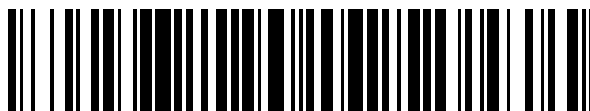


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 385 124**

51 Int. Cl.:
H04L 12/56 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

- 96 Número de solicitud europea: **08832727 .5**
- 96 Fecha de presentación: **17.09.2008**
- 97 Número de publicación de la solicitud: **2191619**
- 97 Fecha de publicación de la solicitud: **02.06.2010**

54 Título: **Planificación de prioridades y control de admisión en una red de comunicación**

30 Prioridad:
17.09.2007 US 973137 P
16.09.2008 US 211718

45 Fecha de publicación de la mención BOPI:
18.07.2012

45 Fecha de la publicación del folleto de la patente:
18.07.2012

73 Titular/es:
QUALCOMM INCORPORATED
ATTN: INTERNATIONAL IP ADMINISTRATION
5775 MOREHOUSE DRIVE
SAN DIEGO, CA 92121, US

72 Inventor/es:
GOGIC, Aleksandar

74 Agente/Representante:
Fàbrega Sabaté, Xavier

ES 2 385 124 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

Planificación de prioridades y control de admisión en una red de comunicación.

5 ANTECEDENTES

Campo

10 La presente divulgación se refiere en general a comunicación, y más específicamente a técnicas para planificar la transmisión de datos y controlar la admisión en una red de comunicación.

Antecedentes

15 Una red de comunicación puede observar grandes fluctuaciones en la carga debido a diferentes motivos. Cuando la carga en la red es alta, puede ser deseable para la red de comunicación servir a ciertos usuarios antes que a otros usuarios. Por ejemplo, cuando ocurre un desastre natural o debido a la mano del hombre, la red de comunicación en el área afectada puede verse presionada por una carga de tráfico excesiva y a veces por fallos de la infraestructura de red provocados por el mismo desastre. Puede ser deseable para la red de comunicación ofrecer personal de asistencia de emergencia, tal como policía o bomberos, por delante del público general. Existe por lo tanto una necesidad en la técnica de técnicas que sirven a diferentes usuarios bajo una carga fuerte de red. Un ejemplo del estado de la técnica se puede encontrar en US2002/0031086.

RESUMEN

25 Esta necesidad es cubierta por el objeto de las reivindicaciones independientes. Las técnicas para llevar a cabo planificación de prioridades y control de admisión en una red de comunicación se describen en este documento. En un aspecto, los flujos de datos pueden ser priorizados y paquetes para distintos flujos de datos pueden colocarse en puntos diferentes en una cola dependiendo de las prioridades de los flujos de datos. En general, se puede soportar cualquier número de niveles de prioridad. En un diseño, paquetes con niveles de prioridad progresivamente mayores pueden colocarse en puntos progresivamente más cerca de la cabeza de la cola y pueden entonces experimentar retrasos de cola progresivamente más cortos. A cada flujo de datos se le puede asignar un nivel de prioridad, y se pueden colocar paquetes para el flujo de datos en un punto en la cola determinado en base al nivel de prioridad de aquel flujo de datos. A un usuario se le puede asignar un nivel de prioridad particular, y todos los flujos de datos que pertenezcan a ese usuario (originarios o que finalicen en el dispositivo de usuario) pueden adherirse al nivel de prioridad del usuario.

35 En otro aspecto, un paquete para un terminal puede transferirse desde una celda fuente a una celda objetivo debido a un traspaso y puede atribuírsele la cantidad de tiempo que el paquete ya ha estado esperando en una cola en la celda de fuente. El paquete puede colocarse en un punto avanzado en una cola en la celda objetivo. El punto puede determinarse en base a la cantidad de tiempo que el paquete ya ha estado esperando. Al no colocar el paquete al final de la cola de la celda objetivo, se puede evitar un retraso de cola excesivo para el paquete.

45 En otro aspecto más, se puede llevar a cabo el control de admisión de forma que reduzca los efectos adversos en flujos de datos prioritarios. En un diseño, la carga de una celda puede determinarse en base a al menos un criterio, por ejemplo, retraso de cola medio de los paquetes a enviar. La carga de celda puede ser considerada como ligera si el retraso medio de cola es menor que un primer umbral, fuerte si el retraso medio de cola es mayor que un segundo umbral o moderado si el retraso medio de cola está entre los primero y segundo umbrales. Todos los flujos de datos prioritarios y los flujos de datos no prioritarios pueden ser admitidos si la carga de celda es ligera. Solo los flujos de datos prioritarios pueden ser admitidos si la carga de celda es fuerte. Todos los flujos de datos prioritarios y ciertos flujos de datos no prioritarios pueden ser admitidos si la carga de celda es moderada. Algunos recursos radio de la celda pueden reservarse en el caso de que terminales con flujos de datos prioritarios en celdas vecinas sean traspasados a la celda.

50 Varios aspectos y características de la divulgación se describen con más detalles a continuación.

Breve descripción de los dibujos

55 La Figura 1 muestra una red de comunicación inalámbrica.
 La Figura 2 muestra un mecanismo de cola a título de ejemplo para cinco clases de tráfico.
 La Figura 3 muestra un diseño de planificación de prioridades con dos niveles de prioridad.
 La Figura 4 muestra un diseño de planificación de prioridades con N niveles de prioridad.
 60 La Figura 5 muestra un proceso para enviar datos con planificación de prioridades.
 La Figura 6 muestra el enrutado y transmisión de un paquete a un terminal sin traspaso.
 La Figura 7 muestra el enrutado y transmisión de un paquete a un terminal con traspaso.
 La Figura 8 muestra un proceso para enviar datos teniendo en cuenta el tiempo previo de encolado.
 La Figura 9 muestra un proceso de control de admisión.
 65 La Figura 10 muestra un diagrama de bloques de un terminal, una estación base y una entidad de red.

Descripción detallada

Las técnicas descritas en este documento pueden ser utilizadas para diferentes redes de comunicación inalámbricas y fijas. Los términos “red” y “sistema” se usan a menudo de forma intercambiable. Por ejemplo, las técnicas pueden usarse para redes de comunicación inalámbrica tales como redes de Acceso Múltiple por División de Código (CDMA), redes de Acceso Múltiple por División de Tiempo (TDMA), redes de Acceso Múltiple por División de Frecuencia (FDMA), redes FDMA Ortogonales (OFDMS), redes de portador única FDMA (SC-FDMA), etcétera. Una red CDMA puede implementar una tecnología radio tal como cdma2000, Acceso Radio Universal Terrestre (UTRA), etcétera. cdma2000 cubre los estándares IS-2000, IS-95 e IS-856. UTRA incluye CDMA de Banda Ancha (WCDMA) y otras variantes de CDMA. Una red TDMA puede implementar una tecnología radio tal como Sistema Global para Comunicaciones Móviles (GSM). Una red OFDMA puede implementar una tecnología radio tal como Banda Ancha Ultra Móvil (UMB), UTRA Evolucionado (E-UTRA), IEEE 802.16 (WiMAX), IEEE 802.20, Flash-OFDM®, etcétera. UTRA y E-UTRA son parte del Sistema de Telecomunicación Universal Móvil (UMTS). Long Term Evolution (LTE) es una versión próxima de UMTS que utiliza E-UTRA. UTRA, E-UTRA, UMTS, LTE y GSM se describen en documentos de una organización llamada “3rd Generation Partnership Project” (3GPP). cdma2000 y UMB se describen en documentos de una organización llamada “3rd Generation Partnership Project 2” (3GPP2). Por claridad, ciertos aspectos de las técnicas se describen a continuación para una red de comunicación inalámbrica.

La Figura 1 muestra una red de comunicación inalámbrica 100, que puede incluir cualquier número de estaciones base 120. Una estación base es generalmente una estación fija que se comunica con terminales y puede también denominarse Nodo B, un Nodo B evolucionado, un punto de acceso, una estación transceptora base (BTS), etcétera. Cada estación base proporciona cobertura de comunicación para un área geográfica particular. El área de cobertura de una estación base puede dividirse en múltiples (por ejemplo tres) áreas más pequeñas. Cada área más pequeña puede ser servida por un subsistema respectivo de estación base. En 3GPP, el término “celda” puede referirse al área más pequeña de cobertura de una estación base y/o subsistema de estación base que sirve a esta área, dependiendo del contexto en el que se utilice el término. En 3GPP2, el término “sector” puede referirse al área de cobertura más pequeña de una estación base y/o subsistema de estación base que sirve a esta área. Por claridad, el concepto de celda en 3GPP se usa en la descripción a continuación.

Un controlador de red 122 puede acoplarse a un conjunto de estaciones base y proporcionar coordinación y control para estas estaciones base. Una pasarela de Protocolo de Internet (IP) 124 puede soportar servicios de datos para terminales y puede ser responsable de establecer, mantener y finalizar sesiones de datos para los terminales. La pasarela IP 124 puede acoplarse a otras redes de datos tales como una red central, redes de datos públicas o privadas, Internet, etcétera. La red 100 puede incluir otras entidades de red no mostradas en la Figura 1.

Los terminales 110 pueden estar dispersos a lo largo de la red, y cada terminal puede ser estacionario o móvil. Un terminal puede también denominarse estación móvil, equipo de usuario, terminal de acceso, unidad de suscriptor, estación, etcétera. Un terminal puede ser un teléfono móvil, un asistente personal digital (PDA), un dispositivo de comunicación inalámbrica, un módem inalámbrico, un dispositivo portátil, un ordenador portátil, etcétera. Un terminal puede comunicarse con la estación base a través de los enlaces delantero y reverso. El enlace delantero (o descendente) se refiere al enlace de comunicación desde la estación base al terminal, y el enlace reverso (o ascendente) se refiere al enlace de comunicación desde el terminal a la estación base. En la descripción en este documento, los términos “terminal” y “usuario” se usan de forma intercambiable. Los términos “estación base” y “celda” se usan también de forma intercambiable.

La red 100 puede utilizar un mecanismo de encolado para priorizar el tráfico de datos y soportar diferentes niveles de calidad de servicio (QoS). Un número de clases de tráfico (TC) puede definirse para diferentes categorías de servicios. Las clases de tráfico también pueden denominarse clases QoS, clases de flujo, categorías de tráfico, categorías de servicio, etcétera. Cada clase de tráfico puede estar asociada con ciertas garantías de QoS tales como un retraso máximo particular para enviar datos. Este retraso máximo también denominarse requisito de retraso, tolerancia de retraso, delimitación de retraso, límite de retraso, retraso máximo permitido, etcétera. El término “requisito de retraso” se usa en gran parte de la descripción a continuación. En general, puede definirse cualquier número de clases de tráfico. Se puede utilizar una cola para almacenar los datos para cada clase de tráfico.

Pueden existir uno o más flujos de datos en un enlace de comunicación entre un terminal y una celda. Un flujo de datos es un flujo de datos entre dos puntos finales específicos. Un flujo de datos también puede denominarse flujo IP, flujo de Control de Enlace Radio (RLC), flujo de Protocolo de Enlace Radio (RLP), etcétera. Un flujo de datos puede estar activo desde el principio hasta el final de la sesión. Por ejemplo, un flujo de datos de música en línea puede estar activo desde el momento en el que un usuario accede a un servidor web-cast hasta que el usuario apaga el reproductor de música en su ordenador. A un flujo de datos se le pueden asignar atributos QoS en el momento de la invocación del servicio o inicio de sesión. Estos atributos QoS pueden incluir descriptores de servicio tales como requisito de retraso, tasa de error de paquete permisible, tasa de datos requerida o esperada, etcétera. Un flujo de datos puede ser para un servicio en particular, y los atributos QoS pueden determinarse en base a los requisitos de datos del servicio. Un flujo de datos puede mapearse a una clase de tráfico específica en base a los atributos de QoS de ese flujo de datos y las garantías QoS de las diferentes clases de tráfico. Los paquetes del flujo de datos pueden almacenarse en la cola para la clase de tráfico a la que pertenece el flujo de datos.

