

# Capítulo 1

## Introducción

### 1.1. Introducción

Las actuales redes de comunicaciones han sido optimizadas a través de los años para soportar un tipo de tráfico específico, como es el caso de las redes telefónicas, redes de transmisión de datos o redes de difusión. Un primer intento de integración de las distintas redes de comunicaciones se realizó con la aparición de la red digital de servicios integrados (RDSI). Esta red permite el acceso con un mismo interfaz, a servicios de conmutación de paquetes y a servicios de conmutación de circuitos. El principal inconveniente que presenta la RDSI es la limitación de ancho de banda ofrecido al servicio, prefijado a 64 kbps para acceso básico y a 2 Mbps en el caso de establecer la conexión a través de un acceso primario.

Los actuales avances en la explotación de las líneas de transmisión han permitido alcanzar niveles de transporte de 2'4 Gbps con portadoras ópticas (OC-48). La estandarización de los soportes lógicos para transmisiones sobre medios ópticos quedó reflejada en la especificación de SONET, basada en una estructura de tramas síncronas que permiten la multiplexación digital [BalChi89]. La especificación de SONET permite anidar en una misma trama diferentes jerarquías síncronas primarias y terciarias, resolviendo, a su vez, el problema de compatibilidad de las tramas plesiocronas de los estándares europeos y americanos. Asimismo, las mejoras en la tecnología Very Large Scale Integration (VLSI) y los avances en la conmutación fotónica han permitido plantear el desarrollo de diferentes elementos de una red digital capaces de manejar grandes volúmenes de información transferidos en los enlaces ópticos.

El estado actual de la tecnología permite evolucionar la RDSI hacia una red con la misma filosofía pero capaz de soportar servicios con un consumo mucho mayor de

ancho de banda. Actualmente, la especificación de la nueva red que contempla estos aspectos, denominada Red Digital de Servicios Integrados de Banda Ancha (RDSI-BA), es llevada a cabo por la Unión Internacional de Telecomunicaciones (ITU) y un consorcio de fabricantes mundiales con el nombre de ATM Forum. Los puntos clave de la concepción de esta red han sido la *integración de servicios* y la *optimización de la utilización de los recursos de la red* [HanHub91].

Las diferentes características de los servicios actuales en cuanto a mecanismos orientados o no orientados a conexión, tasas de transferencia constante o variable y requerimientos temporales, hacen realmente complicada la tarea de acomodarlos en una red con un único mecanismo de transporte. En la RDSI-BA se han definido varios tipos de adaptaciones para soportar los diferentes tipos de servicios que acondicionen, en cada caso, la conexión de cada clase de tráfico según su naturaleza [L.363].

Para conjugar las ventajas de las tecnologías de conmutación de circuitos y de conmutación de paquetes, la RDSI-BA se especificó como una red basada en la conmutación de circuitos virtuales. La necesidad de la gestión eficiente de los recursos de la red, y en particular la velocidad de conmutación, condicionaron a que los paquetes fueran de longitud fija y de tamaño sustancialmente inferior al de otras redes.

El planteamiento de la RDSI-BA es suficientemente flexible como para adaptarse a las innovaciones tecnológicas que vayan apareciendo y permite albergar todo tipo de servicios futuros todavía no contemplados. La RDSI-BA es especialmente adecuada para los nuevos servicios de vídeo como la videotelefonía, la videoconferencia y los servicios de distribución, cuyas necesidades de ancho de banda son bastante superiores a la de los servicios soportados por las redes actuales.

Los nuevos mecanismos de compresión para voz y vídeo, que disminuyen los requerimientos de ancho de banda de estos servicios, junto con las transferencias de datos entre redes locales o metropolitanas generan servicios cuyas necesidades de ancho de banda fluctúan dinámicamente, por ello, se ha adoptado el *modo de transferencia asíncrono* (Asynchronous Transfer Mode, ATM) como el más indicado para la RDSI-BA. El ATM permite la explotación eficiente de los recursos empleando el principio de *ganancia de multiplexación estadística*. Este principio es aplicable a los servicios de tasa binaria variable. A través de la multiplexación estadística se puede explotar la compensación de los diferentes requerimientos de cada servicio en el transcurso de la conexión.

La necesidad de optimizar los recursos de las redes de comunicación ha llevado al desarrollo de distintos algoritmos de compresión para audio y vídeo. Estos mecanismos de compresión fueron inicialmente desarrollados para redes de comunicaciones donde los servicios disponían de una tasa constante de transmisión. El uso de la compresión, en estos casos, obliga a mantener una calidad del servicio variable a fin de conservar el rango de transmisión constante. El ATM de la RDSI-BA ha permitido modificar los algoritmos de compresión para mantener una calidad constante del servicio,

cuantificable en términos de relación señal a ruido (Signal to Noise Ratio, SNR), dado que el ATM soporta las transferencias a tasa variable. Se han realizado algunos estudios para determinar el interés que reporta la transmisión a velocidad variable frente a velocidad constante para estos servicios [COST211].

Los algoritmos propuestos para la compresión de voz pretenden disminuir el consumo de ancho de banda explotando la duración de los intervalos de silencio, del orden del 60% del tiempo, junto con la redundancia temporal. Con el empleo de mecanismos de compresión, el ancho de banda necesario para un servicio de voz es extremadamente pequeño, respecto al de un enlace de la RDSI-BA, por lo que el principio de multiplexación estadística aplica perfectamente. La agrupación en un mismo enlace de un gran número de codificaciones de velocidad variable de servicios de audio permite considerar el tráfico cursado global como un valor prácticamente constante, que se obtiene a través de la adición de las medias de cada uno de los servicios [SriWhi86].

Los servicios de transmisión de vídeo demandan un ancho de banda muy superior a los de audio. A fin de aumentar la efectividad en la transmisión y almacenamiento de las imágenes de vídeo se han desarrollado un conjunto de técnicas de compresión basadas en la explotación de la redundancia temporal y espacial de la imágenes. Las técnicas más efectivas utilizadas para el aprovechamiento de la redundancia espacial de las imágenes trabajan en el dominio de la frecuencia. Los métodos frecuenciales más empleados son la transformada coseno discreta (Discrete Cosinus Transform, DCT), la descomposición en subbandas [KarVet89], incluyendo las denominadas *wavelets* [ArgBen93] y las codificaciones piramidales. Para imágenes en movimiento, se explota además la redundancia temporal a través de la compensación de movimiento, basada en la detección del desplazamiento de objetos. Diferentes estándares han ido apareciendo tanto para imágenes estáticas, JPEG, como para imágenes en movimiento, H261 y MPEG-I [Kou95]. En todos ellos se emplea la DCT [ChaCha94], que se ha revelado como el mejor método para la compresión espacial, y la compensación de movimiento, aplicada, generalmente, sobre elementos rectangulares de la imagen. Para aumentar la eficiencia de la transmisión estos estándares incorporan códigos de longitud variable (códigos de Huffman) para la codificación, obtenidos después de estudios intensivos de codificación empleando dichos algoritmos [H261].

El nuevo estándar MPEG-II [MPEGII] ha sido concebido especialmente para la transmisión de vídeo digital [LagPer95]. El algoritmo de compresión de vídeo MPEG-II se basa en el empleado en el MPEG-I, ampliando diferentes posibilidades de formatos en cuanto al tamaño de la imagen y en cuanto a la presentación entrelazada o progresiva. De esta manera se pretende unificar en un mismo formato de transmisión múltiples resoluciones, permitiendo anidar dentro del formato los estándares JPEG, H261 y MPEG-I que ya estaban estandarizados. La principal innovación que aporta el nuevo estándar MPEG-II es la introducción del concepto de escalabilidad [ChiAna94]

[MitAnd94]. Dentro del formato de transmisión MPEG-II se permiten transmisiones con escalabilidad espacial y temporal, es decir, dependiendo del receptor se decodificará la información en una resolución baja (1/4 de la TV convencional), en una resolución clásica de TV, o en una resolución propia de la televisión de alta definición (HDTV). Asimismo, la cadencia de los cuadros de las secuencias también es seleccionable [SunKwo95]. Los mecanismos de escalabilidad espacial y temporal son especialmente indicados para servicios de difusión o de recuperación de información de bases de datos de vídeo. Un tercer mecanismo de escalabilidad se incorpora para los servicios de vídeo en función de la calidad, cuantificada en términos de SNR, de forma que se definen dos niveles de decodificación para la misma resolución espacial y cadencia temporal.

Para ubicar los servicios de transmisión de imagen sobre la RDSI-BA, en particular según los estándares MPEG, se está elaborando actualmente la especificación para utilizar la capa de adaptación número 5 de la RDSI-BA (AAL5), aunque, previamente se discutió la definición de un nivel de adaptación específico, el número 6, para las transferencias MPEG a velocidad variable.

