



OFICINA ESPAÑOLA DE PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: 2 564 667

61 Int. Cl.:

H04L 12/26 (2006.01) H04L 12/24 (2006.01) H04L 12/721 (2013.01) H04L 12/751 (2013.01)

(12)

TRADUCCIÓN DE PATENTE EUROPEA

T3

- (96) Fecha de presentación y número de la solicitud europea: 15.08.2011 E 11818179 (1)
 (97) Fecha y número de publicación de la concesión europea: 20.01.2016 EP 2608461
- (54) Título: Dispositivo de comunicación, sistema de comunicación, método de comunicación y medio de registro
- (30) Prioridad:

17.08.2010 JP 2010182012

(45) Fecha de publicación y mención en BOPI de la traducción de la patente: 28.03.2016

(73) Titular/es:

NEC CORPORATION (100.0%) 7-1, Shiba 5-chome Minato-ku Tokyo 108-8001, JP

(72) Inventor/es:

SHIKITANI, NAOKI

74) Agente/Representante:

DE ELZABURU MÁRQUEZ, Alberto

DESCRIPCIÓN

Dispositivo de comunicación, sistema de comunicación, método de comunicación y medio de registro

Campo técnico

5

10

35

40

45

50

La presente invención se refiere a una unidad de comunicación, un sistema de comunicación, un método de comunicación y un medio de registro que miden un estado de una ruta de comunicación.

Antecedentes de la técnica

En los últimos años, en la Bibliografía no de patente 1 se ha propuesto, como técnica para realizar el control de rutas en una red de comunicaciones, una técnica denominada "OpenFlow". En la técnica OpenFlow, se considera la comunicación como un flujo de punta a punta, y el control de rutas, la recuperación de fallos, el equilibrio de carga y la optimización se realizan basándose en el flujo. Un "Conmutador OpenFlow" (u OFS, por sus siglas en inglés), que funciona como nodo de transferencia, posee un canal seguro para comunicarse con un "Controlador OpenFlow" (u OFC) y funciona conforme a una tabla de flujos que, además, es escrita o reescrita desde el OFC. En la tabla de flujos se define, para cada flujo, un conjunto de: una regla verificada frente a una cabecera de paquete, una acción que define detalles de tratamiento, y datos estadísticos de flujo.

Por ejemplo, tras recibir un paquete, el OFS busca en la tabla de flujos una entrada que contenga la regla ("Clave de flujo" o, en inglés, "FlowKey") que concuerde con los datos de la cabecera del paquete de recepción. Cuando, como resultado de la búsqueda, se encuentra la entrada que concuerda con el paquete de recepción, el OFS realiza, sobre el paquete de recepción, el tratamiento descrito en un campo "Acción" de la entrada. Por el contrario, si no se encuentra, como resultado de la búsqueda, la entrada que concuerde con el paquete de recepción, el OFS transfiere el paquete de recepción o los datos de cabecera del paquete de recepción al OFC a través del canal seguro. Como resultado de la transferencia, se solicita que se determine una ruta para el paquete basada en un origen y un destino del paquete de recepción y se recibe una entrada de flujo para realizar la determinación, con el fin de actualizar la tabla de flujos.

Listado de citas

25 [Bibliografía no de patente 1] OpenFlow Switch Specification Version 1.0.0 (Wire Protocol 0x01) (31 de diciembre de 2009) [buscado el 20 de julio de 2010] Internet <URL: http://www.openflowswitch.org/documents/openflow-spec-v1.0.0.pdf>

[Bibliografía no de patente 2] ITU-T Recommendation Y.1731

Compendio de la invención

30 En la red que utiliza la técnica OpenFlow descrita en la Bibliografía no de patente 1, cuando se producen un fallo de ruta, una conjunción, etcétera, y se hace necesaria la conmutación de ruta, generalmente se establece que el OFC realice la reconfiguración de topología y el cálculo de ruta para estipular una tabla de flujos de cada OFS.

Sin embargo, en el método antes mencionado representa un problema el hecho de que se tarda tiempo hasta que se estipula la nueva ruta a los OFS y se conmuta la ruta, ya que es necesario realizar la reconfiguración de topología y el cálculo de ruta en el OFC cada vez que falla la ruta y se produce una congestión.

Es un objeto de la presente invención proporcionar una unidad de comunicación, un sistema de comunicación, un método de comunicación y un programa de comunicación que puedan solventar el problema antes mencionado.

En un aspecto de la presente invención, una unidad de comunicación que pertenece a una red incluye: una sección de adición que añade datos de medición del estado de comunicación a una trama de recepción cuando la unidad de comunicación es un nodo de borde de entrada de la red; una sección de medición que mide un estado de comunicación basándose en los datos de medición del estado de comunicación cuando la unidad de comunicación es un nodo de borde de salida de la red; una sección de notificación de resultados de medición que notifica un resultado de medición del estado de comunicación a una unidad de control que controla la red; y una sección de tratamiento que consulta datos de identificación de la trama de recepción para realizar el tratamiento de la trama de recepción basándose en una regla de tratamiento que relaciona los datos de identificación de la trama de recepción con el tratamiento para la trama de recepción.

En otro aspecto de la presente invención, un sistema de comunicación incluye: la unidad de comunicación antes mencionada; una unidad de control que incluye: una sección de cálculo de rutas que calcula una ruta de la trama de recepción a partir del resultado de medición recibido desde la unidad de comunicación; una sección de almacenamiento de rutas que almacena la ruta calculada; y una sección de transferencia de rutas que estipula la regla de tratamiento de la trama a la unidad de comunicación de la ruta calculada, basándose en la ruta calculada que está almacenada en la sección de almacenamiento de rutas.

En otro aspecto de la presente invención, un método de comunicación incluye: añadir datos para medir un estado de

comunicación, a una trama de recepción cuando una unidad de comunicación perteneciente a una red es un nodo de borde de entrada de la red; medir el estado de comunicación basándose en los datos de medición del estado de comunicación cuando la unidad de comunicación es un nodo de borde de salida de la red; notificar un resultado de medición del estado de comunicación a una unidad de control que controla la red; y consultar datos de identificación de la trama de recepción para realizar el tratamiento de la trama de recepción basado en una regla de tratamiento que relaciona los datos de identificación de la trama de recepción con el tratamiento para la trama de recepción.

En otro aspecto de la presente invención, se proporciona un medio de registro no transitorio en el cual se almacena un programa que hace que una unidad de comunicación que pertenece a una red ejecute un proceso de comunicación. El proceso incluye: añadir datos para medir un estado de comunicación a una trama de recepción cuando la unidad de comunicación perteneciente a la red es un nodo de borde de entrada de la red; medir el estado de comunicación basándose en los datos de medición del estado de comunicación cuando la unidad de comunicación es un nodo de borde de salida de la red; notificar un resultado de medición del estado de comunicación a una unidad de control que controla la red; y consultar datos de identificación de la trama de recepción para realizar el tratamiento de la trama de recepción basándose en una regla de tratamiento que relaciona los datos de identificación de la trama de recepción con el tratamiento para la trama de recepción.

Según la presente invención, se hace posible conmutar la ruta a gran velocidad mediante un servidor de control en función del estado de comunicación de la red.