La Figura 2 muestra un mecanismo de encolado a título de ejemplo 200 para cinco clases de tráfico 1 a 5, que se etiquetan TC₁ a TC₅, respectivamente. En el ejemplo que se muestra en la Figura 2, las cinco clases de tráfico tienen requisitos de retraso progresivamente más largos. La clase de tráfico 1 tiene el requisito de retraso más corto de D₁ y puede usarse para Voz sobre IP (VoIP), teleconferencias y otros servicios de tiempo real. Las clases de tráfico 2, 3 y 4 tienen requisitos de retraso progresivamente más largos de D₂, D₃ y D₄ respectivamente. La clase de tráfico 5 tiene el requisito de retraso más largo de D₅ y puede usarse para flujos "lo mejor posible" tales como Protocolo de Transferencia de Archivos (FTP). Los requisitos de retraso de las cinco clases de tráfico pueden tener cualquier valor adecuado y no están dibujadas a escala en la Figura 2. Por ejemplo, el requisito de retraso D₁ de la clase de tráfico 1 puede ser de 50 milisegundos (ms) o menos, el requisito de retraso D₂ de la clase de tráfico 2 puede ser de varios cientos de milisegundos (ms), etcétera.

La Figura 2 muestra un ejemplo de cinco colas para las cinco clases de tráfico en una entidad, la cual puede ser una celda o una pasarela IP. Una celda puede mantener las colas para almacenar datos para enviar en el enlace delantero a diferentes terminales. De forma separada, una celda puede mantener las colas para almacenar datos recibidos de varios terminales por el enlace reverso y puede enviar los datos almacenados a una pasarela IP. Por claridad, la mayoría de la siguiente descripción es para transmisión en el enlace delantero y cualquier desviación para el enlace reverso se analiza de forma separada.

Una celda puede recibir paquetes (por ejemplo, de la pasarela IP 124 en la Figura 1) para transmisión a los diferentes terminales. Los paquetes también pueden denominarse paquetes IP, datagramas, tramas, etcétera. Cada paquete puede estar destinado para un terminal recipiente específico y los paquetes para cada terminal se muestran como cajas con diferentes rellenos en la Figura 2. Los paquetes pueden tener tamaños diferentes, como se ilustra mediante diferentes tamaños de las cajas. La cantidad de tiempo para transmitir cada paquete puede depender de la tasa de datos de transmisión, la cual puede depender de la cantidad de recursos radio asignados al terminal recipiente, las condiciones de canal observadas por el terminal, etcétera.

La Figura 2 muestra una captura de las cinco colas en un instante de tiempo específico. En la Figura 2, el eje horizontal representa el tiempo, y llegan paquetes de entrada desde el lado izquierdo de la Figura 2. Las cinco colas para las cinco clases de tráfico se representan mediante cinco filas horizontales de cajas 210a a 210e para paquetes almacenados en estas cinco colas. La cabeza de cada cola es la caja más a la derecha para esa cola, y el final de cada cola es la caja más a la izquierda para esa cola.

Cada paquete puede pertenecer a un flujo de datos en particular y, cuando se recibe, puede ser colocado al final de la cola para la clase de tráfico a la que le flujo de datos pertenece. Cada cola puede almacenar paquetes para terminales diferentes en el orden en el que se reciben los paquetes. Cada paquete puede moverse desde el final de la cola hacia la cabeza de la cola al irse transmitiendo los paquetes en la cola.

En la Figura 2 una línea vertical fuerte 220 puede representar el instante límite de transmisión para cada paquete en las cinco colas. Las líneas verticales punteadas 222a a 222e pueden representar los puntos en los que se colocan los paquetes que llegan en las cinco colas y pueden estar dibujados a las distancias D₁ a D₅ respectivamente, desde la línea vertical fuerte 220. La distancia desde cada línea vertical punteada 222 a la línea vertical fuerte 220 es determinada por el requisito de retraso de la clase de tráfico asociada. La clase de tráfico 5 puede no tener ningún requisito de retraso, en cuyo caso la línea vertical punteada 222e puede no estar presente.

Cuando se recibe un paquete en la celda, puede ser clasificado y colocado en la cola apropiada en la línea vertical punteada 222 para esa cola. Con el paso del tiempo, los paquetes se mueven de izquierda a derecha en la Figura 2 y se acercan a su instante límite de transmisión en la línea vertical fuerte 220. La distancia desde el extremo puntero/derecho de cada caja a la línea vertical fuerte 220 es la cantidad de tiempo pasado en la cola. A título de ejemplo, cuando un paquete 212 llega a la celda, es clasificado y colocado en la cola para la clase de tráfico 3 en la línea vertical punteada 222c (no mostrada en la Figura 2). Al pasar el tiempo esperando a ser transmitido, el paquete 212 se mueve hacia su instante límite de transmisión en la línea vertical fuerte 220. Un pequeño instante de tiempo después, otro paquete 214 para otro terminal llega a la celda, es clasificado en la misma cola para la clase de tráfico 3, y se coloca de la misma forma en la línea vertical punteada 222c tras el paquete 212.

Los paquetes para cada clase de tráfico pueden transmitirse de forma primero en llegar, primero en salir (FIFO). En la Figura 2 los paquetes en la cola para cada clase de tráfico están numerados secuencialmente empezando por 1 para el paquete en la cabeza de la cola. Para cada cola, el número en cada caja indica el orden en el que llegaron los paquetes. Los paquetes en cada cola pueden ser transmitidos en el orden en el que se recibieron, empezando por el paquete 1, siguiendo por el paquete 2, etcétera. Cada paquete puede ser transmitido al o antes de alcanzar la línea vertical fuerte 220 para cumplir el instante límite de transmisión para ese paquete.

Los paquetes en las cinco colas pueden ser transmitidos de forma que los requisitos de retraso de estos paquetes se puedan cumplir. Un posible orden de transmisión puede ser el siguiente: TC₁(1), TC₅(1), TC₂(1), TC₁(2), TC₄(1), TC₃(1), TC₁(3), TC₂(2), TC₁(4), TC₃(2), TC₅(2), TC₂(3), TC₁(5), etcétera, en donde TC_x(m) denota el paquete m para

la clase de tráfico k. Los paquetes también pueden ser transmitidos en diferentes órdenes.

Si la red está ligeramente cargada, entonces los paquetes pueden ser transmitidos poco después de llegar a la celda. Por ejemplo, el paquete 212 puede ser transmitido en el momento en el que llega el paquete 214. Por lo tanto, el número de paquetes esperando en las colas puede ser bajo, y la mayoría del espacio entre los paquetes que llegaron antes (o más viejos) y el instante límite de transmisión en la línea vertical fuerte 220 puede estar vacío.

Al congestionarse la red, los retrasos de los paquetes se incrementan, y el espacio entre la línea vertical fuerte 220 y las líneas verticales 222A a 222E puede rellenarse. Un planificador puede intentar mantener los retrasos de los paquetes dentro de sus requisitos de retraso y puede intentar planificar cada paquete para transmisión antes de que el paquete supere su instante límite de transmisión en la línea vertical fuerte 220. El planificador puede seleccionar los paquetes para transmisión de forma tal que los paquetes en las cinco clases de tráfico se acercan a sus requisitos de retraso de forma aproximadamente simultánea.

A título de ejemplo, el planificador puede llevar a cabo balance de carga entre dos clases de tráfico X e Y con la clase de tráfico Y teniendo requisitos de retraso más largos (más relajados) que la clase de tráfico X. En un momento dado de tiempo, el retraso de un paquete en la clase de tráfico X puede denotarse $D(X)$ y el retraso de un paquete en la clase de tráfico Y puede denotarse $D(Y)$. El comportamiento a corto plazo del planificador puede seguir uno de los dos casos a continuación.

- Caso 1. En un segmento precedente de tiempo, puede haber más paquetes llegando para la clase de tráfico X que para la clase de tráfico Y. $D(X)$ puede acercarse al requisito de retraso de la clase de tráfico X antes de que $D(Y)$ se acerque al requisito de retraso de la clase de tráfico Y. En un segmento de tiempo posterior, el planificador puede asignar más recursos radio para los paquetes en la clase de tráfico X, y menos o ningún recurso radio para los paquetes en la clase de tráfico Y. $D(X)$ puede reducirse y $D(Y)$ puede empezar a incrementarse, lo cual puede equilibrar $D(X)$ y $D(Y)$ y evitar que $D(X)$ se mueva hacia su límite.
- Caso 2. El caso contrario también se puede aplicar. En el segmento precedente de tiempo, puede haber más paquetes llegando para la clase de tráfico Y que para la clase de tráfico X. En el siguiente segmento de tiempo, el planificador puede asignar más recursos radio para la clase de tráfico Y, y menos o ningún recurso radio para la clase de tráfico X. $D(Y)$ puede reducirse y $D(X)$ puede empezar a aumentar, lo cual puede equilibrar $D(X)$ y $D(Y)$.

El balance de carga descrito anteriormente puede extenderse a cualquier número de clases de tráfico. En una red muy congestionada, el planificador puede asignar más recursos radio a las clases de tráfico con más paquetes, y todas las clases de tráfico pueden acercarse a sus respectivos requisitos de retraso simultáneamente. Cuando las colas están llenas, el planificador puede beneficiarse al máximo de la elasticidad del tráfico, que es la tolerancia de retrasos, esperando al último momento posible para transmitir los paquetes que pertenecen a los flujos más tolerantes al retraso.

El planificador puede mantener los retrasos de los paquetes en cada clase de tráfico dentro de los requisitos de retraso de esa clase de tráfico y puede acercarse simultáneamente a los requisitos de retraso de todas las clases de tráfico. Sin embargo, el planificador puede tener una tendencia hacia la clase de tráfico más alta con el requisito de retraso más corto, por ejemplo, la clase de tráfico 1 en la Figura 2. Por lo tanto, cuando la red está fuertemente cargada, la clase de tráfico más baja con el requisito de retraso más largo puede empezar primero a experimentar retrasos intolerables. Esta clase de tráfico más baja puede ser para servicios "lo mejor posible" y puede abarcar tráfico FTP (tal como correo electrónico) y otro tráfico que puede soportar retrasos más largos. Exceder el requisito de retraso de tal tráfico puede tener un efecto insignificante. Por lo tanto, datos en la clase de tráfico más baja pueden ser mantenidos en la cola hasta que no ocurran desbordamiento de memoria temporal y vencimiento de tiempo de capa superior. El desbordamiento de memoria temporal o vencimiento de tiempo pueden resultar o no en el cese de un flujo de datos. Por ejemplo, un vencimiento de tiempo FTP puede resultar en una retransmisión de paquete, manteniendo por lo tanto el flujo de datos.

Para servicios interactivos tales como navegación web, los paquetes pueden ser mantenidos en colas a pesar de exceder sus requisitos de retraso. Los usuarios pueden empezar a abandonar los servicios cuando experimentan retrasos excesivos. Esto puede reducir las demandas de tráfico así como el balance de carga.

La descripción anterior asume que el planificador puede determinar libremente qué paquetes transmitir. Esta asunción puede no mantenerse de forma completa para todas las tecnologías radio. Además, para algunos servicios en tiempo real como VoIP, la red puede reservar algunos recursos radio para un flujo de datos dado de forma que los paquetes en ese flujo de datos pueden considerarse como preplanificados. La descripción anterior puede entonces aplicarse a clases de tráfico y recursos radio no afectados por estas desviaciones.

En un aspecto, se pueden priorizar flujos de datos, y se pueden colocar paquetes de los flujos de datos en puntos diferentes en una cola dependiendo de las prioridades de los diferentes flujos de datos. En un diseño, un flujo de

datos dado puede mapearse a una clase de tráfico como se describió con anterioridad y puede también asignársele un nivel de prioridad. En general, puede soportarse cualquier número de niveles de prioridad (N) para cada clase de tráfico. El nivel de prioridad 1 puede ser el más alto, el nivel de prioridad N puede ser el más bajo. Diferentes clases de tráfico pueden tener el mismo o diferente número de niveles de prioridad. Los paquetes para todos los flujos de datos en cada clase de tráfico pueden tener el requisito de retraso para cada clase de tráfico. Sin embargo, los paquetes para flujos de datos con niveles de prioridad más altos pueden enviarse con retrasos más cortos en media.

