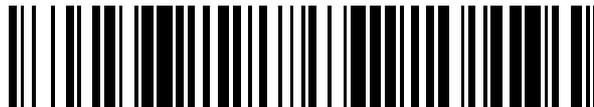


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 711 368**

51 Int. Cl.:

H04L 29/08 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **02.05.2013 PCT/IB2013/053454**

87 Fecha y número de publicación internacional: **07.11.2013 WO13164777**

96 Fecha de presentación y número de la solicitud europea: **02.05.2013 E 13730325 (1)**

97 Fecha y número de publicación de la concesión europea: **10.10.2018 EP 2845372**

54 Título: **Distribución de paquetes de dos niveles con distribución de paquete de primer nivel sin estado a un grupo de servidores y distribución de paquetes de segundo nivel con estado a un servidor dentro del grupo**

30 Prioridad:
04.05.2012 US 201213464608

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
03.05.2019

73 Titular/es:
**TELEFONAKTIEBOLAGET LM ERICSSON (PUBL)
(100.0%)
164 83 Stockholm, SE**

72 Inventor/es:
**ARISOYLU, MUSTAFA;
ARORA, ABHISHEK y
ANAND, PRASHANT**

74 Agente/Representante:
LINAGE GONZÁLEZ, Rafael

ES 2 711 368 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

Distribución de paquetes de dos niveles con distribución de paquete de primer nivel sin estado a un grupo de servidores y distribución de paquetes de segundo nivel con estado a un servidor dentro del grupo

5

Antecedentes

Campo

10 Las realizaciones de la invención se refieren al campo de las redes; y más específicamente, a los paquetes de distribución entre servidores en redes.

Información de antecedentes

15 Existen numerosos ejemplos en redes en los que los clientes o solicitantes de servicios se comunican con servidores o proveedores de servicios o recursos a través de una o más redes. Los clientes solicitan típicamente un recurso o que se realice un servicio y los servidores realizan el servicio o proporcionan el recurso. A menudo, se proporciona un conjunto de servidores para proporcionar una mayor capacidad de servicio, resistencia o alta disponibilidad (por ejemplo, en el caso de que un servidor falle o se desconecte), etc. Como ejemplo, se puede usar una granja de
20 servidores para proporcionar servicios asociados con sitios web, redes de charla interactiva en Internet, sitios de protocolo de transferencia de archivos de gran ancho de banda, servidores de protocolo de transferencia de noticias de red (NNTP), servidores de sistema de nombres de dominio (DNS), por nombrar solo algunos ejemplos.

25 El equilibrio de carga del servidor se usa comúnmente para distribuir las solicitudes de servicio de los clientes entre un conjunto de servidores para equilibrar o distribuir de manera más equitativa las cargas (por ejemplo, cargas de procesamiento o cargas de trabajo) asociadas con el manejo de las solicitudes de servicio entre el conjunto de servidores. El equilibrio de carga del servidor, dentro de ciertas restricciones, generalmente ayuda a evitar que algunos de los servidores se sobrecarguen mientras que otros servidores están siendo infrautilizados. Este equilibrio de carga del servidor generalmente ofrece ventajas tales como un mayor uso de recursos, mayor rendimiento, mayor
30 productividad, menor tiempo de respuesta, menor sobrecarga del servidor, etc. Además, el equilibrio de carga del servidor también se puede usar junto con proporcionar resistencia o alta disponibilidad al redistribuir las cargas cuando uno de los servidores falla, se desconecta, etc. En general, el equilibrio de carga del servidor es transparente para los clientes que solicitan los servicios.

35 Los algoritmos de equilibrio de carga del servidor se usan comúnmente para determinar a qué servidores distribuir las solicitudes de servicio o las cargas de trabajo. Los algoritmos de equilibrio de carga del servidor comúnmente incluyen algoritmos de planificación así como otros algoritmos. Los algoritmos usados para implementar el equilibrio de carga del servidor varían desde algoritmos relativamente simples hasta algoritmos relativamente sofisticados. Los ejemplos de algoritmos relativamente simples incluyen algoritmos de elección aleatoria y algoritmos de turnos rotativos. Los ejemplos de algoritmos más sofisticados incluyen aquellos que toman en cuenta los factores pertinentes para el equilibrio de carga, como la carga actual del servidor, el estatus arriba/abajo de un servidor, etc. A modo de ejemplo, se puede usar un algoritmo de turnos rotativos ponderado y las ponderaciones para los
40 servidores pueden cambiarse dinámicamente basándose en las cargas de trabajo actuales de los servidores.

45 Un tipo común de equilibrio de carga del servidor se conoce como equilibrio de carga del servidor con estado. En un enfoque conocido para el equilibrio de carga del servidor con estado, se detecta un primer paquete de un flujo en un equilibrador de carga del servidor. Por ejemplo, un paquete SYN de una conexión TCP puede detectarse al examinar un encabezado del paquete SYN. El equilibrador de carga del servidor usa un algoritmo de equilibrio de carga del servidor (por ejemplo, un algoritmo de elección aleatoria o un algoritmo de turnos rotativos) para seleccionar o asignar un servidor para manejar la conexión TCP. Una entrada que incluye un identificador para la conexión TCP y un identificador para el servidor asignado para manejar la conexión TCP se agrega como estado a una tabla de estado mantenida por el equilibrador de carga del servidor. Posteriormente, cuando se reciben paquetes adicionales de la conexión TCP, el equilibrador de carga del servidor realiza una búsqueda de tabla en la
50 tabla de estado. La búsqueda en la tabla localiza la entrada que mapea el identificador de conexión TCP al identificador del servidor. El equilibrador de carga del servidor reenvía los paquetes adicionales al servidor identificado. En tal enfoque de estado, el estado que se basa en una decisión anterior de equilibrio de carga (por ejemplo, basado en un cálculo previo con un algoritmo de equilibrio de carga) se almacena o conserva y se usa para tomar las posteriores decisiones de equilibrio de carga.

60 Tales enfoques de equilibrio de carga del servidor con estado tienden a tener ciertos inconvenientes, especialmente cuando el número de conexiones es alto. Como ejemplo, cuando el número de conexiones es alto, el tamaño de la tabla de estado tiende a ser correspondientemente grande. Como resultado, la cantidad de memoria para almacenar la tabla generalmente aumenta. Además, el tiempo y los recursos de procesamiento necesarios para realizar búsquedas de tablas generalmente también aumentan. Tales búsquedas de tabla se pueden realizar para cada
65 paquete recibido para cada conexión establecida. Además, el algoritmo de equilibrio de carga se puede realizar para cada nueva conexión. En consecuencia, especialmente cuando el número de conexiones es alto, tales enfoques

pueden causar que el equilibrador de carga severa sufra el agotamiento de los recursos de la memoria y/o los recursos de procesamiento. Alternativamente, en un esfuerzo por evitar dicho agotamiento de recursos, los equilibradores de carga pueden ser altamente aprovisionados con hardware, lo que generalmente tiende a hacerlos más costosos. Otro inconveniente potencial, si se proporciona un equilibrador de carga del servidor redundante o de copia de seguridad (por ejemplo, para resistencia o alta disponibilidad), es que se puede encontrar una sobrecarga significativa de comunicación y procesamiento para sincronizar el estado (por ejemplo, la tabla de estado) del equilibrador de carga del servidor redundante o de copia de seguridad.

El documento US 7321926 divulga un sistema de asignar un recurso a una solicitud de servicio. Los servidores se asignan virtualmente mediante una jerarquía que incluye grupos de servidores y supergrupos de los grupos de servidores.

El documento US 2006/0233106 divulga una arquitectura que facilita el equilibrio de carga entre una pluralidad de anfitriones y preserva la afinidad de sesión a un anfitrión dado.

Sumario

En un aspecto, un método, que se realiza en uno o más elementos de red que están en comunicación entre clientes que transmiten paquetes y servidores, para distribuir los paquetes entre los servidores que deben procesar los paquetes. Se proporciona la adherencia de los flujos de los paquetes a los servidores asignados para procesar los paquetes de los flujos. El método incluye un paso para recibir un paquete de un flujo en un módulo de distribución de paquetes de primer nivel estático. El método también incluye un paso para seleccionar estáticamente un grupo de servidores, entre una pluralidad de grupos de servidores, para el paquete del flujo con el módulo de distribución de paquetes de primer nivel estático. La selección estática del grupo de servidores no implica el uso de un estado que asigna el paquete del flujo al grupo seleccionado de servidores. El método incluye además un paso para distribuir el paquete del flujo a un sistema de distribución de paquetes de segundo nivel con estado distribuido. También se incluye un paso para seleccionar con estado un servidor del grupo seleccionado de servidores que debe procesar el paquete del flujo con el sistema de distribución de paquetes de segundo nivel con estado distribuido al acceder al estado que asigna el procesamiento de paquetes del flujo al servidor seleccionado. El método incluye un paso para distribuir el paquete del flujo al servidor seleccionado. Las ventajas que pueden lograrse dependiendo de la implementación incluyen la capacidad de manejar eficientemente grandes cantidades de flujos y/o la capacidad de manejar una cantidad de flujos eficientemente sin un suministro excesivo de hardware, la capacidad de evitar el uso del estado en el primer nivel de distribución de paquetes, y la capacidad de dinamización de los servidores sin interrumpir la adherencia de los flujos existentes.

En otro aspecto, un sistema de distribución de paquetes multinivel configurado para ser implementado en uno o más elementos de red en una red entre clientes que transmiten paquetes y servidores. El sistema de distribución de paquetes multinivel es operable para distribuir los paquetes entre los servidores, al tiempo que proporciona la adherencia de los flujos de los paquetes a los servidores. El sistema de distribución de paquetes multinivel incluye un módulo de distribución de paquetes de primer nivel estático configurado para recibir un paquete de un flujo. El módulo de distribución estática de paquetes de primer nivel es operable para seleccionar estáticamente un grupo de servidores, entre una pluralidad de grupos de servidores, para el paquete del flujo. La selección estática del grupo de servidores no implica el uso de un estado que asigna el paquete del flujo al grupo seleccionado de servidores. El sistema de distribución de paquetes multinivel también incluye un sistema de distribución de paquetes de segundo nivel con estado distribuido en comunicación con el módulo de distribución de paquetes de primer nivel estático. El sistema de distribución de paquetes de segundo nivel con estado distribuido incluye una pluralidad de módulos de distribución de paquetes de segundo nivel con estado distribuidos. El sistema de distribución de paquetes de segundo nivel con estado distribuido es operable para seleccionar sin estado un servidor del grupo seleccionado de servidores que debe procesar el paquete del flujo. La selección sin estado del servidor implica el uso de un estado que asigna el flujo al servidor seleccionado. El sistema de distribución de paquetes de segundo nivel con estado distribuido es operable para distribuir el paquete al servidor seleccionado. Las ventajas que pueden lograrse dependiendo de la implementación incluyen la capacidad de manejar eficientemente grandes cantidades de flujos y/o la capacidad de manejar una cantidad de flujos eficientemente sin un suministro excesivo de hardware, la capacidad de evitar el uso del estado en el primer nivel de distribución de paquetes, y la capacidad de dinamización de los servidores sin interrumpir la adherencia de los flujos existentes.

Breve descripción de los dibujos

La invención se puede entender mejor haciendo referencia a la siguiente descripción y los dibujos adjuntos que se usan para ilustrar las realizaciones de la invención. En los dibujos:

La figura 1 es un diagrama de bloques de una red que tiene una realización de un sistema de distribución de paquetes multinivel que es operable para distribuir flujos de paquetes entre servidores que están asignados para procesar los paquetes mientras proporciona la adherencia de los flujos a los servidores.

La figura 2 es un diagrama de bloques de una realización más detallada de un sistema de distribución de paquetes de dos niveles que es operable para distribuir paquetes entre servidores y proporcionar adherencia de flujos de los paquetes a los servidores.

5 La figura 3 es un diagrama de bloques de un sistema de distribución de paquetes de dos niveles que incluye un módulo de distribución de paquetes de primer nivel estático/sin estado que se implementa en hardware que es distinto del hardware usado para implementar un sistema de distribución de paquetes de segundo nivel con estado distribuido.

10 La figura 4 es un diagrama de bloques de un sistema de distribución de paquetes de dos niveles implementado en un único elemento de red.

La figura 5 es un diagrama de flujo de bloques de una realización de un método, que se puede realizar en uno o más elementos de red que están en comunicación entre clientes que transmiten paquetes y servidores, para distribuir los paquetes entre los servidores que deben procesar los paquetes, a la vez que proporciona adherencia de los flujos de los paquetes a los servidores asignados para procesar los paquetes de los flujos.

15 La figura 6 es un diagrama de flujo de bloques de una realización de un método que se puede realizar por cada módulo de distribución de paquetes de segundo nivel con estado distribuido de un grupo seleccionado para flujos de tipo 1 para seleccionar un servidor para procesar un paquete.

La figura 7 es un diagrama de flujo de bloques de una realización de un método que se puede realizar por cada módulo de distribución de paquetes de segundo nivel con estado distribuido de un grupo seleccionado para flujos de tipo 1 o tipo 2 para seleccionar un servidor para procesar un paquete.

25 La figura 8 es un diagrama de bloques de una realización de un conjunto de módulos de distribución de paquetes de segundo nivel con estado distribuidos de un grupo que tiene un conjunto de ejemplos de tablas de flujo correspondiente.

30 La figura 9 es un diagrama de bloques de una realización de un conjunto de módulos de distribución de paquetes de segundo nivel con estado distribuidos de un grupo que tiene un conjunto de ejemplo correspondiente de tablas de flujo y un conjunto de ejemplos correspondiente de tablas de flujo existente en otros.

35 La figura 10 es un diagrama de bloques de una realización de distribución de flujos de tipo 1 y tipo 2 entre grupos de módulos de distribución de paquetes de segundo nivel con estado distribuidos.

La figura 11 es un diagrama de bloques de una realización de un grupo de módulos de distribución de paquetes de segundo nivel con estado distribuidos que son operables para enviar mensajes entre sí en los que los mensajes indican los estados y/o las cargas de los servidores correspondientes.

40 La figura 12 es un diagrama de bloques de una realización de módulos de distribución de paquetes de segundo nivel con estado distribuidos correspondientes a diferentes grupos de servidores que intercambian mensajes entre sí, en los cuales los mensajes indican estados y/o cargas asociados con el grupo de servidores correspondiente.

45 La figura 13 es un diagrama de bloques de una realización de un sistema de distribución de paquetes de dos niveles que es operable para reenviar de manera eficiente paquetes de flujos de mayor volumen (por ejemplo, flujos de elefante) y flujos de menor volumen (por ejemplo, flujos de ratón).

Descripción de las realizaciones

50 En la siguiente descripción, se exponen numerosos detalles específicos, tales como configuraciones específicas de red, algoritmos, secuencias de operaciones, estructuras de datos, opciones de partición/integración y similares. Sin embargo, se entiende que las realizaciones de la invención pueden ponerse en práctica sin estos detalles específicos. En otros casos, los circuitos, estructuras y técnicas bien conocidos no se han mostrado en detalle para no entorpecer la comprensión de esta descripción.

60 Como se usa en el presente documento, un elemento de red (por ejemplo, un enrutador, un conmutador, un puente) es una pieza de equipo de red, que incluye hardware y software, que interconecta comunicativamente a otros equipos en la red (por ejemplo, otros elementos de red, estaciones finales, etc.). Algunos elementos de la red son "elementos de la red de servicios múltiples" que brindan soporte para múltiples funciones de red (por ejemplo, enrutamiento, conexión en puente, conmutación, agregación de capa 2, control de límite de sesión, calidad de servicio y/o gestión de abonados) y/o proporcionan soporte para múltiples servicios de aplicaciones (por ejemplo, datos, voz y video). Las estaciones finales de abonado (por ejemplo, clientes, estaciones de trabajo, portátiles, ultraportátil, miniordenador portátil, teléfonos móviles, teléfonos inteligentes, teléfonos multimedia, teléfonos con protocolo de voz a través de Internet (VOIP), equipos de usuario, terminales, reproductores de medios portátiles, unidades de GPS, sistemas de juegos, decodificadores) acceden a los contenidos/servicios proporcionados a través

de Internet y/o contenidos/servicios proporcionados en redes privadas virtuales (las VPN) superpuestas (por ejemplo, a través de un túnel) en Internet. El contenido y/o los servicios típicamente son proporcionados por una o más estaciones finales (por ejemplo, estaciones finales de servidor) que pertenecen a un servicio o proveedor de contenido o estaciones finales que participan en un servicio de igual a igual, y pueden incluir, por ejemplo, páginas web públicas (por ejemplo, contenido gratuito, escaparates, servicios de búsqueda), páginas web privadas (por ejemplo, páginas web a las que se accede mediante nombre de usuario/contraseña que proporcionan servicios de correo electrónico) y/o redes corporativas a través de las VPN. Típicamente, las estaciones finales de los abonados se acoplan (por ejemplo, a través de equipos de las instalaciones del cliente acoplados a una red de acceso (cableada o inalámbrica)) a elementos de la red de borde, que se acoplan (por ejemplo, a través de uno o más elementos de la red central) a otros elementos de la red de borde, que se acoplan a otras estaciones finales (por ejemplo, estaciones finales de servidor).

La figura 1 es un diagrama de bloques de una red que tiene clientes 101 que son operables para transmitir flujos de paquetes, servidores 112 que son operables para procesar los paquetes, y una realización de un sistema 100 de distribución de paquetes multinivel que es operable para distribuir los paquetes entre los servidores y proporcionan adherencia de los flujos a los servidores asignados para procesar los paquetes de los flujos. En algunas realizaciones, el sistema de distribución de paquetes multinivel puede representar un sistema de equilibrio de carga del servidor multinivel y puede usarse para distribuir o equilibrar de manera aproximadamente uniforme las cargas de trabajo asociadas con el procesamiento de paquetes entre los servidores. Otras realizaciones no están limitadas a dicho equilibrio de carga del servidor.

Los clientes 101 están acoplados, conectados o de otro modo en comunicación a través de una o más redes 102 al sistema 100 de distribución de paquetes multinivel. En la ilustración, se muestra el cliente 1 101-1 a través del cliente P 101-P, donde P puede ser cualquier número entero. Cada cliente puede representar una combinación de hardware y software (por ejemplo, un sistema operativo y una aplicación de cliente) que es capaz de acceder a servicios y/o recursos a través de una o más conexiones de red. Los ejemplos de clientes adecuados incluyen, entre otros, dispositivos de cliente inalámbricos (por ejemplo, teléfonos móviles, portátiles, ordenadores móviles, dispositivos de mano, etc.) y dispositivos de cliente con cable (por ejemplo, ordenadores de escritorio, decodificadores, etc.). Los clientes son operables para transmitir o enviar paquetes de conexiones u otros flujos 105 a través de una o más redes al sistema de distribución de paquetes multinivel.

El sistema 100 de distribución de paquetes multinivel está acoplado, conectado o, de lo contrario, en comunicación con una o más redes 102 para recibir los paquetes de los flujos 105 de los clientes 101. El sistema 100 de distribución de paquetes multinivel también está acoplado, conectado o, de lo contrario, en comunicación con un conjunto de servidores 112. Cada servidor puede representar cualquier combinación de hardware y software (por ejemplo, un sistema operativo y una aplicación de servidor) que sea capaz de proporcionar servicios y/o contenido a los clientes a través de una o más interfaces de red y/o que de otra manera funcione como un servidor. Los ejemplos de servidores adecuados incluyen, entre otros, software ejecutable en hojas (por ejemplo, servidores de hoja), tarjetas de control, tarjetas de servicio, procesadores, núcleos de procesador, etc. Como se muestra en la ilustración, en algunas realizaciones, los servidores se asocian o se agrupan lógicamente en múltiples grupos 13. En la ilustración, los servidores se agrupan lógicamente en el grupo 1 113-1 a través del grupo N 113-N, donde N puede ser cualquier número entero. Al menos uno de los grupos tiene una pluralidad de servidores y los otros grupos pueden tener cada uno uno o más servidores. En algunos casos, cada uno de los grupos tiene una pluralidad de servidores. En algunas realizaciones, un servidor puede ser compartido por más de un grupo (es decir, los grupos pueden estar superpuestos).

El sistema de distribución de paquetes multinivel es operable para distribuir los paquetes de los flujos 105 entre los servidores del conjunto de servidores 112. En algunas realizaciones, el sistema de distribución de paquetes multinivel puede distribuir paquetes de una manera que proporcione o mantenga la adherencia de los flujos de los paquetes a los servidores asignados para procesar los paquetes de los flujos. Tal adherencia de flujo también se conoce como persistencia de flujo. La distribución de paquetes con la adherencia del flujo o la persistencia del flujo puede implicar la distribución de todos los paquetes de un flujo al mismo servidor. El sistema de distribución de paquetes multinivel puede asignar inicialmente un flujo (por ejemplo, basado en un primer paquete del flujo) a uno de los servidores. Por ejemplo, el paquete inicial de un flujo entrante (por ejemplo, el paquete SYN de una conexión TCP, un paquete INIT de una conexión SCTP, etc.) se reenvía a un servidor. Posteriormente, todos los otros paquetes encontrados para la misma conexión o flujo pueden reenviarse al mismo servidor. Comúnmente, cada servidor mantiene el estado (por ejemplo, datos de sesión) que está asociado con los flujos que está asignado a procesar. Este estado a menudo no es mantenido por, o está disponible para, los otros servidores, pero a menudo es necesario para procesar correctamente los paquetes de los flujos. Por ejemplo, este es a menudo el caso de los protocolos de transporte orientados a la conexión como TCP, SCTP, etc.

