de PES 8 se impone mediante los nodos de gestión 4 en la red de forma conjunta. En la figura 2, se generan unas entradas de flujo para los conmutadores  $S_7$ ,  $S_3$ ,  $S_2$ ,  $S_6$ , y  $S_{14}$  de tal modo que la conexión de TCP puede establecerse a través de estos conmutadores con el fin de alcanzar el servidor 2<sub>8</sub>. En una realización específica, los flujos pueden imponerse de forma conjunta mediante múltiples nodos de gestión 4 que se comunican entre sí para fines de coordinación.

## 6. Cálculo de la métrica de utilización

Para cada conmutador de red  $S_1$ - $S_{14}$ , se calcula una métrica de utilización que caracteriza al conmutador de red en términos de eficiencia energética. De forma intuitiva, la métrica expresa la utilización (aproximada) del conmutador de red, por ejemplo en términos del tráfico de red o el número de flujos que está procesando el conmutador de red.

El cálculo de la métrica de utilización puede desencadenarse con cualquier cambio en el estado de un conmutador de red. Este también puede realizarse solo cuando un nuevo trayecto para un flujo se determina mediante el módulo de PES 8.

## 7. Desactivación de Conmutadores

Ha de observarse que, en el contexto de las realizaciones que se ilustran, el PES no realiza la activación y desactivación real de los conmutadores, sino que deja estas acciones a las funciones de gestión existentes que controlan la red.

Con respecto a la desactivación de los conmutadores, esta se ocasiona finalmente por la disminución de las entradas de tabla de encaminamiento en los conmutadores individuales. Típicamente, tales entradas de tabla están asociadas con un tiempo de espera que expira si una entrada de tabla no se usa después de algún momento. Esto conduce a la eliminación de entradas únicas y, finalmente, de todas las entradas del conmutador. Entonces, es sencillo determinar cuándo puede desactivarse un conmutador. En el caso más simple, un conmutador puede desactivarse cuando ya no contiene entrada de tabla de encaminamiento alguna, es decir, cuando este no procesa paquete alguno. Como una realización adicional de la presente invención, las propiedades de tiempo de espera y de duración restante de las entradas de tabla de encaminamiento pueden considerarse al calcular las métricas de utilización (tal como se describe en la sección 6.) prefiriendo los conmutadores que tienen una longitud agregada (o máxima) de los valores de tiempo de espera.

Con respecto a la activación de los conmutadores, esta acción también es completamente transparente para el PES y se desencadena mediante cualquier módulo de cálculo (véase la figura 3) que determine que un conmutador ha de activarse para satisfacer, por ejemplo, los requisitos de QoS.

La activación y desactivación de los conmutadores puede conseguirse mediante diferentes medios. El encendido puede conseguirse, por ejemplo, mediante el uso de métodos de Wake-on-LAN, mediante alguna tecnología de gestión sin supervisión (LOM, *lights-out management*), o mediante el control de las salidas de potencia con las que los conmutadores están conectados directamente. El apagado puede usar LOM o también salidas de potencia, o métodos explícitos ofrecidos por el conmutador, por ejemplo a través de protocolos Web o SNMP.

De acuerdo con una realización preferida, el método tal como se describe en conexión con las figuras 2 y 3 se implementa con la tecnología de OpenFlow. En el presente caso, las funciones que se han descrito anteriormente se asignan tal como sigue:

1. Solicitud inicial: La solicitud inicial no se ve afectada por la presente realización. Esta aún se emite por el servidor de inicio.
2. Intercepción de solicitudes: La solicitud de conexión se recibe por el conmutador de OpenFlow lo más cerca del servidor de inicio. La solicitud se retransmite por el conmutador a su controlador de OpenFlow asociado. El controlador de OpenFlow realiza por lo tanto el módulo de intercepción del nodo de gestión tal como se describe anteriormente.
3. Selección de trayecto: La selección de trayecto no se ve afectada por la presente realización. Algún módulo particular está calculando aún un conjunto de trayectos potenciales a través de la red.
4. Cálculo de PES: En la presente realización, el cálculo de PES ocurre en el interior del controlador de OpenFlow.
5. Imposición de flujo: La imposición de flujo se realiza mediante los conmutadores de OpenFlow. El trayecto seleccionado se transforma en entradas de tabla de encaminamiento de OpenFlow apropiadas mediante el controlador de OpenFlow. Las entradas se propagan entonces desde el controlador de OpenFlow hasta los conmutadores de OpenFlow respectivos.
6. Cálculo de la métrica de utilización: El cálculo de las métricas de utilización es independiente de la intercepción y la imposición con la tecnología de OpenFlow. OpenFlow, sin embargo, posibilita el cálculo de las métricas de utilización que son particulares de la tecnología de OpenFlow. Una métrica de este tipo es el número de entradas de tabla de encaminamiento de OpenFlow (que pueden leerse a partir de los conmutadores). Esta métrica puede usarse directamente para el cálculo de PES, debido a que el número de entradas de tabla de encaminamiento está vagamente correlacionado con la carga de conmutación.