Breve descripción de los dibujos

5

10

15

La Figura 1 es un diagrama de bloques que muestra una configuración de un sistema según una primera realización ilustrativa:

la Figura 2 es un diagrama de bloques que muestra una configuración del sistema según la primera realización ilustrativa;

la Figura 3 es un diagrama de bloques que muestra una configuración de una sección de almacenamiento según la primera realización ilustrativa;

25 la Figura 4 es un diagrama de bloques que muestra una configuración de una sección de almacenamiento de datos de control según la primera realización ilustrativa;

la Figura 5 muestra un formato de trama en la primera realización ilustrativa;

la Figura 6 muestra una tabla de flujos según la primera realización ilustrativa;

la Figura 7 muestra una tabla de PBB según la primera realización ilustrativa;

30 la Figura 8 muestra una tabla de estados de comunicación según la primera realización ilustrativa;

la Figura 9 muestra una tabla de búsqueda de MAC según la primera realización ilustrativa;

la Figura 10 es un diagrama de secuencia que muestra una operación de la primera realización ilustrativa;

la Figura 11 es un diagrama de secuencia que muestra una operación de la primera realización ilustrativa;

la Figura 12 es un diagrama de secuencia que muestra una operación de la primera realización ilustrativa;

35 la Figura 13 es un diagrama de secuencia que muestra una operación de la primera realización ilustrativa;

la Figura 14 es un diagrama de bloques que muestra una configuración del sistema según una segunda realización ilustrativa; y

la Figura 15 es un diagrama de bloques que muestra una configuración del sistema según una tercera realización ilustrativa.

40 Descripción de realizaciones preferidas

[Primera realización ilustrativa]

A continuación se describirá con detalle una primera realización ilustrativa de la presente invención, haciendo referencia a los dibujos adjuntos.

(Configuración global)

La Figura 1 es un diagrama de bloques que muestra una configuración de un sistema según la primera realización ilustrativa. El sistema mostrado en la Figura 1 incluye una red 150, una red 160 y una red 170. La red 150 incluye un conmutador 100, un conmutador 120, un conmutador 130, un conmutador 140 y un servidor 110 de control.

A continuación se describirá en lo fundamental una operación de la red 150. Las flechas de la Figura 1 muestran una ruta de un paquete (el conmutador 100 - el conmutador 120 - el conmutador 130 - el conmutador 140) que se describirá en la realización ilustrativa. El servidor 110 de control y los conmutadores están conectados como se muestra por líneas de puntos. La conexión puede ser realizada por la misma red que conecta los conmutadores o por distintas líneas dedicadas.

Además, en la red 150 el conmutador 100 y el conmutador 140, conectados respectivamente con la red 160 y con la red 170, funcionan respectivamente como un nodo de borde de entrada y como un nodo de borde de salida.

Cuando se aplica la técnica OpenFlow descrita en la Bibliografía no de patente 1 a la red 150 mostrada en la Figura 1, los conmutadores 100 a 140 corresponden a conmutadores OFS y el servidor 110 de control corresponde a un controlador OFC, respectivamente.

A continuación se expondrá la descripción usando como ejemplo la técnica OpenFlow, pero el ámbito de aplicación de la presente realización ilustrativa no está limitado a la técnica OpenFlow. Se puede aplicar una técnica para controlar de forma centralizada una red, como la técnica OpenFlow.

La Figura 2 es un diagrama de bloques que muestra una configuración del conmutador 100 y del servidor 110 de control según la primera realización ilustrativa. Debe señalarse que en la Figura 2 sólo se muestra la configuración del conmutador 100, ya que se ha omitido la descripción de los conmutadores 120, 130 y 140, mostrados en la Figura 1, por tener la misma configuración que el conmutador 100.

(Funciones de secciones del conmutador)

5

10

25

30

35

50

55

El conmutador 100 incluye una sección 101 de medición, una sección 102 de notificación de resultados de medición, una sección 103 de adición, una sección 104 de tratamiento, una sección 105 de almacenamiento, una sección 106 de interfaz de servidor de control, una sección 107 de interfaz de red OpenFlow y una sección 108 de interfaz de redes.

La sección 101 de medición mide un estado de comunicación basándose en un número de secuencia incluido en una trama de recepción, un tiempo de transmisión de trama, etcétera. Más específicamente, se mide un estado de comunicación mediante el cálculo de una tasa de pérdida de tramas, un tiempo medio de retardo y una velocidad media de recepción. Hay que señalar que la medición se realiza cuando el conmutador está situado en un nodo de borde de salida de una red. En el caso de la red 150 mostrada en la Figura 1, la medición se realiza en el conmutador 140.

Cuando en la sección 101 de medición se ha realizado la medición del estado de comunicación, la sección 102 de notificación de resultados de medición transmite un resultado de medición al servidor 110 de control a través de la sección 106 de interfaz de servidor de control.

La sección 103 de adición añade datos necesarios para la transferencia en la red 150, además de los datos de medición del estado de comunicación (un número de secuencia y un tiempo de transmisión de un trama). En la realización ilustrativa, esto se consigue utilizando un método denominado PBB (o de "Puentes de red principal de proveedor", por sus siglas en inglés). Sin embargo, el ejemplo de aplicación de la realización ilustrativa no está limitado al método PBB, y también es posible la aplicación de una técnica de encapsulación de datos de recepción (EoE, o "Ethernet (marca registrada) sobre Ethernet"). Además, también es posible emplear un método para simplemente añadir a una trama de recepción los datos de medición del estado de comunicación y los datos necesarios para la transferencia.

40 En cuanto al método PBB, se está realizando trabajo de normalización en el IEEE (Institute of Electrical and Electronics Engineers) como técnica para que una red principal agrupe redes mediante el uso del método PB ("Puentes de proveedor", por sus siglas en inglés) que es un método de comunicación para operadores de telecomunicaciones.

En general, se provee un nodo de borde en una frontera entre una red PBB y una red PB. En el nodo de borde, una trama recibida desde una red PB se convierte en una trama "MAC-en-MAC" (MAC son las siglas inglesas de "Control de acceso al medio"), y la comunicación mediante trama MAC-en-MAC se realiza en una red PBB.

En la presente realización ilustrativa se describirá un caso en donde se asignan respectivamente las redes 160 y 170 de la Figura 1, y la red 150, a una red PB y a una red PBB. Debe señalarse que las redes 160 y 170 no están limitadas a la red PB, sino que se pueden asignar a cualquier red. La Figura 5 muestra un formato de trama utilizado en la presente realización ilustrativa. La trama mostrada en la Figura 5 se obtiene mediante encapsulación, por la sección 103 de adición, de un trama de recepción (datos originales) que incluye una "MAC dest." (dirección MAC de destino), una "MAC orig." (dirección MAC de origen), un "Tipo" (tipo de protocolo o, en inglés, "Ether Type"), una UDP (Unidad de datos de protocolo) y una SVT (Secuencia de verificación de trama, es decir, detección de errores de trama).