La Figura 3 muestra un diseño de planificación de prioridades para una clase de tráfico k (TC_k) con dos niveles de prioridad 1 y 2. En este diseño, un flujo de datos con nivel de prioridad 1 puede denominarse flujo de datos prioritario, y un flujo de datos con nivel de prioridad 2 puede denominarse flujo de datos no prioritario. La clase de tráfico k puede tener un requisito de retraso de D_k , que puede ser aplicable para todos los flujos de datos en la clase de tráfico k. Los paquetes de flujos de datos no prioritarios pueden denominarse paquetes no prioritarios y pueden tener un tiempo de encolado objetivo de T_2 , en donde en general $T_2 \leq D_k$. Los paquetes de flujos de datos prioritarios pueden denominarse paquetes prioritarios y pueden tener un tiempo de encolado objetivo de T_1 en donde en general $0 \leq T_1 \leq T_2$. El tiempo de encolado objetivo puede también denominarse tiempo de encolado estimado, retraso de transmisión esperado, etcétera. El tiempo de encolado objetivo T_2 puede ser dependiente de la carga de red y de otros factores. El tiempo de encolado objetivo T_1 puede seleccionarse en base a varios factores tales como el número total de niveles de prioridad soportados por el sistema, el retraso de encolado esperado actual o el requisito de retraso de la clase de tráfico k, el retraso esperado actual para este nivel de prioridad, etcétera. En un diseño, se puede seleccionar T_1 de forma que el retraso de transmisión esperado para los paquetes prioritarios no es más que P porcentaje del requisito de retraso de la clase de tráfico k, en donde P podría ser cualquier valor adecuado.

Cuando se recibe un paquete no prioritario para la clase de tráfico k, el paquete puede colocarse al final de la cola para la clase de tráfico k. Cuando se recibe un paquete prioritario (denotado como F en la Figura 3), el paquete puede colocarse en la misma cola. Sin embargo, en vez de colocar el paquete F al final de la cola, el paquete F puede colocarse en un punto dentro de la cola tal que su tiempo de encolado estimado es T_1 . El tiempo real de encolado de un paquete puede no ser conocido con seguridad debido a varios factores relativos a las dinámicas de la red y la compartición de recursos radio. El tiempo de encolado puede estimarse en base a información disponible tal como la carga de red más reciente, etcétera. El paquete F puede colocarse delante en la cola tal que el tiempo estimado de encolado es T_1 .

En general, un paquete prioritario puede colocarse en cualquier punto en una cola antes del final de la cola. El paquete prioritario puede colocarse al principio de la cola (no mostrado en la Figura 3) o en un punto entre el principio a el final de la cola (como se muestra en la Figura 3).

La Figura 4 muestra un diseño de planificación de prioridades para una clase de tráfico k (TC_k) con N niveles de prioridad de 1 a N, en donde en general $N \geq 1$. En este diseño, la clase de tráfico k puede tener un requisito de retraso D_k , el cual puede ser aplicable para todos los flujos de datos en la clase de tráfico k. Los paquetes para flujos de datos con el nivel de prioridad más alto 1 pueden tener un tiempo de retraso objetivo de T_1 , paquetes para flujos de datos con el segundo nivel de prioridad más alto 2 pueden tener un tiempo de retraso objetivo de T_2 , y así sucesivamente, y paquetes para flujos de datos con el nivel de prioridad más bajo N pueden tener un tiempo de retraso objetivo de T_N , en donde en general $0 \leq T_1 < T_2 < \dots \leq D_k$. Los paquetes que llegan con el nivel de prioridad más bajo N pueden colocarse al final de la cola. Los paquetes que llegan con niveles de prioridad más altos 1 a N-1 pueden colocarse en diferentes lugares en la cola, acorde con sus niveles de prioridad, de forma que estos paquetes puedan alcanzar los tiempos de encolado objetivo de T_1 a T_{N-1} , respectivamente.

Los tiempos de encolado objetivo para los niveles de prioridad más altos pueden seleccionarse de varias formas. En un diseño, que puede denominarse diseño de planificación de prioridades A, el tiempo de encolado objetivo T_n para el nivel de prioridad n puede seleccionarse tal que el retraso de transmisión esperado para paquetes con nivel de prioridad n no sea más que P_n porcentaje del tiempo de encolado esperado actual E_k para la clase de tráfico k, en donde $n \in \{1, \dots, N\}$. Los porcentajes para los N niveles de prioridad pueden definirse de tal forma que $0 \leq P_1 < P_2 < \dots < P_N \leq 1$.

En otro diseño, que puede denominarse diseño de planificación de prioridades B, el tiempo de encolado objetivo T_n para el nivel de prioridad n puede seleccionarse tal que el retraso de transmisión esperado para paquetes con nivel de prioridad n no sea más que P_n porcentaje del tiempo de encolado esperado actual E_k para la clase de tráfico k. En este diseño, el tiempo de encolado objetivo T_n puede estar limitado por un valor mínimo particular que puede ser común para todos los niveles de prioridad o puede ser diferente para cada nivel de prioridad.

Por claridad, un ejemplo específico de ambos diseños de planificación A y B se describen a continuación. En este ejemplo, la clase de tráfico k tiene un requisito de retraso de $D_k = 1500$ ms y un retraso de encolado esperado actual de $E_k = 1000$ ms. Ya que $E_k < D_k$, no se cumple la condición de congestión para la clase de tráfico k. Se definen cinco niveles de prioridad de 1 a 5 para la clase de tráfico k. Para el diseño A, los tiempos de encolado objetivo T_1 a T_5 para los cinco niveles de prioridad se definen como 0%, 15%, 30%, 50% y 75% del requisito de retraso D_k para la clase de tráfico k. Para el diseño B, los tiempos de encolado objetivo T_1 a T_5 para los cinco niveles de prioridad se definen como 0%, 15%, 30%, 50% y 75% del requisito de retraso E_k para la clase de tráfico k. La Tabla 1 muestra

ES 2 385 124 T3

los tiempos de encolado objetivo T_1 a T_5 para los cinco niveles de prioridad para ambos diseños A y B.

Nivel de prioridad	Tiempo de encolado objetivo	Planificación de prioridades Diseño A con $D_k = 1500$ ms		Planificación de prioridades Diseño B con $E_k = 1000$ ms	
1	T_1	0% de D_k	0 ms	0% de E_k	0 ms
2	T_2	15% de D_k	225 ms	15% de E_k	150 ms
3	T_3	30% de D_k	450 ms	30% de E_k	300 ms
4	T_4	50% de D_k	750 ms	50% de E_k	500 ms
5	T_5	75% de D_k	1000 ms	75% de E_k	750 ms

5 Para el diseño A, el tiempo de encolado objetivo T_n para cada nivel de prioridad n puede estar delimitado por el más pequeño entre E_k o D_k , de forma que $T_n \leq \min \{E_k, D_k\}$. Cuando la cola no está llena, E_k es menos que D_k y representa el final de la cola. Por lo tanto, si T_n es mayor que E_k para cualquier nivel de prioridad dado n , los paquetes para el nivel de prioridad n pueden colocarse en E_k en vez de en T_n . En el ejemplo mostrado en la Tabla 1, el nivel de prioridad 5 tiene un valor computado de 1125 ms para el 75% de D_k . Ya que 1125 ms es mayor que 1000 ms para E_k , T_5 se fija a 1000 ms en vez de a 1125 ms. Los paquetes con nivel de prioridad 5 pueden por lo tanto
10 colocarse al final de la cola en vez de en 1125 ms.

15 Cuando la red está congestionada y $E_k < D_k$, los diseños A y B pueden proporcionar diferentes tiempos de encolado objetivo para los niveles de prioridad, por ejemplo, como se muestra en la Tabla 1. Sin embargo, cuando la red se acerca al punto de congestión, los diseños A y B pueden proporcionar los mismos tiempos de encolado objetivo para los N niveles de prioridad.

20 En otro diseño, que no se muestra en la Figura 4 y que puede denominarse diseño de planificación de prioridades C, los paquetes con niveles de prioridad más altos, pueden colocarse en el mismo punto en la cola, de forma que $T_1 = T_2 = \dots = T_{N-1}$. Sin embargo, los paquetes pueden ordenarse de acuerdo a su nivel de prioridad, de forma que los paquetes con nivel de prioridad 1 se colocan antes que los paquetes de prioridad 2, que se colocan antes que los paquetes de prioridad 3, etcétera. El punto único T_n puede ser el inicio de la cola o puede ser en algún lugar entre el inicio y el final de la cola.

25 Los tiempos de encolado objetivo para los N niveles de prioridad pueden también definirse de otras maneras. Por ejemplo, algunos niveles de prioridad pueden tener el mismo tiempo de encolado objetivo mientras otros niveles de prioridad pueden tener diferentes tiempos de encolado objetivo.

30 Un diseño de planificación de prioridades puede seleccionarse para su uso, y la misma planificación de prioridades puede llevarse a cabo para cada una de las K clases de tráfico. Se puede usar también una combinación de diseños de planificación de prioridades. Por ejemplo, National Communication Systems (NCS) está desarrollando un conjunto de requisitos para todas las redes en los Estados Unidos para cumplir con un esquema de 5 niveles de prioridad. Se pueden definir cinco o más niveles de prioridad para soportar requisitos NCS. Como otro ejemplo, el Servicio de Prioridades Multimedia (MMPS) en 3GPP2 estipula hasta N niveles de prioridad, en donde N puede ser
35 definido por un operador de red. Se puede utilizar un número configurable de niveles de prioridad para soportar requisitos MMPS.

40 Los diseños de planificación de prioridades descritos anteriormente pueden usarse para cualquier número de niveles de prioridad (N). N puede seleccionarse en base a uno o más estándares soportados por la red y/o otros factores. Por ejemplo, National Communication Systems (NCS) está desarrollando un conjunto de requisitos para todas las redes en los Estados Unidos para cumplir con un esquema de cinco niveles de prioridad. Se pueden definir cinco o más niveles de prioridad para soportar los requisitos NCS. Como otro ejemplo, el Servicio de Prioridad Multimedia (MMPS) en 3GPP2 estipula hasta N niveles de prioridad, en donde N puede ser definido por un operador de red. Se puede utilizar un número configurable de niveles de prioridad para soportar requisitos MMPS.

45 Se le pueden asignar atributos QoS a un flujo de datos en el momento de invocación del servicio o de inicio de sesión. Los atributos QoS pueden ser "modulados" efectivamente por los niveles de prioridad descritos con anterioridad. Por ejemplo, un servicio de correo electrónico puede tener un requisito de retraso de 10 segundos, y los paquetes de correo electrónico pueden almacenarse en una cola hasta 10 segundos sin provocar errores de vencimiento de tiempo. Sin embargo, a un usuario prioritario se le puede asignar el nivel de prioridad 1, y los paquetes de correo electrónico para este usuario pueden colocarse en la cola tal que experimenten retrasos de
50 encolado de dos segundos o menos. Este servicio de correo electrónico no cambia para este usuario prioritario, por ejemplo, no ocurrirían errores de vencimiento de tiempo a no ser que el retraso de encolado es de 10 segundos o más. Por lo tanto, bajo un escenario de funcionamiento normal, el servicio de correo electrónico para un usuario

prioritario no incurriría en un vencimiento de tiempo independientemente de si la red está congestionada o no. La condición de error de vencimiento de tiempo no se cambia para el usuario prioritario y puede ocurrir bajo ciertas circunstancias extraordinarias, por ejemplo, fallo de una estación base. En cualquier caso, la planificación de prioridades descrita con anterioridad puede permitir a los flujos de datos tener retrasos de encolado más cortos y otro tratamiento preferencial mientras aún mantienen los atributos QoS de las clases de tráfico asociadas. Efectivamente, los atributos QoS de estos flujos de datos bajo condiciones de congestión de red serían como si la carga de tráfico fuese baja.

La Figura 5 muestra un diseño de un proceso 500 para enviar datos con planificación de prioridad. El proceso 500 puede ser llevado a cabo por una celda para transmisión de datos en enlace delantero, por un terminal para transmisión de datos en el enlace reverso o por alguna otra entidad de red. Puede recibirse un primer paquete de un primer nivel de prioridad (bloque 512) y puede colocarse al final de la cola (bloque 514). Puede recibirse un segundo paquete de un segundo nivel de prioridad (bloque 516). El segundo paquete puede colocarse en un primer punto en la cola más adelante del final de la cola (bloque 518). En un diseño del bloque 518, se puede determinar un tiempo objetivo de encolado para el segundo nivel de prioridad, por ejemplo, en base a un porcentaje predeterminado de un requisito de retraso de un segundo paquete o un porcentaje predeterminado de un retraso de encolado esperado para los paquetes en la cola. Puede determinarse entonces el primer punto en la cola en base al nivel de prioridad del paquete. Los primero y segundo paquetes pueden pertenecer a diferentes flujos de tráfico en la misma clase de tráfico y pueden tener atributos QoS asociados con la clase de tráfico. Los primero y segundo paquetes pueden estar también destinados para dos terminales.