### **1.2. Modo de Transferencia Asíncrono**

El ATM ha sido normalizado por la ITU como el modo de transferencia para la RDSI-BA [I.113]. El ATM está basado en técnicas de multiplexación estadística por división en el tiempo y conmutación rápida de paquetes que transmite y encamina el tráfico mediante una dirección contenida en lo que se denomina un paquete de información [I.361]. El ATM utiliza pequeños paquetes de longitud fija, denominados celdas o células. Una celda ATM consta de 53 octetos, 5 de los cuales forman la cabecera del paquete (y contienen básicamente una dirección) y los 48 octetos restantes forman el campo de información [I.121].

El tiempo de transmisión de una celda se considera como la longitud de la ranura temporal en el multiplex. El término multiplexación estadística se refiere al hecho de que varias conexiones pueden compartir un enlace de capacidad menor que la suma de sus requisitos de ancho de banda de pico. El término asíncrono se usa para reflejar que las celdas de una misma comunicación pueden aparecer a intervalos irregulares sobre los enlaces de la red [Onv94].

Al igual que en el resto de tecnologías de conmutación de paquetes se consigue mayor eficiencia en la utilización de los recursos que en el caso de conmutación de circuitos, ya que los usuarios acceden al canal de información cuando lo necesitan, es decir, durante el tiempo en que lo necesitan (por supuesto si el canal ya está en uso, un nuevo usuario debe esperar hasta que consiga acceder). En cambio, en el modo de transferencia síncrono (Synchronous Transfer Mode, STM) se identifican los datos de

una determinada comunicación en base a una ranura temporal particular (se usa multiplexación por división en el tiempo) dentro de una trama, de ahí que no sea necesaria una cabecera que contenga la dirección de destino del mensaje. Además, las ranuras temporales en una trama STM no se comparten entre llamadas. Por tanto, incluso cuando la fuente de tráfico está inactiva la ranura temporal correspondiente no puede ser utilizada por otro usuario.

En la elección de paquetes de longitud fija sobre longitud variable intervienen varios factores como, por ejemplo, la eficiencia de transmisión, la velocidad y complejidad de los nodos de conmutación y el retardo que se sufre en la red [Pry93].

El hecho de optar por una longitud corta limita el retardo producido en la red, proporciona una mejor multiplexación estadística de la información sobre el medio que en el caso de elegir paquetes de mayor tamaño y de longitud variable, y permite la integración de tráfico con requisitos de tiempo real. La ventaja en términos de velocidad y complejidad de los conmutadores es también clara, en el caso de paquetes de longitud fija el tiempo para procesar su cabecera es también fijo, sin embargo, si éstos fueran de longitud variable el tiempo para procesar la cabecera debería diseñarse tomando como referencia la longitud más pequeña posible, complicando así la realización física del dispositivo. Por otra parte, el sistema de gestión y control de las memorias tampón, o buffers, presentes en todo conmutador de paquetes es mucho más simple si los paquetes tienen longitud fija. Además, los requisitos de memoria son mucho mayores para el caso de paquetes de longitud variable, ya que en el dimensionado se debe considerar la posibilidad de alojar paquetes de la máxima longitud posible.

Por otra parte, y debido a los enlaces de alta velocidad, el retardo de propagación se hace significativo respecto al tiempo de transmisión y al tiempo de proceso que conlleva la ejecución de los protocolos de comunicación. Por tanto, las redes ATM utilizan protocolos simplificados e intentan desplazar el peso de los protocolos que actualmente se ejecutan de enlace a enlace hacia capas superiores del modelo de referencia de protocolos OSI [Tan89] para que el proceso se lleve a término de extremo a extremo.

Aunque, en principio, ATM puede utilizarse sobre cualquier medio físico, fue desarrollado pensando en los nuevos soportes ópticos, como por ejemplo, la jerarquía digital síncrona JDS [LeeKan93] [SDH93].

El uso de fibra óptica como medio de transmisión, a parte de proporcionar anchos de banda elevados (tasas binarias altas) y mejores relaciones de atenuación (mayor distancia entre repetidores) también se caracteriza por ser un medio de bajo ruido. Ello implica que la tasa de error en el canal sea a su vez baja y, por tanto, que las técnicas de control de errores usadas en la actualidad sean reexaminadas.

Los esquemas de control de errores pueden llevarse a la práctica entre nodos intermedios o entre extremos de la comunicación. Si el control de errores se realiza entre nodos, la retransmisión de celdas erróneas o pérdidas sólo tiene lugar entre nodos

de conmutación adyacentes. Ello involucra un tiempo de proceso en la ejecución de los protocolos correspondientes.

Nótese que si el medio es ruidoso y el control de errores no se realizase en cada nodo, podría presentarse el caso de que ningún paquete llegase a su destino y el caudal se reduciría al mínimo ya que el sistema no haría más que retransmitir paquetes erróneos. Por tanto, cuando el medio es ruidoso, controlar los errores en cada nodo resulta lo más adecuado.

Si el control de errores tiene lugar sólo entre los nodos fuente y destino un paquete erróneo es enviado durante todo el recorrido de la comunicación malgastando recursos, sin embargo, si el medio presenta bajo ruido esta situación no ocurre con frecuencia y se tienen efectos positivos como es el aligerar el tiempo de proceso en los nodos ya que no es necesario verificar en cada nodo la existencia o no de errores y desencadenar las operaciones necesarias para tal menester. Por otra parte, también disminuyen los requisitos de almacenamiento y retardo en los nodos intermedios. Este procedimiento se denomina conmutación rápida de paquetes.

El ATM sigue esta última opción y sólo incorpora control de errores nodo a nodo para la cabecera de las celdas, pero no sobre el campo de información haciendo, por tanto, uso de la mejor tasa de errores que proporcionan los medios de transmisión óptica.

La figura 1.2.1 muestra el modelo de referencia de protocolos (PRM) para la RDSI-BA. En la figura 1.2.2 se ilustran las funciones de cada capa según se describe en [I.321] y [I.413].

El modelo de referencia de protocolos consta de tres planos: el plano de usuario, el plano de control y el plano de gestión. El plano de usuario hace referencia a la transferencia de información de usuario e incluye mecanismos como control de flujo y recuperación de errores. El plano de control se encarga del control de las llamadas y conexiones. Es decir, comporta todas aquellas funciones de señalización necesarias para establecer, supervisar y liberar las conexiones.

El plano de gestión es el responsable de funciones de operación y mantenimiento (Organization and Management, OAM) de la red. La gestión de planos realiza funciones de gestión relativas al sistema global y proporciona la coordinación necesaria entre planos. No posee una estructura de capas. La gestión de capas realiza funciones de gestión que tienen que ver con los recursos y parámetros que residen en sus entidades de protocolo (p.e., meta-señalización). También maneja flujos de información de operación y mantenimiento, pero específicos a cada capa en cuestión.

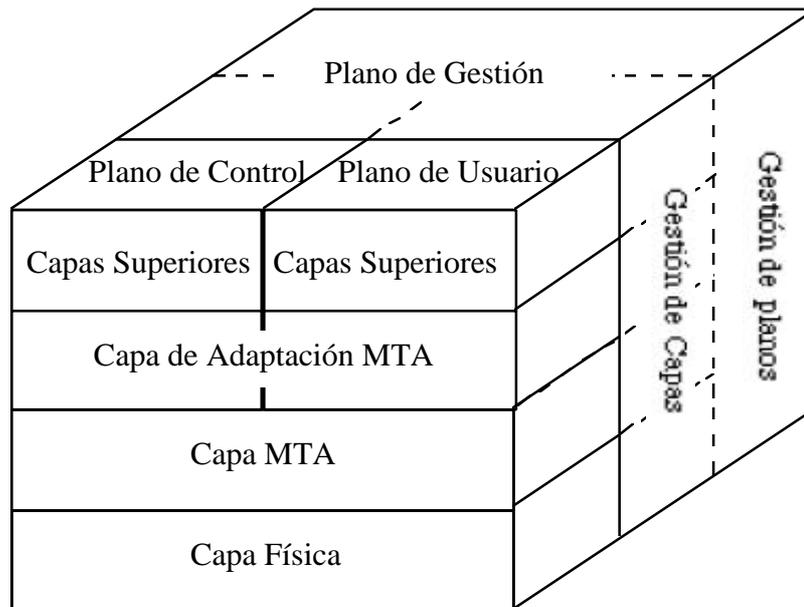


Figura 1.2.1. Modelo de referencia de protocolos para la RDSI-BA

En la actualidad se han definido con claridad tres capas del modelo de referencia: la capa física, la capa ATM y la capa de adaptación ATM. Las capas superiores se encuentran todavía en fase de estudio.