Un flujo puede representar el tráfico de red que se identifica por un conjunto común de atributos que caracterizan el flujo. Como se usa en el presente documento, un flujo representa más que un único paquete (es decir, tiene al menos dos paquetes). Como un primer ejemplo, un flujo puede representar el tráfico de red identificado por cinco atributos de tupla (por ejemplo, dirección IP de origen/destino, protocolo, puertos TCP/UDP de origen/destino). Como segundo ejemplo, un flujo puede representar el tráfico de red identificado por las direcciones de MAC de

origen/destino. Como tercer ejemplo, un flujo puede representar el tráfico de red identificado por los ID de VLAN. Como cuarto ejemplo, un flujo puede representar el tráfico de red identificado por los nodos de origen y destino del tráfico. Como quinto ejemplo, a través de un determinado enlace del nodo a al nodo b, los paquetes que pasan por el enlace con una dirección IP de origen específica (por ejemplo, IP1) pueden representar un flujo identificado por la dirección IP de origen sobre ese enlace con los atributos (IP1, a, b). Como sexto ejemplo, en una red de acceso, el tráfico originado por un abonado puede representar un flujo identificado como el tráfico que pasa a través de un puerto específico de la pasarela residencial. Tales flujos de abonado en las redes de acceso y de borde también pueden identificarse por las direcciones IP del abonado. El flujo del abonado ascendente/descendente (el flujo del lado del abonado/red al lado de la red/abonado) puede tener la dirección IP del abonado como la dirección IP de origen/destino, respectivamente. Un flujo puede incluir un flujo único, subflujos múltiples o muchos subflujos. Un flujo de grano grueso más grande puede incluir subflujos de grano más fino más pequeños. Por ejemplo, un flujo agregado que pasa a través de un enlace (desde el nodo a al nodo b) con múltiples direcciones IP de destino tiene una granularidad más gruesa que un subflujo que pasa por el mismo enlace con una única dirección IP de destino. El flujo anterior se puede denominar flujo de enlace y el último flujo se puede denominar flujo de IP de destino de enlace.

El sistema 100 de distribución de paquetes multinivel, que en este caso es un sistema de distribución de paquetes de dos niveles, incluye un módulo 103 de distribución de paquetes de primer nivel estático y/o sin estado y un sistema 107 de distribución de paquetes de segundo nivel con estado distribuido. El módulo 103 de distribución de paquetes de primer nivel estático/sin estado está acoplado, conectado o, de otro modo, en comunicación con una o más redes 102 para recibir los paquetes de los flujos 105. El módulo de distribución de paquetes de primer nivel estático/sin estado incluye un módulo 104 de selección de grupo estático/sin estado. El módulo de selección de grupo estático/sin estado es operable para seleccionar estáticamente y/o sin estado un grupo de servidores 113, entre los múltiples grupos de servidores 113-1 a 113-N, para los paquetes de los flujos 105. Al seleccionar los grupos de estáticamente/sin estado, se entiende que el módulo no necesita usar el estado para seleccionar los grupos. No es necesario mantener el estado (por ejemplo, una tabla de estado) que asigna o mapea los paquetes de los flujos a grupos de servidores. Especialmente cuando el número de conexiones a manejar es alto, esto puede ayudar a reducir el agotamiento de recursos y/o el aprovisionamiento de hardware. No se requiere que el módulo de distribución de paquetes de primer nivel estático/sin estado mantenga ningún estado, ya que de hecho se puede mantener cierto estado (por ejemplo, otro estado que no sea una tabla de flujo y/o para otros fines además de seleccionar un grupo para un paquete).

Más bien, el módulo puede usar un enfoque estático, como, por ejemplo, un algoritmo hash u otro algoritmo estático (por ejemplo, que determina estáticamente el grupo basándose en el contenido del paquete). Por ejemplo, una función hash puede mapear atributos del paquete a un identificador de grupo del grupo seleccionado. Se puede generar un identificador de flujo a partir del paquete (por ejemplo, basado en ciertos campos del encabezado de paquetes). El ID de flujo puede representar un ID o etiqueta (por ejemplo, un número) que se usa para identificar un flujo. Como ejemplo particular, hay que considerar los servidores N (por ejemplo, N = 10). Una función hash simple puede tomar el último byte de una dirección IP de origen (por ejemplo, 192.168.1.122) del paquete como un entero (por ejemplo, 122) y tomar el módulo N (por ejemplo, N = 10). El número resultante (por ejemplo, 2) está entre 0 y N-1, lo que apunta a un grupo al que se ha de reenviar el paquete. Para mayor claridad, el término estático no implica que la selección deba ser invariable, ya que la selección estática puede cambiar, por ejemplo, si cambia el número de grupos (por ejemplo, se agrega un grupo de tal manera que la función hash se realiza con N = 11, o se elimina un grupo de tal manera que la función hash se realiza con N = 9). Más bien, como se mencionó, el término estático se refiere a la falta de uso del estado (por ejemplo, una tabla de estado) para mapear o asignar el paquete al grupo. El módulo de distribución de paquetes de primer nivel estático/sin estado está acoplado, conectado o, de otro modo, en comunicación para proporcionar una distribución de flujo a grupo de paquetes 106 al sistema 107 de distribución de paquetes de segundo nivel con estado distribuido. El módulo de distribución de paquetes de primer nivel estático/con estado proporciona adherencia de flujo a grupo de tal manera que todos los paquetes del flujo se distribuyen al mismo grupo. El módulo de distribución de paquetes de primer nivel estático/sin estado se puede implementar en hardware, software, firmware o varias combinaciones de los mismos. En algunas realizaciones, el módulo de distribución de paquetes de primer nivel estático/sin estado puede incluir al menos algún hardware, circuitería o lógica en particular (por ejemplo, circuitería de aplicación específica), aunque esto no es necesario.

El sistema 107 de distribución de paquetes de segundo nivel con estado distribuido incluye múltiples módulos 108 de distribución de paquetes de segundo nivel con estado distribuidos. En la ilustración, se muestran el módulo 1 108-1 a través del módulo M 108-M, donde M puede ser cualquier número entero. En algunas realizaciones, cada uno de los módulos 108 de distribución de paquetes de segundo nivel con estado distribuidos puede corresponder a uno o más de los grupos de servidores 113. Cada uno de los módulos 108 de distribución de paquetes de segundo nivel con estado distribuidos incluye un módulo 109 de selección de servidor de segundo nivel con estado distribuido correspondiente y un estado correspondiente 110. En particular, el módulo 1 de distribución de paquetes de segundo nivel incluye un módulo 109-1 de selección de servidor de segundo nivel con estado distribuido y un estado 110-1, y el módulo M de distribución de paquetes de segundo nivel M^{ésimo} incluye un módulo 109-M de selección de servidor de segundo nivel con estado distribuido y un estado 110-M. Cada uno de los módulos de selección de servidor de segundo nivel con estado distribuidos puede ser operable para seleccionar con estado un servidor individual particular, del grupo de servidores seleccionados por el módulo 103 de distribución de paquetes de primer nivel

estático/sin estado, que ha de procesar los paquetes de los flujos 105. En algunas realizaciones, el sistema de distribución de paquetes de segundo nivel con estado distribuido se puede usar para lograr un equilibrio de carga del servidor sustancialmente uniforme en todos los servidores dentro del grupo seleccionado por el primer nivel. Otras realizaciones no se limitan al equilibrio de carga del servidor.

5 Al seleccionar con estado el servidor, el módulo 109 de selección de servidor puede ser operable para acceder y usar el estado correspondiente 110. En algunas realizaciones, el estado puede mapear o asignar el flujo al servidor seleccionado. A modo de ejemplo, cada estado puede representar una tabla de estado que tiene entradas que mapean diferentes conexiones o flujos (por ejemplo, identificadores de flujo) a los servidores correspondientes (por ejemplo, identificadores de servidor) asignados para manejar el procesamiento de los paquetes de las conexiones o flujos. El procesamiento de la distribución de paquetes en el nivel dos tiende a ser más intensivo en computación que el nivel 1, pero se distribuye a través de los múltiples módulos de distribución de paquetes. En consecuencia, a través de dicha distribución, el estado dentro de cada uno de los módulos de distribución de paquetes múltiples generalmente es menor en cantidad que lo que se necesitaría sin dicha distribución. El almacenamiento de cada una de estas cantidades menores de estado generalmente consume menos memoria, y el uso del estado (por ejemplo, realizar búsquedas de tablas) generalmente lleva menos tiempo y consume menos recursos de procesamiento. Ventajosamente, la distribución del sistema de distribución de paquetes de segundo nivel con estado ayuda a hacer que el enfoque sea más ampliable a un gran número de conexiones u otros flujos. El módulo de distribución de paquetes de segundo nivel con estado distribuido está acoplado, conectado o, de otro modo, en comunicación para proporcionar la distribución de flujo a servidor o el reenvío de paquetes 111 a los servidores seleccionados que pueden procesar los paquetes. El módulo de distribución de paquetes de segundo nivel con estado distribuido puede proporcionar adherencia del flujo al servidor, de manera que todos los paquetes del flujo se distribuyan al mismo servidor.

10

15

20

25 Como se muestra, y sin limitación, en algunas realizaciones, el conjunto de servidores 112 puede estar en comunicación con otros servidores 115 a través de una o más redes 114. A modo de ejemplo, los servidores 115 pueden proporcionar servicios o contenido a los clientes. Por ejemplo, los clientes pueden iniciar conexiones con los servidores 115 a través de los servidores 112. Los servidores 112 pueden realizar el procesamiento asociado con las conexiones, como, por ejemplo, la gestión de abonados, la inspección profunda de paquetes o similares.

30 Alternativamente, en otras realizaciones, los servidores 112 pueden ser los puntos finales de conexión o flujo y los servidores 112 pueden proporcionar el conjunto completo de contenido y/o servicios para los clientes.

En algunas realizaciones, el sistema de distribución de paquetes multinivel puede implementarse en un solo elemento de red (por ejemplo, un único enrutador). En otras realizaciones, el sistema de distribución de paquetes multinivel puede implementarse en dos o más elementos de red conectados. En algunas realizaciones, el primer y segundo nivel del sistema de distribución de paquetes multinivel se implementan en hardware diferente.

35 El sistema 100 de distribución de paquetes multinivel puede proporcionar una serie de ventajas, especialmente cuando el número de conexiones y/o la cantidad de tráfico son altos. En los últimos años, ha habido un aumento significativo en la cantidad de dispositivos en red y en el tráfico de red en general que se encuentra en las redes, incluida Internet. Además, muchas de estas conexiones, como las que proporcionan video (por ejemplo, video a demanda, televisión de alta definición, etc.) tienen un ancho de banda relativamente alto. Como se mencionó anteriormente en la sección de antecedentes, algunos enfoques conocidos de equilibrio de carga con estado tienden a sufrir agotamiento de recursos, especialmente cuando la cantidad de conexiones y/o la cantidad de tráfico son altas. Sin embargo, el sistema de distribución de paquetes multinivel puede ayudar a evitar, o al menos reducir, dicho agotamiento de recursos. El módulo de distribución de paquetes de primer nivel implementa un enfoque estático/sin estado y no necesita retener el estado. Las tareas de distribución de paquetes y el mantenimiento del estado más intensivos en computación se realizan en el segundo nivel, que se distribuye en múltiples módulos de distribución de paquetes. Cada uno de los múltiples módulos de distribución de paquetes tiene un estado asociado, que es menor en cantidad que el que se necesitaría sin dicha distribución. El almacenamiento de cada una de estas cantidades menores de estado generalmente consume menos memoria, y el uso del estado (por ejemplo, realizar búsquedas de tablas) generalmente lleva menos tiempo y consume menos recursos de procesamiento.

40 La figura 2 es un diagrama de bloques de una realización más detallada de un sistema 200 de distribución de paquetes de dos niveles. En algunas realizaciones, el sistema de distribución de paquetes de dos niveles se puede usar para el equilibrio de carga del servidor y/o para distribuir sustancialmente las cargas asociadas con el procesamiento de paquetes entre los servidores. Otras realizaciones no están limitadas al equilibrio de carga.

45 El sistema 200 de distribución de paquetes de dos niveles incluye un módulo 203 de distribución de paquetes de primer nivel estático/sin estado y un sistema 207 de distribución de paquetes de segundo nivel con estado distribuido. El módulo de distribución de paquetes de primer nivel estático/sin estado incluye un módulo 204 de selección de grupo estático/sin estado. En algunas realizaciones, el módulo de selección de grupo estático/sin estado puede incluir un módulo hash 216 para seleccionar un grupo para un flujo realizando un hash (por ejemplo, basado en el contenido de un paquete). Como se muestra, en algunas realizaciones, el módulo de selección de grupo estático/sin estado puede omitir el estado 217 que asigna el paquete del flujo al grupo seleccionado de

servidores. El módulo de distribución de paquetes de primer nivel estático/sin estado puede proporcionar una distribución 206 de paquetes de flujo a grupo.

El sistema 207 de distribución de paquetes de segundo nivel con estado distribuido está acoplado, conectado o, de lo contrario, en comunicación con el módulo 203 de distribución de paquetes de primer nivel estático/sin estado. En la realización ilustrada, el sistema de distribución de paquetes de segundo nivel incluye seis módulos 208-1 a 208-6 de distribución de paquetes (PD) de segundo nivel con estado distribuidos (DSSL). Cada uno de los módulos 208-1 a 208-6 de PD de DSSL corresponde a uno respectivo de seis servidores 212-1 a 212-6. Por ejemplo, el módulo 3 208-3 de PD de DSSL corresponde al servidor 3 212-3. Si bien se muestran seis servidores y seis módulos de PD de DSSL en la realización ilustrada, en otras realizaciones puede haber menos o más servidores y/o módulos de PD de DSSL. Los servidores se muestran en líneas discontinuas para indicar que no forman parte del sistema 200 de distribución de paquetes de dos niveles. Cada uno de los seis módulos de PD de DSSL tiene un módulo 209-1 a 209-6 de selección de servidor correspondiente y un estado correspondiente 210-1 a 210-6. Por ejemplo, el módulo 3 208-3 de PD de DSSL tiene un módulo 209-3 de selección de servidor y un estado 210-3 (por ejemplo, una tabla de estado). Los módulos de selección del servidor y el estado pueden ser como se describió anteriormente. Cada uno de los módulos de PD de DSSL realiza cálculos de distribución de paquetes (por ejemplo, cálculos de selección de servidor) y toma decisiones de distribución de paquetes para su servidor correspondiente. El sistema de distribución de paquetes de segundo nivel con estado distribuido y/o los módulos de PD de DSSL proporcionan una distribución 211 de paquetes de flujo a servidor.

En la realización ilustrada, para fines de ilustración, los servidores y sus correspondientes módulos de PD de DSSL se dividen lógicamente en tres grupos. En particular, los servidores 1-3 y los módulos 1-3 de PD de DSSL están en el grupo 1, los servidores 4-5 y los módulos 4-5 de PD de DSSL están en el grupo 2, y los servidores 5-6 y los módulos 5-6 de PD de DSSL están en el grupo 3. El módulo de distribución de paquetes de primer nivel estático/sin estado puede seleccionar uno de estos grupos 1-3 y la distribución 206 de paquetes de flujo a grupo puede distribuir el paquete a uno de estos grupos 1-3. Los módulos de PD de DSSL dentro del grupo seleccionado pueden seleccionar uno de los servidores dentro del grupo seleccionado y la distribución 211 de paquetes de flujo a servidor puede reenviar el paquete al servidor seleccionado.

Para ilustrar mejor una realización del funcionamiento del sistema de distribución de paquetes de dos niveles, hay que considerar la recepción de un primer paquete [F1] de un primer flujo, un segundo paquete [F2] de un segundo flujo y un tercer paquete [F3] de un tercer flujo, en el módulo 203 de distribución de paquetes de primer nivel estático/sin estado. El módulo 203 de distribución de paquetes de primer nivel estático/sin estado, en este ejemplo, puede seleccionar de estáticamente/sin estado el grupo 1 para el primer paquete [F1], el grupo 2 para el segundo paquete [F2] y el grupo 3 para el tercer paquete [F3]. En algunas realizaciones, los paquetes pueden proporcionarse o distribuirse a cada uno de los módulos de PD de DSSL del grupo seleccionado. Como se muestra, el primer paquete [F1] puede distribuirse a cada uno de los módulos 1-3 de PD de DSSL, el segundo paquete [F2] puede distribuirse a cada uno de los módulos 4-5 de PD de DSSL, y el tercer paquete [F3] puede distribuirse a cada uno de los módulos 5-6 de PD de DSSL. En algunas realizaciones, los módulos de PD de DSSL dentro de cada grupo pueden compartir un dominio de multidifusión y/o difusión y los paquetes pueden ser multidifundidos y/o difundidos a los módulos de PD de DSSL dentro de cada grupo. En algunas realizaciones, los módulos de PD de DSSL dentro de cada grupo pueden compartir un identificador de red de área local virtual (VLAN). Por ejemplo, los módulos de PD de DSSL dentro del grupo 1 pueden tener ID de VLAN n.º 1, los módulos de PD de DSSL dentro del grupo 2 pueden tener ID de VLAN n.º 2 y los módulos de PD de DSSL dentro del grupo 3 pueden tener ID de VLAN n.º 3.

En algunas realizaciones, los módulos de distribución de paquetes de segundo nivel dentro de cada grupo pueden realizar colectivamente la distribución de paquetes de manera distribuida para asignar el paquete del flujo a un servidor apropiado dentro de cada grupo. En algunas realizaciones, como consecuencia de la adherencia del flujo, cada paquete para un flujo dado solo puede ser aceptado para ser procesado por un servidor y los otros servidores dentro del grupo pueden descartar el paquete. En la realización ilustrada, por ejemplo, el módulo 2 de PD de DSSL puede seleccionar el servidor 2 y hacer que su copia del paquete [F1] sea aceptada para su procesamiento por el servidor 2, mientras que los módulos 1 y 3 de PD de DSSL no pueden seleccionar su servidor correspondiente y pueden hacer que sus copias correspondientes del paquete [F1] sean descartadas. De manera similar, el módulo 4 de PD de DSSL puede seleccionar el servidor 4 y hacer que su copia del paquete [F2] sea aceptada para su procesamiento por el servidor 4, mientras que el módulo 5 de PD de DSSL no puede seleccionar el servidor 5 y puede hacer que sus copias correspondientes de los paquetes [F2] y [F3] sean descartadas. Continuando, el módulo 6 de PD de DSSL puede seleccionar el servidor 6 y hacer que su copia del paquete [F3] sea aceptada para ser procesada por el servidor 6. Este es solo un ejemplo ilustrativo y el alcance de la invención no se limita a este ejemplo.

La figura 3 es un diagrama de bloques de un sistema de distribución de paquetes de dos niveles que incluye un módulo 304 de distribución de paquetes de primer nivel estático/sin estado que se implementa en el hardware 322 que es distinto del hardware 324 usado para implementar un sistema de distribución de paquetes de segundo nivel con estado distribuido. A modo de ejemplo, el hardware diferente puede ser hardware en diferentes elementos de red, diferentes porciones de hardware de un elemento de red, diferentes placas, diferentes tarjetas, diferentes procesadores, diferentes núcleos de procesador, etc. En algunas realizaciones, además de que un hardware es

distinto, se puede compartir un poco de hardware. En algunas realizaciones, el hardware 324 usado para implementar el sistema 307 de distribución de paquetes de segundo nivel con estado distribuido puede ser el mismo que, o puede incluir, hardware usado para implementar un conjunto de servidores 312 para los cuales el sistema de distribución de paquetes de dos niveles realiza el paquete distribución. A modo de ejemplo, un módulo de distribución de paquetes de segundo nivel con estado distribuido del sistema 307 y uno de los servidores 312 pueden compartir el hardware de una porción de un elemento de red, un hoja, una placa, una tarjeta, un procesador o un núcleo de procesador. Ventajosamente, la mayoría de las tareas de uso intensivo de memoria y computación del sistema general de distribución de paquetes se pueden descargar a la memoria existente y a los recursos de procesamiento de los servidores. Compartir dicho hardware puede ayudar a que la distribución distribuida de paquetes de segundo nivel con estado sea ampliable a un gran número de conexiones u otros flujos sin que sea necesario un suministro excesivo de hardware adicional debido a la capacidad de hacer uso del hardware existente. En algunas realizaciones, el hardware 322 y el hardware 324 pueden estar dentro de un único elemento de red. En otras realizaciones, el hardware 322 y el hardware 324 pueden estar dentro de dos o más elementos de red conectados.