Se puede recibir un tercer paquete de un tercer nivel de prioridad mayor que el segundo nivel de prioridad (bloque 520). El tercer paquete puede colocarse en un segundo punto en la cola más adelante del primer punto (bloque 522). En otro diseño del bloque 522, el tercer paquete puede colocarse en el primer punto en la cola. Sin embargo, pueden colocarse paquetes del tercer nivel de prioridad delante de paquetes del segundo nivel de prioridad en el primer punto cuando estos paquetes se reciben al mismo tiempo.

En un diseño, se puede recibir un paquete de uno de múltiples niveles de prioridad. Los múltiples niveles de prioridad pueden estar asociados con diferentes puntos en la cola, Se puede determinar un punto en la cola para el paquete en base al nivel de prioridad del paquete. Cada paquete puede ser colocado entonces en la cola en un punto diferente. En cualquier caso, los paquetes en la cola pueden ser enviados en orden secuencial (bloque 524).

En un diseño, se pueden mantener colas múltiples para múltiples clases de tráfico, una cola para cada clase de tráfico. Se pueden soportar múltiples niveles de prioridad para cada clase de tráfico y pueden estar asociados con diferentes puntos en la cola para esa clase de tráfico. Cada paquete puede ser colocado en la cola para la clase de tráfico de ese paquete y en el punto en la cola determinado por el nivel de prioridad de ese paquete.

En otro aspecto, un paquete para un terminal puede ser transferido desde una celda fuente a una celda objetivo debido a un traspaso del terminal y puede tenerse en cuenta la cantidad de tiempo que el paquete ha estado ya esperando en la cola en la celda fuente. El paquete puede ser colocado en un punto avanzado en la cola de la celda objetivo. Este punto puede determinarse en base a la cantidad de tiempo que el paquete ha estado ya esperando. Al no colocar el paquete al final de la cola de la celda objetivo, se puede evitar un retraso de encolado excesivo para el paquete.

La Figura 6 muestra el enrutado y transmisión de un paquete a un terminal sin traspaso. En el tiempo t_1 , la red puede recibir el paquete destinado al terminal y puede enrutar el paquete a una celda servidora para el terminal. La celda servidora puede ser la celda con una mayor relación señal a ruido (SNR) en el terminal. La celda servidor puede poner un sello temporal tras la recepción para rastrear el retraso de encolado del paquete. La celda servidora puede determinar la clase de tráfico para el paquete y puede colocar el paquete en una cola apropiada, por ejemplo, o al final de la cola o en un punto adelantado si el paquete tiene una mayor prioridad. El paquete puede progresar en la cola a lo largo del eje temporal (el cual es verticalmente hacia abajo en la Figura 4) y hacia el requisito de retraso de D_k para la clase de tráfico. En el tiempo t_2 , el paquete alcanza la cabeza de la cola. Si el terminal todavía está dentro de la cobertura de la celda servidora, entonces la celda servidora puede transmitir el paquete al terminal en el instante originalmente planificado.

La Figura 7 muestra el enrutado y transmisión de un paquete a un terminal con traspaso. En el tiempo t_1 , la red puede recibir el paquete destinado para el terminal y puede enrutar el paquete a una celda servidora para el terminal. La celda servidora puede poner un sello temporal en el paquete tras su recepción y puede colocar el paquete en un punto adecuado en una cola adecuada. El paquete puede progresar en la cola a lo largo del eje temporal. El terminal puede ser móvil y puede estimar el SNR de las celdas cercanas. El terminal puede determinar que el SNR de otra celda es mejor que el SNR de la celda servidora. En el tiempo t_2 , el terminal puede enviar un informe de medidas SNR a la celda servidora y/o la mejor celda, que es la celda objetiva para traspaso. En el tiempo t_3 , para traspaso iniciado por red, la celda servidora puede enviar un mensaje de dirección de traspaso al terminal para iniciar el traspaso a la celda servidora.

En el tiempo t_4 , la celda servidora puede transferir el paquete a la celda objetivo. La celda objetivo puede determinar

la clase de tráfico del paquete y puede colocar el paquete en una cola apropiada. La celda objetivo puede colocar el paquete en un lugar en la cola tal que el paquete es avanzado en el tiempo en una cantidad de tiempo que el paquete ya ha estado esperando en la cola en la celda servidora. Al paquete se le puede tener en cuenta, por lo tanto, el tiempo previo de espera, de forma que puede ser transmitido de forma puntual. El paquete puede entonces progresar en la cola en la celda objetivo de forma normal y puede ser transmitido por la celda objetivo al terminal cuando el paquete alcanza la cabeza de la cola en el tiempo t_5 .

La Figura 7 muestra un ejemplo en el que el terminal es traspasado una vez que el paquete espera a ser transmitido. El manejo de paquete en la Figura 7 puede extenderse para cubrir cualquier número de traspasos antes de la transmisión del paquete. En general, una celda objetivo puede colocar un paquete en un punto avanzado en una cola adecuada en la celda objetivo de forma que al paquete se le tiene en cuenta parte de o todo el tiempo de espera previo.

Cada celda que maneja el paquete puede también colocar el paquete en un punto apropiado en una cola adecuada para tener en cuenta la prioridad del paquete. La celda que recibe el paquete de la red puede colocar el paquete en un punto determinado en base a cualquiera de los diseños de planificación de prioridades descritos con anterioridad. Cada celda subsiguiente puede (i) determinar el tiempo de encolado restante objetivo para el paquete, el cual puede ser igual al tiempo de encolado objetivo menos el tiempo de espera previo, y (ii) colocar el paquete en la cola tal que puede esperar el tiempo de encolado objetivo restante. El manejo del paquete por diferentes celdas objetivo debido a un traspaso no debería afectar adversamente el manejo prioritario del paquete. La colocación del paquete en la cola de cada celda objetivo puede imitar la colocación del paquete en la cola de la celda que recibió originalmente el paquete de la red.

Un terminal puede ser traspasado a una celda objetivo y luego de vuelta a la celda servidora previa. La celda servidora puede almacenar el paquete por una cantidad predeterminada de tiempo, por ejemplo, hasta que expira el requisito de tiempo del paquete. Esto puede evitar tener que transmitir el paquete desde la celda objetivo otra vez hacia la celda servidora previa.

Los retrasos de encolado se incrementan generalmente durante periodos de congestión de red, y un paquete puede esperar más en una cola debido a la congestión. La probabilidad de que las condiciones de canal cambien y la probabilidad de traspaso pueden incrementarse ambas cuanto más espera el paquete para transmisión. Los paquetes prioritarios pueden generalmente esperar en colas menos tiempo que los paquetes no prioritarios en la misma clase de tráfico. Por lo tanto, la movilidad puede tener menos impacto en los paquetes prioritarios que en los paquetes no prioritarios y el impacto puede ser progresivamente menos para los niveles de prioridad progresivamente mayores.

La Figura 8 muestra un diseño de un proceso 800 para enviar datos para tener en cuenta el tiempo de encolado previo. Se puede recibir un paquete a enviar a un terminal de una primera celda (bloque 812). El paquete puede ser colocado entonces (i) al final de la primera cola en la primera celda o (ii) en un punto en la primera cola determinado en base al nivel de prioridad del paquete. El paquete puede ser retenido por la primera celda hasta que ha pasado el requisito de retraso del paquete.

Se puede determinar la cantidad de tiempo que el paquete ha estado ya esperando en la primera cola en la primera celda (bloque 814). El paquete puede ser colocado en un segundo punto en la segunda cola en una segunda celda para tener en cuenta el tiempo que el paquete ya ha estado esperando en la primera cola (bloque 816). En un diseño del bloque 816, el tiempo de encolado objetivo restante para el paquete puede determinarse en base al tiempo de encolado objetivo para el paquete y la cantidad de tiempo que el paquete ya ha estado esperando en la primera cola. El paquete puede entonces ser colocado en el punto en la segunda cola determinado en base al tiempo de encolado restante para el paquete. El paquete puede ser enviado al terminal cuando alcanza la cabeza de la segunda cola (bloque 818).

Se puede llevar a cabo control de admisión para limitar el número de terminales que admitir en la red, para limitar la cantidad de tráfico y para evitar o mitigar la congestión de red. Cuando la red no está congestionada, se puede admitir cada terminal que desea acceso a la red. Al incrementar el volumen de tráfico y acercarse la red al punto de congestión, se pueden regir por una política de admisión los incrementos adicionales en volumen de tráfico. Se pueden restringir los nuevos flujos de datos y, en un cierto punto, pueden ser parados completamente para mitigar la congestión de red.

Aún en otro aspecto más, se puede llevar a cabo control de admisión y gestión de recursos de forma que reduzcan los efectos adversos en los flujos de datos prioritarios y terminales. En un primer diseño, los flujos de datos pueden ser clasificados como flujos de datos prioritarios o no prioritarios. Un flujo de datos puede ser considerado como un flujo de datos prioritario en base a varios factores tales como atributos QoS del flujo de datos, si el flujo de datos es o no para servicios de emergencia, si el flujo de datos es o no para un usuario con una suscripción de alta calidad, etcétera. Los flujos de datos que no son flujos de datos prioritarios pueden ser considerados como flujos de datos no prioritarios. Los flujos de datos prioritarios y no prioritarios pueden ser admitidos en base a la carga de celda, como se describe a continuación. El primer diseño puede permitir la admisión de flujos de datos en base a cada uno.

En un segundo diseño, los terminales pueden ser clasificados como terminales prioritarios o no prioritarios y pueden ser admitidos en base a la carga de celda. Un terminal puede tener una sesión activa o durmiente y puede tener uno o más flujos de datos en una o más clases de tráfico para un sesión activa. El segundo diseño puede considerarse como un caso especial del primer diseño en el que todos los flujos de datos de un terminal pueden ser o admitidos o no admitidos.

En un diseño, el control de admisión puede ser llevado a cabo para cada celda en base a la carga de esa celda. En un diseño, los flujos de datos prioritarios pueden admitirse y los flujos de datos no prioritarios pueden ser restringidos o bloqueados al completarse la carga de celda. En un diseño, para soportar movilidad de los terminales prioritarios, cada celda puede separar algunos recursos radio o puede controlar la asignación de recursos radio en el caso de que terminales prioritarios en las celdas vecinas sean traspasados a esa celda. Cada celda puede restringir o parar la admisión de terminales no prioritarios en esa celda si hay terminales prioritarios con sesiones activas (por ejemplo, tiene actividad reciente o en curso, tal como sesiones VoIP en curso) en celdas vecinas.

Una entidad de red que lleva a cabo control de admisión (por ejemplo, Función de Política de Control) puede tener acceso a varios tipos de información tal como:

- Atributos QoS de flujos de datos o sesiones de terminales,
- Estados de prioridad de los flujos de datos o sesiones, incluyendo aquellos en las celdas vecinas,
- Estadísticas actuales de retraso de encolado para cada celda de interés, y
- Otra información relevante.

Se pueden definir varias reglas de admisión en base a la información anterior y/u otra información. Se pueden definir diferentes reglas de admisión para diferentes niveles de carga de celda. En general, se puede soportar cualquier número de niveles de carga de celda y los niveles de carga de celda pueden ser definidos de varias maneras, por ejemplo, por un operador de red.

En un diseño, se pueden aplicar tres reglas de admisión a tres niveles de carga de celda, como sigue:

- Carga de celda ligera - no se aplican restricciones de admisión.
- Carga de celda moderada – se admiten flujos de datos no prioritarios en una escala inversamente proporcional al número de terminales prioritarios con sesiones activas en las celdas vecinas, y
- Carga de celda fuerte – se admiten solo flujos de datos prioritarios.

