

REDES Y TRANSMISION DE DATOS

Pable Gil / Jorge Pomares / Francisco Candelas



TEXTOS DOCENTES

*PUBLICACIONES
UNIVERSIDAD DE ALICANTE*

Este libro ha sido debidamente examinado y valorado por evaluadores ajenos a la Universidad de Alicante, con el fin de garantizar la calidad científica del mismo.

Publicaciones de la Universidad de Alicante
Campus de San Vicente s/n
03690 San Vicente del Raspeig
Publicaciones@ua.es
<http://publicaciones.ua.es>
Teléfono: 965903480
Fax: 965909445

© los autores, 2010
© de la presente edición: Universidad de Alicante

ISBN: 978-84-9717-125-0
Depósito legal: MU 1491-2010

Diseño de portada: candela ink.
Composición: Compobell, S.L.
Impresión y encuadernación: Compobell, S.L.



Reservados todos los derechos. Cualquier forma de reproducción, distribución, comunicación pública o transformación de esta obra sólo puede ser realizada con la autorización de sus titulares, salvo excepción prevista por la ley. Diríjase a CEDRO (Centro Español de Derechos Reprográficos, www.cedro.org) si necesita fotocopiar o escanear algún fragmento de esta obra.

ÍNDICE

PRÓLOGO
CAPÍTULO 1. INTRODUCCIÓN A LAS REDES DE COMPUTADORES.....
1.1. Definición y conceptos
1.1.1. Modelo de comunicaciones
1.1.2. Tareas de un sistema de comunicaciones
1.1.3. Objetivos de las redes
1.2. Tipos de redes y topologías.....
1.2.1. Clasificación según topología física
1.2.2. Clasificación según escala geográfica
1.2.3. Clasificación según la forma de establecer la comunicación
1.3. Arquitectura de red
1.3.1. Modelo de referencia OSI/ISO.....
1.3.2. Modelo de referencia TCP/IP
CAPÍTULO 2. TRANSMISIÓN DE SEÑALES
2.1. Introducción. Funciones de la capa física
2.2. Fundamentos de la transmisión de datos
2.2.1. Tipos de señales y parámetros de una señal
2.2.2. Desarrollo en serie de Fourier
2.2.3. Espectro de potencias
2.2.4. Ancho de banda
2.2.5. Ganancia
2.2.6. Prestaciones del medio físico
2.3. Perturbaciones en la transmisión.....
2.4. Velocidades máximas de envío de datos
2.4.1. Teorema de muestreo.....
2.4.2. Velocidad máxima según el teorema de muestreo
2.4.3. Velocidad máxima según la relación señal/ruido.....

2.5.	Filtrado de señales
2.6.	Transmisión analógica y digital. Definición y conceptos básicos
2.7.	Modos de transmisión.....

CAPÍTULO 3. CODIFICACIÓN DE LA INFORMACIÓN.....

3.1.	Señales y tipos de señalización
3.1.1.	Tipos de datos y de señales
3.1.2.	Tipos de señalización en transmisiones digitales.....
3.2.	Codificación Banda Base.....
3.2.1.	Codificación binaria.....
3.2.2.	Codificación Manchester
3.2.3.	Codificación Manchester diferencial.....
3.2.4.	Codificación mBnL.....
3.2.5.	Otras codificaciones
3.3.	Codificación Banda Modulada
3.3.1.	Modulación analógica.....
3.3.2.	Modulación digital.....
3.4.	Multiplexación
3.4.1.	Multiplexación por división en frecuencias
3.4.2.	Multiplexación por división en tiempos.....

CAPÍTULO 4. MEDIOS DE TRANSMISIÓN

4.1.	Introducción
4.2.	Medios Guiados
4.2.1.	Par trenzado
4.2.2.	Cable coaxial
4.2.3.	Fibra óptica
4.3.	Medios no guiados.....
4.3.1.	Ondas de radio.....
4.3.2.	Microondas
4.3.3.	Infrarrojos

CAPÍTULO 5. DISEÑO DEL NIVEL DE ENLACE Y CONTROL DE ERRORES

5.1.	Servicios del nivel de enlace
5.2.	Funciones del nivel de enlace.....
5.3.	Iniciación y terminación de la comunicación
5.4.	Delimitación de tramas.....
5.4.1.	Delimitación temporal
5.4.2.	Delimitación por número de caracteres.....
5.4.3.	Delimitación por caracteres especiales.....
5.4.4.	Delimitación por bits especiales.....

5.4.5. Delimitación por códigos de línea especiales
5.5. Direccionamiento de tramas
5.6. Detección y corrección de errores en tramas
5.6.1. Comprobación de paridad.....
5.6.2. Comprobación de redundancia cíclica (CRC).....
5.6.3. Corrección de errores empleando códigos de bloque
5.6.4. Los códigos Reed-Solomon.....
5.7. Medición de la tasa de errores

CAPÍTULO 6. CONTROL DE FLUJO EN EL NIVEL DE ENLACE.....

6.1. Introducción
6.2. Protocolo de parada y espera.....
6.2.1. Parada y espera en ausencia de errores.....
6.2.2. Parada y espera considerando que se producen errores.....
6.2.3. Parada y espera con confirmaciones negativas
6.2.4. Protocolo bilateral de parada y espera
6.2.5. Análisis de las prestaciones del protocolo de parada y espera.....
6.3. Protocolo de ventana deslizante
6.3.1. Características del protocolo de ventana deslizante.....
6.3.2. Análisis de las prestaciones del protocolo de ventana deslizante

CAPÍTULO 7. PROTOCOLOS ESTANDARIZADOS DEL NIVEL DE ENLACE

7.1. Técnicas de compartición del medio
7.1.1. Técnicas de contienda.....
7.1.2. Selección distribuida.....
7.2. Técnicas de control del enlace: Protocolo HDLC.....
7.2.1. Tipos de estaciones y enlaces.....
7.2.2. Modos de funcionamiento
7.2.3. Estructura de la trama.....
7.2.4. Tipos de tramas.....
7.2.5. Repertorio básico de instrucciones.....
7.2.6. Control del enlace en Modo Normal No Balanceado.....
7.2.7. Control del enlace en Modo Asíncrono No Balanceado
7.2.8. Control del enlace en Modo Asíncrono Balanceado.....
7.2.9. Configuración del medio de transmisión y transferencia de datos.....
7.2.10. Transferencia de datos con rechazo
7.3. Protocolo PPP
7.3.1. Estructura de la trama.....
7.3.2. Tipos de tramas.....
7.3.3. Funcionamiento de PPP.....

CAPÍTULO 8. DISEÑO DEL NIVEL DE RED.....
8.1. Nivel de red.....
8.2. Servicios de la capa de red.....
8.3. Métodos de conmutación.....
8.4. Algoritmos de encaminamiento.....
8.4.1. Encaminamiento fijo o estático
8.4.2. Encaminamiento por inundación
8.4.3. Encaminamiento adaptativo o dinámico
CAPÍTULO 9. ENCAMINAMIENTO DEL NIVEL DE RED.....
9.1. Introducción
9.2. Tipos de protocolos de encaminamiento
9.3. Algoritmos de cálculo de ruta
9.3.1. Algoritmo de Dijkstra.....
9.3.2. Algoritmo de Bellman-Ford
9.4. Protocolo OSPF
9.4.1. Acciones OSPF para construir encaminamiento.....
9.4.2. OSPF encaminando a través de múltiples áreas
9.4.3. Tipos de paquetes OSPF.....
9.5. Protocolo RIP
9.5.1. Características de RIP.....
9.5.2. Acciones RIP para construir encaminamiento
9.6. Protocolos EGP: BGP.....
BIBLIOGRAFÍA

Prólogo

Este libro es adecuado para ser empleado como texto básico de referencia por estudiantes de redes computadores en cualquier titulación de grado en ingeniería. Sus contenidos se han diseñado para cubrir aspectos tanto básicos como específicos de: transmisiones y codificación de datos, multiplexación de datos, control de enlace de datos, conmutación de paquetes, así como algoritmos y protocolos de enrutamiento de las redes LAN más extendidas.

Para facilitar la comprensión, este texto se ha estructurado en nueve capítulos. Cada uno de los cuales presenta un resumen general, un breve índice de contenidos, una lista de objetivos básicos específicos que se pretende alcanzar y por último, un índice de palabras claves para facilitar la búsqueda conceptual. Finalmente, destacar que este libro nace con la intención de ayudar en el proceso de aprendizaje a todos aquellos estudiantes de redes y transmisión de datos. Por este motivo, a lo largo del libro, se proporcionan ilustraciones, gráficos, ejemplos y ejercicios modelo que buscan presentar los contenidos de una manera concisa y con la mayor precisión posible.

Capítulo 1 . INTRODUCCIÓN A LAS REDES DE COMPUTADORES

En este capítulo se proporciona una visión general de las redes de computadores. Así, se presenta una descripción general de las comunicaciones de datos y la tipología de redes que se emplean. Además este capítulo incluye una descripción de las arquitecturas de red OSI y TCP/IP.

CAPÍTULO 1

1-1 Definición y conceptos

- Modelo de comunicaciones
- Tareas de un sistema de comunicaciones
- Objetivos de las redes

1-2 Tipos de redes y topologías

- Clasificación según la topología física
- Clasificación según la escala geográfica
- Clasificación según la forma de establecer la comunicación

1-3 Arquitectura de red

- Modelo de referencia OSI/ISO
- Modelo de referencia TCP/IP

OBJETIVOS

- Explicar los elementos y procesos que intervienen en una comunicación entre dispositivos
- Describir las ventajas que ofrece una red de dispositivos
- Clasificar redes de dispositivos en función de su forma geométrica, modo de conexión, área geográfica que ocupan, ámbito de empleo y gestión o forma de establecer la comunicación
- Describir el concepto de arquitectura de red y la comunicación entre los elementos que lo constituyen
- Presentar dos modelos de arquitectura de red y discutir sus diferencias: OSI/ISO – TCP/IP

TERMINOLOGÍA – PALABRAS CLAVE

- | | | |
|-----------------|----------------------------|-------|
| • Difusión | • Red privada | • ICI |
| • Punto a punto | • Red pública | • SDU |
| • Multipunto | • Conmutación de circuitos | • IDU |
| • Bus | • Conmutación de paquetes | • SAP |
| • Anillo | • ISO/OSI | • PCI |
| • Estrella | • TCP/IP | • PDU |
| • Malla | • Capa o Nivel | |
| • LAN | • Protocolo | |
| • MAN | • Entidad | |
| • WAN | • Interfaz | |

1.1. Definición y conceptos

Una comunicación, sea del tipo que sea, se puede entender como un intercambio de información entre entidades. En la vida cotidiana es posible observar muchos sistemas de comunicaciones. Ejemplos de ello es el intercambio de señales de voz entre dos usuarios empleando dos terminales de telefonía móvil y/o fija. También es un sistema de comunicaciones el que se produce entre dos usuarios que chatean empleando aplicaciones de mensajería instantánea, o cuando un usuario accede a un servidor web para enviar o descargar ficheros de datos.

Para que se produzca este intercambio de información entre entidades (sea del tipo que fuera), es necesario un proceso que involucra la interconexión de dispositivos, es decir la conexión de computadores personales, teléfonos, cableados y medios o dispositivos especiales de interconexión de redes. Por lo tanto, una red de comunicaciones no es más que un conjunto de dispositivos autónomos con capacidad de interconexión.

El proceso de intercambio de datos o información se denomina transmisión de datos. Y además, cualquier sistema de transmisión de datos está formado por cinco componentes básicos: emisor, mensaje, receptor, medio y protocolo.

El mensaje es la información o datos a comunicar. El emisor es el dispositivo que envía los datos del mensaje. El receptor es el dispositivo que recibe el mensaje. El medio es el camino físico por el cual viaja el mensaje desde el emisor al receptor. Y finalmente, el protocolo es un conjunto de reglas que gobiernan la transmisión de datos.

En una comunicación entre dos individuos dotados de terminal de telefonía móvil, el emisor y el receptor serían los dos teléfonos, el mensaje aquello que se quiere comunicar por ejemplo solicitar apuntes de la asignatura, el medio sería el aire, antenas receptoras y/o satélites en su caso, y el protocolo el lenguaje empleado para el entendimiento entre ambos individuos, por ejemplo el castellano.

1.1.1. Modelo de comunicaciones

El modelo simplificado de un sistema de comunicaciones consta de un sistema origen que funciona como emisor y de un sistema destino que hace las funciones de receptor. El sistema origen está formado por dos elementos principales llamados fuente y transmisor, y del mismo modo el sistema receptor, también, está formado por dos elementos. En este caso, los elementos se denominan receptor y destino. La fuente es un dispositivo que genera un conjunto de datos a transmitir y el transmisor es un dispositivo que transforma y codifica los datos que se quieren transmitir para adecuarlos al sistema de transmisión empleado. Esto es así porque, en general, la información y datos no se pueden enviar en el mismo formato que la fuente los genera. Éstos deben ser convertidos a una señal cuyo formato depende del sistema de transmisión que puede ser desde una sencilla línea telefónica a una compleja red de distribución inalámbrica de datos. Además hace falta un sistema destino constituido por un elemento receptor capaz de volver a transformar la señal adecuándola de nuevo a un formato inteligible por el dispositivo destino (Figura 1.1a).

Como sistema de comunicaciones simplificado se puede entender el intercambio de información que se produce entre un computador personal (del inglés PC: 'Personal Computer') y un servidor web. En este sencillo sistema, el computador personal funciona como fuente, el módem instalado en el PC funciona como dispositivo transmisor codificando una secuencia de caracteres, bits de datos, etc. en señales y el sistema de transmisión viene determinado por la compleja red de

telefonía pública. Finalmente, el dispositivo receptor es el módem instalado en la máquina servidora que se encarga de volver a transformar la señal transmitida en un conjunto de datos e información adecuada para ser manejada por el destino. Esto es bits o caracteres. En este caso, el destino es el servidor web (Figura 1.1b).

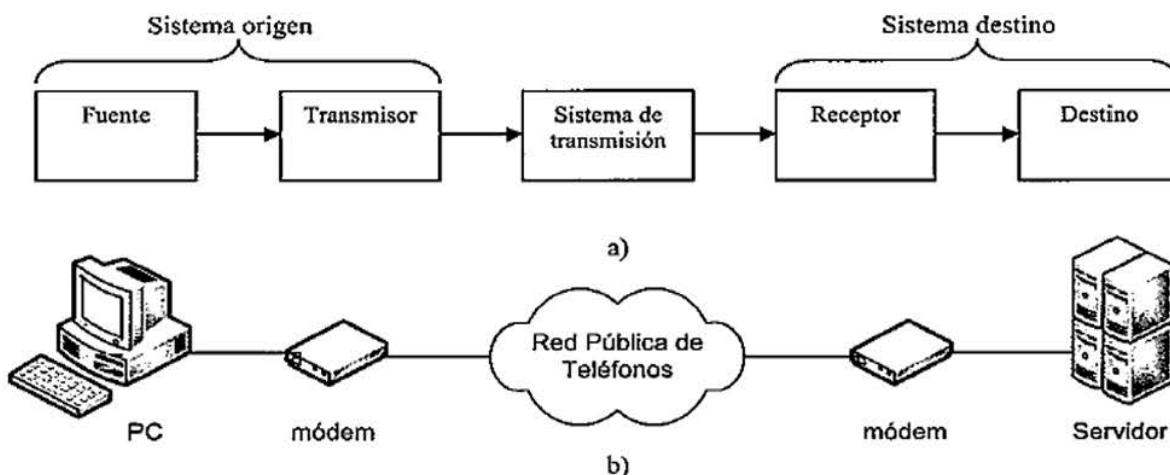


Figura 1-1: Modelo simplificado de comunicaciones. a) Diagrama de bloques general. b) Ejemplo.

Pero para entender mejor las funciones básicas involucradas en un sistema de comunicaciones entre computadores personales es importante descomponer a mayor nivel de detalle el esquema mostrado en la Figura 1.1. Supongamos ahora que queremos describir el proceso de transmisión de datos que se produce entre dos computadores personales mediante el empleo de un software de correo electrónico (Figura 1.2). En dicho proceso, inicialmente, el usuario A activa una aplicación de correo electrónico en el PC y construye el mensaje con el dispositivo de entrada teclado. La cadena de caracteres que forma el mensaje se almacena en memoria del PC como una cadena de bits. Cuando el usuario A pulsa enviar, el PC se conecta a algún medio de transmisión por ejemplo una red local o una línea de telefonía, a través de un dispositivo transmisor como una tarjeta de red o un módem. La cadena de bits en tal caso, se transmite como una secuencia de niveles de tensión $g(t)$ hasta el dispositivo transmisor y a su vez éste la convierte en una señal adecuada al medio físico al que se conecta el dispositivo transmisor. Posteriormente, la señal transmitida es recibida en el dispositivo receptor, que podría ser otro módem u otra tarjeta de red. La señal recibida $r(t)$ puede diferir de la señal transmitida $s(t)$ debido a perturbaciones, perdidas, errores, etc. El receptor estima la señal original $s(t)$ a partir de la recibida $r(t)$, obteniendo una cadena de bits $g'(t)$. Estos bits se envían al PC del usuario B donde se almacenan en memoria y posteriormente se presentarán al usuario a través del dispositivo de salida pantalla en formato de cadena de caracteres.

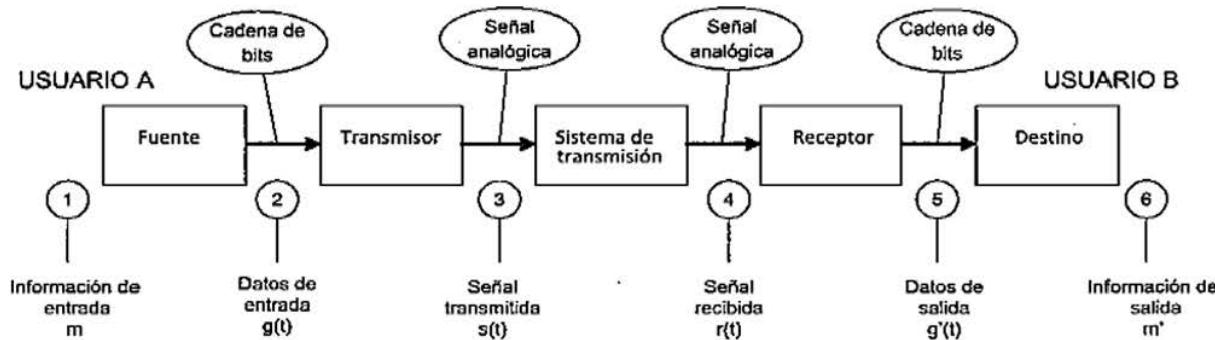


Figura 1-2: Detalle del modelo simplificado de comunicación.

1.1.2. Tareas de un sistema de comunicaciones

Para que sea posible llevar a cabo el proceso de transmisión mostrado en la figura 1.2, es necesario considerar una serie de aspectos fundamentales en las comunicaciones de datos en cualquier red de comunicaciones entre computadores.

El primer aspecto a considerar es la necesidad de *implementar un interfaz* para transmitir la información a través del medio, imprescindible para transformar la información en bits. Posteriormente, un segundo aspecto, es la implementación necesaria para la *generación de señales* adecuadas en intensidad y forma para que se puedan propagar por el medio físico escogido en la red de comunicaciones y puedan ser interpretadas por emisor y receptor adecuadamente. Por lo tanto, debe considerarse la implementación de *mecanismos de sincronismo* que determinen cuando emisor y receptor están preparados para enviar o recibir datos. Además de estas tareas, el sistema de comunicaciones requiere de un conjunto de funcionalidades destinadas a la *gestión del intercambio* de datos para determinar los turnos para transmitir o la cantidad de datos a transmitir. Por otro lado, como se vio en la figura 1.2, es posible que la señal recibida no sea exactamente igual a la señal transmitida por ejemplo porque se hayan producido perturbaciones que hayan provocado errores en la transmisión. En tal caso, otro aspecto a considerar es la implementación de funcionalidades que permitan la *detección y corrección de errores* si fuera necesario. Conviene destacar que es común en muchas ocasiones conectar dispositivos en una red de comunicaciones de distintas velocidades, y que algunos dispositivos emisores son capaces de enviar datos a mayor velocidad que los receptores procesarlos. En estos casos, se producen problemas de congestiónamiento y saturación, por lo tanto otro aspecto a tener en cuenta es la implementación de funciones que permitan *controlar el flujo de intercambio de datos*.

Además, y como no podía ser de otra manera, para que haya comunicación entre dos computadores personales es imprescindible desarrollar *mecanismos de direccionamiento y encaminamiento* para asegurar que el mensaje que se quiere enviar desde un emisor es recibido en receptor adecuado y se ha hecho llegar por el mejor camino posible.

Todos y cada uno de estos aspectos constituyen las funcionalidades básicas o tareas que debe implementar cualquier sistema de comunicaciones entre ordenadores y algunas de ellas serán comentadas a lo largo de la asignatura en los próximos temas. No obstante, no son las únicas tareas que puede realizar un sistema de comunicaciones. En muchos casos, estas tareas se amplian considerando funciones de recuperación por ejemplo cuando la comunicación se ve interrumpida inesperadamente (un fallo durante la descarga de un fichero), funciones de seguridad (la comprobación de direcciones para evitar suplantación de emisor o receptor), funciones de gestión de red (la monitorización del estado de una red, tráfico, etc.), etc.

1.1.3. Objetivos de las redes

La implantación de una red de comunicaciones entre computadores personales en una empresa o negocio tiene como finalidad:

- Compartir recursos como pueden ser impresoras, discos duros o sistemas de archivos.
- Aumentar la tolerancia a fallos: Distribuyendo la información almacenada, o empleando distintos caminos para evitar la pérdida de datos en caso de que algunas máquinas fallen.
- Reducir el coste monetario: Los dispositivos compartidos reducen la inversión económica.

- Potenciar la globalización: Ser capaces de comunicar a cualquier punto del mundo empleando oficinas remotas conectadas en red.
- Acceder a información remota empleando servidores de ficheros, servidores de hipertexto, etc.
- Comunicación de persona a persona o entre grupos de personas haciendo uso de mensajería instantánea, correo electrónico, video llamadas, etc.

1.2. Tipos de redes y topologías

Anteriormente, se ha definido una red como un conjunto de dos o más dispositivos autónomos con la capacidad de interconectarse mediante un enlace de un medio físico. Un enlace no es otra cosa que el medio de comunicación físico que transfiere los datos de un dispositivo a otro. Si se entiende el enlace desde un punto de vista gráfico como una línea que une dos puntos, donde cada punto representa un dispositivo, se pueden clasificar las redes en varios tipos *en función del tipo de conexión al enlace*:

- *Redes de difusión o multipunto*: Cuando más de dos dispositivos comparten el mismo enlace. En este tipo de redes, es posible enviar un mismo mensaje desde un dispositivo al resto (broadcast) o a un conjunto de máquinas (multicast).
- *Redes punto a punto*: Cuando dos dispositivos tienen un enlace directo entre ellos. La conexión de más dispositivos implica el uso de múltiples enlaces punto a punto entre pares de dispositivos. En este caso un mensaje puede tener que visitar multitud de máquinas hasta llegar a su destino.

1.2.1. Clasificación según topología física

Las redes también se pueden clasificar de acuerdo a su *topología física*. La topología física define la representación geométrica de todos los enlaces de una red y los dispositivos físicos que se enlazan entre si. Las topologías más conocidas son: bus, anillo, estrella y malla (Figura 1.3).

- Una topología en *bus* es una configuración donde un único enlace conecta todos los dispositivos de la red constituyendo una red en forma de tronco.
- Una topología en *anillo* es una topología de red donde cada dispositivo tiene una línea de conexión con todos los dispositivos de la red constituyendo una red en forma de anillo.
- Una topología en *estrella* es aquella en la que cada dispositivo sólo tiene un enlace dedicado con un controlador central habitualmente llamado concentrador.
- Una topología en *malla* es una configuración en la que cada dispositivo tiene un enlace punto a punto dedicado con cualquier otro dispositivo. El término dedicado indica que el enlace sólo conduce el flujo de datos entre los dispositivos que interconecta. En la topología en malla, los dispositivos que forman la red pueden ser nodos de reenvío y enrutamiento (router) o equipos finales (PC).

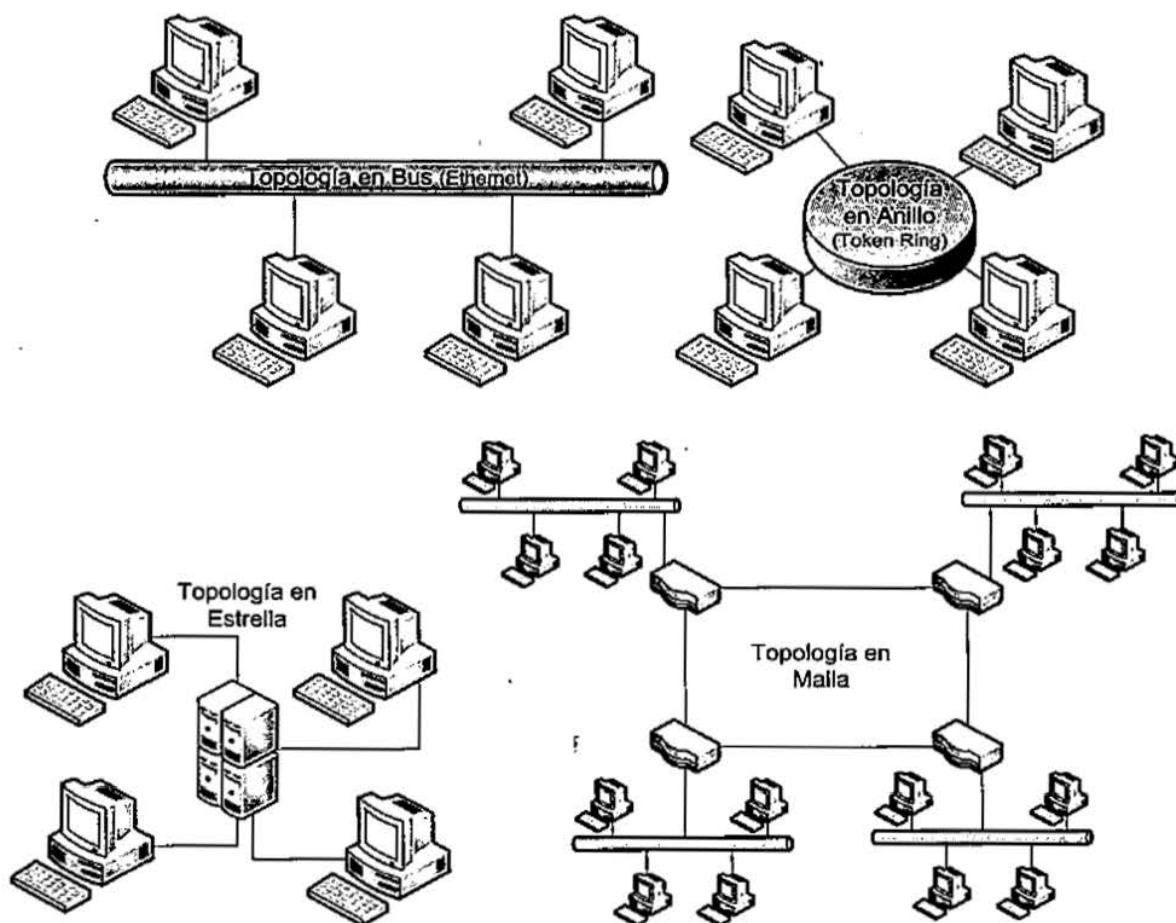


Figura 1-3: Topologías físicas de redes de computadores.

