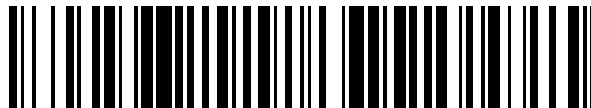


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 424 630**

51 Int. Cl.:

H04L 12/771 (2013.01)

H04L 12/937 (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **29.10.2008** **E 08845163 (8)**

97 Fecha y número de publicación de la concesión europea: **15.05.2013** **EP 2208322**

54 Título: **Un enrutador de rejilla**

30 Prioridad:

30.10.2007 US 927861

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
07.10.2013

73 Titular/es:

**CONTEXTREAM LTD. (100.0%)
AZORIM INDUSTRIAL PARK 94 EM
HAMOSHAVOT ST.
49527 PETACH TIKVA, IL**

72 Inventor/es:

**BARKAI, SHARON;
NOY, ARIEL y
SIDI, RON**

74 Agente/Representante:

VALLEJO LÓPEZ, Juan Pedro

ES 2 424 630 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

Un enrutador de rejilla

Campo de la invención

La presente invención se refiere a enrutadores de red en general y a enrutadores de rejilla en particular.

5 **Antecedentes de la invención**

Las redes de comunicaciones son ubicuas. La Internet está en todas partes y los portadores están intentando proporcionar cada vez más servicios sobre la misma para sus clientes.

10 La solicitud de patente de los Estados Unidos 11/615.977, asignada a los cesionarios comunes de la presente invención, describe una red de límite metropolitano que intenta concentrarse en los servicios que proporciona en lugar de en la topología de la red y/o la configuración de línea. Una red ejemplar de límite metropolitano 40 se muestra en la Fig. 1, a la cual se hace ahora referencia.

15 Para cada servicio, la red de límite metropolitano asigna un buzón 22 cuyo tamaño y calidad del servicio (definido por el ancho de banda, tasas de descarga, control de seguridad, etc.) puede ser una función de la cantidad de datos a transferir a través de los mismos en cualquier momento determinado. Los buzones 22 actúan como memorias intermedias, donde se pueden escribir los datos en un extremo y leerse en el otro extremo.

20 La definición de los buzones se puede promover directamente desde las tablas de servicios que los portadores mantienen de todos modos. Tales tablas existen para la facturación, para el servicio de clientes, para la asignación de ancho de banda, etc. La red de límite metropolitano de la presente invención puede acceder a tales tablas para definir los requisitos de tamaño y de la calidad de servicio (QoS) para sus buzones 22 y pueden almacenar tal información en su propia tabla de servicios 30. La red de límite metropolitano 40 también puede comprender una multiplicidad de enrutadores de rejilla 42, de los cuales se muestran tres de ejemplo (con las letras de referencia A, B, y C) en la Fig. 1. Los enrutadores 42 se pueden conectar juntos con un anillo 41.

25 Cada uno de los enrutadores 42 puede comprender una unidad de interfaz externa (EIU) 44 y una unidad de buzones (MU) 46. Cada una de las unidades de interfaces externas 44 puede proporcionar conexiones a sus clientes asociados y al anillo 41 y cada una de las unidades de buzones 46 puede mantener y gestionar los buzones 22. Sin embargo debido a que la red de límite metropolitano 40 puede ser una red distribuida, los buzones 22 de los clientes asociados de una unidad de interfaz externa 44 puede que no necesariamente se almacenen en la unidad de buzones 46 del mismo enrutador 42. La localización de cada uno de los buzones puede ser una función del tipo de servicio que se proporciona y de si necesita localizarse o no cerca del cliente o cerca del origen para un funcionamiento eficiente. Las unidades de buzones 46 pueden comprender una multiplicidad de memorias intermedias, para los buzones, y unidades de gestión para añadir o borrar buzones cuando se requiera por la tabla de servicios 30.

35 Para cada uno de los paquetes de datos entrantes, cada unidad de interfaz externa 44 puede determinar qué unidad de buzones 46 puede gestionar el buzón 22 para el servicio que se transmite por el paquete de datos. Cada una de las unidades de interfaz externa 44 puede realizar una función de generación de huella digital sobre al menos parte de la información en una cabecera del paquete de datos. El valor de la huella digital resultante puede ser el número del buzón dentro de la red de límite metropolitano 40 para el servicio transportado por el paquete de datos. Con el número del buzón, la unidad de interfaz externa 44 puede escribir directamente el paquete de datos (normalmente usando el DMA remoto (rDMA) en el buzón indicado 22.

40 Por ejemplo, la Familia Smith, etiquetada como 50, puede solicitar un video a demanda desde el VOD etiquetado como 52. La Familia Smith, etiquetada como 50, puede registrar la petición de servicio y la red de límite metropolitano 40 puede crear un buzón 22B, para ese servicio y puede asignar también una unidad de interfaz externa 44, tal como EIU 44A, para el servicio. El VOD 52 puede enviar su flujo de datos de VOD al enrutador 42A, el enrutador más cercano al mismo. La unidad de interfaz externa 44A puede generar la huella digital de la dirección del destino, la de la Familia Smith, y puede determinar que el buzón para VOD 52 de la familia Smith, el buzón 22B₁, puede estar localizado en la unidad de buzones 46B. Por consiguiente, la unidad de interfaz externa 44A puede escribir el flujo de datos en el buzón 22B₁ en la unidad de buzones 46B. Cuando el decodificador de televisión de la Familia Smith se puede conectar al enrutador 42B, el enrutador más próximo al mismo, la unidad de buzones 46B puede enviar los paquetes almacenados en el buzón 22B₁ al decodificador de televisión de la Familia Smith, a través de la unidad de interfaz externa 44B.

Otro servicio es el de dos personas, por ejemplo Dana Smith y su amiga Ann, hablando sobre VoIP. En este ejemplo, el buzón de Dana Smith 22B₂ puede estar localizado en la unidad del buzón 46B, cerca de su casa, mientras que el buzón de Ann 22C puede estar localizado en una unidad de buzones diferente, por ejemplo 46C.

55 Cuando Ann habla sobre su teléfono de VoIP, su ordenador puede enviar su flujo de datos de voz al enrutador 42C, el enrutador con el que está asociada la unidad de interfaz externa 44C puede generar la huella digital de la

dirección de destino, la de Dana Smith, y puede determinar que el buzón de Dana Smith para VoIP, el buzón 22B₂, puede estar localizado en la unidad de buzones 46B. Por consiguiente, la unidad de interfaz externa 44C puede escribir el flujo de datos en el buzón 22B₂ en la unidad de buzones 46B y la unidad de buzones 46B puede redirigir los paquetes al teléfono de VoIP de Dana a través de la unidad de interfaz externa 44B.

5 El documento de Bianco A y otros, Multi-stage Switching Architectures for Software Routers - IEEE Network IEEE Service Center, Nueva York, NY, US, LNKD, DOI: 10.1109 / MNET. 2007.386465 Volumen 21, N° 4, del 1 de julio de 2007 (01/01/2007), páginas 15 - 21, XP01111880561 ISSN: 0890 - 8044, enseña una arquitectura basada en PC multi-etapa que aprovecha los enrutadores basados en PC como elementos de conmutación, para construir un enrutador de alta velocidad, a gran escala, escalable y de software fiable.

10 **Sumario de la presente invención**

En adelante se proporciona, de acuerdo con una realización preferida de la presente invención, un enrutador de rejilla incluyendo una pluralidad de unidades de interfaces externas y una segunda pluralidad de unidades de buzones. Las unidades de interfaces externas reciben paquetes de los servicios desde un punto y proporcionan los paquetes de los servicios a otro punto.

15 Las unidades de buzones almacenan y redirigen los paquetes asociados con los buzones, uno por servicio, y cada unidad de buzones está conectada a las unidades de interfaces externas en una red de tipo CLOS. La unidad de interfaz externa tiene una porción de entrada y una porción de salida y las porciones de entrada y salida y las unidades de buzones se escriben entre sí, en un modo no síncrono.

20 Además, de acuerdo con una realización preferida de la presente invención, las porciones de entrada escriben en las unidades de buzones como la información en los paquetes, las unidades de buzones escriben los paquetes en las porciones de salida asociadas con los buzones en los que se almacenan los paquetes y las porciones de salida escriben los paquetes en la red.

25 Además, de acuerdo con la realización preferida de la presente invención, cada una de las porciones y cada una de las unidades de buzones incluyen una sección de entrada para almacenar los paquetes y las cabeceras relacionadas, un procesador para procesar sólo las cabeceras al menos para tomar las decisiones de encaminamiento y una sección de salida, que almacena en al menos un puntero para la sección de entrada, para organizar los paquetes para la transferencia a la siguiente unidad.

30 Aún más, de acuerdo con una realización preferida de la presente invención, cada una de las secciones de entrada de la porción de entrada está conectada a una pluralidad de tarjetas de interfaz de red, que escriben cada una directamente dentro de la sección de entrada.

Además, de acuerdo con una realización preferida de la presente invención, cada tarjeta de interfaz de red incluye una librería de espacios de usuario y una unidad de DMA de dispersión - recogida para escribir los paquetes desde la librería a la sección de entrada.

35 Adicionalmente, de acuerdo con una realización preferida de la presente invención, cada sección de salida incluye al menos una cola que almacena los punteros a las localizaciones en la sección de entrada.

Además, de acuerdo con una realización preferida de la presente invención, cada una de las secciones de salida incluye una cola de escritura por unidad en la que se escribe.

40 Además, de acuerdo con una realización preferida de la presente invención, la unidad de buzones incluye adicionalmente una sección de buzones que almacena las colas de los buzones, cada una asociada con un buzón. Cada una de las colas de buzones almacena punteros a localizaciones en su sección de entrada cuando se almacenan sus datos.

Aún más, de acuerdo con una realización preferida de la presente invención, el procesador de la unidad de buzones incluye un programador para la programación y el control de la salida de los paquetes.

45 Además de acuerdo con una realización preferida de la presente invención, el procesador de la unidad de buzones incluye un generador de huellas digitales, un calificador, un selector de buzones y un despachador. El generador de huellas digitales genera la huella digital de una cabecera de un paquete para determinar a qué buzón pertenece el paquete. El clasificador controla la entrada del buzón de acuerdo con los requisitos previamente definidos para el buzón. El selector de buzón, almacena los punteros a la cabecera y el paquete en una cola de buzones asociada con el buzón determinado bajo instrucciones del clasificador. El despachador conforma un flujo de salida desde el buzón para almacenar los punteros en una cola de escritura definida anteriormente para el buzón.