En un diseño, se puede llevar a cabo control de admisión de forma separada para cada clase de tráfico. En este diseño, se puede determinar el nivel de carga para cada clase de tráfico en cada celda. En un diseño, el nivel de carga de cada clase de tráfico se puede definir en base a un retraso de encolado medio para esa clase tráfico. Por ejemplo, la carga de la celda ligera (o ninguna congestión) puede ser definida por el retraso medio encolado siendo menos que un primer porcentaje (por ejemplo, 50%) del requisito de retraso de la clase de tráfico. La carga de celda moderada puede ser definida por el retraso medio de encolado estando entre el primer porcentaje y un segundo porcentaje (por ejemplo, 90%) del requisito de retraso. La carga de celda fuerte (o congestión) puede estar definida por el retraso de encolado medio siendo mayor que el segundo porcentaje del requisito de retraso. Los niveles de carga de celda pueden también definirse en base a otros criterios.

Para carga de celda moderada, el número de flujos de datos no prioritarios a admitir puede ser una función del retraso de encolado medio. En un diseño, el retraso de encolado medio puede convertirse a una capacidad reservada, como sigue:

$$C_k = (Z_k - Q_k) * F_k \quad \text{Ecuación (1)}$$

en donde Q_k es el retraso de encolado medio relativo al requisito de retraso para la clase de tráfico k ,

F_k es un factor de escalado para la clase de tráfico k ,

Z_k es el segundo porcentaje para la clase de tráfico k , y

C_k es la capacidad reservada para la clase de tráfico k .

La capacidad reservada puede darse en número de sesiones o terminales o el número de flujos de datos. El valor del factor de escalado F_k y el valor del segundo porcentaje Z_k pueden seleccionarse en base a pruebas empíricas, simulaciones por ordenador, etcétera. El retraso medio de encolado se puede filtrar, por ejemplo, promediándolo a lo largo de un periodo de tiempo. A título de ejemplo, el retraso de encolado medio puede ser el 60% del requisito de retraso, el factor de escalado puede ser igual a 0,5 y el segundo porcentaje puede ser igual al 90%. La capacidad reservada puede entonces computarse como $C_k = (90 - 60) * 0,5 = 15$. Un flujo de datos no prioritario puede ser admitido si se cumplen las siguientes condiciones:

$$Q_k < Z_k \text{ y } C_k > SP_k, \quad \text{Ecuación (2)}$$

donde SP_k es el número de terminales prioritarios para la clase de tráfico K en las celdas vecinas.

5 En el ejemplo anterior con $Q_k = 60\%$ y $F_k = 0,5$, un flujo de datos no prioritario puede ser admitido si el número total de terminales prioritarios activos en las celdas vecinas es menor a 15.

10 El diseño en la ecuación (2) lleva a cabo control de admisión de forma que soporta movilidad de los terminales prioritarios. Una celda x dada puede tener información sobre la presencia de terminales prioritarios activos en celdas vecinas y puede usar esta información para control de admisión. La condición ($C_k > SP_k$) en la ecuación (2) asegura que hay suficiente capacidad reservada en la celda x para manejar el traspaso potencial de terminales prioritarios en las celdas vecinas a la celda x. El control de admisión para la celda x puede severizarse si hay muchos terminales prioritarios con flujos de datos activos en las celdas vecinas y pueden relajarse si no es así.

15 La SNR de los terminales prioritarios en las celdas vecinas puede ser también considerada para el control de admisión en la celda x. Por ejemplo, puede haber muchos terminales prioritarios en una celda vecina grande, pero estos terminales prioritarios pueden estar lejos del área de cobertura de la celda x. La proximidad de cada terminal prioritario a la celda x puede reflejarse en el SNR de la celda x tal y como se mide en ese terminal. En un diseño, la presencia de terminales prioritarios lejanos puede descontarse, por ejemplo, usando $C_k > SP_k * D_k$, donde D_k es un factor de descuento menor que uno. En otro diseño, un terminal prioritario en una celda vecina, puede considerarse al computar SP_k solo si el SNR de la celda x tal y como se mide en el terminal excede un umbral de SNR. En otro diseño más, la "frescura" de las medidas de SNR disponibles para las celdas puede considerarse para control de admisión. Un terminal puede no enviar informes SNR hasta que y a no ser que una SNR de una celda dada exceda un umbral. El terminal puede entonces informar el SNR de la celda que desencadenó el informe SNR así como otras celdas vecinas que pueden ser medidas por el terminal.

25 En otro diseño, se puede llevar a cabo el control de admisión de forma separada para cada grupo de una o más clases de tráfico. Por ejemplo, un grupo puede incluir servicios de tiempo real como VoIP y otro grupo puede incluir las clases de tráfico restantes. En este diseño, el nivel de carga puede ser determinado para grupo de clases de tráfico en cada celda, como se describió con anterioridad, en donde k es ahora un índice para un grupo de clases de tráfico en vez de una clase de tráfico específica. En un diseño, el nivel de carga para un grupo dado puede venir dado por un retraso medio de encolado para todas las clases de tráfico en ese grupo como un porcentaje del requisito de retraso de estas clases de tráfico. El retraso de encolado medio para un grupo dado puede expresarse como una media ponderada de los retrasos medios de encolado para las clases de tráfico en ese grupo. El peso para cada clase de tráfico puede estar determinado por el número de flujos de datos activos en esa clase de tráfico, el número de paquetes en la cola para la clase de tráfico, etc.

30 Para todos los diseños descritos con anterioridad, todas las clases de tráfico deberían acercarse al punto de congestión de forma aproximadamente simultánea. En un instante dado, el retraso de encolado medio para paquetes en todas las clases de tráfico debería ser aproximadamente el mismo porcentaje del requisito de retraso para cada clase de tráfico.

45 La ecuación (1) muestra un ejemplo de mapear retraso medio de encolado con capacidad reservada. La ecuación (2) muestra un ejemplo de admitir un flujo de datos no prioritario en base al retraso medio de encolado y a la capacidad reservada. La computación de la capacidad reservada y/o la admisión de flujos de datos no prioritarios puede también considerar otros factores como el espectro de frecuencia disponible, la tasa de transferencia total, la tasa de datos requerida para un nuevo flujo de datos, etcétera. Por ejemplo, una sesión de video-telefonía puede tener diferentes requisitos de flujo de datos y puede ser tratada diferente que una sesión VoIP para su admisión.

50 Para carga de celda fuerte, el control de admisión puede distinguir entre diferentes niveles de prioridad. Por ejemplo, los flujos de datos de prioridad más alta pueden ser admitidos incluso si el retraso de encolado medio es el 100% del requisito de retraso, los segundos flujos de datos de segunda prioridad más alta pueden ser admitidos solamente si el retraso de encolado medio es el 95% del requisito de retraso, etcétera.

55 En un diseño, un flujo de datos puede mantenerse solo si el requisito de retraso del flujo de datos puede cumplirse y puede ser finalizado si no. Al flujo de datos pueden asignársele atributos QoS que pueden indicar si el flujo de datos puede ser descartado si no se puede cumplir su requisito de retraso. Pueden ser aplicables diferentes tratamientos de flujo para diferentes niveles de prioridad. Por ejemplo, los terminales de prioridad más alta pueden ver mantenidos siempre todos sus flujos de datos, los terminales de segunda prioridad más alta puede ver mantenidos solo sus flujos de datos de alta prioridad, etcétera.

60 Las técnicas descritas en este documento pueden soportar admisión de flujos de datos o terminales prioritarios incluso durante periodos de condiciones de tráfico congestionados. Las técnicas pueden también asegurar que los servicios ofrecidos a los terminales prioritarios se ven mínimamente afectados e interrumpidos debido a carga de red y también como consecuencia de la movilidad de usuario. Las técnicas pueden ser usadas para soportar servicios de emergencia para (i) servicios de tiempo real tales como VoIP y videoconferencia y (ii) servicios multimedia no de tiempo real tales como descargas de información de ruta de escape de emergencia, acceso a páginas web para

información actualizada del tiempo, flujos de tráfico vehicular, etcétera. Las técnicas también pueden utilizarse para soportar diferentes niveles de servicios. Por ejemplo, un usuario con una suscripción de alta calidad puede ser tratado como un terminal prioritario mientras que un usuario con una suscripción general puede ser tratado como un terminal no prioritario.

5 La Figura 9 muestra un diseño de un proceso 900 para controlar la admisión en una red de comunicación. El proceso 900 puede ser llevado a cabo por una celda o una entidad de red. La carga de celda puede determinarse en base a al menos un criterio, por ejemplo, el retraso medio de encolado de paquetes a enviar (bloque 912). En un
10 diseño, la carga de celda puede ser declarada como (i) ligera si el retraso medio de encolado es menor que un primer umbral, (ii) fuerte si el retraso medio de encolado es mayor que un segundo umbral, o (iii) moderado si el retraso medio de encolado está entre los primero y segundo umbrales. Los primero y segundo umbrales pueden determinarse en base a los primero y segundo porcentajes del requisito de retraso de los paquetes a enviar, siendo el segundo porcentaje mayor que el primer porcentaje. La carga de celda puede también determinarse de otras formas y/o en base a otros criterios.

15 Todos los flujos de datos prioritarios y los flujos de datos no prioritarios para terminales pueden ser admitidos si la carga de celda es ligera (bloque 914). Si un flujo de datos es un flujo prioritario o un flujo de datos no prioritario puede determinarse en base a la clase de suscripción, atributos QoS del flujo de datos y/u otra información. Si la carga de celda es fuerte solo pueden ser admitidos flujos de datos prioritarios (bloque 916). Los flujos de datos
20 prioritarios y los seleccionados de entre flujos de datos no prioritarios pueden ser admitidos si la carga de celda es moderada (bloque 918). En un diseño del bloque 918, los seleccionados de entre flujos de datos no prioritarios pueden ser admitidos en base al retraso de encolado medio de los paquetes a enviar y el número de terminales con flujos de datos prioritarios en las celdas vecinas, por ejemplo, como se muestra en la ecuación (2). Algunos recursos de radio de la celda pueden reservarse en el caso de que terminales con flujos de datos prioritarios en celdas
25 vecinas sean traspasados a la celda.

Los bloques 912 a 918 descritos anteriormente pueden ser para una clase de tráfico o un grupo de clases de tráfico agregadas juntas. En un diseño, se puede determinar la carga de celda para cada clase de tráfico en base a al
30 menos un criterio, por ejemplo, el retraso medio de encolado para paquetes en esa clase de tráfico. Los flujos de datos prioritarios y los flujos de datos no prioritarios pueden ser admitidos en base a la carga de celda para esa clase de tráfico, como se describió con anterioridad. En otro diseño, la carga de celda puede determinarse para cada grupo de clases de tráfico. Los flujos de datos prioritarios y los flujos de datos no prioritarios para cada grupo de clases de tráfico pueden ser admitidos en base a la carga de celda para ese grupo.

35 La Figura 10 muestra un diagrama de bloques de un diseño de un terminal 110, una estación base 120, una entidad de red 130. En el terminal 110, un procesador módem 1024 puede recibir datos a enviar al terminal, procesar (por ejemplo, codificar, modular, esparcir y mezclar) los datos, y generar muestras de salida. Un transmisor (TMTR) 1032 puede acondicionar (convertir a analógico, filtrar, amplificar y convertir de forma ascendente en frecuencia) las muestras de salida y generar una señal de enlace reverso, que puede ser transmitida a través de una antena 1034.
40 En el enlace directo, una antena 1034 puede recibir señales de enlace directo de una estación base 120 y/u otras estaciones base. Un receptor (RCVR) 1036 puede acondicionar (por ejemplo, filtrar, amplificar, convertir de forma descendente en frecuencia y digitalizar) la señal recibida de la antena 1034 y proporcionar muestras. El procesador módem 1024 puede procesar (por ejemplo, demodular y decodificar) las muestras y proporcionar datos decodificados. El procesador módem 1024 puede llevar a cabo procesamiento según una tecnología radio (por ejemplo, CDMA 1X, HRPD, WCDMA, GSM, etcétera) utilizada por la red.

Un procesador/controlador 1020 puede dirigir el funcionamiento al terminal 110. El controlador/procesador 1020 puede llevar a cabo o dirigir el proceso 500 en la Figura 5 y/u otros procesos para las técnicas descritas en este documento. Una memoria 1022 puede almacenar códigos de programa y datos para el terminal 110 y puede
50 implementar una o más colas para una o más clases de tráfico. Un procesador de señal digital 1026 puede llevar a cabo varios tipos de procesamiento para el terminal 110. Los procesadores 1020, 1024 y 1026 y a la memoria 1022 pueden implementarse en un circuito integrado de aplicación específica (ASIC) 1010. La memoria 1022 puede también implementarse externa al ASIC.

