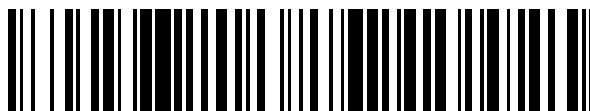


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 878 073**

51 Int. Cl.:

H04W 24/02

(2009.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **24.10.2017** **PCT/EP2017/077164**

87 Fecha y número de publicación internacional: **02.05.2019** **WO19080997**

96 Fecha de presentación y número de la solicitud europea: **24.10.2017** **E 17790753 (2)**

97 Fecha y número de publicación de la concesión europea: **09.06.2021** **EP 3701740**

54 Título: **Gestión de recursos en redes de acceso por radio en la nube**

45 Fecha de publicación y mención en BOPI de la
traducción de la patente:
18.11.2021

73 Titular/es:

NOKIA SOLUTIONS AND NETWORKS OY
(100.0%)
Karakaari 7
02610 Espoo, FI

72 Inventor/es:

SCHEFCZIK, PETER;
HABERLAND, BERND y
KLOTSCHKE, RALF

74 Agente/Representante:

ISERN JARA, Jorge

ES 2 878 073 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

Gestión de recursos en redes de acceso por radio en la nube

5 Campo

Las realizaciones de ejemplo y no limitantes de la invención se relacionan en general con sistemas de comunicación inalámbrica. Realizaciones de la invención se relacionan especialmente con aparatos, métodos, y productos de programas de ordenador en redes de comunicación.

10

Antecedentes

15 La siguiente descripción de la técnica antecedente puede incluir conocimientos, descubrimientos, entendimientos o divulgaciones, o asociaciones en conjunto con divulgaciones no conocidas en la técnica relevante antes de la presente invención pero proporcionadas por la invención. Algunas de tales contribuciones de la invención se pueden señalar específicamente a continuación, mientras que otras de tales contribuciones de la invención serán evidentes a partir de su contexto.

20 Los sistemas de comunicación inalámbrica están bajo constante desarrollo. Actualmente está bajo desarrollo la denominada 5a generación (5G) de red de comunicación móvil. Las futuras redes de comunicación deben tener una arquitectura flexible que sea capaz de adaptarse a una multitud de diferentes casos de uso, usuarios, sensores, requisitos de servicio y despliegues de redes. Esto no se limita a la separación usual de control desde el tráfico de usuarios, sino que también tiene un efecto sobre la estructura del sistema. Debe decidirse dónde procesar ciertas funciones de la red móvil 5G.

25

Una posible implementación de realización de un sistema de comunicación flexible es utilizar cabezales de radio remotos (RRH) y una realización basada en la nube de al menos parte de la infraestructura del sistema. Un cabezal de radio remoto comprende una parte de los elementos de una estación base convencional. Típicamente, un cabezal de radio remoto comprende equipo de radiofrecuencia, convertidores de analógico a digital/digital a analógico y convertidores ascendentes/descendentes, por ejemplo. El resto de la funcionalidad de estación base e infraestructura del sistema pueden ubicarse en otro lugar. En una Red de Acceso por Radio en la Nube (RAN en la Nube) las funciones de red pueden ejecutarse en un grupo de recursos informáticos incluyendo aceleradores de hardware para tareas de capa física (PHY) especial.

30

35 El control eficiente de los recursos del sistema basado en la nube es una tarea desafiante. Las nubes tradicionales están diseñadas para manejar un gran número de tareas de procesamiento paralelo. Sin embargo, los requisitos de un sistema de comunicación son diferentes y controlar un sistema de comunicación basado en la nube no es una tarea trivial, especialmente en los sistemas en desarrollo los retrasos en la programación de recursos deben mantenerse tan pequeños como sea posible.

40

La solicitud de patente WO 2016/074702 A1 divulga una red de acceso basada en la nube en donde se controla el uso de recursos de nivel en diferentes conexiones y servicios.

Resumen

45

La invención está definida por las reivindicaciones anexas.

50 En un primer aspecto, la invención proporciona un método de acuerdo con la reivindicación 1. En un segundo aspecto, la invención proporciona un aparato de acuerdo con la reivindicación 11. En un tercer aspecto, la invención proporciona un producto de programa de ordenador de acuerdo con la reivindicación 18. Realizaciones preferidas están definidas por las reivindicaciones dependientes.

Lista de dibujos

55 Realizaciones de la presente invención se describen a continuación, solo a modo de ejemplo, con referencia a los dibujos acompañantes, en los cuales

La figura 1 ilustra un ejemplo de un entorno de comunicación donde se pueden aplicar algunas realizaciones de la invención

60

La figura 2 es un diagrama de flujo que ilustra una realización de la invención;

La figura 3 ilustra un ejemplo de realización de bucles de control;

65 La figura 4 ilustra un ejemplo de control de recursos de ancho de banda de red media y recursos de procesamiento de banda base;

La figura 5 ilustra un ejemplo de migración de carga; y

La figura 6 ilustra un ejemplo de un aparato que aplica realizaciones de la invención.

Descripción de algunas realizaciones

Las siguientes realizaciones son solo ejemplos. Aunque la especificación puede referirse a "una", "uno", o "algunas" realizaciones en varias ubicaciones, esto no significa necesariamente que cada una de tal referencia sea a las mismas realizaciones, o que la característica solo se aplica a una única realización. También se pueden combinar características únicas de diferentes realizaciones para proporcionar otras realizaciones. Adicionalmente, las palabras "que comprende" y "que incluye" deben entenderse que no limitan las realizaciones descritas para que consistan solo en las características que se han mencionado y tales realizaciones también pueden contener también características, estructuras, unidades, módulos, etc. que no se han mencionado específicamente.

Algunas realizaciones de la presente invención son aplicables a, una estación base, eNodeB, una realización distribuida de una estación base, un elemento de red de un sistema de comunicación, un componente correspondiente, y/o cualquier sistema de comunicación o cualquier combinación de diferentes sistemas de comunicación que soportan la funcionalidad requerida.

Los protocolos usados, las especificaciones de sistemas de comunicación, servidores y equipo de usuario, especialmente en la comunicación inalámbrica, se desarrollan rápidamente. Tal desarrollo puede requerir cambios extra en una realización. Por lo tanto, todas las palabras y expresiones deben interpretarse de manera amplia y están previstas para ilustrar, no restringir, realizaciones.

La figura 1 ilustra una vista simplificada de un entorno de comunicación mostrando solo algunos elementos y entidades funcionales, todos siendo unidades lógicas cuya implementación puede diferir de lo que se muestra. Las conexiones que se muestran en la figura 1 son conexiones lógicas; las conexiones físicas reales pueden ser diferentes. Es evidente para una persona experta en la técnica que los sistemas también comprenden otras funciones y estructuras. Debe apreciarse que las funciones, estructuras, elementos y los protocolos usados en o para la comunicación son irrelevantes para la invención. Por lo tanto, no necesitan ser discutidos con más detalle aquí.