La figura 4 es un diagrama de bloques de un sistema de distribución de paquetes de dos niveles implementado en un único elemento de red. Los elementos de red se suelen separar en un plano de control y un plano de datos (a veces denominado plano de reenvío o plano de medios). En el caso de que el elemento de red sea un enrutador (o esté implementando la funcionalidad de enrutamiento), el plano de control generalmente determina cómo se han de enrutar los datos (por ejemplo, los paquetes) (por ejemplo, el próximo salto para los datos y el puerto de salida para esos datos), y el plano de datos se encarga de reenviar esos datos. Por ejemplo, el plano de control típicamente incluye uno o más protocolos de enrutamiento que se comunican con otros elementos de la red para intercambiar rutas y seleccionar esas rutas basándose en una o más métricas de enrutamiento. Las rutas y adyacencias se almacenan en una o más estructuras de ruta en el plano de control. El plano de control programa el plano de datos con información basada en la estructura o estructuras de enrutamiento. El plano de datos usa estas estructuras de reenvío y adyacencia cuando reenvía tráfico. Para el reenvío de capa 2, el elemento de red puede almacenar una o más tablas de puente que se usan para reenviar datos basándose en la información de capa 2.

En algunas realizaciones, un elemento de red incluye un conjunto de una o más tarjetas de línea, un conjunto de una o más tarjetas de control y, opcionalmente, un conjunto de una o más tarjetas de servicio (a veces denominadas tarjetas de recursos). En otras realizaciones, un elemento de red incluye un conjunto de una o más tarjetas de línea y un conjunto de una o más tarjetas de servicio. Estas tarjetas se acoplan juntas mediante uno o más mecanismos (por ejemplo, una primera malla completa que acopla las tarjetas de línea y una segunda malla completa que acopla todas las tarjetas, una estructura de conmutador, un plano posterior, etc.). El conjunto de tarjetas de línea constituye el plano de datos, mientras que el conjunto de tarjetas de control y/o tarjetas de servicio proporcionan el plano de control e intercambian paquetes con elementos de red externos a través de las tarjetas de línea. El conjunto de tarjetas de servicio puede proporcionar procesamiento especializado (por ejemplo, servicios de capa 4 a capa 7 (por ejemplo, firewall, IPsec, IDS, P2P), controlador de límite de sesión VoIP, pasarelas inalámbricas móviles (GGSN, pasarelas de sistema de paquetes evolucionados (EPS)). A modo de ejemplo, se puede usar una tarjeta de servicio para terminar los túneles IPsec y ejecutar los algoritmos de cifrado y autenticación del asistente.

Con referencia nuevamente a la figura 4, el elemento de red incluye una o más tarjetas 422 de línea, una estructura 420 de conmutación y un conjunto de tarjetas 424 de control y/o servicio. Las tarjetas pueden estar completamente interconectadas entre sí de manera convencional. En la realización ilustrada, se muestran N tarjetas de control y/o servicio, donde N puede ser cualquier número entero. En la realización ilustrada, el sistema 404 de distribución de paquetes de primer nivel estático/sin estado se implementa en las tarjetas 422 de línea. Un sistema 408 de distribución de paquetes de segundo nivel con estado distribuido se implementa en el conjunto de tarjetas 424 de control y/o servicio. Las tarjetas de control y/o servicio también se usan para implementar servidores 412. En particular, el módulo 1 408-1 de PD de DSSL se implementa en la tarjeta 1 424-1 de control/servicio que también implementa el servidor 1 412-1, el módulo 2 408-2 de PD de DSSL se implementa en la tarjeta 2 424-2 de control/servicio que también implementa el servidor 2 412-2, y el módulo N 408-N de PD de DSSL se implementa en la tarjeta N 424-N de control/servicio que también implementa el servidor N 412-N. El módulo de distribución de paquetes de primer nivel puede incluir información sobre un grupo seleccionado (por ejemplo, un ID de grupo) de las tarjetas de control/servicio en un paquete y reenviar el paquete a la estructura de conmutador. La estructura de conmutador, que en una realización incluye un conjunto de tarjetas de conmutación, puede usar la información sobre el grupo seleccionado (por ejemplo, el ID de grupo) para realizar multidifusión y/o difusión a todas las tarjetas de control/servicio del grupo seleccionado.

La figura 5 es un diagrama de flujo de bloques de una realización de un método 530, que se puede realizar en uno o más elementos de red que están en comunicación entre clientes que transmiten paquetes y servidores, para distribuir los paquetes entre los servidores que han de procesar los paquetes, a la vez que proporciona adherencia de los flujos de los paquetes a los servidores asignados para procesar los paquetes de los flujos. En algunas realizaciones, el método 530 puede realizarse mediante el sistema 100 de distribución de paquetes multinivel de la figura 1 o el sistema 200 de distribución de paquetes de dos niveles de la figura 2. En otras realizaciones, el método puede ser realizado por otros sistemas de distribución de paquetes multinivel o de dos niveles diferentes. Además, el

sistema 100 de distribución de paquetes multinivel y el sistema 200 de distribución de paquetes de dos niveles pueden realizar métodos iguales o diferentes al método 530.

El método incluye recibir un paquete de un flujo en un módulo de distribución de paquetes de primer nivel estático, en el bloque 531. El paquete puede ser de cualquier tipo y protocolo conocido. Una ventaja de los enfoques divulgados en el presente documento es que no están limitados a ningún tipo de paquete o protocolo de paquete en particular (por ejemplo, son indiferentes al protocolo). La adherencia del flujo se puede lograr incluso en protocolos que no permiten y/o sin necesidad de incluir información sobre un servidor asignado a un flujo dentro de un paquete o encabezado de paquete.

Un grupo de servidores se selecciona estáticamente para el paquete del flujo, entre una pluralidad de grupos de servidores, con el módulo de distribución de paquetes de primer nivel estático, en el bloque 532. Al seleccionar el grupo de servidores, no hay acceso ni uso del estado (por ejemplo, una tabla de flujo) que asigna el paquete del flujo al grupo seleccionado de servidores. En algunas realizaciones, el grupo de servidores se selecciona realizando un hash o similar con contenido del paquete.

El paquete del flujo se distribuye a un sistema de distribución de paquetes de segundo nivel con estado distribuido, en el bloque 533. En algunas realizaciones, el paquete del flujo es multidifundido, difundido o, de lo contrario, se distribuye a cada uno de un conjunto de módulos de distribución de paquetes de segundo nivel con estado distribuidos. En algunas realizaciones, los módulos de segundo nivel comparten un ID de VLAN. En algunas realizaciones, cada uno de los conjuntos de módulos distribuidos corresponde a un servidor diferente del grupo seleccionado. En algunas realizaciones, el paquete del flujo se distribuye desde el hardware usado para implementar el módulo de distribución de paquetes de primer nivel hasta el hardware separado o distinto que se usa para implementar el sistema de distribución de paquetes de segundo nivel con estado distribuido, que en algunos casos también se usa para implementar los servidores.

Un servidor del grupo seleccionado de servidores que ha de procesar el paquete del flujo se selecciona con estado con el sistema de distribución de paquetes de segundo nivel con estado distribuido, en el bloque 534. Seleccionar con estado el servidor puede incluir el acceso y el uso del estado que asigna el procesamiento de paquetes del flujo al servidor seleccionado. En algunas realizaciones, cada uno de un conjunto de módulos de distribución de paquetes de segundo nivel con estado distribuidos, que corresponden a un servidor diferente del grupo seleccionado, determina si hacer que el paquete correspondiente sea aceptado para su procesamiento por el servidor correspondiente o descartado. En un aspecto, solo uno de los módulos para el grupo seleccionado determina que el paquete ha de aceptarse, mientras que todos los demás módulos para el grupo seleccionado determinan que el paquete ha de descartarse.

El paquete del flujo se distribuye al servidor seleccionado, en el bloque 535. El servidor seleccionado puede procesar el paquete del flujo. En algunas realizaciones, posteriormente, otros paquetes del flujo se dirigen al mismo servidor seleccionando primero el mismo grupo y luego seleccionando el mismo servidor de manera que se proporcione la adherencia del flujo.

Para ciertos tipos de flujos, es posible detectar el inicio del flujo o el primer paquete del flujo basándose solo en los bits o el contenido del primer paquete del flujo sin necesidad de consultar otra información. Por ejemplo, un paquete inicial de una conexión o flujo puede identificarse como tal en el contenido del paquete inicial. Es decir, el paquete inicial de un flujo puede tener bits o contenido que indica que el paquete es el primer paquete del flujo. Los ejemplos de tales tipos de paquetes incluyen, entre otros un paquete TCP que tiene un indicador SYN para indicar que es el primer paquete en un flujo, un paquete SCTP que tiene un indicador para indicar que es el primer paquete en un flujo. De manera similar, otros protocolos orientados a la conexión comúnmente tienen una forma de informar a otros elementos de la red a través del contenido de un paquete inicial de que el paquete es el primer paquete de una conexión o flujo. Como un ejemplo adicional, un flujo de nivel de abonado puede identificarse por la dirección IP del abonado (por ejemplo, la dirección IP de origen/destino del tráfico ascendente/descendente). En tal caso, una solicitud de inicio RADIUS o una solicitud DHCP puede indicar el inicio del flujo de nivel de abonado. Como el ID de flujo se basa en la dirección IP de origen, el nuevo flujo para un abonado se detectará percibiendo el paquete RADIUS o el paquete DHCP que se genera para establecer la sesión del abonado. Los tipos de conexiones o flujos para los cuales es posible detectar o identificar el inicio del flujo o el primer paquete del flujo basándose solo en los bits o el contenido del primer paquete del flujo sin necesidad de consultar otra información se conocen como conexiones o flujos de tipo 1.

La figura 6 es un diagrama de flujo de bloques de una realización de un método 638 que puede ser realizado por cada módulo de distribución de paquetes de segundo nivel con estado distribuido de un grupo seleccionado para flujos de tipo 1 para seleccionar un servidor para procesar un paquete. Se recibe un paquete en un módulo de distribución de paquetes de segundo nivel con estado distribuido, en el bloque 639. Se realiza una determinación, en el bloque 640, si el paquete es para un nuevo flujo al examinar el paquete (por ejemplo, un encabezado del paquete). Por ejemplo, el paquete o encabezado de paquete puede incluir un marcador o indicador que indica que es un paquete inicial de un nuevo flujo. El módulo de distribución de paquetes de segundo nivel puede tener un módulo de examen de paquetes para hacer esta determinación. Si el paquete es para un nuevo flujo (es decir, "sí"

es la determinación en el bloque 640), el método avanza al bloque 641. En el bloque 641, se selecciona un servidor para procesar el paquete con un nuevo algoritmo de distribución de paquetes de flujo que es consistente entre todos los módulos del grupo seleccionado. Por el nuevo algoritmo de distribución de paquetes de flujo que es consistente entre todos los módulos del grupo seleccionado, se entiende que las instancias de los algoritmos en cada uno de los módulos del grupo seleccionado son operables para seleccionar el mismo servidor idéntico para procesar el paquete (es decir, todos seleccionan consistentemente el mismo servidor). Por ejemplo, todos los módulos del grupo seleccionado seleccionarán el servidor 2 en lugar de cualquiera de los módulos que seleccionan un servidor diferente. En algunas realizaciones, los algoritmos son todos idénticos. En otras realizaciones, los algoritmos pueden ser diferentes pero, sin embargo, producen selecciones consistentes. Los tipos de algoritmos descritos anteriormente son adecuados, como, por ejemplo, de turnos rotativos, de selección aleatoria, de turnos rotativos ponderados, selección aleatoria ponderada, funciones hash, etc., así como otros tipos conocidos de algoritmos más sofisticados. Como ejemplo específico, cada módulo puede usar una función hash coherente (por ejemplo, una misma función hash) que mapea un ID de flujo extraído del paquete a los ID del servidor, de modo que se seleccionará el mismo servidor y solo ese servidor aceptará el paquete para su procesamiento. En algunas realizaciones, este algoritmo puede intentar distribuir o equilibrar aproximadamente de manera igual las cargas asociadas con el procesamiento de paquetes entre los servidores (por ejemplo, usarse para equilibrar la carga del servidor) por ejemplo a través de ponderaciones. Otras realizaciones no se limitan al equilibrio de carga del servidor.

Se realiza una determinación, en el bloque 642, si el servidor seleccionado, que es seleccionado por el módulo que realiza el método (es decir, el servidor seleccionado en el bloque 641), corresponde al módulo que realiza el método. Por ejemplo, la determinación sería "sí" si un módulo 2 correspondiente a un servidor 2 seleccionó el servidor 2, o la determinación sería "no" si el módulo 2 correspondiente al servidor 2 seleccionara cualquier otro servidor además del servidor 2. El segundo nivel puede tener un módulo de determinación para hacer esta determinación. Si el servidor seleccionado no corresponde al módulo (es decir, "no" es la determinación en el bloque 642), el método avanza al bloque 643. En el bloque 643, el módulo que realiza el método hace que el paquete sea descartado (por ejemplo, el módulo descarta el paquete, el módulo no reenvía el paquete al servidor o el módulo hace que el servidor descarte el paquete, etc.). Alternativamente, si el servidor seleccionado corresponde al módulo (es decir, "sí" es la determinación en el bloque 642), el método avanza al bloque 644. En el bloque 644, se registra un nuevo identificador de flujo (ID de flujo) en una tabla de flujo del módulo que realiza el método. En algunas realizaciones, la tabla de flujo puede incluir un conjunto completo de ID de flujo para todos los flujos asignados al servidor correspondiente al módulo que realiza el método. En un aspecto, la tabla de flujo puede incluir pares de ID de flujo mapeados al ID de servidor del servidor correspondiente. Alternativamente, en otro aspecto, la tabla de flujo solo puede incluir explícitamente los ID de flujo y el ID del servidor puede ser implícito o entendido. El módulo hace que el paquete sea aceptado para su procesamiento por el servidor correspondiente, en el bloque 645 (por ejemplo, el módulo acepta el paquete, el módulo reenvía el paquete al servidor o el módulo hace que el servidor acepte el paquete, etc.).

Con referencia nuevamente al bloque 640, si el paquete no es para un nuevo flujo (es decir, la determinación es "no"), el método avanza al bloque 646. En el bloque 646, se determina si un identificador de flujo (ID de flujo) para el paquete está en una tabla de flujo del módulo que realiza el método. El segundo nivel puede tener un módulo de verificador de tabla de flujo para hacer esta determinación. Si el ID de flujo no está en la tabla de flujo, el módulo hace que el paquete sea descartado, en el bloque 647. Normalmente, dado que el paquete no es para un nuevo flujo, y el ID de flujo no está en la tabla de flujo del módulo, se asumirá que el ID de flujo está en una tabla de flujo de uno de los otros módulos del grupo seleccionado. Alternativamente, si el ID de flujo está en la tabla de flujo, el módulo hace que el paquete sea aceptado para su procesamiento por el servidor correspondiente, en el bloque 648. Dado que el paquete es para un flujo existente listado en la tabla de flujo del módulo que realiza el método, para mantener la persistencia del flujo o la adherencia del flujo, el paquete debería ser aceptado para su procesamiento. El segundo nivel puede tener un módulo de determinación para determinar si aceptar o descartar. Por lo general, es posible que todos los demás módulos que realizan el método no encuentren el ID de flujo para el paquete en sus tablas de flujo (por ejemplo, solo una tabla de flujo incluye el ID de flujo) y puede hacer que el paquete sea descartado para que no sea procesado por sus servidores correspondientes.

Este es solo un ejemplo de un método adecuado. Si bien el método se ha mostrado en una forma relativamente básica, las operaciones se pueden agregar y/o eliminar opcionalmente del método. Además, aunque se ha mostrado y descrito un orden particular de operaciones, debe entenderse que ese orden particular es de ejemplo. Realizaciones alternativas pueden realizar opcionalmente las operaciones en orden diferente, combinar ciertas operaciones, superponer ciertas operaciones, etc. Por ejemplo, otra realización puede hacer la determinación en el bloque 646 antes de la determinación en el bloque 640.

Para otros tipos de conexiones o flujos, no es posible detectar el inicio del flujo, o el primer paquete del flujo, basándose solo en los bits o el contenido del primer paquete del flujo sin necesidad de consultar otra información. Tales tipos de conexiones o flujos se denominan en el presente documento conexiones o flujos de tipo 2.

La figura 7 es un diagrama de flujo de bloques de una realización de un método 750 que puede ser realizado por cada módulo de distribución de paquetes de segundo nivel con estado distribuido de un grupo seleccionado para flujos de tipo 1 o tipo 2 para seleccionar un servidor para procesar un paquete. El método 750 es un enfoque

genérico que no se limita a ningún protocolo en particular o la necesidad de detectar un nuevo flujo solamente desde el contenido del paquete. Este enfoque puede usarse para flujos de tipo 1, flujos de tipo 2 o una combinación de flujos de tipo 1 y tipo 2. Se recibe un paquete en un módulo de distribución de paquetes de segundo nivel con estado distribuido, en el bloque 751. Se realiza una determinación, en el bloque 752, si un ID de flujo para el paquete está en una tabla de flujo del módulo que realiza el método. Si el ID de flujo está en la tabla de flujo (es decir, "sí" es la determinación en el bloque 752), el método avanza al bloque 753. En el bloque 753, el módulo hace que el paquete sea aceptado para su procesamiento por el servidor correspondiente. Alternativamente, si el ID de flujo no está en la tabla de flujo (es decir, "no" es la determinación en el bloque 752), el método avanza al bloque 754. En el bloque 754, se determina si el ID de flujo para el paquete se encuentra en una tabla de flujo existente en otros del módulo que realiza el método. En algunas realizaciones, la tabla de flujo existente en otros puede enumerar o incluir los ID de flujo asignados a los servidores del grupo seleccionado que no sea el servidor correspondiente al módulo que realiza el método. La combinación de los ID de flujo incluidos en la tabla de flujo del módulo y los diferentes ID de flujo incluidos en la tabla de flujo existente en otros del módulo representa la identificación de todos los flujos existentes a los que da servicio el grupo de servidores. Los ID de flujo en la tabla de flujo y la tabla de flujo existente en otros de un módulo dado se excluyen mutuamente. Si el ID de flujo se encuentra en la tabla de flujo existente en otros (es decir, "sí" es la determinación en el bloque 754), el módulo hace que el paquete sea descartado, en el bloque 755.

Alternativamente, si el ID de flujo no se encuentra en la tabla de flujo existente en otros (es decir, "no" es la determinación en el bloque 754), el método avanza al bloque 756. El hecho de que el ID de flujo del paquete no exista ni en la tabla de flujo del módulo ni en la tabla de flujo existente en otros del módulo indica de manera inferencial que el paquete es para un nuevo flujo. Como tal, se puede tomar una nueva decisión de distribución de paquetes para el paquete del nuevo flujo. En algunos casos, se puede tomar esta decisión para lograr el equilibrio de carga del servidor, aunque esto no es necesario. Se selecciona un servidor para procesar el paquete con un nuevo algoritmo de distribución de paquetes de flujo que es consistente entre todos los módulos del grupo seleccionado, en el bloque 756. Como se mencionó anteriormente, esto se refiere a las instancias de los algoritmos en cada uno de los módulos del grupo seleccionado que selecciona el mismo servidor para procesar el paquete. Se determina si el servidor seleccionado corresponde al módulo que realiza el método, en el bloque 757. Si el servidor seleccionado corresponde al módulo que realiza el método (es decir, "sí" es la determinación en el bloque 757), se registra o almacena un nuevo ID de flujo para el paquete en una tabla de flujo del módulo que realiza el método, en el bloque 758. El módulo hace que el paquete sea aceptado para su procesamiento por el servidor correspondiente, en el bloque 759. Con referencia nuevamente al bloque 757. Si el servidor seleccionado no se corresponde con el módulo que realiza el método (es decir, "no" es la determinación), se registra o almacena un nuevo ID de flujo para el paquete en la tabla de flujo existente en otros del módulo que realiza el método, en el bloque 760. El módulo hace que el paquete sea descartado, en el bloque 761.

De forma similar, todos los demás módulos del grupo pueden mantener una tabla de flujo existente en otros. Cada una de las tablas de flujo existente en otros permite al módulo correspondiente saber cuáles y cuántos flujos están manejando los otros servidores del grupo. En algunas realizaciones, esta información en el conjunto de tablas de flujo existente en otros se puede usar para informar o modificar el nuevo algoritmo de planificación de paquetes (por ejemplo, para realizar el equilibrio de carga del servidor). Por ejemplo, el nuevo algoritmo de distribución de paquetes de flujo puede distribuir paquetes para nuevos flujos a un servidor con el menor número de flujos. En algunas realizaciones, los módulos del grupo pueden intercambiar mensajes para sincronizar la información sobre sus tablas de flujo y las tablas de flujo existente en otros para mantener una vista coherente de los flujos dentro del grupo.