La topología en malla ofrece varias ventajas respecto al resto de topologías de red. Así, la topología en malla proporciona una mayor tolerancia a fallos y fiabilidad. El uso de enlaces dedicados garantiza que cada conexión transporte únicamente los mensajes entre los dispositivos que interconecta, eliminando el problema de que en caso de fallo se pierdan todos los datos que transporta la red, ya que sólo se perderán los que transporta ese enlace. Además, se puede evitar los enlaces que han tenido fallos o problemas de tráfico y, por otro lado, es fácil para el administrador de la red localizar y detectar las causas de la pérdida y dar soluciones a éstas. También cabe destacar que estas redes, por su topología, tienen la ventaja de la privacidad y seguridad ya que únicamente pueden acceder a un mensaje, que es transportado por un enlace determinado, los dispositivos conectados a él. Pero no obstante, la topología en malla también sufre de ciertas desventajas relacionadas principalmente con la cantidad de cable y el número de dispositivos de interconexión (puertos de entrada y salida necesarios). Esto obliga a costes monetarios elevados, a limitaciones de espacio disponible y a dificultades en la instalación y reconfiguración de la red.

Para evitar estos problemas, se hace más adecuado en algunas ocasiones la instalación de redes estrella cuyo coste es más barato que la red en malla. En una topología en estrella, cada dispositivo necesita sólo de un enlace de un puerto de entrada y salida para conectarse a un nodo concentrador, por lo tanto esto hace que se facilite la instalación y reconfiguración de la red, además de reducir el número de cables y conexiones. La red en estrella también se caracteriza por la robustez y tolerancia a fallos, ya que si falla un enlace solamente afecta a los dispositivos conectados a ese enlace, permaneciendo todos los demás enlaces activos. El principal problema radica en caso de que falle el

concentrador, en cuyo caso todos los enlaces de la red fallarán. Además, el número de dispositivos conectados al concentrador determina la capacidad de proceso de éste.

Al igual que en la topología en malla, la topología en estrella requiere gran cantidad de cableado. Otras topologías como el bus o el anillo reducen la cantidad de éste. Por ejemplo, la topología en bus tiene la principal ventaja de la sencillez de instalación. Sin embargo, se caracteriza también por su dificultad para su reconfiguración (añadir o quitar equipos) o localizar y detectar fallos. Además, en topologías en bus, la rotura del cable del bus provoca la imposibilidad de transmisión entre cualquier dispositivo de la red.

Finalmente, cabe comentar de la topología en anillo que es fácil de instalar y de reconfigurar si se compara con una topología en bus mediante cable coaxial. No obstante, sigue siendo más difícil de instalar y reconfigurar si se compara con una topología en bus implementada con dispositivos HUB. También destaca porque los fallos son localizables fácilmente, aunque no tanto como en un bus con HUBs. Siendo su principal desventaja, el tráfico que es del tipo unidireccional siempre en un sentido del anillo.

En la actualidad es frecuente encontrarse con redes de *topología híbrida*, es decir de redes que entremezclan varias topologías convencionales. Así se puede tener varias redes con topología en bus que se unen a un concentrador formando una topología en estrella.

1.2.2. Clasificación según escala geográfica

Por otro lado, en la actualidad y sobre todo desde el punto de vista de negocio, las redes se pueden clasificar en tres categorías dependiendo de la *escala geográfica* de su implantación. Así, en función de su extensión geográfica se disponen de redes: LAN, MAN y WAN.

- Las Redes de *Área Local* (LAN, Local Area Network)
- Las Redes de *Área Metropolitana* (MAN, Metropolitan Area Network).
- Las Redes de *Área Extendida* (WAN, Wide Area Network).

En la Figura 1-4 se muestran algunas de las características que determina el diseño de una red como red LAN, MAN o WAN.

	LAN	MAN	WAN
Espacio Geográfico	Edificio o campus	Ciudad o Región	País o Continente
Propiedad	Entidad Privada	Privada o Pública	Típicamente Pública
Velocidades	[10Mbps, 10Gbps]		
Topologías físicas más usadas	Bus, Anillo, Estrella	Anillo, Malla, Bus	Malla, Punto a punto
Ejemplos	Red Ethernet con impresora compartida y base de datos compartida. WiFi.	Televisión por cable, Redes xDSL entre LANs, WiMay, WiFi-Mesh y ATM.	ATM, Frame Relay, ADSL, GPRS y 3G.

Figura 1-4: Escala geográfica.

Actualmente, es difícil clasificar las redes como LAN, MAN o WAN porque la barrera que determina el tipo de red en función de la escala se hace cada vez más difusa. Esto es así porque es muy raro encontrarse en el mundo empresarial o de negocio redes LAN, MAN o WAN aisladas (Figura 1.5). Generalmente, las redes LAN se interconectan entre sí mediante MAN o WAN formando lo que se conoce como interred, también denominado en el mundo anglosajón la internet.

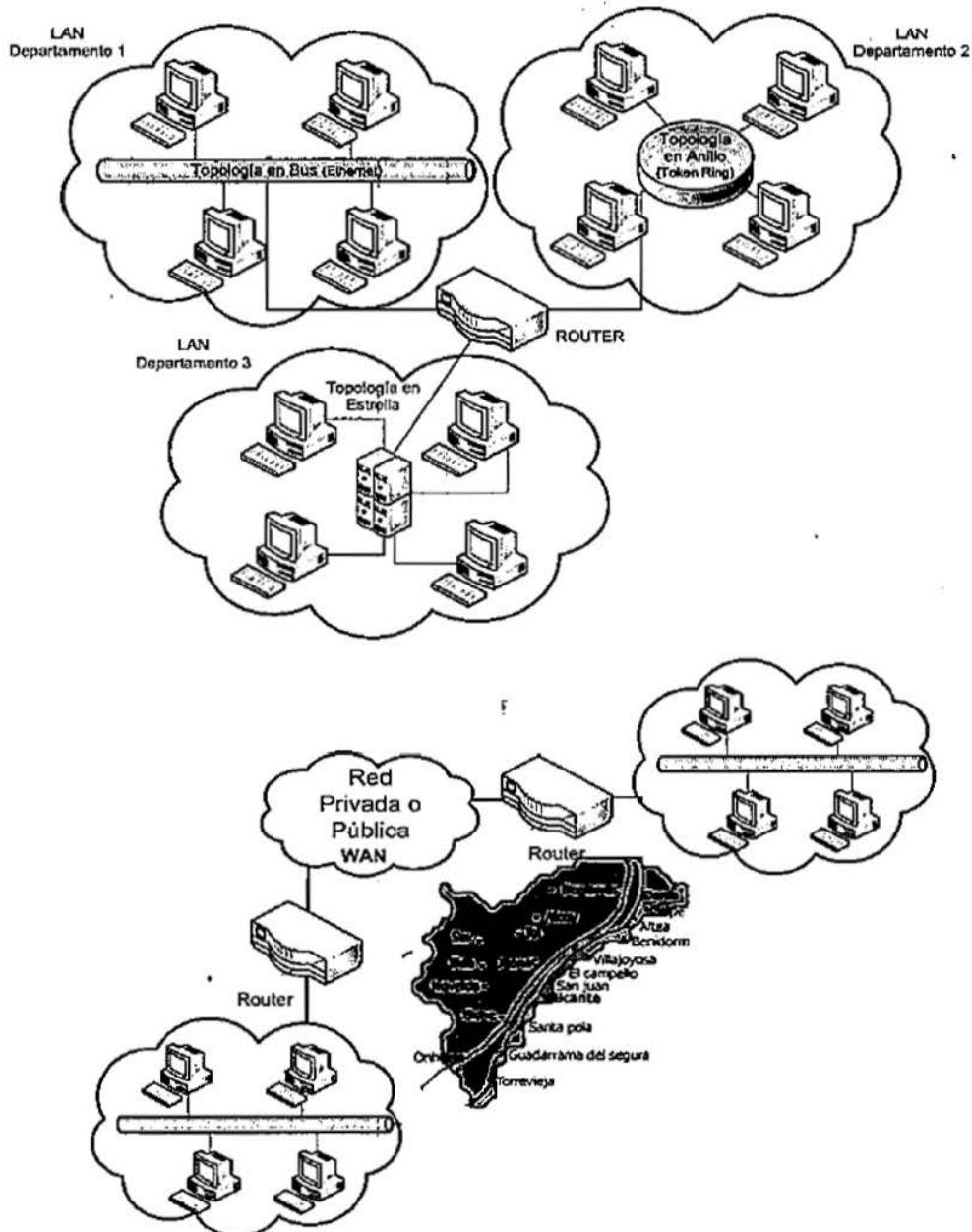


Figura 1-5: Escala geográfica.

En la figura 1.4 se ha comentado la propiedad de la redes según la escala geográfica. Se determina la propiedad de una red en función del ámbito de los datos y del organismo propietario que gestiona y administra los dispositivos y topologías que la constituyen y/o los datos que se transportan. De este modo, las redes se pueden catalogar como redes de ámbito privado o público.

- **Red privada.** El administrador y propietario de la red es una entidad privada y la red se emplea para fines propios.
- **Red pública.** Una empresa u organización propietaria de la red alquila parte de su capacidad a múltiples usuarios o bien vende partes de la red a operadores que a su vez ofrecen servicios públicos. Las redes públicas suelen estar reguladas por un organismo

público, como es la comisión de Telecomunicaciones del estado (CTT). Ejemplo de ello es la red de telefonía RTC, que operan Telefónica, Orange, Ya.com, etc.

Recientemente, ha aparecido el concepto de Redes de *Área Personal* (PAM, Personal Area Network). Las redes PAM hacen referencia a redes de comunicaciones entre dispositivos móviles muy cercanos. Un claro ejemplo de este tipo de redes se produce cuando un dispositivo telefónico bluetooth crea una conexión con otro dispositivo con bluetooth.

1.2.3. Clasificación según la forma de establecer la comunicación

Frecuentemente, las redes de área extendida, WAN, proporcionan medios de transmisión a largas distancias y se extienden geográficamente por muchos centenares o miles de kilómetros. Éstas se pueden clasificar, en función de la forma en que se establece la comunicación, en redes de dos tipos: redes de conmutación de circuitos y redes de conmutación de paquetes. Habitualmente, las redes WAN conmutadas conectan LAN entre sí mediante dispositivos enruteadores. Ejemplos de tecnologías de redes WAN conmutadas son X.25, la red de alta velocidad de retransmisión de tramas Frame Relay o ATM (Asynchronous Transfer Mode).

- Red *conmutación de circuitos*. La conmutación se realiza empleando un camino físico, fijo, único y dedicado para una misma comunicación de datos. Por ejemplo, este es el caso de la telefonía clásica analógica.
- Red *conmutación de paquetes*. La información a transmitir se divide en fragmentos o unidades de información, denominados paquetes, segmentos o tramas. Existen dos tipos de conmutación de paquetes: circuitos virtuales y datagramas. En el primero de ellos, se establece un camino virtual y los paquetes siguen una misma ruta hasta que se libera dicho camino. En el segundo, no se establece camino virtual previo y cada paquete puede seguir un camino distinto. En este caso, por lo tanto los paquetes pueden llegar desordenados.

Estos dos tipos de redes en función de la manera de establecer la comunicación se tratarán en mayor profundidad en el capítulo 8.

1.3. Arquitectura de red

Una *arquitectura de red* se puede definir como el conjunto de capas y protocolos que constituyen un sistema de comunicaciones. Cada capa o nivel es un consumidor de servicios ofrecidos por el nivel inferior y proveedor de servicios del nivel superior. Además cada capa se implementa mediante un conjunto de entidades. Se entiende como *entidades*, aquellos elementos de un nivel que dialogan con otros elementos del mismo nivel y se entiende como *servicio* un conjunto de funciones.

La comunicación entre entidades de una misma capa, en distintos dispositivos, es gobernada por un conjunto de reglas denominadas *protocolos*. Sin embargo, si la comunicación se produce entre entidades de capas distintas de un mismo dispositivo, al conjunto de reglas que gobiernan dicho intercambio de información se le denomina *interfaz* (Figura 1.6).

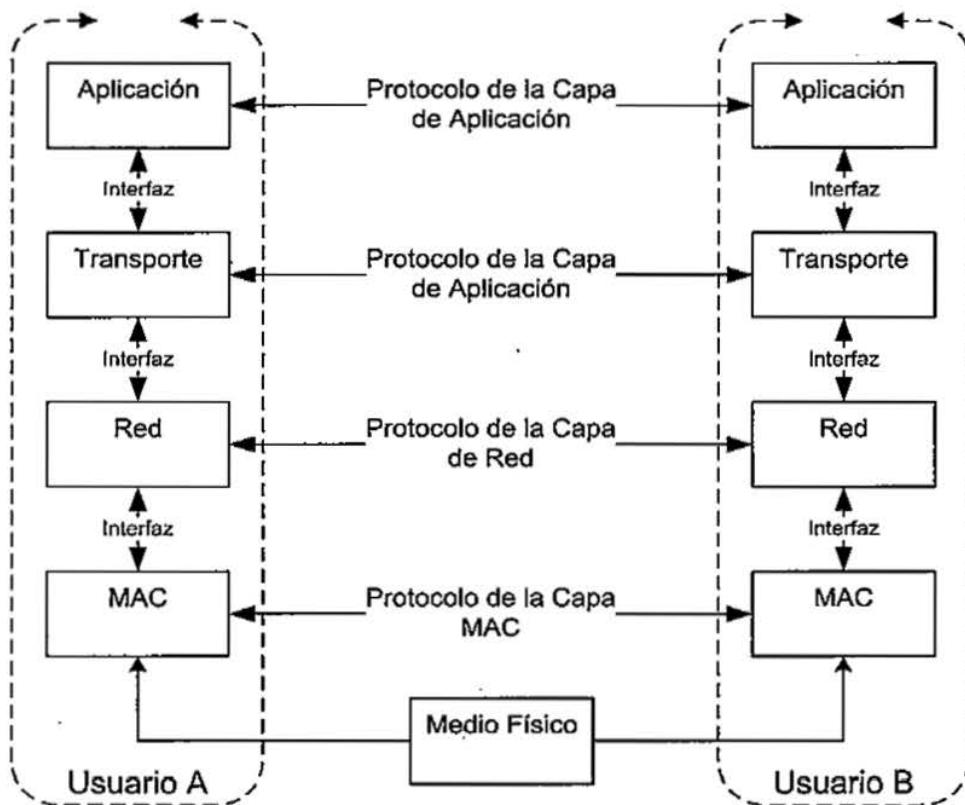


Figura 1-6: Modelo de comunicación horizontal y vertical entre dos dispositivos.

Una arquitectura estratificada en capas permite definir funciones específicas en un sistema grande y complejo. Cuando un sistema tiene una estructura estratificada es mucho más fácil realizar modificaciones en una parte de éste sin que estas modificaciones obliguen a modificar todo el sistema. Así es más fácil añadir funcionalidades, eliminarlas o actualizarlas con nuevas modificaciones. Un ejemplo de arquitectura diseñada e implementada en capas es la arquitectura que permite el manejo de un PC. Así un PC, se puede definir como una arquitectura estratificada en niveles, dónde se dispone de un hardware, un firmware para acceder a los dispositivos software vía un software de bajo nivel específico, un software de gestión de procesos como es el sistema operativo que se encarga de gestionar la interacción entre dispositivos, memoria y procesos. Y finalmente un software de servicios, formado por las aplicaciones de usuario (Figura 1.7).

Existen dos arquitecturas de red que han sido fundamentales en el desarrollo de estándares de comunicaciones. Se trata de los modelos de referencia TCP/IP y OSI/ISO. El modelo de referencia OSI (Open System Interconexión) es una normativa internacional de la ISO (Internacional Standards Organization). El modelo de referencia TCP/IP constituye, hoy por hoy, la arquitectura de red más empleada en cualquier sistema de comunicaciones que requiera interconexión entre sistemas diversos. En ambos casos, ambas arquitecturas de red se organizan en un conjunto de capas. Así el modelo de referencia OSI/ISO se compone de siete capas y el modelo de referencia TCP/IP en cuatro: Enlace o Host-Red, Red, Transporte y Aplicación. En la capa de enlace algunos autores diferencian 2 subcapas, Física y Acceso a red, hasta constituir las 5 capas mostradas en la Figura 1.7. De ahí que muchos autores llamen a la capa de enlace capa Host-Red (Física y Acceso a Red).

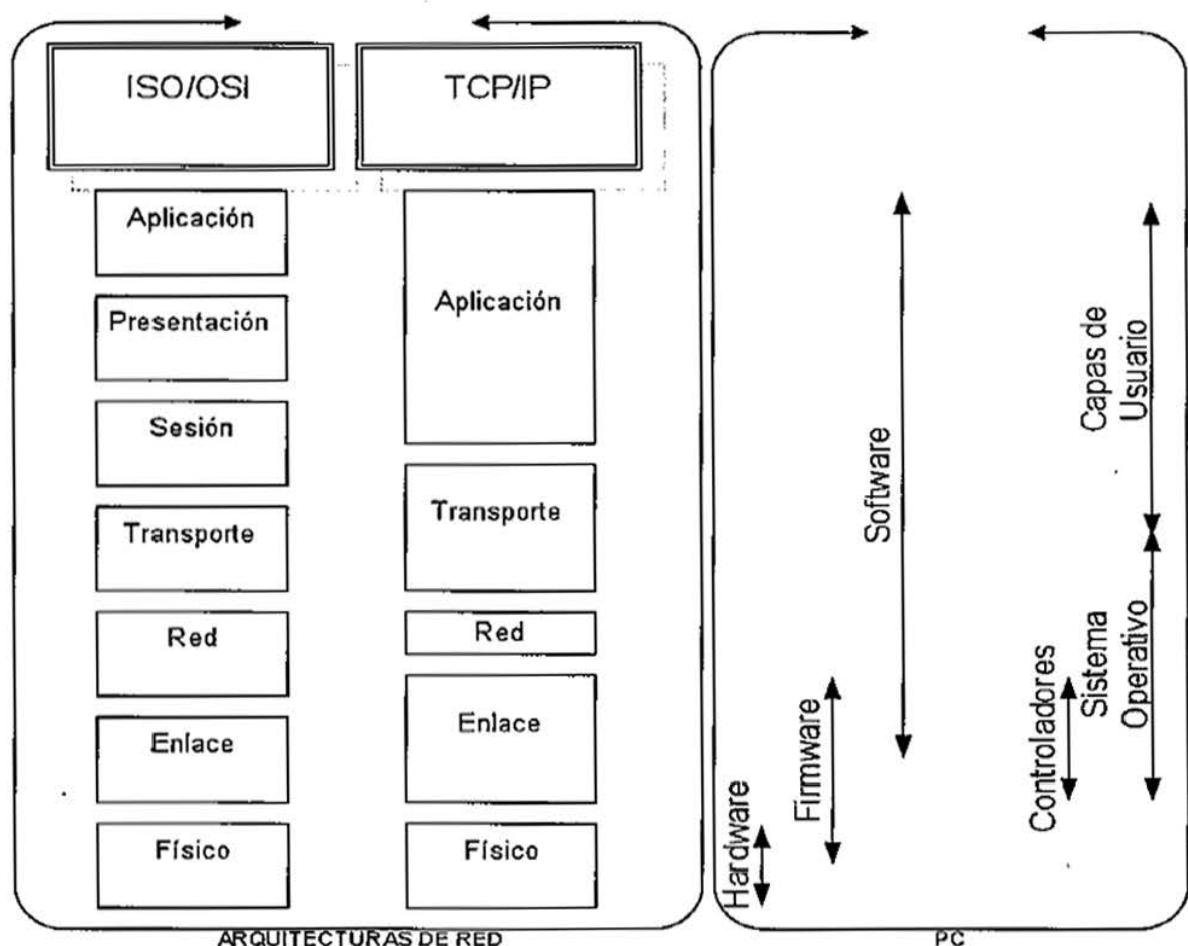


Figura 1-7: Comparación sobre distintas arquitecturas estratificadas en capas.

En una arquitectura de red estratificada en capas, existe un conjunto de protocolos por cada capa, por ejemplo el modelo de referencia TCP/IP soporta los siguientes protocolos:

- Nivel Físico y Acceso a red: Ethernet, Token Ring, Inalámbrica 802.11g, etc.
- Nivel de Red: IP y protocolos de soporte ARP, RARP, ICMP, IGMP, etc.
- Nivel de Transporte: TCP, UDP, etc.
- Nivel de Aplicación: HTTP, FTP, DNS, SNMP, Telnet, etc.

Un protocolo, en función del nivel que lo implementa, se puede encargar de tareas tan diversas como direccionamiento (IP, Ethernet, etc.), control de errores (Ethernet, TCP, etc.), multiplexación y demultiplexación (RDSI, etc.), codificación (Ethernet, Token Ring, etc.), encriptación (SSL, TSL, etc.), etc. Pero, ¿cómo se produce el flujo de información y la comunicación entre las distintas capas de una arquitectura de red de un dispositivo? y ¿cómo se produce el intercambio de información entre arquitecturas de dos dispositivos distintos (emisor y receptor)? (Figura 1.6). Para dar respuesta a la primera de las preguntas conviene introducir varios conceptos: La unidad de datos de protocolo o *PDU*, la unidad de datos de interfaz o *IDU*, la información de control de interfaz o *ICI*, la unidad de datos de servicio o *SDU* y el punto de acceso a los servicios o *SAP* que es un identificador único (Figura 1.8)

Supóngase ahora que se produce flujo de información para una comunicación vertical entre dos capas adyacentes de una arquitectura de red. Se entiende por capas adyacentes, capas contiguas de un mismo dispositivo. Por ejemplo, un usuario emisor a través de software de aplicación quiere redactar un correo electrónico para enviarlo a otro usuario de la red (receptor). Si se sabe que el número de capas de la arquitectura es un valor genérico n , se requiere que cada capa de nivel superior acceda a los servicios de la capa de nivel inferior para transmitir la información entre capas a través del SAP. Así, la capa de nivel superior $n+1$ necesita acceder a la capa de nivel inferior n para enviar un paquete de datos IDU (IDU porque la comunicación es entre entidades de capas distintas). La IDU que se recibe en el nivel n , se compone de dos unidades de datos, la SDU procedente del nivel $n+1$ y la ICI que ha añadido el mismo nivel $n+1$ antes de enviarla al nivel n (Figura 1.8).

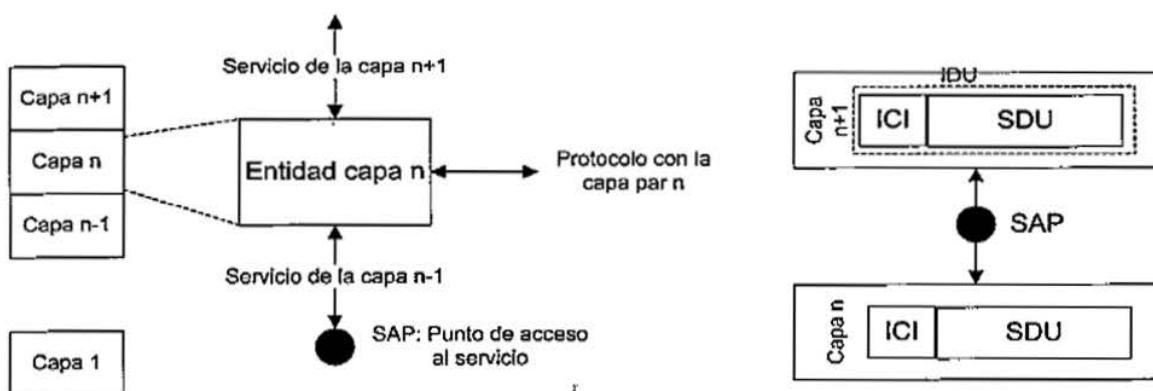


Figura 1-8: Elementos que intervienen en la transferencia de información entre dos capas.

Cuando la capa n recibe el IDU de la capa $n+1$, ésta elimina la cabecera ICI de la IDU quedándose únicamente con la SDU que encapsula. A esa SDU procedente de la capa $n+1$ le añade una cabecera PCI propia de la capa n . Este nuevo bloque de datos que se ha formado en la capa n constituye la PDU de la capa n . Posteriormente, si la capa n quiere transmitir los datos a la capa $n-1$, añadirá a esta PDU de la capa n , la ICI de la capa n . Por lo tanto la PDU de la capa n no es más que la SDU de la capa n que se enviará a la capa $n-1$. Y así, el proceso se irá repitiendo hasta alcanzar la capa de más bajo nivel (Figura 1.9).

En la práctica, desde un punto de vista de implementación software, la ICI viene a ser el conjunto de parámetros que se pasa en las llamadas a la función que implementa el servicio de nivel inferior, junto con los datos a enviar o recibir, SDU. La PCI está constituida por los campos de cabecera y colas añadidos dentro de las tramas y paquetes enviados a través de la red por el protocolo de una determinada capa.

Para dar respuesta a la segunda de las preguntas formuladas conviene observar el esquema de la figura 1.10, donde se muestra el encapsulamiento de información entre capas de un mismo dispositivo y el sentido del flujo de datos entre distintas capas y entre distintos dispositivos. El proceso que se sigue para realizar el intercambio de información entre entidades de distintas capas es el mostrado en la figura 1.9 que ya ha sido comentado anteriormente.

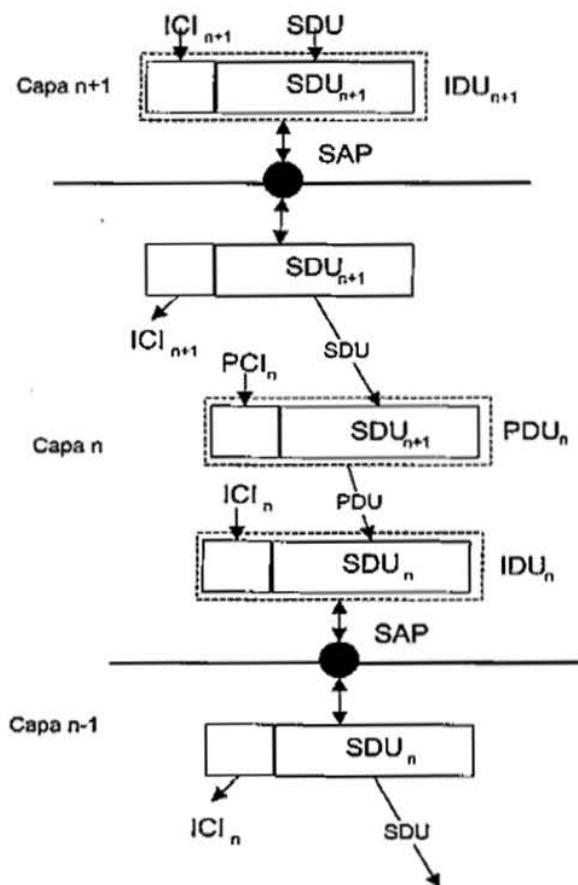


Figura 1-9: Transferencia de información entre capas.