En el ejemplo de la figura 1, se muestra un sistema basado en Red de Acceso por Radio en la Nube (RAN en la Nube). Sin embargo, las realizaciones descritas en estos ejemplos no se limitan al sistema ilustrado en la figura 1 que es simplemente un ejemplo de posible sistema.

El ejemplo simplificado de una red de la figura 1 se puede dividir en tres partes, o niveles: una Nube 100 Central, una o más Nubes de Borde, EC, 102, 104, y un número de Unidades de Extremo Frontales Flexibles, FEU, 106, 108, 110, a las cuales están conectados un conjunto de Cabezales de Radio Remotos, RRH, 112, 114, 116. En una realización, las Unidades de Extremo Frontales Flexibles se pueden conectar a las Nubes de Borde a través de redes definidas por Software, SDN. Las conexiones entre FEUs y Nubes de Borde pueden denotarse como red media. Las Nubes de Borde se pueden conectar a Nube Central a través de una red de retorno. Las conexiones entre RRHs y FEUs pueden denotarse como red frontal.

En una realización, diferentes partes o niveles del sistema realizan diferentes funciones. Los Cabezales de Radio Remotos se encargan de la interfaz aérea y de la operación de radiofrecuencia. Típicamente las Unidades de Extremo Frontales Flexibles gestionan servicios de latencia críticos. Una Unidad de Extremo Frontal es un grupo de recursos controlados con métodos en la nube. La FEU es para proporcionar servicios ultra fiables o de baja latencia para ser asignados. El hardware usado para realizar FEU puede ser un servidor con hardware dedicado similar a placas de aceleradores. Las FEUs pueden ubicarse relativamente cerca de los RRHs, pero las ubicaciones deben ser de fácil acceso, mantenimiento, refrigeración, y alimentación. Típicamente en una realización, las conexiones entre RRHs y FEUs se realizan con fibra y la longitud del enlace de conexión es de pocos 100 metros.

En una realización, una Nube de Borde es una nube que proporciona servicios de radio dedicados asignados lógicamente al borde de red, pero de una forma centralizada localmente. Típicamente el área de servicio puede ser un área de alrededor de 20km de diámetro alrededor de RRHs, por ejemplo, dependiendo de la topología de red de transporte. Mediante esta agregación son logrables ganancias de grupo. Como en la FEU también aquí pueden ser necesarios aceleradores de hardware, pero los componentes principales pueden ser servidores ordinarios (disponibles para la venta, OTS). El hardware de Nube de Borde es fácil de acceder, mantener, reemplazar o actualizar, enfriar, y alimentar; pero a mayor escala, impulsando de este modo las ganancias de grupo.

En las redes 5G se propone un concepto de sección. Las secciones se usan para aislamiento de recursos lógicos. Una instancia de sección de red es un conjunto de funciones de red y los recursos para estas funciones de red que se disponen y configuran para formar una red lógica completa para cumplir con ciertas características de red y calidad de servicio (QoS). Cada sección puede estar dedicada a un tipo dado de servicios que tienen características de red similares. Por ejemplo, una sección puede encargarse de la Banda Ancha Móvil, MBB, o MBB mejorada, eMBB, y otra

sección puede ser responsable del tráfico de Internet de las Cosas, IoT. En la Nube Central, el Descriptor 120 de Servicio se encarga de la abstracción de descripciones de secciones con Indicadores Clave de Rendimiento, KPIs e información de ubicación. El orquestador 122 es responsable de la Gestión de extremo a extremo de Sección y MultiAlquiler. La Nube Central también comprende funciones 124 de Núcleo de Paquete Mejorado, EPC.

En una realización, cada Nube de Borde comprende un grupo 126, 128 de unidades de banda base, BBUs. Cada FEU también puede tener un grupo 130, 132, 134 de unidades de banda base. Una BBU resume todo el procesamiento del plano de datos y el control relacionado antes de la modulación final (en la dirección de enlace descendente, y viceversa en dirección de enlace ascendente). Un grupo de BBU es un conjunto de recursos (tal como una unidad central de procesamiento, CPUs, aceleradores, memorias) que se proporcionan en una nube para ofrecer el procesamiento de una funcionalidad de BBU. Una BBU virtual, vBBU, es un conjunto de funcionalidad de BBU que flota en la nube, compuesta por recursos del grupo de BBU para proporcionar uno o más servicios a un conjunto de usuarios o clientes. Típicamente, varias vBBUs pertenecen a una celda.

El Encadenador de Servicio de Banda Base, BBSC, 136, 138 está configurado para controlar el procesamiento de banda base flexible y el equilibrio de carga. Los Orquestadores 140, 142 de RAN, son responsables del despliegue de MultiSección/Alquiler en la Nube de Borde de acuerdo con solicitudes de Orquestador 122.

La figura 1 ilustra un ejemplo donde la red tiene tres partes o niveles separados. En general, el número de partes o niveles puede ser distinto de tres, lo cual se usa aquí como un ejemplo. Los requisitos de latencia y temporización de las partes o niveles se relajan cuanto más lejos está la parte o nivel desde el RRH.

La realización del sistema de acceso por radio basado en la nube descrito anteriormente puede utilizar los denominados microservicios. Un microservicio es un pequeño componente que se ejecuta en su propio proceso y que se comunica con otros microservicios a través de mecanismos ligeros como una API de HTTP (interfaz de programación de aplicaciones de Protocolo de Transferencia de HiperTexto). Los diferentes microservicios se pueden basar en diferentes arquitecturas de software, se pueden implementar en diferentes lenguajes, se pueden iniciar en diferentes cantidades, pero tomados en conjunto se pueden usar para formar una aplicación o servicio más complejo. En el contexto de un sistema de acceso por radio, los bloques de vBBU representados en la figura 1 pueden realizarse como un microservicio y ampliarse para las diferentes secciones de red de acuerdo con la capacidad necesaria. Por ejemplo, una vBBU (MBB) es un microservicio que puede procesar componentes de apilamiento del portador de radio de datos de Banda ancha Móvil. La vBBU (IoT Masivo), la vBBU (IoT V2x) y la vBBU (IoT ind 4.0) son microservicios que pueden procesar componentes de apilamiento del servicio de IoT. El orquestador de RAN y el BBSC también pueden ser microservicios separados. Lo mismo ocurre con el microservicio de Funciones de Celda y el microservicio de Plano de Control de RAN. Por supuesto, tal microservicio único todavía contiene diferentes funcionalidades que podrían separarse además en microservicios más pequeños en el futuro.