Dado que el método 750 de la figura 7 usa las tablas de flujo existente en otros, que deben almacenarse, tiende a consumir más memoria que el método 638 de la figura 6. Además, mantener y usar las tablas de flujo existente en otros tiende a requerir recursos de procesamiento adicionales. Por ejemplo, es posible que deba realizarse una búsqueda en la tabla de flujo existente en otros al llegar cada paquete cuyo ID de flujo no esté representado en la tabla de flujo del módulo. Por consiguiente, la realización del método 638 puede ofrecer ciertas ventajas sobre la ejecución del método 750 para flujos de tipo 1. Afortunadamente, muchos de los tipos más comunes de conexiones o flujos son conexiones o flujos de tipo 1 (por ejemplo, TCP, SCTP, GTP, otros protocolos orientados a la conexión, etc.). En algunas realizaciones, el método 638 se puede usar para tales tipos de flujos sin necesidad de realizar el método 750. En otras realizaciones, el método 750 se puede realizar para todos los flujos independientemente de si son de tipo 1 o de tipo 2. En aún otras realizaciones, el método 638 puede realizarse para los flujos que son de tipo 1 y el método 750 puede realizarse para los flujos que son de tipo 2 y/o para flujos que son de ambos tipos.

Este es solo un ejemplo de un método adecuado. Si bien el método se ha mostrado en una forma relativamente básica, las operaciones se pueden agregar y/o eliminar opcionalmente del método. Además, aunque se ha mostrado y descrito un orden particular de operaciones, ha de entenderse que ese orden particular es de ejemplo. Realizaciones alternativas pueden realizar opcionalmente las operaciones en orden diferente, combinar ciertas operaciones, superponer ciertas operaciones, etc. Por ejemplo, otra realización puede hacer la determinación en el bloque 754 antes de la determinación en el bloque 752. Como otro ejemplo, los dos métodos se pueden combinar en un solo método híbrido.

La figura 8 es un diagrama de bloques de una realización de un conjunto de módulos 808 de distribución de paquetes de segundo nivel con estado distribuido (PD de DSSL, por sus siglas en inglés), de un grupo que tiene un conjunto correspondiente de tablas 810 de flujo. En la realización ilustrada, para fines de ilustración, el grupo tiene tres módulos 808-1, 808-2, 808-3 de PD de DSSL que tienen tres tablas 810-1, 810-2, 810-3 de flujo correspondientes, aunque el alcance de la invención no es tan limitada. Además, en la realización ilustrada, para fines de ilustración, seis flujos que tienen el ID de flujo n.º 1-6 son servidos por el grupo y representados en las tablas de flujo, aunque el alcance de la invención no está tan limitado. Un primer servidor correspondiente al módulo 1 de PD de DSSL se asigna para dar servicio a un primer flujo con ID de flujo n.º 1, un tercer flujo con ID de flujo n.º 3 y un quinto flujo con ID de flujo n.º 5. En consecuencia, los ID de flujo n.º 1, n.º 3 y n.º 5 se incluyen en la tabla 810-1 de flujo del módulo 1 de PD de DSSL. Se asigna un segundo servidor correspondiente al módulo 2 de PD de DSSL para dar servicio a un segundo flujo con ID de flujo n.º 2. Por consiguiente, el ID de flujo n.º 2 se incluye en la tabla 810-2 de flujo del módulo 2 de PD de DSSL. Se asigna un tercer servidor correspondiente al módulo 3 de PD de DSSL para dar servicio a un cuarto flujo con ID de flujo n.º 4 y un sexto flujo con ID de flujo n.º 6. En consecuencia, los ID de flujo n.º 4 y n.º 6 se incluyen en la tabla 810-3 de flujo del módulo 3 de PD de DSSL.

La figura 9 es un diagrama de bloques de una realización de un conjunto de módulos 908 de PD de DSSL de un grupo que tiene un conjunto correspondiente de tablas 910 de flujo y un conjunto correspondiente de tablas 962 de flujo existente en otros. En la realización ilustrada, para fines de ilustración, el grupo tiene tres módulos 908-1, 908-2, 908-3 de PD de DSSL que tienen tres tablas 910-1, 910-2, 910-3 de flujo correspondientes, aunque el alcance de la invención no es tan limitada. Además, en la realización ilustrada, para fines de ilustración, seis flujos que tienen el ID de flujo n.º 1-6 son servidos por el grupo y representados en las tablas de flujo, aunque el alcance de la invención no está tan limitado. Un primer servidor correspondiente al módulo 1 de PD de DSSL se asigna para dar servicio a un primer flujo con ID de flujo n.º 1, un tercer flujo con ID de flujo n.º 3 y un quinto flujo con ID de flujo n.º 5. En consecuencia, los ID de flujo n.º 1, n.º 3 y n.º 5 se incluyen en la tabla 810-1 de flujo del módulo 1 de PD de DSSL. Se asigna un segundo servidor correspondiente al módulo 2 de PD de DSSL para dar servicio a un segundo flujo con ID de flujo n.º 2. Por consiguiente, el ID de flujo n.º 2 se incluye en la tabla 810-2 de flujo del módulo 2 de PD de DSSL. Se asigna un tercer servidor correspondiente al módulo 3 de PD de DSSL para dar servicio a un cuarto flujo con ID de flujo n.º 4 y un sexto flujo con ID de flujo n.º 6. En consecuencia, los ID de flujo n.º 4 y n.º 6 se incluyen en la tabla 810-3 de flujo del módulo 3 de PD de DSSL.

El ID de flujo n.º 2 correspondiente al servidor 2, el ID de flujo n.º 4 correspondiente al servidor 3 y el ID de flujo n.º 6 correspondientes al servidor 3 se incluyen en la tabla 962-1 de flujo existente en otros del módulo 1 de PD de DSSL. El ID de flujo n.º 1 correspondiente al servidor 1, el ID de flujo n.º 3 correspondiente al servidor 1, el ID de flujo n.º 5 correspondiente al servidor 1, el ID de flujo n.º 4 correspondiente al servidor 3 y el ID de flujo n.º 6 correspondiente al servidor 3 se incluyen en la tabla 962-2 de flujo existente en otros del módulo 2 de PD de DSSL. El ID de flujo n.º 1 correspondiente al servidor 1, el ID de flujo n.º 3 correspondiente al servidor 1, el ID de flujo n.º 5 correspondiente al servidor 1 y el ID de flujo n.º 2 correspondiente al servidor 2 se incluyen en la tabla 962-3 de flujo existente en otros del módulo 3 de PD de DSSL. Como se muestra en la realización ilustrada, las tablas de flujo existente en otros pueden, en algunas realizaciones, identificar los servidores que corresponden a (por ejemplo, el servicio) cada ID de flujo. Ventajosamente, la inclusión de los ID de servidor en la tabla de flujo existente en otros ayuda a proporcionar una imagen más completa del grupo y, en algunas realizaciones, se puede usar para fines de alta disponibilidad y/o para realizar el equilibrio de carga del servidor. Alternativamente, en otras realizaciones, las tablas de flujo existente en otros pueden simplemente enumerar los ID de flujo correspondientes a los otros servidores.

La figura 10 es un diagrama de bloques de una realización de la distribución de flujos de tipo 1 y tipo 2 entre los grupos 1013 de los módulos 1008 de PD de DSSL. El grupo 1 1013-1 está dedicado a flujos de solo tipo 1. Los ejemplos de tales flujos incluyen, entre otros flujos TCP y SCTP. Cada uno de un módulo 1 1008-1 de PD de DSSL, un módulo 2 1008-2 de PD de DSSL y un módulo 3 1008-1 de PD de DSSL tiene un módulo 1064 de selección de servidor que es operable para seleccionar servidores para flujos de tipo 1. A modo de ejemplo, cada uno de los módulos de selección de servidor en el grupo 1 puede ser operable para implementar el método de la figura 6.

El grupo 2 1013-2 se usa tanto para flujos de tipo 1 como de tipo 2. Por ejemplo, el grupo 2 se puede usar para la gestión de flujo de nivel de abonado. Un flujo de nivel de abonado puede incluir subflujos de tipo 1 y tipo 2. Por ejemplo, un abonado puede tener diferentes aplicaciones ejecutándose en una estación final, y algunas de las aplicaciones pueden usar flujos de tipo 1 mientras que otros usan flujos de tipo 2. Como otro ejemplo, el grupo 2 puede usarse para algunos flujos de tipo 1 y otros flujos de tipo 2 independientemente de si son flujos de abonados. Cada uno de un módulo 4 1008-4 de PD de DSSL y un módulo 5 1008-5 de PD de DSSL tiene un módulo 1064 de selección de servidor que es operable para seleccionar servidores para flujos de tipo 1, y también un módulo 1066 de selección de servidor que es operable para seleccionar servidores para flujos de tipo 2. A modo de ejemplo, cada uno de los módulos 1064 de selección de servidor de tipo 1 puede ser operable para implementar el método de la figura 6, y cada uno de los módulos 1066 de selección de servidor de tipo 2 puede ser operable para implementar el método de la figura 7. Como se muestra, cada uno de los módulos de PD de DSSL en el grupo 2 también puede incluir un conmutador lógico 1065 que es operable para examinar paquetes entrantes, determinar si los paquetes son para flujos de tipo 1 o tipo 2 y, en consecuencia, hacer que se use el módulo 1064 de selección de servidor tipo 1 o el módulo 1066 de selección de servidor de tipo 2.

El grupo 3 1013-3 está dedicado a flujos de tipo 2. Un módulo 6 1008-6 de PD de DSSL tiene un módulo 1066 de selección de servidor que es operable para seleccionar servidores para flujos de tipo 2.

5 La figura 11 es un diagrama de bloques de una realización de un grupo 1113 de módulos 1108 de PD de DSSL que son operables para enviar mensajes entre sí en los que los mensajes indican los estados y/o las cargas de los servidores correspondientes 1112. Un módulo 1 1108-1 de PD de DSSL corresponde a un servidor 1 1112-1, un módulo 2 1108-2 de PD de DSSL corresponde a un servidor 2 1112-2 y un módulo 3 1108-3 de PD de DSSL corresponde a un servidor 3 1112-3. Los servidores y los módulos de PD de DSSL pertenecen al mismo grupo 1113. El módulo 1 de PD de DSSL envía un primer mensaje (m1) que indica un primer estado (S1) y una primera carga (L1) del servidor 1 al módulo 2 de PD de DSSL y al módulo 3 de PD de DSSL. El módulo 2 de PD de DSSL envía un segundo mensaje (m2) que indica un segundo estado (S2) y una segunda carga (L2) del servidor 2 al módulo 1 de PD de DSSL y al módulo 3 de PD de DSSL. El módulo 3 de PD de DSSL envía un tercer mensaje (m3) que indica un tercer estado (S3) y una tercera carga (L3) del servidor 3 al módulo 1 de PD de DSSL y al módulo 2 de PD de DSSL. En un aspecto, cada módulo de PD de DSSL puede difundir o multidifundir un paquete que indica el estado y/o carga a los otros módulos dentro de su grupo. Los módulos de PD de DSSL pueden incluir cada uno un módulo de envío de mensajes y un módulo de recepción de mensajes para enviar y recibir respectivamente estos mensajes. En algunas realizaciones, cada módulo de PD de DSSL también puede enviar el mismo o un tipo diferente de mensaje que indica un estado y/o una carga a un módulo de distribución de paquetes de primer nivel. El paquete puede tener cualquiera de varios formatos o protocolos adecuados. En otra realización, los mensajes pueden incluir información de estado sin incluir información de carga. En otra realización más, los mensajes pueden incluir información de carga sin incluir información de estado.

Los estados representan estados y/o estatus de los servidores. Los ejemplos de estatus/estado adecuado incluyen, entre otros vivo (por ejemplo, activo y capaz de procesar paquetes), apagado sin problemas (por ejemplo, capaz de procesar paquetes para conexiones existentes pero no aceptando nuevas conexiones), dejando el grupo, uniéndose al grupo, reiniciando (por ejemplo, en proceso de recuperación de un fallo o reinicio), y similares, y combinaciones de los mismos. En un aspecto, un mensaje recibido de un módulo de PD de DSSL puede indicar implícitamente, en lugar de explícitamente, el estado. Por ejemplo, cuando se recibe un mensaje de un módulo de PD de DSSL, se puede suponer que el servidor correspondiente está vivo y/o activo. Si no se recibe un mensaje de un módulo de PD de DSSL dentro de un período de tiempo determinado, se puede suponer que el servidor correspondiente no está disponible. Las cargas pueden representar cargas asociadas con los servidores. Los ejemplos de cargas adecuadas incluyen, entre otros carga de trabajo actual, número de conexiones, número de flujos, carga en un procesador, carga térmica (por ejemplo, temperatura), cantidad de memoria consumida, tamaño de la tabla de flujo, etc.).

35 El módulo 1 de PD de DSSL tiene un nuevo módulo 1109-1 de ajuste de algoritmo de distribución de paquetes de flujo, el módulo 2 de PD de DSSL tiene un nuevo módulo 1109-2 de ajuste de algoritmo de distribución de paquetes de flujo y el módulo 3 de PD de DSSL tiene un nuevo módulo 1109-3 de flujo de ajuste de algoritmo de distribución de paquetes. En algunas realizaciones, cada uno de los nuevos módulos de ajuste de algoritmo de distribución de paquetes de flujo puede ser operable para ajustar dinámicamente un algoritmo de distribución de paquetes (por ejemplo, un nuevo algoritmo de distribución de paquetes de flujo) basado en los estados y/o cargas. Por ejemplo, en algunas realizaciones, las ponderaciones de un nuevo algoritmo de distribución de paquetes de flujo ponderado, como una elección aleatoria ponderada o un algoritmo de turnos rotativos ponderado, pueden ajustarse dinámicamente basándose en el conjunto de cargas indicadas en los mensajes intercambiados dentro del grupo para tener en cuenta las cargas actuales en los servidores. También se contemplan otras formas de modificar otros tipos de algoritmos conocidos en las técnicas. Como otro ejemplo, los módulos de PD de DSSL de un grupo pueden modificar sus algoritmos de distribución de paquetes para considerar un servidor dado que ya no está disponible y/o para considerar un servidor determinado que se está agregando al grupo.

50 En diferentes realizaciones, la información de estado y/o carga se puede usar para diferentes fines. En algunas realizaciones, la información de estado y/o carga se puede usar para distribuir paquetes entre los servidores de acuerdo con una política u objetivo de distribución de paquetes. Los ejemplos incluyen, entre otros, equilibrio de carga severa, calidad de garantía de servicio, equilibrio de carga térmica, gestión de temperatura, etc. En algunas realizaciones, la información de estado y/o carga se puede usar junto con la implementación de alta disponibilidad. La alta disponibilidad a veces se denomina resistencia, redundancia, protección contra fallos, etc. La alta disponibilidad se usa comúnmente para prevenir o al menos reducir la interrupción del servicio que, de lo contrario, puede ocurrir cuando uno o más servidores dejan de estar disponibles. Los servidores pueden dejar de estar disponibles debido a varias razones potenciales, como, por ejemplo, un fallo de hardware, un fallo de software, apagándose o desconectándose para cambiar el software y/o hardware por parte de un administrador, o similares. Los mensajes intercambiados entre los módulos de PD de DSSL se pueden usar, o pueden ayudar, para detectar cuándo un servidor deja de estar disponible. Posteriormente, los servicios del servidor no disponible pueden transferirse a uno o más servidores para que no haya interrupciones en la prestación de los servicios.

65 Varias configuraciones conocidas de alta disponibilidad o redundancia son adecuadas. Los ejemplos de configuraciones de alta disponibilidad adecuadas incluyen, entre otros configuraciones activas/pasivas en las que un servidor activo tiene un servidor pasivo o en espera completamente redundante que se puede activar si el servidor activo ya no está disponible, configuraciones activas/activas en las que uno o más servidores activos son operables

para hacer una copia de seguridad de otro servidor activo y manejar ese procesamiento de servidores en caso de que ese servidor ya no esté disponible, las configuraciones pasivas M activas N donde los servidores pasivos M (por ejemplo, uno o más) hacen una copia de seguridad de los servidores activos N (por ejemplo, una pluralidad), y similares, y combinaciones de los mismos. En varias realizaciones de ejemplo, se puede proporcionar redundancia por servidor, por grupo, por conjunto de grupos o globalmente (por ejemplo, para todos los grupos).

Hay que considerar un ejemplo de cómo agregar un nuevo servidor activo a un grupo. Los nuevos algoritmos de distribución de paquetes de flujo de los módulos de PD de DSSL pueden modificarse para tener en cuenta el nuevo servidor activo. Por ejemplo, suponiendo que un nuevo algoritmo de distribución de paquetes de flujo basado en hash mapee un módulo n a una dirección IP de origen a n servidores en un grupo, cuando el nuevo servidor activo se agregue al grupo, el módulo puede tomarse n + 1 en lugar de n para considerar el nuevo servidor activo.

Ventajosamente, la adherencia de flujo al servidor para los flujos existentes se puede mantener cuando se agrega un servidor adicional dado a un grupo (es decir, no es necesario cambiar los servidores que manejan los flujos existentes). El nuevo algoritmo de distribución de paquetes de flujo se invoca solo para flujos nuevos y no para flujos existentes. Hay que considerar un flujo de tipo 1 que se maneja con el método de la figura 6. Cuando un módulo de PD de DSSL recibe un paquete para un flujo existente, el módulo de PD de DSSL determina que el paquete es para un flujo existente mediante el examen del encabezado del paquete (por ejemplo, en el bloque 640). El módulo de PD de DSSL determina si el ID de flujo para el flujo existente está en la tabla de flujo del módulo (por ejemplo, en el bloque 646). El servidor que estaba sirviendo al flujo encontrará el ID de flujo asociado en su tabla de flujo correspondiente y aceptará el paquete (por ejemplo, en el bloque 648). Otros servidores no encontrarán el ID de flujo en sus correspondientes tablas de flujo y descartarán el paquete (por ejemplo, en el bloque 647). Por lo tanto, la adherencia del flujo al servidor se mantiene para los flujos existentes cuando se agrega un nuevo servidor a un grupo.

De manera similar, hay que considerar un flujo de tipo 2 que se maneja mediante el método de la figura 7. Cuando se agrega un servidor dado a un grupo, se puede mantener la adherencia de flujo al servidor para los flujos existentes (es decir, no es necesario cambiar los servidores que manejan los flujos existentes). Inicialmente, la tabla de flujo y la tabla de flujo existente de otros correspondientes al servidor agregado están vacías. La tabla de flujo solo se rellena cuando el servidor correspondiente acepta un nuevo flujo para procesarlo. Cuando se recibe un paquete para un flujo existente, el módulo de PD de DSSL correspondiente al servidor agregado determinará que el ID de flujo para el paquete no está en su tabla de flujo correspondiente (por ejemplo, en el bloque 752). La tabla de flujo existente en otros que corresponde al servidor agregado puede actualizarse con los flujos existentes para el grupo antes de tomar decisiones para aceptar o descartar paquetes. Después de que se haya actualizado la tabla de flujo existente en otros, ya que el paquete es para un flujo existente, el módulo de PD de DSSL determinará que el ID de flujo para el paquete se encuentra en su tabla de flujo existente en otros (por ejemplo, en el bloque 754) y hará que el paquete sea descartado (por ejemplo, en el bloque 755).

Un servidor activo también puede ser eliminado de un grupo. Por ejemplo, en una configuración activa/activa, uno de los servidores activos puede fallar o desconectarse. En algunas realizaciones, con el fin de proporcionar una alta disponibilidad, el estado de un servidor dado puede ser mantenido por uno o más servidores activos o pasivos que están destinados a proporcionar resistencia o alta disponibilidad para el servidor dado. Por ejemplo, hay que considerar un flujo de tipo 1 y una configuración de alta disponibilidad activa/pasiva. El servidor pasivo o en espera puede mantener el estado y un conjunto de flujos (por ejemplo, una tabla de flujo) de cada servidor activo del que es responsable para realizar una copia de seguridad. Como otro ejemplo, en una configuración de alta disponibilidad activa/activa, un servidor activo puede mantener al menos esa parte del estado y un conjunto de flujos de otro servidor activo que es responsable de realizar una copia de seguridad. Cuando falla un servidor activo, el servidor responsable de realizar una copia de seguridad de ese servidor puede mover los ID de flujo a su tabla de flujo. Los nuevos algoritmos de distribución de paquetes de flujo también pueden modificarse para reflejar que el servidor se ha eliminado del grupo. Por ejemplo, un algoritmo basado en hash se puede realizar módulo n-1 después de eliminar el servidor, mientras que el módulo n se realizó antes de la eliminación del servidor, donde n es el número de servidores activos en el grupo.

En algunas realizaciones, los mensajes que indican el estatus también pueden proporcionarse a un módulo de distribución de paquetes de primer nivel estático/sin estado. Mientras los mensajes recibidos por el módulo de distribución de paquetes de primer nivel indiquen que al menos un servidor del grupo está activo y disponible para procesar paquetes, el módulo de distribución de paquetes de primer nivel no necesita ajustar necesariamente su enfoque de distribución de paquetes. Más bien, puede seguir distribuyendo paquetes al grupo como antes. Si no se reciben mensajes para un grupo en un período de tiempo determinado, el módulo de distribución de paquetes de primer nivel puede inferir que todos los servidores del grupo no están disponibles y puede cambiar su mapeo de flujo a grupo para que los paquetes se proporcionen a uno o más grupos distintos.