50 Adicionalmente, de acuerdo con una realización preferida de la presente invención, el procesador de la porción de entrada incluye un creador de la cabecera, un generador de huellas digitales y un selector de la unidad de buzones. El creador de la cabecera crea una nueva cabecera con un puntero a la localización de almacenamiento del paquete. El generador de huellas digitales lee la nueva cabecera y determina en qué unidad de buzones almacenar

el paquete. El selector de la unidad de buzones escribe punteros para la nueva cabecera y para la localización de almacenamiento en una cola de escritura para la unidad de buzones determinada.

Además, de acuerdo con una realización preferida de la presente invención, el procesador de la porción de salida incluye un solicitador de escritura para actualizar una cabecera de un paquete con una dirección de destino del paquete y para escribir un puntero para el paquete dentro de una cola de transmisión.

También se proporciona, de acuerdo con una realización de la presente invención, un enrutador de rejilla que incluye una pluralidad de unidades de interfaces externas y una segunda pluralidad de unidades de buzones. Las unidades de interfaces externas reciben paquetes de los servicios desde un punto y proporcionan los paquetes de los servicios a otro punto. Las unidades de buzones almacenan y retransmiten los paquetes asociados con los buzones, uno por servicio, y cada una de las unidades de buzones se conecta a cada una de las unidades de interfaces externas. Las unidades se implementan sobre PC (ordenadores personales) que tienen controladores de memoria y mueven los paquetes a través de la misma a la tasa de E/S (entrada/salida) de los controladores de memoria o próxima a la misma.

Además, de acuerdo con una realización preferida de la presente invención, cada uno de los ordenadores personales incluye una sección de entrada para almacenar los paquetes y las cabeceras relacionadas, un procesador para procesar solo las cabeceras y al menos para tomar decisiones de encaminamiento y una sección de salida, que almacena al menos un puntero a la sección de entrada para organizar los paquetes para la transferencia al siguiente ordenador personal.

Además, de acuerdo con una realización preferida de la presente invención, la sección de entrada de una porción de entrada de cada una de las unidades de interfaces externas se conecta a una pluralidad de tarjetas de interfaz de red, que escriben cada una directamente dentro de la sección de entrada. El procesador de la porción de entrada determina el encaminamiento de los paquetes a una de las unidades de buzones.

Aún más, de acuerdo con una realización preferida de la presente invención, el enrutador de rejilla también incluye una unidad de colocación de datos directa para transferir paquetes desde una unidad a otra. La unidad de colocación de datos directa puede ser una unidad de rDMA, un dispositivo de descarga de TCP, una tarjeta de interfaz de red habilitada para RDMA o una tarjeta interfaz de red habilitada para DDP.

Además, de acuerdo con una realización preferida de la presente invención, la unidad de buzones incluye adicionalmente una sección de buzones que almacena las colas de los buzones, asociada cada una con un buzón. Cada una de las colas de buzón apunta a las localizaciones en su sección de entrada donde se almacenan sus datos.

Además, de acuerdo con una realización preferida de la presente invención, cada una de las unidades de interfaces externas tiene una porción de entrada para escribir en cada una de las unidades de buzones y una porción de salida para recibir los paquetes de la unidad de buzones. Las porciones de entrada y salida y las unidades de buzones se escriben entre sí en un modo asíncrono.

También se proporciona, de acuerdo con una realización preferida de la presente invención, un procedimiento que incluye la transferencia de paquetes y cabeceras, almacenadas y apuntados en áreas de memoria separadas de un ordenador personal en una cola de escritura del ordenador personal para separar áreas de memoria de otro ordenador personal a través de una unidad de colocación de datos directa a una tasa de entrada/salida de los controladores de memoria de los ordenadores o próxima a la misma.

Adicionalmente, de acuerdo con una realización preferida de la presente invención, el procedimiento también incluye el procesamiento de cada una de las cabeceras para determinar dentro de qué cola de escritura se colocará un puntero a la cabecera y su paquete asociado.

Finalmente, se proporciona, de acuerdo con una realización preferida adicional de la presente invención, un procedimiento que incluye el almacenamiento de paquetes y cabeceras de una pluralidad de buzones de una unidad de buzones en un área de almacenamiento general, procesando cada una de las cabeceras para determinar a qué buzón pertenece cada uno de los paquetes y su cabecera asociada, colocando punteros al paquete y su cabecera asociada dentro de una cola de buzón para el buzón y el procesamiento de la cola del buzón para controlar la entrada y la salida del buzón.

Breve descripción de los dibujos

El tema objeto con respecto a la invención se apunta particularmente y se reivindica distintivamente en la porción de conclusión de la memoria descriptiva. La invención, sin embargo, tanto la organización como el procedimiento de funcionamiento, junto con los objetos, características y ventajas de la misma, se pueden entender del mejor modo por referencia a la siguiente descripción detallada cuando se lee con los dibujos adjuntos en los que:

la Fig. 1 es una ilustración esquemática de una red de límite metropolitano;

la Fig. 2 es una ilustración esquemática de un enrutador de rejilla novedoso, construido y operativo de acuerdo con una realización preferida de la presente invención;

5 las Fig. 3A, 3B y 3C son ilustraciones esquemáticas de una porción de entrada de la unidad de interfaz externa, una unidad de buzones y una porción de salida de la unidad de interfaz externa, respectivamente, que forman parte del enrutador de rejilla de la Fig. 2;

las Fig. 4A, 4B, 4C y 4E son ilustraciones esquemáticas del movimiento de datos a través de los elementos de las Fig. 3A, 3B y 3C; y

la Fig. 5 es una ilustración esquemática del funcionamiento de los elementos de las Fig. 3A, 3B y 3C en el registro de un nuevo buzón.

10 Se apreciará que por simplicidad y claridad de la ilustración, los elementos mostrados en las figuras no necesariamente se han dibujado a escala. Por ejemplo, las dimensiones de algunos de los elementos pueden estar exageradas con relación a otros elementos por claridad. Además, donde se considere apropiado, las referencias numéricas se pueden repetir entre las figuras para indicar los elementos correspondientes o análogos.

Descripción detallada de la presente invención

15 En la siguiente descripción detallada, se muestran numerosos detalles específicos para proporcionar un entendimiento completo de la invención. Sin embargo se entenderá por los expertos en la materia que la presente invención se puede poner en práctica sin estos detalles específicos. En otros casos, métodos, procedimientos y componentes bien conocidos no se han descrito con detalle de modo que no obscurezcan la presente invención.

20 A menos que se establezca específicamente lo contrario, como es evidente de las siguientes discusiones, se apreciará que a través de toda la memoria descriptiva, las exposiciones que usan términos tales como "procesamiento", "computación", "cálculo", "determinación" o similares, se refieren a la acción y/o procedimientos de un ordenador, sistema de computación o dispositivo de computación electrónica similar que manipula y/o transforma los datos representados como cantidades físicas, tal como electrónicas dentro de los registros del sistema de computación y/o las memorias en otros datos representados de forma similar como cantidades físicas dentro de las
25 memorias del sistema de computación, registros u otro almacenamiento de tal información, dispositivos de transmisión o de representación.

Los solicitantes se han dado cuenta de que, para la mayoría de las implementaciones, puede ser necesario acceder un número muy grande de buzones. Puede que una unidad de buzones (MU) no sea suficiente para mantenerlos a todos.

30 Ahora se hace referencia a la Fig. 2, que ilustra un novedoso enrutador de rejilla 50, construido y operativo de acuerdo con una realización preferida de la presente invención, que comprende una multiplicidad de unidades de buzones MU 54. De acuerdo con una realización preferida de la presente invención, las unidades buzones MU 54 se pueden conectar en un lado a las unidades de interfaces externas de la izquierda (EIU) 52L y en el otro lado a las unidades de interfaces externas de la derecha EIU 52R.

35 De acuerdo con una realización preferida de la presente invención, las unidades 52 y 54 se pueden conectar en una red de tipo Clos, a través de la cual cada una de las unidades de buzones 54 se puede conectar con cada una de las EIU 52. De este modo, cada unidad de buzones 54 se puede acceder por cada una de las unidades de interfaces externas 52 al mismo tiempo.

40 Esto puede ayudar al enrutador 50 a operar a plena tasa y una carga equilibrada (es decir, cualquier paquete recibido se puede colocar dentro de un buzón y cualquier paquete en un buzón se puede leer). Hay una pequeña o ninguna congestión, al menos en parte debido al hecho de que el procedimiento de generación de huella digital difunde el tráfico relativamente de forma uniforme entre las unidades de buzones 54 y al hecho de que las unidades de buzones 54 son accesibles desde todas las EIU 52.

45 Si una red de límite metropolitano solo requiere un enrutador 50, entonces las unidades de interfaces externas EIU 52 se pueden conectar directamente a las unidades de buzones 54. Sin embargo, si la red de límite metropolitano requiere más de un enrutador 50, entonces, de acuerdo con una realización preferida de la presente invención, el enrutador 50 también comprende conmutadores 56 que agregan salidas desde las EIU 52 y las MU 54 sobre los enlaces de alta velocidad 58, tal como un 10G Ethernet. Hay dos conjuntos de enlaces 58L y 58R, en los lados, izquierdo y derecho de las unidades de buzones 54.

50 Por ejemplo, en la Fig. 2 hay 12 EIU a la izquierda 52L alojados en un chasis de EIU 53L, 12 unidades de buzones 54 alojados en un chasis de buzones 55, y 12 EIU a la derecha 52R alojados en un chasis de EIU a la derecha 53R. Cada uno de los chasis de EIU 53 aloja dos conmutadores (para el chasis de EIU de la izquierda 53L, los conmutadores 56A y 56B están a la salida, mientras que para el chasis de EIU de la derecha 53R, los conmutadores 56G y 56H están a la entrada). El chasis de buzones 55 aloja cuatro conmutadores, dos a la izquierda (56C y 56D) y dos a la derecha (56E y 56F).

