

Diseño de redes de acceso en sistemas móviles UMTS con soporte de calidad de servicio

A. B. García, M. Álvarez-Campana, E. Vázquez, J. Berrocal
Dept. Ingeniería de Sistemas Telemáticos, Universidad Politécnica de Madrid
abgarcia@dit.upm.es

Resumen

El soporte eficiente de múltiples clases de tráfico con diferentes requisitos de calidad de servicio plantea nuevos problemas de ingeniería a la hora de diseñar una red de comunicaciones móviles de tercera generación (3G). El problema es más crítico en el caso de la red de acceso, donde habitualmente los recursos radio y de transmisión suelen ser escasos. En este artículo se aborda la problemática del diseño de la red de acceso radio UMTS (Universal Mobile Telecommunications Systems) desde una perspectiva práctica. El objetivo es proporcionar una visión de los principales parámetros y alternativas de diseño que se pueden plantear en un escenario de despliegue real.

Palabras clave

UMTS, Red de acceso, Calidad de Servicio, Modelos de tráfico, Dimensionado, ATM

I. INTRODUCCIÓN

A pesar de los recientes anuncios señalando la inminente llegada de los sistemas de comunicaciones móviles de 3G, la realidad ha venido a demostrar que son muchos los aspectos por resolver antes de acometer su despliegue a gran escala. A la hora de justificar el retraso, suele argüirse la carencia de terminales como el principal motivo. No obstante, existen otros factores a tener en cuenta como la relativa inmadurez de los estándares y el salto tecnológico trascendental que los sistemas 3G suponen con respecto a los actuales sistemas 2G. Esta observación tan tajante merece una explicación.

Centrándonos en el caso de UMTS, es de justicia reconocer el vertiginoso ritmo de trabajo desempeñado por el 3GPP (Third Generation Partnership Project), organismo responsable de elaborar las especificaciones de UMTS. Aún así, debido a lo ambicioso del proyecto y la extraordinaria complejidad que subyace bajo el mismo, son muchos los detalles que aún no han sido cubiertos con la suficiente

profundidad. A ello han de sumarse los numerosos aspectos de implementación que dejan abiertas las especificaciones del 3GPP (puesto que no forman parte de la misión de dicho organismo) y para los cuales no existe una experiencia consolidada en el sector. De entre todos estos aspectos, quizás el ejemplo más significativo sea el delicado tema de cómo conseguir el soporte eficiente de múltiples aplicaciones con los requisitos de calidad de servicio (QoS, Quality of Service) que exige cada una de ellas.

La concepción del sistema UMTS viene condicionada por los requisitos de los servicios que ofrecerá esta nueva generación de redes móviles. Un primer aspecto a tener en cuenta es la gran diversidad de servicios que deberán ser soportados [1], algunos ni siquiera conocidos hoy en día. El soporte de aplicaciones de voz, datos, vídeo y, en general, servicios multimedia, con requisitos muy variados en lo que respecta al ancho de banda necesario y la tolerancia a factores como el retardo o las pérdidas, lleva a la necesidad de dotar a UMTS de mecanismos de QoS. El problema no es trivial, especialmente cuando se plantea la necesidad de obtener una utilización eficiente de los recursos de la red. Este objetivo es esencial en aquellas partes de la red donde habitualmente es previsible la escasez de recursos, como es el caso del interfaz radio y de la red de acceso (UTRAN, UMTS Terrestrial radio Access Network).

La mayoría de los trabajos publicados sobre UTRAN se centran en aspectos de planificación radio (por ejemplo, [2], [3] y [4]). La gestión eficiente de los recursos radio es de capital importancia en toda red celular, y más aún si cabe en UMTS debido a la utilización de WCDMA [5] (Wideband Code Division Multiple Access). Aún así, no ha de olvidarse que la red de acceso también abarca la infraestructura de transmisión entre las estaciones base y el núcleo de red, siendo este tramo justamente uno de los más costosos en la redes celulares. El objeto del presente artículo es precisamente abordar la problemática asociada al dimensionado de los recursos de transmisión de la red de acceso UTRAN.

El resto del artículo se ha organizado de la siguiente manera. La sección II proporciona una visión general de los principios de arquitectura de UMTS, haciendo especial énfasis en UTRAN. En el apartado III se analizan los principales factores a considerar en el dimensionado de UTRAN, prestando especial atención al modelado de tráfico, los requisitos de QoS, y las alternativas de multiplexión de tráfico. En la sección IV se analizan las alternativas existentes para el despliegue de la red de transmisión en términos de interfaces de transmisión y topologías de red. Finalmente, en el último apartado se resumen las principales conclusiones de nuestro trabajo y las líneas de trabajo futuras.

II. FUNDAMENTOS DE UMTS

A. Arquitectura UMTS

En la figura 1 se representa la arquitectura básica de una red UMTS [6], compuesta por sus tres partes fundamentales: los equipos de usuario, la red de acceso y el núcleo de red.

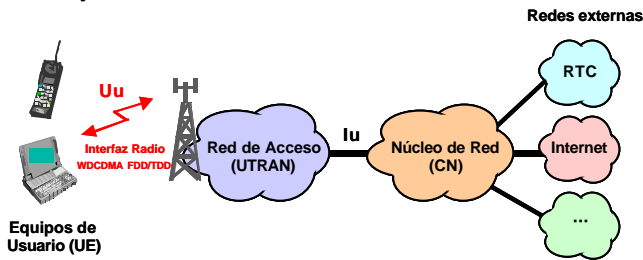


Fig. 1: Arquitectura genérica UMTS

Los equipos de usuario acceden a la red a través del interfaz radio (Uu), basado en WCDMA. La red de acceso radio (UTRAN) se encarga de transportar el tráfico de usuario (voz, datos, señalización móvil-red) hasta el núcleo de red (CN, Core Network), con el que se comunica a través del interfaz Iu. Dentro del núcleo de red se encuentran los recursos de conmutación y transmisión necesarios para completar el trayecto de la comunicación hacia el abonado remoto, que puede pertenecer a la red UMTS o a una red externa (RTC, Internet, ...).

El núcleo de red en UMTS se plantea como la evolución del existente en las actuales redes 2G basadas en GSM/GPRS. En la primera fase de normalización (Release 99 [7]), se propone la reutilización de la infraestructura disponible en dichas redes. La evolución del CN propiamente dicha se deja para fases posteriores (Release 4 [8] y Release 5 [9]).

La red de acceso en UMTS, por el contrario, difiere desde el primer momento con respecto a las

redes 2G. Con objeto de acomodar de manera flexible y eficiente los distintos tipos de aplicaciones posibles en UMTS, se recurre al empleo de técnicas de conmutación de paquetes. Así, la Release 99 establece el empleo de ATM (Asynchronous Transfer Mode) como tecnología de transporte en UTRAN¹ [11]. La selección inicial de ATM se justifica por el hecho de ser una de las tecnologías más flexibles y maduras (al menos en la fecha en la que se tomó la decisión) para el despliegue de redes multiservicio con QoS.