La figura 12 es un diagrama de bloques de una realización de módulos 1210 de PD de DSSL correspondientes a diferentes grupos de servidores que intercambian mensajes entre sí en los que los mensajes indican estados y/o cargas asociados con el grupo de servidores correspondiente. Un módulo 11210-1 de PD de DSSL corresponde a un grupo 1 de servidores 1213-1. Un módulo 2 1210-2 de PD de DSSL corresponde a un grupo 2 de servidores

1213-2. El módulo 1 de PD de DSSL envía un primer mensaje (M1) que indica un primer estado (S1) y una primera carga (L1) asociada con el grupo 1 de servidores. El módulo 2 de PD de DSSL envía un segundo mensaje (M2) que indica un segundo estado (S2) y una segunda carga (L2) asociada con el grupo 2 de servidores. Los módulos de PD de DSSL pueden incluir cada uno un módulo de envío de mensajes y un módulo de recepción de mensajes para enviar y recibir respectivamente estos mensajes. En un aspecto, los estados y las cargas pueden reflejar todo el grupo (por ejemplo, un número total de servidores activos y una carga media). En otro aspecto, los estados y las cargas pueden incluir diferentes estados y cargas correspondientes para cada servidor del grupo. En algunas realizaciones, uno de los módulos de PD de DSSL de un grupo, o un módulo de PD de DSSL dedicado adicional compartido por un grupo, puede recibir información de estado y/o carga sobre otros servidores del grupo a través de los mensajes intercambiados como se describe anteriormente junto con la figura 11, y luego puede promediar y/o combinar los estados y las cargas en un mensaje que se intercambia con los otros grupos.

El módulo 1 de PD de DSSL incluye un módulo 1268-1 de reasignación de grupo de servidores que es operable para reasignar un servidor entre grupos de servidores (por ejemplo, del grupo 1 al grupo 2 y/o del grupo 2 al grupo 1). De manera similar, el módulo 2 de PD de DSSL incluye un módulo 1268-2 de reasignación de grupo de servidores que es operable para reasignar un servidor entre grupos de servidores. En algunas realizaciones, un servidor puede ser reasignado entre los grupos para lograr un objetivo dado y/o de acuerdo con los criterios (por ejemplo, para lograr el equilibrio de carga del servidor, para lograr una calidad de servicio garantizada, para moderar las temperaturas en los servidores, etc.). Por ejemplo, un servidor puede ser reasignado dinámicamente y sin problemas entre grupos para equilibrar de manera más uniforme las cargas entre los servidores de los grupos (por ejemplo, cuando un grupo tiene más recursos del servidor de los que necesita y el otro no tiene suficientes). En el contexto del equilibrio de carga del servidor, la capacidad de reasignar servidores entre grupos permite que se logre un equilibrio de carga del servidor muy uniforme mediante los dos esquemas de nivel divulgados en el presente documento cuando se desee. Como se muestra, en una realización, basándose en los mensajes intercambiados, un servidor 1269 puede eliminarse del grupo 2 y agregarse como servidor 1270 al grupo 1.

Es habitual que un gran porcentaje del tráfico en las redes, incluida Internet, sea atribuible a flujos de volumen relativamente un poco mayores. Estos flujos de mayor volumen a veces se denominan flujos de elefante. Un porcentaje relativamente pequeño del tráfico es atribuible a un número relativamente grande de flujos de menor volumen. Estos flujos de volumen más pequeños a veces se denominan flujos de ratón.

La figura 13 es un diagrama de bloques de una realización de un sistema 1300 de distribución de paquetes de dos niveles que es operable para reenviar de manera eficiente paquetes de flujos de mayor volumen (por ejemplo, flujos de elefante) y flujos de menor volumen (por ejemplo, flujos de ratón). Como se usa en el presente documento, los términos flujos de mayor y menor volumen son términos relativos y no absolutos (es decir, un flujo de mayor volumen es más grande que un flujo de menor volumen). El sistema de distribución de paquetes de dos niveles incluye un módulo 1303 de distribución de paquetes de primer nivel. El módulo de distribución de paquetes de primer nivel incluye un módulo 1371 de distribución de paquetes de mayor volumen de primer nivel con estado que es operable para distribuir paquetes para flujos de volumen relativamente mayores ya identificados. Como se muestra en (1), supóngase que se recibe un paquete entrante de un flujo de mayor volumen aún no identificado en el módulo de distribución de paquetes de primer nivel. El módulo de distribución de paquetes de mayor volumen de primer nivel con estado incluye una tabla 1373 de flujo de mayor volumen que mapea los ID de flujo para flujos de mayor volumen ya identificados (por ejemplo, los flujos de elefante) a los servidores correspondientes ya asignados para manejar los flujos de mayor volumen. La tabla de flujo de mayor volumen representa el estado que se mantiene para los flujos de mayor volumen. Dado que los flujos de mayor volumen son generalmente pocos en número, la cantidad de estado que se necesita almacenar y utilizar no es excesiva. Un módulo 1372 de verificador de tabla de flujo de mayor volumen es operable para verificar la tabla de flujo de mayor volumen para un ID de flujo correspondiente al paquete entrante. Dado que el paquete proviene de un flujo de mayor volumen aún no identificado (por ejemplo, es el primero o uno de los paquetes iniciales del flujo de mayor volumen), aún no se ha determinado que el flujo sea un flujo de mayor volumen y el ID de flujo aún no se ha registrado en la tabla de flujo de mayor volumen.

Como se muestra en (2), al determinar que el ID de flujo del paquete entrante no se encuentra en la tabla de flujo de mayor volumen, se consulta un módulo 1304 de selección de grupo de mayor volumen no identificado/menor volumen estático/sin estado. Este módulo 1304 que es operable para distribuir paquetes para flujos de menor volumen (por ejemplo, flujos de ratón), como su nombre lo indica, así como para flujos de mayor volumen (por ejemplo, flujos de elefante) que aún no se han identificado y registrado como tales en la tabla de flujo de mayor volumen. El módulo puede usar un enfoque estático/sin estado similar a los descritos en otra parte del presente documento para seleccionar un grupo de servidores. Como se muestra en (3), el paquete puede ser multidifundido, difundido o de otra manera distribuido a un conjunto de módulos 1308-1, 1308-2 de distribución de paquetes de segundo nivel distribuidos en el grupo seleccionado.

Un módulo 1308-1 de distribución de paquetes de segundo nivel distribuido incluye un módulo 1374 de desviación de gran volumen. Si el ID de flujo para el paquete estaba en la tabla de flujo de mayor volumen, y el paquete se distribuyó a un servidor seleccionado, entonces el módulo de desviación de mayor volumen es operable para hacer que el paquete sea proporcionado al servidor 1312 seleccionado para su procesamiento. Sin embargo, como este no es el caso para el paquete actualmente considerado de un flujo de mayor volumen aún no identificado, como se

muestra en (4) se usa un módulo 1309 de selección de servidor de segundo nivel distribuido para seleccionar un servidor en el grupo seleccionado al que se distribuirá el paquete. Se puede usar un enfoque similar o igual a los descritos en otro lugar en el presente documento para la selección del servidor. Como se muestra en (5), un módulo 1375 de detección de mayor volumen es operable para detectar flujos de mayor volumen. Se conocen varios enfoques en la técnica para detectar flujos de elefante u otros flujos de mayor volumen. Los enfoques simples, por ejemplo, son usar el historial pasado, para designar ciertas direcciones IP de origen como flujos de mayor volumen, etc. Si el flujo se detecta como un flujo de ratón, no es necesario tomar medidas adicionales. Sin embargo, supóngase, para fines de ilustración, que el módulo de detección de mayor volumen determina que el paquete es para un mayor flujo de volumen. Como se muestra en (6), se invoca un módulo 1376 de modificación de la tabla de flujo de mayor volumen. Como se muestra en (7), el módulo de modificación es operable para hacer que se modifique la tabla 1373 de flujo de mayor volumen. Por ejemplo, el módulo de modificación puede indicar a la tabla de flujo de mayor volumen que agregue una entrada que asigne el ID de flujo para el paquete al servidor seleccionado.

Como se muestra en (8), cuando un paquete entrante subsiguiente del ahora identificado flujo de mayor volumen se recibe en el módulo de distribución de paquetes de primer nivel, el módulo de verificador de la tabla de flujo de mayor volumen comprobará la tabla de flujo de mayor volumen. El ID de flujo para el ahora identificado flujo de mayor volumen se conserva ahora como estado en la tabla de flujo de mayor volumen junto con el servidor correspondiente seleccionado para procesar los paquetes del flujo de mayor volumen. Como se muestra en (9), el módulo de distribución de paquetes de mayor volumen de primer nivel con estado puede unidifundir o distribuir el paquete al servidor seleccionado indicado en la tabla de flujo de mayor volumen. Por consiguiente, el módulo de distribución de paquetes de primer nivel es operable para realizar tanto la distribución de paquetes de flujo de unidifusión de estado a servidor para una porción de los flujos (por ejemplo, flujos de mayor volumen), como una distribución de paquetes de multidifusión de flujo a grupo estático/sin estado para una porción de los flujos (por ejemplo, flujos de menor volumen). En otras realizaciones, tal enfoque híbrido se puede usar para otros criterios además de flujos de mayor/menor volumen (por ejemplo, tipos particulares de tráfico o conexiones).

El enfoque descrito anteriormente puede ofrecer ciertas ventajas potenciales. Por un lado, una gran parte del tráfico (por ejemplo, la de los flujos de elefante) se puede detectar y unidifundir al servidor correspondiente en lugar de tener que multidifundir los paquetes a todos los servidores de un grupo. Dependiendo de la cantidad de servidores dentro del grupo, esto puede ayudar a reducir significativamente la cantidad de tráfico de paquetes en la interconexión o el acoplamiento entre el primer y el segundo nivel. Además, algunas de las tareas de computación relativamente más intensivas que de otro modo se realizarían en el segundo nivel con estado pueden evitarse. Debido a que la tabla de flujo de mayor volumen puede almacenar el estado de los flujos que son relativamente pequeños, la cantidad de estado no es excesiva (por ejemplo, la tabla de flujo de mayor volumen puede ser relativamente pequeña), por lo que la memoria para almacenar el estado y el tiempo y el procesamiento para acceder al estado tampoco son excesivos. Además, las tareas relativamente más intensivas en computación de detectar flujos de mayor volumen, así como la distribución de paquetes con estado, se han descargado al módulo de distribución de paquetes de segundo nivel distribuido.

Habiendo descrito varias realizaciones, ahora se describirán algunas realizaciones alternativas adicionales. Se han descrito realizaciones con respecto a los sistemas de distribución de paquetes de dos niveles. Sin embargo, otras realizaciones pueden tener tres o más niveles. Por ejemplo, un primer nivel puede mapear sin estado flujos a grupos más grandes, un segundo nivel puede mapear sin estados los flujos a subgrupos más pequeños dentro de los grupos más grandes, y un tercer nivel puede mapear sin estado los flujos a servidores individuales dentro de los subgrupos.

Como otro ejemplo, en algunas realizaciones, se puede proporcionar un tercer nivel de distribución de paquetes para soportar el encadenamiento de servicios entre los servidores y/o el hardware en el que se implementan los servidores. En el encadenamiento de servicios, un paquete puede distribuirse a un primer servidor y/o primer hardware de implementación del servidor usando el enfoque de distribución de paquetes de dos niveles descrito anteriormente. Posteriormente, después del procesamiento de paquetes en el primer servidor y/o primer hardware de implementación del servidor, se puede usar un tercer nivel de distribución de paquetes para distribuir el paquete a un segundo servidor y/o segundo hardware de implementación del servidor del mismo elemento de red. El primer servidor y/o el primer hardware de implementación del servidor, después de procesar el paquete, pueden reenviar o distribuir el paquete al segundo servidor y/o al segundo hardware de implementación del servidor del mismo o diferente grupo dentro del mismo elemento de red. Como un ejemplo simple, supongamos que un primer servidor implementado en una primera tarjeta de servicio pertenece al grupo 1, y un $N^{\text{ésimo}}$ servidor implementado en una $N^{\text{ésima}}$ tarjeta de servicio pertenece al grupo N, y un paquete ha de procesarse primero en el primer servidor, y luego posteriormente se enviará al servidor N para su posterior procesamiento. También supongamos que el flujo es un flujo de tipo 1, que no es necesario. Después de procesar el paquete en el primer servidor, un módulo de distribución de paquetes de tercer nivel estático/sin estado correspondiente al primer servidor y/o al primer hardware de implementación del servidor puede seleccionar un grupo (en este caso, el grupo N). Se puede usar un enfoque estático/sin estado similar al descrito anteriormente para el nivel 1. El paquete puede ser difundido o multidifundido al grupo N. Se puede realizar una segunda instancia de la distribución de paquetes de segundo nivel distribuida en un módulo de distribución de paquetes de segundo nivel distribuido correspondiente al servidor $N^{\text{ésimo}}$ para

seleccionar el servidor N^{ésimo} para procesar el paquete para su procesamiento. El servidor N^{ésimo} puede procesar el paquete. Si se desea, se puede realizar un encadenamiento adicional repitiendo generalmente el proceso descrito anteriormente.

- 5 Las realizaciones se han descrito con respecto a los servidores que proporcionan servicios a los clientes. Otras realizaciones son aplicables a la distribución de paquetes entre un conjunto de clientes. Por ejemplo, un conjunto de clientes en uno o más elementos de red (por ejemplo, cada uno en una hoja de servidor diferente) puede iniciar conexiones con nodos externos (por ejemplo, servidores). Los sistemas de distribución de paquetes de dos niveles divulgados en el presente documento pueden usarse para reenviar paquetes desde los nodos externos (por ejemplo, servidores) a los clientes que mantienen la adherencia del flujo.

15 En la descripción anterior, a los efectos de la explicación, se han expuesto numerosos detalles específicos para proporcionar una comprensión completa de las realizaciones de la invención. Sin embargo, será evidente para un experto en la técnica que una o más realizaciones adicionales pueden ponerse en práctica sin algunos de estos detalles específicos. Las realizaciones particulares descritas no se proporcionan para limitar la invención sino para ilustrarla. Los expertos en la técnica reconocerán que la invención puede ponerse en práctica con modificaciones y alteraciones dentro del espíritu y alcance de las reivindicaciones adjuntas. Por lo tanto, la descripción debe considerarse como ilustrativa en lugar de limitativa. El alcance de la invención no debe determinarse por los ejemplos específicos proporcionados anteriormente, sino únicamente por las siguientes reivindicaciones.

20 En otros casos, se han mostrado circuitos, estructuras, dispositivos y operaciones bien conocidos en forma de diagrama de bloques o sin detalles para evitar que se entorpezca la comprensión de la descripción. Además, cuando se consideran apropiados, los números de referencia o las porciones terminales de los números de referencia se han repetido entre las figuras para indicar elementos correspondientes o análogos, que pueden tener opcionalmente características similares.

30 Las referencias en la especificación a "una realización", "una realización", "una realización de ejemplo", etc., indican que la realización descrita puede incluir una característica, estructura o característica particular, pero cada realización no necesariamente incluye la característica, estructura, o característica particular. Además, tales frases no se refieren necesariamente a la misma realización. Además, cuando una característica, estructura o característica particular se describe en relación con una realización, se afirma que es del conocimiento de un experto en la técnica afectar dicha característica, estructura o característica en relación con otras realizaciones, ya sea descrita explícitamente o no.

35 En la descripción y las reivindicaciones, se pueden usar los términos "acoplado" y "conectado", junto con sus derivados. Debe entenderse que estos términos no están destinados a ser sinónimos entre sí. "Acoplado" se usa para indicar que dos o más elementos, que pueden estar o no en contacto físico o eléctrico directo entre sí, cooperan o interactúan entre sí. "Conectado" se usa para indicar el establecimiento de comunicación entre dos o más elementos que están acoplados entre sí.

40 Se han descrito varias operaciones y métodos. Algunos de los métodos se han descrito en una forma básica en los diagramas de flujo, pero las operaciones se pueden agregar y/o eliminar opcionalmente de los métodos. Además, mientras que los diagramas de flujo muestran un orden particular de las operaciones de acuerdo con realizaciones de ejemplo, debe entenderse que ese orden particular es de ejemplo. Realizaciones alternativas pueden realizar opcionalmente las operaciones en orden diferente, combinar ciertas operaciones, superponer ciertas operaciones, etc. Se pueden hacer muchas modificaciones y adaptaciones a los métodos y se contemplan.

50 Se pueden implementar diferentes realizaciones de la invención usando diferentes combinaciones de software, firmware y/o hardware. Por lo tanto, las técnicas mostradas en las figuras se pueden implementar usando el código y los datos almacenados y ejecutados en uno o más dispositivos electrónicos (por ejemplo, una estación final, un elemento de red). Tales dispositivos electrónicos almacenan y comunican (internamente y/o con otros dispositivos electrónicos a través de una red) el código y los datos usando medios legibles por ordenador, tales como medios de almacenamiento tangibles, no transitorios legibles por ordenador y comunicaciones transitorias no tangibles legibles por ordenador o medios de transmisión. Algunos ejemplos representativos de medios de almacenamiento tangibles, no transitorios legibles por ordenador incluyen, entre otros, discos magnéticos; discos ópticos; memoria de acceso aleatorio (RAM), como RAM estática (SRAM) y RAM dinámica (DRAM), memoria de solo lectura (ROM), como ROM programable (PROM), ROM borrrable y programable (EPROM) y ROM borrrable y programable eléctricamente (EEPROM), dispositivos de memoria flash; memoria de cambio de fase, y similares. El medio de almacenamiento tangible puede incluir uno o más materiales físicos sólidos o tangibles, como, por ejemplo, un material semiconductor, un material de cambio de fase, un material magnético, etc. Las realizaciones pertenecen a un artículo de fabricación que incluye un medio de almacenamiento legible por ordenador no transitorio tangible que almacena una secuencia de instrucciones que, si son ejecutadas por una máquina (por ejemplo, un elemento de red, un conmutador, un enrutador, una estación final, un anfitrión, un sistema informático o un dispositivo electrónico con al menos un microprocesador causa o resulta en que la máquina realice una o más operaciones o métodos divulgados en el presente documento. Algunos ejemplos representativos de medios de transmisión legibles por ordenador transitorios no tangibles incluyen, entre otros, señales eléctricas, ópticas, acústicas u otras formas

5 propagadas, tales como ondas portadoras, señales infrarrojas y señales digitales. Además, dichos dispositivos electrónicos suelen incluir un conjunto de uno o más procesadores acoplados a uno o más componentes, como uno o más dispositivos de almacenamiento (medios de almacenamiento no transitorios legibles por máquina), dispositivos de entrada/salida del usuario (por ejemplo, un teclado, una pantalla táctil, y/o una pantalla), y conexiones de red. El acoplamiento del conjunto de procesadores y otros componentes se realiza normalmente a través de uno o más buses y puentes (también denominados controladores de bus). Por lo tanto, el dispositivo de almacenamiento de un dispositivo electrónico dado generalmente almacena el código y/o los datos para su ejecución en el conjunto de uno o más procesadores de ese dispositivo electrónico.