55 En la estación base 120, el transmisor/receptor (TMTR/RCVR) 1046 puede soportar comunicación radio con el terminal 110 y/u otros terminales. Un controlador/procesador 1040 puede también llevar a cabo varias funciones para comunicación con los terminales. El controlador/procesador 1040 puede también llevar a cabo o dirigir el proceso 500 en la Figura 5, el proceso 800 en la Figura 8, el proceso 900 en la Figura 9, y/u otros procesos para las técnicas descritas en este documento. Una memoria 1042 puede almacenar códigos de programa y datos para la
60 estación base 120. La memoria 1042 puede implementar una o más colas para una o más clases de tráfico. Una unidad de comunicación (Comm) puede soportar comunicación con otras entidades de red, por ejemplo, entidad de red 130. En general, la estación base 120 puede incluir cualquier número de controladores, procesadores, memorias, transmisores, receptores, unidades de comunicación, etcétera.

65 La entidad de red 130 puede ser el controlador de red 122 o la pasarela IP 124 en la Figura 1 o puede ser otra entidad de red. Dentro de la entidad de red 130, un controlador/procesador 1050 puede llevar a cabo o dirigir el

proceso 500 en la Figura 5, el proceso 900 en la Figura 9 y/u otros procesos para las técnicas descritas en este documento. Una memoria 1052 puede almacenar códigos de programa y datos para la entidad de red 130. Una unidad de comunicación 1054 puede soportar comunicación con otras entidades de red, por ejemplo, la estación base 120 puede incluir cualquier número de controladores, procesadores, memorias, transmisores, receptores, unidades de comunicación, etcétera.

Aquellos expertos en la técnica entenderían que se pueden representar información y señales utilizando cualquiera de una variedad de diferentes tecnologías y técnicas. Por ejemplo, datos, instrucciones, comandos, información, señales, bit, símbolos y chips que pueden ser referenciados a lo largo de la descripción anterior pueden representarse mediante voltajes, corrientes, ondas electromagnéticas, campos o partículas magnéticas, campos o partículas ópticas o cualquier combinación de los mismos.

Aquellos expertos en la técnica entenderían además que los varios ilustrativos bloques lógicos, módulos, circuitos y pasos de algoritmo descritos en conexión con la divulgación en este documento pueden implementarse como hardware electrónico, software de ordenador o combinaciones de ambos. Para ilustrar claramente esta intercambiabilidad de software y hardware, varios ilustrativos componentes, bloques, módulos, circuitos y pasos se han descrito anteriormente en términos de su funcionalidad. Si tal funcionalidad se implementa como hardware o software depende de la aplicación particular y restricciones de diseño impuestas por el sistema global. Los expertos en la técnica pueden implementar la funcionalidad descrita de formas diferentes para cada aplicación en particular, pero tales decisiones de implementación no deberían interpretarse como fuera del alcance de la presente divulgación.

Los varios ilustrativos bloques lógicos, módulos y circuitos descritos en conexión con la divulgación en este documento pueden implementarse o llevarse a cabo con un procesador de propósito general, un procesador digital de señal (DSP), un circuito integrado de aplicación específica (ASIC), una matriz programable de puertas (FPGA) u otro dispositivo lógico programable, puerta discreta o lógica de transistor, componentes hardware discretos o cualquier combinación de los mismos diseñado para llevar a cabo las funciones descritas en este documento. Un procesador de propósito general puede ser un microprocesador, pero de forma alternativa, el procesador puede ser cualquier procesador convencional, controlador, microcontrolador o máquina de estados. Un procesador también puede implementarse como una combinación de dispositivos de comunicación, por ejemplo, una combinación de un DSP y un microprocesador, una pluralidad de microprocesadores, uno o más microprocesadores en conjunción con un núcleo DSP o cualquier otra configuración semejante.

Las etapas de un método o algoritmo descrito en conexión con la divulgación de este documento pueden realizarse directamente en hardware, en un módulo software ejecutado por un procesador o en una combinación de ambos. Un módulo software puede residir en una memoria RAM, memoria flash, memoria ROM, memoria EPROM, memoria EEPROM, registros, disco duro, un disco extraíble, un CD-ROM o cualquier otra forma de medio de almacenamiento conocida en la técnica. Un medio de almacenamiento a título de ejemplo se acopla al procesador tal que el ordenador puede leer información de, y escribir información en, el medio de almacenamiento. De forma alternativa, el medio de almacenamiento puede ser integral al procesador. El procesador y medio de almacenamiento pueden residir como componentes discretos en un terminal de usuario.

En uno o más diseños de ejemplo, las funciones descritas se pueden implementar en hardware, software, firmware, o cualquier combinación de los mismos. Si se implementan en software, las funciones se pueden almacenar en o transmitir sobre como una o más instrucciones o código en un medio legible por ordenador. Los medios legibles por ordenador incluyen tanto medios de almacenamiento de ordenador y medios de comunicación incluyendo cualquier medio que facilita la transferencia de un programa de ordenador de un lugar al otro. Un medio de almacenamiento puede ser cualquier medio disponible que puede ser accedido por un ordenador de propósito general o de propósito especial. A título de ejemplo, y no limitante, tal medio legible por ordenador puede comprender RAM, ROM, EEPROM, CD-ROM u otro almacenamiento de disco óptico, almacenamiento de disco magnético u otros dispositivos de almacenamiento magnético, o cualquier otro medio que puede usarse para transportar o almacenar medios de código de programa en forma de instrucciones o estructuras de datos y que puede ser accedido por un ordenador de propósito general o de propósito específico. También, cualquier conexión se denomina apropiadamente un medio legible por ordenador. Por ejemplo, si se transmite software desde un sitio web, servidor, u otra fuente remota usando un cable coaxial, cable de fibra óptica, par trenzado, línea de abonado digital (DSL), o tecnologías inalámbricas tales como infrarrojas, radio, y microondas entonces el cable coaxial, cable de fibra óptica, par trenzado, DSL, o tecnologías inalámbricas tales como infrarrojas, radio, microondas se incluyen en la definición de medio. Disco (del inglés "disk" o "disc"), tal y como se usa en este documento, incluye disco compacto (CD), disco láser, disco óptico, disco digital versátil (DVD), disco flexible y disco blu-ray en donde discos (del inglés "disks") usualmente reproducen datos magnéticamente, mientras discos (del inglés "discs") reproducen datos ópticamente con láser. Se deberían incluir combinaciones de los anteriores dentro del alcance de medios legibles por ordenador.

La descripción previa de la divulgación se proporciona para permitir a cualquier experto en la materia hacer o usar la divulgación. Varias modificaciones de la divulgación serán automáticamente evidentes a aquellos expertos en la materia, y los principios genéricos definidos en este documento pueden aplicarse a otras variaciones sin alejarse del alcance de la divulgación. Por lo tanto, la divulgación no pretende limitarse a los ejemplos y diseños descritos en

este documento si no que se les debe otorgar el alcance más amplio consistente con los principios y características novedosas descritas en este documento.

REIVINDICACIONES

1. Un método de enviar datos en una red de comunicación que tiene estaciones base, que comprende:
 - 5 recibir (812) de una primera estación base un paquete para enviar a un terminal;
determinar (814) una cantidad de tiempo que el paquete ya ha esperado en una primera cola en la primera estación base;
 - 10 colocar (816) el paquete en una segunda cola en la segunda estación base,
estando el paquete colocado en un punto en la segunda cola para tener en cuenta la cantidad de tiempo que el paquete ya ha esperado en la primera cola;
 - y
 - enviar (818) el paquete al terminal cuando el paquete alcanza la cabeza de la segunda cola.

2. El método según la reivindicación 1, en el que colocar el paquete en la segunda cola comprende
 - 15 determinar el tiempo de encolado objetivo restante para el paquete en base a un tiempo de encolado objetivo para el paquete y la cantidad de tiempo que el paquete ya ha esperado en la primera cola, y colocar el paquete en el punto en la segunda cola determinado en base al tiempo de encolado objetivo para el paquete.

3. El método según la reivindicación 1, en el que el paquete se coloca al final de la primera cola cuando es recibido por la primera estación base.

4. El método según la reivindicación 1, en el que el paquete se coloca en un punto en la primera cola determinado en base a un nivel de prioridad del paquete.

5. El método según la reivindicación 1, en el que el paquete es retenido por la primera estación base hasta que el requisito de retraso del paquete ha pasado.

6. El método según la reivindicación 1, que comprende además:
 - 30 admitir el paquete en la primera cola en base a la prioridad del paquete y la carga de la primera estación base;
 - y
 - 35 admitir el paquete en la segunda cola en base a la prioridad del paquete y la carga de la segunda estación base.

7. El método según la reivindicación 6, en el que admitir está basado en al menos uno de atributos de Calidad de Servicio, estado de prioridad de los flujos de datos o sesiones de la primera estación base y la segunda estación base y estadísticas actuales de retraso de encolado de la primera estación base y la segunda estación base.

8. Un aparato de comunicación en una red de comunicación que tiene estaciones base, que comprende:
 - 45 medios para recibir de una estación base un paquete a enviar a un terminal, para determinar la cantidad de tiempo que el paquete ya ha estado esperando en una primera cola en la primera estación base, para colocar el paquete en una segunda cola en la segunda estación base, siendo el paquete colocado en un punto en la segunda cola para tener en cuenta la cantidad de tiempo que el paquete ya ha estado esperando en la primera cola y para enviar el paquete al terminal cuando el paquete alcanza la cabeza de la segunda cola.

9. El aparato según la reivindicación 8, en el que los medios están formados por al menos un procesador.

10. El aparato según la reivindicación 9, en el que el al menos un procesador está configurado para determinar el tiempo restante objetivo de encolado para el paquete en base un tiempo de encolado objetivo para el paquete y la cantidad de tiempo que el paquete ya ha estado esperando en la primera cola y colocar el paquete en el punto en la segunda cola determinado en base al tiempo restante objetivo de encolado para el paquete.

11. Un producto de programa de ordenador, que comprende:
 - 60 un medio legible por ordenador que comprende:
 - 65 código para hacer que al menos un ordenador lleve a cabo un método según una de las reivindicaciones 1 a 7 al ejecutarse.