Como resultado de la encapsulación se añade una cabecera de PBB a los datos originales, tal como se muestra en

la Figura 5. La cabecera de PBB incluye una "MAC dest. de red principal" (dirección MAC de destino de red principal, que en lo que sigue se denominará "dirección B-MAC de destino", donde la "B" corresponde al inglés "Backbone", por "Red principal"), una "MAC orig. de red principal" (dirección MAC de origen de red principal, que en lo que sigue se denominará análogamente "dirección B-MAC de origen"), una "B-TAG" (o "Etiqueta de VLAN de red principal", donde "VLAN" son las siglas inglesas de "Red de área local virtual"), y una "I-TAG" (o "Etiqueta de instancia de servicio"). La dirección B-MAC de destino indica una dirección MAC de destino de una trama utilizada en la red 150, y la dirección B-MAC de origen indica una dirección MAC de origen de una trama en la red 150. La B-TAG almacena un "B-VID" (abreviatura de "B-VLAN ID", o "Identificador de red de área local virtual de red principal"), que es un identificador de una ruta utilizada en la red 150.

- Se describirá la I-TAG utilizada en la presente realización ilustrativa. Tal como se muestra en la Figura 5, la I-TAG incluye un "I-TAG TPID" (abreviatura del inglés "I-TAG Protocol ID", o "Identificador de protocolo de I-TAG"), un "I-PCP" (abreviatura de "I-TAG Priority Code Point", o "Punto de prioridad de código de I-TAG"), un "I-DEI" (abreviatura de "I-TAG Drop Elligible Indication", o "Indicación elegible de entrega de I-TAG"), reserva y un "I-SID" (abreviatura de "Service Instance ID", o "Identificador de instancia de servicio").
- En la presente realización ilustrativa, a un campo del I-SID se añaden, tal como se muestra en la Figura 5, un "ID de flujo" ("Identificador de flujo"), un número de secuencia y datos de tiempo de transmisión. El ID de flujo, que se añade en el nodo de borde de entrada de la red 150 (el conmutador 100 de la Figura 1), es, en la técnica OpenFlow, un identificador basado en el flujo. El servidor 110 de control controla el ID de flujo.
- El número de secuencia se añade en el nodo de borde de entrada de la red 150. El número de secuencia es un valor numérico que se incrementa en una unidad cada vez que se transmite una trama que recorre el mismo camino (una trama con el mismo flujo). El nodo de borde de salida de la red 150 (el conmutador 140 de la Figura 1) inspecciona el valor numérico, y mide la velocidad de recepción y la pérdida de tramas en una ruta.
 - Los datos de tiempo de transmisión se añaden en el nodo de borde de entrada de la red 150. Los datos de tiempo de transmisión se toman de un "RTC" ("Reloj de tiempo real", por sus siglas en inglés; no se muestra en los dibujos) que mide el tiempo en un dispositivo, por ejemplo. En el nodo de borde de salida, el tiempo de retardo en una ruta se mide comparando los datos de tiempo de transmisión añadidos a una trama y los datos de hora real tomados del RTC del dispositivo.

25

30

35

40

- Es posible seleccionar y añadir, o bien un número de secuencia, o bien datos de tiempo de transmisión, para que corresponden a lo estipulado acerca de cuál de los parámetros: tasa de pérdida de tramas, tiempo medio de retardo y velocidad media de recepción, se mide. Por ejemplo, la medición es factible mediante la adición del número de secuencia cuando sólo se mide la tasa de pérdida de tramas, y de los datos de tiempo de transmisión cuando se miden el tiempo medio de retardo o la velocidad media de recepción.
- La sección 104 de tratamiento trata la trama de recepción según una regla (entrada) de tratamiento correspondiente a la trama de recepción almacenada en una tabla 105-1 de flujos de la sección 105 de almacenamiento. Debe señalarse que la regla de tratamiento corresponde a una entrada de la tabla de flujos de la técnica OpenFlow. Los detalles de la tabla 105-1 de flujos se describirán más adelante.
- Más concretamente, en primer lugar se busca en la tabla 105-1 de flujos una regla de tratamiento que corresponda a una trama de recepción. Cuando existe dicha regla de tratamiento en la tabla 105-1 de flujos, se realiza el tratamiento indicado. El tratamiento corresponde a una acción ("Action") en OpenFlow. En la presente realización ilustrativa se supone típicamente que el tratamiento se transfiere al siguiente conmutador dentro de una ruta de transferencia de la trama de recepción, pero no está limitado a la transferencia. Son ejemplos de tratamiento distintos de la transferencia el "unicast", el "multicast", el control de descarte, el control de equilibrio de carga, el control de recuperación de fallos, el control de transferencia por túnel a puerto virtual y el encriptado.
- Cuando no existe dicha regla de tratamiento en la tabla 105-1 de flujos, a través de la sección 106 de interfaz de servidor de control se transmite una "pregunta" al servidor 110 de control acerca del tratamiento de la trama de recepción. Esta operación corresponde a "Packet-in" en OpenFlow.
 - La Figura 3 es un diagrama de bloques que muestra detalles de la sección 105 de almacenamiento. En la sección 105 de almacenamiento se almacenan la tabla 105-1 de flujos que se muestra en la Figura 6 y una tabla 105-2 de PBB que se muestra en la Figura 7.
- 50 En primer lugar, en la tabla 105-1 de flujos se almacena una entrada en la que se relacionan una clave de búsqueda para cada flujo y el tratamiento (la acción). "ID de flujo" indica un identificador de un flujo como se ha descrito más arriba. "Puerto de entrada" indica un puerto de entrada de un trama. "MAC dest." indica una dirección MAC de destino de una trama. "Ether ID" indica un tipo de protocolo ("Ether Type") de una trama. "VLAN ID" indica un identificador de VLAN de una trama. "Prioridad VLAN" indica una prioridad de una trama. "IP orig." indica una dirección IP (Protocolo de Internet) de origen de un trama. "IP dest." indica una dirección IP de destino de una trama. "IP prot." indica un tipo de protocolo IP de una trama. Los bits "IP ToS bits" indican el ToS ("Tipo de servicio", por sus siglas en inglés) de IP de una trama. "Puerto orig. PCT/PDU" indica un número de puerto de origen de PCT/PDU (Protocolo de control de transmisión/Protocolo de

datos de usuario") de un trama. "Puerto dest. PCT/PDU" indica un número de puerto de destino de PCT/PDU de un trama. "Acción" indica los detalles de tratamiento para un trama correspondiente.

Por ejemplo, se pueden seleccionar una dirección MAC de origen, una dirección MAC de destino y un VLAN ID (Identificador de red local virtual) como clave de búsqueda de un flujo. Debe tenerse en cuenta que la tabla 105-1 de flujos corresponde a una tabla de flujos en OpenFlow.

Tal como se muestra en la Figura 7, en la tabla 105-2 de PBB se almacenan un ID de flujo, una dirección B-MAC de destino, una dirección B-MAC de origen y un B-VID.

La sección 106 de interfaz de servidor de control es una interfaz para comunicación entre el conmutador 100 y el servidor 110 de control. En OpenFlow, la comunicación se lleva a cabo a través de un canal seguro.