El conmutador 56A agrega la salida de las EIU 52L sobre el enlace 58La que, en este ejemplo, es un enlace de 10 Gbps. El enlace 58La conecta al conmutador 56C que, a su vez conecta con todas las 12 unidades de buzones 54. El conmutador 56E agrega la salida de las unidades de buzones 54 sobre el enlace 58R que conecta con el conmutador 56G. El conmutador 56G conecta con todas las EIU 52R. Los conmutadores 56B, 56D, 56F y 56H conectan con otros chasis de buzones 55 que forman parte de los otros enrutadores 50.

Cada uno de los enlaces 58 conecta entre las 12 EIU 52 y las 12 unidades de buzones 54 y de este modo, transporta $12 \times 12 = 144$ conexiones. De este modo, cada una de las conexiones tiene una tasa de bits de 10 Gbps / 144 o aproximadamente 70 Mbps. En cualquier momento determinado, pueden llegar 144 paquetes en cualquiera de los conmutadores 56 (escritos en los mismos por las unidades rDMA). Durante el tiempo en el que llegan, se pueden conmutar 12 paquetes a su destino. De este modo, $144 - 12 = 132$ paquetes pueden colisionar en cualquier momento determinado. Este es un número de paquetes relativamente bajo considerando la velocidad de las conexiones.

Con un número tan bajo de colisiones, el enrutador de rejilla 50 puede tener pocas pérdidas de paquetes y una perturbación de oscilación relativamente pequeña, y de este modo, una calidad de servicio relativamente alta. La tasa de colisión mínima también posibilita que los conmutadores 56 sean mucho más simples, ya que el tráfico es relativamente equilibrado, con pocas ráfagas, si las hay.

Los solicitantes se han dado cuenta además de que cada buzón opera de forma unidireccional. Esto es, un proveedor de servicio envía datos a su unidad de interfaz externa, que a su vez, escribe los datos dentro del buzón asociado con el servicio. La unidad de interfaz externa del abonado toma los datos del buzón y los envía al abonado. La presente invención puede ser un enrutador de rejilla no síncrono que escribe separadamente a y desde una pluralidad de buzones, pero a una velocidad relativamente alta y con una calidad de servicio relativamente alta.

Ahora se hace referencia a las Fig. 3A, 3B y 3C, que ilustran respectivamente los elementos de un enrutador de rejilla 50, y a las Fig. 4A, 4B, 4C y 4D que ilustran su operación. De acuerdo con una realización preferida de la presente invención, cada unidad de interfaz externa EIU 52 tiene dos secciones, una sección de entrada 57 y una sección de salida 59, que escribe separadamente a y desde las unidades de buzones 54, respectivamente.

Cada una de las unidades 54, 57 y 59 se pueden implementar sobre un hardware de PC (ordenador personal) normalizado, en particular uno que tiene una arquitectura X86. Tal arquitectura tiene una unidad de procesamiento central (CPU), una RAM (memoria de acceso aleatorio), un controlador de memoria para leer y escribir los datos desde la RAM a la CPU y una cantidad más pequeña de memoria de operación ("caché") fácilmente disponible para la CPU.

Cada una de las unidades 54, 57 y 59 pueden tener una sección de datos de entrada 60, una unidad de procesamiento central (CPU) 62 y una sección de datos de salida 64. Cada una de las secciones de datos de entrada 60 puede recibir los paquetes escritos en la misma, cada una de las CPU 62 puede realizar procedimientos relativamente simples para determinar cómo y dónde redirigir los paquetes y cada una de las secciones de datos de salida 64 puede escribir los paquetes a la siguiente unidad. Esta estructura puede posibilitar a un enrutador 50 operar a la tasa de entrada/salida máxima o cerca de esta, como se define por la tasa de funcionamiento de su controlador de memoria, ya que, como se trata más adelante en este documento, las CPU 62 no están involucradas en el movimiento de datos desde una localización de memoria a otra.

La sección de entrada 57 puede comprender un área de almacenamiento de datos 70 y un área de almacenamiento de cabeceras de paquetes 72 en su sección de datos de entrada 60W. Ejecutándose sobre la CPU 62W puede estar un creador de cabecera 74, un generador de huellas digitales 76 y un selector de la unidad de buzones 78 y una sección de datos de salida 64W puede comprender una multiplicidad de colas de escritura 80, una por unidad de buzón 54, y una unidad de rDMA 81.

Como se muestra en la Fig. 4A, los datos se pueden escribir (flecha 82) dentro del área de almacenamiento de datos 70, normalmente dentro de una sección de memoria disponible, tal como la sección k. Esta operación de escritura puede ocurrir cada vez que una tarjeta de interfaz de red (NIC) (no mostrada) puede tener datos para escribir. La (NIC) puede escribir en la sección de memoria k usando una operación de "dispersión recogida". La NIC se puede modificar a través de su "librería de espacio - usuario", para escribir la lista de dispersión recogida directamente dentro de la siguiente localización disponible en el área de almacenamiento de datos 70, en una operación de tipo DMA. Como los datos entrantes pueden ser un paquete y como los paquetes pueden ser de longitudes variables, cada sección de memoria se puede identificar por la localización de su primer bit.

Los paquetes normalmente comprenden cabeceras y colas, con datos entre las mismas, donde las cabeceras normalmente incluyen como mínimo la dirección de los dispositivos de origen y de destino y el tipo de servicio transportado en el paquete. El creador de la cabecera 74 puede leer (flecha 84) la cabecera en su totalidad y puede crear una nueva cabecera, que añade un puntero a la sección de memoria k dentro de la cabecera original. El creador de la cabecera 74 puede almacenar (flecha 85) la nueva cabecera en el área de almacenamiento de cabeceras de paquetes 72, en un contenedor de ejemplo v. Cada contenedor en el área de almacenamiento 72 puede ser relativamente pequeño en comparación con las secciones de memoria del área de almacenamiento de

datos 70, y los contenedores pueden ser generalmente del mismo tamaño.

El generador de huellas digitales 76 puede revisar las cabeceras almacenadas en el área de almacenamiento 72, tomando una cada vez. Puede leer cada una de las cabeceras (flecha 86), tal como la cabecera almacenada en el contenedor v, y puede generar un valor de huella digital H_1 a partir de los datos almacenados en la misma. El valor de la huella digital H_1 puede indicar dentro de qué unidad de buzón 54 transferir el paquete. Como se trata en la Solicitud de Patente de los Estados Unidos 11/615.977, la entrada de la huella digital puede variar dependiendo del tipo de servicio transportado por el paquete. Normalmente, el paquete puede incluir una indicación del tipo de servicio transportado en el mismo.

El selector de la unidad de buzones 78 puede actualizar la siguiente dirección de destino de la cabecera a la de la unidad de buzones seleccionada 54, puede añadir el valor de la huella digital H_1 , y puede escribir (flecha 87) punteros al paquete y a la cabecera actualizados dentro de la cola de escritura 80 para la unidad de buzones MU indicada por el valor de la huella digital H_1 . Por ejemplo, el selector de la unidad de buzones 78 puede escribir los punteros k y v dentro de la cola de escritura 80 para la unidad de buzones MUh. Cada uno de los contenedores en las colas 80 puede ser bastante pequeño ya que solo almacenan dos punteros, uno para el paquete y otro para la cabecera actualizada.

Volviendo a la Fig. 4B, la unidad de rDMA 81 puede leer cada una de las colas de escritura 80 en una forma de operación por turnos (es decir un paquete desde una cola, otro paquete desde la cola siguiente, etc.). Para cada una de las colas 80, la unidad de rDMA 81 puede acceder al siguiente contenedor de la cola y puede leer los dos punteros (por ejemplo, k y v) almacenados en el mismo. La unidad de rDMA 81 puede acceder a continuación (flecha 88) al paquete almacenado en la dirección del paquete (por ejemplo k) y puede acceder (flecha 89) a la cabecera almacenada en la dirección de cabecera (por ejemplo, v) y puede escribirlas, como la dirección de destino añadida, en el siguiente paquete disponible y el almacenamiento de cabeceras en la unidad de buzones MU asociada con la cola de escritura. Se apreciará que la unidad de rDMA 81 normalmente puede operar a medida que los datos se hacen disponibles para la escritura, no está sincronizado con la NIC que escribe dentro de la sección de datos de entrada 70.

Por ejemplo, la cola de escritura MUh puede estar asociada con la unidad de buzones MUh y de este modo, el rDMA 81 puede escribir (flecha 90) el paquete almacenado en la sección de memoria k a la sección de memoria, tal como la sección de memoria q, en el área de almacenamiento de paquetes 70M de la unidad de buzones MUh y puede escribir (flecha 91) la cabecera almacenada en el contenedor v, en un contenedor tal como el contenedor a, en un área de almacenamiento de cabeceras 72M de la unidad de buzones MUh. Antes de escribir la cabecera dentro de su contenedor (por ejemplo el contenedor a), el rDMA 81 puede actualizar la cabecera para incluir la nueva sección de memoria (por ejemplo, q) del paquete.

La sección de datos de entrada 60M de cada unidad de buzones 54 puede comprender el área de almacenamiento de paquetes 70M y el área de almacenamiento de cabeceras 72M como para las secciones de entrada 57, el área de almacenamiento de paquetes 70M puede tener un tamaño variable de las secciones de memoria mientras que el área de almacenamiento de cabeceras 72M puede tener contenedores de tamaño fijo.

Ejecutándose sobre una CPU 62M puede estar un generador de huellas digitales 92, un selector de buzones 94, un clasificador 96 y un despachador 98. La sección de datos de salida 64M puede comprender una multiplicidad de colas de escritura 100, una por sección de salida 59 y una unidad de rDMA 99.

Además, cada una de las unidades de buzones 54 puede comprender una multiplicidad de colas de buzones 102. Estas colas 102 pueden representar buzones 22 pero, de acuerdo con una realización preferida de la presente invención, no almacenan propiamente los datos; en su lugar, almacenan los punteros a las cabeceras y los datos almacenados en la sección de datos de entrada 60M. Por consiguiente, cada una de las colas de buzón 102 puede ser bastante pequeña. Se apreciará que, teniendo colas separadas 102 para cada buzón, en lugar de almacenar los paquetes en sus buzones, posibilita que las colas sean muy pequeñas y de este modo puede ser relativamente fácil almacenarlas bien en una memoria de acceso aleatorio o en memoria caché para un procesamiento rápido.