Layer Management	Higher layer functions	Higher layers	
	Convergence	CS	AAL
Segmentation and reassembly	SAR		
Generic flow control Cell header generation/extraction Cell VPI/VCI traslation Cell multiplex and demultiplex	ATM		
Cell rate decoupling HEC sequence generation/verification Cell delination Transmission frame adaptation Transmission frame generation/recovery	TC	Physical layer	
Bit timing	PM		
Physical medium			

CS: Convergence Sublayer  
 PM: Physical Medium  
 SAR: Segmentation and Reassembly sublayer  
 TC: Transmission Convergence

Figura 1.2.2. Funciones de la RDSI-BA en relación al modelo de referencia

La capa física (Physical Layer, PL) se divide en dos subcapas: la subcapa del medio físico (Physical Medium, PM) y la subcapa de convergencia a la transmisión (Transmission Convergence, TC) [I.432]. La subcapa PM incluye todas aquellas funciones dependientes del medio físico en cuestión (p.e., conversión electro-óptica). La subcapa TC efectúa todas las funciones necesarias para transformar un flujo de celdas en un flujo de entidades de datos (por ejemplo bits), compatible con el esquema de multiplexación del sistema de transmisión.

La capa de adaptación ATM (ATM Adaptation Layer, AAL) [I.362] [I.363] tiene como función básica el aislar las capas superiores de las características propias de la capa ATM. Se encarga de adaptar los datos procedentes de niveles superiores a un formato que pueda ser manipulado por la capa ATM. La capa AAL se organiza en dos subcapas. La subcapa de convergencia (Convergence Sublayer, CS) y la subcapa de segmentación y reensamblado (Segmentation And Reassembly, SAR). La subcapa SAR tiene como misión el segmentar los datos de las capas superiores a un formato compatible con el campo de información de usuario de una celda ATM (48 octetos), o recíprocamente, el reensamblar dichos campos de información a unidades de datos de protocolo (Protocol Data Unit, PDU) de la capa superior.

La subcapa CS es independiente del servicio y realiza funciones necesarias para soportar aplicaciones específicas (p.e., recuperación de reloj en servicios de vídeo insertando palabras de sincronización). Los posibles mecanismos de adaptación definidos para el transporte de información en redes ATM [I.362] se han planteado según el tipo de servicio que deben soportar. Los parámetros que clasifican un tipo de servicio son [Pry93]:

- i) Necesidad de relación temporal entre fuente y destino.
- ii) Tasa generada.
- iii) Modo de conexión.

A partir de estos parámetros se definen cuatro tipos de servicios [I.211]:

i) Servicios de Clase A. Tienen requerimientos temporales, su tasa binaria es constante y están orientados a conexión. Para estos servicios la red ATM establece una conexión que emula un circuito.

ii) Servicios de Clase B. Tienen las mismas características que los de clase A salvo que la tasa generada es variable. Estos son los servicios clásicos de vídeo y audio comprimidos con calidad semi-constante.

iii) Servicios de Clase C. No tienen requerimientos temporales, su tasa es variable y están orientados a conexión. Estos servicios son clásicos en la conexión entre redes locales.

iv) Servicios de Clase D. Son servicios de tasa variable, sin requerimientos temporales y no orientados a conexión. Son servicios de tráfico de datos no orientados a la conexión.

Los distintas capas AAL definidas se corresponden con los diferentes servicios de la siguiente forma:

- AAL-1: definida para soportar aplicaciones de clase A.
- AAL-2: definida para soportar aplicaciones de clase B.
- AAL-3/4: cuando se aceptó que un sólo protocolo AAL podía usarse para soportar servicios de datos orientados a conexión y en modo datagrama (no orientados a conexión) se especificó la capa AAL-3/4 para tratar ambos servicios, o sea, servicios clase C y D.
- AAL-5: como resultado de la complejidad asociada con la capa AAL-3/4, se propuso el nivel AAL-5, también conocido como capa de adaptación simple y eficiente (Simple and Efficient Adaptation Layer, SEAL). Proporciona funciones más limitadas (detección de errores pero no recuperación) y posee menores requisitos en cuanto al proceso que implica y al ancho de banda que necesita. Se puede aplicar para servicios de clase B, C o D.

La capa ATM es independiente del medio físico y de los servicios que transporta. Se encarga de las funciones relacionadas con la información presente en la cabecera de la celda ATM y que, por tanto, son necesarias para el tratamiento lógico de dicha celda ATM [I.363]. La identificación de canales virtuales y la detección de errores en la cabecera de la celda son ejemplos de dichas funciones.

Cabe destacar que no existe una correspondencia clara entre el modelo de referencia de protocolos para la RDSI-BA y el modelo de protocolos OSI de 7 niveles.

### 1.2.1. Estructura de la Celda ATM

Como ya se enunció anteriormente, la celda ATM tiene 53 octetos de longitud, 5 de los cuales forman la cabecera de la celda que consta de varios campos de información tal como se aprecia en la figura 1.2.3. Ello es así para el interfaz usuario-red [UNI3.0], en el caso del interfaz red-red o nodo-nodo, el campo GFC se dedica también a VPI.

La funcionalidad asociada a los distintos campos de la cabecera está reducida al máximo (básicamente encaminamiento) para garantizar un proceso rápido en los nodos de la red. De los cinco octetos que forman la cabecera, tres contienen información de encaminamiento, dos para lo que se denomina identificador de canal virtual (Virtual Channel Identifier, VCI) y un octeto para el identificador de camino virtual (Virtual Path Identifier, VPI). El resto de la cabecera consiste en un campo de control de flujo

genérico (Generic Flow Control, GFC), tres bits para un identificador del tipo de información útil que transporta la celda (Payload Type Identificador, PTI), un bit de prioridad frente a pérdida de celdas (Cell Loss Priority, CLP) y ocho bits para un campo de control de errores en la cabecera (Header Error Control, HEC).

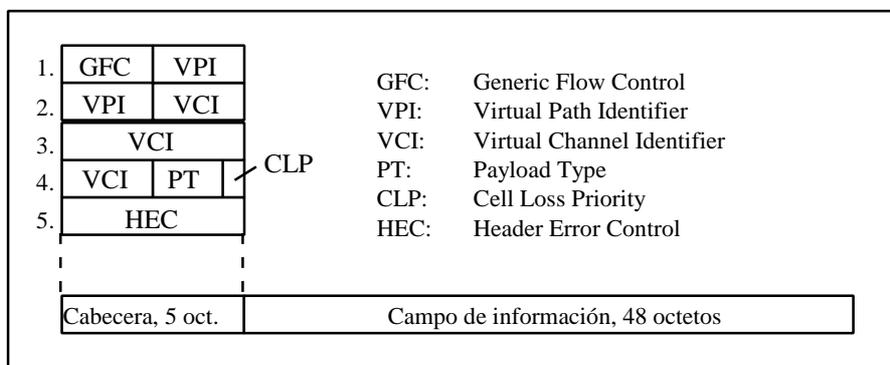


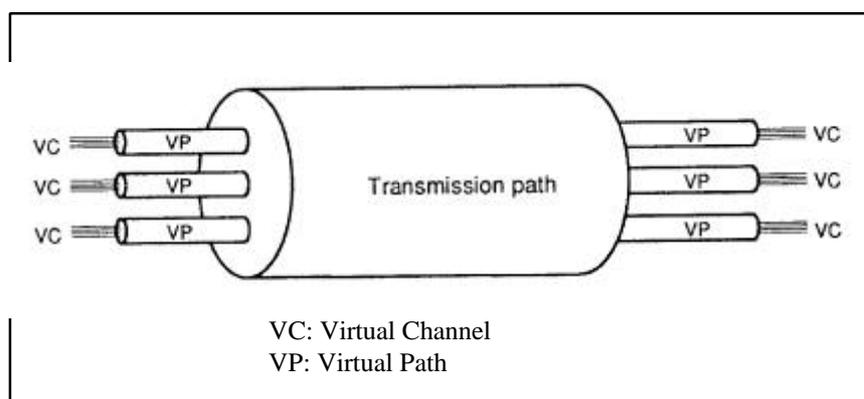
Figura 1.2.3. Estructura de la celda ATM

Tal y como corresponde a una tecnología de conmutación de paquetes, la cabecera contiene principalmente información para transportar las celdas ATM de un nodo al siguiente. Sin embargo, en lugar de especificarse explícitamente las direcciones de fuente/destino, las celdas ATM se etiquetan mediante números identificadores de conexión virtual (VCI/VPI), ello permite identificar las conexiones mediante un número menor de bits y conseguir un encaminamiento universal. El nodo destino se especifica mediante una secuencia de pasos en el proceso de encaminamiento que se determina por procedimientos de señalización en el momento de establecer la conexión.