En la figura 2 se muestra la arquitectura de UTRAN, en la que pueden observarse los elementos que la componen y los interfaces definidos entre ellos.

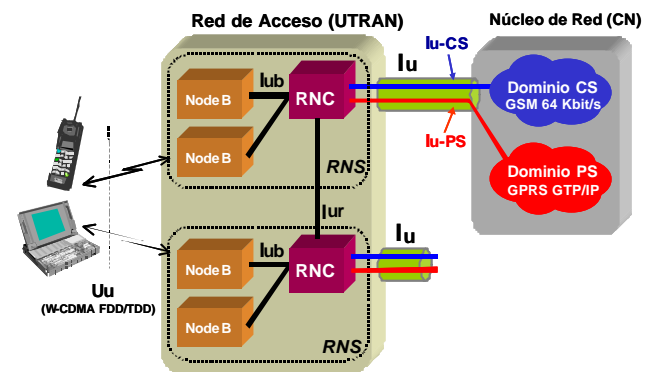


Fig. 2: Arquitectura de UTRAN

La red de acceso UMTS consta de uno o más subsistemas RNS (Radio Network Subsystem). Cada RNS cubre un conjunto de células UMTS, siendo responsable de la gestión de los recursos asociados a ellas. Un RNS está formado por un controlador RNC (Radio Network Controller) y un conjunto de estaciones base (Nodos-B).

Dentro de la red radio se definen dos tipos de interfaces: el interfaz Iub entre cada Nodo-B y el RNC que lo controla y el interfaz Iur entre RNCs. Este último interfaz, sin equivalente en las redes 2G, permite la comunicación directa entre RNCs para el soporte de trasposos suaves (Soft-Handover) entre estaciones base pertenecientes a distintos RNCs.

La red radio también posee dos interfaces externos: el interfaz radio Uu, basado como ya se ha dicho en WCDMA, y el interfaz Iu con el núcleo de red. Este último se subdivide lógicamente en dos interfaces: Iu-CS hacia el dominio de conmutación de circuitos e Iu-PS hacia el dominio de conmutación de paquetes.

¹ Actualmente el 3GPP trabaja en la definición de una solución alternativa para UTRAN basada en IP[10].

B. Protocolos UTRAN

En la figura 3 se muestran, de manera simplificada, los diversos bloques que componen la arquitectura de protocolos de UTRAN [11]. En dicha figura se han omitido algunos detalles por motivos de claridad (por ejemplo, no se muestra el interfaz Iur).

Los protocolos de UTRAN se estructuran en dos capas: la capa de red radio (RNL, Radio Network Layer) y la capa de red de transporte (TNL, Transport Network Layer). Esta descomposición tiene como objetivo aislar las funciones que son específicas del sistema UMTS (encuadradas dentro de RNL), de aquellas otras que dependen de la tecnología de transporte utilizada (ubicadas dentro de TNL).

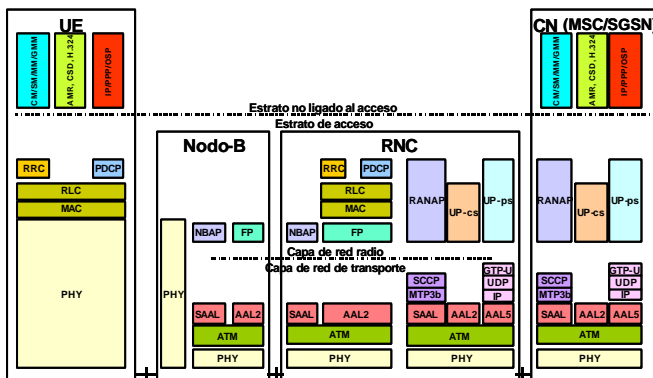


Fig. 3: Arquitectura de protocolos de UTRAN

Como se ha indicado, en la Release 99 la capa TNL se articula en torno al uso de ATM. La información que se transporta sobre ATM es básicamente de dos tipos:

- Información móvil-red: Se trata de la información, señalización o tráfico de usuario, que intercambian entre sí los móviles y el nodo de entrada al núcleo de red. Este nodo será un MSC (Mobile Switching Center) en caso de una comunicación con el dominio CS, o un SGSN (Serving GPRS Support Node) en caso del dominio PS.
- Señalización UTRAN: Se trata de información intercambiada entre nodos-Bs y RNCs (protocolo NBAP, Node-B Application Part), entre RNCs y Núcleo de Red (protocolo RANAP, Radio Access Network Protocol) y entre RNCs (RNSAP, Radio Network Subsystem Application Part).

En la figura 4 se muestran con detalle los protocolos involucrados en el plano de usuario de los interfaces Iub, Iu-CS e Iu-PS. Los protocolos de adaptación ATM utilizados en cada caso son AAL2 para Iub e Iu-CS (también para Iur, aunque no se muestre en la figura) y AAL5 para Iu-PS.

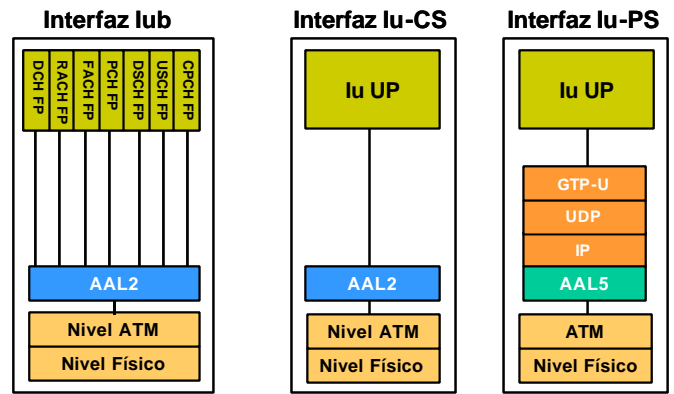


Fig. 4: Protocolos UTRAN del plano de usuario

Desde la perspectiva del plano de usuario, el interfaz Iub puede considerarse como una prolongación del interfaz radio. De hecho, si se observa con detenimiento la figura 3 se puede observar que los protocolos radio (RRC/RLC/MAC) terminan en el RNC. El transporte de las tramas radio (MAC) entre un nodo-B y su correspondiente RNC se basa en la utilización del protocolo AAL2, de modo que cada canal radio emplea una miniconexión AAL2.

El protocolo AAL2, definido en la recomendación Q.363.2 [12] del ITU-T, permite la multiplexión eficiente de varios flujos de datos (hasta 248 ‘miniconexiones’) sobre un mismo circuito virtual ATM.

El funcionamiento del protocolo AAL2 se ilustra en la figura 5. En un primer paso, cada flujo de datos se convierte en un flujo de paquetes CPS (Common Part Sublayer), genéricamente denominados minicélulas AAL2, con una cabecera de tres octetos y una carga útil de 1 a 45 (ó 64) octetos (el tamaño que más convenga en cada caso). Los flujos de minicélulas resultantes son multiplexados (entrelazados) a continuación formando bloques de 47 octetos. Añadiendo a cada bloque un octeto adicional (un puntero que facilita la recuperación ante pérdidas en recepción), se forman los segmentos de 48 octetos (CPS PDUs) con los que se rellena la carga útil de las células ATM de la conexión.