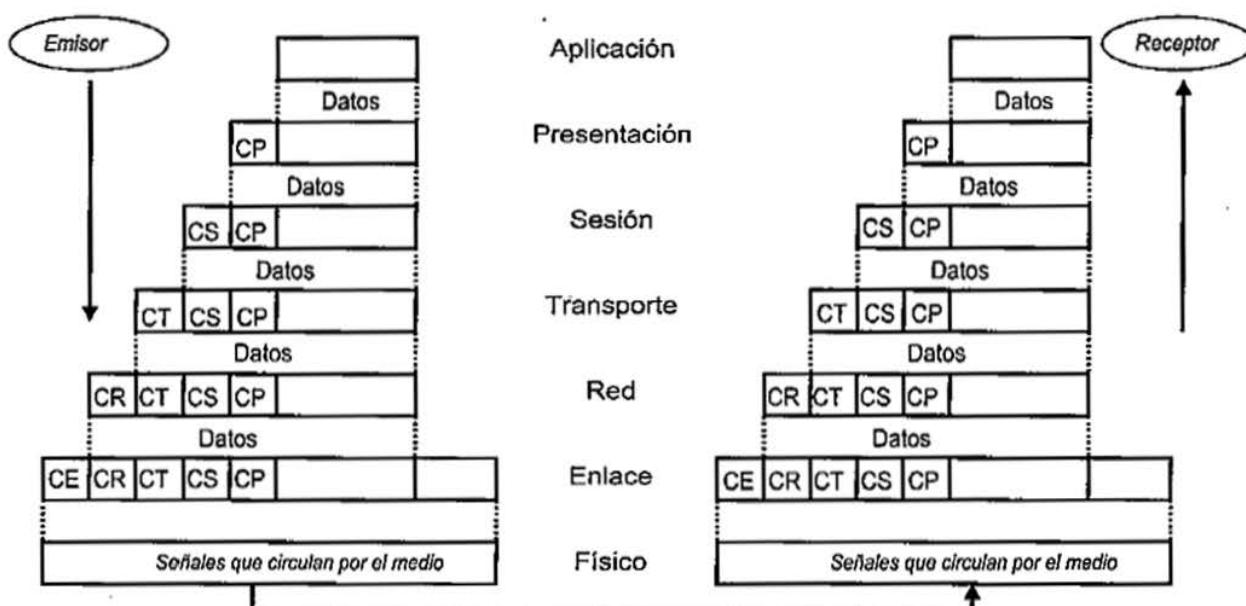


Figura 1-10: Comparación sobre distintas arquitecturas estratificadas en capas.

1.3.1. Modelo de referencia OSI/ISO

La ventaja de los modelos de referencia para las arquitecturas de red radica en que suponen un principio de normalización necesario para permitir la interoperabilidad entre equipos que se quieren comunicar. El inconveniente es que cualquier proceso de estandarización tiende a congelar la tecnología y reduce la implantación de nuevos desarrollos futuros.

Como se ha comentado anteriormente, el modelo de referencia OSI consta de siete capas. De la capa más alta a la más baja estas son: Aplicación, Presentación, Sesión, Transporte, Red, Enlace y Física.

- Nivel *Físico*: Se encarga de la transmisión de cadenas de bits de datos a lo largo de un canal de comunicación soportado por un medio físico. En él se definen las especificaciones mecánicas y eléctricas según el medio de transmisión para que sea posible la comunicación. Es en este nivel donde se establece el modo y tecnología de transmisión, se define la topología física empleada, mecanismos de sincronización entre emisor y receptor, la representación de los bits mediante técnicas de codificación, técnicas de modulación, etc. (Ver Capítulos del 2 al 4).
- Nivel de *Enlace*: Se encarga de conseguir una transmisión entre estaciones de un mismo enlace sin que haya errores. Hace que el nivel físico aparezca ante el nivel de red como un medio libre de errores. Este nivel es el responsable del movimiento de datos entre dispositivos. Así, es función de este nivel organizar los datos en bloques de datos denominados tramas, controlar el flujo de información, el direccionamiento físico, control de acceso al medio y control de errores. (Ver Capítulos del 5 al 7).
- Nivel de *Red*: Se encarga de manejar los bits de datos en bloques denominados paquetes. Además tiene como función gestionar el encaminamiento de dichos paquetes a través de la red. Mientras que el nivel de enlace supervisa la entrega de bloques de datos (en este nivel llamados tramas) entre dos dispositivos unidos por un enlace, el nivel de red asegura la entrega de bloques de datos (en este nivel llamados paquetes) entre dos dispositivos cualesquiera unidos por enlaces distintos. (Ver Capítulos 8 y 9).
- Nivel de *Transporte*: Es el responsable de la entrega de un mensaje completo (el mensaje está compuesto por todos los bits de datos organizados en paquetes) entre dos procesos (extremo a extremo), uno en el emisor y otro en el receptor (origen y destino). Es por lo tanto, en el nivel de transporte del emisor donde un mensaje se divide en un conjunto de bloques de datos llamados segmentos y se reensambla posteriormente en el nivel de transporte del receptor hasta volver a constituir el mensaje. También, es en este nivel donde se establecen controles de conexión, controles de flujo extremo a extremo y controles de errores extremo a extremo.
- Nivel de *Sesión*: Las funcionalidades y servicios implementados en los cuatro primeros niveles no son suficientes para algunos procesos, por este motivo el nivel de sesión tiene como finalidad establecer, mantener y sincronizar la interacción entre sistemas de comunicación. Gracias a este nivel, es posible establecer diversas sesiones entre máquinas o insertar puntos de verificación en el flujo de datos, conceptos ampliamente empleados en procesos de descarga de ficheros peer to peer.
- Nivel de *Presentación*: Es la capa que se ocupa de los aspectos de sintaxis y semántica de la información y datos que se transmiten. En este nivel se implementan funciones de criptografía para asegurar la privacidad de los datos transmitidos, funciones de compresión para reducir el número de bits a transmitir y funciones de traducción para adecuar el formato de representación de la información de modo que sea inteligible por emisor y receptor.

- Nivel de *Aplicación*: Es el nivel más alto de la arquitectura, y es el nivel en que interactúan los usuarios (emisor-receptor). Es en este nivel donde se ubica el software, procesos, clientes o servidores (transferencia de ficheros, correo electrónico, navegadores web, etc.) que se utilizan para enviar los mensajes.

La elección de las siete capas se basó en una serie de criterios entre los que destacan:

- Cada capa debe tener una función de comunicación diferente a la de las otras.
- El número de capas debe ser suficientemente elevado para que éstas sean lo más independientes posible y lo más pequeño posible para que las capas sean de fácil manejo.
- Cada capa debe realizar una función bien definida.
- La función que realiza cada capa debe seleccionarse de modo que minimice el flujo de información entre los interfaces.

1.3.2. Modelo de referencia TCP/IP

El modelo de referencia TCP/IP define una arquitectura de comunicaciones estructurada en cuatro niveles o capas.

- Nivel de *Host-Red*: Este nivel se subdivide en dos capas, nivel de *interfaz de acceso a la red* y *nivel físico*. El nivel de interfaz de acceso a la red constituye el nivel software más bajo de la arquitectura TCP/IP. Es esta capa la responsable de añadir a los datos, información de control para transmitirlos a través de una red específica (bus, anillo, etc.). La segunda de las capas, el *nivel físico*, define las características físicas y de hardware, es decir tipo de conectores, número de pines de cada conector, especificaciones eléctricas para las señales que se intercambian a través de los pines de los conectores, etc.
- Nivel de *Red*: Es el nivel de interconexión y ésta es la capa encargada de encaminar los datos que forman los mensajes de una máquina a otra, a lo largo de todas las conexiones que hacen posible la comunicación entre emisor y receptor.
- Nivel de *Transporte*: Este nivel intermedio que proporciona mecanismos para regular adecuadamente el intercambio de mensajes entre procesos del dispositivo emisor y procesos del dispositivo receptor, asegurando que los datos que constituyen dichos mensajes se entregan libres de errores, en orden y sin pérdidas ni duplicaciones.
- Nivel de *Aplicación*: Es el nivel más alto de la arquitectura, y es el nivel en que interactúan los usuarios (emisor-receptor). Es en este nivel donde se ubica el software, procesos, clientes o servidores (transferencia de ficheros, correo electrónico, navegadores web, etc.) que se utilizan para enviar los mensajes.

En el modelo inicial de referencia TCP/IP no disponía funcionalidades de encriptación de datos y/o autentificación de identidades. Posteriormente, este tipo de funcionalidades se han incluido en la arquitectura como capas o niveles auxiliares intermedios mediante la incorporación de nuevos protocolos. Por ejemplo el protocolo IPSec añade un nivel de seguridad entre el de Red y el nivel de Transporte, y el protocolo SSL añade un nivel de seguridad entre el nivel de Transporte y el de Aplicación.

Capítulo 2. TRANSMISIÓN DE SEÑALES

En este capítulo se describen distintos aspectos relativos a las señales que se envían, habitualmente, por un sistema de comunicación. El capítulo que se enmarca dentro del nivel físico del modelo OSI, muestra un estudio de distintas propiedades de los medios de transmisión y que afectan directamente a las prestaciones de las redes de computadores tales como la velocidad máxima de transmisión, tipo de información a enviar, etc. Se describirán características de los medios de transmisión como ancho de banda o como las perturbaciones que limitarán y condicionarán el resto de niveles superiores del modelo OSI.

CAPÍTULO 2

2-1 Introducción. Funciones de la capa física

2-2 Fundamentos de la transmisión de datos

- Tipos de señales y parámetros de una señal
- Desarrollo en serie de Fourier. Espectro de potencias
- Ancho de banda. Ganancia.
- Prestaciones del medio físico

2-3 Perturbaciones en la transmisión

2-4 Velocidades máximas de envío de datos

- Teorema de muestreo. Velocidad máxima
- Velocidad máxima según la relación señal/ruido

2-5 Filtrado de señales

2-6 Transmisión analógica y digital. Definición y conceptos básicos

2-7 Modos de transmisión

OBJETIVOS

- Comentar las funciones básicas de la capa física de una arquitectura de red
- Mostrar los elementos básicos, desde un punto de vista la capa física, intervienen en una comunicación
- Clasificar las señales atendiendo a su naturaleza
- Mostrar el modelado de señales que representan datos digitales mediante desarrollo en serie de Fourier
- Exponer la definición de conceptos básicos como ancho de banda, ruido, ganancia, filtrado o velocidad.
- Definir los parámetros básicos que determinan las prestaciones de un medio físico
- Explicar los modos de transmisión en un medio físico.

TERMINOLOGÍA – PALABRAS CLAVE

- | | | |
|------------------------|----------------------------|-------------------------|
| • DTE/DCE | • Velocidad de modulación | • Transmisión sincrona |
| • Medio físico | • Velocidad de transmisión | • Transmisión asíncrona |
| • Canal de datos | • Estado/Nivel | • Señal de reloj |
| • Circuito de datos | • Atenuación | • Línea de datos |
| • Parámetros de señal | • Ruido/Distorsión | • Filtrado |
| • Serie de Fourier | • Muestreo | • Teorema de Shannon |
| • Espectro de potencia | • Ancho de Banda | • Teorema de Nyquist |
| • Armónico | • Ganancia | |

2.1. Introducción. Funciones de la capa física

Un aspecto fundamental y del cual depende directamente la calidad de la red es el medio de transmisión empleado. El medio de comunicación limita y condiciona propiedades básicas del sistema de transmisión tales como la velocidad máxima, tipo de señales que pueden transmitirse, condiciones de funcionamiento etc. Es por tanto, necesario realizar un estudio no solo de las propiedades del medio de transmisión sino también del tipo de señal que se está introduciendo por el medio. Así, se considerará como señal a la entidad física que se transmite a través de un canal para unir dos elementos independientes. Dicha señal ha de transportar la información en forma de bits generada por las capas de nivel superior de la arquitectura de red utilizada. A lo largo de este tema se van a describir las funciones del nivel físico definido por el OSI. Este nivel físico, que también es empleado por otras arquitecturas como TCP/IP, se encarga de realizar la transmisión de una serie de bits por el canal de comunicación en forma de señales físicas.

El nivel físico recibe una trama binaria que se debe convertir a una señal eléctrica, electromagnética u otra dependiendo del medio de transmisión, de tal forma que, a pesar de la degradación que pueda sufrir en el medio, vuelva a ser interpretable correctamente en el receptor. El nivel físico será, por lo tanto, la capa encargada de coordinar las acciones necesarias para la transmisión de un flujo de bits a través de un medio físico, de manera que dentro de las especificaciones propias al nivel físico cabría citar algunas como:

- Definir el tipo de medio. Por ejemplo, si se va a transmitir por fibra óptica, par trenzado o si se trata de una comunicación inalámbrica.
- Velocidad de transmisión. Es decir, la cantidad de información que es capaz de transmitir un sistema de comunicación por unidad de tiempo
- Configuración de la línea (punto a punto o multipunto).
- Señalización / Modulación. Se encargará de adecuar las señales al medio físico por el que se desea realizar la transmisión.
- Sincronización emisor-receptor. Define los protocolos para el envío de datos por parte del emisor y confirmación por parte del receptor.
- Topología. Estructura física de la red empleada.
- Interconexiones mecánicas. Dispositivos que aparecen en cualquier sistema de comunicación tales como conectores, concentradores así como sus características eléctricas y mecánicas.

En la Figura 2.1 se muestran los elementos básicos implicados en una comunicación. En esta figura se representa un esquema en el que dos ordenadores desean comunicarse entre sí empleando la RTC (Red Telefónica Conmutada). Los elementos básicos que conforman una comunicación son los siguientes:

- DTE (Equipo Terminal de Datos). Se trata del dispositivo que quiere enviar y recibir información a través de la red. En el caso de la Figura 2-1 los dos ordenadores que desean comunicarse.
- DCE (Equipo Comunicador de Datos). Este es el dispositivo que adapta o transforma la información generada por el DTE para hacerla adecuada a la soportada por el medio físico (p. ej. modem, tarjeta de red, etc.).
- Medio Físico. Lugar por donde se encamina la señal para llegar a su destino (p. ej. cable, espectro de radio, etc.).

- Canal de datos. Define una transmisión unidireccional en el medio físico.
- Circuito de datos. Define la comunicación entre los dos DCE (puede comportar más de un canal de datos). A lo largo de este tema se estudiarán los circuitos de datos.
- Enlace de datos. Define la comunicación entre los dos DTE.

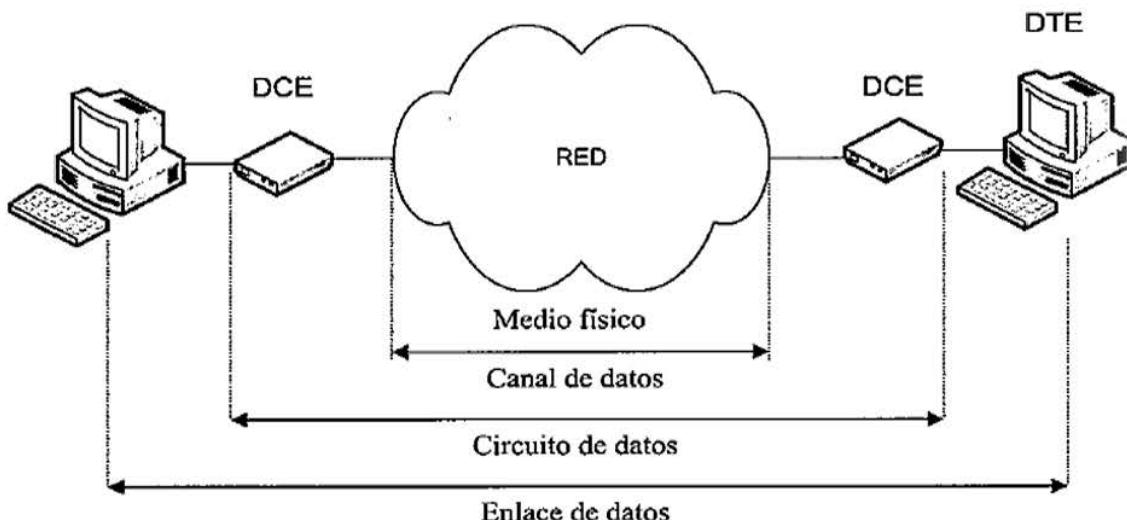


Figura 2-1. Modelo simplificado de comunicaciones.

2.2. Fundamentos de la transmisión de datos

2.2.1. Tipos de señales y parámetros de una señal

Como se ha indicado anteriormente, se considera como señal a la variación en el tiempo de una magnitud física (p. ej. señal electromagnética, óptica, etc.). Es posible clasificar las señales atendiendo a distintos criterios como los mostrados en la Figura 2-2 atendiendo a su naturaleza:

- Analógica. Señal continua que varía suavemente en el medio en tiempo y en amplitud.
- Digital. Señal discreta que sólo puede tener un número definido de valores en instantes de tiempo determinados.
- Periódica: formada por un patrón que se repite continuamente.
- Aperiódica: no posee patrón repetitivo.
- Simples: basadas en funciones seno o coseno.
- Compuestas: basadas en una composición de señales simples.

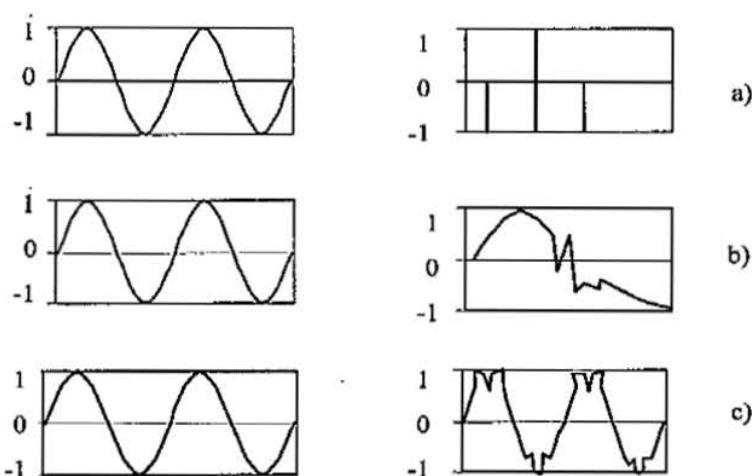


Figura 2-2. a) Señal Analógica – Digital. b) Señal periódica-aperiódica. c) Señal simple-compuesta.

Además las señales pueden clasificarse atendiendo a la magnitud física de la señal en: señales eléctricas, que son aquellas en las que su magnitud hace referencia a voltaje o amperaje, señales ópticas, en las que su magnitud es intensidad luminosa, y señales electromagnéticas, en las que varía con el tiempo la fase, amplitud y/o la frecuencia.

Como se ha descrito anteriormente, las señales periódicas presentan un determinado patrón que se repite cada cierto tiempo. Una típica señal periódica es la onda senoidal:

$$s(t) = A \operatorname{sen}\left(2\pi \frac{1}{T} t + \phi\right) \quad (2.1)$$

La onda senoidal anterior y, en general, cualquier señal periódica, $s(t)$, puede ser caracterizada por los siguientes parámetros (véase Figura 2-3):

- **Amplitud, A:** Valor de la señal en un instante t .
- **Periodo, T:** Tiempo en segundos que tarda la señal en completar un ciclo, de forma que se cumple que: $s(t+T)=s(t)$.
- **Frecuencia:** Inversa del periodo $f=1/T$. La frecuencia se mide en Hertzios (Hz) o Ciclos por segundo.

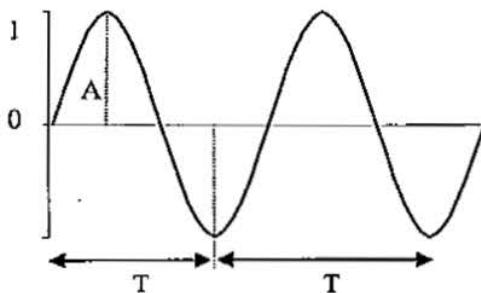


Figura 2-3. Parámetros que definen una señal.

Adicionalmente, otro parámetro que define una señal periódica es la fase, ϕ . La fase es una medida de la posición relativa del inicio de la señal dentro de un periodo de la misma. En la Figura 2-4 se observa el efecto de variar cada uno de estos parámetros en una señal senoidal. En la Figura 2-4.a se ha representado una señal con periodo $T=50$ mseg., amplitud $A=0.5$ y fase=0; en la Figura 2-4.b se aumenta la amplitud $A=1$; en la Figura 2-4.c el periodo se ha aumentado a 100 mseg. y, finalmente, en la Figura 2-4.d la fase vale $\pi/2$ con respecto a la señal de la Figura 2-4.c.

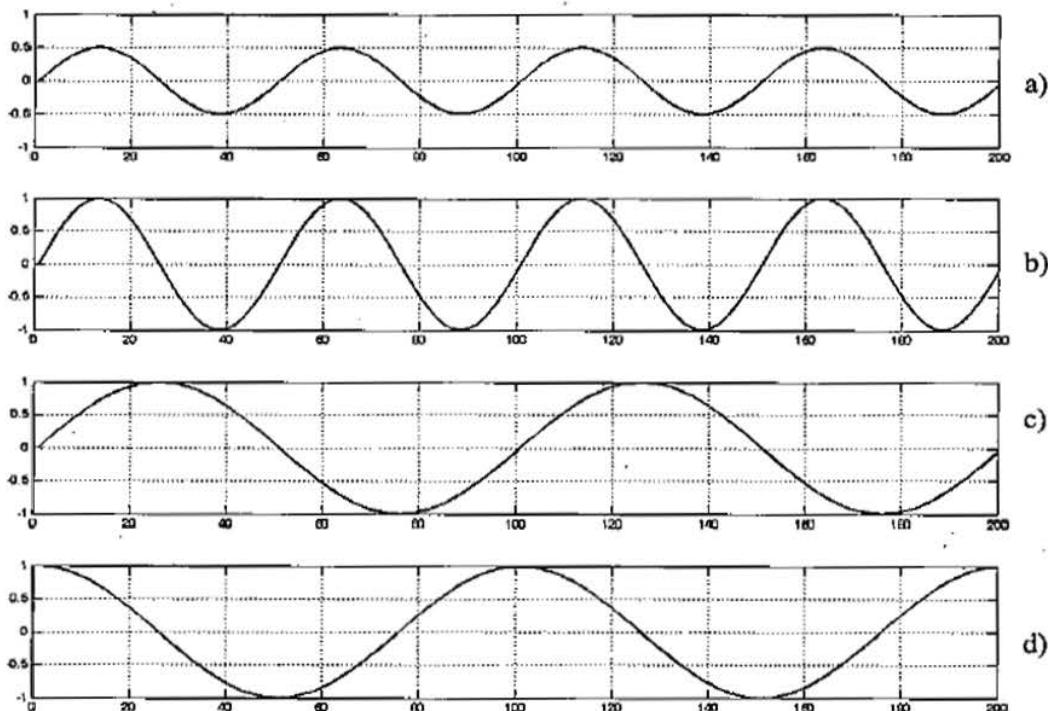


Figura 2-4. Efecto de variar los parámetros de una onda senoidal.

2.2.2. Desarrollo en serie de Fourier

Cualquier señal, por compleja que sea puede ser descompuesta en partes más sencillas. En concreto, las ondas electromagnéticas típicas empleadas para la transmisión de información en redes de computadores pueden ser descompuestas en señales de distinta frecuencia. Esto se puede conseguir aplicando el desarrollo en serie de Fourier. De esta manera, como se verá posteriormente, será más sencillo realizar la descomposición de una señal compleja en muchas señales senoidales cada una de ellas con una frecuencia dada, y estudiar la distorsión en el medio de las mismas, que estudiar la distorsión directamente de la señal compuesta.

Dada una señal compuesta periódica cualquiera, el matemático *Fourier* demostró que es posible descomponerla en una serie, en general infinita, de funciones simples (senos y cosenos) de diferente amplitud y frecuencia.

$$s(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos(2\pi f_0 t) + \sum_{n=1}^{\infty} b_n \sin(2\pi f_0 t) \quad (2.2)$$

donde:

$$a_n = \frac{2}{T} \int_0^T s(t) \cdot \cos(2\pi n f_0 t) dt \quad n=0 \dots \infty \quad (2.3)$$

$$b_n = \frac{2}{T} \int_0^T s(t) \cdot \sin(2\pi n f_0 t) dt \quad n=1 \dots \infty \quad (2.4)$$

Los valores a_n y b_n se denominan amplitud del seno y coseno del n -ésimo armónico respectivamente. De manera que al par $(a_n \cos(2\pi n f_0 t), b_n \sin(2\pi n f_0 t))$ se le denomina armónico de orden n . Además, la frecuencia $f_0 = 1/T$ será la frecuencia fundamental. Cada armónico en el que se descompone la señal original presenta una frecuencia que será múltiplo de la frecuencia fundamental.

2.2.3. Espectro de potencia

Como se observa de la ecuación correspondiente al desarrollo en serie de Fourier para obtener la descomposición de una señal compleja $s(t)$ es necesario sumar infinitos armónicos. Sin embargo, los primeros armónicos (los de menor orden) tienen mayor importancia ya que las ondas senoidales correspondientes tienen mayor amplitud. Esto último implica que no es necesario disponer de los infinitos armónicos para poder posteriormente reconstruir una señal, sino que bastará con los n primeros armónicos, ya que la mayor parte de la información se encuentra almacenada en ellos. Esto último se puede estudiar analizando el denominado espectro de potencias que mide la contribución de cada término armónico a la reconstrucción de la señal original:

$$P_n = \sqrt{a_n^2 + b_n^2} \quad (2.5)$$

Asimismo, a $a_0/2$ se le conoce como componente continua de la señal y representa el desplazamiento del nivel medio de la señal respecto a la amplitud 0.

2.2.4. Ancho de banda

Otro concepto importante a la hora de analizar el proceso de transmisión de señales es el de ancho de banda, B . El ancho de banda es una característica natural de todos los medios y se define como el rango de frecuencias de las señales que se pueden transmitir por el medio físico.

En la Figura 2.5 se muestra un ejemplo de espectro de potencia de una señal electromagnética en el que se ha representado el valor del ancho de banda, B , del medio por el cual se está realizando la transmisión. Como se ha indicado anteriormente cuanto menor es el orden del armónico la potencia correspondiente tiende a ser mayor. Además, debido al ancho de banda limitado del medio, únicamente se podrán transmitir los armónicos con frecuencia menor o igual a B .