Pueden resumirse ventajas importantes de la presente invención tal como sigue:

- 1) Aprovechamiento del potencial de ahorro de energía en redes de comunicación sin comprometer el rendimiento de la red



- 2) Algoritmo (localizado) sumamente eficiente para la concentración de carga, conduciendo a a) la minimización del consumo de energía del propio algoritmo de gestión y b) la conservación de los requisitos de sincronismo que pueden existir con respecto a las operaciones de gestión anteriormente ejecutadas
- 3) Adaptabilidad a una variedad de granularidades de escenario a través de métricas, incluyendo específicas de puerto de conmutador, específicas de conmutador de red, específicas de segmento de red
- 4) Adaptabilidad a escenarios híbridos a través de métricas, incluyendo la consideración de
  - a. parámetros de QoS
  - b. Equipo de cálculo (IT), por ejemplo servidores en un centro de datos
  - c. Requisitos de aplicación.

En la actualidad, los enfoques son o bien de conocimiento pleno e ineficientes en la producción de configuraciones, o bien de conocimiento parcial e incompletos, ya que estos podrían comprometer el rendimiento de la red. El enfoque de la presente invención ni requiere cálculo complejo alguno ni compromete el rendimiento de la red, pero es aún efectivo dentro de los ahorros de energía alcanzables que son factibles en la red.

Si las métricas se seleccionan de forma no adecuada, los ahorros de energía podrían ser subóptimos. No obstante, es fácilmente posible intercambiar las métricas sobre la marcha, y evaluar la efectividad de las métricas para, por ejemplo, un centro de datos particular con su carga. Además, si la carga en la red es demasiado grande, en primer lugar no puede ahorrarse energía y, naturalmente, el algoritmo no será capaz de mejorar esta situación. Ha de observarse, no obstante, que esta segunda limitación solo implica una tara de gestión adicional pero baja (en forma de cálculos) por el algoritmo, debido a que este está aún en ejecución. No obstante, esta desventaja puede evitarse en el caso más simple mediante la desactivación del algoritmo de selección a una carga de red global determinada.

Muchas modificaciones y otras realizaciones de la invención que se expone en el presente documento acudirán a la mente el experto en la materia con la que está relacionada la invención, teniendo el beneficio de las enseñanzas que se presentan en la descripción precedente y los dibujos asociados. Por lo tanto, ha de entenderse que la invención no ha de limitarse a las realizaciones específicas divulgadas y que se pretende incluir modificaciones y otras realizaciones dentro del alcance de las reivindicaciones adjuntas. A pesar de que en el presente documento se emplean expresiones específicas, estas se usan solo en un sentido genérico y descriptivo y no para fines de limitación.