La sección 107 de interfaz de red OpenFlow es una interfaz de comunicación con un nodo (el conmutador 120 de la Figura 1) de una red (la red 150 de la Figura 1) a la que se aplica OpenFlow.

La sección 108 de interfaz de redes es una interfaz de comunicación con redes (la red 160 y la red 170 de la Figura 1) distintas de una red OpenFlow.

La sección 101 de medición, la sección 102 de notificación de resultados de medición, la sección 103 de adición, la sección 104 de tratamiento, la sección 106 de interfaz de servidor de control, la sección 107 de interfaz de red OpenFlow y la sección 108 de interfaz de redes pueden implementarse en *hardware*, o bien pueden implementarse en forma de una combinación de *software* y una unidad procesadora para ejecutar el *software*. El *software* a implementar se puede instalar en el conmutador 100 mediante el uso de un medio 100a de registro no transitorio para almacenar el *software*. La sección 105 de almacenamiento puede implementarse en forma de cualquier dispositivo de almacenamiento, por ejemplo una memoria de semiconductor.

(Funciones de secciones del servidor de control)

5

15

20

35

45

El servidor 110 de control incluye una sección 111 de cálculo de rutas, una sección 112 de transferencia de rutas y una sección 113 de almacenamiento de datos de control.

La sección 111 de cálculo de rutas consulta la sección 113 de almacenamiento de datos de control y calcula una ruta de un flujo basándose en un algoritmo dado, conforme a los datos recibidos. Más concretamente, la sección 111 de cálculo de rutas recibe en primer lugar un resultado de medición de un estado de comunicación o bien un mensaje "packet-in" desde el conmutador 100. A continuación, la sección 111 de cálculo de rutas consulta los datos de topología, etcétera, de la red 150 almacenados en la sección 113 de almacenamiento de datos de control, y calcula una ruta de flujo apropiada basándose en el algoritmo dado. Para el cálculo de ruta se puede utilizar cualquier algoritmo.

A través de la sección 106 de interfaz de servidor de control, la sección 112 de transferencia de rutas notifica una regla de tratamiento que corresponde a una ruta calculada por la sección 111 de cálculo de rutas a la sección 105 de almacenamiento en cada conmutador de la ruta. Al mismo tiempo, la sección 112 de transferencia de rutas notifica datos necesarios para la encapsulación de una cabecera de PBB (por ejemplo, una dirección B-MAC de destino, una dirección B-MAC de origen y un B-VID) al conmutador 100. Esta operación corresponde a "Flow_mod" en OpenFlow.

La Figura 4 es un diagrama de bloques que muestra detalles de la sección 113 de almacenamiento de datos de control. La sección 113 de almacenamiento de datos de control incluye una tabla 113-1 de topologías, una tabla 113-2 de estados de comunicación y una tabla 113-3 de búsqueda de MAC.

En la tabla 113-1 de topologías se almacenan datos de topología de la red 150 controlada por el servidor 110 de control. En la tabla de estados de comunicación se almacena un estado de comunicación para cada ruta de comunicación (flujo). No se describirá con detalle la tabla 113-1 de topologías, ya que puede tener una estructura de datos cualquiera.

La Figura 8 muestra detalles de la tabla 113-2 de estados de comunicación. En la tabla 113-2 de estados de comunicación se almacenan una "Pérdida de tramas", un "Tiempo medio de retardo" y una "Velocidad media", además de una dirección B-MAC de destino, una dirección B-MAC de origen y un B-VID. La pérdida de tramas, que se expresa en unidades de porcentaje, indica la tasa de pérdida de tramas de una ruta correspondiente. El tiempo medio de retardo, que se expresa en µs (microsegundos), indica el tiempo medio de retardo de trama de una ruta correspondiente. La velocidad media indica la velocidad media de recepción de tramas, expresada en "tps" (tramas por segundo), en una ruta correspondiente.

La Figura 9 muestra detalles de la tabla 113-3 de búsqueda de MAC. La tabla 113-3 de búsqueda de MAC relaciona y almacena MAC incluidas en una trama de recepción procedente de fuera de la red 150 y una dirección B-MAC para el PBB utilizado en la red 150. La tabla 113-3 de búsqueda de MAC se genera cuando se establece la red 150.

La sección 111 de cálculo de rutas y la sección 112 de transferencia de rutas pueden implementarse en *hardware*, o bien pueden implementarse en forma de una combinación de *software* y una unidad procesadora para ejecutar el

software. El software a implementar se puede instalar en el servidor 110 de control mediante el uso de un medio 110a de registro no transitorio para almacenar el software. La sección 113 de almacenamiento de datos de control puede implementarse en forma de cualquier dispositivo de almacenamiento, por ejemplo una HDD (unidad de disco duro, por sus siglas en inglés) y una memoria de semiconductor.

5 (Operación)

10

15

20

25

30

45

Se describirá con detalle una operación según la presente realización ilustrativa, haciendo referencia a las Figuras 10 a 13. Las Figuras 10 y 11 son diagramas de secuencia que muestran la operación del conmutador 100 como nodo de borde de entrada y del servidor 110 de control. La Figura 12 es un diagrama de secuencia que muestra la operación del conmutador 120, del conmutador 130 y del servidor 110 de control. La Figura 13 es un diagrama de secuencia que muestra la operación del conmutador 140 como nodo de borde de salida y del servidor 110 de control.

(Operación de nodo de borde de entrada)

Se describirá en primer lugar la operación del conmutador 100 como nodo de borde de entrada de la red 150, haciendo referencia a las Figuras 10 y 11. En primer lugar, el conmutador 100 recibe una trama desde un nodo de la red 160 (no mostrado) (Figura 10: paso S101). El formato de una trama recibida aquí (denominada en adelante un "trama de recepción") corresponde a los "datos originales" del formato de trama mostrado en la Figura 5.

A continuación, la sección de tratamiento 104 busca en la tabla 105-1 de flujos una entrada correspondiente a la trama de recepción, utilizando como clave los datos almacenados en un encabezado de la trama de recepción (Figura 10: paso S102). Cuando el resultado de la búsqueda es que existe la entrada correspondiente, se realiza a continuación un paso S108 de la Figura 11 que se describirá más adelante. Cuando el resultado de la búsqueda es que no existe ninguna entrada correspondiente, se transfieren los datos de cabecera de la trama de recepción al servidor 110 de control (Figura 10: paso S103).

A continuación, se busca en el servidor 110 de control una dirección B-MAC de destino correspondiente, utilizando como clave una dirección MAC de destino (MAC dest.) almacenada en los datos de cabecera de la trama de recepción (Figura 10: paso S104).

A continuación, la sección 111 de cálculo de rutas determina una ruta basándose en los datos de cabecera de la trama de recepción, un estado de comunicación de una ruta relevante, etcétera, utilizando un algoritmo dado. La ruta determinada se almacena en la tabla 113-2 de estados de comunicación (Figura 10: paso S105).

La sección 112 de transferencia de rutas notifica a la tabla 105-1 de flujos de cada conmutador de la ruta una regla de tratamiento correspondiente a la ruta determinada. Al mismo tiempo, la sección 112 de transferencia de rutas notifica una dirección B-MAC de origen, una dirección B-MAC de destino, un B-VID y un ID de flujo de la ruta determinada a la tabla 105-2 de PBB (Figura 10: paso S106).