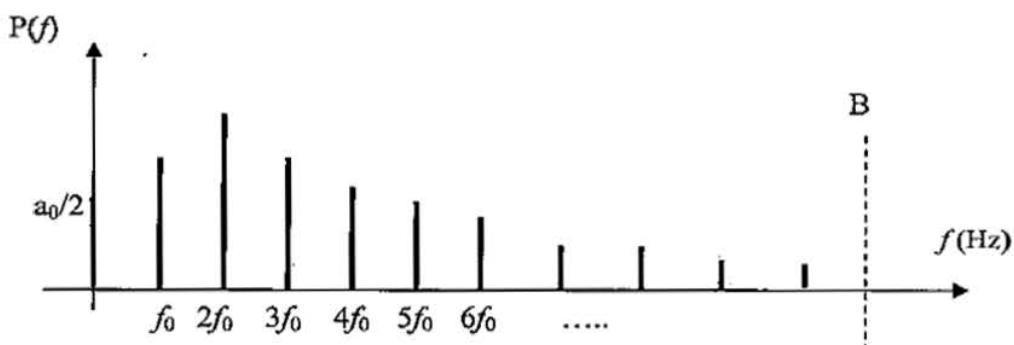


Figura 2-5. Espectro de potencia y ancho de banda.

Como ejemplo de descomposición de una señal compleja en armónicos simples, en la Figura 2.6 se muestra como puede reconstruirse al carácter ASCII 'a' a partir de sus armónicos simples. La señal de pulsos es analógica y en principio aperiódica. Sin embargo, y como se ha indicado con anterioridad, el desarrollo en serie de Fourier se aplica únicamente a señales periódicas. Para soslayar este problema se considera la transmisión del byte correspondiente de forma continua, por lo que el periodo de la señal se considera como el tiempo que se tarda en transmitir 8 bits.

En la parte superior de la Figura 2.6 se representa la señal de pulso original. A continuación se muestra cuál sería la señal reconstruida al considerar 1, 2, 3, 4, 5, 8 y 20 armónicos respectivamente. Como se observa en esta figura, al aumentar el número de armónicos en el medio la señal posee mejor definición. Así, cuanto mayor sea el ancho de banda más armónicos admitirá el medio y, por lo tanto, más fácil será realizar la reconstrucción de la señal original en el otro extremo de la comunicación. En esta figura se muestra que con menos de tres armónicos el sistema no es capaz de interpretar correctamente la señal recibida.

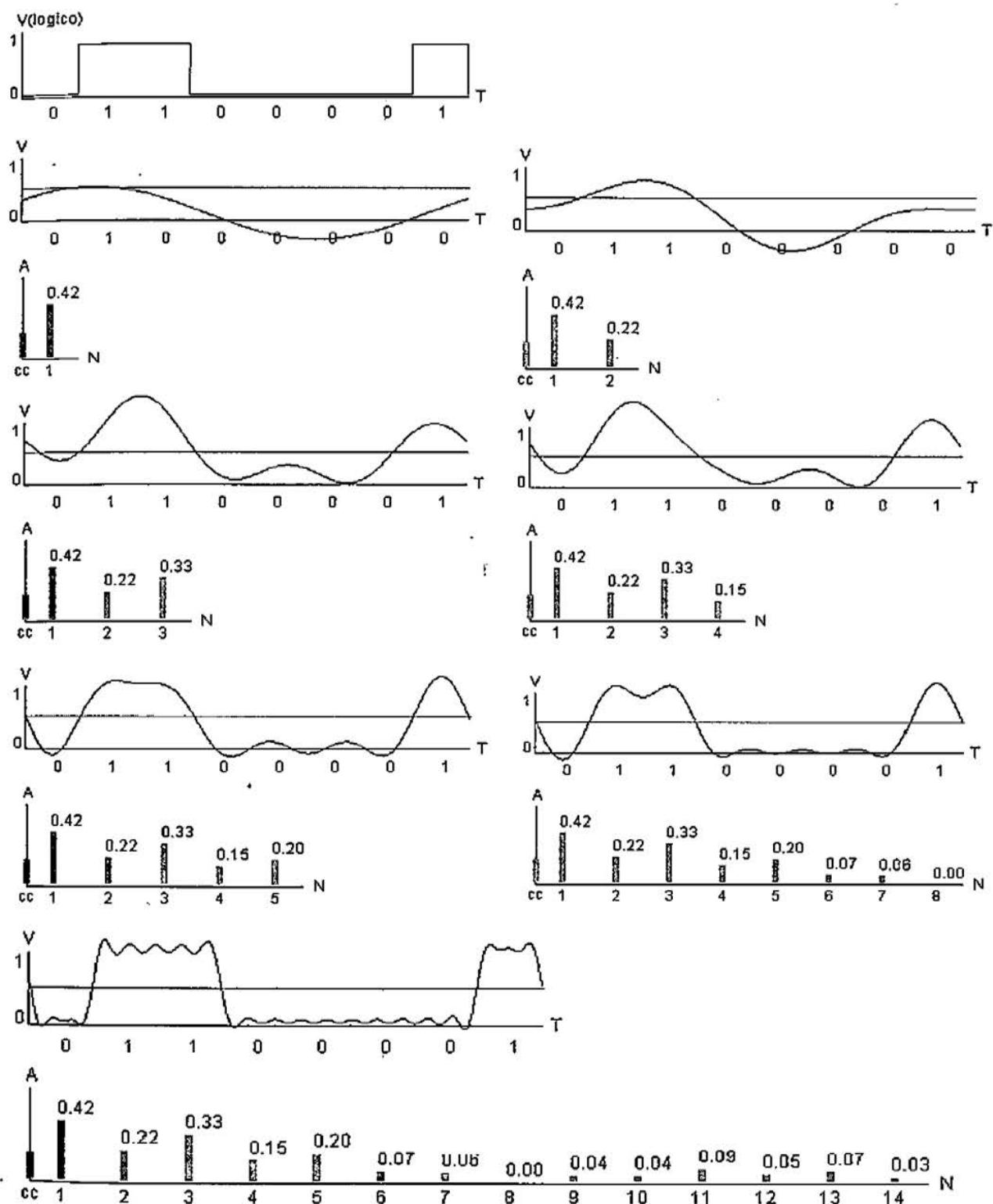


Figura 2-6. Reconstrucción del carácter "a" dependiendo del número de armónicos considerado.

2.2.5. Ganancia

Dado un determinado medio de transmisión, se define la ganancia del mismo como la relación entre la amplitud de la salida (V_s , tensión de salida) y la de la entrada (V_e , tensión de entrada). El comportamiento ideal es que la ganancia del medio sea de 1 para todo el rango de frecuencias, sin embargo, en general la evolución de la ganancia con respecto a la frecuencia de los medios de transmisión es similar al mostrado en la Figura 2.7. En dicho esquema se ha representado con f_i la frecuencia inicial, es decir, para que un determinado armónico pueda transmitirse por el medio es necesario que como mínimo presente esta frecuencia. De forma similar con f_c se indica la frecuencia superior de corte, de forma que los armónicos de mayor frecuencia no serán transmitidos por el medio. En la Figura 2.7 se han definido f_i y f_c para una atenuación de 3 dB ($G = -3\text{dB}$). Con ganancia de amplitud esto implica un valor de 0,707:

$$G = \frac{V_s}{V_e} \rightarrow G(\text{dB}) = 20 \log G = -3\text{dB} \rightarrow G = 0,707 \quad (2.6)$$

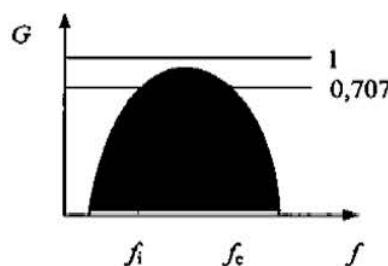


Figura 2.7. Ganancia de un medio de transmisión.

Ejercicio 2.1

Considérese que se tarda 1 mseg. en transmitir la señal correspondiente al carácter 'a' representada en la Figura 2.6 por un medio de transmisión telefónico. Se pide determinar cuántos armónicos se transmitirán por el citado medio. Para resolver este ejercicio es necesario tener en cuenta que el ancho de banda telefónico es de 4 KHz. Además en este caso se considera que $f_i=400\text{ Hz}$ y $f_c=4400\text{ Hz}$.

El primer paso para resolver el problema consiste en determinar el valor de la frecuencia fundamental que nos indicará cuánto estarán separados en frecuencia los armónicos.

$$f_0 = \frac{1}{T} = \frac{1}{10^3 \text{ seg}} = 1000\text{Hz.}$$

Esto quiere decir que se transmitirá el primer armónico que presenta una frecuencia de 1KHz (su frecuencia es mayor a 400 Hz), el segundo a 2 KHz, el tercero a 3 KHz y el cuarto a 4KHz. El quinto no se enviará ya que presentaría una frecuencia mayor a la frecuencia superior de corte f_c .

2.2.6. Prestaciones del medio físico

La velocidad de modulación V_m se define como la cantidad de veces por segundo que una señal puede cambiar su valor. Este parámetro indica lo rápido que puede cambiar la información y se mide en baudios.

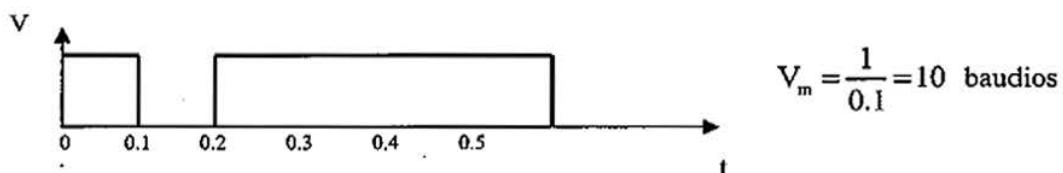


Figura 2-8. Velocidad de modulación.

Un medio de transmisión de b baudios podría transportar varios bits en cada instante. En este caso, la velocidad de transmisión (en bits por segundo) es igual a:

$$V_t (\text{bps}) = V_m \log_2 N \quad (2.7)$$

donde N es el número de estados significativos de la señal. En la Figura 2.9 se muestra una comparativa entre velocidad de transmisión y de modulación.

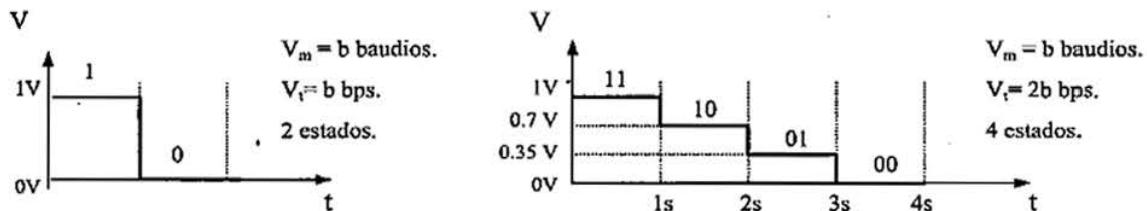


Figura 2-9. Velocidad de transmisión y de modulación

Si se aumenta el número de estados posibles de la señal, aumenta la velocidad de transmisión (sin tener que variar la velocidad de modulación):

- 2 estados: 1 bit en cada estado (0,1).
- 4 estados: 2 bits en cada estado (00, 01, 10, 11) → se duplica la velocidad inicial.
- 8 estados: 3 bits en cada estado (000, 001, 010, 011, 100, 101, 110, 111) → se triplica la velocidad inicial.
- 16 estados: 4 bits en cada estado (0000, 0001, ..., 1110, 1111) → se cuadriplica la velocidad inicial.

Por otra parte, existe una relación entre la velocidad de transmisión y el número de armónicos que pueden transmitirse. Para obtener esta relación considérese que se transmite un total de λ bits en T segundos. Por lo tanto, la velocidad de transmisión será de:

$$V_t = \frac{\lambda}{T} = \lambda \cdot f_0 \rightarrow f_0 = \frac{V_t}{\lambda} \quad (2.8)$$

Por lo tanto, si aumenta la velocidad de transmisión, también aumentará la frecuencia fundamental y, consecuentemente, la separación entre los armónicos. Considerando un medio de transmisión con un ancho de banda B y una frecuencia inicial $f_i=0$, se cumplirá lo siguiente:

$$n \cdot f_0 \leq B \rightarrow n \frac{V_t}{\lambda} \leq B \rightarrow n \leq \frac{\lambda B}{V_t} \quad (2.9)$$

Según esta expresión, al aumentar la velocidad de transmisión, menor será la cantidad de armónicos transmitidos para un ancho de banda, B , dado. Por lo tanto, el ancho de banda limita la velocidad de transmisión máxima que puede alcanzarse, ya que si se aumenta excesivamente puede ocurrir que el número de armónicos transmitidos no sean suficientes para la posterior reconstrucción de la señal como se mostró en el Apartado 2.2.4. En la Tabla 2.1, se muestra para el carácter 'a' de la Figura 2.6 cuantos armónicos se transmitirían por un medio de ancho de banda de 1KHz según se aumenta la velocidad de transmisión.

Bps	Frecuencia fundamental	Armónicos en canal	Frecuencia del último armónico (Hz)
100	12,5	80	1000
300	37,5	26	975
600	75	13	975
1200	150	6	900
2400	300	3	600
4800	600	1	600
9600	1200	0	0

Tabla 2-1: Armónicos transmitidos dependiendo de la velocidad de transmisión. $B=1\text{KHz}$.

2.3. Perturbaciones en la transmisión

A la hora de caracterizar un medio de transmisión no sólo hay que tener en cuenta su ancho de banda que limita la calidad de la información reconstruida, sino que existen otras fuentes de error que son denominadas perturbaciones. Como se ha indicado anteriormente, la energía de la señal recibida debe ser suficiente para que la electrónica en el receptor pueda detectar la señal adecuadamente. Dicho de otro modo, los armónicos que lleguen al receptor deben contener suficiente información como para poder reconstruir la señal. Además, la señal recibida debe conservar un nivel suficiente en relación a la perturbación. Es necesario, por lo tanto, analizar qué perturbaciones pueden aparecer en el medio y que de alguna forma van a limitar y distorsionar los armónicos que lleguen al receptor.

Las perturbaciones más comunes que aparecen a menudo en cualquier medio de transmisión, son las siguientes:

- **Atenuación.** Se trata de un decremento en la amplitud de los diferentes armónicos transmitidos. La atenuación depende de la frecuencia a la que es transmitido un determinado armónico, por lo tanto, algunos armónicos sufrirán mayor atenuación que otros. En general la energía de la señal transmitida decrece con la distancia y la atenuación es habitualmente una función creciente de la frecuencia. La atenuación se mide en decibelios y se cumple que:

$$A(\text{db}) = 10 \log_{10} \frac{P_s}{P_e}, \quad V(\text{db}) = 20 \log_{10} \frac{V_s}{V_e} \quad (2.10)$$

donde P_s y P_e es la potencia de salida y entrada respectivamente, mientras que V_s y V_e es la tensión de salida y entrada.

- Distorsión de retardo. Se trata de un retraso en la llegada de algunos armónicos al otro lado de la línea. Para una señal que se transmite por un medio con ancho de banda limitado, la velocidad de propagación tiende a ser mayor cerca de la frecuencia central y disminuye al acercarse a los extremos f_i y f_c . Por lo tanto, los distintos armónicos llegarán al receptor en instantes de tiempo diferentes, dando lugar a desplazamientos de fase entre las diferentes frecuencias.
- Ruido. Se trata de una perturbación aleatoria que se encuentra en cualquier medio de transmisión. Esta perturbación se mide mediante la denominada relación señal-ruido en decibelios. Un valor alto de esta relación implicará una buena transmisión ya que el medio se verá afectado en menor medida por el ruido.

$$\left(\frac{PS}{PN} \right)_{\text{dB}} = 10 \log_{10} \frac{PS}{PN} \quad (2.11)$$

donde PS y PN representan la potencia de la señal y el ruido respectivamente.

2.4. Velocidades máximas de envío de datos

En este apartado se va a describir la manera de obtener la velocidad máxima que puede alcanzarse en un medio físico según dos criterios distintos: el teorema de muestreo y la relación señal/ruido en el medio.

2.4.1. Teorema de muestreo

El teorema de muestreo o teorema de Nyquist afirma que al hacer pasar una señal por un medio con ancho de banda B , dicha señal puede reconstruirse con muestras tomadas a una frecuencia $f_m = 2B$, no siendo necesario una frecuencia de muestreo mayor.

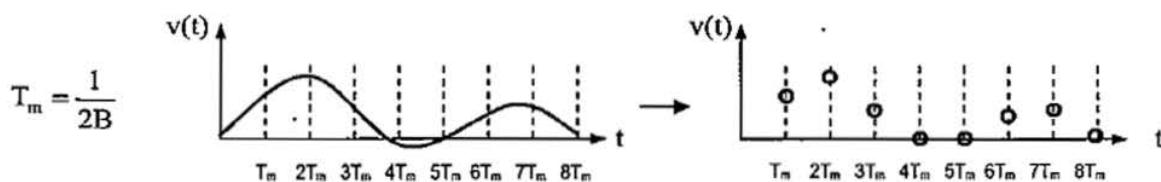


Figura 2-10. Proceso de muestreo

A mayor frecuencia de muestreo f_m se capturan más muestras de la señal original por unidad de tiempo. Esto implica que las muestras detectan variaciones más pequeñas, por lo tanto, en principio a mayor frecuencia de muestreo, más información se dispondrá de la señal original y más fácil será su

posterior reconstrucción. Sin embargo, a partir de $f_m = 2B$ no se conseguirá mayor calidad porque el medio elimina las componentes de alta frecuencia (es decir, frecuencias por encima de B).

2.4.2. Velocidad máxima según el teorema de muestreo

De lo dicho en el Apartado 2.4.1 se puede concluir que la variación de señal más pequeña detectable dura $T_m=1/2B$, de manera que la velocidad de modulación máxima es de $V_{mMAX}=1/T_m=2B$. En general cuando se utilizan N niveles o estados y considerando un medio sin ruido, la velocidad máxima teórica a la que es posible transmitir cualquier carácter en un medio de ancho de banda B es:

$$V_t(\text{bps}) = 2B \log_2 N \quad (2.12)$$

2.4.3. Velocidad máxima según la relación señal/ruido

En el caso de un medio con ruido, *Shannon* estimó que la velocidad de envío de datos máxima en un canal ruidoso con relación de potencias señal/ruido (PS/PN) y cuyo ancho de banda es B Hertzios es:

$$V_t(\text{bps}) = B \log_2 \left(1 + \frac{PS}{PN} \right) = B \frac{\log_{10} \left(1 + \frac{PS}{PN} \right)}{\log_{10} 2} \quad (2.13)$$

donde PS/PN no se expresa en decibelios y:

$$\left(\frac{PS}{PN} \right)_{\text{db}} = 10 \log_{10} \left(\frac{PS}{PN} \right) \quad (2.14)$$

El límite establecido por Shannon es un límite físico de los medios de comunicación debido al efecto del ruido. Como se observa de la Ecuación (2.14), la velocidad máxima establecida por el teorema de Nyquist puede incrementarse aumentando progresivamente el número de niveles. Sin embargo, la velocidad máxima obtenida por el teorema de Nyquist no podría alcanzarse si supera el límite máximo establecido por Shannon. Esto es debido a que el efecto del ruido provocaría que los niveles de señal serían irreconocibles.

Ejercicio 2.2

Una empresa tiene la sede central en Alicante y varias sedes localizadas en otras ciudades: Elche, Alcoy, Denia, Villena y Orihuela. En cada sede se dispone de una red IEEE 802.3, que conectan varias computadoras, un servidor y un router que proporciona conectividad con el exterior. Si además se sabe que los routers disponen de un MODEM que funciona a 3200 baudios, y la relación señal/ruido de la línea telefónica es de 60dB. Calcula la máxima velocidad a la que puede transmitir el router.

$$\left\{ \begin{array}{l} V_t = 2B \cdot \log_2(N) \\ V_t = B \cdot \log_2\left(1 + \frac{PS}{PN}\right) \end{array} \right\} \rightarrow 2 \cdot \log_2(N) = \log_2\left(1 + \frac{PS}{PN}\right) \rightarrow N = \sqrt{1 + \frac{PS}{PN}}$$

$$\left(\frac{PS}{PN}\right)_{dB} = 10 \log_{10}\left(\frac{PS}{PN}\right) \rightarrow \frac{PS}{PN} = 10^6 = 1000000$$

$$N = \sqrt{(1 + 1000000)} \approx 1000$$

$$V_t = V_m \cdot \log_2 N = 3200 \cdot \log_2 1000 = 3200 \cdot 10 = 32000 \text{ bps}$$

2.5. Filtrado de señales

Como se ha descrito anteriormente, en los sistemas de comunicaciones se desea manejar información que debe estar dentro de un determinado rango de frecuencias, por lo tanto, a menudo surge la necesidad de permitir o eliminar ciertos grupos de frecuencias de la señal original. Esta función es realizada por los filtros. De hecho, los medios de transmisión funcionan como filtros que únicamente permiten transmitir en el rango de frecuencias correspondiente a su ancho de banda. Así, el estudio de los filtros, nos permitirá interpretar las características de los enlaces físicos que se comportan como cierto tipo de filtros.

Los filtros constituyen una parte de las redes de comunicaciones que presentan características selectivas de frecuencias. Esto significa que la atenuación en ellos es variable con la frecuencia, lo cual permite discriminar las señales que pasarán libremente a través del filtro. Si la señal que se aplica a la entrada del filtro contiene armónicos en un amplio rango de frecuencias, dicho filtro actuará de manera que solamente algunos armónicos en determinadas frecuencias aparecerán a la salida. En base a la función principal de los filtros, que es permitir el paso de armónicos en las frecuencias que se desea y presentar una atenuación elevada para las frecuencias indeseables, estos pueden ser:

- *Filtros pasa bajo:* dejan pasar las frecuencias que están por debajo de la seleccionada por el filtro. Sólo permite el paso de las frecuencias inferiores a la frecuencia de paso o de corte f_c y atenúa o suprime todas las frecuencias superiores a la frecuencia de corte.
- *Filtros pasa alto:* dejan pasar las frecuencias que están por encima de la seleccionada por el filtro. Sólo Permite el paso de las frecuencias superiores a la frecuencia de corte o de paso f_c y atenúa todas las frecuencias menores a la frecuencia de corte.
- *Filtros pasa banda:* son una combinación de los dos filtros anteriores. Estos eliminan las componentes cuyas frecuencias están por encima o por debajo de unos límites o frecuencias de corte de cada filtro. Sólo las frecuencias comprendidas entre ellas pasan a su través. Esta banda de frecuencias permitidas se llama banda de paso. Por lo tanto, permite el paso de un rango de frecuencias definido entre una frecuencia de corte inferior f_{ci} y una frecuencia de corte superior f_{cs} .
- *Filtros banda eliminada:* estos filtros realizan la función inversa de los filtros pasa banda, es decir, dejan pasar las frecuencias de la señal que se encuentren por debajo de la frecuencia de corte inferior, f_{ci} , o por encima de la frecuencia de corte superior, f_{cs} , pero no las que se encuentren entre este rango de frecuencias.

Ejemplo*Distribución del ancho de banda en ADSL y filtrado.*

El principio de funcionamiento de la tecnología ADSL se basa en que en un cable telefónico se puede llegar a transmitir un espectro de más de 1MHz, y para las conexiones analógicas solo se usa el espectro de la voz (4KHz). Por lo tanto, el ancho de banda aprovechable del medio físico será mayor a 1 MHz y se divide de la siguiente manera (ver Figura 2.11):

Las frecuencias inferiores se utilizan para el canal de la señal de voz analógica ya existente.

En medio está el espectro dedicado al enlace ascendente o del usuario a la red.

En las frecuencias superiores está el enlace descendente o desde la red al usuario.

Un filtro pasa bajo con frecuencia de corte $f_c = 22$ KHz permitiría aislar el espectro de frecuencias correspondiente a la voz del resto de información que circula por el medio. A este filtro se le conoce como Splitter. Además, un filtro pasa alto con frecuencia de corte $f_c = 230$ KHz podría emplearse para obtener la información procedente del canal descendente, es decir, los datos que se están descargando de la red. Por último, para aislar los datos del canal ascendente podría emplearse un filtro pasa banda con una frecuencia de corte inferior $f_{ci} = 23$ KHz y una frecuencia de corte superior $f_{cs} = 220$ KHz. En la Figura 2-13 se ha representado todo el proceso de filtrado de señales en ADSL.

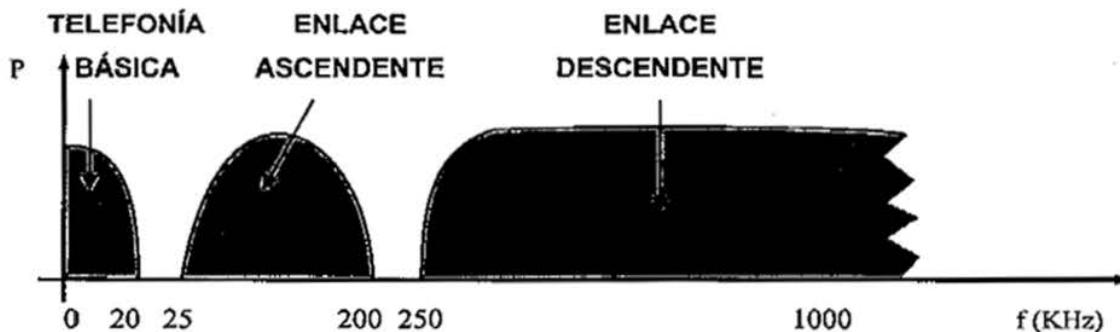


Figura 2-11. Distribución del ancho de banda en ADSL

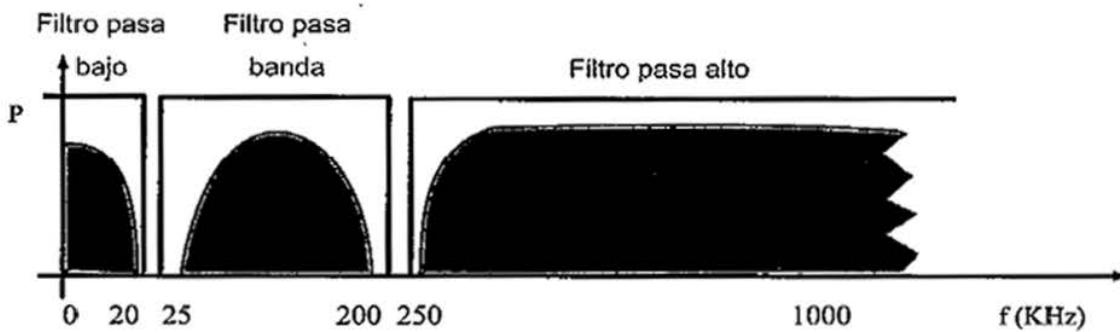


Figura 2-12. Filtrado de la información en ADSL.