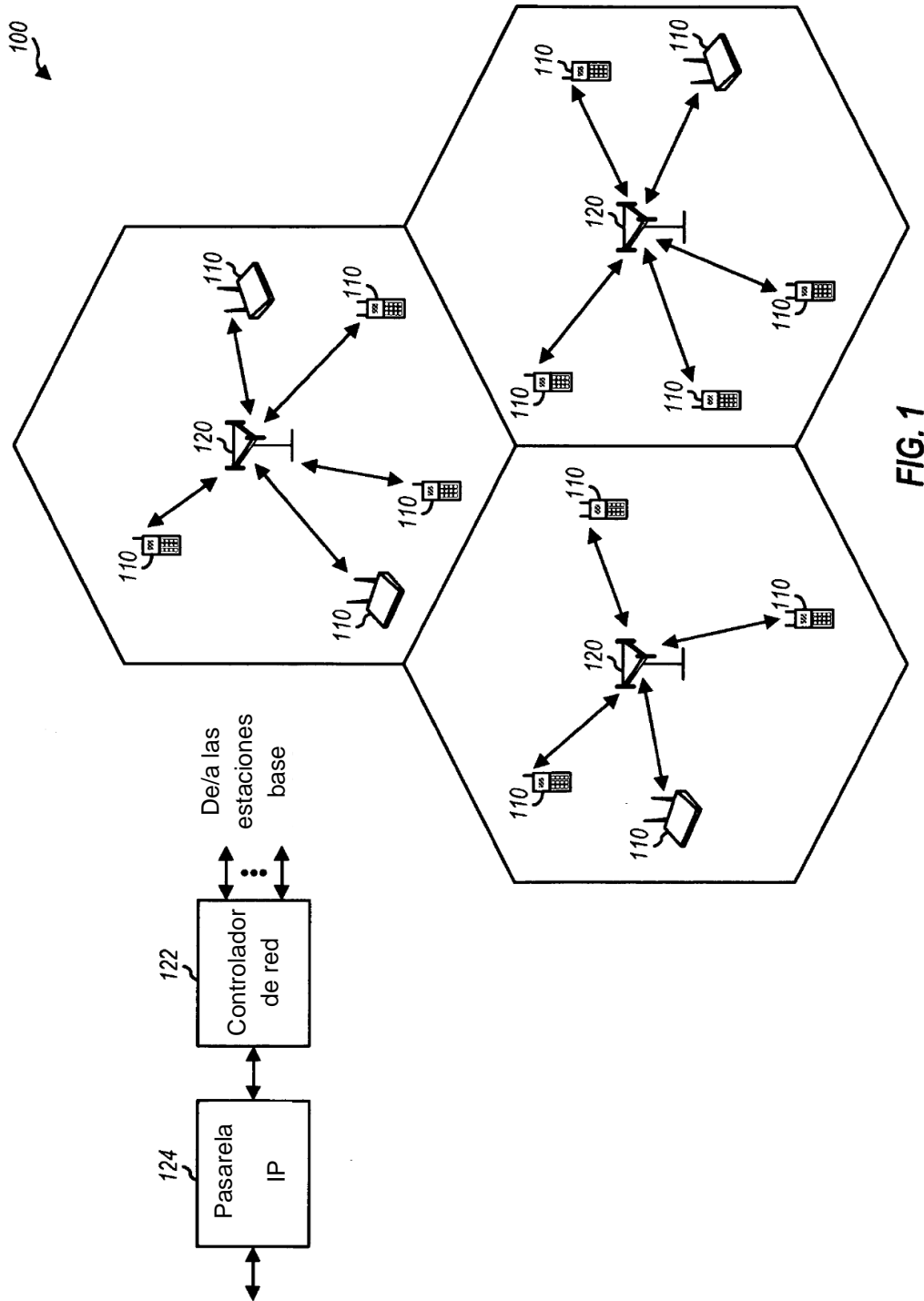


FIG. 1

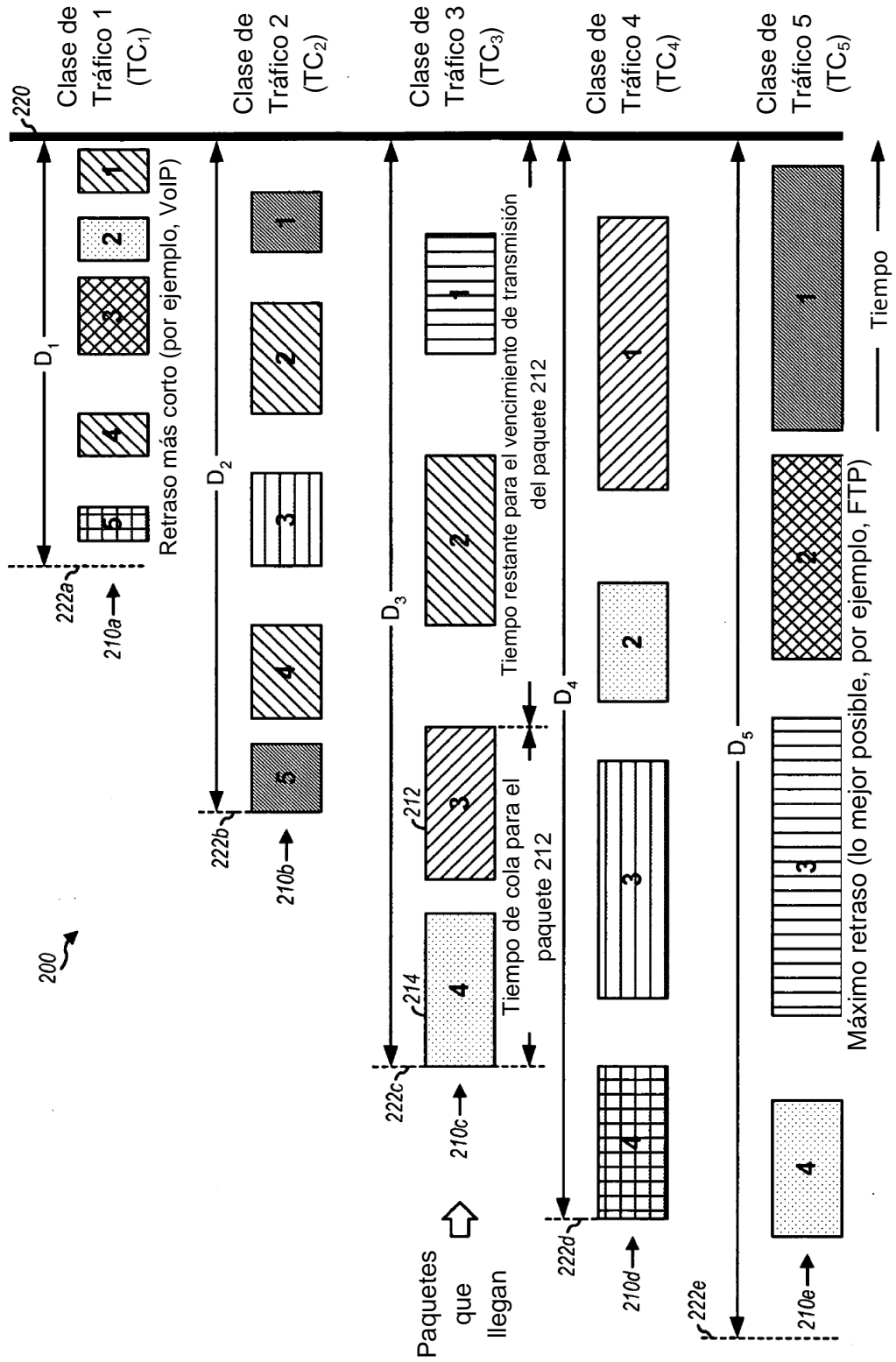


FIG. 2

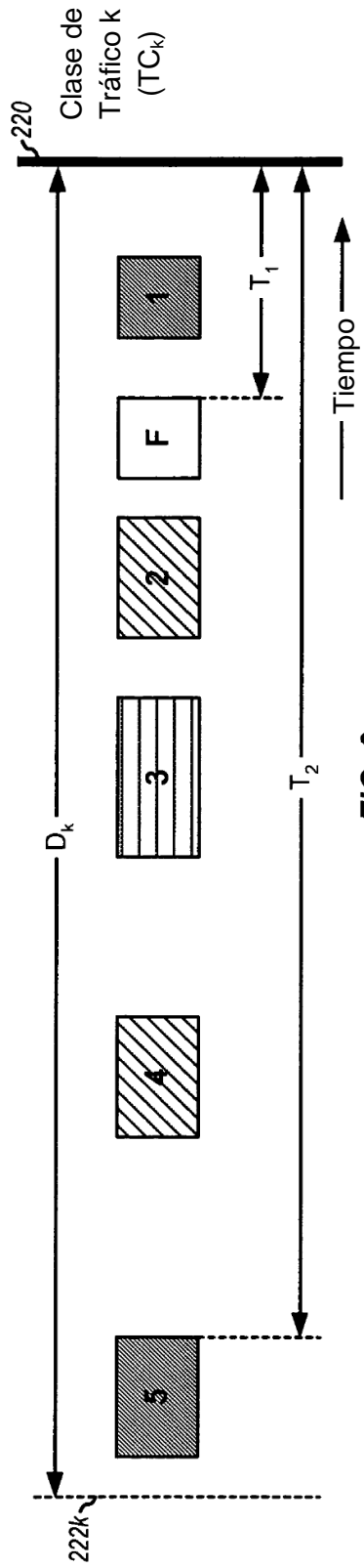


FIG. 3

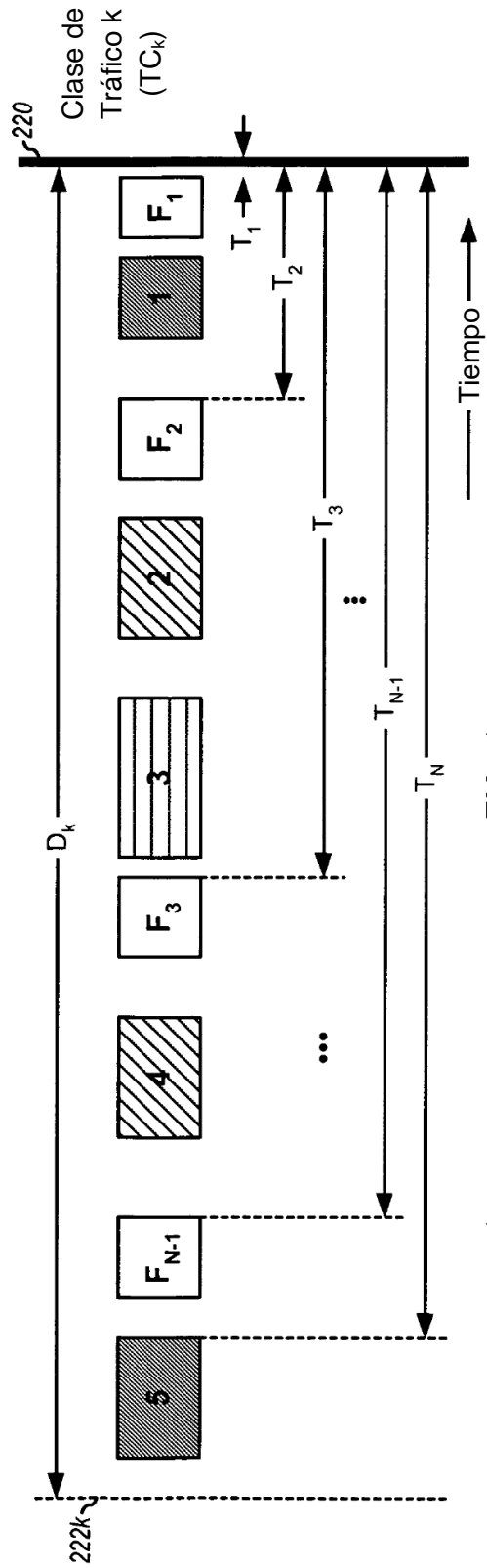


FIG. 4

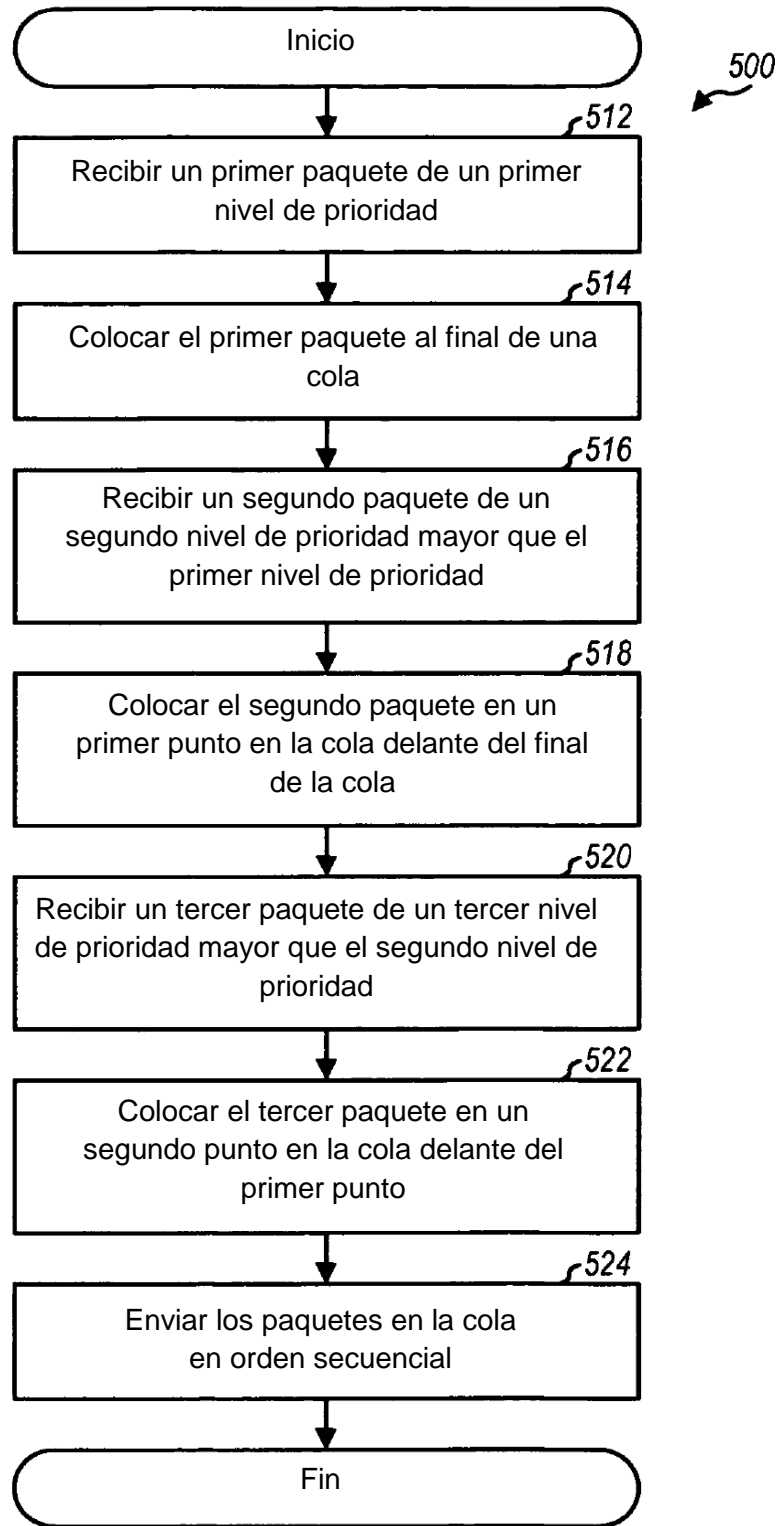


FIG. 5