### 1.2.2. Concepto de Conexión Virtual

El desarrollo de una conexión en ATM se divide en tres fases: la fase de establecimiento donde se reservan los recursos necesarios (en el caso en que éstos estén disponibles, en caso contrario la conexión simplemente se rechaza), la fase de transferencia de la información y la fase de liberación de recursos y de la conexión. Las celdas ATM de una misma comunicación viajan por la misma ruta mientras dura la transmisión (ello resulta favorable frente a un modo no orientado a conexión, evitando el resecuenciamiento de paquetes en servicios de tiempo real). Dicha ruta se especifica en la fase de establecimiento de la llamada. La cabecera de la celda ATM contiene en cada momento la información que la red necesita para encaminar la celda sobre la ruta preestablecida.

Un camino virtual consta de varios canales virtuales (ver figura 1.2.4). La distinción en la etiqueta de direccionamiento entre VCI y VPI permite a la red usar una notación más corta y compacta para las rutas con más tráfico (p.e., enlaces entre grandes ciudades) gracias al identificador VPI, a la vez que con el identificador VCI se preserva la identidad de cada canal individual dentro del camino de comunicación. Ello permite que los equipos utilizados en la transmisión traten las llamadas sólo en base al campo VPI, sin la necesidad de operar con el resto del campo de direccionamiento (VCI) hasta que el enlace general llega a su destino final, donde el tráfico correspondiente a cada canal se distribuye de acuerdo a su VCI.



*Figura 1.2.4. Relación entre camino virtual y canal virtual*

El resto de campos en la cabecera se usa para controlar el flujo del tráfico generado en la parte de red de usuario (sólo en el interfaz usuario-red), para discernir el tipo de información que transportan las celdas (celdas de usuario o celdas propias de la red), y para marcar cuál es la prioridad de las celdas frente a pérdidas y poder, así, ejercer alguna acción sobre las mismas en caso de congestión. Finalmente, se utiliza un octeto para asegurar la integridad de la cabecera (básicamente de la dirección) a través de la red.

El ATM usa el concepto de conexión virtual [Lan94], estableciendo conexiones virtuales entre cada par de nodos de conmutación intermedios en la transmisión desde un extremo fuente a un extremo destino. Estas conexiones se denominan virtuales para distinguirlas de los circuitos o canales dedicados como en el caso del STM. En ATM el enlace no se reserva únicamente a un usuario sino que se comparte temporalmente con otros usuarios.

Las conexiones ATM sólo existen como conjuntos de tablas de encaminamiento que se mantienen en cada conmutador y que están indexadas en función de la etiqueta de direccionamiento (VCI/VPI) de la cabecera de las celdas. Las etiquetas de direccionamiento en ATM tienen sólo un significado local, en cuanto que sólo son

relevantes entre dos equipos de conmutación ATM adyacentes. Cuando se establece un camino virtual, a cada conmutador se le proporciona un conjunto de tablas de búsqueda (indexadas por los identificadores de camino/canal virtual) para identificar una celda de entrada por su etiqueta de direccionamiento presente en la cabecera, encaminarla a través de la red de conmutación hasta su puerto de salida destino, reescribiendo la dirección de entrada por una nueva etiqueta que será usada por el siguiente conmutador en la ruta de transmisión y que de nuevo la reconocerá como un índice de su propia tabla (Fig. 1.2.5).

La celda pasará de conmutador a conmutador por la ruta preestablecida en la fase de establecimiento de la conexión, pero dicha ruta es virtual en cuanto a que las facilidades de transporte sólo se le dedican mientras la celda las atraviese.

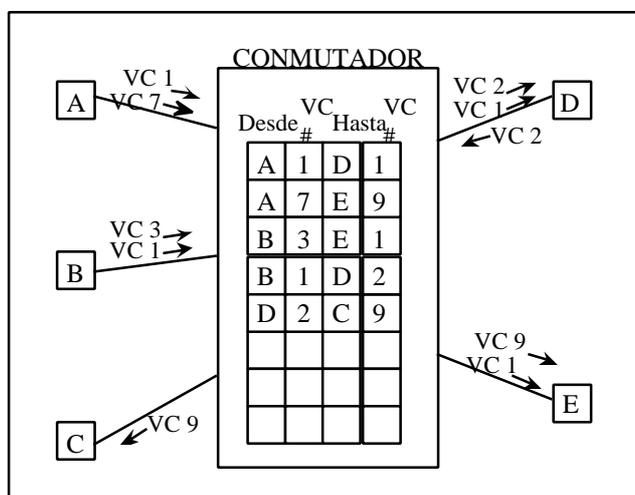


Figura 1.2.5. Tablas de búsqueda en el conmutador

### 1.3. Gestión del tráfico en redes ATM

La adopción del ATM ha permitido la incorporación de los servicios de velocidad variable en la nueva RDSI-BA. De esta forma se ha conseguido una maximización del uso de los recursos aplicando el principio de ganancia por multiplexación estadística. Como inconveniente, se introducen dos nuevos efectos respecto a las redes de conmutación de circuitos ya existentes: *pérdidas de paquetes* por desbordamiento de las colas de espera de los elementos de la red y un *retardo variable* entre llegadas de paquetes al receptor.

La ubicación de ancho de banda en las redes ATM para servicios CBR es prácticamente igual a la tasa constante de generación. En los servicios VBR se pueden emplear diferentes técnicas de ubicación de ancho de banda [DecFac91]:

i) *Ancho de banda de pico*. En esta asignación de ancho de banda se reserva el máximo valor alcanzado por la generación de la tasa binaria, o tasa de pico. Esta asignación es muy ineficiente, ya que, en el transcurso de la conexión, el ancho de banda consumido será muy inferior al reservado. No obstante, esta asignación de recursos asegura una probabilidad de pérdida de información muy baja.

ii) *Ancho de banda equivalente*. Es una asignación de ancho de banda inferior a la anterior. Depende de las características del servicio y del trayecto que recorrerá la información en la red, de forma que, se explotarán las ventajas de la multiplexación estadística al compartir diferentes conexiones los mismos recursos de la red. El ancho de banda equivalente asignado a una conexión será el mayor de los calculados en la compartición de recursos que emplee, a fin de ofrecer un grado de servicio mínimo por parte de la red. El aumento en la eficiencia de la explotación de los recursos de la red, a través de la multiplexación estadística, tiene como contrapartida las pérdidas de información debidas a la capacidad limitada de los buffers de los nodos y las variaciones del retardo extremo a extremo.

Durante el establecimiento de una conexión se estipula un contrato de tráfico entre el servicio y la red. En este contrato se especifican las características del tráfico generado por la fuente y la calidad de servicio ofrecida por la red (Quality of Service, QoS). A través de los descriptores de tráfico especificados en el contrato usuario-red se determina el encaminamiento de la información y los recursos reservados para la conexión [ReiBer92]. Los descriptores del tráfico empleados en la conexión son [UNI3.0]:

i) *Tasa de pico* ( $R_p$ , peak rate). Es el intervalo de emisión mínimo, especificado como el inverso del menor tiempo entre la generación de dos celdas.

ii) *Tolerancia de la variación del retardo entre celdas* (CDVT, cell delay variation tolerance). Algunos servicios son susceptibles a las variaciones de la distancia entre celdas resultantes del tránsito en la red. Esta variación se introduce en la espera en las colas de los multiplexores y conmutadores y depende de las tasas de generación instantáneas del conjunto de conexiones que comparten los mismos recursos. En el contrato de tráfico se especifica, a través del parámetro CDVT, el valor máximo de variación admisible por el servicio.

iii) *Tasa sostenida de celdas* (Sustainable Cell Rate, SCR) y tolerancia de ráfaga (Burst tolerance,  $\tau_s$ ). Estos parámetros son opcionales y permiten mejorar la explotación de los recursos. La SCR es la expresión de la tasa media generada ( $R_s$ ) durante un intervalo de tiempo ( $\tau_s$ ). A través de este parámetro se puede controlar la generación sostenida de celdas a tasas inferiores a  $R_p$ . Así, el número máximo de celdas

que respetaría el contrato establecido, también denominado Tamaño Máximo de Ráfaga (Maximum Burst Size, MBS), vendría expresado por:

$$\text{MBS} = \left\lceil 1 + \frac{\tau_s}{T_s - T} \right\rceil \quad (1.3.1)$$

donde  $T$  es el tiempo entre llegadas a la tasa de pico. Los parámetros  $\tau_s$  y  $T_s=1/R_s$  permiten controlar el cumplimiento de un valor medio en la tasa de generación en períodos más o menos largos de tiempo.

En el contrato establecido entre el servicio y la red se acuerda también la calidad de servicio (QoS) dada por la red en términos de:

- i) Probabilidad de pérdida de celdas.
- ii) Retardo máximo extremo a extremo.
- iii) Variabilidad máxima del retardo entre celdas consecutivas (*jitter*).