A continuación, el conmutador 100 actualiza la tabla 105-1 de flujos y la tabla 105-2 de PBB basándose en los datos notificados en el paso S106 (Figura 11: paso S107).

Después de actualizar cada tabla, la sección 103 de adición toma la hora actual del RTC (Figura 11: etapa S108) y encapsula una B-MAC, una B-TAG y una I-TAG, para transmitir una trama (Figura 11: etapa S109).

Por último, se incrementa un número de secuencia y se almacena en el conmutador 100.

(Operación de nodo de retransmisión)

Se describirá ahora una operación del conmutador 120 como nodo de retransmisión en la red 150, haciendo referencia a la Figura 12. Se ofrece a continuación la descripción de la operación del conmutador 120, y el conmutador 130 opera de la misma manera.

En primer lugar, el conmutador 120 recibe una trama desde el conmutador 100 (Figura 12: paso S111). Después, la sección 104 de tratamiento busca en la tabla 105-1 de flujos (Figura 12: paso S112). Cuando existe una entrada correspondiente, se realiza el tratamiento de la trama de recepción conforme a la entrada (Figura 12: paso S113). En la presente realización ilustrativa, se proporciona el tratamiento para transferir la trama de recepción al siguiente conmutador (conmutador 130) de una ruta de transferencia. Si no existe la entrada correspondiente, se formula una pregunta al servidor 110 de control. En este caso, puesto que se presupone que ya está estipulada la entrada correspondiente, se omite la descripción,

(Operación de nodo de borde de salida)

50 Se describirá por último, haciendo referencia a la Figura 13, una operación del conmutador 140 como nodo de borde de salida de la red 150.

En primer lugar, el conmutador 140 recibe una trama desde el conmutador 130 (Figura 13: paso S121).

A continuación, la sección 101 de medición compara un número de secuencia almacenado en la trama de recepción con un historial de números de secuencia almacenados en el conmutador 140, y almacena el valor de (diferencia - 1) (Figura 13: paso S122).

La sección 101 de medición compara los datos de tiempo de transmisión almacenados en la trama de recepción con los datos de hora actual tomados del RTC, y almacena una diferencia (Figura 13: paso S123).

La sección 101 de medición compara los datos de tiempo almacenados en el momento de la anterior recepción de trama con los datos de hora actual tomados del RTC, y almacena una diferencia (Figura 13: paso S124).

A continuación, la sección 104 de tratamiento desencapsula la cabecera de PBB de la trama de recepción, y transfiere la cabecera de PBB a un nodo correspondiente de la red 170 conforme a una dirección MAC de destino almacenada en los datos originales después de la desencapsulación (Figura 13: paso S125).

Tras la operación hasta el paso S125, el conmutador 140 determina si se han recibido N tramas del mismo flujo (Figura 13: paso S126). N es un valor opcional, y puede ser introducido por un operador de la red 150 que maneja el servidor 110 de control, por ejemplo. Cuando se han recibido las N tramas del mismo flujo, se lleva a cabo un paso S127. Si no se han recibido N tramas del mismo flujo, no se realiza el paso S127 hasta recibir de nuevo una trama.

Cuando en el paso S126 se determina que se han recibido las N tramas del mismo flujo, se realiza el paso S127. En el paso S127, la sección 101 de medición calcula una tasa de pérdida de tramas, un tiempo medio de retardo y una velocidad media de recepción a partir de un número de secuencia y datos de tiempo de transmisión.

La tasa de pérdida de tramas se calcula aquí dividiendo la suma de valores, en el historial, de (diferencia - 1) de números de secuencia almacenados en el paso S122, por el número N de tramas de recepción. El tiempo medio de retardo se calcula dividiendo la suma de las diferencias de tiempo, en el historial, almacenadas en el paso S123, por el número N de tramas de recepción. La velocidad media de recepción se calcula dividiendo el inverso de la diferencia de tiempo almacenada en el paso S124, por el número N de tramas de recepción.

Por último en el conmutador 140, la sección 102 de notificación de resultados de medición transfiere el resultado de medición calculado en el paso S127 al servidor 110 de control (paso S128).

En el servidor 110 de control que ha recibido el resultado de medición, la sección 111 de cálculo de rutas calcula una ruta basándose en un algoritmo dado, mediante consulta de la sección 113 de almacenamiento de datos de control, y almacena la ruta calculada en la tabla 113-1 de topologías y en la tabla 113-2 de estados de comunicación (paso S129).

Se describirá aquí un método para determinar si un conmutador cualquier es un nodo de borde de entrada o un nodo de borde de salida.

Existen diversos métodos para averiguar si un conmutador determinado es un nodo de borde de entrada. Por ejemplo, existe un método en el que se consulta un campo específico de una trama de recepción. Por ejemplo, tal como ocurre en la presente realización ilustrativa, cuando se aplica el método PBB se asigna una etiqueta denominada S-TAG (o "Etiqueta de VLAN de servicio") (no mostrada en la Figura 5) a los datos originales de la Figura 5. Mediante la consulta de la etiqueta, un conmutador puede determinar si el conmutador es un nodo de borde de entrada. También existe otro método para determinar si un destino de conexión de un puerto que recibe la trama es el conmutador que tiene la misma función, mediante el uso de un LLDP (Protocolo de descubrimiento de capa de enlace, por sus siglas en inglés). También es posible emplear otro método para avisar de que un conmutador correspondiente es un nodo de borde de entrada, desde el servidor de control 110, utilizando el hecho de que el servidor 110 de control conoce la topología.

Después, se debe consultar una dirección B-MAC de destino de una trama de recepción para averiguar si un conmutador determinado es o no un nodo de borde de salida. Cuando la dirección B-MAC de destino de la trama de recepción coincide con una dirección MAC del conmutador, el conmutador puede determinar que el propio conmutador es un nodo de borde de salida.

45 (Efecto)

5

10

20

35

40

Según la presente realización ilustrativa como se ha descrito en lo que antecede, cuando el nodo de borde de entrada de una red recibe una trama se añaden datos de medición del estado de comunicación, y se mide un estado de comunicación en un nodo de borde de salida para notificar al servidor de control, y el servidor de control actualiza una ruta

50 Mediante esta operación el servidor de control puede supervisar el estado de una red. En consecuencia, es posible realizar conmutación de rutas a gran velocidad en respuesta a la aparición de un fallo y el descenso de la calidad de línea.

[Segunda realización ilustrativa]

(Configuración y operación)

10

15

20

25

30

35

A continuación se describirá con detalle una segunda realización ilustrativa de la presente invención, haciendo referencia a la Figura 14. La Figura 14 es un diagrama de bloques que muestra una configuración de un sistema según la presente realización ilustrativa.

5 Un sistema según la presente realización ilustrativa incluye un conmutador 200 y un servidor 210 de control. La configuración de la red es la misma que en la Figura 1, y los conmutadores y el servidor de control han sido reemplazados por los de la Figura 14.