El generador de huellas digitales 92 puede leer las cabeceras en orden, generando la huella digital de los datos de la huella digital original de cada una de las cabeceras para determinar a qué cola de buzón 102 asociará los paquetes. El generador de huellas digitales 92 puede proporcionar un valor de huella digital H_2 al selector de buzones 94 que, a su vez, puede seleccionar la cola de buzón 102 indicada por el valor de huella digital H_2 . El selector de buzones 94 puede proporcionar (flecha 103) el valor de buzón, listado en este caso como B, al clasificador 96 para determinar cómo manejar el paquete de entrada.

El clasificador 96 puede controlar el acceso a la cola de buzón 102 basándose en la calidad de servicio u otros requisitos de canal para ese buzón. Tales requisitos pueden estar asociados con cada buzón 22 (indicado en la Fig. 4B por líneas discontinuas 101 en un extremo de la cola de buzón 102) y pueden estar disponibles para revisión por el clasificador 96. Otra información, tal como la contabilidad abierta, la seguridad, la cadena de sesión, etc., también puede estar asociada con cada cola de buzón 102, así como la sección de salida 59 a través de la cual transmitir los

paquetes a su destino, y la dirección de destino original.

5 El clasificador 96 puede intentar mantener una tasa de tráfico prefijada para el buzón actual. El clasificador 96 puede desechar cualesquiera paquetes de entrada cuando la cola de buzón actual 102 está llena. El desecho de tales paquetes puede permitir que protocolos tales como el TCP sincronicen sobre la tasa apropiada para ese buzón. El tamaño máximo de buzón se puede definir bien por el número de paquetes, medido a través de un contador de paquetes, o por el tamaño total de paquetes.

10 La cola de buzón 102 puede ser una cola cíclica y de este modo, el clasificador 96 puede determinar que la cola del buzón 102 puede estar llena cuando un puntero a la cabeza de la cola del buzón 102 está cerca de un puntero al final de la cola del buzón 102. Cuando el clasificador 96 desecha un paquete, el clasificador 96 puede aumentar un contador de 'paquetes desechados' para ese buzón. El clasificador 96 puede proporcionar (flecha 105) sus instrucciones de vuelta al selector de buzón 94 el cual, si no se desechó el paquete, puede escribir (flecha 107) los punteros a la cabecera y las direcciones de los paquetes dentro del siguiente contenedor de la cola del buzón seleccionado 102. En la realización de ejemplo de la Fig. 4B, estas direcciones son a y q respectivamente.

15 El despachador 98 puede revisar (Fig. 4C) cada una de las colas de buzón 102 para determinar la tasa a la que extraer los paquetes desde el mismo y para conformar el flujo de salida. El despachador 98 puede utilizar un algoritmo de conformación tal como el algoritmo "leaky bucket" ("cubo agujereado") como se conoce en la técnica, que extrae paquetes desde una cola mientras que cumple con el parámetro de QOS / conformación definido para cada buzón. Se apreciará además que también se pueden usar otras fórmulas de conformación en la presente invención.

20 Cuando el despachador 98 determina que un paquete puede estar listo para su liberación desde un buzón actual, el despachador 98 puede actualizar la cabecera con la dirección de destino de la sección de salida 59 asociada con la cola de buzón actual 102. El despachador 98 puede copiar (flecha 109) la información del puntero almacenada en el contenedor actualmente en la cabeza de la cola del buzón 102 a la cola de escritura 100 asociada con la cola del buzón 102. La Fig. 4B muestra que la sección de salida etiquetada como 59n es la asociada con la cola de buzón de ejemplo 102.

25 Volviendo a la Fig. 4D, la unidad de rDMA 99 puede leer cada cola de escritura 100 en una forma por turnos. Para cada cola 100, la unidad de rDMA 99 puede acceder al siguiente contenedor de la cola y puede leer los dos punteros (por ejemplo, a y q) almacenados en el mismo. La unidad de rDMA 99 puede acceder a continuación (flecha 110) a la cabecera almacenada en la dirección de la cabecera (por ejemplo, a) y puede acceder (flecha 111) al paquete almacenado en la dirección del paquete (por ejemplo q) y escribirlas, como la dirección de destino actualizada, dentro de la siguiente cabecera disponible y el almacenamiento de paquetes en la sección de salida 59 asociada con la cola de escritura.

30 Por ejemplo, la cola de escritura 59n puede estar asociada con la sección de salida 59n y de este modo, el rDMA 102 puede escribir (flecha 112) la cabecera almacenada en un contenedor a, en un contenedor tal como el contenedor m, en el área de almacenamiento de cabeceras 72R de la sección de salida 59n y puede escribir (flecha 113) el paquete almacenado en la sección de memoria q a una sección de memoria, tal como la sección de memoria s, en el área de almacenamiento de paquetes 70R de la sección de salida 59n. Antes de escribir la cabecera dentro de su contenedor (por ejemplo, el contenedor m), el rDMA 102 puede actualizar la cabecera para incluir la nueva sección de memoria (por ejemplo, s) del paquete. Se apreciará que la unidad de rDMA 99, normalmente puede operar a medida que los datos se hacen disponibles para su escritura; no está sincronizada con la unidad de rDMA 81 (Fig. 4B) o cualquier otra unidad de escritura.

35 La sección del área de datos de entrada 60R de cada sección de salida 59 puede comprender el área de almacenamiento de paquetes 70R y el área de almacenamiento de cabeceras 72R. Como en las otras unidades, el área de almacenamiento de paquetes 70R puede tener un tamaño variable de las secciones de memoria mientras que el área de almacenamiento de cabeceras 72R puede tener contenedores de tamaño fijo.

40 Ejecutándose sobre la CPU 62R puede estar un solicitador de escritura 120. La sección de datos de salida 64R puede comprender una única cola de transmisión 122 y una unidad de DMA 124.

45 El solicitador de escritura 120 puede revisar (flecha 126) cada una de las nuevas cabeceras, comprobando la dirección de destino original y puede buscar la siguiente dirección de destino para la dirección de destino original y su copia local de la tabla de ARP, etiquetada como 130. El solicitador de escritura 120 puede actualizar la cabecera almacenada en el paquete para incluir la dirección MAC al siguiente destino para el paquete (es decir, el destino externo) y puede escribir (flecha 128) un puntero a la dirección del paquete dentro del siguiente contenedor disponible de la cola de transmisión 122. En la realización de ejemplo de la Fig. 4D, esta dirección es s.

50 Volviendo a la Fig. 4E, la unidad de DMA 124 puede acceder al siguiente contenedor de la cola de transmisión 122 y puede leer el puntero (por ejemplo, s) almacenado en el mismo. La unidad de DMA 124 puede acceder a continuación (flechas 132 y 133) al paquete almacenado en la dirección del paquete (por ejemplo, s) y puede escribirla, como la dirección de destino actualizada, fuera hacia la red. Se apreciará que la unidad de DMA 124 normalmente puede operar a medida que los datos se hacen disponibles para la escritura; no está sincronizada con

la unidad de rDMA 81 o 99.

Se apreciará que el enrutador 50 (Fig. 2) puede escribir sólo una vez por unidad, realizando todo su procesamiento sobre las cabeceras almacenadas separadamente de los paquetes. Además, rellenando las colas de transmisión y los buzones con punteros en lugar de con los datos reales, el enrutador 50 minimiza las operaciones de lectura y escritura. De hecho, las CPU 62 no copian los datos desde un lugar al otro en memoria, y, como resultado, procesan los datos mucho más rápidamente. Las CPU 62 meramente revisan las cabeceras, una cantidad relativamente pequeña de datos que se pueden llevar a la memoria caché de trabajo y procesarse relativamente rápidamente.

Además, la entrada y la salida están separadas, de modo que se puede escribir en un buzón y leerse desde el mismo al mismo tiempo. Además, las operaciones de entrada y salida no tienen que estar necesariamente sincronizadas; ocurren cada vez que hay datos disponibles en las diversas colas de escritura. A pesar de esto, el enrutador de rejilla 50 puede mover datos a través del mismo rápidamente.

Se apreciará que el rDMA y las unidades de DMA realizan el transporte del tráfico del servicio agregado desde una unidad del enrutador 50 a otra unidad segura. El protocolo del rDMA es un protocolo seguro con control de flujo y, como resultado, no genera colisiones. Se apreciará además que la presente invención puede usar otros protocolos de colocación de datos directa que soportan el acceso a memoria directo del ordenador y dan acceso al sistema operativo a los datos recibidos desde la red sin realizar ninguna copia de memoria basada en la CPU. Ejemplos de otros tipos de unidades de colocación directa de datos pueden ser los dispositivos de descarga de TCP, RNIC (NIC con RDMA habilitado), NIC con DDP habilitado, y otros dispositivos de descarga.

Se observa que el flujo anterior es operativo para los servicios existentes. Ahora haremos referencia a la Fig. 5, que ilustra el procedimiento para un nuevo paquete. Cuando llega un nuevo paquete a una sección de entrada 57, el creador de cabeceras 74 puede comprobar la dirección de origen para ver si el servicio representado es conocido. En una realización, el creador de cabeceras 74 puede comprobar una tabla de ARP local (procedimiento de registro de direcciones) 141. Si existe el servicio, el creador de cabeceras 74 puede enviar a la dirección de origen una "respuesta de ARP". Sin embargo, si el servicio no está listado en la tabla de ARP, el creador de cabeceras 74 puede comenzar un nuevo procedimiento de registro de buzón. En primer lugar, el creador de cabeceras 74 puede solicitar al generador de huellas digitales 76 que calcule el valor de la huella digital H_1 a partir de los datos en el paquete. Basándose en los resultados de la huella digital, el selector de la unidad de buzones 78 puede determinar la unidad de buzones 54 para el servicio y puede enviar un comando de Registro de Buzón, que puede incluir una referencia para la identificación del servicio y para un identificador de la sección de entrada 57, para la unidad de buzones seleccionada 54.

Una nueva unidad de servicio 140 puede recibir el comando de Registro de Buzón y puede emitir una petición a la tabla 30 (Fig. 1) de servicios y usuarios. Con la contestación desde la tabla 30, la nueva unidad de servicio 140 puede generar un nuevo buzón 22 que tiene los parámetros del servicio recibidos desde la tabla 30 y que tiene un valor de huella digital apropiada generada de acuerdo con la fórmula para ese tipo de servicio. La nueva unidad de servicio 140 también puede determinar la sección de salida 59, normalmente a partir de la huella digital de la dirección de origen y puede almacenar esta, así como la dirección de origen listada en el paquete, en una tabla de ARP local 142.