Sobre el interfaz Iu, el protocolo AAL utilizado depende de si se trata de una comunicación modo circuito (Iu-CS), en cuyo caso se usa AAL2, o una comunicación modo paquete (Iu-PS), en cuyo caso se usa AAL5. En el caso del interfaz Iu-CS, el empleo de AAL2 constituye la mejor opción para el transporte eficiente de flujos de datos modo circuito (paquetes de voz, datos modo circuito, fax, ...). En el caso de Iu-PS, la solución adoptada consiste en el empleo de un túnel IP (protocolo GTP, GPRS Tunneling Protocol) por

sesión de datos, recurriéndose al protocolo AAL habitual para el transporte de IP, esto es, el protocolo AAL5.

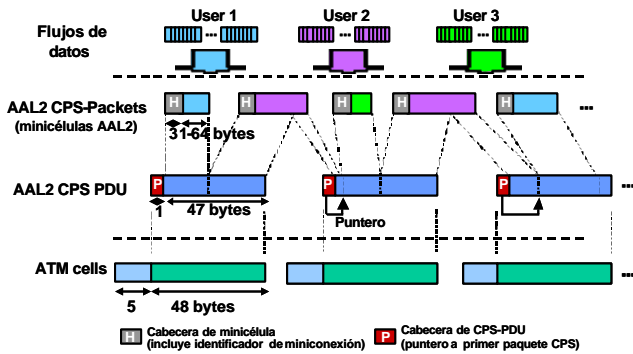


Fig. 5: Protocolo de adaptación ATM tipo 2 (AAL2)

C. Calidad de servicio en UMTS

El soporte de QoS en UMTS se basa en la arquitectura jerárquica definida en la especificación TS 23.107 [13], representada en la figura 6. La calidad de servicio extremo a extremo se sustenta en la calidad que proporcionan los servicios portadores subyacentes: el servicio portador local, el servicio portador UMTS y el servicio portador externo. Esta primera descomposición tiene como objeto no limitar innecesariamente los equipos terminales a emplear (ej. un PC) y las posibles redes destino (ej. Internet) con las que comunicarse. Es por este motivo que la especificación deja fuera del ámbito de la calidad de servicio de UMTS a los servicios portadores local y externo, centrándose exclusivamente en la normalización del servicio portador UMTS.

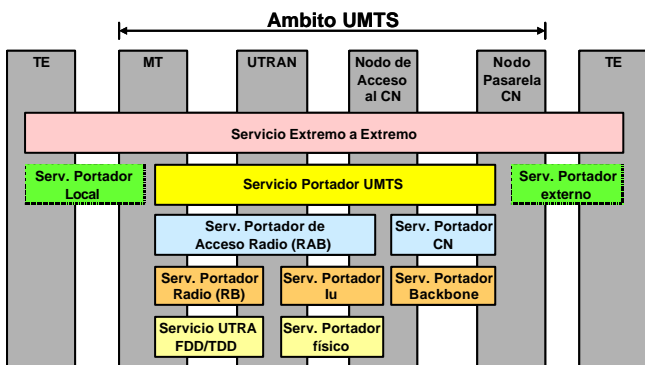


Fig. 6: Arquitectura de calidad de servicio UMTS

En una segunda descomposición, el servicio portador UMTS se sustenta en la QoS que le proporciona el servicio portador de acceso radio (RAB) y el servicio portador del núcleo de red. El primero abarca el trayecto comprendido entre la

terminación móvil (el terminal radio del usuario) y el nodo de acceso al núcleo de red (un MSC o un SGSN, según el caso), pasando a través del interfaz radio (Uu), la red de acceso radio (UTRAN), y el interfaz Iu. El concepto de RAB es determinante en la provisión de servicios UMTS con distintos perfiles de calidad de servicio, puesto que implica la utilización de recursos sobre el interfaz radio y la red de acceso, precisamente donde se presentan las mayores limitaciones de ancho de banda.

El servicio portador del núcleo de red, por su parte, abarca el trayecto comprendido entre el nodo de acceso (MSC o SGSN) hasta el nodo pasarela (GMSC o GGSN) hacia la red destino de interés (ej. RTC o Internet). La QoS en este trayecto se apoya en la que proporciona el correspondiente backbone (de circuitos o de paquetes).

Desde el punto de vista de los requisitos de QoS, y atendiendo fundamentalmente al criterio de su tolerancia al retardo, en UMTS se han definido cuatro clases de tráfico:

- *Conversacional*. Dentro de esta clase se encuadran la comunicaciones de audio y vídeo en tiempo real entre personas. Este tipo de comunicaciones se caracteriza por exigir un retardo extremo a extremo muy reducido, con objeto de que los usuarios no pierdan la sensación de interactividad. Ejemplos de aplicaciones conversacionales son la telefonía, la videotelefonía o la videoconferencia.
- *Streaming (afluente)*. En esta categoría se incluyen las aplicaciones que permiten a los usuarios la descarga de contenidos multimedia (audio y video clips) para su reproducción on-line, con una sensación que, sin serlo, se aproxima a la de tiempo real. El hecho de que la transferencia de información sea unidireccional permite retrasar el instante de inicio de la reproducción posibilitando el empleo de "buffers" relativamente grandes en el extremo receptor para absorber las fluctuaciones de retardo. Ello permite relajar significativamente los requisitos de retardo con respecto a los servicios conversacionales.
- *Interactivo*. Esta clase de tráfico engloba las aplicaciones de acceso remoto a información en la modalidad online, donde el usuario (o una máquina) envía peticiones hacia el equipo remoto esperando que éste le devuelva las respuestas en un tiempo razonablemente reducido. Ejemplos de aplicaciones bajo esta categoría son la navegación web, las consultas a bases de datos o el acceso remoto a ordenadores (telnet).

- *Background (diferible)*. Esta última clase da cabida a un número considerable de aplicaciones de datos en las que el usuario no exige una respuesta inmediata por parte de la red, admitiendo retardos que oscilan desde unos pocos segundos hasta incluso varios minutos. Ejemplo de tales aplicaciones son el correo electrónico o la descarga de ficheros, por citar algunas.

La formalización del concepto de calidad de servicio en UMTS recurre a la clásica solución consistente en la definición del correspondiente conjunto de parámetros o atributos. Así, en la especificación TS 23.107 se define una docena de atributos de calidad de servicio, así como el rango de valores que éstos pueden tomar. Como se observa en la tabla 1, no todos los atributos son aplicables a las cuatro clases de tráfico definidas.