Dependiendo del tipo de tarea de procesamiento, se pueden establecer diferentes requisitos para los microservicios relacionados (BBUs virtuales) que llevan a cabo el procesamiento para la función necesaria. El procesamiento puede ser un procesamiento de programador relacionado con usuario o celda, procesamiento de usuario extendido (xUP), o funciones de celda comunes que se ejemplifican y procesan desde el principio incluso si ningún usuario es manejado por la RAN. Por ejemplo, las funciones de programación deben llevarse a cabo y señalizarse dentro de una cantidad dada de tiempo (1ms en 4G, presumiblemente 125µs en 5G). Cuando un nuevo usuario ingresa al sistema con una Configuración de Portador de Radio de Datos, se permite que el recurso de procesamiento para el nuevo usuario que ingresa tome pocas docenas de milisegundos. El despliegue de nuevos microservicios para el manejo de próximas cargas, tales como futuras Configuraciones de Portadores de Radio de Datos, debe habilitarse dentro de unos pocos segundos. El problema ahora es encontrar y controlar recursos (capacidad de procesamiento y ancho de banda de red x) para los microservicios por sección de red, que permitan el despliegue y gestión de estos componentes virtualizados (microservicios) que puedan hacer frente a los diferentes requisitos para las tareas de procesamiento descritas anteriormente. De esa manera se debe dirigir especial atención hacia las situaciones de sobrecarga que pueden suceder en tal sistema.

Los sistemas de control en la nube del estado de la técnica pueden proporcionar el nivel más alto de la funcionalidad necesaria, pero una red de acceso por radio, especialmente en la 5G emergente, requerirá restricciones de temporización que no pueden ser cubiertas por las soluciones conocidas. Tradicionalmente los orquestadores están trabajando con cargas de CPU y no pueden hacer referencia a la Calidad de Experiencia, QoE, parámetros de usuarios individuales. Se conoce que las mediciones de CPU son propensas a errores, retrasos y a menudo la degradación de servicio puede no estar ampliamente correlacionada con la carga de la CPU que realiza el servicio. Por ejemplo, las mediciones de carga de CPU nunca pueden seguir en el procesamiento de tiempo que se produce sobre una base de ms o sub ms, que es el caso con los servicios de baja latencia.

Por lo tanto, se propone un método donde el control de las partes rápidas del sistema se separa del control de las partes más lentas del sistema de acceso por radio. El diagrama de flujo de la figura 2 ilustra una realización.

El gráfico de la figura 2 ilustra un ejemplo para operar una red de acceso por radio, donde la red de acceso por radio comprende dos o más niveles de recursos, tales como FEUs, Nubes de Borde y Nube Central, donde cada nivel tiene

diferentes requisitos de latencia/temporización. En una realización en la etapa 200, la asignación de recursos para la red de acceso por radio se realiza con dos o más bucles de control, donde cada bucle de control está configurado para asignar recursos con diferentes requisitos de latencia.

- 5 En la etapa 202, los recursos de un nivel más interno se controlan por un bucle de control más interno con asignación de recursos de baja latencia.

10 En la etapa 204, los recursos de niveles externos se controlan por uno o más bucles de control externos para una asignación de recursos críticos con menos latencia, teniendo cada bucle de control externo un requisito de latencia menos crítico en comparación con los bucles más cercanos al bucle más interno.

15 Adicionalmente, en una realización, los datos se intercambian entre los bucles de control que sirven a diferentes niveles. A saber si un bucle de control inferior (que sirve a un proceso de latencia más baja) no puede hacer frente a la situación actual envía una alarma al siguiente bucle de control superior, con un requisito de temporización menos restrictivo. Al utilizar este tipo de jerarquía se pueden evitar fácilmente decisiones contradictorias en cada nivel. Por ejemplo, si una protección de sobrecarga rápida realizada por el bucle de control más interno no remedia un nivel de servicio deteriorado se puede realizar una migración de carga dentro del siguiente bucle de control externo. Adicionalmente, si esto tampoco está mejorando la situación (por ejemplo todos los recursos informáticos disponibles están completamente ocupados), entonces se puede ejemplificar un microservicio adicional a través de un bucle de control externo adicional, por ejemplo.

20 En una realización, dependiendo del tipo de función de procesamiento que va a ser ejecutada, un microservicio asociado puede estar ya sea ubicado en una FEU o en una Nube de Borde o en ambas. Esto se ilustra en la figura 1 donde las vBBUs pueden ubicarse en ambos niveles. Las vBBUs en FEUs pueden procesar conexiones de latencia baja y heredada mientras que las VBBUs en Nubes de Bordes pueden procesar solo conexiones de latencia heredadas (más lentas).

25 En una realización, un bucle de control puede gestionar los recursos del siguiente bucle de control interno sobre la base de las indicaciones transmitidas por el siguiente bucle de control interno.

30 En una realización, cuando un bucle de control detecta que no puede cumplir con sus tareas de asignación de recursos relacionadas con una o más conexiones, puede transmitir una indicación a un bucle externo. De este modo, un bucle de control puede gestionar recursos de un bucle de control interno sobre la base de indicaciones transmitidas por el bucle de control interno.

35 En una realización, cuando detecta que los microservicios existentes no pueden proporcionar servicio a una o más conexiones de la red de acceso por radio, un bucle de control puede inicializar dinámicamente nuevos microservicios o más recursos a microservicios existentes sobre la base de la detección.

40 La figura 3 ilustra una realización de ejemplo, donde la arquitectura de gestión dinámica de recursos consiste principalmente en tres bucles de control, bucle interno (bucle de control 3) 300, bucle medio (bucle de control 2) 302 y bucle externo (bucle de control 1) 304, apilados uno encima del otro controlando los microservicios distribuidos sobre los recursos en la FEU y la EC. De esa manera ciertas funciones, tales como función PHYCell, solo están presentes en una FEU y se pueden colocar en aceleradores de hardware especiales, mientras que las otras funciones de BBU (microservicios) son todas adecuadas para ejecutarse en hardware de plataforma de propósito general (GPP).

45 Debe anotarse aquí, que el número de bucles de control no está limitado a tres. Dependiendo del sistema el número de diferentes niveles y bucles de control puede variar, como bien sabrá un experto en la técnica.

50 Un punto en el uso de múltiples bucles de control es también las interacciones entre diferentes bucles de control y disponibles para cada sección de red para diferentes tipos y para escenarios de multi-inquilinos: recursos de interacciones 306 de bucle externo entre el orquestador de RAN y BBSC para el despliegue de diversos microservicios, e interacción entre secciones de red, interacciones 310 de bucle medio entre Control de Sobrecarga basado en BBSC y TTI (Intervalo de Tiempo de Transmisión) para habilitar el equilibrio de carga de grupo de recursos y las interacciones 55 312 de bucle interno entre el Control de Sobrecarga basado en TTI y vBBUs para recortar las ráfagas de uso de recursos dentro de una configuración dada.

60 Como se mencionó anteriormente, cada bucle de control puede configurarse para actuar con diferentes requisitos de latencia: bucle 304 externo sobre la base de varios segundos, bucle 302 medio típicamente 0,5 segundos de temporización y bucle 300 interno con temporización menor que 1 ms.

65 Estos bucles de interfuncionamiento habilitan reacciones que son lo suficientemente rápidas para mantener estable la QoE mientras que permiten que la nube supervise los recursos mientras que habilita ganancias de multiplexación estadística garantizadas en un entorno de nube muy dinámico con respecto al uso de recursos.