## REIVINDICACIONES

1. Red de comunicación, que incluye  
 5 una pluralidad de elementos de red - nodos de comunicación - que están habilitados para solicitar el establecimiento de conexiones - flujos - entre sí,  
 una pluralidad de elementos de red - conmutadores (3) - que soportan el encaminamiento de dichos flujos,  
 unos medios para interceptar solicitudes de establecimiento de flujo,  
 por lo menos un módulo de cálculo (7) que está configurado para seleccionar previamente, para cada flujo, un  
 10 conjunto de trayectos de red posibles que son compatibles con los requisitos de rendimiento y/o las restricciones de calidad de dicho flujo, en la que cada trayecto de red incluye un conjunto específico de conmutadores (3) que ha de recorrer dicho flujo, y  
 por lo menos un módulo de selección (8) que está configurado para aplicar un algoritmo de selección que analiza la carga de los conmutadores (3) que están contenidos en cualquier trayecto de red de dicho conjunto previamente  
 15 seleccionado de trayectos de red para determinar un único trayecto de red a partir de dicho conjunto previamente caracterizada por que dicho módulo de selección (8) está configurado para realizar dicho algoritmo de selección tal como sigue:
- 20 1) Crear el conjunto de conmutadores (3) que contiene todos los conmutadores (3) que son parte de cualquier trayecto de red de dicho conjunto previamente seleccionado de trayectos de red y apostillar cada uno de los conmutadores (3) con una métrica de utilización,
  - 2) Desmarcar todos los conmutadores (3) en todos los trayectos de red,
  - 3) Seleccionar, a partir del conjunto creado en 1), el conmutador (3) con el valor más alto de su métrica de utilización,
  - 25 4) Eliminar, con respecto al conjunto creado en 1), el conmutador seleccionado (3),
  - 5) Marcar el conmutador seleccionado (3) en todos los trayectos de red que contienen el mismo, y
  - 6) Si todos los conmutadores (3) en un trayecto de red están marcados, emitir como salida ese trayecto de red, en caso contrario, continuar con 3).
- 30 2. Red de acuerdo con la reivindicación 1, que incluye por lo menos un nodo de gestión (5) que funciona como entidad de coordinación para un dominio de gestión (4) que incluye por lo menos un subconjunto de dichos conmutadores (3).
- 35 3. Red de acuerdo con la reivindicación 2, en la que dicho por lo menos un nodo de gestión (5) está equipado con un módulo de intercepción (6) que está configurado para recibir dichas solicitudes interceptadas de establecimiento de flujo.
- 40 4. Red de acuerdo con la reivindicación 2 o 3, en la que dicho por lo menos un módulo de cálculo (7) y/o dicho por lo menos un módulo de selección (8) están incluidos en dicho por lo menos un nodo de gestión (5).
- 45 5. Red de acuerdo con cualquiera de las reivindicaciones 1 a 4, en la que dicho por lo menos un módulo de selección (8) está configurado para recibir como entrada dicho conjunto previamente seleccionado de trayectos de red a partir de dicho por lo menos un módulo de cálculo (7).
- 50 6. Red de acuerdo con cualquiera de las reivindicaciones 2 a 5, en la que dicho por lo menos un módulo de selección (8) está configurado para emitir como salida dicho único trayecto de red seleccionado a un módulo de control (9) de dicho por lo menos un nodo de gestión (5),  
 en la que dicho módulo de control (9) puede configurarse para imponer el flujo respectivo a través de dicho único trayecto de red seleccionado.
- 55 7. Red de acuerdo con cualquiera de las reivindicaciones 1 a 6, en la que dichos nodos de comunicación son servidores (2) de un centro de datos (1), y/o  
 en la que por lo menos uno de dichos conmutadores (3) se realiza como un conmutador de OpenFlow.
- 60 8. Red de acuerdo con cualquiera de las reivindicaciones 1 a 7, en la que dicho módulo de intercepción (6) de dicho por lo menos un nodo de gestión (5) se realiza como un controlador de OpenFlow, y/o  
 en la que dichos conmutadores (3) se realizan como puertos de conmutación, conmutadores de red o segmentos de red completos.
- 65 9. Método para accionar una red de comunicación, en particular una red de comunicación de acuerdo con cualquiera de las reivindicaciones 1 a 8, incluyendo dicha red una pluralidad de elementos de red - nodos de comunicación - que están habilitados para solicitar el establecimiento de conexiones - flujos - entre sí, y una pluralidad de elementos de red - conmutadores (3) - que soportan el encaminamiento de dichos flujos, comprendiendo el método las etapas de  
 interceptar solicitudes de establecimiento de los flujos,  
 para cada flujo, seleccionar previamente un conjunto de trayectos de red posibles que son compatibles con los