La nueva unidad de servicio 140 puede enviar una confirmación de vuelta al creador de cabeceras 74 de la sección de entrada 57 que, a su vez, puede enviar una contestación de ARP a la dirección de origen.

REIVINDICACIONES

1. Un enrutador de rejilla que comprende:

5 una pluralidad de unidades de interfaces externas (44) configuradas para recibir paquetes de servicios desde un punto y para proporcionar los paquetes de los servicios a otro punto; y
 una segunda pluralidad de unidades de buzones (46) configurados para almacenar y redirigir dichos paquetes asociados con los buzones respectivos, conectada cada una de las unidades de buzones (46) a dichas unidades de interfaces externas (44),
 en el que cada una de dichas unidades de interfaces externas (44) tiene una porción de entrada y una porción de salida, y

10 cada una de dichas porciones y cada una de las unidades de buzones, comprende:

una sección de entrada configurada para almacenar dichos paquetes y las cabecera relacionadas;
 un procesador de buzón configurado al menos para tomar decisiones de encaminamiento; y
 una sección de salida configurada para organizar dichos paquetes para transferir a una unidad siguiente, en el que:

15 dicha sección de entrada de dicha porción de entrada está conectada a una pluralidad de tarjetas de interfaz de red, y se **caracteriza porque**:

20 cada una de las tarjetas de interfaz de red está configurada para escribir directamente en dicha sección de entrada,
 cada una de dichas tarjetas de interfaz de red comprende una librería de espacios de usuario, y se asigna un buzón por servicio y porque dicha sección de salida está configurada además para almacenar al menos un puntero a dicha sección de entrada.

25 2. El enrutador de rejilla de acuerdo con la reivindicación 1 y en el que dichas porciones de entrada comprenden medios para escribir en dichas unidades de buzones según la información en dichos paquetes, dichas unidades de buzones comprenden medios para escribir paquetes en porciones de salida asociadas con dichos buzones en los que están almacenados dichos paquetes y dichas porciones de salida comprenden medios para escribir paquetes a dicho otro punto.

3. El enrutador de rejilla de acuerdo con la reivindicación 1 o la reivindicación 2, y en el que dicho procesador de buzón está configurado para procesar solo dichas cabeceras para tomar dichas decisiones de encaminamiento.

30 4. El enrutador de rejilla de acuerdo con la reivindicación 3 en el que cada una de las unidades de buzón está conectada a dichas unidades de interfaces externas en una red de tipo CLOS.

5. El enrutador de rejilla de acuerdo con una cualquiera de las reivindicaciones anteriores, y en el que dicha librería comprende una unidad de DMA de dispersión-recogida para escribir dichos paquetes desde dicha librería a dicha sección de entrada.

35 6. El enrutador de rejilla de acuerdo con una cualquiera de las reivindicaciones anteriores, y en el que al menos una de dichas secciones de salida comprende al menos una cola para almacenar dicho al menos un puntero a las localizaciones en dicha sección de entrada, o una cola de escritura por unidad en la que se escribe.

40 7. El enrutador de rejilla de acuerdo con la reivindicación 3 y en el que dicha unidad de buzones comprende adicionalmente una sección de buzones configurada para almacenar colas de buzones, asociada cada una con un buzón, en el que cada una de dichas colas de buzones está configurada para almacenar punteros a localizaciones en dicha sección de entrada donde se almacenan sus datos, o en el que dicho procesador de buzón también comprende un programador para programar y controlar la salida de dichos paquetes, o en el que dicho procesador de buzón comprende:

45 un generador de huellas digitales para generar una huella digital de una cabecera de un paquete para determinar a qué buzón pertenece dicho paquete;
 un calificador para controlar la entrada a dicho buzón de acuerdo con los requisitos definidos previamente para dicho buzón;
 un selector de buzón para almacenar punteros a dicha cabecera y a dicho paquete en una cola de buzón asociada con dicho buzón determinado bajo instrucción de dicho calificador; y un despachador configurado para conformar un flujo de salida desde dicho buzón y para almacenar dichos punteros en una cola de escritura definida previamente para dicho buzón.
 50

8. El enrutador de rejilla de acuerdo con cualquiera de las reivindicaciones anteriores, y en el que dicha porción de entrada comprende además un procesador de la porción de entrada, comprendiendo dicho procesador de la porción de entrada;
 un creador de cabeceras configurado para crear una nueva cabecera con un puntero a la localización de

- almacenamiento del paquete;
un generador de huellas digitales configurado para leer dicha nueva cabecera y para determinar en qué unidad de buzones almacenar dicho paquete; y
un selector de unidad de buzones configurado para escribir punteros a dicha nueva cabecera y a dicha localización de almacenamiento en una cola de escritura para dicha unidad de buzones determinada.
- 5
9. El enrutador de rejilla de acuerdo con cualquiera de las reivindicaciones anteriores, teniendo dicha porción de salida un procesador, comprendiendo dicho procesador de la porción de salida:
- un solicitador de escritura configurado para actualizar una cabecera de un paquete con una dirección de destino de dicho paquete y para escribir un puntero a dicho paquete dentro de una cola de transmisión.
- 10
10. El enrutador de rejilla de cualquiera de las reivindicaciones anteriores, en el que dichas unidades se implementan sobre PC (ordenadores personales) que tienen controladores de memoria y en el que dichas unidades se configuran para mover dichos paquetes a través del mismo sustancialmente a la tasa de E/S (Entrada / Salida) de dichos controladores de memoria.
- 15
11. El enrutador de rejilla de acuerdo con la reivindicación 10 y que también comprende una unidad de colocación directa de datos configurada para transferir paquetes desde una unidad a otra.
12. El enrutador de rejilla de acuerdo con la reivindicación 11 y en el que dicha unidad de colocación de datos directa es una de: una unidad de rDMA, un dispositivo de descarga de TCP, una tarjeta de interfaz de red con RDMA habilitado y una tarjeta de interfaz de red con DDP habilitado.
13. Un procedimiento de redirigir el tráfico a través de enrutadores de rejilla, comprendiendo el procedimiento:
- 20
- en un enrutador de rejilla:
almacenar paquetes y cabeceras de una pluralidad de buzones de una unidad de buzones (46) en un área de almacenamiento general;
procesar cada una de dichas cabeceras para determinar a qué buzón pertenece cada uno de los paquetes y la cabecera asociada;
- 25
- colocar punteros a dicho paquete y su cabecera asociada dentro de una cola de buzón para dicho buzón; y
procesar dicha cola de buzón para controlar la entrada y la salida de dicho buzón,
y **caracterizado porque** dicho procesamiento de dicha cola de buzón comprende transferir paquetes y cabeceras, almacenadas en áreas separadas de memoria de un ordenador personal y apuntadas en dicha cola de escritura de dicho ordenador personal a áreas de memoria separadas de otro ordenador personal a través
- 30
- de una unidad de colocación directa de datos.