A continuación, se estudia con más detalle un ejemplo de la operación de diferentes bucles de control.

El bucle 304 de control externo es responsable del despliegue/redespliegue dinámico de microservicios de acuerdo con las estadísticas de uso de recursos desde el BBSC activo en el bucle 302 de control medio. Esto incluye el manejo de solicitudes de recursos a través de secciones. Para esto el BBSC recopila y procesa mediciones de uso de recursos, tal como con qué frecuencia se necesitaba en el pasado una migración de usuarios desde un microservicio a otro recurso de procesamiento (microservicio). Si el BBSC detecta que no hay microservicios disponibles para procesar usuarios adicionales ya sea indica al orquestador de RAN que ejemplifique nuevos microservicios o que dé más recursos de procesamiento (núcleos) a un microservicio ya existente. Debe anotarse que el orquestador de RAN proporciona recursos de procesamiento de banda base para diferentes secciones de red, por ejemplo, para IoT, para servicios de MBB o para eMBB. Por supuesto un redespliegue de microservicios se ejecuta sobre la marcha durante el sistema en ejecución sin alterar los servicios en curso.

El bucle 302 de control medio con el BBSC como un elemento activo es responsable del Equilibrio de Carga Intra Sección entre microservicios (vBBU) mediante la Migración de Carga en caso de KPIs de QoE incorrectos dentro de una vBBU detectada por el bucle 300 de control interno. Esto se hace comunicando cuellos de botella, por ejemplo, en ancho de banda de red media o latencia de red media por el bucle de control interno al BBSC. El BBSC puede ser configurado para realizar migraciones para que uno o más usuarios quiten la carga desde un microservicio sobrecargado. Después de la migración se puede retirar la sobrecarga del microservicio en cuestión. La migración de carga en las bases de Portador de Radio de Datos (DRB) está configurada para que suceda sin interrupciones de servicio para los usuarios relacionados.

El bucle 300 de control interno, que puede estar integrado en el proceso de Programador consiste en el Control de Sobrecarga basado en TTI, TTI-OLC, para la capacidad de procesamiento y el uso de Ancho de Banda de Red Media. Este bucle constituye un tratamiento instantáneo de una situación de sobrecarga dentro del tiempo de Intervalo de Tiempo de Transmisión TTI (1ms o 0.125ms) y está configurado para garantizar un eficiente aprovisionamiento de recursos de un microservicio evitando el sobre aprovisionamiento.

En una realización, el BBSC señala al Orquestador de RAN (RANO) si se deben iniciar nuevos microservicios, o se deben eliminar microservicios existentes, por ejemplo, si el equilibrio de carga solo no remedia la situación. Sin embargo, el RANO puede tomar este consejo o no, dependiendo de su visión de los recursos o políticas de operador, etc. Además, el RANO puede realizar su propio control de recursos, tal vez con mediciones desde el BBSC. La Nube Central puede a su vez controlar el RANO en un bucle externo adicional (cuarto bucle) de acuerdo con las políticas de Nube Central. El RANO puede estar ya sea ubicado en la Nube de Borde (por ejemplo, en una situación de multiprovedores) o en la Nube Central.

Se da una implementación de ejemplo en la figura 4 para controlar los recursos de ancho de banda de red media y recursos de procesamiento de banda base (BB) asignados a un microservicio. En una realización, puede haber varios programadores 400 de equipo de usuario (UE) compitiendo en un módulo de control de sobrecarga por recursos. Los programadores de UE pueden estar ubicados en vBBUs ya sea en FEU o Nube de Borde de acuerdo con la ubicación de la parte en tiempo real de la vBBU. Sin embargo, en la arquitectura de programador hay solo un programador 440 de dominio de frecuencia (FD) y un único módulo 402 de control de sobrecarga. El programador 440 de FD puede estar ubicado en FEU y el módulo 402 de control de sobrecarga ya sea en FEU o en Nube de Borde de acuerdo con la ubicación del programador de UE.

Cada programador 400 de UE en la Nube de Borde o FEU está configurado para estimar el ancho de banda de red media y el esfuerzo de procesamiento de BB con concesiones 408 de parámetro de programador desde el programador 404 de FD. Estos Programadores de UE luego envían solicitudes 410 de recursos al módulo 402 de control de sobrecarga. En caso de que el control de sobrecarga detecte una situación de sobrecarga se configura para modificar los parámetros de programación u omitir usuarios para secciones de tiempo posteriores y retornar 412 los parámetros de programación nuevos o no cambiados de vuelta a cada Programador de UE por separado. Además el control de sobrecarga reporta los KPIs de QoE al BBSC para propósitos de migración de carga. Finalmente, cada programador de UE aplica 414 los nuevos parámetros de programación a la capa 406 en tiempo real de Capa 2 (L2 RT).

Como se mencionó el uso de tres bucles es simplemente una posible implementación de una jerarquía de bucles de control. Se pueden diseñar bucles de control adicionales por ejemplo para gestionar recursos de ancho de banda o recursos de CPU entre diferentes tecnologías de acceso como redes de acceso por radio 4G versus 5G. También se podrían usar menos bucles de control, por ejemplo, omitiendo el bucle de control externo descrito anteriormente (Orquestador de RAN).

Las soluciones existentes implementan una disposición fija de funciones de procesamiento independientemente de la carga y los requisitos de servicio. En las soluciones existentes la red de acceso por radio debe estar equipada para el tráfico máximo esperado. De esa manera la composición de funciones se fija entre el acceso por radio y los recursos en la nube de red y en caso de una sobrecarga no se puede hacer mucho para remediar la situación. La solución propuesta tiene muchas ventajas sobre las soluciones existentes.

La solución propuesta soporta la creación y la gestión de secciones de red. El bucle más interno puede manejar los requisitos de servicio basados en TTI y de este modo habilita una latencia ultrabaja en las futuras redes 5G. Se pueden realizar considerables ganancias de grupo con menos equipo, y las situaciones de sobrecarga se remedian evitando el sobre aprovisionamiento de recursos. La colocación de procesamiento se puede hacer de manera diferente para el usuario y el plano de control y también de acuerdo con necesidades de operador. La descomposición flexible de funciones de red móvil entre el acceso por radio y las infraestructuras en la nube de red permite soportar arquitecturas totalmente distribuidas, parcialmente distribuidas, y totalmente centralizadas. Se pueden aplicar fácilmente nuevos servicios y características de red. Por ejemplo, la adición de microservicios de IoT para nuevos esquemas de acceso a IoT se puede hacer de acuerdo con la sensibilidad de latencia del servicio requerido.

La figura 5 ilustra un ejemplo de cómo el bucle de control interno detecta un problema que no puede resolver y solicita recursos al siguiente bucle externo. El ejemplo se relaciona con la migración de carga para un Portador de Radio de Datos de banda Ancha Móvil, DRB de MBB. El ejemplo se refiere al bucle 500 de control interno, donde el elemento activo es TTI_OLC 502 y el siguiente bucle 504 de control externo donde el elemento activo es BBSC 506. Aquí se asume que un microservicio de MBB está activo en vBBU1 508 con tráfico de vídeo en un Portador de Radio de Datos, DRB 514. Suponer además que aumenta la carga del microservicio que tiene el DRB.