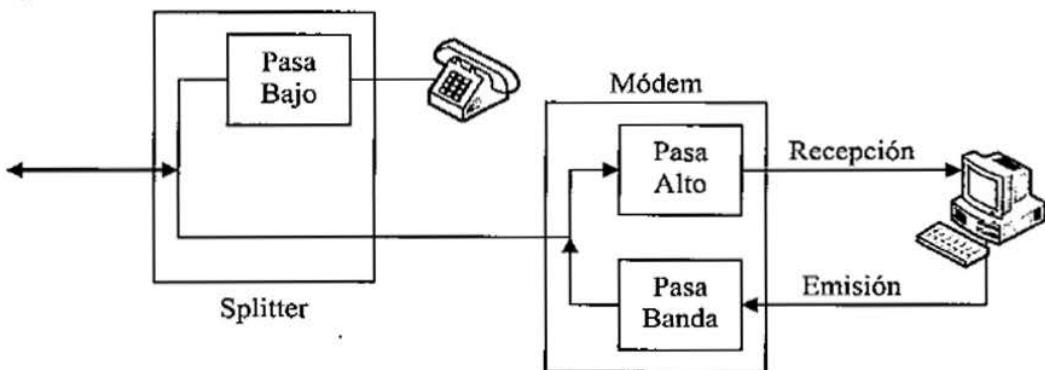


Figura 2-13. Esquema del proceso de filtrado de información en ADSL.

2.6 Transmisión analógica y digital. Definiciones y conceptos básicos

Como se ha indicado en el apartado 2.2.1, a la hora de realizar la transmisión de información se puede optar por emplear información analógica o bien discretizar esta información y transmitirla en forma de señales digitales. Tanto la transmisión analógica como la digital presentan una serie de propiedades que es necesario analizar con el objetivo de conocer cual será el desempeño de la misma. Así, por ejemplo, cuando se emplea transmisión analógica la señal tiende a debilitarse con la distancia lo cual implica en ocasiones tener que emplear amplificadores con el objetivo de aumentar su energía. Este tipo de dispositivos, aunque regeneran la señal, también amplifican el ruido que haya podido aparecer durante la transmisión. En medios digitales también afecta dicha atenuación, sin embargo, es más fácil diferenciar entre 0 y 1 por parte del receptor. En general la transmisión digital presenta notables ventajas frente a la analógica dentro de las cuales cabe mencionar:

- Disminución del coste en las tecnologías utilizadas. Los dispositivos empleados (p. ej. multiplexores, codificadores, commutadores, etc.) presentan menor coste. Además el tendido de líneas de transmisión de banda ancha es más económico.
- Se consigue una transmisión de datos a distancias mayores utilizando líneas de calidad inferior.
- Cuando se transmiten datos digitales se puede alcanzar un alto grado de multiplexión como se verá en el capítulo 3 y de una forma más sencilla que empleando datos analógicos.
- Se pueden aplicar técnicas de seguridad y privacidad como puede ser encriptación de la información digital transmitida.
- Es mucho más fácil restaurar y procesar las señales mediante software fácilmente configurable y actualizable.



2.7 Modos de transmisión

En primer lugar es necesario diferenciar entre transmisión serie y paralela. En una transmisión serie los bits que componen la trama de datos son transmitidos secuencialmente sobre una línea. Sin embargo, una transmisión paralela consiste en transmisiones simultáneas de N cantidades de bits. Estos bits se envían simultáneamente a través de N canales diferentes. Estos canales pueden ser:

- N líneas físicas: en cuyo caso cada bit se envía en una línea física.
- Una línea física dividida en varios subcanales. Estos canales se obtienen según la división del ancho de banda, por lo que por cada línea se transmite a una frecuencia distinta.

Actualmente, se emplea transmisión serie en la mayoría de los casos, utilizándose la paralela únicamente para buses internos de una arquitectura de procesador. La transmisión serie es más económica y sencilla de instalar. También presenta un mayor alcance (en el caso de transmisión paralela la longitud es de varios metros mientras que en transmisión serie se pueden alcanzar cientos de metros sin modulación). Hoy incluso se emplea la transmisión serie para dispositivos internos (discos duros, USB), e incluso entre procesadores.

Un aspecto que se ha de tener en cuenta a la hora de realizar la transmisión de bits en serie es la posibilidad de utilizar transmisión síncrona o asíncrona. Esta clasificación se hace atendiendo a la forma en la que el receptor determina donde se encuentran los bits intercambiados en una transmisión.

En comunicación síncrona se requiere emplear una señal de reloj. Esta señal, también denominada reloj de bit, presenta una forma de pulsos de manera que cada uno de ellos indica el momento en el que se debe hacer la lectura de la señal de datos. Atendiendo a como se especifica la señal de reloj se puede encontrar comunicación síncrona empleando una línea con la señal de reloj de bit independiente de la línea de datos (Figura 2.14) o comunicación síncrona con reloj de bit junto con los datos (Figura 2.15). La primera aproximación es empleada habitualmente en las interfaces DTE-DCE actuales mientras que la segunda es más utilizada en las LAN.

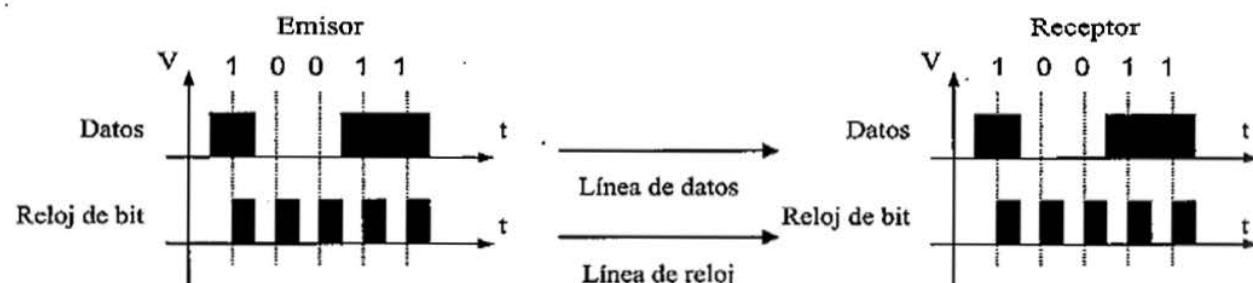


Figura 2-14. Esquema de transmisión síncrona. Datos y reloj en líneas distintas.

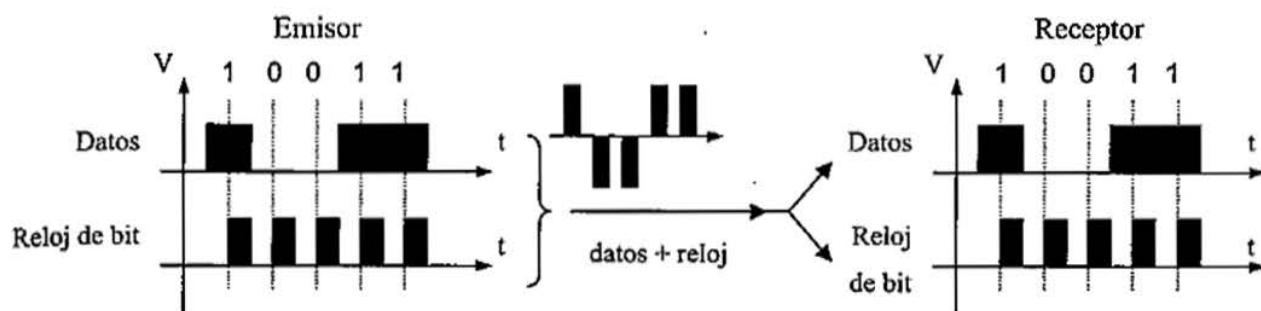


Figura 2-15. Esquema de transmisión síncrona. Datos y reloj en la misma línea.

En comunicación asíncrona no se envía la señal de reloj de bit, sino que se envía un pulso de sincronización cada cierto número de bits. Este tipo de comunicación se emplea en interfaces DTE-DCE lentas (ver Figura 2.16) o medios que no facilitan la sincronización.

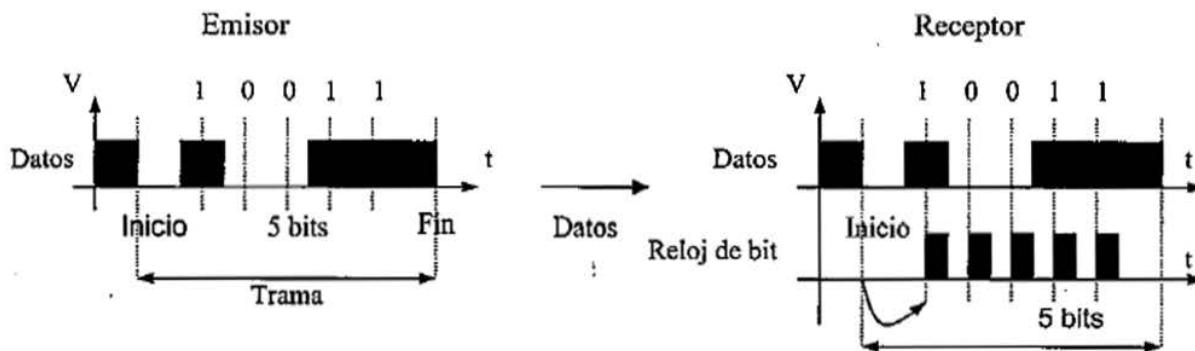


Figura 2-16. Esquema de transmisión asíncrona.

Capítulo 3. CODIFICACIÓN DE LA INFORMACIÓN

En una red de computadores se necesita que la información, que se intercambian dos o más dispositivos, sea transformada para adecuarse al medio físico de transmisión que los conecta. En este capítulo se trata la conversión de señales digitales y analógicas, así como las técnicas de modulación que se emplean para llevar a cabo la codificación de datos en señales, en función de su naturaleza digital o analógica. Además, en este capítulo, se explica como conseguir una adecuada eficiencia del medio físico haciendo uso de técnicas de multiplexación. Con estas técnicas se permite transmitir simultáneamente varias señales a través de un mismo enlace de datos.

CAPÍTULO 3

3-1 Señales y tipos de señalización

- Tipos de datos y de señales
- Tipos de señalización en transmisiones digitales

3-2 Codificación Banda Base

- Codificación binaria RZ y RNZ
- Codificación Manchester y Manchester diferencial
- Codificación mBnL

3-3 Codificación Banda Modulada

- Modulación analógica: ASK, FSK, PSK, QPSK, QAM y otras
- Modulación digital: PCM, PCM diferencial y DM

3-4 Multiplexación

- Multiplexación por división en frecuencias
- Multiplexación por división en tiempos

OBJETIVOS

- Describir la naturaleza de los tipos de datos que se pueden transportar en un mensaje y de las señales que se emplean para transmitirlos
- Comentar los mecanismos para construir las señales que transportan datos dependiendo de la naturaleza de los datos y del medio por el cual se van a transmitir
- Mostrar ejemplos de codificaciones y ámbitos de uso en redes de computadores
- Clasificar y explicar el concepto de multiplexación y los métodos de multiplexación más empleados en comunicaciones analógicas y digitales
- Mostrar ejemplos de procesos de multiplexación y el uso en redes de comunicación

TERMINOLOGÍA – PALABRAS CLAVE

- | | | |
|---------------------|---------------------------|-------------------------|
| • Datos analógicos | • Multinivel | • ASK/FSK/PSK |
| • Datos digitales | • RZ/RNZ | • QPSK/QAM |
| • Señal analógica | • Manchester | • PCM |
| • Señal digital | • 2B1Q/8B6T/4B5B | • PCM
diferencial/DM |
| • Elemento de dato | • MLT-3/HDB3 | • FDM |
| • Elemento de señal | • Señal portadora | • TDM |
| • Banda Base | • Señal moduladora | • WDM |
| • Banda Modulada | • Señal modulada | • CDMA |
| • Unipolar | • Modulación
analógica | • mBnL |
| • Bipolar | • Modulación digital | |

3.1. Señales y tipos de señalización

En el capítulo anterior, se hizo una clasificación de tipos de señales en función de su naturaleza. Además, se hizo hincapié en como las señales que se transmiten a través de un medio de transmisión tienen la finalidad de transportar secuencias de bits correspondientes a la información que se desea comunicar e intercambiar entre emisor y receptor. Sin embargo, no se comentó como esa información binaria se transporta en forma de señal y si es necesario que esta información sea codificada o no. Pero antes de comentar estos dos aspectos, conviene hacer una breve clasificación de los tipos de datos que pueden componer la información que se desea enviar, y el tipo de señales que se pueden emplear para transportar estos datos.

3.1.1. Tipos de datos y de señales

Los mensajes de información que intercambian un emisor y un receptor en un sistema de comunicación pueden estar compuestos por datos digitales o por datos analógicos. El empleo de un tipo u otro de datos para componer el mensaje depende de los recursos y dispositivos disponibles en la comunicación. Así, si emisor y receptor son dos seres humanos, los datos los compondrán con naturaleza analógica, sin embargo si emisor y receptor son dos PCs, los datos los compondrán con naturaleza digital.

Del mismo modo, ya sean analógicos o digitales, esos datos en muchos casos es necesario adaptarlos al medio físico que se empleará en el proceso de transmisión. Por lo tanto, estos datos se someterán a lo que se conoce como proceso de señalización o codificación para llevar a cabo esta adaptación.

En función de los datos que constituyen el mensaje y las señales que se emplean para transportar dichos datos, se dispone de las siguientes cuatro situaciones:

- *Datos digitales+señales digitales:* Es la forma más simple de codificar, ya que datos y señales tienen la misma naturaleza digital. Este tipo de codificación se conoce como codificación de línea. La manera más sencilla para realizar la codificación de línea es asignar un nivel de tensión al '1' binario y otro nivel de tensión distinto al '0' binario. Sin embargo, en muchos casos para mejorar las prestaciones conviene emplear bloques de datos mayores a un bit, para ser codificados en niveles de tensión, así como mecanismos de sincronismo. A este proceso de codificación se le conoce como *banda base*.
- *Datos analógicos+señales digitales:* En este caso, los datos son analógicos como por ejemplo la voz humana o el video y televisión convencional. En tales casos, es necesario someter a los datos a un proceso de digitalización para poder ser transmitidos a través de sistemas digitales. La técnica más empleada en los procesos de digitalización es la que se emplea para almacenar música en soportes digitales como es la codificación PCM y sus variantes. El dispositivo empleado se denomina *códec*.
- *Datos digitales+señales analógicas:* Este otro caso es el caso más común que nos encontramos en nuestros hogares cuando queremos comunicar nuestro PC con otro a través de la red telefónica conmutada. Los datos almacenados en memoria de los computadores personales son secuencias de bits que representan ficheros, imágenes, texto, video, etc. y éstos se quieren transportar a través de líneas de cableado que

soportan señales analógicas. En tal caso, un dispositivo DCE del tipo módem convertirá los datos digitales en señales analógicas. Las técnicas que emplean los módems para adecuar la señal son las basadas en modulación por desplazamiento de fase (PSK), modulación por variación de amplitud (ASK) y modulación por variación de frecuencias (FSK). A este proceso de codificación se le denomina *banda modulada*.

- *Datos analógicos+señales analógicas:* El último de los casos es cuando los datos son analógicos, como la voz humana, y para ser transmitido por líneas analógicas no deben alterar su naturaleza. Estos datos simplemente se modulan mediante una señal cuya banda de frecuencias (FM), amplitudes (AM) o fases (PM) son las adecuadas para transmitir por el medio físico escogido.

Este capítulo se centrará principalmente en comentar las técnicas de codificación que se emplean cuando los datos son digitales. Aunque también se comentará como se realiza la transmisión de datos analógicos a través de señales digitales.

3.1.2. Tipos de señalización en transmisiones digitales.

Como ya se ha comentado anteriormente, existen dos posibles tipos de señalización en transmisiones digitales: aquellos que emplean datos analógicos y los que emplean datos digitales, en ambos, las señales generadas serán digitales. El proceso de señalización se denomina *codificación de línea*. Para hacer más entendible este proceso se van a definir dos conceptos: *elemento de señal* y *elemento de datos*. Ambos constituyen posibles elementos de señalización en el proceso de codificación.

- *Elemento de datos:* Es la unidad más pequeña que representa un elemento de información a transmitir.
- *Elemento de señal:* Es la unidad más pequeña en duración de una señal digital que se transmite.

De ahí se puede determinar que el elemento de datos es la información que se transporta en la señal y el elemento de señal son los portadores de los datos.

De estos dos conceptos se puede extraer la relación existente entre velocidad de modulación y velocidad de transmisión que se comentó en el capítulo 2. Así, la velocidad de transmisión mide la tasa de datos que define el número de elementos de datos (bits) enviados por unidad de tiempo (segundo). Sin embargo, la velocidad de modulación determina el número de elementos de señal enviados por unidad de tiempo (segundo), es decir determina la tasa de señal.

Muchas veces, uno de los objetivos de un sistema de comunicaciones es aumentar la tasa de datos al mismo tiempo que se busca reducir la tasa de señales. Este hecho supone un aumento en la velocidad de transmisión sin necesidad de aumentar el ancho de banda que se necesita para transmitir la señal. Y algunas otras, se emplea una tasa de elementos de señal mayor que la velocidad de transmisión, por motivos de seguridad y fiabilidad.

Las técnicas de codificación se pueden catalogar en métodos de señalización del tipo:

- *Banda base:* Cuando se emplean estas técnicas, la información a transmitir (secuencia de bits) se envía empleando una codificación de línea, es decir sin necesidad de realizar modificaciones para enviarla por el medio.
- *Banda modulada:* En este tipo de técnicas, la información a transmitir (secuencia de bits) tiene que ser adaptada para ser enviada a través del medio físico.

El empleo de unas técnicas de codificación u otras depende del medio físico que se va a utilizar en la transmisión y de las especificaciones de éste, así como de otros factores que influyen en el intercambio de datos entre dispositivos. Así, factores como el ancho de banda del medio, sincronización entre emisor y receptor, inmunidad al ruido y a las interferencias en el medio físico, el coste de complejidad de implementación de la técnica de codificación o la necesidad de incorporar mecanismos para la detección de errores, son algunos de los factores que determinan la técnica a escoger e implementar.

- *Sincronismo:* Se requiere que los métodos de codificación, en muchos casos, incorporen mecanismos para interpretar correctamente las señales que el receptor recibe procedente del emisor, es decir dónde empieza y terminan los datos, dónde comienza y termina cada bit de datos, etc.
- *Ruido e Interferencias:* Se requiere que el método de codificación sea lo más inmune a perturbaciones.
- *Complejidad:* Los métodos que necesita mayor número de niveles de señal son más costosos de implementar que los sencillos.
- *Detección de errores:* Es importante que los métodos de codificación incorporen mecanismos para la detección de errores en el código generado y transmitido.

3.2. Codificación Banda Base

Existen varios mecanismos de codificación banda base. Estos se pueden catalogar de acuerdo a una serie de características que determinan su espectro de señal:

- *Unipolar:* Cuando en la señal todos los niveles de amplitud se encuentran a un mismo lado del eje de tiempo. Es decir, todos los niveles tienen valores de tensión, mayores o iguales a cero, o todos son menores o iguales a cero.
- *Bipolar:* Cuando en los niveles de amplitud se pueden encontrar a ambos lados del eje de tiempo. Es decir, pueden tener valores positivos o negativos indistintamente.
- *Multinivel:* Cuando en la señal se pueden tener diferentes tipos de elementos de señal, permitiendo diferentes niveles de amplitud de señal para codificar los elementos de datos. Esto permite aumentar la velocidad de transmisión, aumentando el número de bits por baudio que se codifican por elemento de señal.

3.2.1. Codificación binaria

Son mecanismos de codificación unipolar o bipolar. En estos casos, se suele asignar niveles de tensión asociados a cada valor de bit ('0' o '1').

De entre ellos destacan las codificaciones NRZ (No retorno a cero) y RZ (Retorno a cero). Tradicionalmente, el esquema NRZ y RZ fueron diseñados como unipolares.

En el NRZ, el dato '1' se codifica como un nivel de tensión positivo y el dato '0' como un nivel de tensión negativo. Se denomina NRZ porque la señal no regresa a nivel de tensión cero en mitad del elemento de señalización que define el bit en la señal.

Sin embargo, el RZ nace para solventar algunos de los problemas que tiene el método de codificación NRZ. Así, en el método NRZ, el principal problema surge cuando emisor y receptor no

están sincronizados para el envío y recepción de datos. Esto es así porque el receptor no sabe dónde termina un bit y comienza el siguiente con ese mecanismo de señalización. El RZ utiliza tres valores de tensión (positivo, negativo y cero) para codificar los datos y la señal no cambia entre bits sino durante la propia duración del bit. La señal vuelve a cero en la segunda mitad de la duración de cada bit (Figura 3.1).

Generalmente, éste tipo de métodos se suele emplear en la grabación magnética de datos digitales y no suelen ser muy atractivos para la codificación previa necesaria para transmisión de señales.

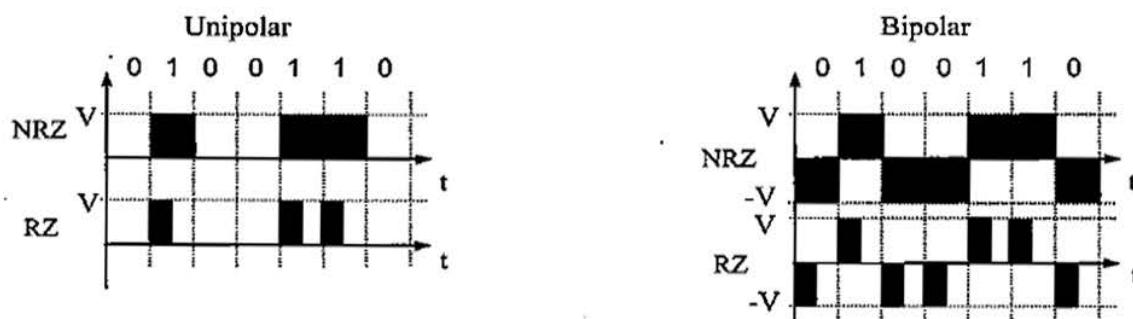


Figura 3-1: Ejemplos codificaciones binarias.

3.2.2. Codificación Manchester

Son mecanismos de codificación bipolar. En estos casos, a cada valor de bit ('0' o '1') se le asocian transiciones de nivel entre dos tensiones. El nivel de tensión permanece en un valor durante la primera mitad y se mueve a otro nivel durante la segunda mitad de duración del bit. Esta transición en mitad de la duración del bit permite utilizarse como elemento de sincronización entre emisor y receptor (Figura 3.2).

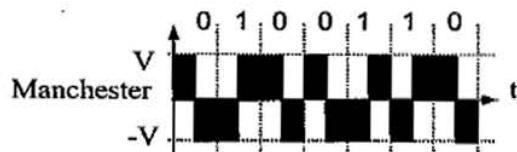


Figura 3-2: Ejemplo codificación Manchester.

En este método es posible incorporar mecanismos de detección de errores. Mecanismos que consistirían en comprobar la presencia o ausencia de transiciones en mitad del intervalo de tiempo. En caso de ausencia, se estaría produciendo un error.

Este método se suele emplear para la codificación de señales que se transmiten en redes LAN con topología en bus sobre par trenzado o cable coaxial y empleando método de acceso al medio CSMA/CD, es decir, en concreto las redes del tipo IEEE 802.3 (Consultar Capítulo 4).

3.2.3. Codificación Manchester diferencial

También son mecanismos de codificación bipolar. En estos casos, a cada valor de bit ('0' o '1') se le asocian transiciones de nivel entre dos tensiones. El nivel de tensión permanece en un valor durante la primera mitad y se mueve a otro nivel durante la segunda mitad de duración del bit. La diferencia con el método Manchester viene dada por la forma de determinar los valores del bit al comienzo de éste. Así, si un bit se codifica con una transición dada, y el siguiente bit es un '0' entonces se repetirá la misma transición, pero si es un '1' entonces se cambia la transición (Figura 3.3).

Este método se suele emplear para la codificación de señales que se transmiten en redes LAN con topología en anillo y empleando método de acceso al medio basado en testigo, es decir, en concreto las redes del tipo IEEE 802.5 (Consultar Capítulo 4).

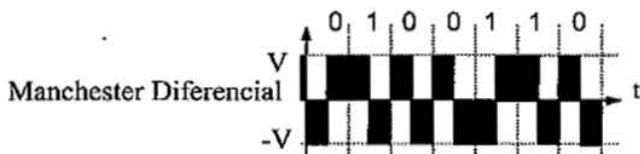


Figura 3-3: Ejemplo codificación Manchester Diferencial.

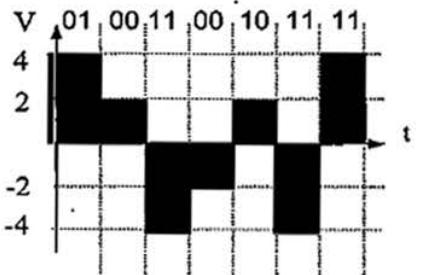
3.2.4. Codificación mBnL

Las técnicas de codificación mBnL tienen como objetivo incrementar el número de bits por baudio, es decir por elemento de señal. De este modo, se consigue un aumento en la velocidad de transmisión sin necesidad de aumentar el ancho de banda necesario para transmitir la señal. Esto es así, porque por cada baudio se están transmitiendo más bits y por lo tanto la tasa de bits a transmitir por un ciclo de señal. En estas técnicas de codificación, m es la longitud en bits del patrón de datos a enviar, B significa dato binario, n es la longitud en elementos de señal del patrón de señal codificado y L es el número de niveles de señal. Para implementar estas técnicas, se codifican bloques de datos de m bits de datos en n elementos de señal. Esto significa que m bits de datos pueden producir 2^m patrones de datos. En el caso de que $2^m=L^n$ cada patrón de datos se codifica con el mismo número de elementos de señal. Frecuentemente, el parámetro L que determina el número de niveles de señal cambia su nomenclatura por otras letras representativas. Así, cuando L toma el valor 2 se sustituye por la letra B de binario, cuando toma el valor 3 por la letra T de ternario o de modo similar, cuando $L=4$ es sustituido por la Q de cuaternario.

En estas técnicas, la asignación entre patrones de señal y patrones de datos se hace mediante el empleo de una tabla de codificación.

Codificación 2B1Q

Se trata de un mecanismo multinivel. Su nombre indica que codifica patrones de $m=2$ elementos de datos en un patrón de $n=1$ elemento de señal, donde se tienen $L=Q=4$ niveles de señal diferentes. Los dos tipos de elementos de datos son '0' o '1' y por lo tanto se pueden producir $2^m=2^2=4$ patrones de datos. Cada elemento de señal presenta 2 bits, y el conjunto de todos los pares de bits posibles representan los 4 patrones de datos. Además $L=4$ el número de niveles de tensión que puede adoptar la señal a transmitir (Figura 3.4). La asignación entre patrones de señal y patrones de datos se hace con una tabla que considera el estado previo de la señal. En este caso $2^m=L^n$, ya que $2^2=4^1$.



Codificación	Nivel Anterior (+)	Nivel Anterior (-)
00	+2	-2
01	+4	-4
10	-2	+2
11	-4	+4

Figura 3-4: Ejemplo de codificación 2B1Q.