- requisitos de rendimiento y/o las restricciones de calidad de dicho flujo, en el que cada trayecto de red incluye un conjunto específico de conmutadores (3) que ha de recorrer dicho flujo, y aplicar un algoritmo de selección que analiza la carga de los conmutadores (3) que están contenidos en cualquier trayecto de red de dicho conjunto previamente seleccionado de trayectos de red para determinar un único trayecto de red a partir de dicho conjunto previamente seleccionado de trayectos de red que es el más adecuado en términos de eficiencia energética,
- 5 caracterizado por que dicho algoritmo de selección incluye las siguientes etapas:  
 apostillar cada uno de dichos conmutadores (3) que están contenidos en cualquier trayecto de red de dicho conjunto previamente seleccionado de trayectos de red con un valor escalar que está relacionado con su métrica de utilización,
- 10 a partir del conjunto de conmutadores (3) que están contenidos en cualquier trayecto de red de dicho conjunto previamente seleccionado de trayectos de red, seleccionar el conmutador (3) con la métrica de utilización más alta, eliminar dicho conmutador seleccionado (3) con respecto a dicho conjunto de conmutadores (3),
- 15 marcar el conmutador seleccionado (3) en todos los trayectos de red de dicho conjunto previamente seleccionado de trayectos de red que contienen dicho conmutador seleccionado (3),  
 en el caso en el que todos los conmutadores (3) de un trayecto de red preseleccionado están marcados, seleccionar ese trayecto de red preseleccionado como trayecto de red para el flujo respectivo.
- 20 10. Método de acuerdo con la reivindicación 9, en el que dicho algoritmo de selección se desencadena mediante un cambio de la configuración de flujo en dicha red de comunicación.
11. Método de acuerdo con la reivindicación 9 o 10, en el que la interceptación de solicitudes de establecimiento de flujo se realiza por el conmutador (3) que se encuentra lo más cerca del nodo de comunicación solicitante.
- 25 12. Método de acuerdo con cualquiera de las reivindicaciones 9 a 11, en el que las solicitudes interceptadas de establecimiento de flujo se reencaminan hacia un nodo de gestión (5) de dicha red de comunicación.
- 30 13. Método de acuerdo con cualquiera de las reivindicaciones 9 a 12, en el que dicha métrica de utilización de un conmutador (3) se calcula basándose en el tráfico de red y/o el número de flujos que se están procesando mediante dicho conmutador (3), y/o  
 en el que el cálculo de dicha métrica de utilización se desencadena mediante cualquier cambio en el estado de dicho conmutador (3).
- 35 14. Método de acuerdo con cualquiera de las reivindicaciones 9 a 13, en el que dicho algoritmo de selección se ejecuta sobre un conjunto de trayectos alternativos que tienen diferentes puntos de extremo de trayecto, y/o en el que dicho algoritmo de selección está configurado para seleccionar a partir de dicho conjunto previamente seleccionado de trayectos de red como trayecto de red para el flujo respectivo un trayecto de red que incluye solo conmutadores (3) con una métrica de utilización por encima de un umbral predefinido.
- 40 15. Método de acuerdo con cualquiera de las reivindicaciones 9 a 14, en el que dicho algoritmo de selección puede aplicarse a diferentes granularidades, en particular a puertos de conmutación, conmutadores de red o segmentos de red.