El dimensionado y ubicación de recursos en redes ATM hace necesario un estudio del comportamiento de los servicios en diferentes escalas temporales, realizando una aproximación jerárquica de tres niveles [CasCav91]:

i) *Nivel de llamada*. Se consideran intervalos entre segundos y horas. Este nivel está especialmente indicado para el dimensionado de redes ATM. El dimensionado dependerá de la clase de servicios soportados por la red y la cantidad de éstos que simultáneamente puedan estar en curso. Es imprescindible disponer de las características de estos servicios, relacionadas con el tiempo entre llegadas de nueva petición, la estadística de la duración de la conexión y la dependencia de los puntos de interconexión solicitados en función de la situación geográfica [HofWeb85].

ii) *Nivel de ráfaga*. Se sitúa en el rango de décimas de segundos a pocos segundos. Tiene especial influencia en el proceso de aceptación de nuevas conexiones por la repercusión en la calidad de los servicios. Este nivel está directamente relacionado con la ubicación de recursos, ya que, dependiendo de la naturaleza de los tráficos que compartan los mismos dispositivos de la red, se alcanzará un grado de servicio ofrecido por la red. El nivel de ráfaga de los diferentes tráficos determinará la probabilidad de que, instantáneamente, se alcancen niveles de utilización de los enlaces superiores a 1 y, por tanto, sea necesario el almacenamiento de las celdas en los buffers de los nodos. El grado de servicio ofrecido dependerá directamente de la capacidad de almacenamiento de los buffers de los dispositivos, estableciéndose una relación de compromiso entre el tiempo máximo de espera, o retardo admisible, y la probabilidad de pérdida [FuhBou91].

iii) *Nivel de celda*. Se encuentra en el rango milisegundos. Determina los valores mínimos en el dimensionado de las colas de espera en los nodos, así como, los rangos de variación del retardo entre celdas [ChaLeo94].

Las funciones de control de tráfico en las redes ATM tienen como propósito la gestión y regulación del tráfico en la red. Estos mecanismos son empleados para el control de congestión. La congestión en una red ATM se alcanza cuando uno de sus recursos es explotado a un nivel próximo a su máximo rendimiento. En general, este nivel de utilización se traduce en niveles de ocupación muy elevados de algunas colas de espera de algunos nodos. El control de congestión debe regular el tráfico entrante en la red, de forma que no se sobrepasen los niveles máximos de ocupación dimensionados. El control de congestión se ejerce de dos maneras diferentes:

i) *Control de congestión preventivo*. Las técnicas preventivas de congestión pretenden controlar el tráfico entrante en la red para evitar situaciones de congestión. Para llevar a cabo este control se emplean dos mecanismos:

a) *Control de admisión* (Call Admission Control, CAC). Representa el conjunto de acciones tomadas en la fase de negociación del establecimiento de la conexión para aceptar o rehusar una nueva llamada. En base a los recursos disponibles para ubicar el nuevo tráfico, se evalúa la viabilidad de alcanzar la calidad del servicio demandada por la nueva conexión manteniendo las calidades de las conexiones en curso. La aceptación dependerá del contrato usuario-red que se pretenda establecer y del nivel de explotación de los recursos en el momento de la negociación del contrato.

b) *Control de los parámetros de usuario* (Usage Parameter Control, UPC). Este control de tráfico pretende asegurar que el tráfico entrante en la red cumple con el contrato establecido en el proceso de aceptación de la llamada. También recibe la denominación de control de policía. Los mecanismos de policía tienen como función monitorizar el tráfico entrante en la red para que se cumpla con el contrato usuario-red establecido en la fase de aceptación de la llamada. La acción del mecanismo de policía se traduce en descartar o marcar las celdas que violen el contrato. La marcación de celdas permite establecer una política de selección cuando un dispositivo está congestionado, a través de la asignación de niveles de prioridad. Así, en estados de congestión, la celdas de prioridad más baja no son almacenadas en las colas de espera.

ii) *Control de congestión reactivo*. La regulación del tráfico se ejerce cuando la red ha alcanzado un estado de ocupación próximo a la saturación de un recurso. Llegada esta situación, la red puede hacer uso de tres mecanismos:

- a) Señalización a las fuentes de tráfico que comparten los recursos gestionados para que disminuyan su tasa de generación.
- b) Encaminamiento alternativo de las conexiones sobre otros recursos con un grado de congestión inferior.
- c) Descarte de celdas almacenadas en las colas de espera en base al nivel de prioridad y el grado de servicio acordado en la conexión.

### 1.3.1. Control preventivo

Las técnicas de control preventivo propuestas en las redes ATM se basan en el CAC y en el UPC. Las estrategias de aceptación de nuevas conexiones en ATM son actualmente un importante campo de investigación. El CAC está directamente relacionado con la ubicación de recursos en la red y depende de la caracterización de las fuentes de tráfico soportadas por la red. Dentro de las diferentes técnicas propuestas para la determinación de los recursos necesarios de una conexión han recibido gran acogida las desarrolladas en los trabajos de [Hui88], [AniMit82] y [GueGun92]. Una asociación directa entre los parámetros aceptados como descriptores de tráfico y la asignación de recursos se plantea en [ShiSai95].

Para asegurar que el servicio cumple con los parámetros descriptores de tráfico prefijados durante el transcurso de la llamada, la red dispone de un dispositivo denominado mecanismo de policía. La función del mecanismo de policía es monitorizar en el acceso a la red el tráfico cursado y actuar eliminando celdas cuando se superen las cotas de consumo de ancho de banda acordadas. Dos tipos de mecanismos de policía se están estudiando unos basados en una ventana deslizante y otros basados en sistemas de créditos como el Leaky Bucket y todas sus variantes [Rath91]. En [ZamMat93] se realizó un estudio comparativo de las principales propuestas de funciones de policía. Los mecanismos analizados son:

i) *Cubo agujereado* (Leaky Bucket, LB). Este mecanismo fue propuesto en [Tur86] y se basa en una generación de créditos a intervalos regulares de tiempo ( $I$ ) que se almacenan en una cola de dimensión finita ( $L$ ), de forma que, cada celda generada por el servicio consume un crédito. Habitualmente, este mecanismo se implementa con un simple contador, cuyo valor inicial, y máximo, es  $L$ . El contador se incrementa periódicamente en una unidad cada intervalo  $I$ . Si la tasa de generación de celdas es superior a la tasa de llegada de los créditos, el número de créditos del contador se decrementa a la tasa diferencial entre la tasa de las llegadas y la de la generación de créditos. Cuando una celda es transmitida y no se dispone de créditos, el mecanismo entonces actúa marcándola o descartándola.

ii) *Mecanismo de ventana saltante* (Jumping Window, JW). Se habilita un tiempo de observación en el enlace (T) que abarca un número de slots del enlace. Durante este período se cuentan el número de celdas transmitidas por el servicio. Todas las celdas que superan el número de orden fijado (N) son marcadas o descartadas. Una vez transcurrido el intervalo T, se procede a un nuevo conteo en el intervalo adyacente. Una variante de este mecanismo es el Triggered Jumping Window (TJW), cuyos intervalos no son consecutivos sino que se inician con la transmisión de la siguiente celda fuera del anterior.

iii) *Mecanismo de ventana deslizante* (Sliding Window, SW). Consiste en la monitorización del flujo de celdas que atraviesa un intervalo de tiempo, o ventana de observación, que se desplaza slot a slot. Solo un número máximo de celdas (N) es admitido, o no marcado, en el intervalo T. Este mecanismo permite una tasa máxima de  $N/T$  celdas. Esta función de policía requiere mayor complejidad en su implementación que las anteriores por su continua necesidad de recálculo.

iv) *Mecanismo de ponderación exponencial de la media móvil* (Exponentially Weighted Moving Average, EWMA). En este caso se habilita una ventana saltante para la cual se especifica dinámicamente un número máximo de celdas admisibles. Así, el número máximo de celdas aceptable en la ventana  $i$ -ésima ( $N_i$ ) depende del número medio de celdas admisibles (N) junto con una media ponderada de las celdas aceptadas en las ventanas anteriores ( $X_j$ ). La relación aplicada puede expresarse:

$$N_i = \frac{N - \gamma S_{i-1}}{1 - \gamma} \quad \gamma \in [0, 1] \quad (1.3.2)$$

donde,

$$S_{i-1} = (1 - \gamma)X_{i-1} + \gamma S_{i-2} \quad (1.3.3)$$

El factor  $\gamma$  depende de la relación entre la tasa de pico y la tasa media especificada en el contrato usuario-red. El EWMA también es un algoritmo con un alto coste de implementación, respecto a los presentados anteriormente.

v) *Mecanismo controlador-espaciador* (Spacer Controller, SC). A diferencia de las otras funciones de policía, este mecanismo no sólo monitoriza el flujo de celdas generadas por la fuente sino que interviene activamente sobre el flujo. Realiza la operación de espaciado cuando dos celdas superan la tasa instantánea máxima permitida, aunque, esta actuación sólo se produce sobre un número consecutivo máximo de celdas. Este mecanismo se basa en dos etapas: una de supervisión del contrato establecido y otra de espaciado cuando sea necesario. El funcionamiento se regula con dos parámetros:

a) Tiempo mínimo de separación de dos celdas consecutivas (Theoretical Reemission time, TRT). A través de este parámetro se determina si la celda se ha de retransmitir inmediatamente, se debe espaciar o se debe descartar.

b) Instante real de reemisión (Actual Reemision Time, ART). Este parámetro fija el instante en que se debe transmitir la celda en el medio ranurado. Se especifica su valor en slots.