El conmutador 200 incluye una sección 201 de medición, una sección 202 de notificación de resultados de medición, una sección 203 de adición, una sección 204 de tratamiento, una sección 205 de almacenamiento, una sección 206 de interfaz de servidor de control, una sección 207 de interfaz de red OpenFlow, una sección 208 de interfaz de redes y una sección 209 de OAM ("Operación, Administración, Mantenimiento"). Las secciones que no son la sección 209 de OAM son las mismas que las secciones del conmutador 100 en la primera realización ilustrativa mostrada en la Figura 2, y se omite su descripción. La sección 209 de OAM se describirá más adelante. Al igual que en la primera realización ilustrativa, la sección 201 de medición, la sección 202 de notificación de resultados de medición, la sección 203 de adición, la sección 204 de tratamiento, la sección 206 de interfaz de servidor de control, la sección 207 de interfaz de red OpenFlow, la sección 208 de interfaz de redes y la sección 209 de OAM pueden implementarse en hardware, o bien pueden implementarse en forma de una combinación de software y una unidad procesadora para ejecutar el software. Mediante el uso de un medio 200a de registro no transitorio para almacenar el software se puede instalar, en el conmutador 200, un programa de software a implementar. La sección 205 de almacenamiento puede implementarse en forma de cualquier dispositivo de almacenamiento, por ejemplo una memoria de semiconductor.

El servidor 210 de control incluye una sección 211 de cálculo de rutas, una sección 212 de transferencia de rutas y una unidad 213 de almacenamiento de datos de control. Como en el caso del conmutador 200, la configuración del servidor 210 de control es la misma que la configuración del servidor 110 de control de la primera realización ilustrativa mostrada en la Figura 2, y se omite su descripción. La sección 211 de cálculo de rutas y la sección 212 de transferencia de rutas pueden implementarse en *hardware*, o bien pueden implementarse en forma de una combinación de *software* y una unidad procesadora para ejecutar el *software*. Mediante el uso de un medio 210a de registro no transitorio para almacenar el programa de *software* se puede instalar, en el servidor 210 de control, un programa de *software* a implementar. La unidad 213 de almacenamiento de datos de control puede implementarse en forma de cualquier dispositivo de almacenamiento, por ejemplo una HDD (unidad de disco duro) y una memoria de semiconductor.

A continuación se describirá la sección 209 de OAM. La sección 209 de OAM supervisa un estado de comunicación de una ruta que no está registrada en la sección 205 de almacenamiento. Para la sección 209 de OAM se puede emplear la "Ethernet Operations, Administration, Maintenance" ("Ether-OAM") descrita en la Bibliografía no de patente 2, por ejemplo. Tal como se describe en la Bibliografía no de patente 2, la Ether-OAM está propuesta como recomendación de la UIT-T (Unión Internacional de Telecomunicaciones - Sector de Normalización de las Telecomunicaciones). Más concretamente, se puede conseguir la funcionalidad de la sección 209 de OAM mediante el uso de una función denominada CC ("Verificación de continuidad", por sus siglas en inglés) de la Ether-OAM.

Se describirá a continuación la operación de la sección 209 de OAM cuando se utiliza la Ether-OAM. La sección 209 de OAM transmite y recibe periódicamente una trama de CC desde y hacia una ruta especificada por el servidor 210 de control a través de la sección 206 de interfaz de servidor de control. En la sección 204 de tratamiento se encapsula la trama de CC, y se transmite a través de la sección 207 de interfaz de red OpenFlow. Al gestionar a la trama de CC de la misma manera que a la trama de PBB en la primera realización ilustrativa, la sección 209 de OAM mide un estado de comunicación de una ruta que no está registrada en la sección 205 de almacenamiento.

45 La operación después de recibir la trama de CC es casi la misma que la operación mostrada en las Figuras 10 a 13 de la primera realización ilustrativa, y se omite su descripción.

(Efecto)

Según la presente realización ilustrativa como se ha descrito en lo que antecede, la sección 209 de OAM mide el estado de comunicación de la ruta que no está registrada en el conmutador 200.

Mediante esta operación el servidor de control puede conocer el estado de comunicación de la ruta que no está registrada en cada conmutador de la red. En consecuencia, el servidor de control puede supervisar el estado de comunicación de la ruta opcional de la red, y se puede realizar un cambio de ruta a gran velocidad en respuesta a la aparición de un fallo o el descenso de la calidad de línea.

[Tercera realización ilustrativa]

55 (Configuración y operación)

A continuación se describirá con detalle una tercera realización ilustrativa de la presente invención, haciendo

referencia a la Figura 15.

10

15

30

35

40

La Figura 15 es un diagrama de bloques que muestra una configuración de sistema de la presente realización ilustrativa.

El sistema según la presente realización ilustrativa incluye una unidad 1000 de comunicación y una unidad 1100 de control. Aunque no se muestra en la Figura 15, la unidad 1000 de comunicación y la unidad 1100 de control pertenecen a una red.

La unidad 1000 de comunicación incluye una sección 1001 de medición, una sección 1002 de notificación de resultados de medición, una sección 1003 de adición y una sección 1004 de tratamiento.

La sección 1001 de medición mide un estado de comunicación basándose en datos de medición del estado de comunicación cuando la unidad 1000 de comunicación es un nodo de borde de entrada de la red.

La sección 1002 de notificación de resultados de medición notifica el resultado de medición del estado de comunicación a la unidad 1100 de control por medio de la sección 1001 de medición, cuando la unidad 1000 de comunicación es un nodo de borde de salida de la red

La sección 1003 de adición añade datos de medición del estado de comunicación a un trama que la unidad 1000 de comunicación recibe.

La sección 1004 de tratamiento consulta datos de identificación de la trama recibida y realiza el tratamiento de la trama recibida, conforme a una regla de tratamiento que relaciona los datos de identificación de una trama y el tratamiento para dicha trama.

Debe señalarse que, como en el caso de la primera realización ilustrativa, la sección 1001 de medición, la sección 1002 de notificación de resultados de medición, la sección 1003 de adición y la sección 1004 de tratamiento pueden implementarse en *hardware*, o bien pueden implementarse en forma de una combinación de un programa de *software* y una unidad procesadora para ejecutar el programa de *software*. El programa de *software* a implementar se puede instalar en la unidad 1000 de comunicación mediante el uso del medio 1000a de registro no transitorio para almacenar el programa de *software*.

La unidad 1100 de control incluye una sección 1101 de cálculo de rutas, una sección 1102 de transferencia de rutas y una sección 1103 de almacenamiento de rutas.

La sección 1101 de cálculo de rutas calcula una ruta a partir del resultado de medición recibido desde la unidad 1000 de comunicación.

La sección 1102 de transferencia de rutas establece una regla de tratamiento de una trama para transferir unidades de una ruta basándose en una ruta almacenada en la sección 1103 de almacenamiento de rutas.

La unidad 1103 de almacenamiento de rutas almacena una ruta calculada por la sección 1101 de cálculo de rutas.

La sección 1101 de cálculo de rutas y la sección 1102 de transferencia de rutas se pueden implementar en *hardware*, o bien se pueden implementar como una combinación de un programa de *software* y una unidad procesadora para ejecutar el programa de *software*. El programa de *software* a implementar se puede instalar en la unidad 1100 de control mediante el uso de un medio 1100a de registro no transitorio para almacenar el programa de *software*. La sección 1103 de almacenamiento de rutas se puede implementar en forma de un dispositivo de almacenamiento opcional tal como una HDD (unidad de disco duro) y una memoria de semiconductor.