REIVINDICACIONES

- 1.- Un método, que se realiza en uno o más elementos de red que están en comunicación entre los clientes que transmiten paquetes y servidores, para distribuir los paquetes entre los servidores que han de procesar los paquetes, al tiempo que proporciona la adherencia de los flujos de los paquetes a los servidores asignados a procesar los paquetes de los flujos, el método comprende los pasos de:
- 5 recibir un paquete de un flujo en un módulo (531) de distribución de paquetes de primer nivel estático;
- 10 seleccionar estáticamente un grupo de servidores, de entre una pluralidad de grupos de servidores, para el paquete del flujo con el módulo de distribución de paquetes de primer nivel estático, en el que la selección estática del grupo de servidores no implica el uso de un estado que asigna el paquete del flujo al grupo seleccionado de servidores (532);
- 15 distribuir el paquete del flujo a un sistema (533) de distribución de paquetes de segundo nivel con estado distribuido;
- seleccionar con estado un servidor del grupo seleccionado de servidores que ha de procesar el paquete del flujo con el sistema de distribución de paquetes de segundo nivel con estado distribuido al acceder al estado que asigna el procesamiento de paquetes del flujo al servidor seleccionado (534); y
- 20 distribuir el paquete del flujo al servidor seleccionado (535), caracterizado porque
- el paso de distribución del paquete del flujo al sistema de distribución de paquetes de segundo nivel con estado distribuido comprende un paso de distribución del paquete del flujo a cada uno de una pluralidad de módulos de
- 25 distribución de paquetes de segundo nivel con estado distribuidos que cada uno corresponde a un servidor diferente del grupo seleccionado de servidores.
- 2.- El método de la reivindicación 1, en el que el sistema de distribución de paquetes de segundo nivel con estado distribuido se implementa en un hardware que es distinto del hardware usado para implementar el módulo de
- 30 distribución de paquetes de primer nivel estático.
- 3.- El método de la reivindicación 1, en el que el sistema de distribución de paquetes de segundo nivel con estado distribuido se implementa en el hardware que incluye el hardware usado para implementar el grupo seleccionado de
- 35 servidores.
- 4.- El método de la reivindicación 1, en el que cada uno de los módulos de distribución de paquetes de segundo nivel con estado distribuidos realiza los pasos de:
- 40 determinar que el paquete no es para un nuevo flujo al examinar un encabezado del paquete;
- determinar si un identificador de flujo correspondiente al flujo para el paquete se incluye en una tabla de flujo del módulo de distribución de paquetes de segundo nivel con estado distribuido que realiza la determinación, en el que la tabla de flujo asigna identificadores de flujo al servidor que corresponde al módulo de distribución de paquetes de
- 45 segundo nivel con estado distribuido que está haciendo la determinación; y
- uno de: (1) hacer que el paquete sea descartado si el identificador de flujo no está incluido en la tabla de flujo; y (2) hacer que el paquete sea aceptado para su procesamiento por el servidor correspondiente si el identificador de flujo está incluido en la tabla de flujo.
- 50 5.- El método de la reivindicación 4, en el que cada uno de los módulos de distribución de paquetes de segundo nivel con estado distribuido realiza los pasos de:
- recibir un segundo paquete;
- 55 determinar que el segundo paquete es para un nuevo flujo al examinar un encabezado del segundo paquete;
- seleccionar un servidor para procesar el segundo paquete con un algoritmo de distribución de paquetes, en el que el algoritmo de distribución de paquetes es operable para seleccionar el servidor de manera consistente con los algoritmos de distribución de paquetes de otros módulos de distribución de paquetes de segundo nivel con estado
- 60 distribuidos para el grupo seleccionado;
- determinar si el servidor seleccionado para procesar el segundo paquete corresponde al módulo de distribución de paquetes de segundo nivel con estado distribuido que seleccionó el servidor; y
- 65 uno de (1) hacer que el paquete sea descartado si el servidor seleccionado no corresponde al módulo de distribución de paquetes de segundo nivel con estado distribuido que seleccionó el servidor; y (2) registrar un nuevo identificador

de flujo en una tabla de flujo del módulo de distribución de paquetes que seleccionó el servidor y hacer que el servidor seleccionado acepte el paquete para procesarlo si el servidor seleccionado se corresponde con el módulo de distribución de paquetes de segundo nivel con estado distribuido que seleccionó el servidor.

5 6.- El método de la reivindicación 1, en el que cada uno de los módulos de distribución de paquetes de segundo nivel con estado distribuidos realiza los pasos de:

10 determinar si un identificador de flujo correspondiente al flujo para el paquete se incluye en una tabla de flujo del módulo de distribución de paquetes de segundo nivel con estado distribuido, en el que la tabla de flujo asigna identificadores de flujo al servidor que corresponde al módulo de distribución de paquetes de segundo nivel con estado distribuido; y

15 uno de: (1) hacer que el paquete sea aceptado para su procesamiento por el servidor correspondiente si el identificador de flujo está incluido en la tabla de flujo; y (2) hacer que el paquete sea descartado si el identificador de flujo no está incluido en la tabla de flujo y el identificador de flujo está incluido en una tabla de flujo existente en otros del módulo de distribución de paquetes de segundo nivel con estado distribuido, en el que la tabla de flujo existente en otros enumera los identificadores de flujo asignados a los servidores del grupo que no sea el servidor correspondiente al módulo de distribución de paquetes de segundo nivel con estado distribuido.

20 7.- El método de la reivindicación 1, en el que el sistema de distribución de paquetes de segundo nivel con estado distribuido incluye una pluralidad de módulos de distribución de paquetes de segundo nivel con estado distribuidos que corresponden cada uno a un servidor diferente del grupo seleccionado de servidores, y en el que un módulo de distribución de paquetes de segundo nivel con estado distribuido realiza los pasos de:

25 enviar un mensaje a otros módulos de distribución de paquetes de segundo nivel con estado distribuidos que corresponden a los servidores del grupo seleccionado, en el que el mensaje indica un estado y una carga de un servidor correspondiente al módulo de distribución de paquetes de segundo nivel con estado distribuido;

30 recibir mensajes de los otros módulos de distribución de paquetes de segundo nivel con estado distribuidos, en el que los mensajes recibidos indican los estados y las cargas de los servidores correspondientes a los otros módulos de distribución de paquetes de segundo nivel con estado distribuidos; y

35 ajustar un algoritmo de distribución de paquetes usado por el módulo de distribución de paquetes de segundo nivel con estado distribuido basándose en los estados y las cargas indicadas en los mensajes recibidos.

40 8.- El método de la reivindicación 1, en el que el sistema de distribución de paquetes de segundo nivel con estado distribuido incluye al menos un módulo de distribución de paquetes de segundo nivel con estado distribuido correspondiente al grupo seleccionado de servidores y al menos un módulo de distribución de paquetes de segundo nivel con estado distribuido correspondiente a un segundo grupo de servidores, y en el que el método comprende además los pasos de:

45 recibir un mensaje en al menos un módulo de distribución de paquetes de segundo nivel con estado distribuido correspondiente al grupo seleccionado de servidores de al menos un módulo de distribución de paquetes de segundo nivel con estado distribuido correspondiente a un segundo grupo de servidores, en el que el mensaje recibido indica al menos una carga asociada con el segundo grupo de servidores; y

reasignar un servidor del segundo grupo al grupo seleccionado basándose al menos en parte en al menos una carga indicada en el mensaje recibido y al menos una carga asociada con el grupo seleccionado de servidores.

50 9.- El método de la reivindicación 1, en el que el método se realiza en un único elemento de red que tiene una pluralidad de tarjetas de línea y una pluralidad de tarjetas seleccionadas de las tarjetas de control, las tarjetas de servicio y las tarjetas de recursos, en el que hardware usado para implementar el módulo de distribución de paquetes de primer nivel estático comprende el hardware de las tarjetas de línea, y en el que el hardware a través del cual se distribuye el sistema de distribución de paquetes de segundo nivel con estado distribuido comprende el hardware de las tarjetas seleccionadas de las tarjetas de control, las tarjetas de servicio y las tarjetas de recursos.

10.- El método de la reivindicación 1, que comprende además los pasos de:

60 recibir un segundo paquete de un segundo flujo en el módulo de distribución de paquetes de primer nivel estático, en el que el segundo flujo es un flujo de mayor volumen;

65 seleccionar con estado un servidor para el segundo paquete del segundo flujo con el módulo de distribución de paquetes de primer nivel estático, en el que seleccionar con estado el servidor para el segundo paquete incluye el uso de un estado que asigne flujos de mayor volumen a los servidores; y

distribuir el segundo paquete del segundo flujo al servidor seleccionado.

11.- El método de la reivindicación 1, en el que cada uno de la pluralidad de grupos de servidores forma una red de área local virtual (VLAN) diferente.

5 12.- El método de la reivindicación 1, en el que seleccionar con estado el servidor del grupo seleccionado de servidores que ha de procesar el paquete del flujo comprende seleccionar el servidor para equilibrar las cargas en los servidores del grupo seleccionado, en el que la selección estática del grupo de servidores comprende realizar una función hash con información del paquete para seleccionar estáticamente el grupo de servidores, en el que cada uno de los servidores del grupo seleccionado comprende uno de una hoja de servidor diferente, una tarjeta de control diferente, una tarjeta de servicio diferente, un procesador diferente y un núcleo de procesador diferente.

13.- El método de la reivindicación 1, que agrega además un servidor al grupo seleccionado de servidores, en el que agregar el servidor al grupo seleccionado no altera la adherencia de los flujos existentes a los servidores del grupo seleccionado.

15 14.- El método de la reivindicación 1, que comprende además, después de procesar el paquete con el servidor seleccionado:

20 seleccionar estáticamente un segundo grupo de servidores de entre la pluralidad de grupos de servidores; y distribuir el paquete al segundo grupo de servidores.

25 15.- Un sistema (100) de distribución de paquetes multinivel, el sistema de distribución de paquetes multinivel configurado para ser implementado en uno o más elementos de la red en una red entre clientes que transmiten paquetes y servidores, de distribuir los paquetes entre los servidores (112), al mismo tiempo que proporciona una adherencia de los flujos de los paquetes a los servidores, el sistema de distribución de paquetes multinivel comprendiendo:

30 un módulo (103) de distribución de paquetes de primer nivel estático configurado para recibir un paquete de un flujo, el módulo (103) de distribución de paquetes de primer nivel estático operable para seleccionar estáticamente un grupo de servidores, de entre la pluralidad de grupos de servidores, para el paquete del flujo, en el que la selección estática del grupo de servidores no implica el uso de un estado que asigna el paquete del flujo al grupo seleccionado de servidores; y

35 un sistema (107) de distribución de paquetes de segundo nivel con estado distribuido en comunicación con el módulo (103) de distribución de paquetes de primer nivel estático, en el que el sistema (107) de distribución de paquetes de segundo nivel con estado distribuido incluye una pluralidad de módulos (108) de distribución de paquetes de segundo nivel con estado distribuidos, el sistema de distribución de paquetes de segundo nivel con estado distribuido operable para seleccionar con estado un servidor del grupo seleccionado de servidores que ha de procesar el paquete del flujo, en el que seleccionar con estado el servidor implica el uso de un estado que asigna el flujo al servidor seleccionado, el sistema de distribución de paquetes de segundo nivel con estado distribuido operable para distribuir el paquete al servidor seleccionado, caracterizado porque

45 la pluralidad de módulos (108) de distribución de paquetes de segundo nivel con estado distribuidos comprende una pluralidad de módulos (108) de distribución de paquetes de segundo nivel con estado distribuidos que corresponden cada uno a un servidor diferente del grupo seleccionado de servidores.

50 16.- El sistema de distribución de paquetes multinivel de la reivindicación 15, en el que el sistema de distribución de paquetes de segundo nivel con estado distribuido se implementa en un hardware que es distinto del hardware usado para implementar el módulo de distribución de paquetes de primer nivel estático.

55 17.- El sistema de distribución de paquetes multinivel de la reivindicación 15, en el que el sistema de distribución de paquetes de segundo nivel con estado distribuido se implementa en un hardware que incluye el hardware usado para implementar el grupo seleccionado de servidores.

60 18.- El sistema de distribución de paquetes multinivel de la reivindicación 15, en el que el sistema de distribución de paquetes multinivel está configurado para implementarse en un único elemento de red que tiene una pluralidad de tarjetas de línea y una pluralidad de tarjetas seleccionadas de las tarjetas de control, las tarjetas de servicio y las tarjetas de recursos, en el que el hardware usado para implementar el módulo de distribución de paquetes de primer nivel estático comprende el hardware de las tarjetas de línea, y en el que el sistema de distribución de paquetes de segundo nivel con estado distribuido se distribuye por hardware de las tarjetas seleccionadas de las tarjetas de control, las tarjetas de servicio y las tarjetas de recursos.

65 19.- El sistema de distribución de paquetes multinivel de la reivindicación 15, en el que un módulo de distribución de paquetes de segundo nivel con estado distribuido correspondiente a un servidor comprende:

- un módulo de examen de paquetes operable para determinar que el paquete no es para un nuevo flujo al examinar un encabezado del paquete;
- 5 una tabla de flujo que asigna identificadores de flujo al servidor correspondiente al módulo de distribución de paquetes de segundo nivel con estado distribuido;
- un módulo de verificador de tabla de flujo operable para determinar si un identificador de flujo correspondiente al flujo para el paquete está incluido en la tabla de flujo; y
- 10 un módulo de determinación operable a uno de: (1) hacer que el paquete sea descartado si el identificador de flujo no está incluido en la tabla de flujo; y (2) hacer que el paquete sea aceptado para su procesamiento por el servidor correspondiente si el identificador de flujo está incluido en la tabla de flujo.
- 15 20.- El sistema de distribución de paquetes multinivel de la reivindicación 19, en el que el módulo de examen de paquetes es operable para recibir un segundo paquete y operable para determinar que el segundo paquete es para un nuevo flujo al examinar un encabezado del segundo paquete, y que comprende además:
- un algoritmo de distribución de paquetes operable para seleccionar un servidor para procesar el segundo paquete, en el que el algoritmo de distribución de paquetes es operable para seleccionar el servidor de manera consistente con los algoritmos de distribución de paquetes de otros módulos de distribución de paquetes de segundo nivel con estado distribuidos para el grupo seleccionado;
- 20 un módulo de determinación operable para determinar si el servidor seleccionado para procesar el segundo paquete es el mismo que el servidor correspondiente al módulo de distribución de paquetes de segundo nivel con estado distribuido; y
- 25 un módulo de determinación operable a uno de: (1) hacer que el paquete sea descartado si el servidor seleccionado no es el mismo que el servidor correspondiente al módulo de distribución de paquetes de segundo nivel con estado distribuido; y (2) registrar un nuevo identificador de flujo en una tabla de flujo del módulo de distribución de paquetes y hacer que el servidor seleccionado acepte el paquete para su procesamiento si el servidor seleccionado es el mismo que el servidor correspondiente al módulo de distribución de paquetes de segundo nivel con estado distribuido.
- 30 21.- El sistema de distribución de paquetes multinivel de la reivindicación 15, en el que un módulo de distribución de paquetes de segundo nivel con estado distribuido correspondiente a un servidor comprende:
- 35 una tabla de flujo que asigna identificadores de flujo al servidor que corresponde al módulo de distribución de paquetes de segundo nivel con estado distribuido;
- 40 un módulo de verificador de tabla de flujo operable para determinar si un identificador de flujo correspondiente al flujo para el paquete está incluido en la tabla de flujo; y
- una tabla de flujo existente en otros, en el que la tabla de flujo existente en otros incluye identificadores de flujo asignados a los servidores del grupo seleccionado que no sea el servidor correspondiente al módulo de distribución de paquetes de segundo nivel con estado distribuido; y
- 45 un módulo de determinación operable a uno de: (1) hacer que el paquete sea aceptado para su procesamiento por el servidor correspondiente si el identificador de flujo está incluido en la tabla de flujo; y (2) hacer que el paquete sea descartado si el identificador de flujo no está incluido en la tabla de flujo y el identificador de flujo está incluido en la tabla de flujo existente en otros.
- 50 22.- El sistema de distribución de paquetes multinivel de la reivindicación 15, en el que el sistema de distribución de paquetes de segundo nivel con estado distribuido incluye una pluralidad de módulos de distribución de paquetes de segundo nivel con estado distribuidos que cada uno corresponde a un servidor diferente del grupo seleccionado de servidores, y en el que un módulo de distribución de paquetes de segundo nivel con estado distribuido correspondiente a un servidor comprende:
- 55 un módulo de envío de mensajes operable para enviar un mensaje a otros módulos de distribución de paquetes de segundo nivel de estado distribuidos del grupo seleccionado, en el que el mensaje indica un estado y una carga del servidor correspondiente;
- 60 un módulo de recepción de mensajes operable para recibir mensajes de los otros módulos de distribución de paquetes de segundo nivel con estado distribuidos, en el que los mensajes recibidos indican estados y cargas de los servidores correspondientes; y
- 65

un módulo de ajuste de algoritmo de distribución de paquetes operable para ajustar un algoritmo de distribución de paquetes usado por el módulo de distribución de paquetes de segundo nivel con estado distribuido basándose en los estados y las cargas indicados en los mensajes recibidos.

- 5 23.- El sistema de distribución de paquetes multinivel de la reivindicación 15, en el que el sistema de distribución de
paquetes de segundo nivel con estado distribuido incluye al menos un módulo de distribución de paquetes de
segundo nivel con estado distribuido correspondiente al grupo seleccionado de servidores y al menos un módulo de
distribución de paquetes de segundo nivel con estado distribuido correspondiente a un segundo grupo de servidores,
y en el que el al menos un módulo de distribución de paquetes de segundo nivel con estado distribuido
10 correspondiente al grupo seleccionado de servidores comprende:

un módulo de recepción de mensajes operable para recibir un mensaje en al menos dicho módulo de distribución de
paquetes de segundo nivel con estado distribuido correspondiente al grupo seleccionado de servidores de al menos
15 un módulo de distribución de paquetes de segundo nivel con estado distribuido correspondiente a un segundo grupo
de servidores, en el que el mensaje recibido indica al menos una carga asociada con el segundo grupo de
servidores; y

un módulo de reasignación de grupo de servidores operable para reasignar un servidor del segundo grupo al grupo
seleccionado basándose al menos en parte en al menos dicha carga indicada en el mensaje recibido y al menos una
20 carga asociada con el grupo seleccionado de servidores.