Una versión simplificada del algoritmo de funcionamiento del controlador-espaciador se muestra en la figura 1.3.1. Obsérvese que se admite superar el tiempo mínimo entre llegadas un valor  $\tau$ . Este tiempo depende de la variación del retardo admisible especificado en el contrato usuario-red.

El mecanismo de policía adoptado como estándar para las redes ATM es una variación del Leaky Bucket que considera, con valores no necesariamente enteros, los parámetros de llegada de los créditos (I) y la capacidad de almacenamiento de éstos (L) [I.371]. Esta versión de estado continuo del LB se denomina algoritmo de tasa genérica de celdas (Generic Cell Rate Algorithm, CGRA). Existen dos expresiones algorítmicas equivalentes de este mecanismo de policía presentadas en la recomendación [UNI3.0]. En la figura 1.3.2 se presenta una versión del leaky bucket de estado continuo, GCRA(I, L).

Junto con los mecanismos de policía, se ha considerado como control preventivo la conformación de tráfico (Traffic Shaping). En algunos esquemas de control de policía LB se ha añadido un buffer a la entrada del UPC para el almacenamiento de las celdas que violen el contrato establecido. La inserción de este buffer pretende reducir la probabilidad de descarte de celdas a través del suavizado del tráfico entrante a la red.

De forma más genérica, se ha contemplado en [UNI3.0] la conformación de tráfico como un mecanismo que se puede emplear sobre el flujo de celdas de una o varias conexiones. Su propósito es reducir la variabilidad del tráfico generado por las fuentes a través de técnicas de suavizado. La conformación de tráfico se aplica para la reducción de la tasa de pico, la limitación de la longitud de las ráfagas y la minimización de la CDV, a través de un adecuado espaciado de las celdas. La aplicación de las técnicas de suavizado tiene notables efectos como la reducción de la probabilidad de pérdida de las celdas o la disminución de la ubicación de recursos necesarios para cada conexión [MitAnd94] [RosFra94] [PanZar93]. La desventaja que presentan las técnicas de suavizado es el incremento del retardo extremo a extremo, ya que, se requiere de un tiempo de almacenamiento para la conformación. Las técnicas empleadas en el suavizado de tráfico de vídeo se pueden clasificar en:

i) *Conformación por almacenamiento*. Se basan en el almacenamiento de un conjunto de celdas generadas por la fuente durante un intervalo de tiempo. Posteriormente, las celdas son entregadas a la red a una tasa constante igual al valor medio obtenido en el intervalo de almacenamiento.

ii) *Conformación predictiva*. A través de la caracterización de la fuente de tráfico, con el conocimiento de la correlación de la tasa binaria se prevee la tasa binaria media

del próximo intervalo de tiempo. Esto reduce el tiempo de almacenamiento notablemente, aunque, la continua corrección de la tasa media provoca que el flujo de celdas sufra variaciones en su espaciado. Esta variación se reduce cuanto mejor sea la predicción y menor sea la variabilidad del tráfico de la fuente.