Tabla 1: Atributos de calidad de servicio en UMTS

	Conversacional	Streaming	Interactivo	Background
Tasa de bit máxima	x	x	x	x
Entrega ordenada	x	x	x	x
Tamaño máximo SDU	x	x	x	x
Información de formato de SDU	x	x		
Ratio de SDUs erróneos	x	x	x	x
Ratio de error de bit residual	x	x	x	x
Entrega de SDUs erróneas	x	x	x	x
Retardo de transferencia	x	x		
Tasa de bit garantizada	x	x		
Prioridad de tráfico			x	
Prioridad de asignación/retención	x	x	x	x

La especificación TS 23.107 propone también una arquitectura donde se identifican el tipo de funciones requeridas para el soporte de QoS en una red UMTS. Se trata de una modelo genérico, que deja libertad absoluta en lo referente a los detalles de implementación. Puede decirse, en definitiva, que la arquitectura para el soporte de QoS definida por el 3GPP es incompleta, siendo muchos los aspectos que deja sin concretar. En los siguientes apartados se abordan algunas alternativas para la solución de dichos aspectos.

III. DIMENSIONADO DE UTRAN

El empleo de una tecnología de conmutación de paquetes, como lo es ATM, en UTRAN supone un cambio trascendental en relación con los actuales sistemas 2G, basados en el paradigma de conmutación de circuitos. Este cambio afecta a la metodología de dimensionado, puesto que los clásicos modelos de Erlang utilizados en las redes 2G no son de aplicación directa en el nuevo escenario.

El dimensionado de UTRAN debe, más bien, inspirarse en los modelos desarrollados dentro del ámbito de las redes multiservicio con garantías de QoS. Desafortunadamente, y a pesar de los logros alcanzados durante las últimas décadas, no existe aún una experiencia suficientemente consolidada en este terreno. Un tratamiento riguroso del problema requiere la consideración de múltiples factores, entre los que cabe considerar los siguientes:

- La carga de tráfico a soportar en cada interfaz, teniendo en cuenta las distintas contribuciones (tráfico de usuario, señalización, gestión, sobrecarga de cabeceras etc.)
- Los parámetros de tráfico de fuentes (tasa de pico, tasa media, factor de actividad, etc.)
- Los requisitos de QoS (tolerancia a pérdidas y retardos) asociados a cada tipo de tráfico.
- Las diferentes estrategias de multiplexión de tráfico sobre ATM y el ahorro de ancho de banda que éstas pueden proporcionar (ganancia por multiplexión estadística).
- Las políticas de gestión de tráfico (control de admisión, priorización de tráfico, control de congestión, ...)

La consideración detallada de todos estos factores puede conducir a modelos analíticos muy complejos. Es por ello que en la práctica suele recurrirse a modelos simplificados (modelos con varios estados de actividad, modelos de fluidos, etc.) o al empleo de técnicas de simulación. En los siguientes apartados se proporcionan una serie de pautas acerca de cómo abordar en la práctica los diferentes factores arriba mencionados.

A. Carga de tráfico

Uno de los puntos de partida esenciales en el dimensionado de cualquier red de comunicaciones es el conocimiento acerca de la demanda de tráfico a cursar. En el caso de UMTS, su naturaleza multiservicio requiere el desglose de las previsiones de tráfico por servicio o, al menos, por clases de tráfico.

Las consideraciones de tráfico a aplicar en UTRAN deben estar en consonancia con las utilizadas en el proceso de planificación radio. El resultado principal de este proceso es la determinación del número de Nodos-B necesarios para cubrir una determinada zona geográfica según criterios de cobertura, previsiones de tráfico y grado de servicio. Puesto que, por otro lado, el cuello de botella en las redes celulares normalmente está en el interfaz radio, cabe plantearse como criterio inicial para el

dimensionado de UTRAN su capacidad para cursar todo (o casi todo) el tráfico que el interfaz aire pueda admitir.

Además del tráfico debido a las aplicaciones utilizadas por los usuarios, es necesario tener en cuenta otros: señalización móvil-red, señalización UTRAN (NBAP, RNSAP, RANAP), y tráfico de gestión (OAM, Operation and Maintenance). En el caso de los interfaces Iub e Iur también debe considerarse la contribución adicional debida a trasposos suaves. Habitualmente estas contribuciones suelen considerarse mediante la aplicación de un porcentaje sobre el tráfico de usuario o una aportación fija por elemento (Nodo-B o RNC, según el caso).

A la hora de efectuar los cálculos es necesario tener en cuenta también las sobrecargas debidas a las cabeceras de los distintos protocolos UTRAN.

B. Modelos de Tráfico

Existen diversos estudios que proponen o utilizan modelos de tráfico en redes de acceso de sistemas 3G (referencias [14] a [20]). Sin embargo, no se han encontrado, en la literatura revisada, estudios en los que se aborden de manera conjunta las cuatro clases de tráfico definidas en UMTS (conversacional, afluente, interactivo y diferible). En este apartado se propone un modelo de tráfico que se estima apropiado para el dimensionado de UTRAN. Se trata de un modelo aplicable a las cuatro clases de tráfico (aunque los parámetros concretos variarán de una a otra), y capaz de representar la tasa variable que muchas fuentes de tráfico exhiben en la realidad, posibilitando la consideración y evaluación del efecto de la multiplexión estadística en la fase de dimensionado.

Puesto que en nuestro caso estamos interesados en el estudio de la red de transporte ATM de UTRAN, los modelos de tráfico considerar no sólo el comportamiento de la aplicación en sí, sino las características relevantes de los protocolos de la capa de red radio. En el caso de aplicaciones asimétricas, se supondrá implícitamente que se modela el sentido descendente de la comunicación, por ser éste el que presentará en un caso típico más cantidad de tráfico.

El modelo que proponemos incluye tres niveles, concepto que está en consonancia con varios de los estudios de caracterización de tráfico revisados (por ejemplo [18], [20] y [21]). Tal como se indica en la figura 7, los tres niveles considerados son los siguientes:

- *Nivel de sesión.* Mediante este nivel se modela el inicio y final de una sesión de usuario (por ejemplo, una llamada de voz, o una sesión de

navegación Web). A este nivel los parámetros relevantes son los patrones de llegada y de duración de las sesiones.. En caso de que, como resultado de la fase de planificación radio, se disponga de información acerca del número de usuarios activos simultáneamente por cada Nodo-B, este nivel podría abstraerse y suponer, a la hora de dimensionar UTRAN, que todos los usuarios se encuentran dentro de una sesión activa.

- *Nivel de ráfaga.* Cuando un usuario se encuentra activo, dentro de una sesión, su patrón de generación de tráfico se modela mediante dos estados (alto y bajo) con características diferenciadas de generación de paquetes. Los datos necesarios para la completa caracterización de este nivel son las distribuciones estadísticas de la duración de cada uno de los dos estados.
- *Nivel de paquete.* A este nivel se especifica, dentro de cada uno de los dos estados del nivel anterior, el proceso estadístico de generación de paquetes. Las distribuciones del tiempo entre llegadas de paquetes y del tamaño de los paquetes son los dos aspectos significativos.