El TTI_OLC 502 dentro del bucle 500 de control más interno (bucle de control 3) está configurado para detectar la sobrecarga y señalar una necesidad de migración al BBSC 506 en el siguiente bucle 504 de control externo (bucle de control 2).

El BBSC puede ser configurado para seleccionar un microservicio de MBB de destino en vBBU1 510 para migrar 512 el procesamiento de DRB 514. El BBSC transmite una solicitud a vBBU1 510 para la instanciación del DRB 514 en la vBBU1 y una solicitud a vBBU1 para la eliminación del DRB 514 en la vBBU1.

La figura 6 ilustra una realización. La figura ilustra un ejemplo simplificado de un aparato que aplica realizaciones de la invención. En algunas realizaciones, el aparato puede ser un elemento de red de una Red de Acceso por Radio en la Nube.

Debe entenderse que el aparato se representa en este documento como un ejemplo que ilustra algunas realizaciones. Es evidente para una persona experta en la técnica que el aparato también puede comprender otras funciones y/o estructuras y no se requieren todas las funciones y estructuras descritas. Aunque el aparato se ha representado como una entidad, se pueden implementar diferentes módulos y memoria en una o más entidades físicas o lógicas.

El aparato 600 del ejemplo incluye una circuitería 602 de control configurada para controlar al menos parte de la operación del aparato.

El aparato puede comprender una memoria 604 para almacenar datos. Adicionalmente la memoria puede almacenar software 606 ejecutable por la circuitería 602 de control. La memoria puede estar integrada en la circuitería de control.

El aparato comprende además una o más circuiterías 608, 610 de interfaz configuradas para conectar el aparato a otros dispositivos y elementos de red del sistema. La interfaz puede proporcionar una conexión por cable o inalámbrica.

En algunas realizaciones, el aparato puede comprender además una circuitería 612 de acelerador diseñada específicamente para realizar tareas dadas en relación con la circuitería de control.

En una realización, el software 606 puede comprender un programa de ordenador que comprende medios de código de programa adaptados para hacer que la circuitería 602 de control del aparato realice diferentes microservicios de la Red de Acceso por Radio en la Nube.

La Red de Acceso por Radio en la Nube puede comprender varios aparatos de la figura 6, y los aparatos pueden estar conectados entre sí y realizar diversas funciones de la Red de Acceso.

Las etapas y funciones relacionadas descritas en las figuras anteriores y adjuntas no están en orden cronológico absoluto, y algunas de las etapas pueden realizarse simultáneamente o en un orden diferente al dado. También se pueden ejecutar otras funciones entre las etapas o dentro de las etapas. Algunas de las etapas también se pueden excluir o reemplazar con una etapa correspondiente.

Los aparatos o controladores capaces de realizar las etapas descritas anteriormente pueden implementarse como un ordenador digital electrónico, sistema de procesamiento o una circuitería que puede comprender una memoria de trabajo (memoria de acceso aleatorio, RAM), una unidad central de procesamiento (CPU), y un reloj de sistema. La CPU puede comprender un conjunto de registros, una unidad aritmética lógica, y un controlador. El sistema de procesamiento, controlador o la circuitería están controlados por una secuencia de instrucciones de programa transferidas a la CPU desde la RAM. El controlador puede contener un número de microinstrucciones para operaciones básicas. La implementación de microinstrucciones puede variar dependiendo del diseño de CPU. Las instrucciones de programa pueden estar codificadas por un lenguaje de programación, que puede ser un lenguaje de programación

de alto nivel, tal como C, Java, etc., o un lenguaje de programación de bajo nivel, tal como un lenguaje de máquina, o un ensamblador. El ordenador digital electrónico también puede tener un sistema operativo, que puede proporcionar servicios de sistema a un programa de ordenador escrito con las instrucciones de programa.

5 Como se usa en esta solicitud, el término 'circuitaría' se refiere a todo de lo siguiente: (a) implementaciones de circuitos solo de hardware, tales como implementaciones en circuitaría solo analógica y/o digital, y (b) combinaciones de circuitos y software (y/o firmware), tales como (según sea aplicable): (i) una combinación de procesadores o (ii) porciones de procesadores/software incluyendo procesadores de señales digitales, software, y memorias que trabajan
10 en conjunto para hacer que un aparato realice diversas funciones, y (c) circuitos, tales como unos microprocesadores o una porción de unos microprocesadores, que requieren software o firmware para operación, incluso si el software o firmware no está físicamente presente.

Esta definición de 'circuitaría' se aplica a todos los usos de este término en esta solicitud. Como ejemplo adicional, como se usa en esta solicitud, el término 'circuitaría' también cubriría una implementación de simplemente un procesador (o múltiples procesadores) o una porción de un procesador y su (o sus) software y/o firmware
15 acompañantes. El término 'circuitaría' también cubriría, por ejemplo y si fuera aplicable al elemento particular, un circuito integrado de banda base o circuito integrado de procesador de aplicaciones para un teléfono móvil o un circuito integrado similar en un servidor, un dispositivo de red celular, u otro dispositivo de red.

20 Una realización proporciona un programa de ordenador incorporado en un medio de distribución, que comprende instrucciones de programa que, cuando se cargan en un aparato electrónico, se configuran para controlar el aparato para ejecutar las realizaciones descritas anteriormente.

El programa de ordenador puede estar en forma de código fuente, forma de código objeto, o en alguna forma
25 intermedia, y puede almacenarse en algún tipo de portador, que puede ser cualquier entidad o dispositivo capaz de portar el programa. Tales portadores incluyen un medio de registro, memoria de ordenador, memoria de solo lectura, y un paquete de distribución de software, por ejemplo. Dependiendo de la potencia de procesamiento necesaria, el programa de ordenador puede ejecutarse en un único ordenador digital electrónico o puede distribuirse entre un número de ordenadores.

30 El aparato también puede implementarse como uno o más circuitos integrados, tales como circuitos integrados de aplicación específica ASIC. También son factibles otras realizaciones de hardware, tal como un circuito construido con componentes lógicos separados. También es factible un híbrido de estas diferentes implementaciones. Cuando se selecciona el método de implementación, una persona experta en la técnica considerará los requisitos establecidos
35 para el tamaño y consumo de potencia del aparato, la capacidad de procesamiento necesaria, costes de producción, y volúmenes de producción, por ejemplo.

Será obvio para una persona experta en la técnica que, a medida que avanza la tecnología, el concepto inventivo puede implementarse de diversas formas. La invención y sus realizaciones no están limitadas a los ejemplos descritos
40 anteriormente sino que pueden variar dentro del alcance de las reivindicaciones.