(Efecto)

Según la presente realización ilustrativa como se ha descrito en lo que antecede, la sección 1003 de adición de la unidad 1000 de comunicación añade datos de medición del estado de comunicación, y la sección 1001 de medición mide un estado de comunicación, para notificar a un resultado de medición a la unidad 1100 de control.

Mediante la operación precedente, es posible la conmutación de rutas a gran velocidad por el servidor de control, en función del estado de comunicación de la red.

REIVINDICACIONES

1. Una unidad de comunicación para uso en una red OpenFlow, que comprende:

una sección (103) de adición adaptada para añadir datos de medición del estado de comunicación a un trama de recepción;

5 una sección (101) de medición que mide un estado de comunicación basándose en los datos de medición del estado de comunicación;

una sección (102) de notificación de resultados de medición que notifica un resultado de medición del estado de comunicación a una unidad de control que controla dicha red; y

- una sección (104) de tratamiento que consulta datos de identificación de la trama de recepción para realizar el tratamiento de la trama de recepción definido en una regla de tratamiento que relaciona los datos de identificación de la trama de recepción y el tratamiento para la trama de recepción.
 - 2. La unidad de comunicación según la reivindicación 1, en donde el resultado de medición del estado de comunicación incluye al menos uno de una tasa de pérdida de tramas, un tiempo medio de retardo y una velocidad media de recepción sobre una ruta de comunicación de la trama de recepción.
- 3. La unidad de comunicación según la reivindicación 1 o 2, en donde los datos de medición del estado de comunicación incluyen un número de secuencia de la trama de recepción y los datos de tiempo de transmisión de la trama de recepción en una unidad de comunicación en calidad de origen de transmisión de la trama de recepción.
 - 4. La unidad de comunicación según cualquiera de las reivindicaciones 1 a 3, en donde dicha sección (103) de adición añade los datos de medición del estado de comunicación a la trama de recepción cuando se recibe la trama de recepción desde una red externa.
 - 5. La unidad de comunicación según cualquiera de las reivindicaciones 1 a 4, en donde dicha sección (103) de adición añade datos de identificador de un nodo de borde de entrada de dicha red a la que pertenece dicha unidad de comunicación, y datos de identificador de un nodo de borde de salida de dicha red, a la trama de la recepción.
- 6. La unidad de comunicación según la reivindicación 5, en donde dicha sección (101) de medición mide el estado de comunicación cuando los datos de identificador de dicho nodo de borde de salida de dicha red a la que pertenece dicha unidad de comunicación, son los datos de identificador de dicha unidad de comunicación en la está presente que dicha medición sección.
 - 7. La unidad de comunicación según cualquiera de las reivindicaciones 1 a 6, en donde dicha sección (103) de adición encapsula los datos de medición del estado de comunicación como una cabecera de PBB (Puente de red principal de proveedor).
 - 8. Un sistema de comunicación que comprende:

20

30

la unidad de comunicación según cualquiera de las reivindicaciones 1 a 7;

una unidad (110) de control que comprende:

una sección (111) de cálculo de rutas que calcula una ruta de la trama de recepción a partir del resultado de medición recibido desde dicha unidad de comunicación;

una sección (113) de almacenamiento de rutas que almacena la ruta calculada; y

una sección (112) de transferencia de rutas que estipula la regla de tratamiento de la trama a dicha unidad de comunicación de la ruta calculada, basándose en la ruta calculada que está almacenada en dicha sección de almacenamiento de rutas.

40 9. Un método de comunicación para uso en una red OpenFlow, que comprende:

adición de datos para medir un estado de comunicación a una trama de recepción, por un nodo de borde de entrada de dicha red:

medición del estado de comunicación basándose en los datos de medición del estado de comunicación, por un nodo de borde de salida de dicha red;

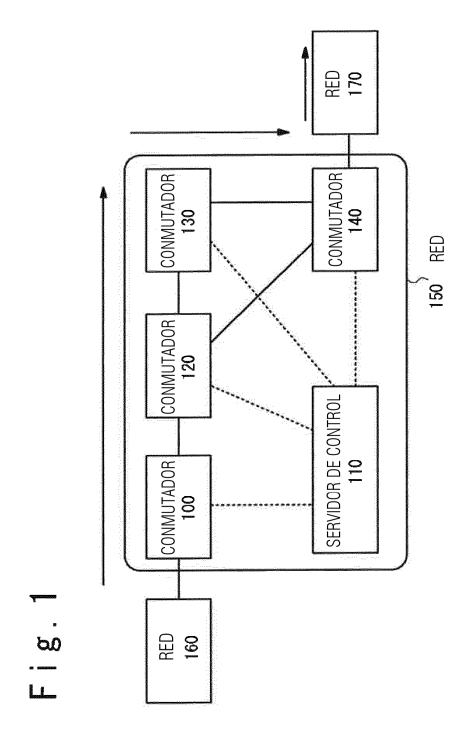
45 notificación de un resultado de medición del estado de comunicación a una unidad de control que controla dicha red;
v

consulta de datos de identificación de la trama de recepción para realizar el tratamiento de la trama de recepción definido en una regla de tratamiento que relaciona los datos de identificación de la trama de recepción y el

ES 2 564 667 T3

tratamiento para la trama de recepción.

- 10. Un medio de registro no transitorio en el cual se almacena un programa que hace que una unidad de comunicación para uso en una red OpenFlow, ejecute un proceso de comunicación que comprende:
- adición de datos para medir un estado de comunicación a una trama de recepción;
- 5 medición del estado de comunicación basada en los datos de medición del estado de comunicación;
 - notificación de un resultado de medición del estado de comunicación a una unidad de control que controla dicha red; v
- consulta de datos de identificación de la trama de recepción para realizar el tratamiento de la trama de recepción definido en una regla de tratamiento que relaciona los datos de identificación de la trama de recepción con el tratamiento para la trama de recepción.