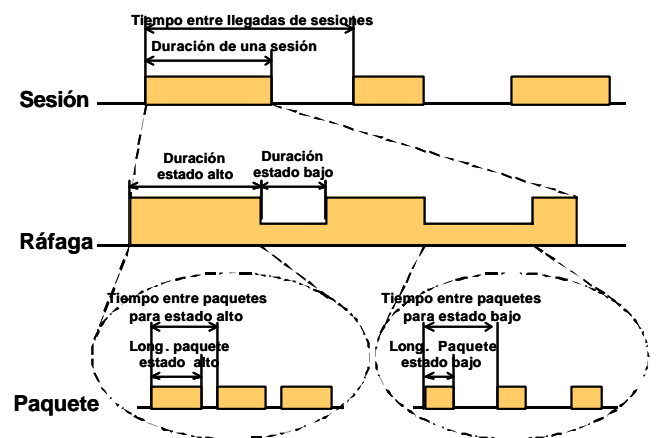


Fig. 7: Modelo de fuente en tres niveles

Para la concreción de los parámetros relevantes de cada clase de tráfico, se propone la elección de una aplicación lo suficientemente significativa de cada tipo. Así, por ejemplo, haciendo uso de la información contenida en algunas de las especificaciones del 3GPP (entre ellas TS 23.107 [13] y TS 22.105 [22]), cabría elegir las siguientes aplicaciones representativas: telefonía (código AMR [23], Adaptive Multirate) para la clase conversacional, video streaming para la clase afluente, navegación Web para la interactiva, y correo electrónico como ejemplo de tráfico diferible.

A título ilustrativo, en la tabla 2 se sugieren posibles valores para los parámetros de tráfico de voz y navegación Web, aplicables a los interfaces Iub e Iur (incluyendo la sobrecarga de todos los protocolos por encima de AAL2).

Tabla 2: Parámetros de caracterización de tráfico

	Voz	Web
Nivel de ráfaga		
Duración estado ALTO	exponencial media: 3 s.	geométrica media: 0,19 s.
Duración estado BAJO	exponencial media: 3 s.	geométrica media: 412 s.
Nivel de paquete		
Tiempo entre paquetes (estados ALTO y BAJO)	constante valor: 20 ms.	constante valor: 40 ms.
Tamaño de paquete en estado ALTO	constante valor: 40 octetos.	constante valor: 325 octetos.
Tamaño de paquete en estado BAJO	constante valor: 13 octetos.	constante valor: 0 octetos.

Los valores de la tabla han sido calculados a partir de las indicaciones presentes en los documentos del 3GPP, y las recomendaciones de ETSI en su informe técnico TR 101 112 [21]. El nivel de sesión no se modela asumiendo que ya ha sido considerado en la fase de planificación radio.

C. Parámetros de QoS

Cada clase de tráfico impone un conjunto de requisitos de QoS para cuya satisfacción la red de transporte debe ser cuidadosamente dimensionada en cada uno de sus tramos. De todos los parámetros de QoS que las especificaciones de UMTS definen, dos de los que más claramente influyen en el dimensionado de UTRAN son el ratio de errores y el retardo máximo de transferencia (este último para las dos clases de tiempo real: conversacional y afluente).

Los valores de partida para los dos parámetros de QoS seleccionados pueden derivarse a partir de las indicaciones para el servicio portador RAB definidas en TS 23.107 [13]. Estos parámetros deben ser traducidos o adaptados a requisitos parciales en cada uno de los interfaces de la red de acceso, a fin de obtener valores que puedan ser directamente utilizados para el dimensionado de los mismos.

Así, para el retardo de transferencia, el informe técnico TR 25.853 [24] proporciona una buena base sobre la que establecer la metodología de adaptación ya que muestra un desglose de los retardos involucrados en el servicio RAB, para los servicios de

tiempo real. Nuestra propuesta consiste en utilizar dichos valores de retardo adaptándolos al caso concreto de estudio cuando sea oportuno (por ejemplo, teniendo en cuenta el número de conmutadores ATM existentes en cada enlace), salvo en lo referente a los retardos de multiplexión y de-multiplexión. Esta exclusión se debe a que dichos retardos dependen de la mezcla concreta de tráfico, la capacidad de los enlaces, el tipo de circuitos ATM, y las estrategias de multiplexión de tráfico utilizadas.

El límite en el retardo de multiplexión y de-multiplexión en cada enlace será el que, junto con el resto de los retardos estimados en el informe del 3GPP, provoque que el retardo total en el servicio RAB sea el máximo permisible dependiendo de la clase de tráfico. A la hora de adaptar el retardo admisible se debe tener en cuenta la posible utilización de mecanismos de recuperación de errores (mediante retransmisiones) por parte del protocolo RLC (Radio Link Control), entre la terminación móvil y el RNC que le da servicio. Este mecanismo no se utiliza en el caso de la clase conversacional, donde los límites de retardo son muy estrictos. Sin embargo, en [13] se indica que los retardos permisibles para las portadoras de tipo *afluente*, en determinados márgenes, podrían permitir la presencia de retransmisiones en RLC.

En cuanto al segundo de los parámetros de QoS considerados, el ratio de errores, el objetivo es encontrar el valor de CLR (Cell Loss Ratio) a respetar en cada interfaz, puesto que se trata de una red ATM. Partiendo de la tasa máxima de errores permisible en el servicio RAB (TS 23.107) se deben ir realizando las adaptaciones oportunas teniendo en cuenta los múltiples factores que entran en juego, entre los que cabe destacar los siguientes:

- En ocasiones se debe desglosar una determinada tasa de errores en dos tramos físicos adyacentes. En este caso, si se supone que las probabilidades de error son mucho menores que uno (único caso en que el funcionamiento del sistema será de alguna utilidad), la probabilidad conjunta será aproximadamente la suma de las de cada tramo.
- En caso de que un protocolo realice segmentación y/o concatenación de SDUs (Service Data Units) de nivel superior, un error o una pérdida de una PDU (Protocol Data Unit) del protocolo provoca en general la pérdida de un número mayor de SDUs de la capa usuaria. En la figura 8 se puede ver mediante un ejemplo este efecto, típico por ejemplo del procesado llevado a cabo por la capa AAL2 para la transmisión de SDUs sobre células ATM.

- Evidentemente, los mecanismos de detección o corrección de errores utilizados permiten relajar los requisitos de pérdidas a imponer a capas inferiores. En el caso del servicio RAB, la posible utilización de retransmisiones en RLC debe ser tenida en cuenta a este respecto.

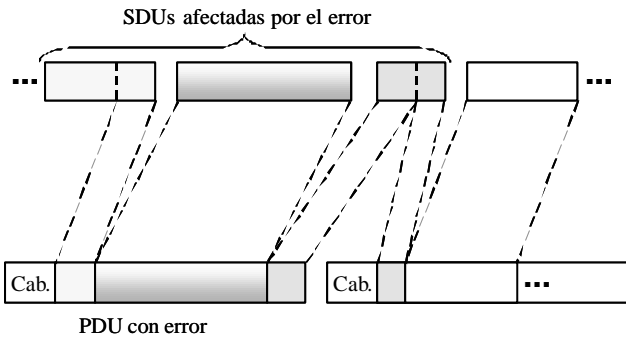


Fig. 8: Efecto de la segmentación y concatenación de unidades de datos sobre la tasa de pérdidas

D. Estrategias de Multiplexión y Tipos de Conexiones

A partir de lo expuesto en los apartados precedentes, puede deducirse que la QoS en UTRAN viene determinada por la capacidad para diferenciar tipos de tráfico en los niveles ATM y AAL2.