REIVINDICACIONES

1. Un método para operar una red de acceso por radio, donde la red de acceso por radio comprende dos o más niveles de recursos, teniendo cada nivel de recursos diferentes requisitos de latencia/temporización, siendo las funcionalidades de la red de acceso por radio realizadas utilizando secciones de red y microservicios, comprendiendo el método:

realizar asignación de recursos para la red de acceso por radio con tres o más bucles (300, 302, 304) de control, donde cada bucle de control de los tres o más bucles de control está configurado para asignar recursos con diferentes requisitos de latencia, y configurado para controlar recursos entre los niveles de recursos,

controlar recursos de un nivel de recursos más interno de los tres o más niveles de recursos por un bucle (300) de control más interno de los tres o más bucles de control que tienen la asignación de recursos de latencia más baja de los tres o más bucles de control, y

controlar recursos de niveles de recursos externos de los tres o más niveles de recursos por uno o más bucles (304) de control externos de los tres o más bucles de control para asignación de recursos críticos con menos latencia, teniendo cada uno de los uno o más bucles de control externos de los tres o más bucles de control un requisito de latencia menos crítico en comparación con los bucles de control más cercanos al bucle más interno, en donde

los uno o más bucles de control externos son responsables del despliegue/redespliegue dinámico de microservicios de acuerdo con las estadísticas de uso de recursos desde un bucle (302) de control medio de los tres o más bucles de control, estando el bucle de control medio más cercano al bucle de control más interno que el uno o más bucles de control externos,

el bucle de control medio es responsable del equilibrio de carga intra sección entre microservicios de acuerdo con el ancho de banda y latencia detectados por el bucle de control más interno y

el bucle más interno programa recursos dentro de un Intervalo de Tiempo de Transmisión usado en el nivel más interno.

2. El método de la reivindicación 1, que comprende además:

detectar por un bucle de control de los tres o más bucles de control que no puede cumplir con sus tareas de asignación de recursos

transmitir por el bucle de control una indicación a un bucle de control externo de los tres o más bucles de control.

3. El método de la reivindicación 2, en donde un bucle de control de los tres o más bucles de control gestiona recursos de un bucle de control interno de los tres o más bucles de control sobre la base de indicaciones transmitidas por el bucle de control interno.

4. El método de cualquier reivindicación precedente, en donde cada sección de red está dedicada a un tipo dado de servicios que tienen características de red similares.

5. El método de la reivindicación 4, que comprende además: realizar funcionalidades de la red de acceso por radio dentro de cada sección de red como microservicios.

6. El método de la reivindicación 5, en donde las funcionalidades de la red de acceso por radio comprenden una o más conexiones de la red de acceso por radio, comprendiendo además el método:

detectar que los microservicios existentes no pueden proporcionar servicio a la una o más conexiones de la red de acceso por radio;

inicializar dinámicamente por un bucle de control de los tres o más bucles de control nuevos microservicios o más recursos a microservicios existentes sobre la base de la detección.

7. El método de la reivindicación 5, que comprende además:

gestionar solicitudes de recursos a través de secciones de red.

8. El método de la reivindicación 1, en donde cuando el bucle de control más interno detecta sobrecarga, transmite información sobre la sobrecarga a un bucle de control externo.

9. El método de la reivindicación 8, en donde el bucle de control externo recibe información sobre la sobrecarga desde el bucle de control más interno, controla el uso de recursos asignado al bucle de control más interno.

10. El método de la reivindicación 5, que comprende además: asignar microservicios a diferentes niveles con base en los requisitos de latencia de las conexiones servidas por los microservicios.

5 11. Un aparato en una red de acceso por radio donde la red de acceso por radio comprende tres o más niveles de recursos, teniendo cada nivel de recursos diferentes requisitos de latencia/temporización, siendo funcionalidades de la red de acceso por radio realizadas utilizando secciones de red y microservicios, comprendiendo el aparato:

al menos un procesador; y

10 al menos una memoria incluyendo código de programa de ordenador,

la al menos una memoria y el código de programa de ordenador configurados para, con el al menos un procesador, hacer que el aparato al menos:

15 realice asignación de recursos para la red de acceso por radio con tres o más bucles (300, 302, 304,) de control

donde cada bucle de control de los tres o más bucles de control está configurado para asignar recursos con diferentes requisitos de latencia, y configurado para controlar recursos entre los niveles de recursos,

20 controle recursos del nivel más interno de los tres o más niveles de recursos por un bucle (300) de control más interno de los tres o más bucles de control que tienen la asignación de recursos de latencia más baja de los tres o más bucles de control, y

25 controle recursos de los niveles de recursos externos de los tres o más niveles de recursos por uno o más bucles (304) de control externos de los tres o más bucles de control para asignación de recursos críticos con menos latencia, teniendo cada uno de los uno o más bucles de control externos de los tres o más bucles de control un requisito de latencia menos crítico en comparación con los bucles de control más cercanos al bucle más interno, en donde

30 el uno o más bucles de control externos son responsables del despliegue/redespliegue dinámico de microservicios de acuerdo con las estadísticas de uso de recursos desde un control (302) medio de los tres o más bucles de control, estando el bucle de control medio más cercano al bucle de control más interno que el uno o más bucles de control externos,

35 el bucle de control medio es responsable del equilibrio de carga intra sección entre microservicios de acuerdo con el ancho de banda y latencia detectados por el bucle de control más interno y

el bucle más interno programa recursos dentro de un Intervalo de Tiempo de Transmisión usado en el nivel más interno.

40 12. El aparato de la reivindicación 11, la al menos una memoria y el código de programa de ordenador configurados para, con el al menos un procesador, hacer que el aparato además

45 detecte por un bucle de control de los tres o más bucles de control que no puede cumplir con sus tareas de asignación de recursos,

transmita por el bucle de control una indicación a un bucle de control externo de los tres o más bucles de control.

50 13. El aparato de la reivindicación 11, la al menos una memoria y el código de programa de ordenador configurados para, con el al menos un procesador, hacer que el aparato además

gestione recursos de un bucle de control interno con base en indicaciones transmitidas por el bucle de control interno.

55 14. El aparato de cualquier reivindicación precedente 11 - 13, la al menos una memoria y el código de programa de ordenador configurados para, con el al menos un procesador, hacer que el aparato además: dedique cada sección de red a un tipo dado de servicios que tienen características de red similares.

60 15. El aparato de la reivindicación 14, la al menos una memoria y el código de programa de ordenador configurados para, con el al menos un procesador, hacer que el aparato además: realice funcionalidades de la red de acceso por radio como microservicios.