Por ejemplo, supóngase que se quiere codificar la secuencia '01' en un elemento de señal. Así, como la señal en un instante anterior tenía un nivel positivo de tensión, el '01' se codifica con una tensión de +4 voltios. Sin embargo, si la señal en el instante anterior hubiese tenido un nivel negativo de tensión se hubiera codificado como -4 voltios (Figura 3.4).

Este tipo de codificación se emplea en las líneas xDSL (líneas de abonado digital) para ofrecer una conexión de alta velocidad en Internet empleando líneas telefónicas convencionales. Además, la codificación en varios niveles mediante tablas permite realizar codificaciones pseudo-aleatorias evitando acoplamientos entre señales, es decir, facilitando diferenciar unas señales de otras.

Codificación 8B6T

Otro mecanismo multinivel es la codificación 8B6T. Su nombre indica que codifica patrones de $m=8$ elementos de datos en un patrón de $n=6$ elementos de señal. Los dos tipos de elementos de datos son '0' o '1' y por lo tanto se pueden producir $2^m=2^8=256$ patrones de datos. Sin embargo, considerando 3 niveles de tensión, ya que $L=T=3$, (positivo, negativo y cero), es posible codificar $3^n=3^6=729$ patrones de señal distintos. Es decir, cada elemento de señal representa 8 bits, y el conjunto de todos los octetos de bits posibles representan los 256 patrones de datos o niveles de tensión que puede adoptar la señal a transmitir (Figura 3.5) y disponemos de 729 patrones de señal para codificarlos.

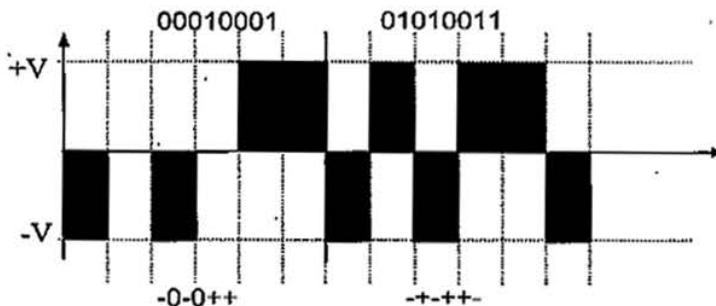


Figura 3-5: Ejemplo de codificación 8B6T de 2 patrones de datos de 8 bits cada uno codificados con 6 elementos de señal.

Este tipo de codificación se emplea en los medios físicos 100Base-4T de redes LAN que se comentarán en el Capítulo 4.

3.2.5. Otras codificaciones

Las técnicas de codificación banda base que se han expuesto en este capítulo no son las únicas existentes. En la actualidad, existen otras muchas técnicas de codificación que aunque no se hayan presentado en detalle si son de especial interés como esquemas de codificación de muchas de las tecnologías actuales y futura, por lo tanto, merecen al menos ser mencionadas. Entre ellas, caben destacar la codificación MLT-3, la codificación 4B5B y la codificación HDB3.

- *MLT-3*: Es una codificación multilínea. A diferencia de las codificaciones NRZ, RZ y Manchester, codifica más de dos niveles de tensión. En concreto codifica tres niveles de tensión, y emplea más de una línea para transmitir los datos. Esta codificación se emplea para enviar a 100Mbps por un cable de cobre convencional.
- *4B5B*: Es una codificación de bloques. A diferencia de las codificaciones 2B1Q y 8B6T comentadas anteriormente, en este caso se cambian 4 bits por bloques de 5 elementos de señal. Por lo tanto, permiten añadir información redundante para asegurar el proceso de sincronización y detección de errores. Existen más variantes de la codificación por bloques, así otras a destacar son 8B10B donde se cambian bloques de 8 bits por bloques de 10 elementos de señal, la 3B4B que cambia bloques de 3 bits por bloques de 4 elementos, y la 5B6B que cambia bloques de 5 bits por 6 elementos de señal.
- *HDB3*: Es una codificación bipolar con sustitución de varios ceros consecutivos por una secuencia determinada de voltajes. Así, en HDB3, si hay que transmitir una secuencia de 4 ceros consecutivos, éstos se sustituyen por 000V o B00V dependiendo del número de pulsos distintos de cero después de la última sustitución, donde V indica mantener la polaridad respecto al último pulso distinto de cero y B indica cambiar la polaridad respecto al último pulso distinto de cero.

3.3. Codificación Banda Modulada

La codificación banda modulada se emplea cuando la información tiene que ser adaptada para poder ser transmitida por un medio físico. Esto ocurre cuando los datos a transmitir y el medio que se emplea en la transmisión no soportan la misma naturaleza de señales. Por ejemplo, los datos tienen naturaleza digital, sin embargo el medio físico sólo soporta señales analógicas y, por lo tanto, se requiere que esos datos digitales sean codificados en forma de señales analógicas (Ej. módem). Para llevar a cabo esta adaptación se emplea lo que se conoce como proceso de modulación. El proceso de modulación emplea tipos de señales distintas que se denominan:

- *Portadora*: Es la señal que se va a transmitir y que debe ser modificada para codificar los datos que se desea enviar.
- *Moduladora*: Es la señal que se emplea para modificar algún parámetro de la portadora, para que esta última pueda ser transmitida. Representa los datos digitales que se quieren hacer llegar al receptor.
- *Modulada*: Es la señal que se obtiene después de modificar la portadora. Esta es la señal que se transmite por el medio físico e incorpora la codificación de los datos digitales.

Existen varias técnicas de codificación banda modulada. Éstas se catalogan según la naturaleza digital o analógica de los tipos de señal portadora y moduladora. Por lo tanto, si analizamos los tipos de situación en función de la naturaleza de las señales portadora y moduladora se tienen:

Señal Moduladora	Señal Portadora	Señal Modulada	Codificación
Analógica	Analógica	Analógica	AM, FM, PM
Digital	Analógica	Analógica	ASK, FSK, PSK
Analógica	Digital	Digital	PCM y variantes
Digital	Digital	Digital	Banda Base

Una posible clasificación de las técnicas de codificación, para enviar datos digitales (señal moduladora digital), en función de la naturaleza del medio y por lo tanto de la señal portadora que se va a transmitir por él, es:

- *Modulación analógica*: Cuando se quiere transmitir por un medio analógico en cuyo caso se requiere de una señal portadora analógica.
- *Modulación digital*: Cuando se quiere transmitir por un medio digital en cuyo caso se requiere de una señal portadora digital.

3.3.1. Modulación analógica

La modulación analógica se puede catalogar de acuerdo al parámetro de la señal portadora que se modifica para codificar los datos de la señal moduladora.

Así, si la señal portadora es analógica, ésta se describe a partir de tres parámetros básicos: frecuencia, amplitud y fase (ver Apartado 2.2). Por lo tanto, si el parámetro que se modifica es la amplitud la modulación se conoce como *ASK* o desplazamiento en amplitud. Si el parámetro que se modifica es la frecuencia, entonces la modulación se denomina *FSK* o desplazamiento en frecuencia. Y finalmente, si el parámetro es la fase entonces se denomina modulación *PSK* o desplazamiento en fase. A continuación se mostrará con cierto detalle como se realiza cada una de estas modulaciones.

Modulación ASK

La modulación ASK (Amplitude Shift Keying) se caracteriza por ser una modulación analógica en amplitud, donde la señal portadora modifica su amplitud en función del valor de la amplitud en la señal moduladora. Así, finalmente, los valores binarios de la señal moduladora estarán representados en la señal modulada por dos amplitudes diferentes de la señal portadora (Figura 3.6).

La señal moduladora que contiene los datos digitales, matemáticamente, se puede representar como una señal compuesta periódica cualquiera. Señal que se puede descomponer como una serie infinita de funciones simples (senos y cosenos) de diferente amplitud y frecuencia.

$$u_m(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos(2\pi n f_0 t) + \sum_{n=1}^{\infty} b_n \sin(2\pi n f_0 t) \quad (3.1)$$

Además, la señal portadora que el medio físico soporta para transmitir se representa como una señal senoidal cualquiera de amplitud máxima un valor de tensión V.

$$u_p(t) = V \sin(2\pi f_p t + \phi) \quad (3.2)$$

Por lo tanto, la señal modulada se puede calcular a partir de una composición de las señales de las Ecuaciones (3.1) y (3.2), como se indica en la Ecuación (3.3).

$$u_M(t) = u_p(t) \circ u_m(t) = \begin{cases} V_1 \sin(2\pi f_p t + \varphi) & \text{si } u_m = 0 \\ V_2 \sin(2\pi f_p t + \varphi) & \text{si } u_m = 1 \end{cases} \quad (3.3)$$

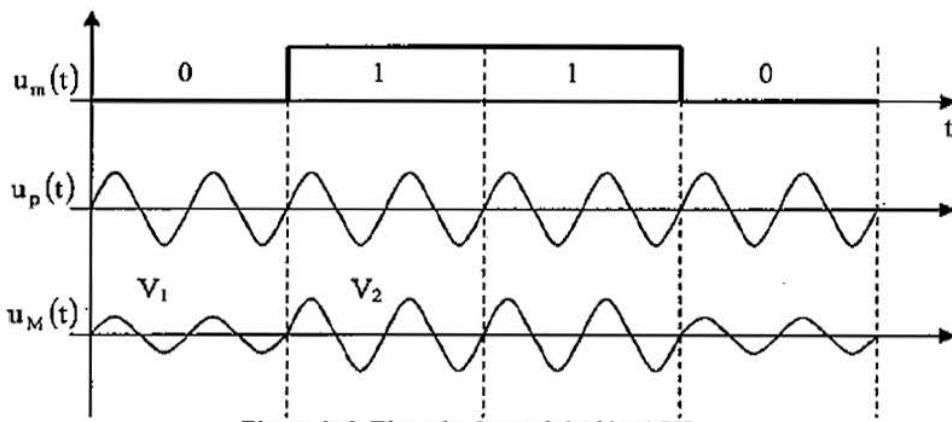


Figura 3-6: Ejemplo de modulación ASK.

Modulación FSK

La modulación FSK (Frequency Shift Keying) se caracteriza por ser una modulación analógica en frecuencias, donde se emplean dos señales portadoras con frecuencias distintas. De este modo, los valores binarios de la señal moduladora estarán representados en la señal modulada por dos frecuencias diferentes, cada una de ellas correspondientes a cada una de las dos señales portadoras (Figura 3.7).

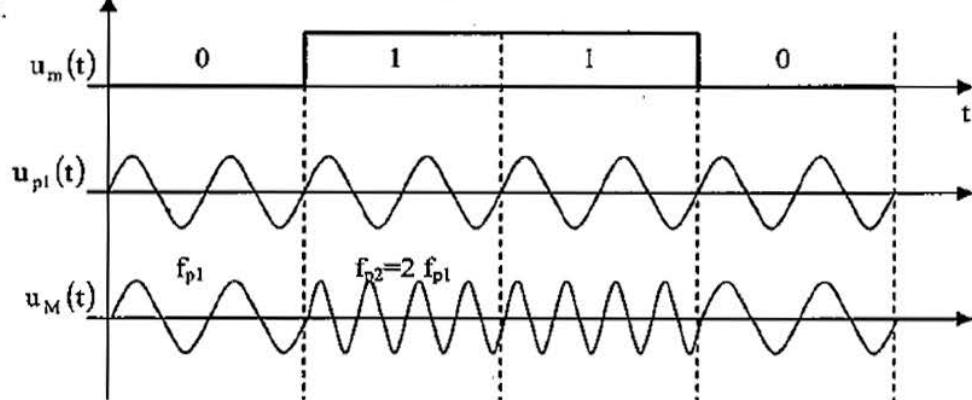


Figura 3-7: Ejemplo de modulación FSK.

Para este tipo de modulación, la señal moduladora se vuelve a representar por la Ecuación (3.1). Sin embargo, se emplean dos señales portadoras de igual fase y amplitud, pero con distinta frecuencia, como se indica en las Ecuaciones (3.4) y (3.5). En concreto en el ejemplo de la Figura 3.7 se ha empleado una frecuencia $f_{p2} = 2(f_{p1})$.

$$u_{p1}(t) = V \sin(2\pi f_{p1} t + \varphi) \text{ si '0' ó } u_{p1}(t) = 0 \text{ si '1'} \quad (3.4)$$

$$u_{p2}(t) = V \sin(2\pi f_{p2} t + \varphi) \text{ si '1' ó } u_{p2}(t) = 0 \text{ si '0'} \quad (3.5)$$

En este caso, la señal modulada se puede calcular a partir de una composición de las señales de las Ecuaciones (3.1), (3.4) y (3.5), como se indica en la Ecuación (3.6).

$$u_M(t) = \begin{cases} V \sin(2\pi f_{p1} t + \varphi) & \text{si '0'} \\ V \sin(2\pi f_{p2} t + \varphi) & \text{si '1'} \end{cases} \quad (3.6)$$

Modulación PSK

La modulación PSK (Phase Shift Keying) se caracteriza por ser una modulación analógica en fase, donde se emplea una portadora cuyo parámetro de fase se modifica en función del valor de la señal moduladora que codifica '0' y '1'. De este modo, los valores binarios de la señal moduladora estarán representados en la señal modulada por dos fases diferentes de la señal portadora empleada en el proceso. Por ejemplo en la Figura 3.8, la señal portadora codifica un '1' modificando su fase 180° y codifica '0' manteniendo su fase original en 0° .

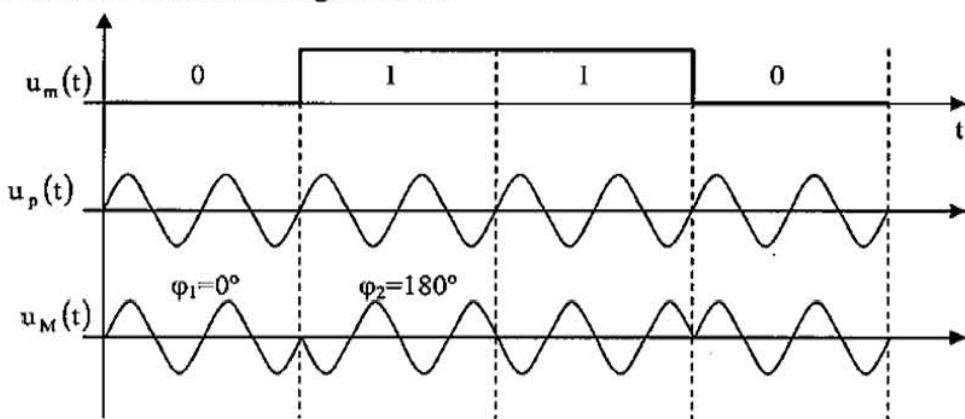


Figura 3-8: Ejemplo de modulación PSK.

En la modulación PSK, la señal moduladora de nuevo se representa por la Ecuación (3.1). La señal portadora es una señal senoidal, cuya representación matemática ya se indicó en la ecuación (3.2). No obstante, la señal modulada (3.7) es la señal portadora adelantada o atrasada en fase dependiendo de los datos que modela la señal moduladora. Para el ejemplo de la Figura 3.8, $\varphi_1 = 0^\circ$ y $\varphi_2 = \pi \text{ rad} = 180^\circ$.

$$u_M(t) = \begin{cases} V \sin(2\pi f_p t + \varphi_1) & \text{'0'} \\ V \sin(2\pi f_p t + \varphi_2) & \text{'1'} \end{cases} \quad (3.7)$$

Modulación QPSK

La modulación QPSK (Quadrature Phase Shift Keying) es una variante de la modulación PSK. También se conoce con el nombre de PSK con desplazamiento de fase en cuadratura ó 4-PSK.

En este tipo de modulación, en lugar de emplear un desplazamiento en fase de 180° se emplean desplazamientos de fase de la señal portadora correspondientes a múltiplos de 90° . Así, la señal modulada adquiere la representación matemática de la Ecuación (3.8).

$$u_M(t) = \begin{cases} V \sin(2\pi f_p t + 0) & '00' \\ V \sin(2\pi f_p t + \pi/2) & '01' \\ V \sin(2\pi f_p t + \pi) & '11' \\ V \sin(2\pi f_p t + 3\pi/2) & '10' \end{cases} \quad (3.8)$$

Este tipo de modulación permite que cada elemento de señal represente 2 bits. Este tipo de señalización permite aumentar la velocidad de modulación en el proceso de transmisión.

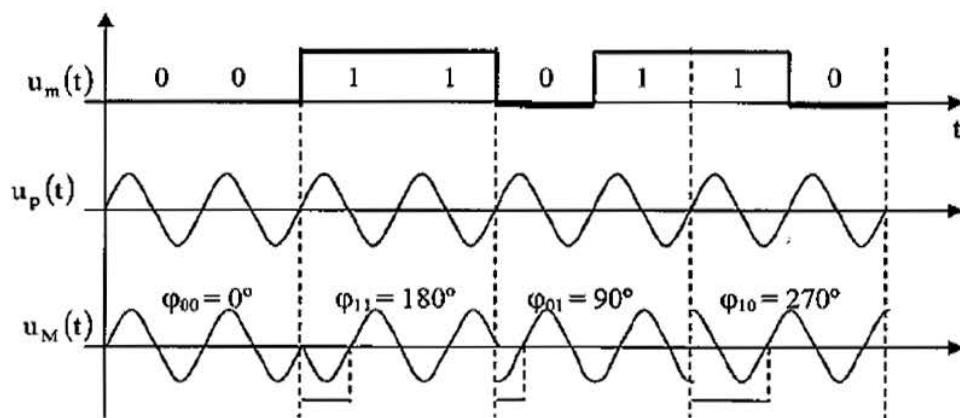


Figura 3-9: Ejemplo de modulación QPSK.

Modulación QAM

La modulación *QAM* (Quadrature Amplitude Modulation) es una variante de la modulación ASK y PSK. En este tipo de modulación se emplean distintas amplitudes y desplazamientos de fase de la señal portadora.

Existen numerosas variaciones posibles en la modulación *QAM* en función del número de niveles de señalización que se emplean. De ahí que se haya diseñado una herramienta gráfica, conocida como '*Diagrama de Fase*', que permite relacionar niveles de señalización con número de bits que se pueden enviar por cada nivel de señalización (Figura 3.10).

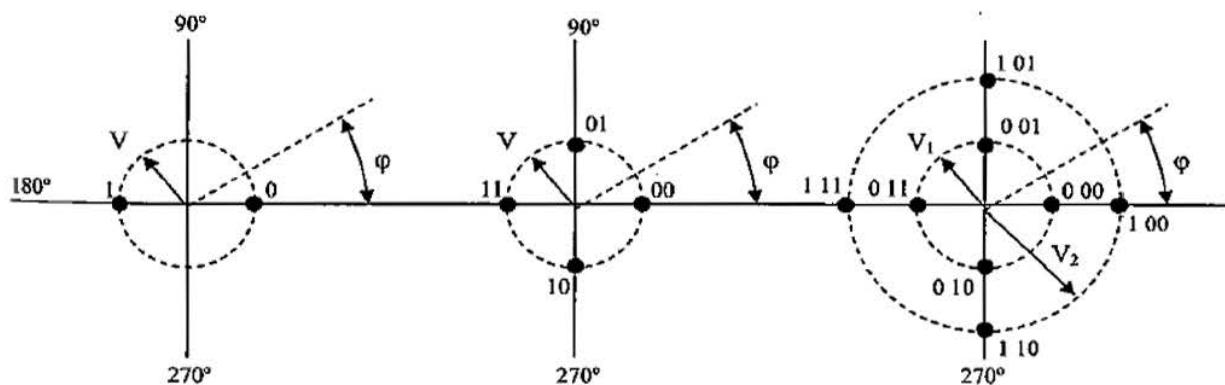


Figura 3-10: Diagramas de Fase.

Otras modulaciones

Además, a parte de las modulaciones anteriormente mencionadas, existen modulaciones que se pueden implementar como modificaciones de las anteriores. Así, se conoce la modulación *MFSK* como una modulación multínivel de la modulación *FSK*. Es decir, la modulación *MFSK* emplea más de dos señales portadoras. Por ejemplo, cuatro señales portadoras todas ellas iguales pero con frecuencias distintas. Con ello se consigue, al igual que en *QPSK*, enviar 2 bits al mismo tiempo. Por lo tanto, del mismo modo, si se desean enviar 3 bits, se necesitarían ocho señales portadoras de frecuencias distintas, y así sucesivamente. No obstante, es importante constatar que se requiere que las frecuencias estén lo suficientemente separadas entre sí para que los sistemas moduladores y demoduladores funcionen adecuadamente.

3.3.2. Modulación digital

Si anteriormente, la modulación analógica se ha considerado como aquella que convierte datos digitales a señales analógicas, la modulación digital consiste en convertir señales analógicas a datos digitales. Por lo general, cualquier dato analógico está representado directamente por una señal analógica, de ahí que la modulación digital, por lo tanto, consista en convertir señales analógicas a datos digitales. Como representación de datos analógicos mediante señales analógicas se tiene la información creada por un micrófono a través de un equipo de sonido, la creada por una cámara de televisión, etc. En algunos ámbitos se conoce a este proceso de modulación digital como proceso de digitalización de información.

Entre las técnicas de modulación digital destacan dos: Modulación por codificación de pulsos o *PCM* y Modulación por codificación diferencial de pulsos o *PCM diferencial*. Además, destaca la modulación Delta o *DM* como una técnica de *PCM diferencial*.

Una vez que las señales analógicas se han convertido en datos digitales, es posible emplear cualquiera de las técnicas de codificación Banda Base (ver Apartado 3.2) para transformar esos datos en señales digitales.

Modulación *PCM*

La técnica *PCM* (Pulse Code Modulation) permite transformar una señal analógica en una señal digital a partir de tres procesos: muestreo, cuantificación y codificación.

- *Muestreo*: Proceso que permite adquirir un conjunto de muestras de la señal analógica en ciertos instantes de tiempo. La señal se muestrea con una frecuencia de muestreo igual a dos veces el ancho de banda del medio, de acuerdo al teorema de Nyquist.
- *Cuantificación*: La señal analógica muestreada es un conjunto de pulsos con valores de amplitud comprendidos entre la mínima y la máxima amplitud de la señal. El proceso de cuantificación consiste en dividir el rango de valores de amplitud en un conjunto discreto de niveles q , tomando un valor fijo de incremento. Posteriormente, se asigna un nivel concreto a cada valor de amplitud.
- *Codificación*: Proceso que permite asignar a cada muestra cuantificada una secuencia de bits que representan el nivel de cuantificación. Así, a partir del número de niveles q que cuantifican una señal muestreada, se puede calcular el número de bits necesarios para codificar el valor de cuantificación, como $q=2^n$, siendo n el número de bits empleados para la codificación.

En la práctica, la cuantificación y codificación se hacen conjuntamente mediante un dispositivo denominado *ADC* (Analog to Digital Converter) o conversor de analógico a digital.

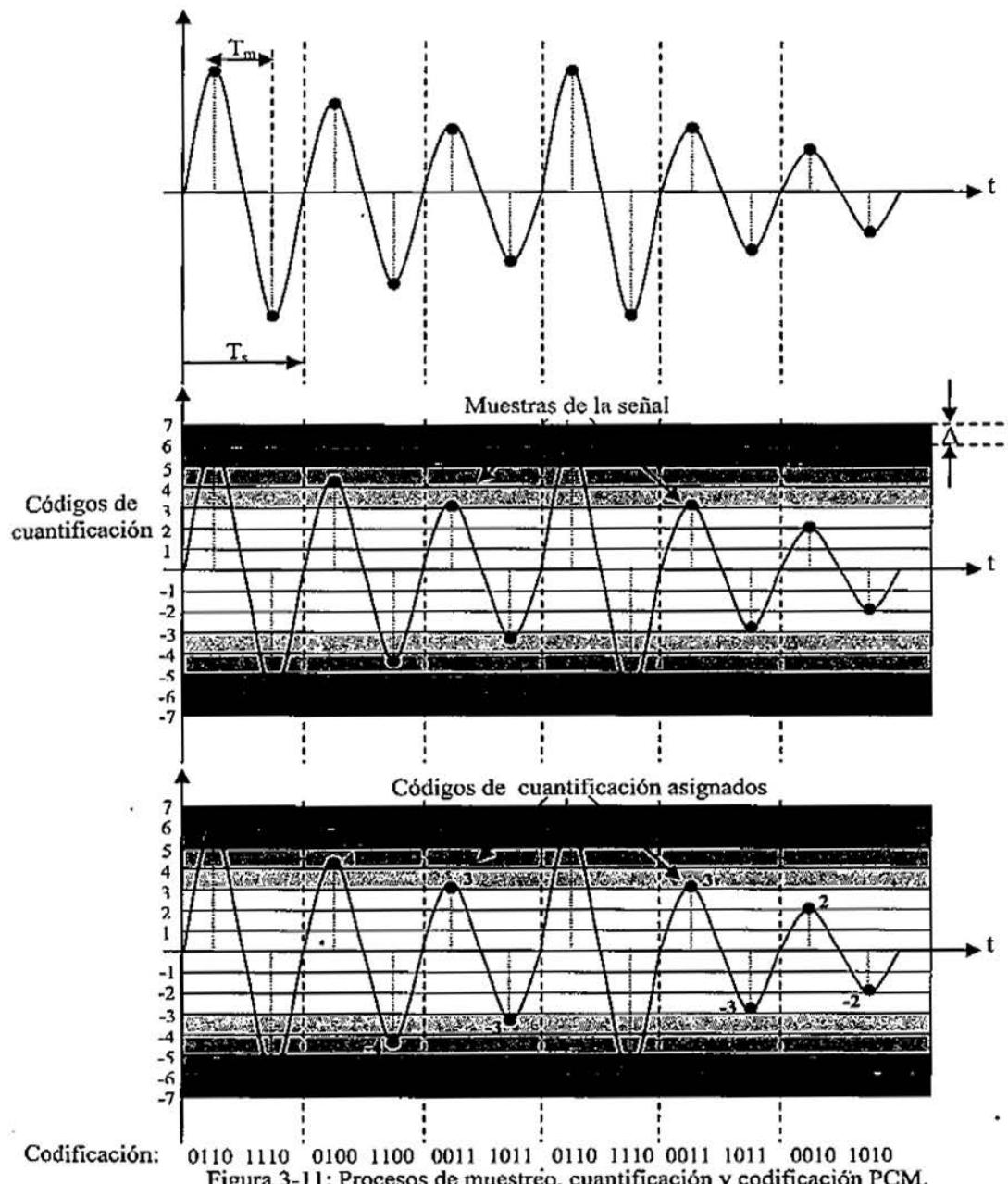


Figura 3-11: Procesos de muestreo, cuantificación y codificación PCM.