Así, por ejemplo, en [16] se exponen los dos métodos básicos para proporcionar diferenciación de tráfico en la red de transporte AAL2/ATM, que son: diferenciación de tráfico a nivel AAL2 y a nivel ATM. En el caso de UTRAN, cabe considerar la aplicación de dichos métodos en los interfaces Iub, Iur e Iu-CS. La figura 9 ilustra los dos métodos mencionados suponiendo sólo dos tipos de servicios.

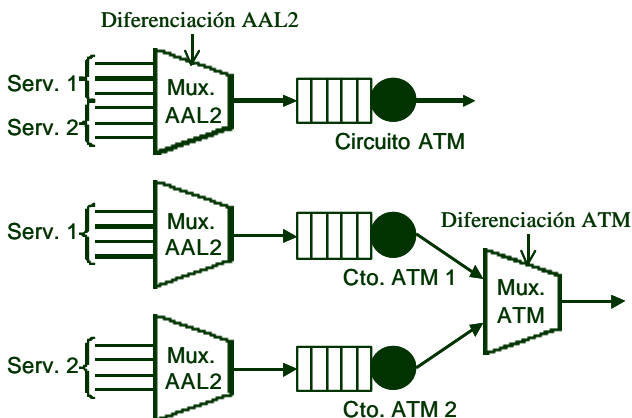


Fig. 9: Métodos básicos de diferenciación de tráfico

Resumidamente, el funcionamiento de cada uno de los métodos es:

- Sólo a nivel AAL2. El multiplexor AAL2 incluye algún método de diferenciación del tráfico, como un planificador capaz de dar distinto tratamiento a los paquetes de distintos servicios. Posteriormente todo el tráfico se envía por el mismo circuito ATM (por lo tanto, en este caso a nivel ATM no se utiliza gestión de tráfico). Tal y como se indica en el informe TR 25.934 [25] (que surge en la Release 4), no hay actualmente una manera estandarizada de proporcionar QoS a nivel AAL2 en la recomendación del ITU-T. Sin embargo se pueden utilizar mecanismos no estandarizados o propietarios, y de hecho hay artículos en que se proponen y analizan técnicas concretas (ver por ejemplo [15], [17], [26] y [27]).
- Sólo a nivel ATM. En este caso conceptualmente se tiene un multiplexor AAL2 dedicado a cada tipo de servicio, con lo que el tráfico de entrada de cada uno es homogéneo, y no es necesario realizar diferenciación a este nivel. Sin embargo, la salida de cada multiplexor será transportada por un circuito ATM distinto, cada uno del tipo y parámetros adecuados al tratamiento que requiera el servicio. Es decir, se utilizan los mecanismos de gestión de tráfico que proporciona ATM.
- Hay autores (una minoría) que mencionan o proponen la posibilidad de realizar diferenciación a los dos niveles: AAL2 y ATM (ver [16] y [17]).

Como primera aproximación a la provisión de QoS en UTRAN proponemos utilizar el segundo de los métodos, es decir, diferenciación a nivel ATM, por tratarse de la única solución normalizada a fecha de hoy. La figura 10 muestra de manera esquemática la aplicación de este enfoque en el contexto del interfaz Iub.

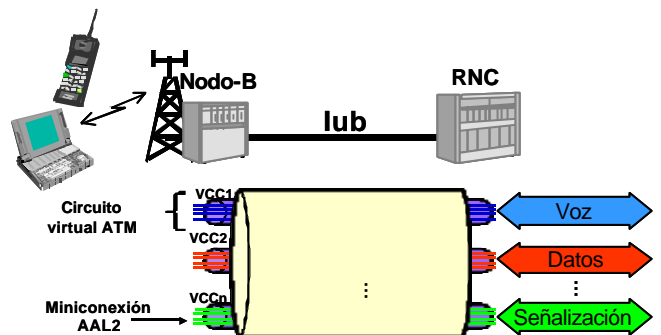


Fig. 10: Diferenciación de QoS a nivel ATM

Puesto que los diversos tipos de circuitos proporcionan garantías de QoS distintas, una posible decisión es dedicar un circuito ATM adecuadamente dimensionado para cada clase de tráfico de usuario. Adicionalmente, y puesto que se trata de un tipo de tráfico vital para el funcionamiento de la red, se podría decidir la utilización de un circuito exclusivo para el tráfico de señalización.

El tipo de circuito (de contrato de tráfico) a utilizar para cada clase de tráfico es otro de los aspectos a determinar. Hay múltiples posibilidades a este respecto. Por poner algún ejemplo, parece adecuado utilizar circuitos de tipo CBR para el tráfico de voz o que requiera de un aislamiento especial debido a sus requisitos de retardo (ver [15]), y se podrían utilizar circuitos de tasa variable (VBR-rt o VBR-nrt) para datos de las clases afluente o interactiva. La clase diferible, al carecer de requisitos de QoS, parece una buena candidata para ser transportada sobre un circuito UBR.

Cada circuito virtual ATM (VCC, Virtual Channel Connection) puede transportar hasta 248 miniconexiones AAL2, correspondientes a otras tantas comunicaciones de usuarios que, en caso de ofrecer un comportamiento de tasa variable, podrían exhibir ganancia por multiplexión estadística. A la hora de estimar el ancho de banda requerido, cabe la aplicación de algún método aproximado como, por ejemplo, los modelos de fluidos [28] y [29]). En la figura 11 se muestra gráficamente los resultados que proporcionan este tipo de modelos.

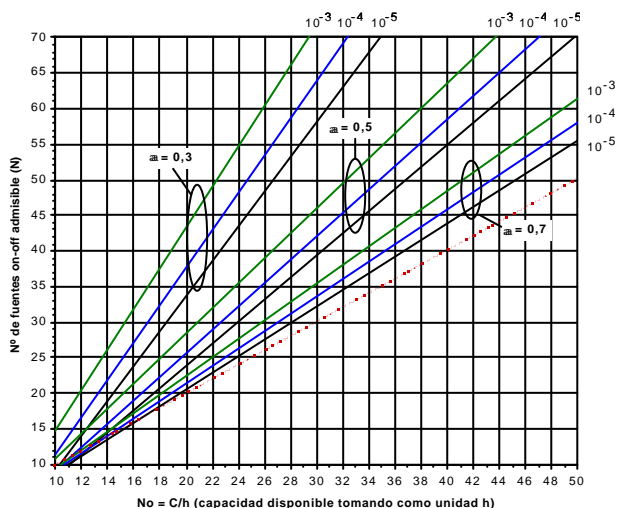


Fig. 11: Dimensionado de enlaces mediante el método de fluidos

La gráfica muestra el número N de fuentes on-off que puede cursar un enlace en función del cociente C/h , siendo C la capacidad del enlace y h la tasa de

pico de las fuentes, para diferentes valores del factor de actividad (α) y el factor de pérdidas.