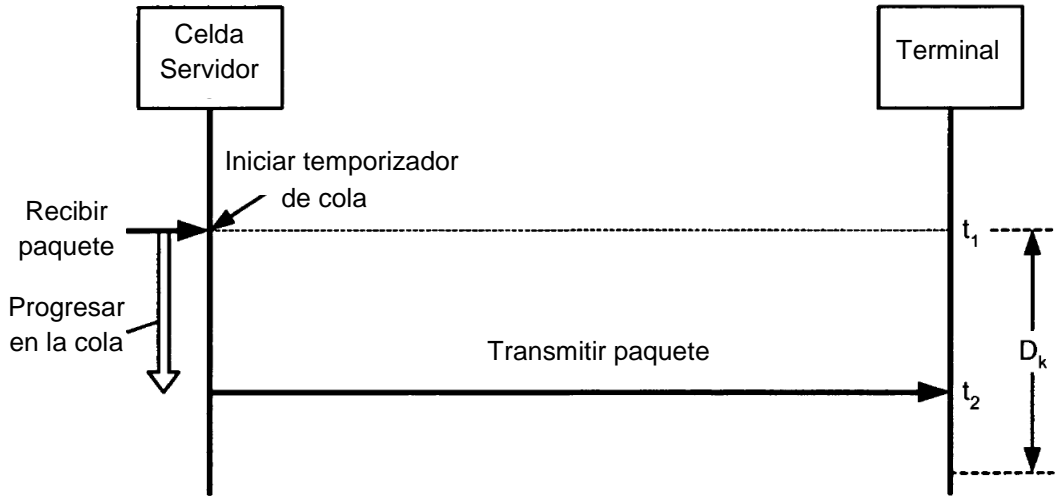


FIG. 6

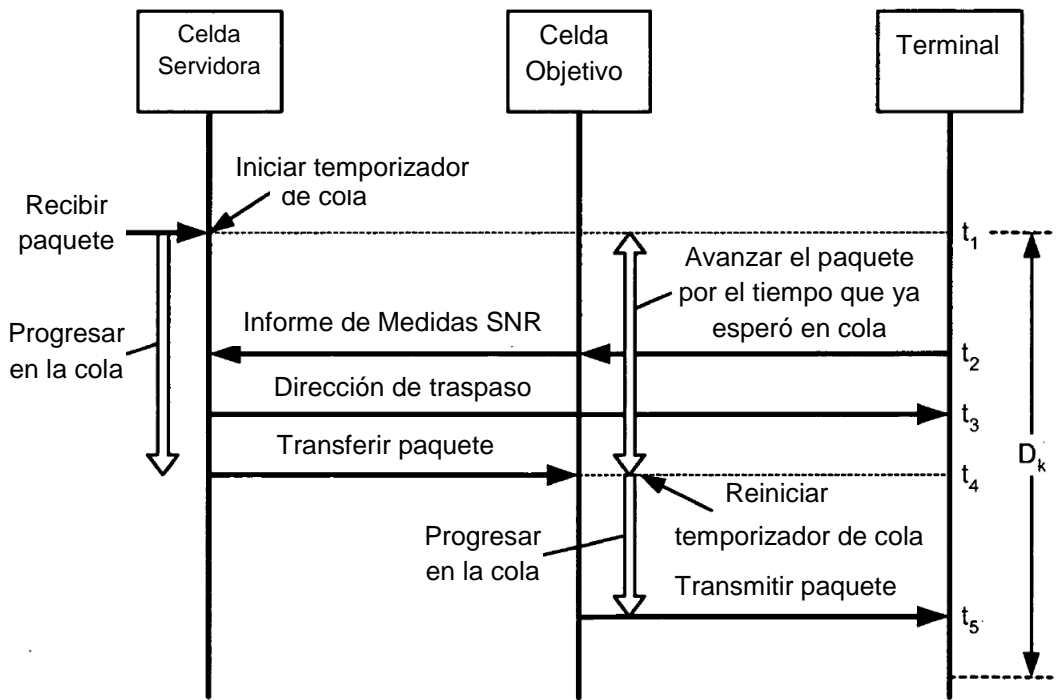


FIG. 7

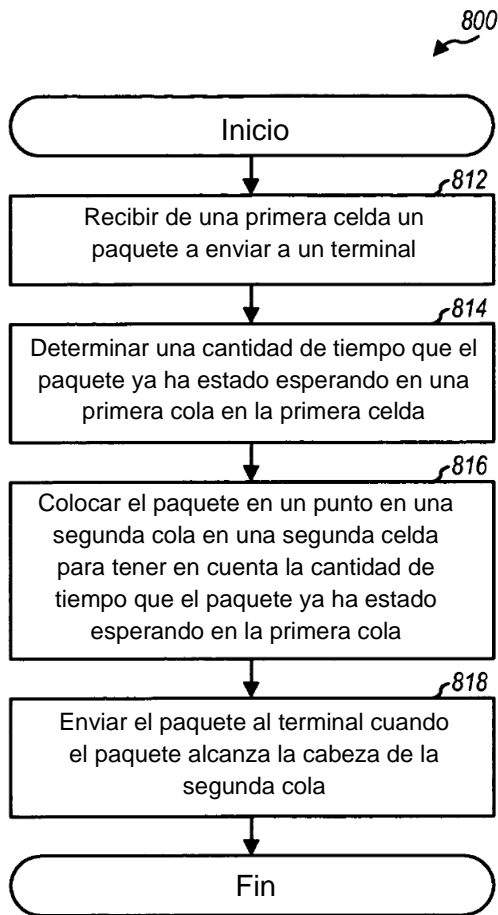


FIG. 8

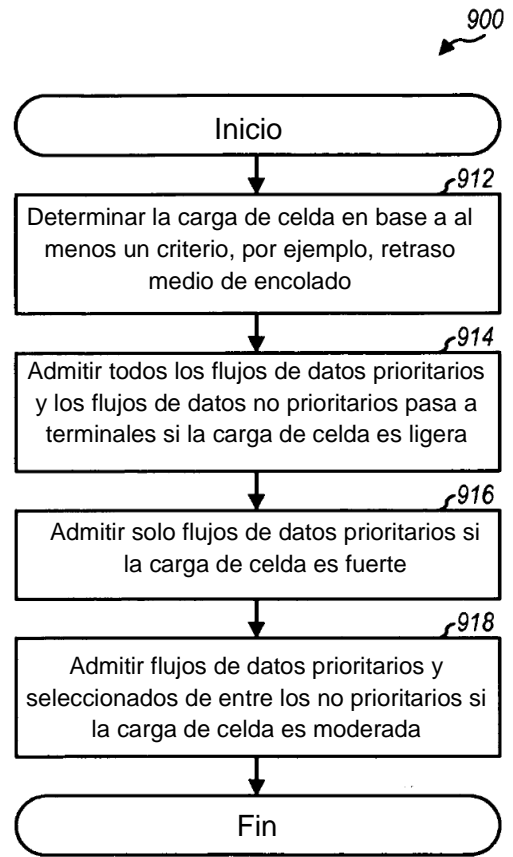


FIG. 9

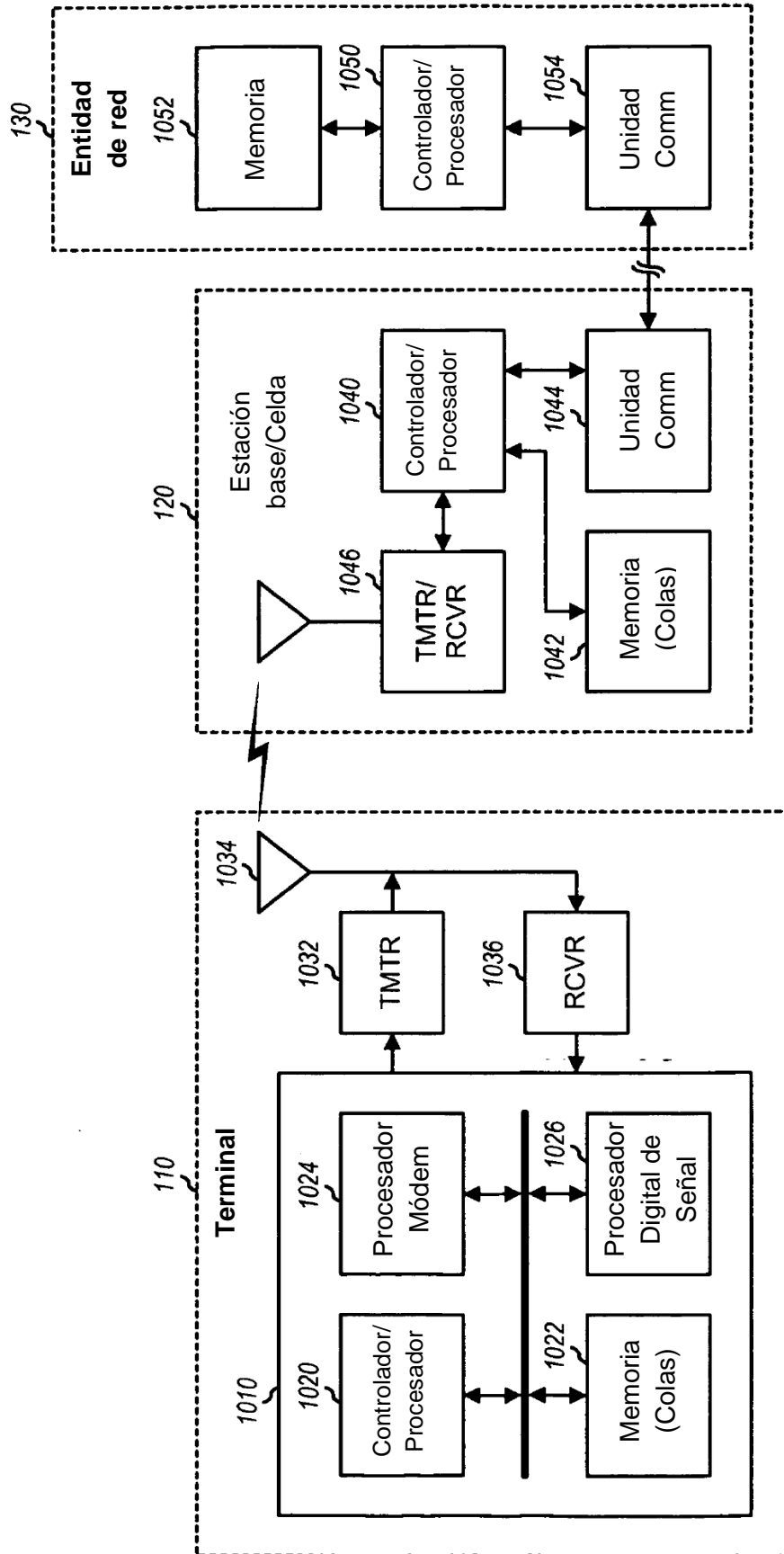


FIG. 10