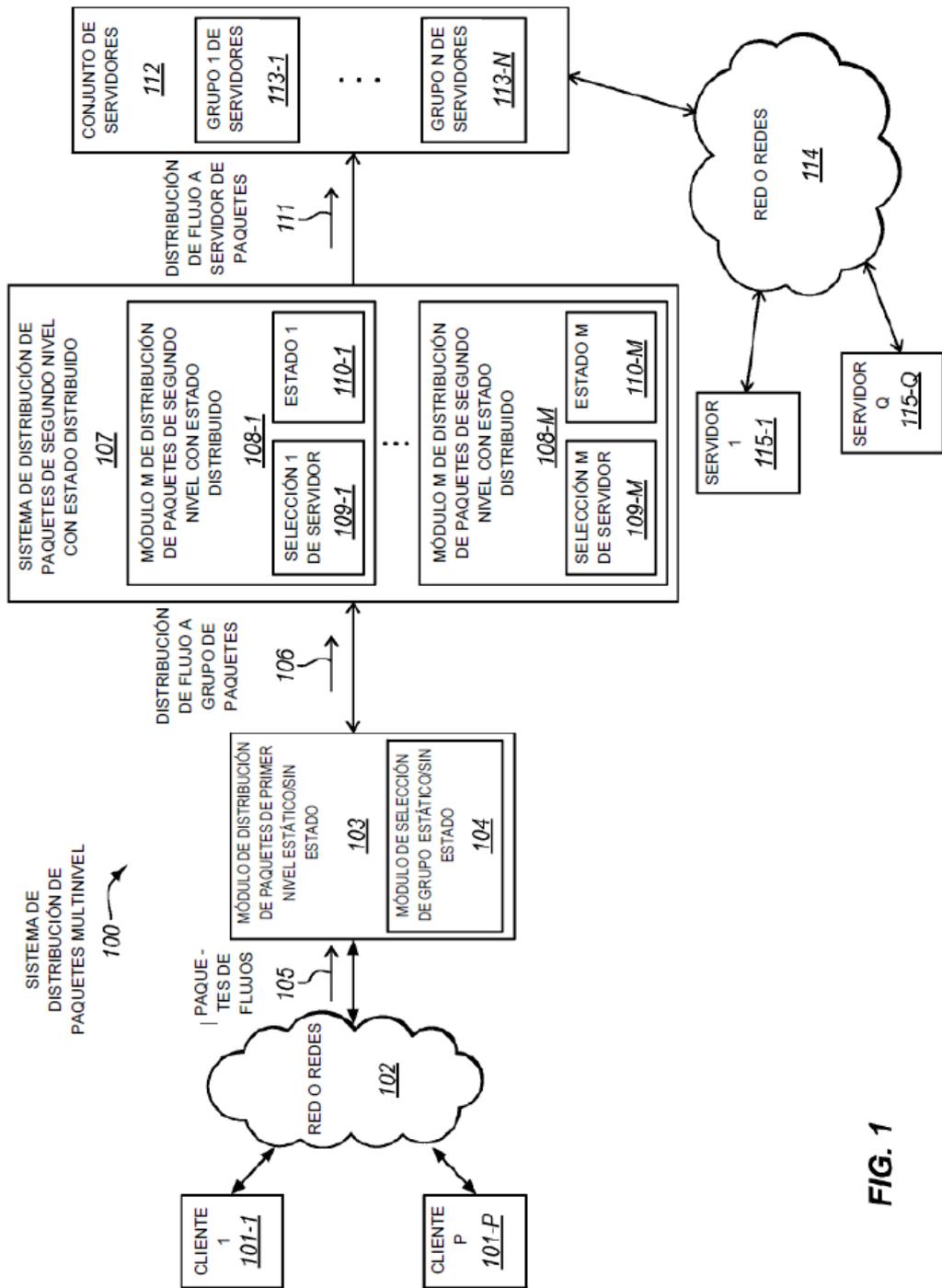


FIG. 1

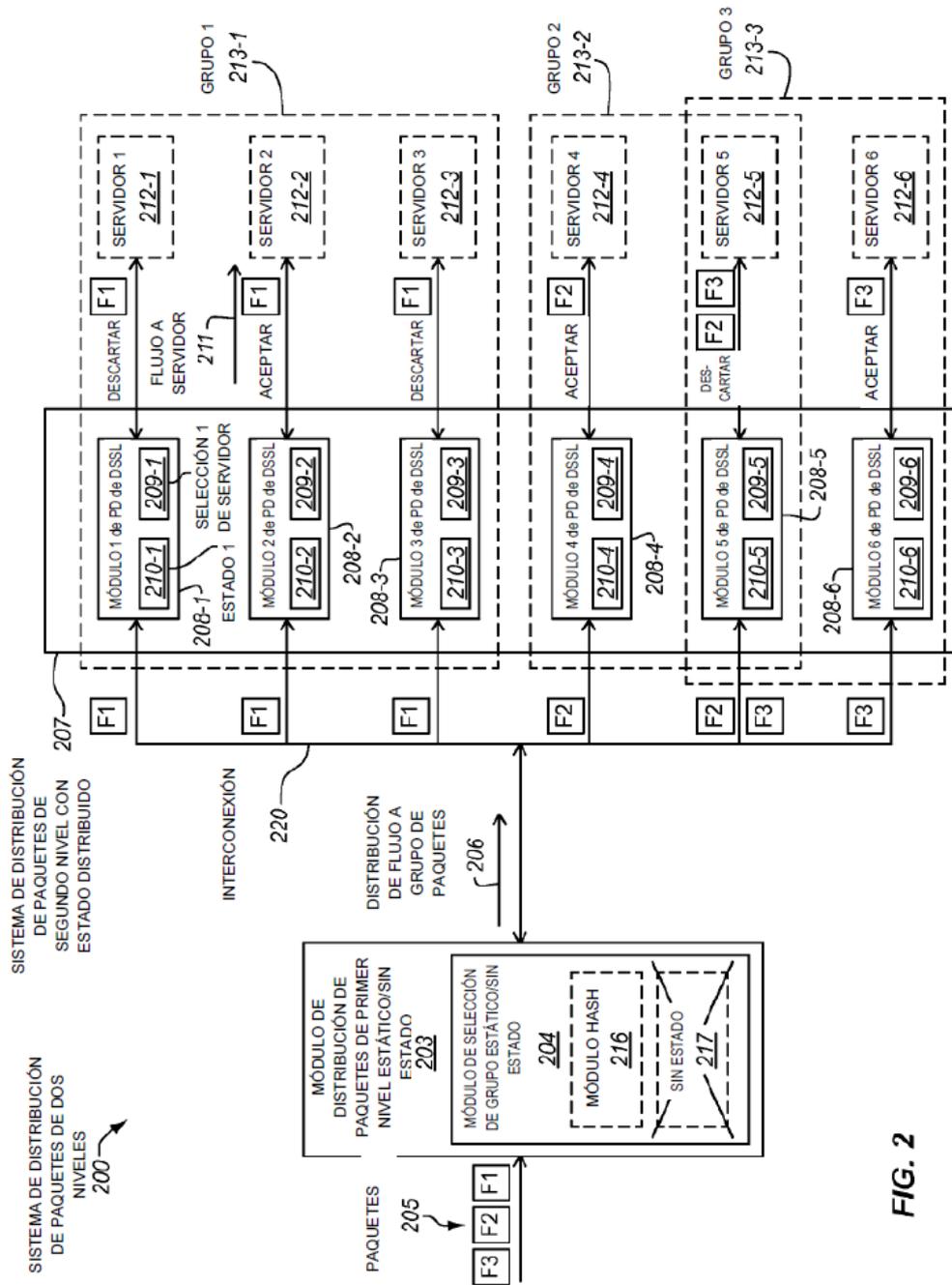


FIG. 2

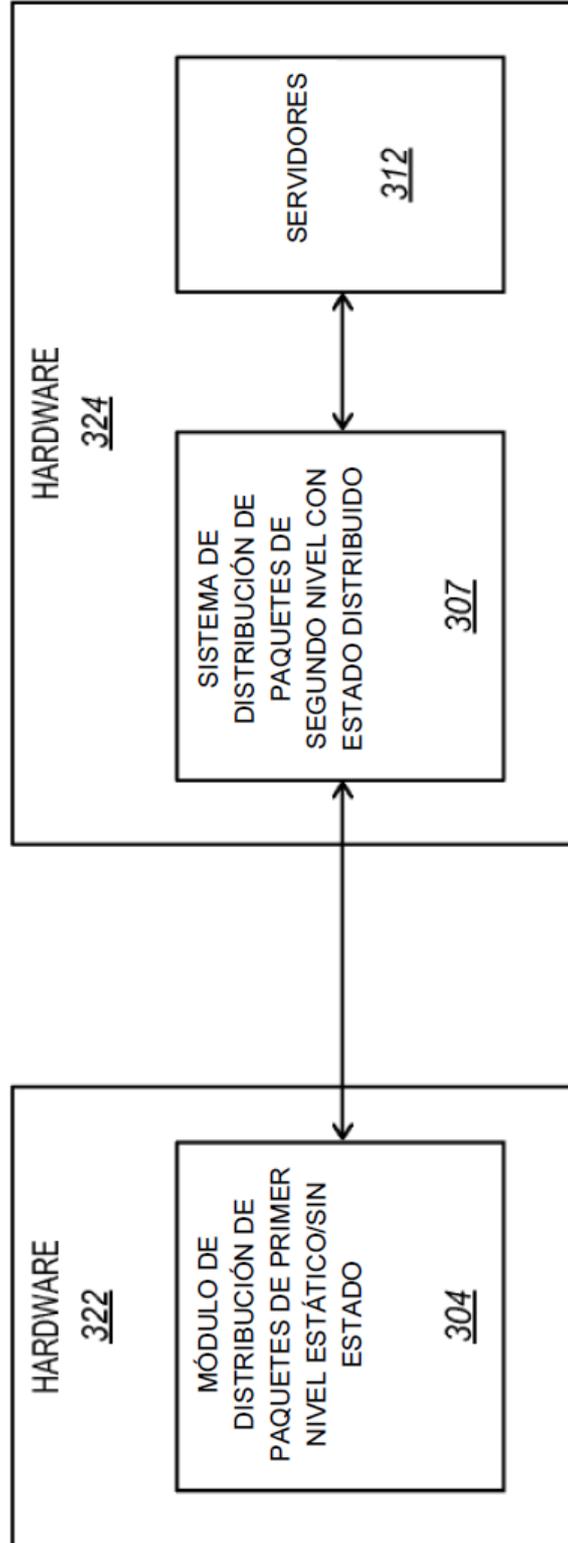


FIG. 3

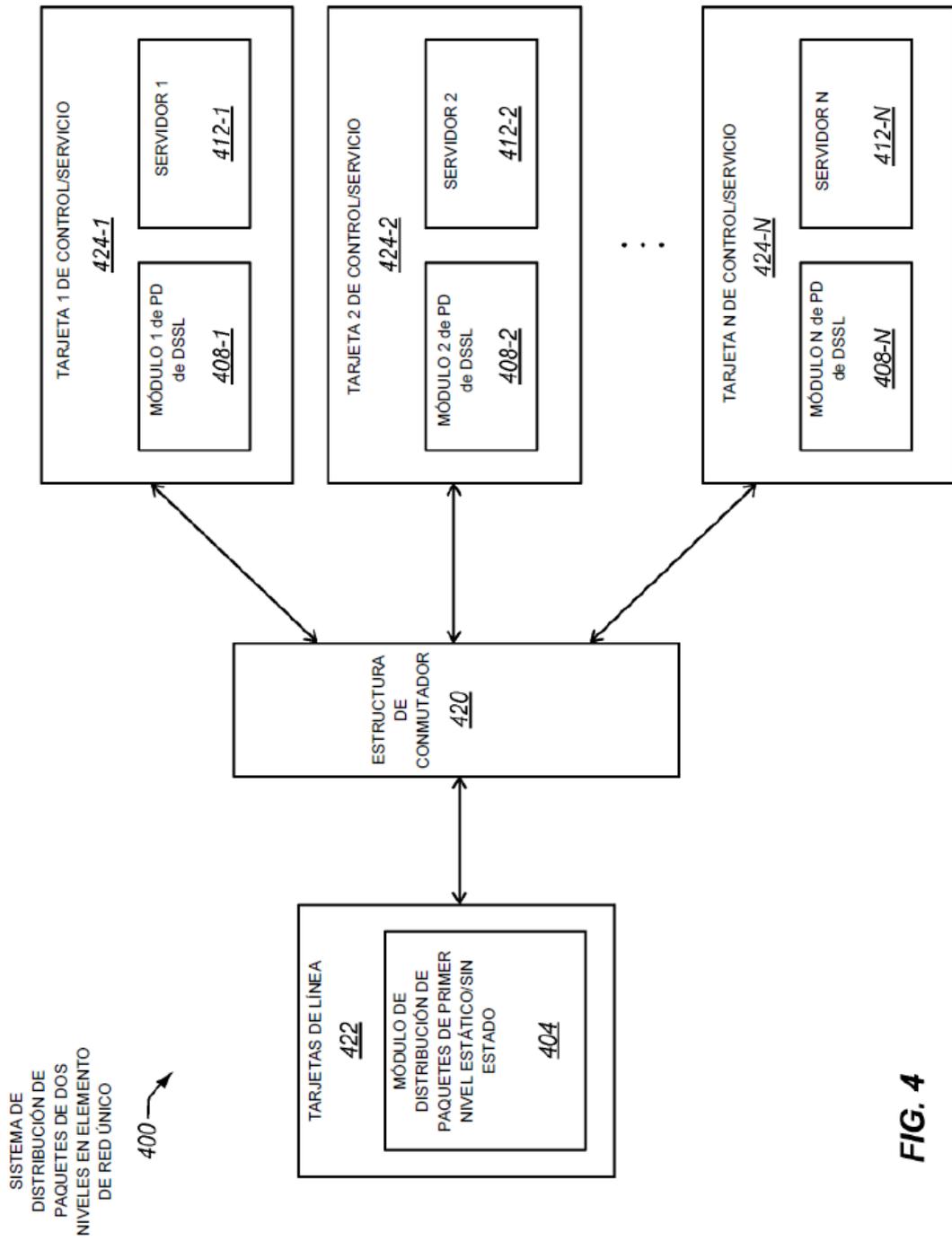
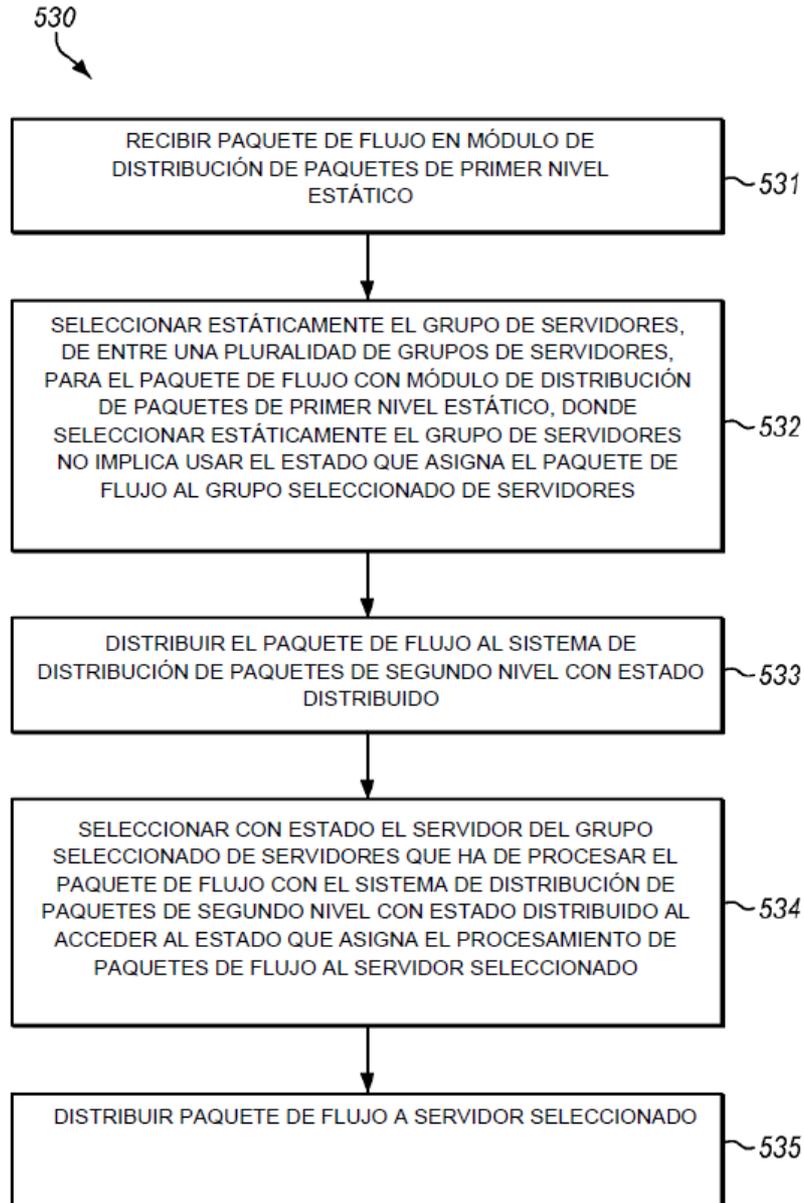


FIG. 4

FIG. 5

MÉTODO PARA DISTRIBUIR PAQUETES
ENTRE SERVIDORES MIENTRAS SE
PROPORCIONA ADHERENCIA DE FLUJOS
DE LOS PAQUETES A SERVIDORES
ASIGNADOS PARA PROCESAR LOS
PAQUETES DE LOS FLUJOS