De la gráfica puede deducirse que, para un factor de actividad dado y un cierto objetivo de pérdidas, la capacidad requerida es menor que la suma de la tasa de pico de las fuentes (representada mediante una línea discontinua). El ahorro de ancho de banda (la ganancia por multiplexión estadística) es mayor cuanto menor es el factor de actividad y menores son las exigencias en términos de pérdidas.

A modo de ejemplo, supóngase un Nodo-B con 1500 abonados telefónicos cuyo tráfico medio en la hora cargada es de 20 mE (mili-Erlang). Fijando una probabilidad de bloqueo máxima del 1% en el interfaz radio, el modelo de Erlang-B conduce a la necesidad de 42 circuitos de voz. Considerando un factor de actividad $\alpha=50\%$ y fijando un ratio de pérdidas menor que 10^{-4} (equivalente a una trama de voz perdida cada 200 segundos), las curvas anteriores nos indican que basta dimensionar una capacidad sobre I_{ub} de unas 30 veces la tasa de pico de una llamada. El ahorro de ancho de banda por el efecto de la multiplexión estadística en este caso sería de un 29%.

IV. DISEÑO DE LA RED DE TRANSMISIÓN

A. Interfaces de transmisión

A la hora de seleccionar la infraestructura de transmisión a emplear en UTRAN, cabe recurrir a cualquiera de las soluciones tradicionales: coaxial, fibra óptica o radioenlaces. Gracias a la existencia de numerosos interfaces ATM normalizados, los esquemas de transmisión sobre dichos medios pueden ser los habituales PDH (ej. E1, E3, ...) o SDH (ej. STM-1), si bien explotados en ATM.

Uno de los principales tipos de interfaces a utilizar en la fase inicial de despliegue de UMTS pueden ser los interfaces PDH. En particular, para células con poco tráfico, cabe la posibilidad de pensar en la utilización de interfaces E1 (2048 Kbit/s) explotados en ATM.

Si el tráfico en la célula es mayor, cabe pensar en la utilización del siguiente interfaz en la jerarquía PDH, el interfaz E3 (34 Mbit/s). Sin embargo, es muy posible que la capacidad de este interfaz sea demasiado elevada en la mayoría de los casos. Como alternativa, se puede recurrir a la combinación de varios E1s mediante la técnica IMA (Inverse Multiplexing ATM).

El interfaz IMA permite la utilización de varios interfaces E1 ATM, de manera que el conjunto se comporta como un único interfaz ATM cuya

capacidad es la suma (aproximadamente) de la que ofrecen cada uno de los interfaces E1 de manera aislada. El funcionamiento de este tipo de interfaces se ilustra en la figura 12.

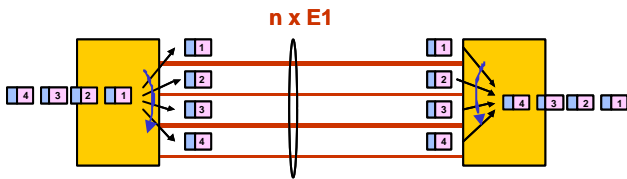


Fig. 12: Funcionamiento de un interfaz IMA

En caso de necesitar interfaces ATM de mayor capacidad, es posible recurrir al empleo de interfaces ATM sobre SDH. Así, por ejemplo, una posibilidad es el empleo de interfaces STM-1 (155Mbit/s) no canalizados. La necesidad de este tipo de interfaces puede aparecer conforme va aumentando el grado de concentración de tráfico en la red de interconexión entre Nodos-B y RNCs, así como para la conexión de estos con el núcleo de red (interfaz Iu).

B. Topología de la red de transmisión

En cuanto a la topología de la red de acceso, a la hora de interconectar los Nodos-B con los RNCs, así como éstos al núcleo de red (MSC/SGSN), es posible considerar varias alternativas. En la figura 13 se representan algunas de las configuraciones más habituales utilizadas en redes celulares.

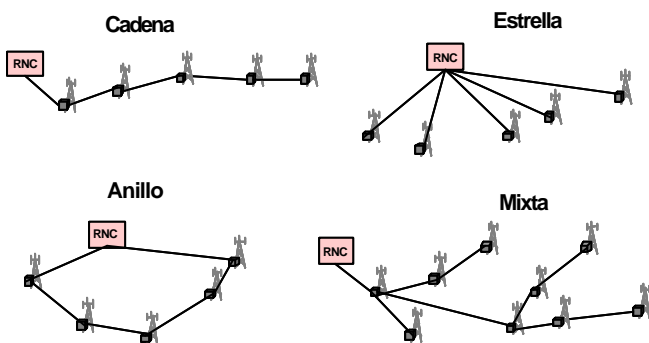


Fig. 13: Posibles configuraciones topológicas de UTRAN

Desde una óptica económica, resultan especialmente atractivas las topologías de interconexión que favorecen la concentración de tráfico. Así, por ejemplo, la elección de una configuración en cadena para interconectar varios Nodos-B a un RNC puede conducir a un ahorro considerable de recursos de transmisión,

especialmente en aquellos casos en los que la capacidad de los nodos-B sea relativamente pequeña. Este podría ser el caso de células UMTS en entornos rurales, e incluso de una gran mayoría de las células durante la fase inicial de despliegue de red.

La concentración de tráfico en UTRAN puede efectuarse mediante equipos auxiliares independientes o integrados dentro de los nodos-B. Una primera posibilidad es efectuar la concentración a nivel PDH, mediante DXXs (Digital Crossconnects). Esta solución, sin embargo, no permite aprovechar las ventajas que ofrece ATM. Para sacar partido de éstas, resulta más conveniente recurrir a dispositivos capaces de multiplexar el tráfico a nivel de circuitos virtuales ATM, como multiplexores ATM o conmutadores ATM. En este sentido, puede ser interesante que los nodos-B incluyan este tipo de funcionalidades a fin de evitar el empleo de equipos auxiliares.

En la figura 14 se muestra una posible topología de red de transmisión para el despliegue inicial de la UTRAN. Suponiendo que las necesidades iniciales de tráfico no son muy elevadas, los Nodos-B utilizan interfaces E1. En ocasiones, un nodo-B puede actuar de concentrador de tráfico de otros Nodos-B, pudiendo el número de E1s necesarios para llevar el tráfico agregado hacia el RNC menor que el que haría falta si se usara una topología en estrella.

Como se indica en la figura, cabe también la posibilidad de utilizar conmutadores ATM intermedios para efectuar concentrar tráfico a mayor escala. Nótese que el uso de estos conmutadores permite concentrar el tráfico de varios árboles de Nodos-B que pueden estar lejos del RNC que les controla. Finalmente, en la figura se muestra cómo el uso de conmutadores ATM permite concentrar el tráfico de varios RNCs hacia el núcleo de red.