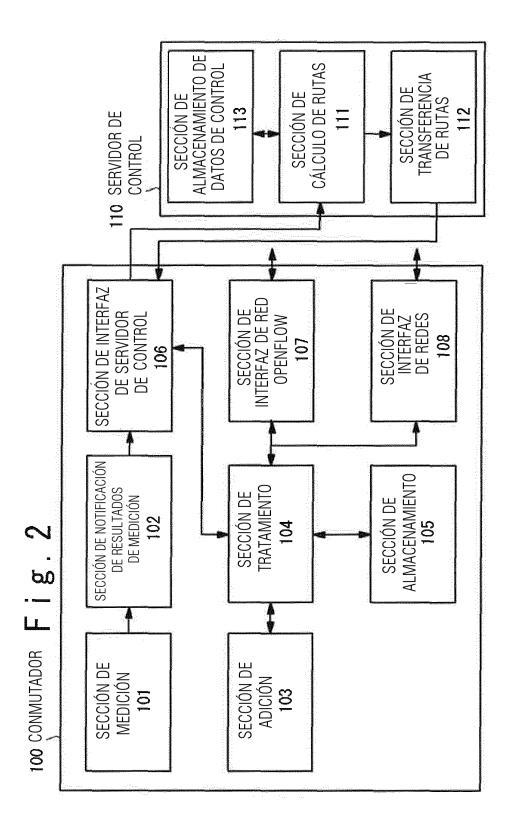


Fig. 3

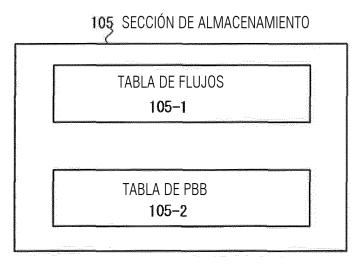
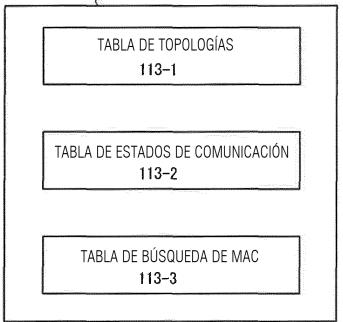
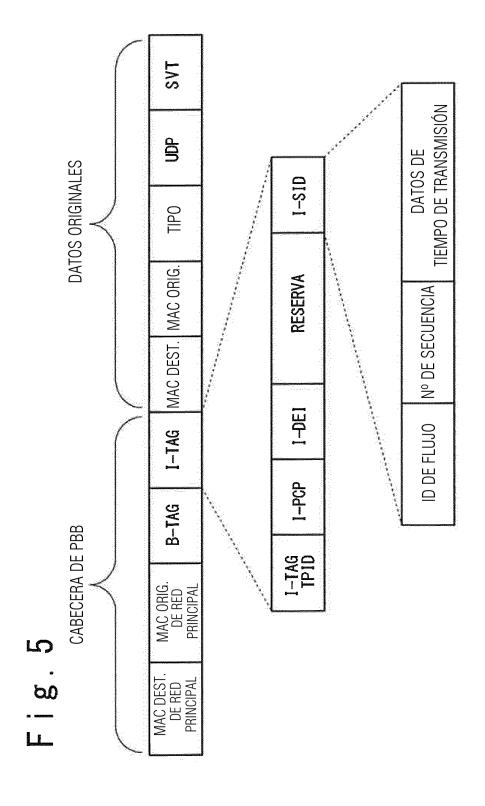


Fig. 4

113 SECCIÓN DE ALMACENAMIENTO DE DATOS DE CONTROL

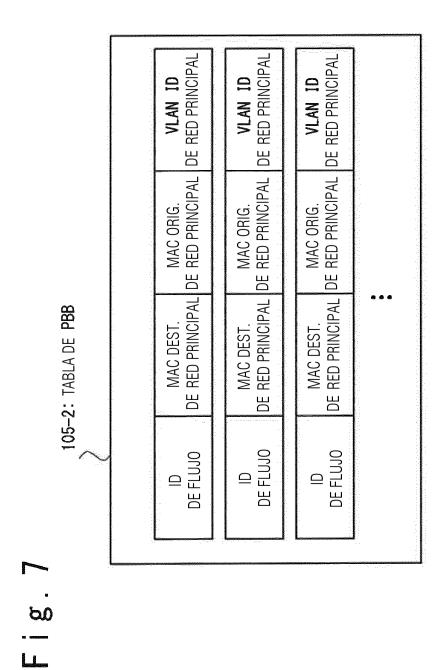




ACCIÓN ACCIÓN ACCIÓN PUERTO DEST. PCT/PDU PUERTO DEST. PCT/PDU PUERTO DEST. PCT/PDU PUERTO ORIG. PCT/PDU PUERTO ORIG. PCT/PDU PUERTO ORIG. PCT/PDU IP TOSBITS IP TOSBITS IP TOSBITS IP PROT. IP PROT. IP PROT. IP Dest. IP Dest. **IP** DEST. **IP** ORIG. **FP** ORIG. **IP** ORIG. PRIORIDAD VLAN PRIORIDAD VLAN PRIORIDAD VLAN **3** e MLAN ID 田田口 MAC ORIG. MAC ORIG. MAC ORIG. MAC DEST. MAC DEST. MAC DEST. PUERTO DE ENTRADA PUERTO DE ENTRADA PUERTO DE ENTRADA ID DE I ID DE FLUJO ID DE FLUJO

Q

105-1 TABLA DE FLUJOS



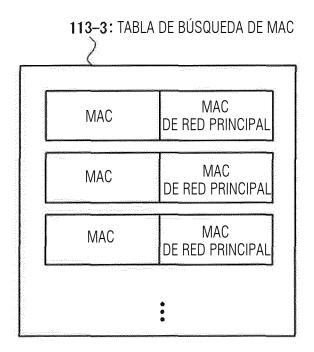
 ∞ <u>_</u>

113-2: TABLA DE ESTADOS DE COMUNICACIÓN

TIEMPO MEDIO DE VELOCIDAD MEDIA RETARDO (μs) (tps) TIEMPO MEDIO DE VELOCIDAD MEDIA RETARDO (μ s) (tps) TIEMPO MEDIO DE VELOCIDAD MEDIA RETARDO (μ s) (tps) PÉRDIDA DE TRAMAS (%) PÉRDIDA DE TRAMAS (%) PÉRDIDA DE TRAMAS (%) DE RED PRINCIPAL DE RED PRINCIPAL DE RED PRINCIPAL VLAN ID VLAN 1D MLAN ID DE RED PRINCIPAL DE RED PRINCIPAL MAC ORIG. DE RED PRINCIPAL MAC ORIG. MAC ORIG. MAC DEST. DE RED PRINCIPAL DE RED PRINCIPAL DE RED PRINCIPAL MAC DEST. MAC DEST.

(tps)

Fig. 9



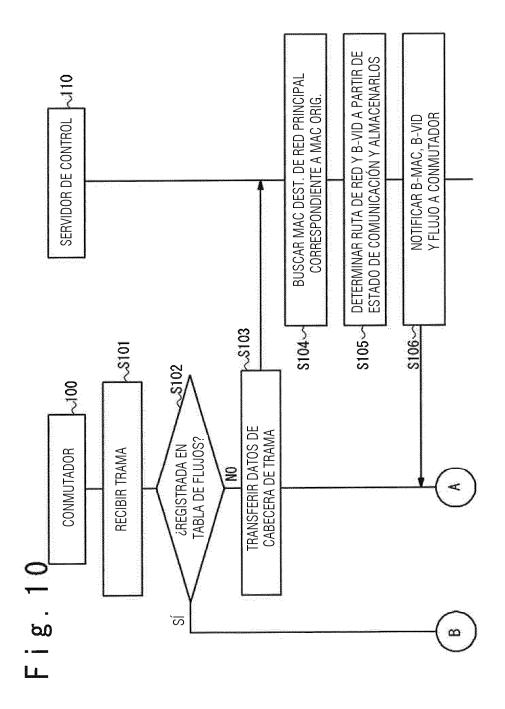


Fig. 11