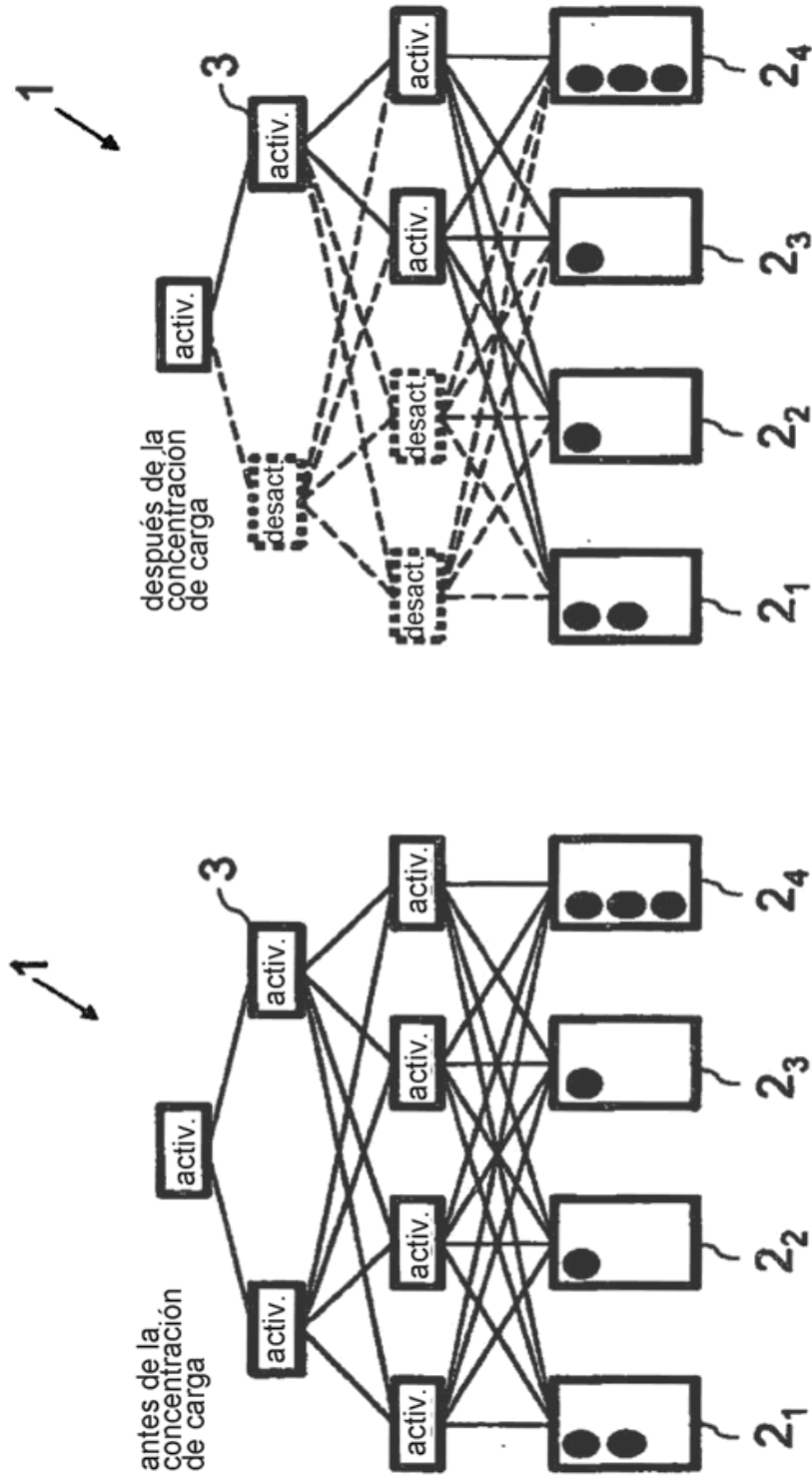


Fig. 1

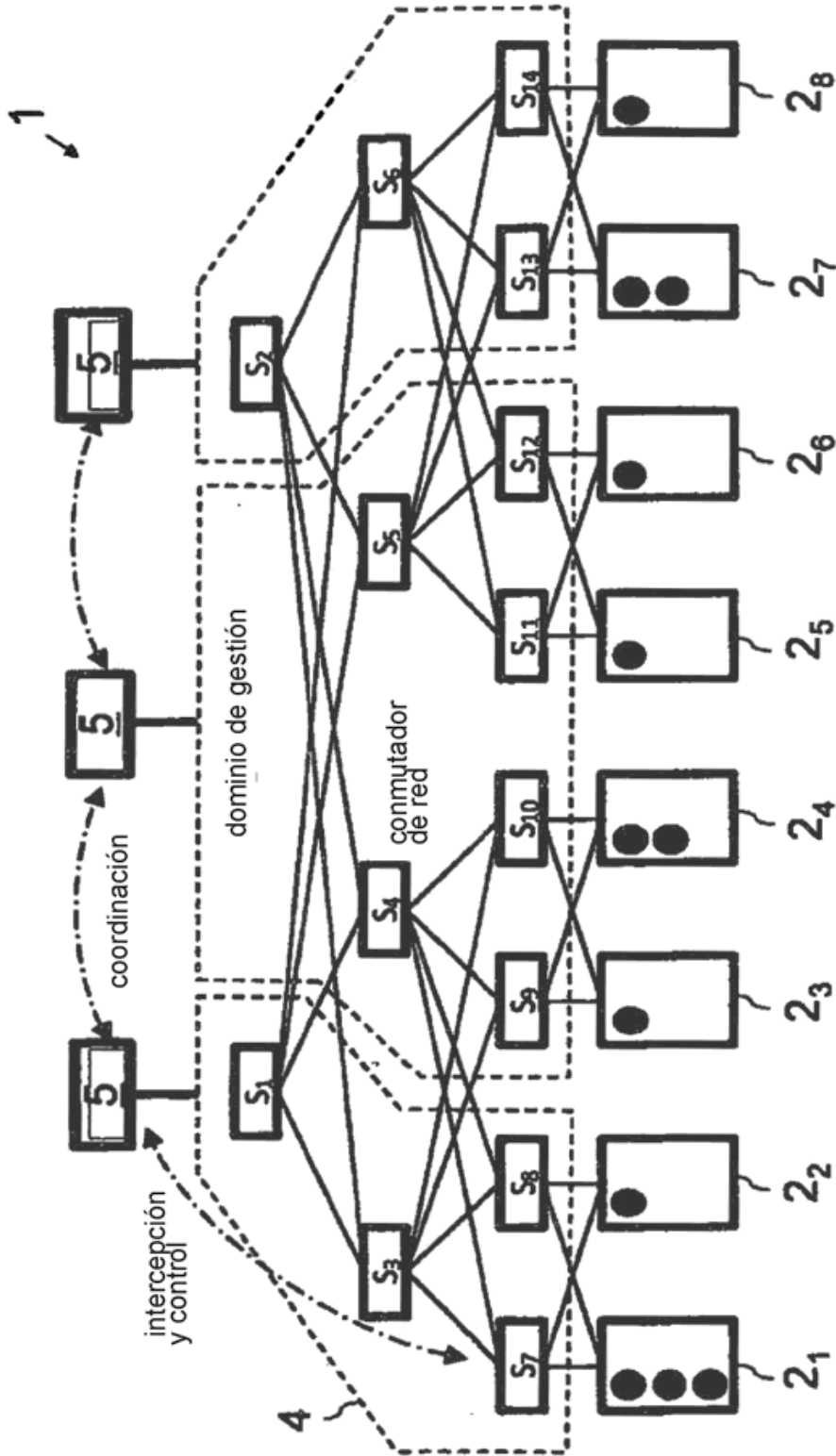