En la figura 3.11, se ha realizado un proceso de modulación PCM. Inicialmente, se ha sometido a una señal analógica a un proceso de muestreo ideal, tomando muestras cada T_m segundos, donde T_m se conoce como periodo de muestreo. Al proceso de muestreo también se le conoce como PAM o modulación por amplitud de pulsos. En este proceso se ha obtenido un conjunto de pulsos con valores de amplitud comprendidos entre la mínima y máxima amplitud de la señal. Durante el proceso de muestreo, las muestras obtenidas pertenecen a un rango infinito de números reales. Posteriormente, se ha aplicado un proceso de cuantificación, en el cual se le asigna a cada valor de muestreo o PAM un código de cuantificación. Para ello, el primer paso es dividir el rango de amplitudes en zonas de una anchura Δ , y a continuación en un segundo paso se asignan valores cuantificados de acuerdo a la zona a la que más se aproxime el valor de muestreo o PAM. La elección del número de niveles y de la anchura Δ , dependen de la precisión con la que se desee recuperar la señal analógica original.

En el caso mostrado en la figura 3.11 se ha dividido el rango de amplitudes en $q=15$ posibles códigos de cuantificación comprendidos entre $[-7, 7]$. Y a cada una de las 12 muestras se les ha asignado un valor cuantificado dentro de ese intervalo, en función del valor de cuantificación más próximo. En concreto, los valores asignados son 6, -6, 4, -4, 3, -3, 6, -6, 3, -3, 2 y -2.

Todo proceso de cuantificación tiene errores. La cuantificación es un proceso de aproximación, puesto que los valores de amplitud de las muestras del proceso PAM pueden no ser idénticos a los niveles establecidos y es necesario redondearlos, aproximándolos al valor de cuantificación más cercano. Un ejemplo de este error se observa en la cuarta muestra de la figura 3.11, donde el valor PAM es algo menor que -4, y sin embargo se aproxima por el valor -4 por ser el más cercano, con el consiguiente error de redondeo. No obstante, en la misma figura, la quinta muestra tiene error de cuantificación cero. Además es fácil comprobar que el error de cuantificación depende del número de niveles. Los errores de cuantificación se añaden como una señal ruido a la señal que se genera. Esta señal de ruido se conoce con el nombre de ruido de cuantificación del ADC.

Por último, el proceso de codificación supone la última etapa de la técnica de Modulación por Codificación de Pulses (PCM). En esta etapa, cada una de las muestras cuantificadas se codifica con n bits. El número de bits para codificar cada muestra depende del número de niveles de cuantificación como ya se había comentado previamente, así como de la forma de codificar valores negativos. Para el caso de la figura 3.11, y de acuerdo a la fórmula $q=2^n$, se tiene que $15=2^4$ y por lo tanto $n=4$. Una vez determinado el número de bits, es necesario asociar a todos los códigos de cuantificación un código binario. En el ejemplo mostrado en la figura 3.11, los 15 niveles comprendidos entre $[-7, 7]$ se codifican con los 4 bits calculados. Una manera de proceder típica es asignar el código binario correspondiente al valor decimal del nivel de cuantificación asignado a la muestra. Así, la primera muestra tiene asignado el valor +6, por lo tanto, se codifica con su valor binario de 4 bits, 0110, y se emplea el complemento a dos para los valores negativos. No obstante, es posible asignar otros códigos preestablecidos para codificar cada nivel de cuantificación.

En la práctica los ADC que emplean codificación PCM emplean un número determinado de bits y dependiendo del número de bits empleado fijan el número de niveles de cuantificación. Por supuesto, a mayor número de bits para representar la señal, mayor será el número de niveles de cuantificación y más nivel de detalle tendrá la señal digital resultante.

Ejercicio 3.1

Un dispositivo de red envía una señal analógica como la que se muestra en la figura 3.12. Se quiere transformar mediante la técnica PCM para enviarla a través de un medio físico digital. Realiza el proceso de conversión considerando que la señal analógica posee una frecuencia inferior de 0 Hz, una frecuencia superior de 100Hz y que se propaga en 0.05s y contesta a las siguientes cuestiones:

- Determina la frecuencia de muestreo del proceso PAM. ¿Cuántas muestras se extraerán de la señal exactamente?
- Atendiendo a todos los valores posibles de amplitud, determina el número de bits necesarios para la cuantificación de las muestras, si sólo se pueden representar valores enteros.
- Determina la secuencia de bits que hay que enviar si se codifica la señal mediante PCM, asignando un código binario a cada uno de los valores de amplitud.

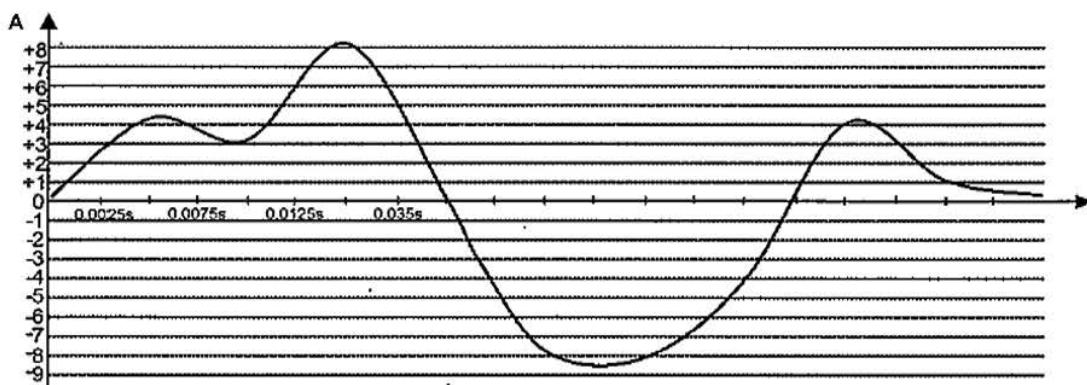


Figura 3-12: Señal analógica.

- a) $B=100\text{Hz}-0\text{Hz}=100\text{Hz}$, $f=2B=200\text{Hz} \rightarrow T=1/f=1/200=0,005\text{s}$ y por lo tanto una muestra cada $0,005\text{ segundos} \rightarrow 0,05/0,005=10$ muestras
- b) Hay que codificar 18 valores entre -9 y +8. Se necesitarán, por lo tanto, 5 bits para codificar los 18 valores posibles de tensión, ya que el número de bits viene dado por $n=\log_2 N$ donde $N=18$. Así, $n=\log_2 18=4.17$ bits. Y puesto que con 4 bits sólo se podrían codificar 16 valores, hay que optar por el valor entero inmediatamente superior a 4.17 que es 5.

- c) Un posible código binario para cada uno de los valores de tensión podría ser:

8	01000	-1	10001
7	00111	-2	10010
6	00101	-3	10011
5	00101	-4	10100
4	00100	-5	10101
3	00011	-6	10110
2	00010	-7	10111
1	00001	-8	11000
0	00000	-9	11001

Los valores de tensión a codificar para el periodo de muestreo, son: +4, +3, +8, 0, -8, -8, -4, +4, +1

La codificación sería: 00100 00011 01000 00000 11000 11000 10100 00100 00001

Modulación PCM diferencial

La técnica PCM diferencial es una variación de la técnica PCM convencional. Esta técnica también permite transformar una señal analógica en una señal digital a partir de tres procesos: muestreo, cuantificación y codificación. La diferencia respecto a la técnica PCM reside en el proceso de codificación de la señal muestreada. Y la ventaja frente a PCM es que suele requerir de menor número de bits para codificar la señal. Sin embargo, también tiene algunas desventajas como la necesidad de memoria temporal para almacenar el estado anterior empleado en la codificación.

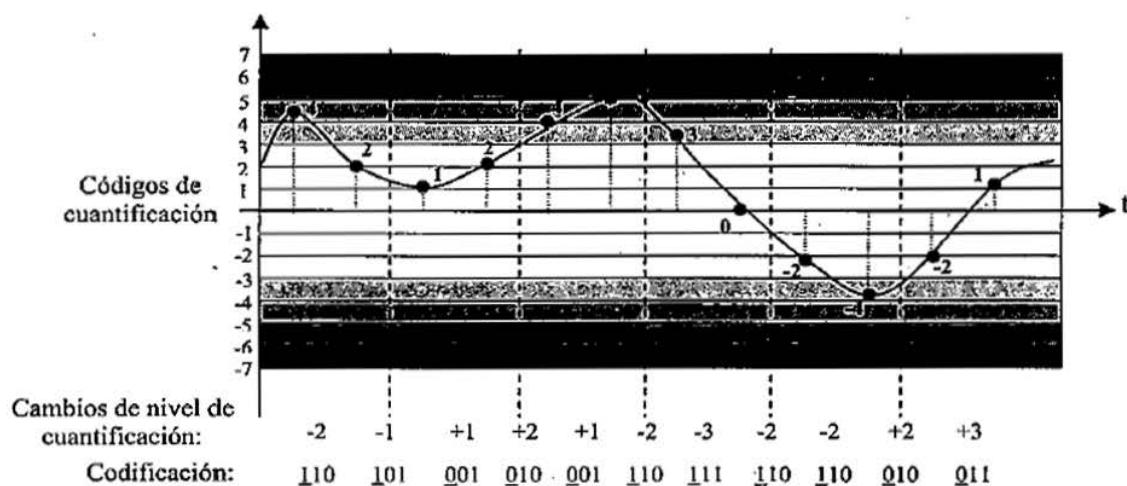


Figura 3-13: Proceso de codificación PCM diferencial.

En el proceso de codificación, la técnica PCM diferencial no codifica la muestra en función del nivel de cuantificación más próximo como ocurría en PCM, sino que codifica el incremento o decremento entre dos muestras cuantificadas consecutivas. Así, se codificará con el bit '1' cuando el cambio entre dichos valores de cuantificación sea de decrecimiento y se codificará con el bit '0' cuando el cambio entre esos valores de cuantificación sea de crecimiento. Y a ese bit de signo, que determina si hay crecimiento o decrecimiento, se le añaden tantos bits como sean necesarios para codificar en binario el valor absoluto del número de niveles de cuantificación en el que difieren dos muestras consecutivas. Así, observando la figura 3.13, se comprueba que entre la primera muestra +4 y la segunda muestra +2, se ha producido un decrecimiento de 2 niveles de cuantificación. Por lo tanto, la codificación en este caso es 110 donde el primer bit '1' indica decrecimiento y los otros dos bits '10' indican el número de niveles que decrece la señal.

El principal inconveniente de la técnica PCM diferencial radica en que imposibilita la codificación de señales constantes, ya que en este tipo de señales no habría crecimientos ni decrecimientos de señal para ser codificados. Además, la codificación de cambios bruscos en la señal requiere de un elevado número de bits. Sin embargo, disfruta de varias ventajas frente a la técnica PCM, entre las que destacan dos:

- PCM diferencial permite aumentar la velocidad de transmisión en algunos casos, puesto que generalmente disminuye el número de bits a transmitir.
- PCM diferencial permite emplear medios físicos que soportan un número limitado de bits para el proceso de codificación de niveles de cuantificación.

Ejercicio 3.2

Un dispositivo inalámbrico envía una señal analógica como la que se muestra en la figura 3.14. Se quiere transformar mediante la técnica PCM diferencial para enviarla a través de un medio físico digital. Realiza el proceso de conversión considerando que la señal analógica posee una frecuencia

inferior de 0 Hz, una frecuencia superior de 125Hz y que se propaga en 0.064s y contesta a las siguientes cuestiones:

- Determina la frecuencia de muestreo del proceso PAM. ¿Cuántas muestras se extraerán de la señal exactamente?
- Determina la secuencia de bits que hay que enviar si se codifica la señal mediante PCM diferencial.
- ¿Es posible ahorrar bits a transmitir codificando la señal de la figura 3.14 con esta técnica con respecto al haber usado la PCM? Si es posible calcula cuántos bits se han ahorrado en la transmisión de la señal, y sino es posible justifica la razón de qué sea así.

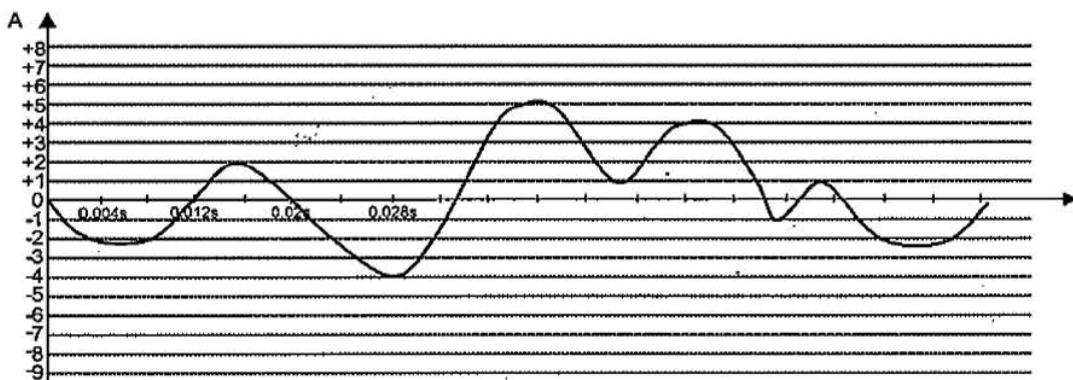


Figura 3-14: Señal Analógica.

- $B=125\text{Hz}-0\text{Hz}=125\text{Hz}$, $f=2B=250\text{Hz} \rightarrow T=1/f=1/250=0,004\text{s}$ y por lo tanto una muestra cada 0,004 segundos. $\rightarrow 0,064/0,004=16$ muestras
- Incrementos a codificar: -2,0,+2,+2,-2,-2,+3,+4,+2,-2,-2,+2,-1,-4,+1. Se necesitarán, por lo tanto, 3 bits para indicar el incremento o decremento, y 1 bit adicional para indicar si es incremento o decremento, en total 4 bits. La codificación sería: 1010 0000 0010 0010 1010 1010 1010 0011 0100 0010 1010 1010 0011 1001 1100 0001
- Con PCM diferencial se han empleado 4 bits, con PCM hubieran sido necesarios bits para codificar los 10 niveles posibles desde -4 a 5. Y para codificar 10 niveles son necesarios 4 bits. Por lo tanto, en este caso no hay un ahorro de bits a transmitir.

Modulación Delta

La técnica Modulación Delta o DM es un caso particular de la técnica de modulación PCM diferencial. Esta técnica además de reducir la complejidad respecto a la PCM, reduce la información a transmitir y precisa un ancho de banda más pequeño para realizar la transmisión por un medio físico.

En el proceso de codificación, la técnica DM codifica únicamente los cambios positivos o negativos entre dos muestras cuantificadas consecutivas. Es decir, cuando entre dos muestras consecutivas se produzca un incremento en los valores de cuantificación, se codificará con el bit '1' y cuando el cambio entre dichos valores de cuantificación sea un decremento se codificará con el bit '0'. En esta técnica no es necesario codificar cuántos niveles de variación supone ese incremento o decremento, sino únicamente el hecho de que haya variación independientemente de cuál sea su valor.

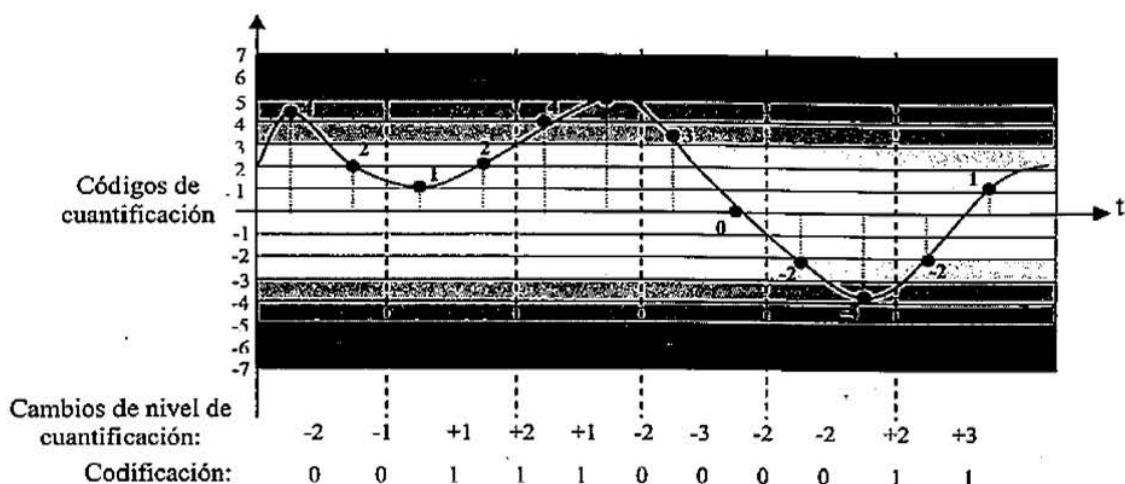


Figura 3-15: Proceso de codificación DM.

Para llevar a cabo una modulación DM en la práctica (Figura 3.16), el sistema modulador construye una segunda señal denominada señal escalera, y compara la señal analógica a transmitir con esta señal escalera. Así, el modulador para cada muestra es capaz de determinar los incrementos o decrementos comparando el valor de la señal analógica con el último valor de la señal escalera. Cuando la amplitud de la señal analógica es mayor que la de escalera el siguiente bit será un '1' y cuando sea menor será un '0'.

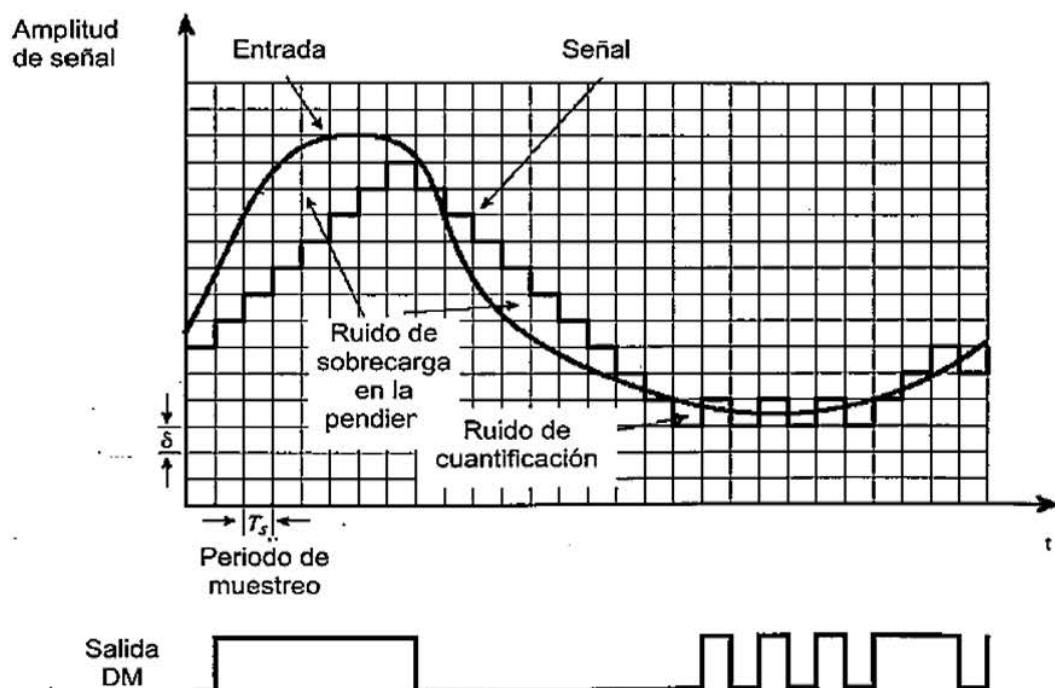


Figura 3-16: Proceso real de codificación DM usando señal escalera (Cortesía [Stallings-2004]).

3.4. Multiplexación

Como ya se comentó en el Capítulo 2, el ancho de banda determina las prestaciones de un medio físico. Además, las especificaciones de un medio físico se rigen por ancho de banda limitado, además de por una velocidad máxima de transmisión, entre otros muchos factores.

En muchas ocasiones, los dispositivos que se conectan a un medio físico requieren de menor ancho de banda del máximo que ofrece el medio en cuestión. En estos casos, el medio físico puede ser compartido para realizar transmisiones simultáneamente por varios dispositivos. Por lo tanto, se define *multiplexación* como el conjunto de técnicas que permiten transmitir varias señales a la vez a través de un mismo medio físico que forma parte de un único enlace de datos. El objetivo de la multiplexación consiste en realizar un empleo eficiente de medios físicos con gran ancho de banda, permitiendo que éstos transporten señales de distintos dispositivos a la vez. La multiplexación, por lo tanto, es importante para que medios físicos con gran ancho de banda como son el cable coaxial, la fibra óptica o las microondas terrestres puedan ser empleados con eficiencia para transmitir varias señales simultáneamente. De este modo, se evita que se malgaste el uso de medios físicos, con gran ancho de banda, con transmisiones que no requieren tanto ancho de banda.

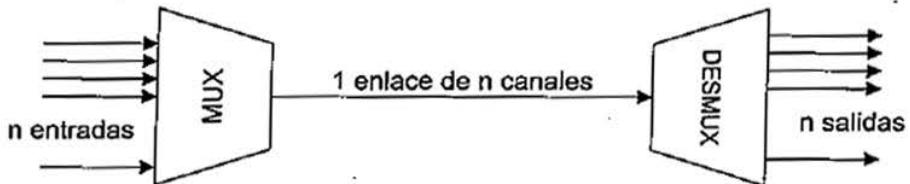


Figura 3-17: Proceso de multiplexación y desmultiplexación.

Así, para hacer un uso eficiente de los enlaces de comunicaciones se emplean técnicas de multiplexación. Existen cuatro tipos de técnicas básicas de multiplexación:

- *Multiplexación por división en frecuencias (FDM: Frequency Division Multiplexing)*: Es una técnica de multiplexación analógica que consiste en transmitir varias señales por un mismo medio, asignando una banda de frecuencias diferente para cada señal. Se emplea cuando el rango de frecuencias del ancho de banda útil del medio supera el ancho de banda requerido por cada señal que se desea transmitir.
- *Multiplexación por división en tiempos (TDM: Time Division Multiplexing)*: Es una técnica de multiplexación digital que consiste en transmitir datos digitales procedentes de distintos dispositivos por un mismo medio. En TDM se divide el enlace de datos en porciones de tiempo, y se asigna una porción de tiempo a cada dispositivo durante la cual puede transmitir datos.
- *Multiplexación por longitud de onda (WDM: Wave Division Multiplexing)*: Es una técnica de multiplexación analógica que consiste en transmitir varias señales luminosas sobre un mismo enlace de fibra óptica. La técnica es conceptualmente similar a la FDM, consiste en combinar señales con frecuencias distintas. La diferencia es que en WDM las frecuencias de las señales son muy altas comparadas con las que se emplean en FDM.
- *Multiplexación por división de código (CDMA: Code Division Multiple Access)*: Es una técnica de multiplexación que consiste en transmitir varias señales por un mismo medio compartiendo frecuencias y tiempos, pero usando codificaciones distintas para cada señal. Las codificaciones se llevan a cabo empleando patrones de bits válidos o códigos de línea. Este tipo de multiplexación se emplea en transmisiones inalámbricas (Wifi 802.11n, GSM, 3G, etc.).

3.4.1. Multiplexación por división en frecuencias

La multiplexación por división del medio en frecuencias se aplica cuando el ancho de banda útil de un enlace de datos es mayor que los anchos de banda combinados de las señales que se desea transmitir. Esta técnica permite transmitir varias señales simultáneamente si cada una de ellas se modula con señales portadoras de frecuencias distintas. Además, se requiere que las frecuencias de las señales portadoras se encuentren separadas entre sí para evitar solapamientos e interferencias. Posteriormente, las señales moduladas se combinan formando una señal compuesta que se transporta por el enlace de datos.

Todo sistema de multiplexación requiere de un sistema de desmultiplexación que separe la señal compuesta formada en el proceso de multiplexación en un conjunto de señales distintas (Figura 3.17). El proceso de multiplexación se lleva a cabo en el dispositivo emisor (Figura 3.18) y el proceso de desmultiplexación se realiza en el dispositivo receptor (Figura 3.19).

Proceso de Multiplexación

Observando la Figura 3.18a se muestra un esquema general de multiplexación por división en frecuencias. Así, varias señales $u_m(t)$ se quieren enviar a través de un mismo enlace de datos. Para ello, cada señal $u_m(t)$ se modula mediante una señal portadora $u_p(t)$ y posteriormente, cada señal modulada $u_M(t)$ resultante se suma para dar lugar a una señal modulada de banda compuesta, $u_{MB}(t)$.

Como se observa en la Figura 3.18b, el espectro de la señal modulada en banda compuesta, $u_{MB}(t)$, está representado por la suma de los espectros de las señales $u_{mi}(t)$ desplazados hasta quedar centrados en cada una de las frecuencias de las señales portadoras $u_{pi}(t)$. Y las frecuencias f_{pi} de las señales portadoras $u_{pi}(t)$ se han escogido de modo que los anchos de bandas de las distintas señales, B_i , no se solapen entre sí.

En el esquema, se puede observar como la suma de los anchos de banda de cada una de las distintas señales, B_i , es inferior al ancho de banda del medio físico, B .

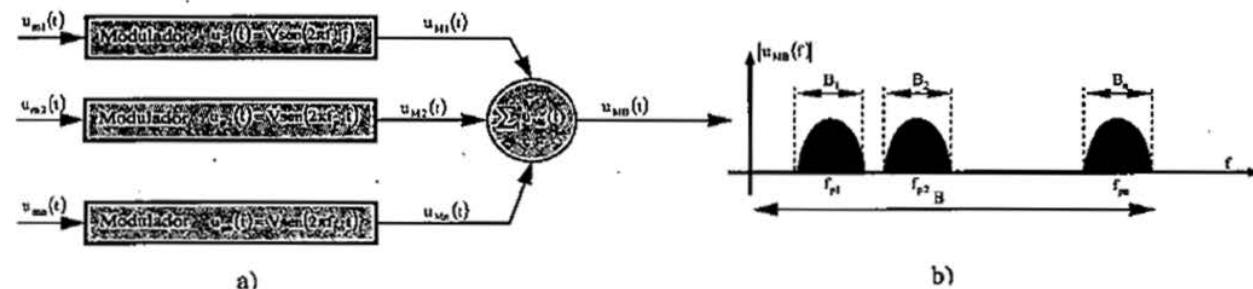


Figura 3-18: Multiplexación FDM. a) Modulación. b) Señales multiplexadas

Proceso de Desmultiplexación

El proceso desmultiplexador emplea un conjunto de filtros paso banda que permiten descomponer la señal multiplexada, es decir, la señal en banda base compuesta, en un conjunto de señales simples. Una vez los filtros han permitido obtener las señales moduladas individuales procedentes de diferentes dispositivos, éstas se hacen pasar por un sistema demodulador que separa las señales moduladoras de sus portadoras (Figura 3.19).