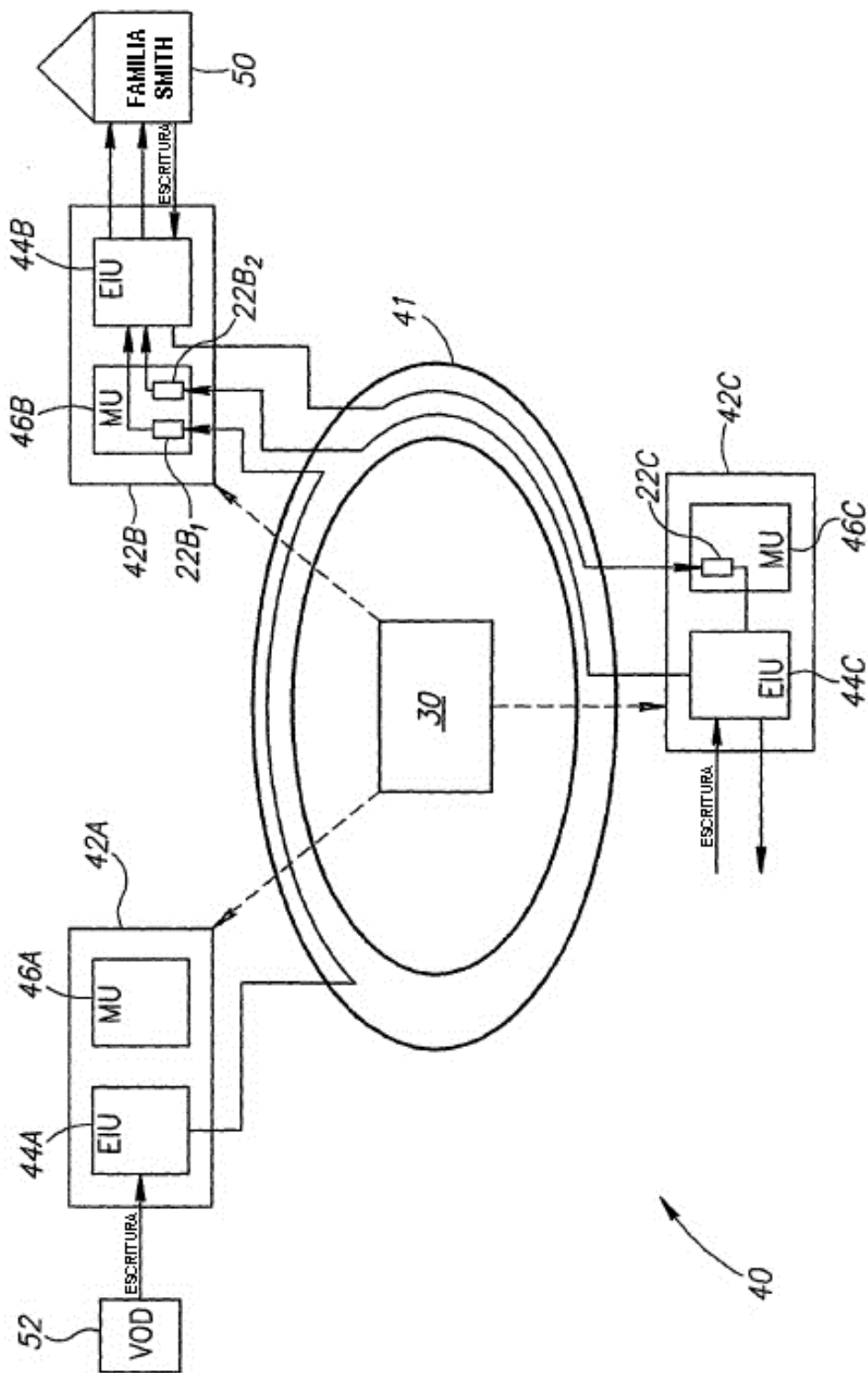


FIG.1

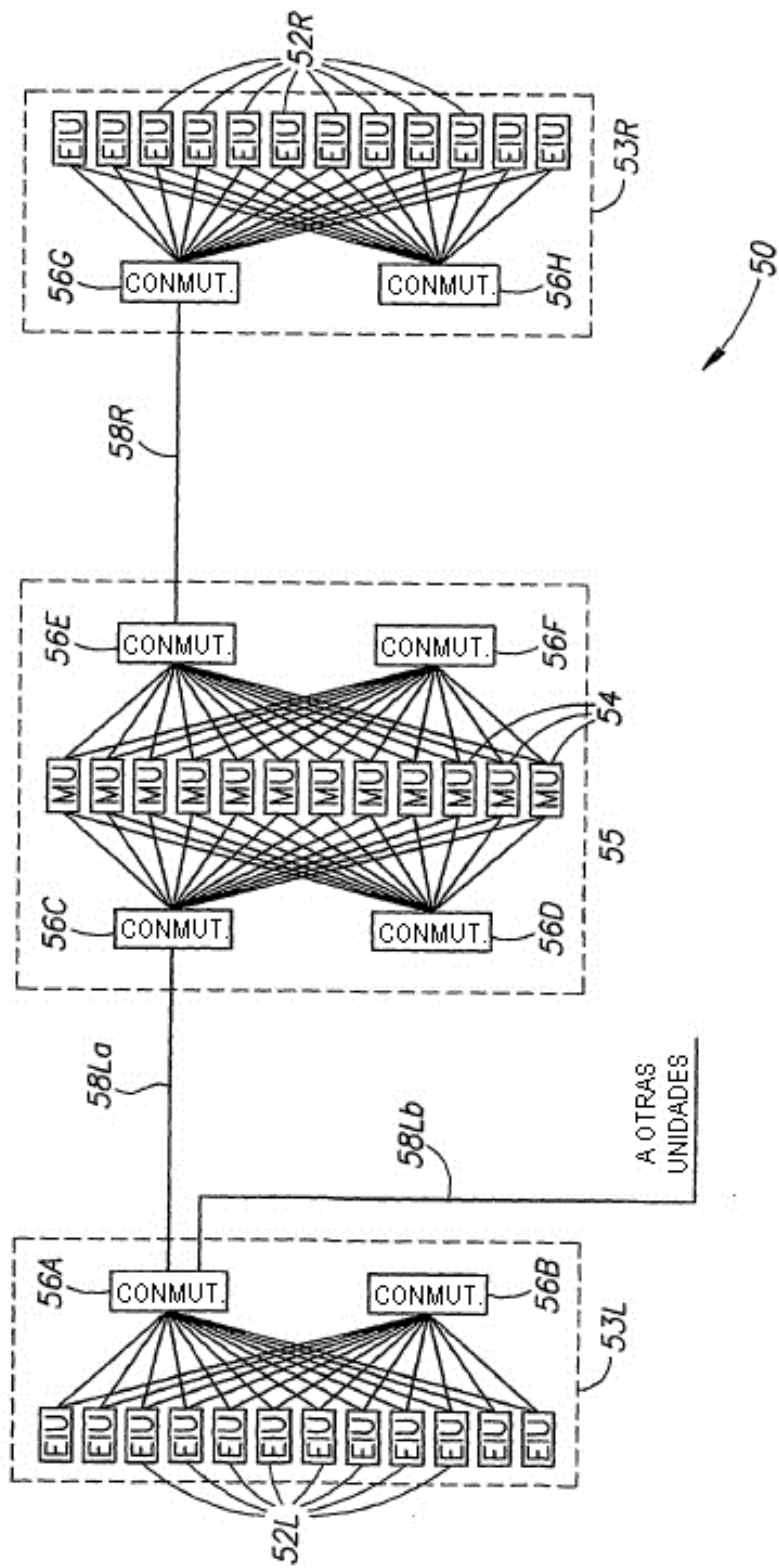


FIG.2

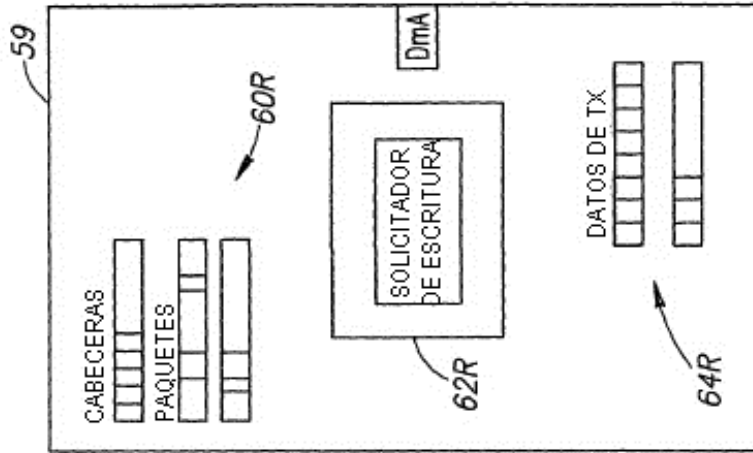


FIG. 3C

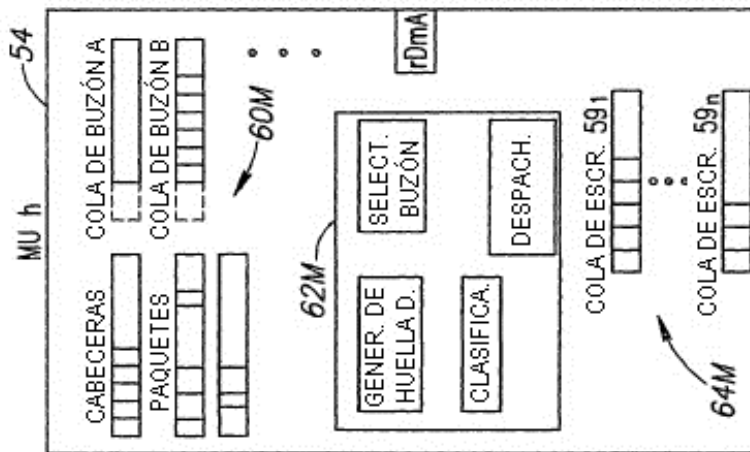