16. El aparato de la reivindicación 15, en donde las funcionalidades de la red de acceso por radio comprenden una o más conexiones de la red de acceso por radio, la al menos una memoria y el código de programa de ordenador configurados para, con el al menos un procesador, hacer que el aparato además:

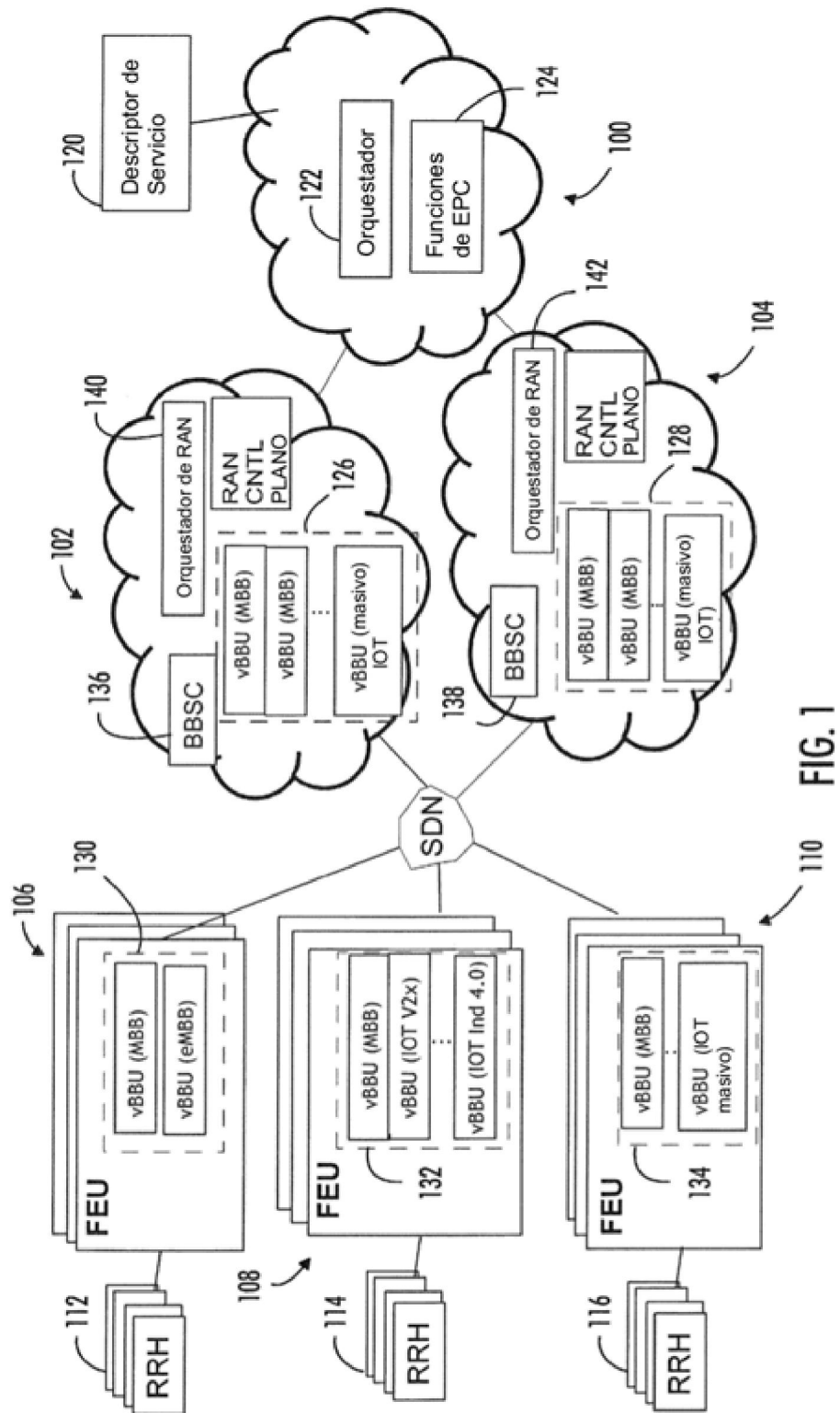
65

detecte que los microservicios existentes no pueden proporcionar servicio a la una o más conexiones de la red de acceso por radio;

5 inicialice dinámicamente por un bucle de control de los tres o más bucles de control nuevos microservicios o más recursos a microservicios existentes con base en la detección.

10 17. El aparato de la reivindicación 11, la al menos una memoria y el código de programa de ordenador configurados para, con el al menos un procesador, hacer que el aparato además: asigne microservicios a diferentes niveles de recursos con base en los requisitos de latencia de las conexiones servidas por los microservicios.

18. Un producto de programa de ordenador incorporado en un medio de distribución legible por un ordenador y que comprende instrucciones de programa que, cuando se cargan en un aparato, ejecutan el método de acuerdo con cualquiera de las reivindicaciones 1 a 10.