Fig. 2

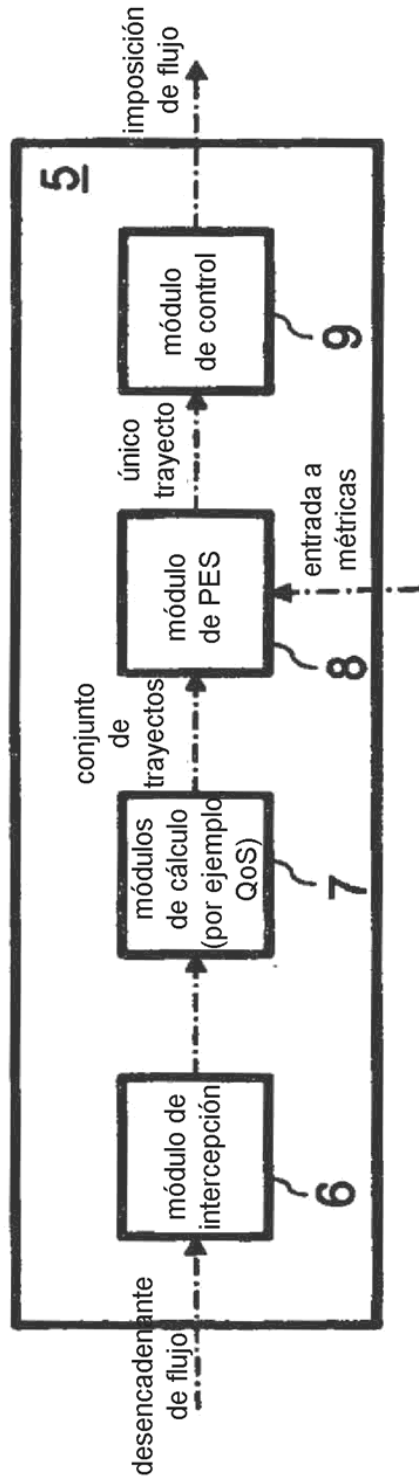


Fig. 3