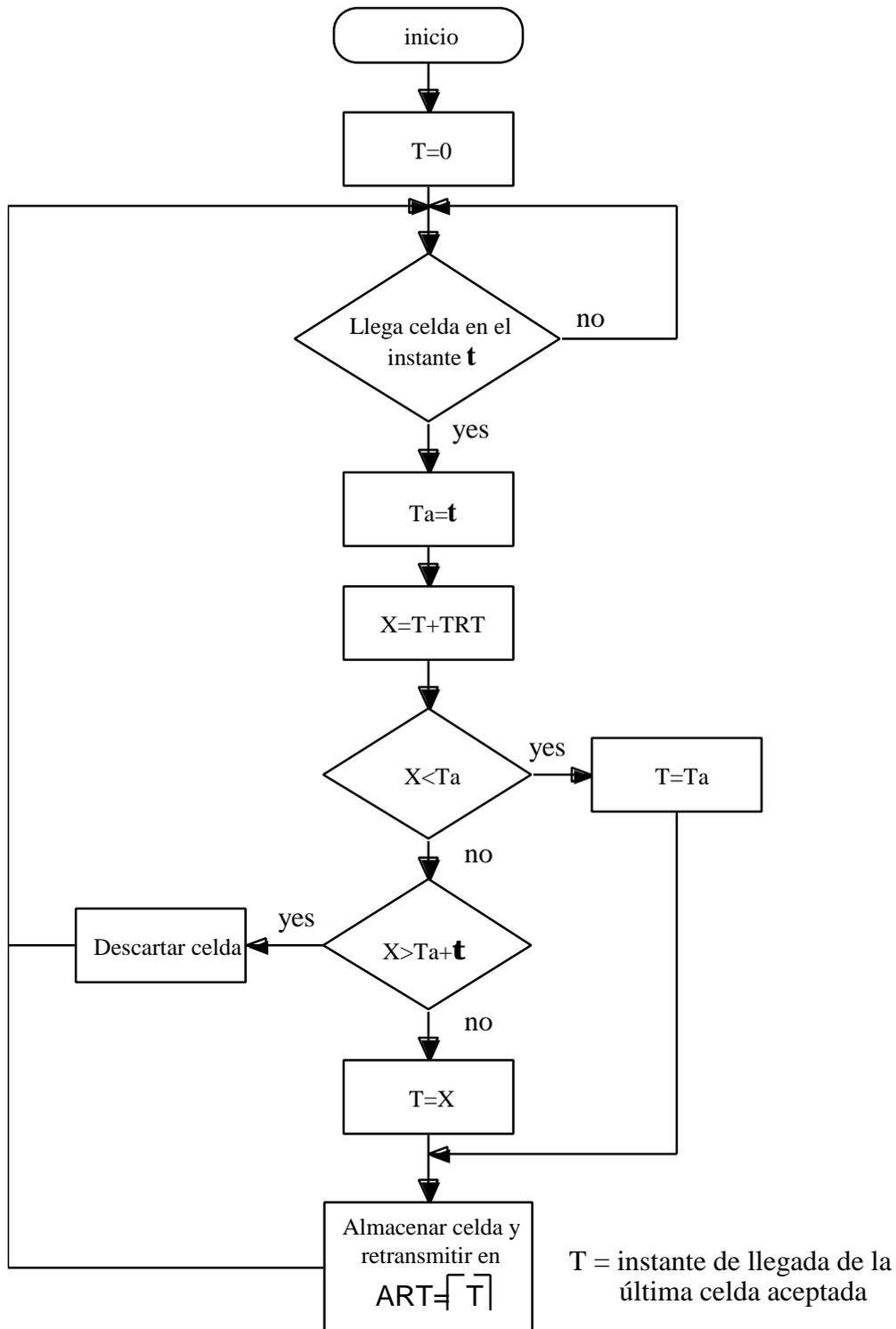


Figura 1.3.1. Mecanismo de funcionamiento del controlador-espaciador

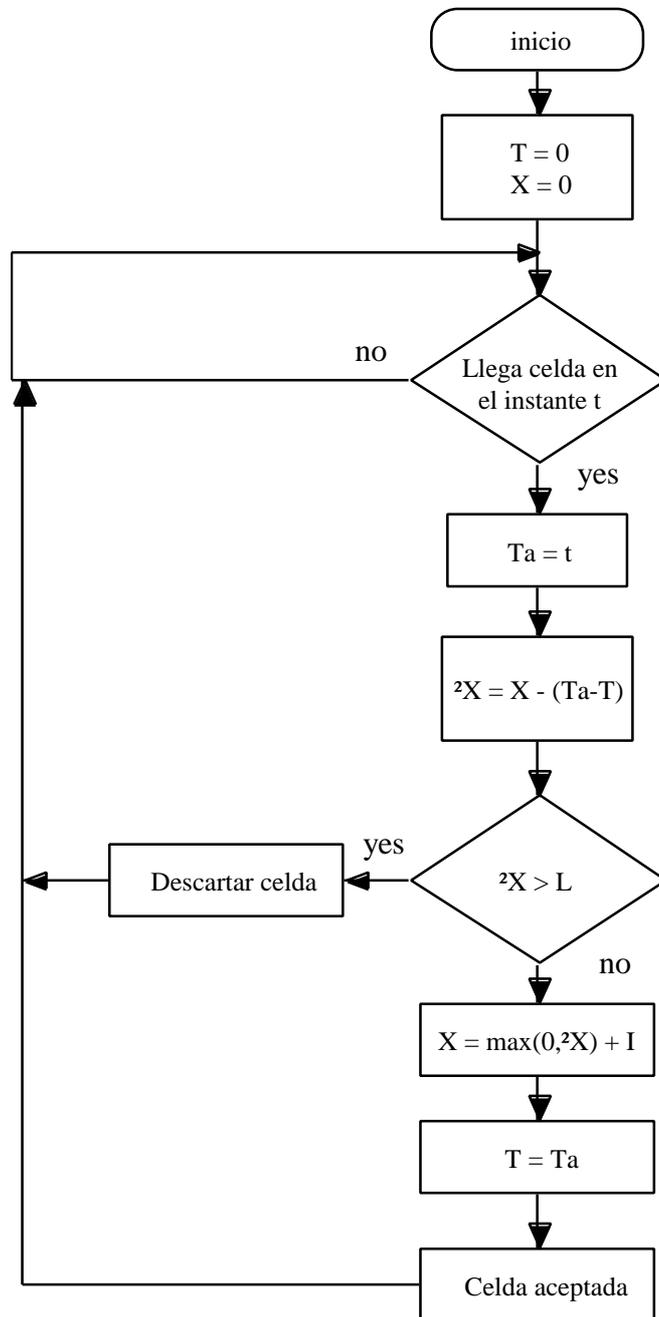


Figura 1.3.2. Algoritmo de funcionamiento del mecanismo de policía GCRA en la versión Leaky Bucket de estado continuo

### 1.3.2 Control reactivo

El comportamiento estadístico de las distintas fuentes de tráfico que comparten los recursos de la red provoca la aparición de estados de congestión en la red a pesar de emplear adecuados criterios en el control de admisión. Ante la aparición de tales situaciones es imprescindible una continua monitorización del estado de los dispositivos de la red. El control de tráfico ejercido por la red pretende evitar que se alcancen niveles de congestión en los dispositivos. Dentro de las funciones disponibles por la red para evitar las situaciones de congestión la regulación del tráfico de la fuente se puede realizar a través de técnicas reactivas [GaiPuj96]. Asimismo, se pueden emplear técnicas de descarte de celdas, según su nivel de prioridad y grado de servicio acordado por la red, o reencaminamiento de las conexiones.

Las técnicas reactivas por señalización se basan en indicar al servicio la variación de la tasa que debe realizar. Los mecanismos propuestos para llevar a cabo este control reactivo en redes ATM han sido:

- i) Control extremo a extremo basado en tasa.
- ii) Control punto a punto basado en créditos.

El control por tasa es mucho más adecuado para las redes de alta velocidad debido a la reducción de la complejidad de los nodos de la red. El control por tasa se realiza notificando a la fuente el aumento o reducción de su tasa. Para evitar oscilaciones en el tráfico de la red se suele emplear un aumento lineal y un decremento exponencial de la tasa de la fuente. El control basado en créditos emplea el clásico mecanismo de ventana deslizante para cada enlace lo cual lo hace desaconsejable dada la complejidad introducida en cada nodo de la red cuando éste soporta un elevado número de conexiones.

Dentro de las técnicas reactivas por tasa se han propuesto dos modos de señalización:

i) *Señalización hacia adelante* (Explicit Forward Congestion Indication, EFCI). La señalización se realiza en el nodo congestionado marcando las celdas de la conexión que pasan por el dispositivo congestionado. El destino de la conexión envía celdas de gestión en sentido opuesto para comunicar a la fuente el estado de congestión.

ii) *Señalización hacia atrás* (Backward Explicit Congestion Notification, BECN). En este caso el nodo congestionado informa directamente a la fuente. Esta señalización no está estandarizada actualmente para las redes ATM.

Originalmente, el mecanismo EFCI se planteó para servicios VBR aunque se ha observado su adecuación para los servicios ABR [CheLiu96]. El ATM Forum adoptó el mecanismo de control reactivo por tasa extremo a extremo Enhanced Proportional Rate

Control Algorithm (EPRCA) para las redes ATM en septiembre de 1994. El mecanismo adoptado para el control reactivo se basa en la marcación de las celdas de información enviadas desde la fuente al destino. Para la señalización, el nodo fuente inyecta celdas de gestión hacia el nodo destino. Estas celdas incorporan un campo de sentido en el que circulan y un bit de indicación de congestión. Tanto el receptor como los nodos intermedios pueden alterar el indicador de señalización cuando la celda retorna hacia el emisor. El nodo fuente sólo puede aumentar linealmente la tasa de generación cuando la celda de gestión llega a su origen y no señala un estado de congestión. En caso contrario, la fuente decremente multiplicativamente su tasa.

### 1.4. Transmisión de vídeo MPEG sobre redes ATM

Los nuevos servicios que soporta la RDSI-BA incorporan, en su mayoría, información de vídeo digital. La información de vídeo previamente se comprime antes de ser entregada a la red. El MPEG es el mecanismo de compresión estandarizado mayormente empleado para este propósito. El proceso llevado a cabo en la transmisión de vídeo digital en redes ATM puede ser expresado a través de la arquitectura de referencia mostrada en la figura 1.4.1 [HanOro95]. El equipo terminal se conecta a través del nodo de acceso a la red ATM. El nodo de acceso realiza las funciones de multiplexación en el nivel ATM, añadiendo los campos de información necesarios para realizar funciones de corrección de errores, detección de celdas pérdidas y asignación de prioridades.

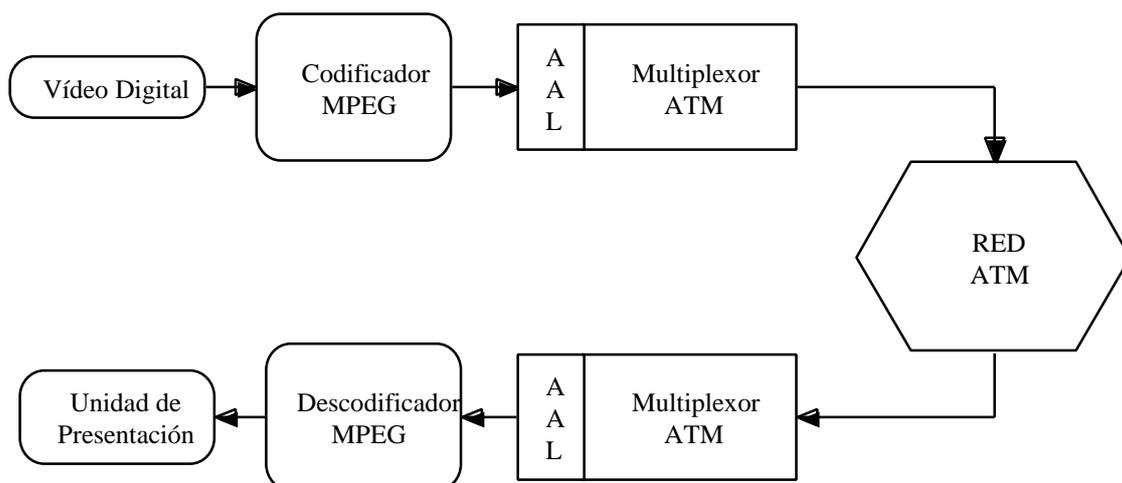


Figura 1.4.1. Arquitectura de referencia de la transmisión de vídeo digital sobre ATM

Tras la aparición del MPEG-II, el ATM Forum inició los estudios del transporte de información multimedia sobre redes ATM. Las primeras discusiones se centraban entre emplear la sintaxis del entramado MPEG-II, para confeccionar la multiplexación de los diferentes tipos de información, o establecer diferentes conexiones, de forma que la multiplexación se realizará en la propia red. Los parámetros que permitían evaluar las ventajas de cada técnica de multiplexación eran: los mecanismos de corrección de errores, la variación del retardo de celdas y la sincronización [ATMFORUM/94-0570].

Como resultado previo, se llegó a la conclusión de que era más adecuado realizar la transmisión empleando el entramado MPEG-II ( MPEG-II Transport Stream, MTS) por las ventajas de sincronización que ofrecía en recepción, aprovechando las referencias temporales (Program Clock Reference, PCR) que incorpora [VerSam96]. En cuanto a la recuperación de reloj se propuso el mecanismo basado en la detección de fase de un PLL digital planteado anteriormente en [VerPin88].