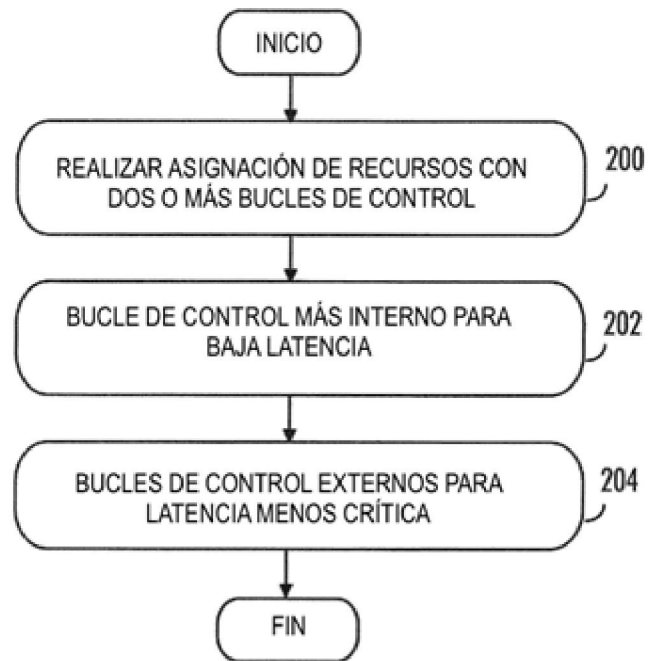


FIG. 2

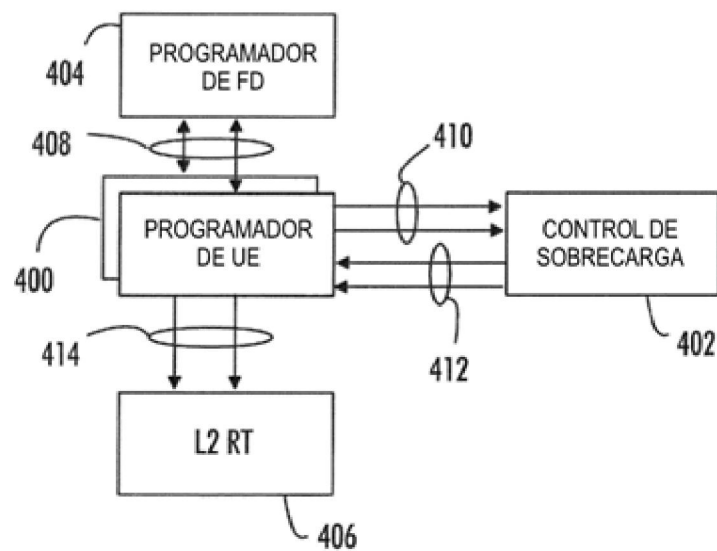


FIG. 4

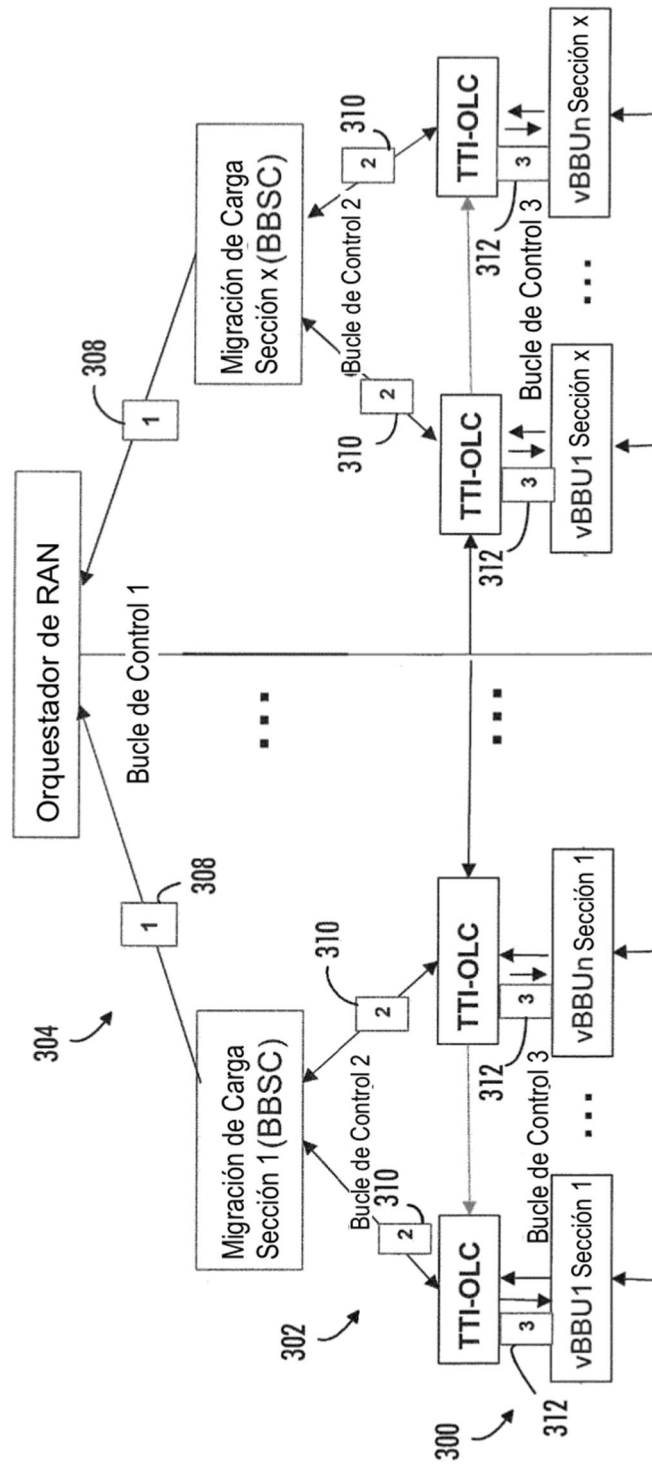


FIG. 3

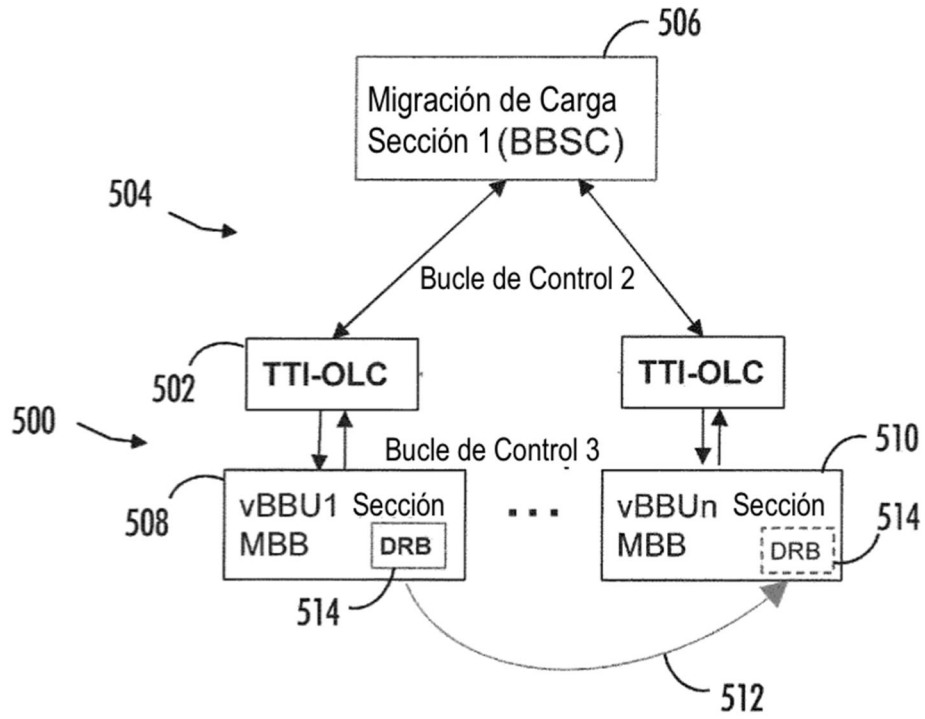


FIG. 5

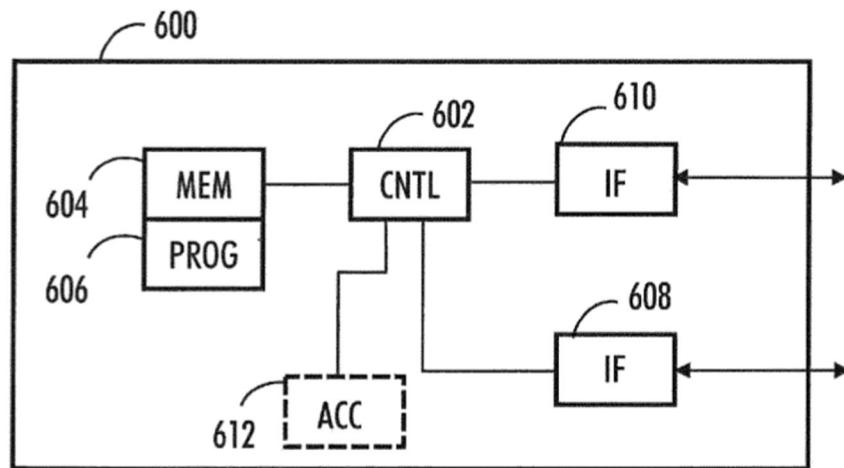


FIG. 6