FIG. 3B

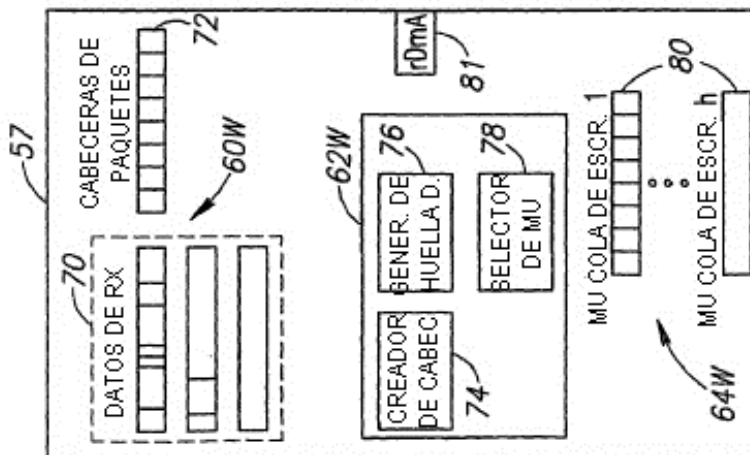


FIG. 3A

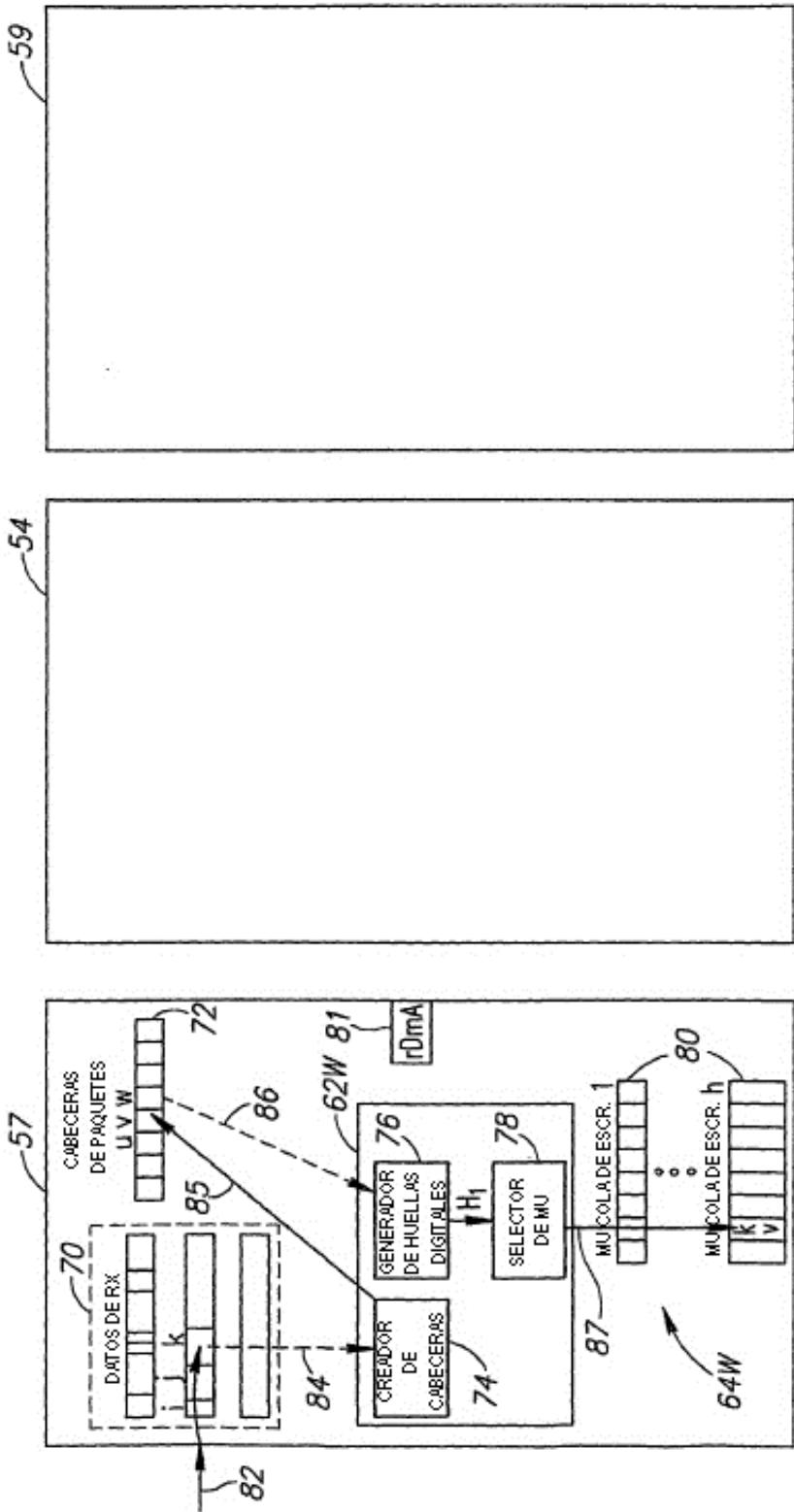


FIG. 4A

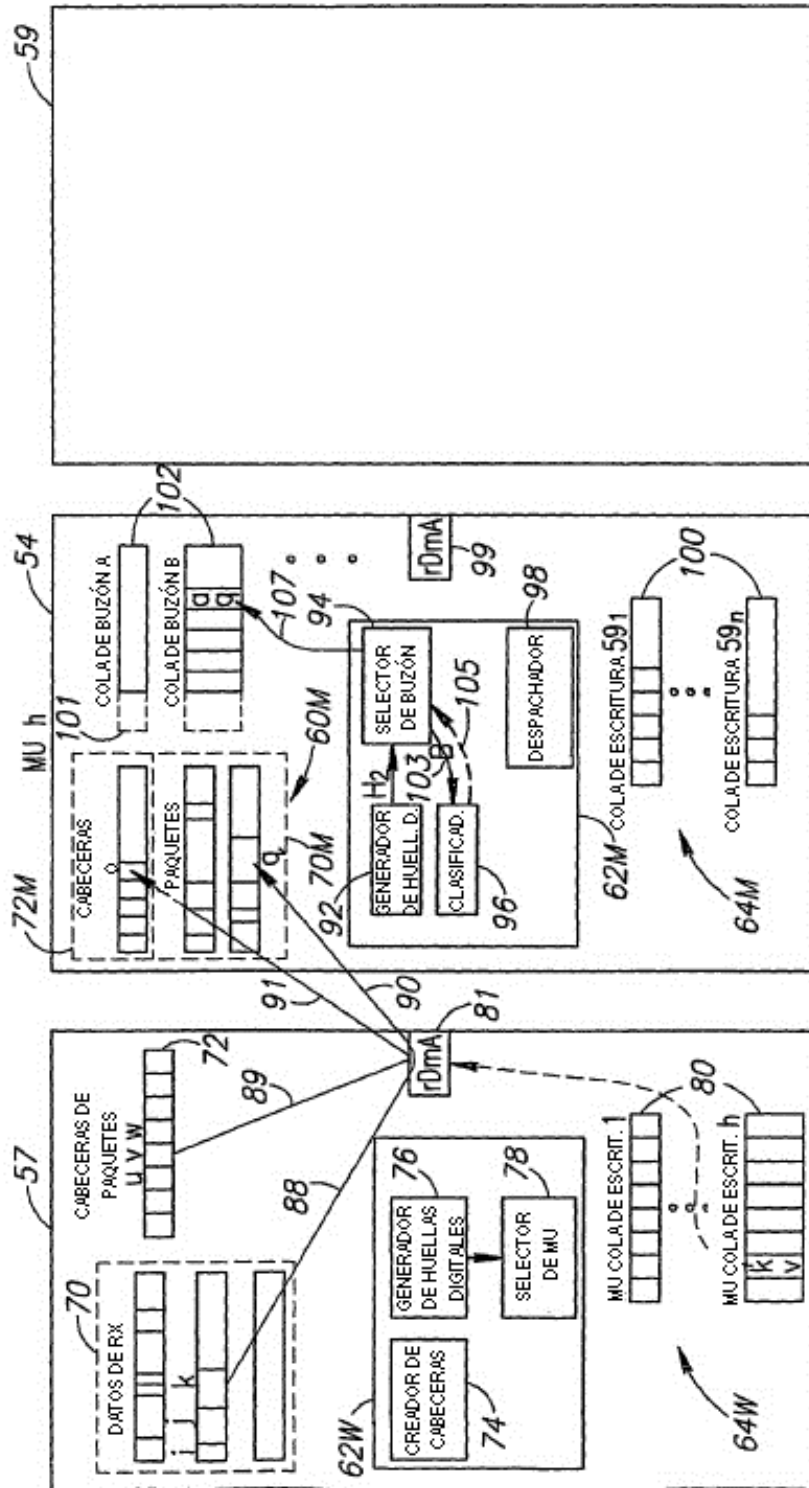


FIG.4B