De entre los mecanismos de adaptación que permiten multiplexación se estudió el AAL3/4. Esta clase de adaptación soporta, sobre una misma conexión, todos los flujos de información. Este mecanismo estaba diseñado inicialmente para tráficos de clases C y D. Tras el análisis del MTS realizado en [ATM-FORUM/93-1010] y [ATM-FORUM/93-797] para ubicarlo en redes ATM, se llegó a la conclusión de emplear el modo de adaptación AAL5. Este mecanismo permite mapear directamente las unidades básicas de 188 octetos generados por el MTS en dos celdas ATM con un remanente de señalización de 8 octetos de las unidades de datos AAL-5. A su vez, esta adopción facilita las labores de interconexión de redes cuando los usuarios del servicio están conectados en sistemas de comunicación heterogéneos.

Otra de las ventajas que ofrece la sintaxis MPEG-II es la especificación de diferentes niveles que pueden ser empleados para distinguir resoluciones, espaciales o temporales, y diferentes calidades de imagen. Este concepto, denominado escalabilidad, es extraordinariamente útil para compatibilizar la recepción simultánea en diferentes equipos de presentación. A su vez, se permite, con facilidad, la elección del nivel de prioridad para las celdas, de forma que las celdas que transportan información fundamental para obtener una calidad básica en la imagen se marcan con un nivel de prioridad superior a las celdas que se emplean para transferir la información complementaria [AigFaz95] [RiyLam95].

Esta técnica de escalabilidad de la relación señal a ruido (signal to noise ratio, SNR) también puede ser empleada, distinguiendo, en vez de niveles de prioridad, dos posibles conexiones con diferentes calidades de servicio y parámetros. Este tipo de transmisión recibe el nombre de bi-nivel, dado que, en una conexión se inserta la información básica y en la otra la complementaria. Se han realizado diferentes estudios sobre las ventajas de la transmisión bi-nivel y con señalización de prioridad. Las características de la transmisión bi-nivel permiten emplear para cada flujo de información un tipo de adaptación diferente y una calidad de servicio distinta. El tipo de servicio más adecuado

para el nivel básico es un servicio de clase A, manteniendo una tasa constante, y para el nivel complementario es un servicio de clase B, con tasa variable. Dentro de los diferentes estudios realizados se aconseja la ubicación de un 20% del ancho de banda en el nivel básico y el resto en el nivel complementario. Obviamente, la transmisión bi-nivel también puede realizarse con dos conexiones de tasa variable, aunque, esta transmisión plantea mayores problemas de sincronización e incrementa la probabilidad de pérdida por la multiplexación estadística del nivel básico. En la figura 1.4.2 se plantea el esquema básico de transmisión bi-nivel.

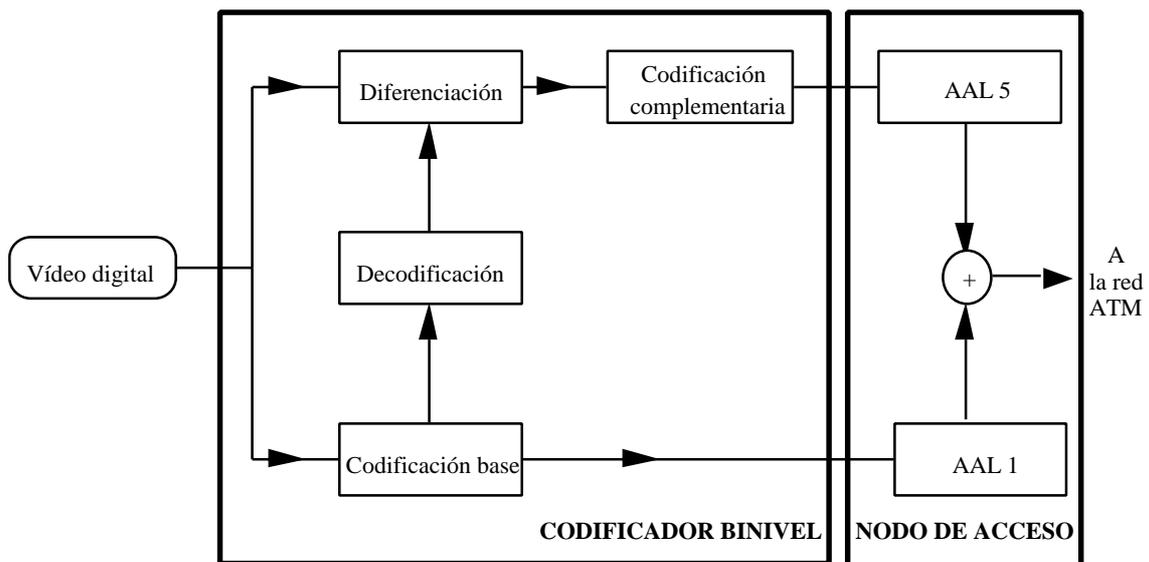


Figura 1.4.2. Esquema de codificación y transmisión binivel

## 1.5. Objetivo y presentación de la tesis

La evaluación del comportamiento de los distintos dispositivos y enlaces que componen las redes se realiza actualmente a través de la confección de modelos de los distintos elementos que intervienen en la comunicaciones. En particular, es necesaria la determinación de modelos de fuentes de tráfico de vídeo de velocidad variable, dado que, este tipo de tráfico será el principal consumidor de los recursos de la red. La aportación de un modelo, que capture con precisión la generación del tráfico de este tipo de fuentes, permite la evaluación de los distintos dispositivos de la red a través de análisis matemáticos, cuando sea tratable el problema, o mediante el desarrollo de simulaciones [RigFra91]. La aportación de modelos de tráfico contribuye a:

- i) El desarrollo de técnicas para la asignación de recursos de la red y el desarrollo de algoritmos de encaminamiento.

- ii) La determinación de los parámetros descriptores de tráfico y el diseño de mecanismos de policía que monitoricen dicho parámetros.
- iii) La cuantificación de las ventajas de la multiplexación estadística y el desarrollo de estrategias de multiplexación atendiendo a la QoS contratada por las conexiones que comparten un recurso.
- iv) La evaluación de arquitecturas de conmutación.
- v) La definición de los umbrales de actuación de los controles reactivos en estados de congestión.
- vi) La maximización de la explotación de los recursos de la red a través de la renegociación dinámica, durante la conexión, de los recursos asignados.

La inferencia de los modelos de tráfico para fuentes de vídeo MPEG VBR pasa primero por una etapa de análisis del algoritmo de codificación. En el capítulo 2, se evalúa la influencia de los distintos parámetros controlables del algoritmo MPEG en la calidad de la imagen y en el rango de compresión. Este análisis se realiza empleando criterios objetivos y subjetivos, determinando una adecuada elección del valor de los parámetros en función del nivel de calidad deseado. Interpretando el tráfico generado como una variable aleatoria, se realiza un estudio exhaustivo de los distintos parámetros y funciones que lo pueden caracterizar. Este análisis nos permitirá establecer las relaciones entre el valor de las variables que controlan el algoritmo de compresión MPEG y el comportamiento estadístico del tráfico generado por el codificador. Algunos de estos resultados fueron adelantados en [MatSal94].

La confección de modelos de tráfico de vídeo MPEG VBR se discute en el capítulo 3. A partir de los resultados obtenidos en el capítulo 2, se ajustan distintos modelos ligados a la tasa binaria generada por el codificador en tiempos de cuadro y de grupo de imágenes. Asimismo, se desarrolla un predictor de la tasa binaria generada en un cuadro para la suavización del tráfico entregado a la red. Los modelos propuestos son autoregresivos y markovianos. Los modelos autoregresivos tienen la ventaja de ser fácilmente programables a la hora de realizar simulaciones pero no ofrecen la posibilidad de resultados analíticos cuando se estudian junto con los modelos de otros elementos de la red. Los modelos basados en procesos de Markov, aunque son difíciles de programar, presentan la ventaja de que, en algunos casos, se pueden obtener resultados analíticos al estudiarlos junto a modelos de otros dispositivos. Parte de los resultados de este capítulo se presentaron en [MatSal95].

El análisis de la transmisión de vídeo MPEG VBR sobre redes ATM se desarrolla en el capítulo 4. En este capítulo se cuantifican las ventajas de la multiplexación estadística para el tráfico MPEG VBR. A través de la caracterización del tráfico multiplexado se propone una nueva técnica de modelado de tráfico agregado, basada en un modelo autoregresivo. Algunos de los resultados obtenidos fueron avanzados en [MatSal96] y [PagMat95]. A partir de los resultados obtenidos en el modelado de fuentes simples

mediante un modelo markoviano, se plantea un nuevo esquema de reserva dinámica de recursos. A través de la caracterización del comportamiento del codificador MPEG VBR realizado en el capítulo 2, también se considera en el esquema el control reactivo para fuentes de tráfico de este tipo.

Finalmente, en el capítulo 5 se presentan la conclusiones y algunos de los posibles trabajos futuros que se han derivado del desarrollo de la presente tesis.