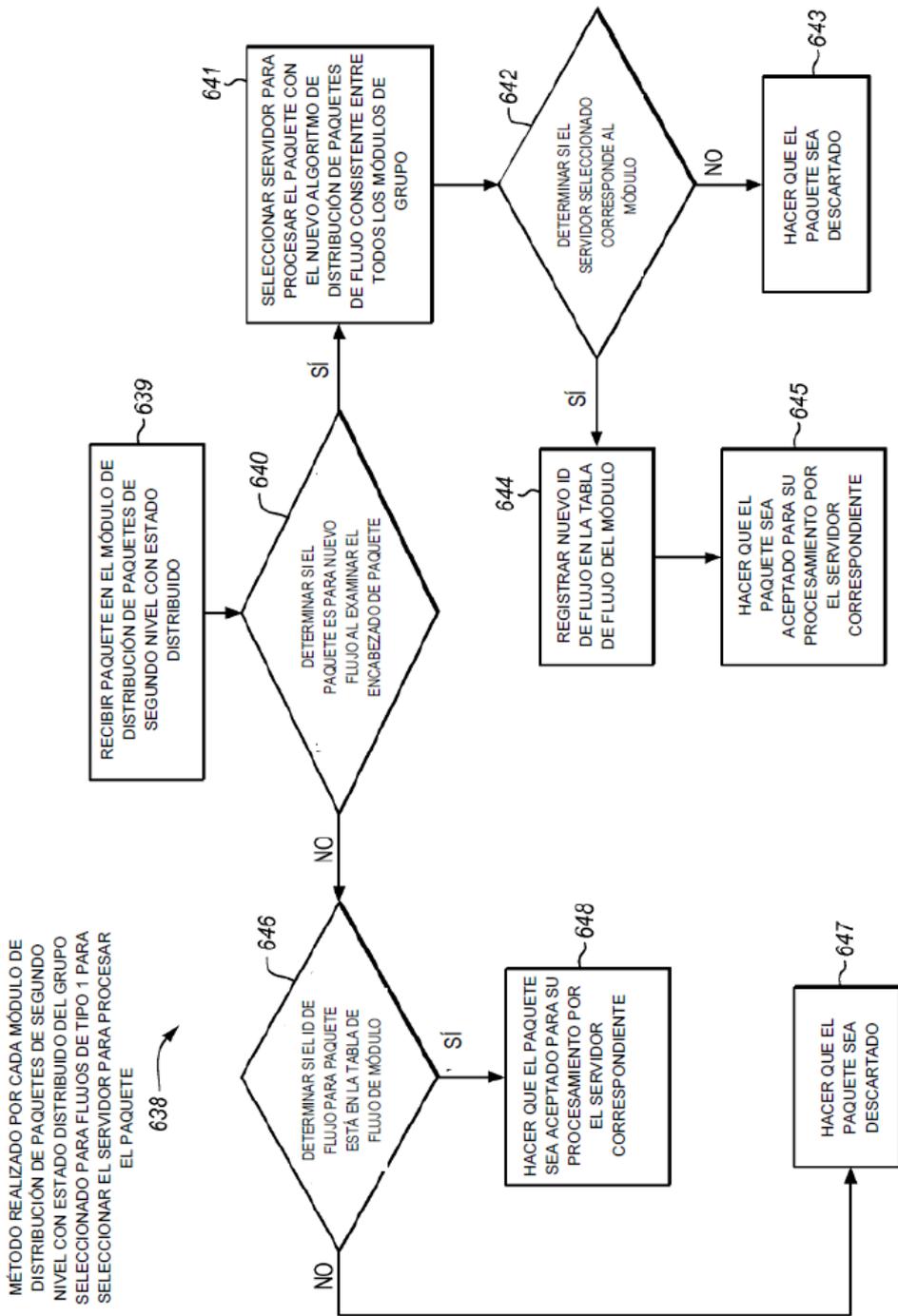


FIG. 6

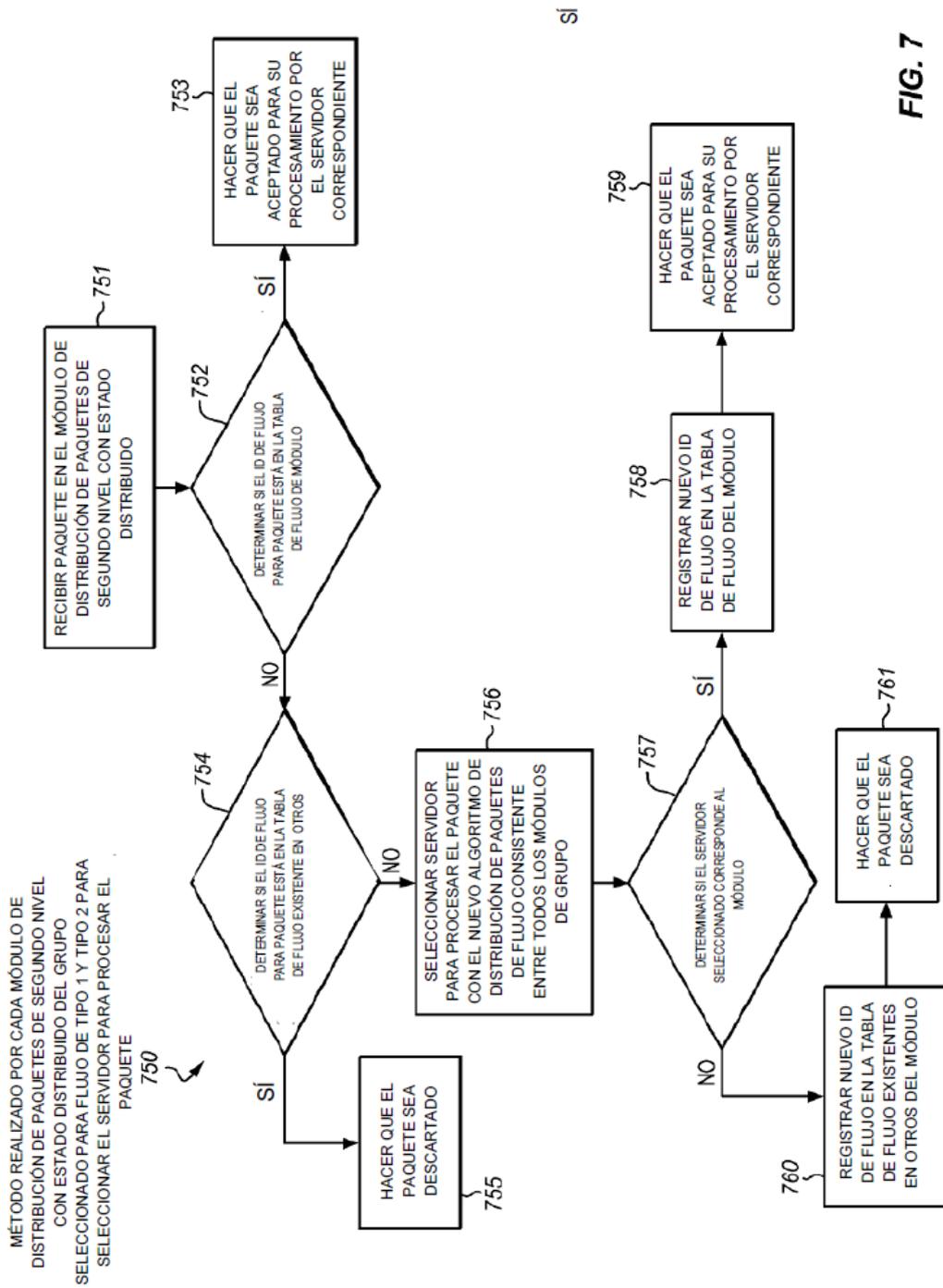


FIG. 8

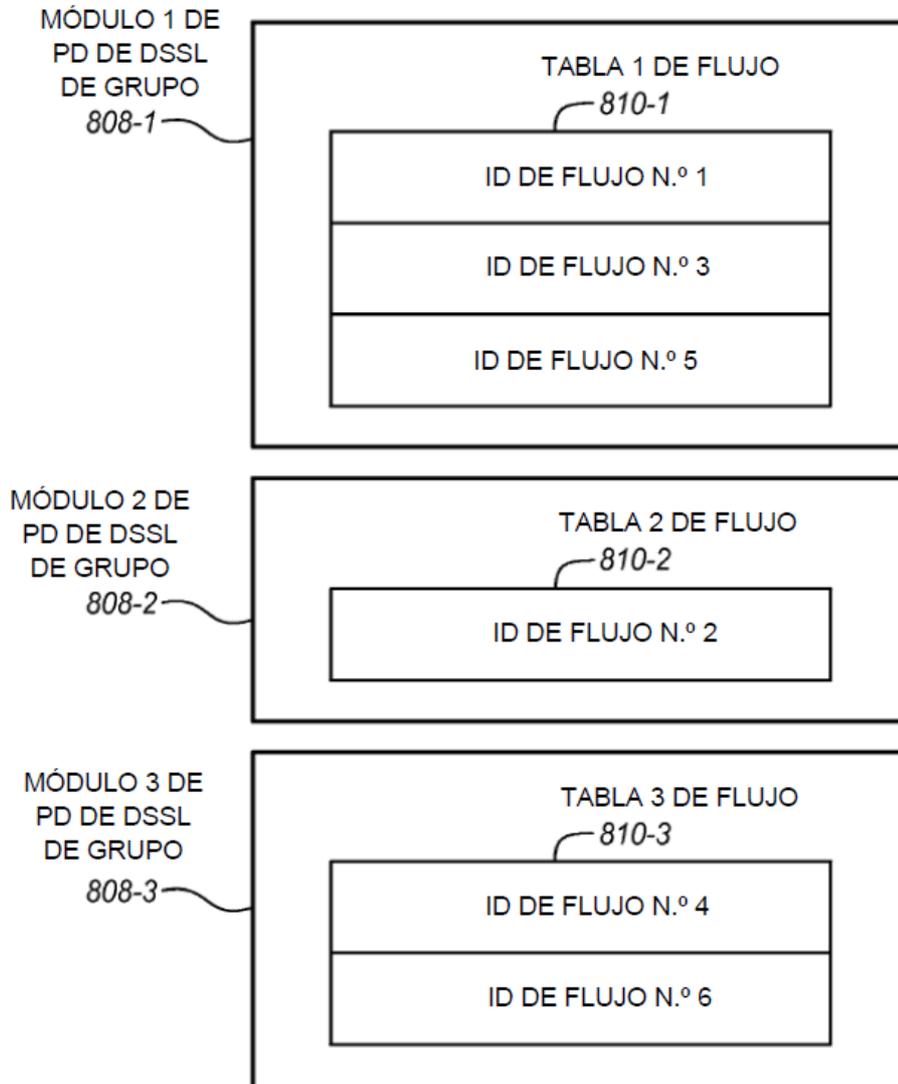


FIG. 9

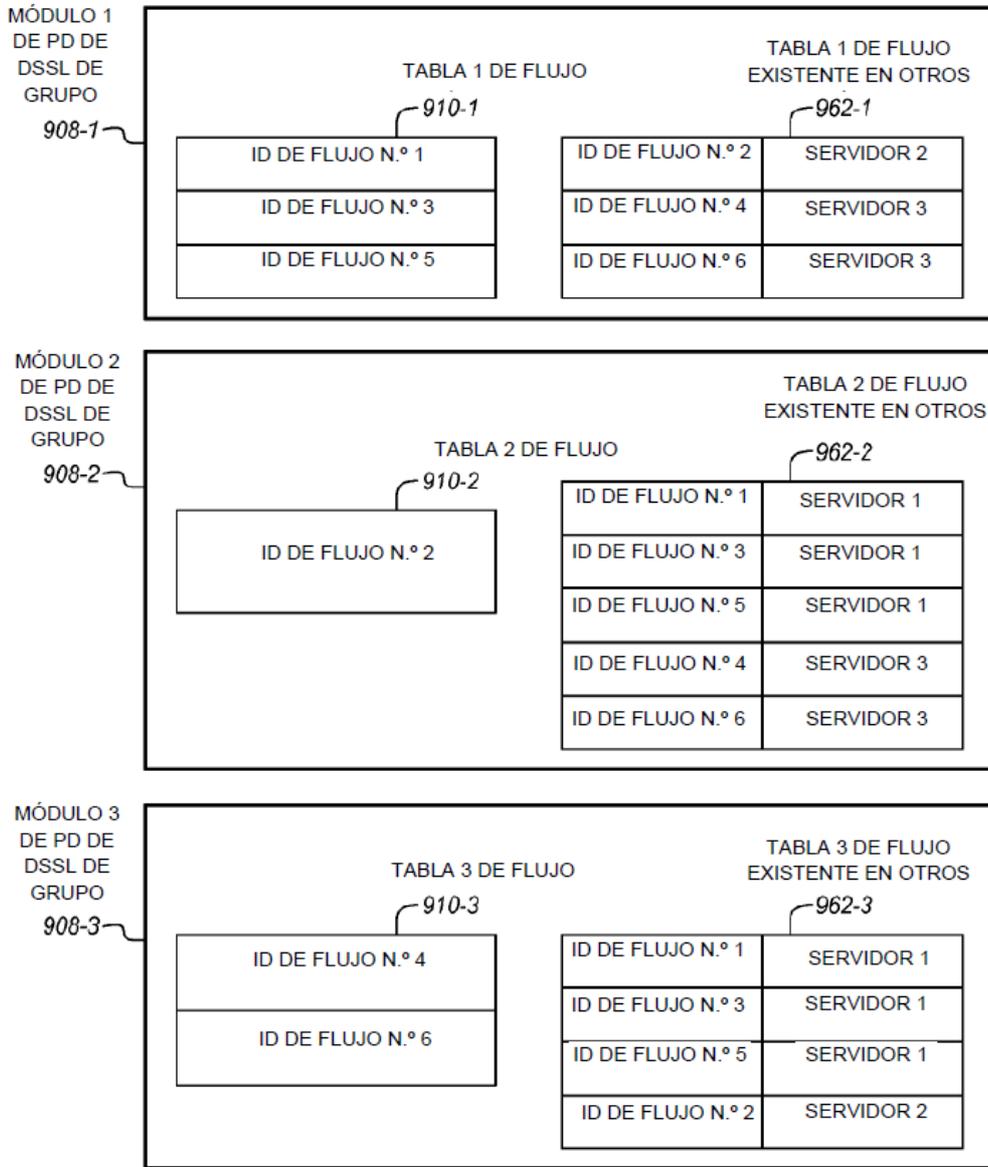


FIG. 10

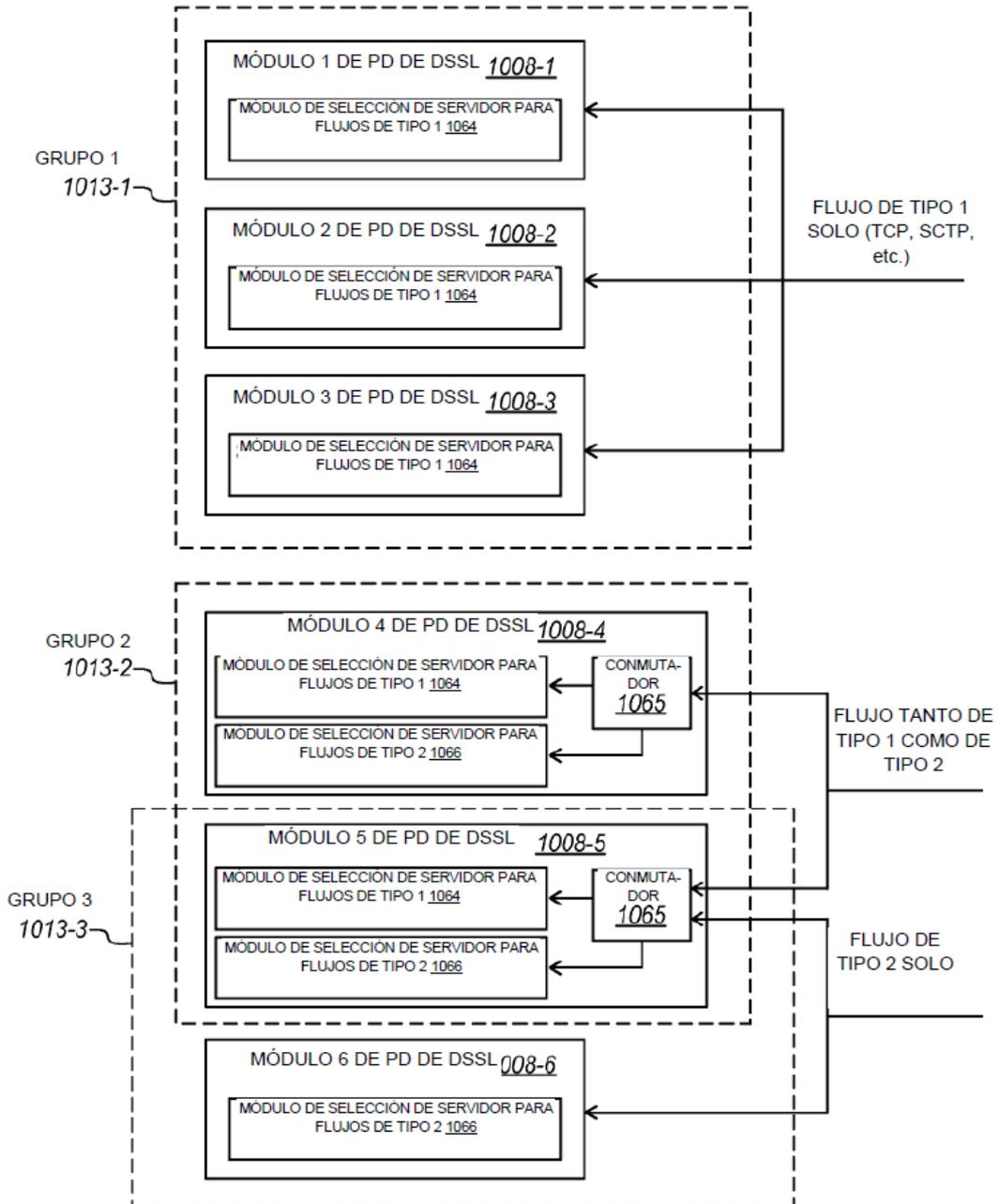
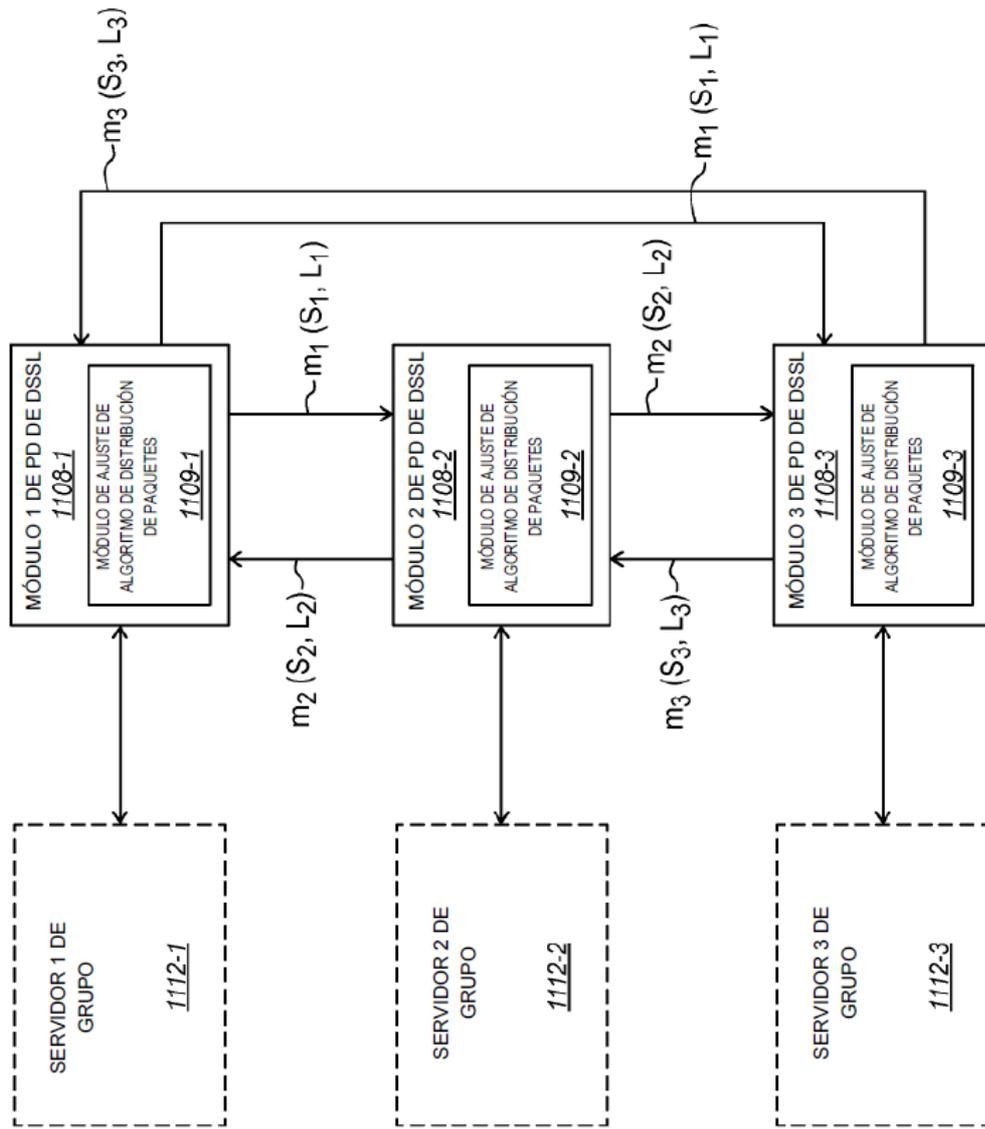


FIG. 11
1113



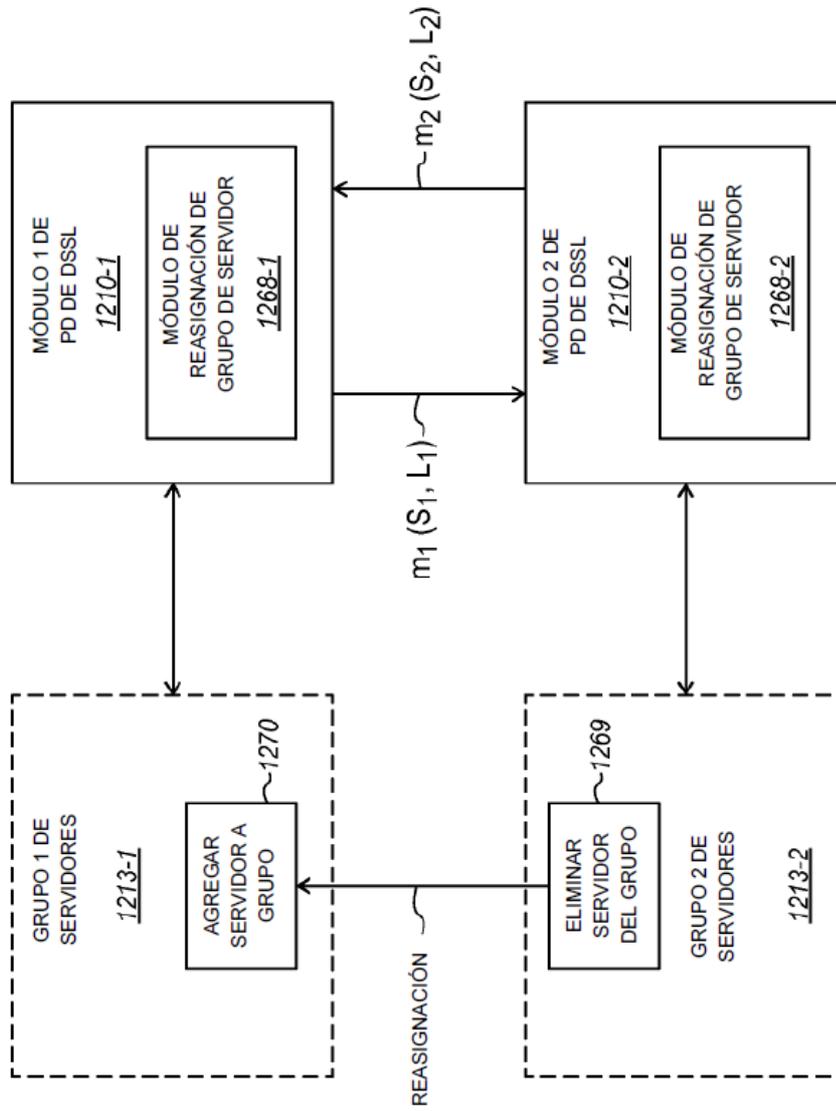


FIG. 12

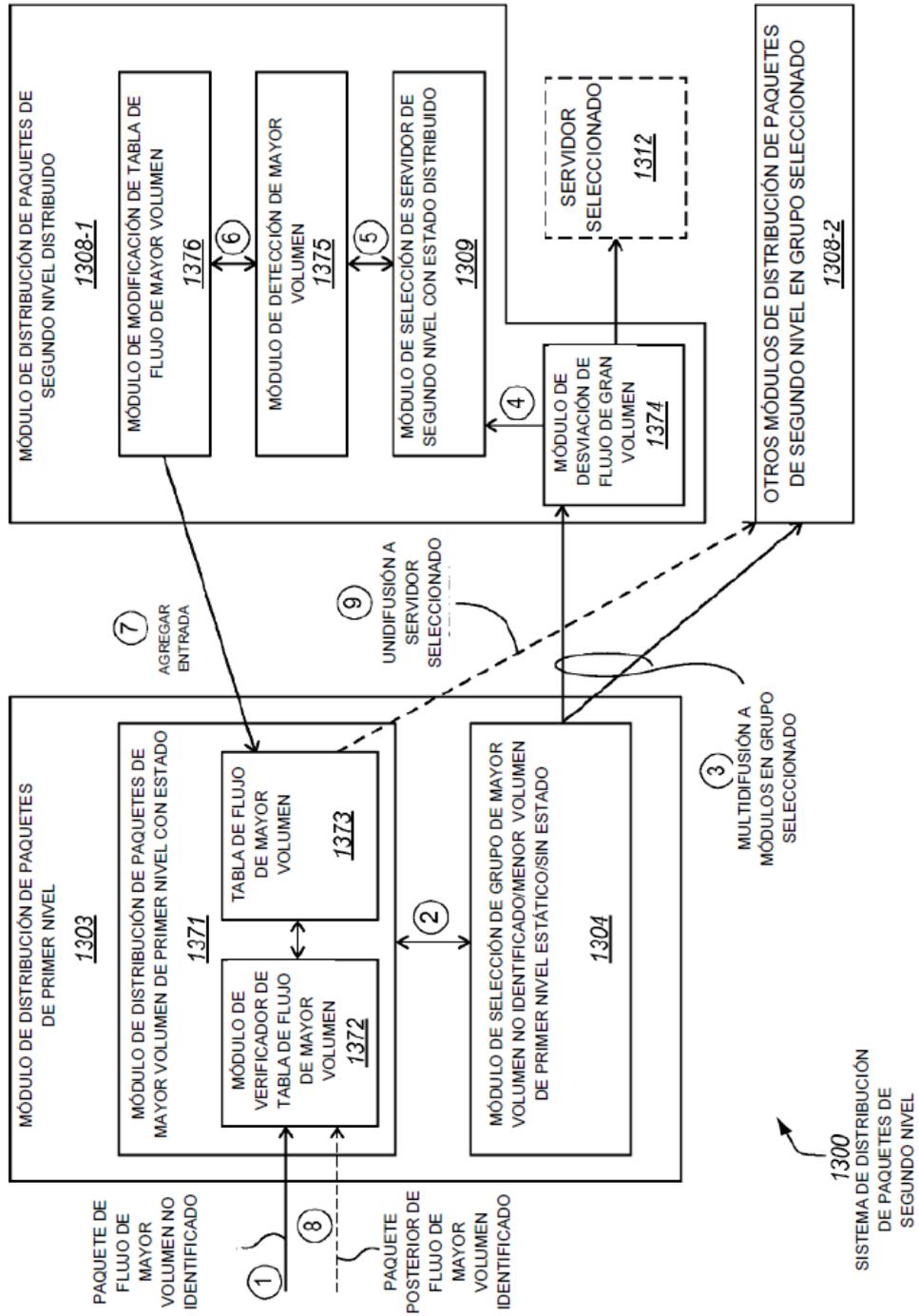


FIG. 13