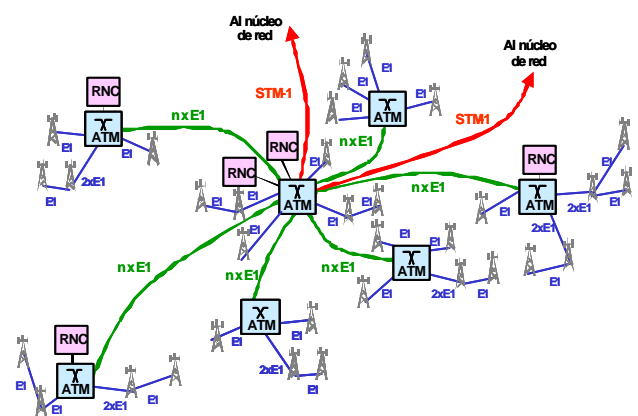


Fig. 14: Ejemplo de topología UTRAN

V. CONCLUSIONES Y TRABAJOS FUTUROS

En este artículo se ha abordado desde un punto de vista práctico el problema del dimensionado de la red de acceso radio UTRAN. Fijando como objetivos de fundamentales la optimización de recursos de transmisión y el soporte de QoS, se han analizado los principales parámetros y alternativas de diseño a considerar en un escenario de despliegue real.

A lo largo de nuestro análisis se ha puesto de manifiesto la extraordinaria complejidad asociada al dimensionado de UTRAN. Así, aunque nuestro trabajo puede servir como primera aproximación al problema, creemos necesario seguir profundizando en la elaboración de una metodología más precisa. En la misma línea, consideramos conveniente acometer un estudio más exhaustivo del impacto que puedan tener los mecanismos de gestión de tráfico y estrategias de multiplexión. Debido a la dificultad que plantea un tratamiento analítico del problema, actualmente se trabaja en el desarrollo de un modelo de simulación para el dimensionado y evaluación de prestaciones de UTRAN.

REFERENCIAS

- [1] UMTS Forum. "The UMTS Third Generation Market Study Update". Informe del UMTS Forum Núm. 17, Agosto 2001.
- [2] Dimitriou, N.; Tafazolli, R.; Sfikas, G. "Quality of service for multimedia CDMA". IEEE Communications Magazine , Vol. 38 Núm. 7, págs. 88 –94, Julio 2000.
- [3] Lister, D.; Dehghan, S.; Owen, R.; Jones, P. "UMTS capacity and planning issues". First International Conference on 3G Mobile Communication Technologies, págs. 218 –223, Marzo 2000.
- [4] Parsa, Kourosh; Ghassemzadeh, Saeed S.; Kazeminejad, Saied. "Systems Engineering of Data Services in UMTS WCDMA Systems". 2001 IEEE International Conference on Communications (ICC2001). Helsinki (Finlandia), Junio 2001.
- [5] Holma, H.; Toskala, A. "WCDMA for UMTS: Radio Access for Third Generation Mobile Communications". John Wiley & Sons, 2000.
- [6] 3GPP. "General UMTS Architecture". 3GPP TS 23.101. Diciembre 2000.
- [7] 3GPP. "3rd Generation mobile system Release 1999 Specifications", 3GPP TS 21.101, Marzo 2002.
- [8] 3GPP. "3rd Generation mobile system Release 4 Specifications", 3GPP TS 21.102, Marzo 2002.
- [9] 3GPP. "3rd Generation mobile system Release 5 Specifications", Draft 3GPP TS 21.103, Marzo 2002.
- [10] 3GPP. "IP Transport in UTRAN". 3GPP TR 25.933. Marzo 2002.
- [11] 3GPP, "UTRAN General Overview", 3GPP TS 25.401, Marzo 2002.
- [12] ITU-T Recommendation I.363.2. "B-ISDN ATM Adaptation layer specification: Type 2 AAL". Noviembre 2000.
- [13] 3GPP. "QoS Concept and Architecture". 3GPP TS 23.107, Marzo 2002.
- [14] Valko, A. G.; Racz, A.; Fodor, G. "Voice QoS in Third-Generation Mobile Systems". IEEE Journal on Selected Areas in Communications, Vol. 17, Núm. 1, págs. 109-123, Enero 1999.
- [15] Isnard, O.; Calmel, J.-M.; Beylot, A.-L.; Pujolle, G. "Handling Traffic Classes at AAL2 / ATM layer over the Logical Interfaces of the UMTS Terrestrial Access Network". 11th IEEE International Symposium on Personal, Indoor and Mobile Radio Communication, Vol. 2, págs. 1464-1468, Londres, Septiembre 2000.
- [16] Eneroth, Göran; Fodor, Gábor; Leijonhufvud, Gösta; Rácz, András; Szabó, István. "Applying ATM/AAL2 as a Switching Technology in Third-Generation Mobile Access Networks". IEEE Communications, Vol. 37 Núm. 6, págs. 112-122, Junio 1999.
- [17] Yoo, Sang-Kyung; Park, Hong-Shik. "Quality-of-Service Provisioning for Mobile Voice and Data Services over ATM Network using AAL2". 3rd ICACT, Muju (Corea), Febrero 2001.
- [18] Reyes-Lecuona, A.; González-Parada, E.; Casilari, E.; Casasola, J. C.; Díaz-Estrella, A. "A page-oriented WWW traffic model for wireless system simulations". Proceedings of the 16th International Teletraffic Congress (ITC'16), págs. 1271-1280, Edinburgo, (Reino Unido), Junio 1999.
- [19] Staehle, Dirk; Leibnitz, Kenji; Tran-Gia, Phuoc. "Source Traffic Modeling of Wireless Applications". International Journal of Electronics and Communications, Vol. 55, Núm. 1, 2001.
- [20] Klemm, Alexander; Lindemann, Christoph; Lohmann, Marco. "Traffic Modeling and Characterization for UMTS Networks". Proc. of the Globecom, Internet Performance Symposium, San Antonio TX, Noviembre 2001.
- [21] ETSI. "Universal Mobile Telecommunications System (UMTS); Selection procedures for the

- choice of radio transmission technologies of the UMTS”. TR 101 112 V3.2.0. Abril 1998.
- [22] 3GPP. “Services and Service Capabilities” .3G TS 22.105.
- [23] 3GPP. “AMR Speech Codec; General Description”. 3G TS 26.071.
- [24] 3GPP. “Delay Budget within the Access Stratum”. 3GPP TR 25.853.
- [25] 3GPP. “QoS optimization for AAL type 2 connections over Iub and Iur interfaces”. 3G TR 25.934.
- [26] Lim, Huhnkuk; Lee, Suwon; Lee, Dongwook; Kim, Kiseon; Song, Kwangsuk; Oh, Changhwan. “A New AAL2 Scheduling Algorithm for Mobile Voice and Data Services over ATM”. ITC-CSCC 2000, vol. 1, págs. 229-232, Pusan (Corea), Julio 2000.
- [27] Wu, J.-L.C.; Huang, Chi-Hong; Sheu, Rong-Tsung. “Performance study of AAL2 protocol for low-bit-rate multimedia services”. 15th International Conference on Information Networking, págs. 793 –798, 2001.
- [28] McDysan, David E.; Spohn, Darren L. “ATM Theory and Applications”, McGraw-Hill, 1999.
- [29] Hersent, Olivier; Gurle, David; Petit, Jean-Pierre. “IP Telephony. Packet-based multimedia communications systems”, Addison-Wesley, 2000.