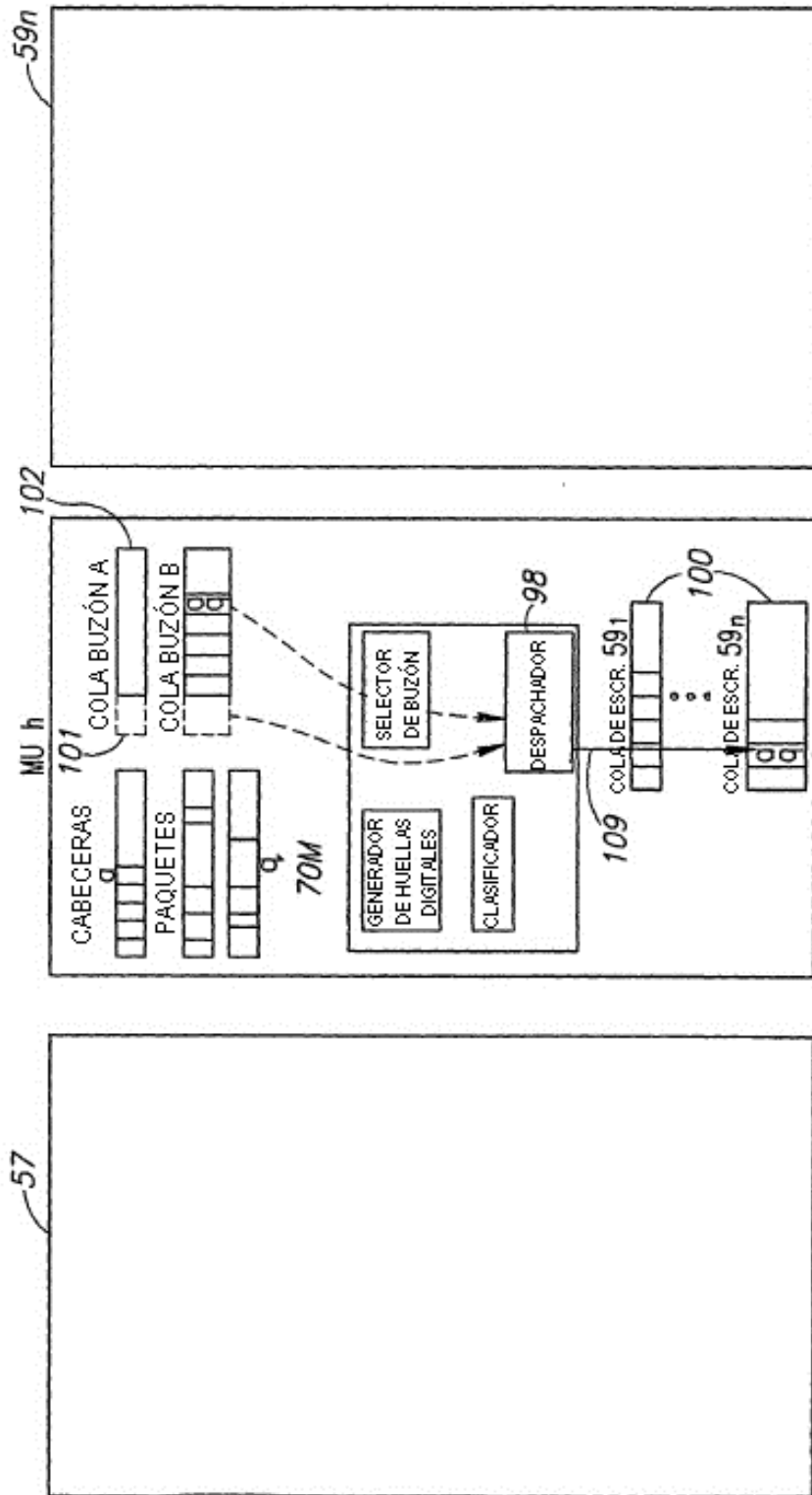


FIG. 4C

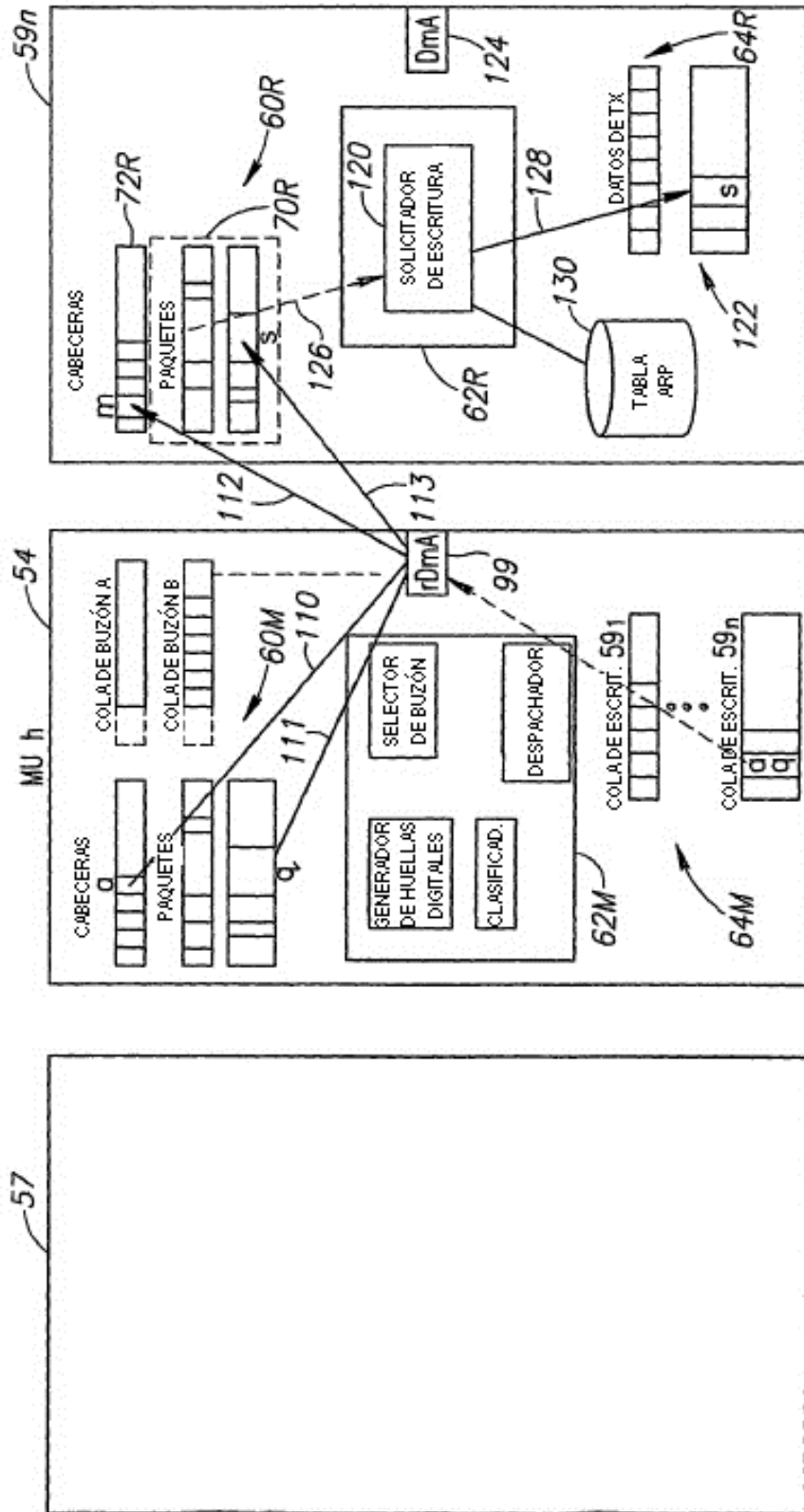


FIG.4D

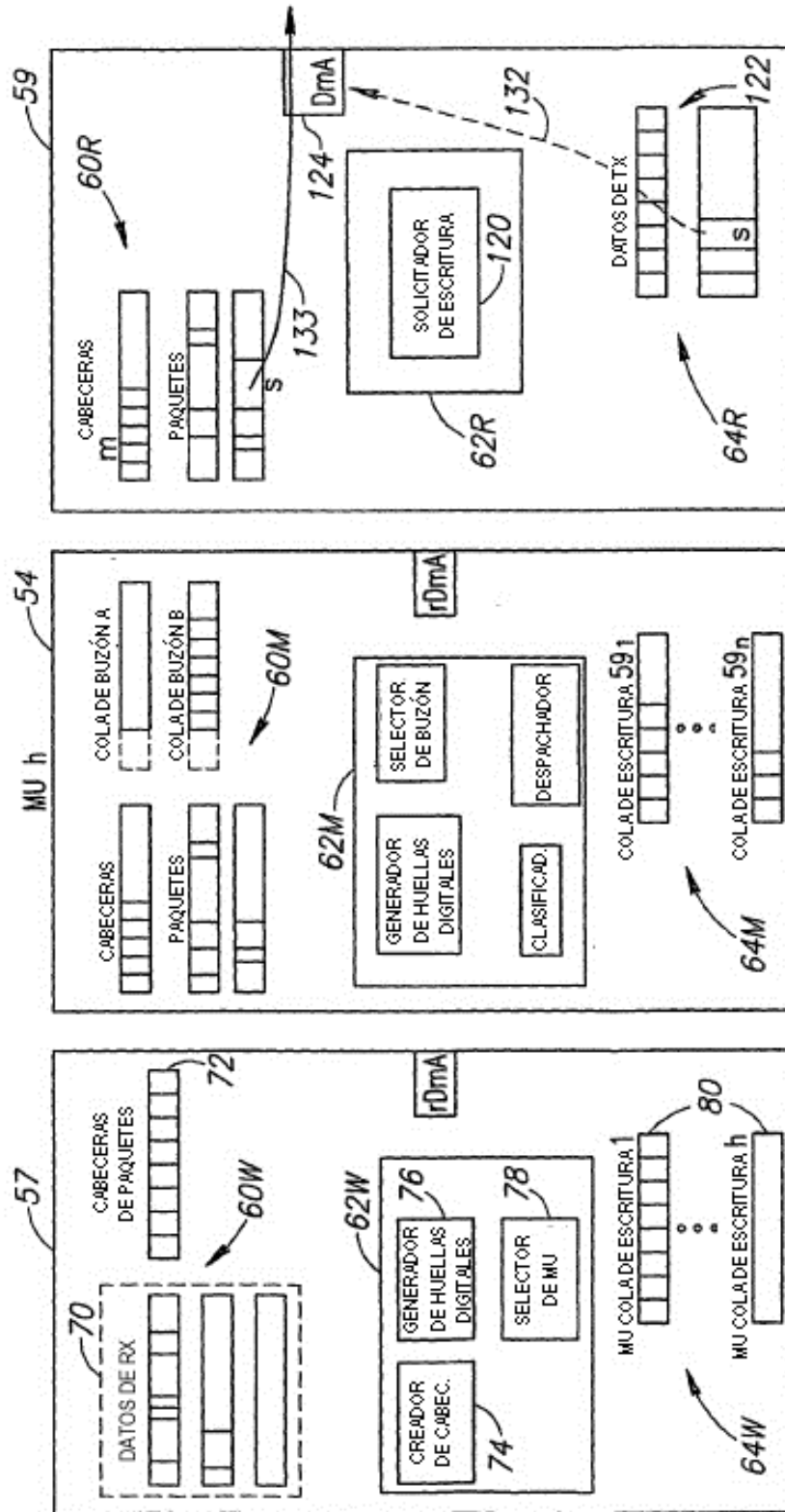


FIG.4E

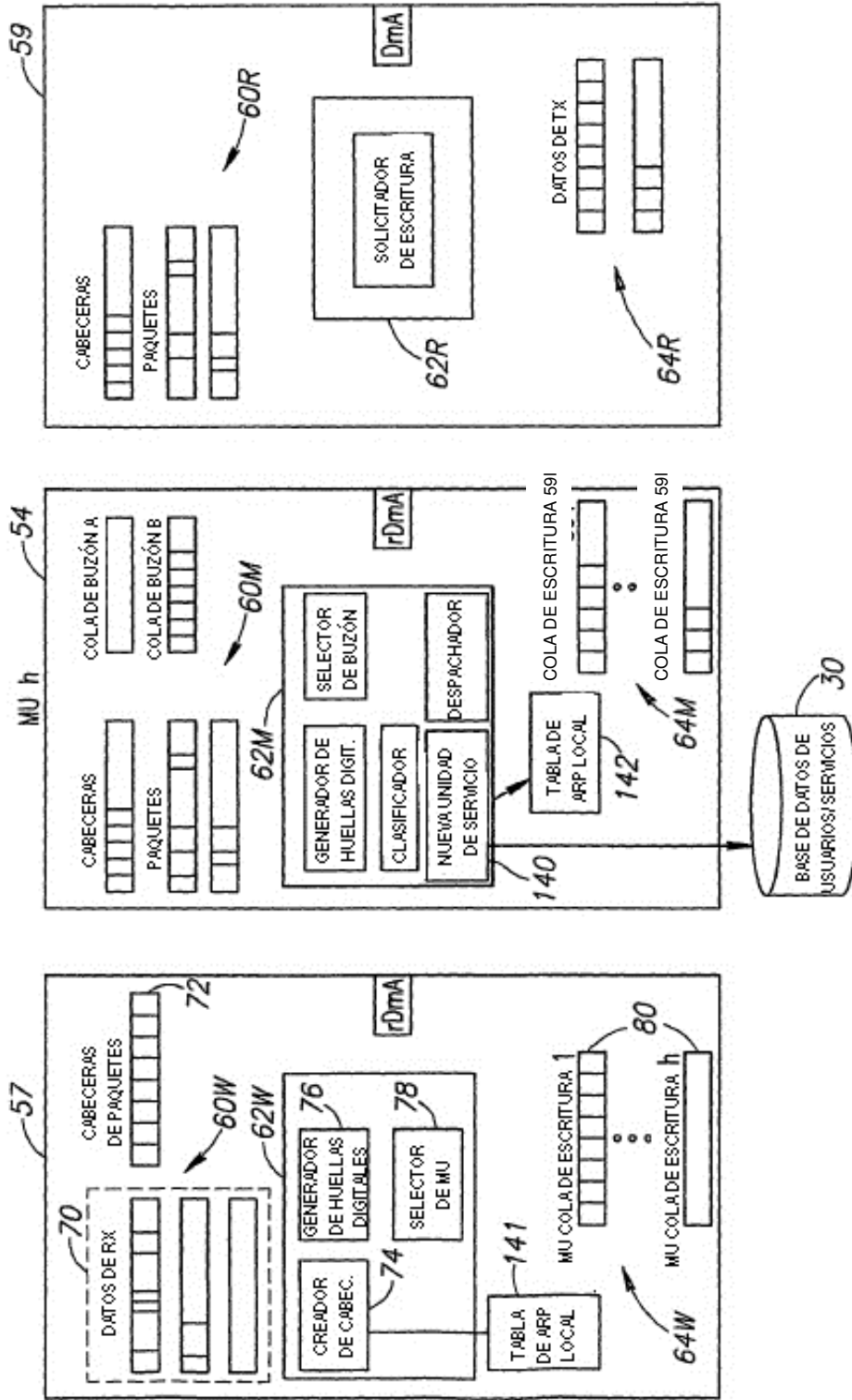


FIG.5