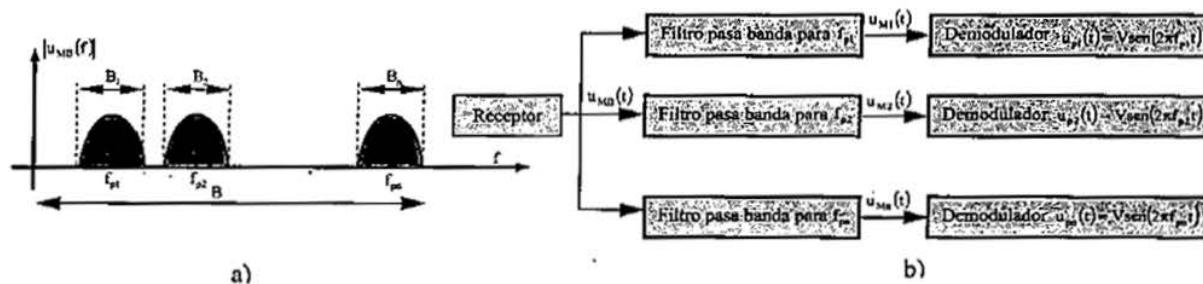


Figura 3-19: Desmultiplexación FDM.

Ejercicio 3.3

Se quiere transmitir información de diversas fuentes F_1 , F_2 y F_3 por un mismo medio físico de transmisión basado en tecnología analógica. Las señales que transmite cada uno son: F_1 envía datos a 20Khz, F_2 a 4Khz y F_3 a 1Mhz. Si el medio de transmisión sólo soporta transmisión analógica y hace uso de la técnica FDM para la transmisión de varias fuentes, entonces:

- Se pide calcular el ancho de banda, en modo simplex, que tiene que soportar el medio de transmisión para las tres comunicaciones F_1 , F_2 y F_3 .
- Se pide determinar el ancho de banda necesario para que las comunicaciones sean full-duplex, implementando canales de subida y de bajada y conociendo que los canales de bajada ocupan $\frac{3}{4}$ del ancho de banda del medio.
- Determina el ancho de banda de cada uno de los canales (full-duplex: subida y bajada calculados en el apartado b), asignados a las fuentes para que no haya ruido de intermodulación.
- Calcula la frecuencia de las señales portadoras necesarias para multiplexar cada una de las comunicaciones full-duplex.

$$a) B_{\text{medio}} = B_{F_1} + B_{F_2} + B_{F_3} = 20\text{Khz} + 4\text{Khz} + 1000\text{Khz} = 1024\text{Khz}$$

b)

$$B_{\text{medio}} = B_{\text{subida}} + B_{\text{bajada}}$$

$$B_{\text{bajada}} = (3/4)B_{\text{medio}} = 3 \cdot 1024\text{Khz} / 4 = 768\text{Khz}$$

$$B_{\text{subida}} = B_{\text{medio}} - B_{\text{bajada}} = 1024\text{Khz} - 768\text{Khz} = 256\text{Khz}$$

- El canal de subida se tiene que dividir en 3 subcanales, uno para cada fuente. Del mismo modo, el canal de bajada se tiene que dividir también, 3 subcanales. Además, siempre que $\Delta B \geq 0$ no habrá ruido de intermodulación. Por lo tanto:

$$B_{\text{medio}} = B_{\text{subida}} + B_{\text{bajada}}$$

$$B_{\text{bajada}} = (B + \Delta B) \cdot n = 768\text{Khz} \rightarrow n = 768\text{Khz} / 3 = 256\text{Khz}$$

$$B_{\text{subida}} = (B + \Delta B) \cdot n = 256\text{Khz} \rightarrow n = 256\text{Khz} / 3 = 85.3\text{Khz}$$

- Las señales portadoras para los canales de bajada estarán centradas en: 128Khz, 384Khz y 640Khz. Y para los canales de subida estarán centradas en 42.66Khz, 128Khz y 213.33Khz.

3.4.2. Multiplexación por división en tiempo

La multiplexación por división del medio en tiempos se aplica cuando la velocidad de transmisión máxima que soporta el medio físico es superior a la velocidad de transmisión de la suma de todas las señales digitales que se desea transmitir. Este tipo de multiplexación es adecuado para un procesamiento digital de señales.

Existen dos tipos básicos de multiplexación por división de tiempos: *TDM sincrónico* y *TDM estadístico*. A continuación se presentan ambos tipos.

Multiplexación TDM sincrona

Esta técnica permite transmitir varias señales simultáneamente si las distintas partes de cada una de ellas se entremezclan. Este proceso de mezcla se puede realizar empleando partes de cada señal de un bit, de bloques de octetos u otras cantidades de bits superiores. Al número de bits que constituyen estos bloques o porciones de señal se les llama *ranuras temporales*. Así, cada trama que se enviará por el medio consta de un conjunto de ranuras temporales, y se asigna a cada dispositivo emisor una o más ranuras temporales por trama. El resultado es una trama de bits formada por un conjunto de bits de datos procedentes de distintos dispositivos y todos ellos mezclados entre sí.

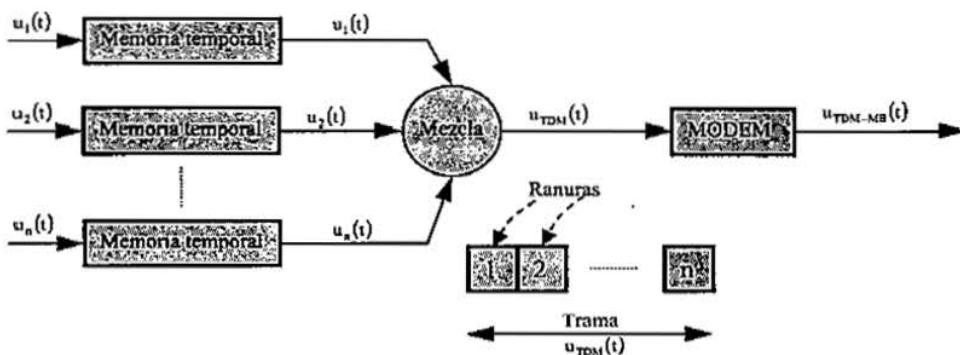


Figura 3-20: Multiplexación TDM sincrona.

Observando la Figura 3.20 se muestra un esquema general de multiplexación por división en tiempos. Así, varias señales $u_i(t)$ se quieren enviar a través de un mismo enlace de datos. Para ello, cada señal $u_i(t)$ se almacena temporalmente en un buffer. Posteriormente, se comprueban secuencialmente los buffers para extraer los datos almacenados y componer una señal de datos mezclada $u_{TDM}(t)$ y finalmente, cada señal modulada $u_{TDM}(t)$ puede ser transmitida directamente por un medio digital o quizás necesite transformarse a señal analógica, $u_{TDM-MB}(t)$, modulándose con un MODEM para transmitirse por un medio analógico.

Por supuesto, en el dispositivo receptor, deberá procederse de modo inverso para extraer las señales $u_i(t)$ y deshacer todos y cada uno de los procesos ejecutados en la fase de multiplexación. A este proceso inverso se le conoce con el nombre de *desmultiplexación*.

Por lo general, en una multiplexación TDM que se emplea para multiplexar n dispositivos conectados, donde todos y cada uno de ellos necesita enviar la misma tasa de datos, se emplean n ranuras de tiempo. De este modo, se asigna una ranura a cada dispositivo. Si la duración en tiempo de la trama de datos que componen los datos entremezclados es T , entonces la duración de cada ranura de tiempo será de T/n .

En el caso en que los n dispositivos conectados, necesiten enviar distinta tasa de datos, la multiplexación TDM tiene que emplear una de estas tres posibles estrategias. Estas estrategias se emplearán en función del tipo de tasa de datos que se necesita enviar:

- Tasa de datos distinta en los dispositivos, pero la tasa de un dispositivo sea múltiplo de la de los otros.
- Tasa de datos distinta en los dispositivos, y esas tasas no correspondan a valores múltiplos enteros unos de otros.

Para el primer caso se pueden emplear las estrategias TDM de:

- *Asignación de Múltiples ranuras*: Esta estrategia consiste en asignar más de una ranura en una trama a un mismo dispositivo de entrada. Por ejemplo, supóngase que se tienen que enviar datos de tres dispositivos A, B y C. Además, se sabe que los dispositivos A y B tienen una tasa de 10Kbps y el dispositivo C una tasa del doble, es decir, 20Kbps. Se puede asignar una ranura a A, otra ranura a B y dos ranuras a C. De este modo, en vez de emplear 3 ranuras que corresponderían a los 3 dispositivos, se emplean 4 ranuras, para conseguir múltiplos de 10Kbps.
- *Multiplexación multinivel*: Esta estrategia consiste en multiplexar en varios niveles. Por ejemplo, se tienen que enviar datos de tres dispositivos A, B y C. Además, se sabe que los dispositivos A y B tienen una tasa de 10Kbps y el dispositivo C una tasa del doble, es decir, 20Kbps. Entonces, se puede multiplexar A y B y posteriormente, como si formasen una única fuente AB, luego multiplexar C con la fuente AB que se obtiene como resultado de la multiplexación anterior.

Para el segundo de los casos se puede emplear la estrategia TDM de *inserción de pulsos*, también llamada *inserción de bits*. En estos casos se opta por comprobar qué dispositivo tiene la tasa de bits más grande, y se añaden bits de relleno a las tasas de bits del resto de dispositivos para que todos tengan la misma tasa de datos. Esta tercera técnica se combina con una de las dos estrategias anteriores para no emplear demasiados bits de relleno.

Multiplexación TDM estadística

La multiplexación síncrona tiene algunos inconvenientes y no es tan eficiente como cabría pensar, a pesar de las estrategias adoptadas cuando la tasa de datos no es la misma en todos los dispositivos. Un ejemplo de esta inefficiencia es como se desaprovechan ranuras temporales de la trama a transmitir cuando un dispositivo de entrada no tiene datos para cargar en la ranura, porque no está en disposición de enviar datos. Esto es así, porque TDM síncrona siempre asigna ranuras temporales a los dispositivos conectados al enlace tenga o no que transmitir datos en ese momento.

Este problema puede ser solucionado haciendo que el sistema de multiplexación asigne ranuras dinámicamente, basándose en la demanda. Es decir, sólo cuando el dispositivo de entrada tenga datos que transmitir recibe la asignación de una ranura. A esta estrategia de multiplexación se le conoce con el nombre de *TDM estadística*. Generalmente, cuando el número de dispositivos conectados, n , es mayor que el número de ranuras de tiempo, m , es decir $m < n$, se suele emplear multiplexación estadística, así el sistema multiplexor comprueba cada dispositivo de entrada cíclicamente para determinar si es necesario asignarle una ranura. Si el dispositivo tiene datos para enviar se asigna ranura y sino pasa al siguiente dispositivo y realiza el mismo proceso, y así sucesivamente.

La multiplexación estadística requiere de un procedimiento de direccionamiento, ya que al no haber ranuras reservadas y preasignadas a cada dispositivo de entrada, no se sabe en qué orden van los datos, y qué ranura corresponde a cada dispositivo de entrada. Este hecho obliga a insertar bits adicionales en la cabecera de cada ranura, que permitan identificar el dispositivo.

Comparación de multiplexaciones TDM síncronas con distintas estrategias

En la figura 3.21, se muestra el funcionamiento de las distintas estrategias de TDM síncrono. Así, en la Figura 3.21a se observa como se multiplexan 4 canales de 50Kbps cada uno utilizando una ranura de tiempo de 2 bits y un TDM síncrono porque la tasa de datos es la misma en todos los dispositivos de entrada. Sin embargo, en la figura 3.21b y 3.21c las tasas de los dispositivos de entrada son distintas. En concreto en la figura 3.21b se dispone de 3 canales donde uno de ellos es el doble de tasa que los otros dos, y se ha adoptado una estrategia de multiplexación multinivel de 2 niveles de multiplexación. Primero se multiplexan las fuentes A y B, y el enlace resultante AB se emplea para multiplexarlo con el dispositivo C. De ese modo la trama queda dividida en 2 ranuras de 4 bits, una ranura para el resultado de la multiplexación de AB y otra para C. Y finalmente, en la figura 3.21c se dispone de 4 canales, dos de ellos tienen el triple de tasa de datos que los otros dos. Se ha optado por emplear la técnica de múltiples ranuras, para asignar más de una ranura a la fuente A y B. En concreto se han asignado ranuras de un bit para C y D, y tres ranuras de un bit para A y B. En todos estos ejemplos se ha presupuesto que el tamaño de trama a emplear es de 8 bits, y que la máxima velocidad que soporta el enlace es de 200Kbps.

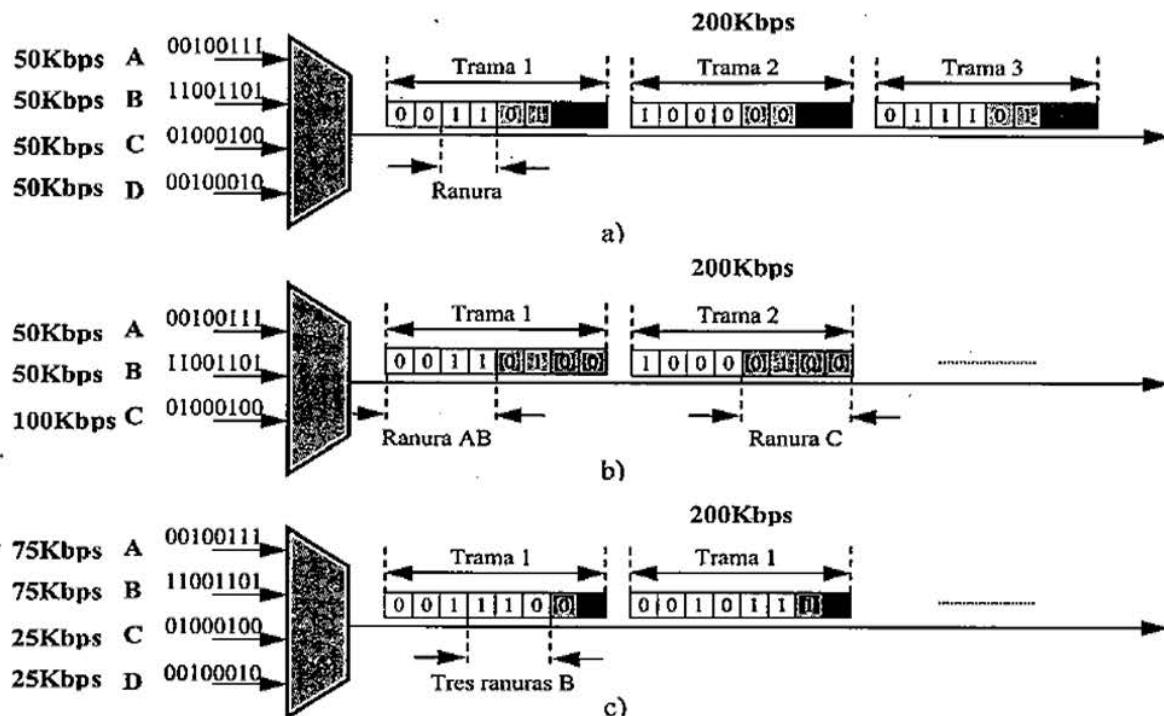


Figura 3-21: Ejemplos de Multiplexación. a) TDM síncrono. b) TDM síncrono multinivel. c) TDM síncrono con múltiples ranuras.

En la figura 3.22, se muestra un ejemplo de multiplexación en el que la tasa de bits de los dispositivos de entrada no son múltiplos enteros unos de otros. En este caso, como se comentó anteriormente, se hace necesario el empleo de la técnica conocida como inserción de pulsos o bits. Un mecanismo sencillo de implementar la técnica de inserción de bits es tomar como tasa de referencia la tasa de datos más alta de los dispositivos de entrada, en el ejemplo de la figura 50Kbps. A continuación, se añaden bits adicionales al resto de líneas con tasas más bajas, esto es a C y D que tienen tasas de 45 y 40Kbps. En concreto, se añaden tantos bits como sea necesario para incrementar su tasa hasta la tasa de referencia que para este ejemplo es de 50Kbps.

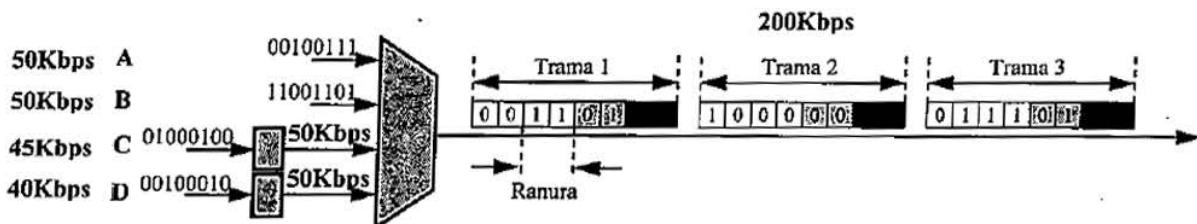


Figura 3-22: Ejemplo de Multiplexación con inserción de pulsos o bits.

Si las fuentes C y D han incrementado su tasa hasta 50Kbps, la tasa de datos de salida después de multiplexar los cuatro dispositivos de entrada es de 200Kbps, ya que hay 4 fuentes de 50Kbps. Las tasas de C y D se han incrementado en 5Kbps y 10Kbps respectivamente.

Comparación de multiplexaciones TDM síncrona y estadística

En la figura 3.23, se compara los métodos de multiplexación síncrona y estadística cuando alguno de los dispositivos no tiene datos para enviar. En tal caso, en el TDM síncrono aparecen ranuras vacías, que muestran la ineficiencia de este método. Sin embargo, el TDM estadístico como se observa es mucho más eficiente con el medio físico, enviando los datos en menos tiempo.

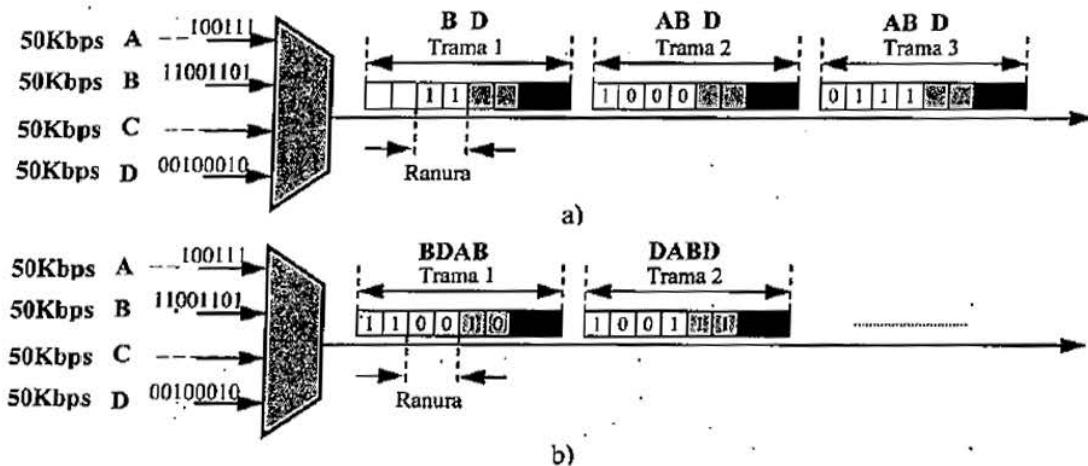


Figura 3-23: a) TDM síncrono con ranuras vacías. b) TDM estadística.

Ejercicio 3.4

Se quiere transmitir información de diversas fuentes F1, F2, F3 Y F4 por un mismo medio físico de transmisión basado en tecnología digital. Las señales que transmite cada uno son: F1 y F2 envían datos con una señal digital a 10Kbps, F3 a 5Kbps y F4 a 15Kbps. Si el medio de transmisión sólo soporta transmisión digital y hace uso de la técnica TDM para la transmisión de varias fuentes, empleando mecanismos de señalización de 2 bits/muestra (es decir 2 bits/baudio código-linea), entonces:

- Se pide calcular el ancho de banda que tendría que soportar el medio de transmisión para las cuatro comunicaciones F1, F2, F3 y F4.

- b) Teniendo en cuenta que la duración de la trama TDM es de 0.4ms. ¿Cuántos bits ocupa la trama? ¿Cuántos bits procedentes de cada fuente se almacenan en la trama de datos?
- c) ¿Cuánto dura una ranura temporal de la fuente F3?
- d) ¿Es posible implementar una multiplexación TDM empleando el método de multiplexación multinivel, sin combinarlo con ningún otro tipo de método de multiplexación?

a)

$$V_{t\text{-medio}} = V_{t\text{-F1}} + V_{t\text{-F2}} + V_{t\text{-F3}} + V_{t\text{-F4}} = \\ = 10\text{Kbits/s} + 10\text{Kbits/s} + 5\text{Kbit/s} + 15\text{Kbit/s} = 40\text{Kbit/s}$$

$$V_{t\text{-medio}} = 2B_{\text{medio}} \log_2 N \rightarrow B_{\text{medio}} = \frac{V_{t\text{-medio}}}{2 \log_2 N} = \frac{40\text{Kbit/s}}{2 \cdot 2\text{bit/muestra}} = 10\text{KHz}$$

b)

$$\frac{1\text{trama}}{4 \cdot 10^{-4}\text{s}} = 2500\text{tramas/s} \rightarrow \text{bits/trama} = \frac{40\text{Kbits/s}}{2500\text{trama/s}} = 16\text{bits/trama}$$

$$\text{FlyF2} \rightarrow \text{bits/trama} = \frac{10\text{Kbit/s} \cdot 16\text{bit}}{40\text{Kbit/s}} = 4\text{bits}$$

$$\text{F3} \rightarrow \text{bits/trama} = \frac{5\text{Kbit/s} \cdot 16\text{bit}}{40\text{Kbit/s}} = 2\text{bits}$$

$$\text{F4} \rightarrow \text{bits/trama} = \frac{15\text{Kbit/s} \cdot 16\text{bit}}{40\text{Kbit/s}} = 6\text{bits}$$

c)

$$\text{F3} \rightarrow \text{bits/trama} = \frac{4 \cdot 10^{-4}\text{s} \cdot 2\text{bits}}{16\text{bits}} = 0,5 \cdot 10^{-4}\text{s}$$

- d) No es posible. Inicialmente, se podría pensar en emplear 2 niveles de multiplexación. De modo que en el primer nivel F1 y F2 se podrían multiplexar porque tienen la misma tasa, pero F3 y F4 necesitarían de una multiplexación por asignación múltiple de ranuras ya que no tienen la misma tasa. Por lo tanto, no se puede multiplexar únicamente empleando varios niveles de multiplexación síncrona normal.

Ejercicio 3.5

Se multiplexan cuatro fuentes utilizando TDM. Si cada fuente transmite por un canal y envía 100bytes/s y se multiplexa 1 byte por canal. Se pide:

- a) Identificar el número de bits de la trama que se transmite por el enlace.
- b) Determina la tasa de tramas.
- c) Calcula la duración de la trama en segundos.
- d) Determina la tasa de bits para el enlace.
- a) Si se multiplexa 1 byte por canal, entonces cada trama está transportando 1 byte por cada fuente. Por lo tanto, el tamaño de cada trama es de 4 bytes (32 bits) puesto que hay 4 fuentes.

- b) Si cada canal envía 100byte/s y una trama transporta 1 byte por canal, entonces una trama enviará 400byte/s y transportará 4 bytes por canal, por lo tanto:

$$\frac{1\text{trama} \cdot 400\text{bytes/s}}{4\text{bytes}} = 100\text{tramas/s}$$

- c) La duración de una trama es, entonces:

$$\frac{1\text{trama}}{100\text{tramas/s}} = 10^{-2}\text{s}$$

- d) Si el enlace transporta 100tramas/s, y puesto que cada trama contiene 32 bits, entonces:

$$100\text{tramas/s} \cdot 32\text{bits} = 3200\text{bits/s} = 3200\text{bps}$$

De esos 3200bps, como hay 4 canales que se multiplexan con la misma tasa, entonces la tasa en bits para cada canal es: $3200\text{bps}/4=800\text{bps}$



Capítulo 4 . MEDIOS DE TRANSMISIÓN

En este capítulo se realiza una descripción de las principales características de los medios de transmisión empleados en las redes actuales. Estos medios de comunicación pueden clasificarse en guiados (por ejemplo, cable coaxial, par trenzado, fibra óptica) y no guiados (por ejemplo, el espectro de radiofrecuencia). De las características del medio físico empleado depende directamente el desempeño de la red que lo utiliza ya que influye directamente en el número máximo de ordenadores que se puede interconectar, la distancia entre ellos, la velocidad máxima de transmisión, etc.

CAPÍTULO 4

4-1 Introducción

4-2 Medios guiados

- Par trenzado
- Cable coaxial
- Fibra óptica

4-3 Medios no guiados

- Ondas de radio
- Microondas
- Infrarrojos

OBJETIVOS

- Comentar las principales características de los medios de transmisión empleados en redes de computadores
- Describir el principio de funcionamiento y principales tipos de medios de transmisión guiados: par trenzado, cable coaxial y fibra óptica.
- Describir las principales características y modos de transmisión en los medios no guiados
- Explicar los modos de propagación de las ondas electromagnéticas
- Describir la utilización de ondas de radio, microondas e infrarrojos para la transmisión inalámbrica en redes de datos
- Exponer los principales medios de transmisión empleados en las tecnologías de red Ethernet actuales

TERMINOLOGÍA – PALABRAS CLAVE

- | | | |
|------------------------|---|------------------------------|
| • Medio de transmisión | • Fibras ópticas multimodo de índice discreto | • Propagación en línea vista |
| • Medio guiado | • Fibras ópticas multimodo de índice gradual | • Ondas de radio |
| • Medio no guiado | • Fibras ópticas monomodo | • Microondas |
| • Par trenzado | • Propagación superficial | • Infrarrojos |
| • Cable coaxial | • Propagación troposférica | • Redes WiFi |
| • Fibra óptica | • Propagación ionosférica | • UTP/STP |

4.1 Introducción

El medio de transmisión empleado en una red de computadores constituye el canal por el cual fluye la información desde el origen hasta el destino de la comunicación. En cualquier red de computadores, el medio de comunicación empleado adquiere un papel fundamental ya que de las características del mismo va a depender directamente aspectos de la red tales como la velocidad de transmisión máxima que se pueda alcanzar, el número de equipos conectados, la distancia máxima entre los equipos que interconecta o los errores que puedan generarse a la hora de comunicar los datos. Por ejemplo, en la Tabla 4-1 se muestran distintas tecnologías de red Ethernet junto con el cable empleado, las velocidades máximas y el alcance máximo.

Tecnología	Velocidad de transmisión	Tipo de cable	Distancia máxima	Topología
10Base2	10 Mbps	Coaxial fino (50Ω)	200 m	Bus (Conector BNC-T)
10Base5	10 Mbps	Coaxial grueso (50Ω)	500 m	Bus (Conector AUI)
10BaseT	10 Mbps	2 pares trenzado (categoría UTP3)	100 m	Estrella (Hub o Switch)
10BaseF	10 Mbps	Fibra óptica	2000 m	Estrella (Hub o Switch)
100BaseT4	100Mbps	4 pares trenzado (categoría UTP3)	100 m	Estrella. Half Duplex(hub) y Full Duplex(switch)
100BaseTX	100Mbps	2 pares trenzado (UTP5)	100 m	Estrella. Half Duplex(hub) y Full Duplex(switch)
100BaseFX	100Mbps	Fibra óptica (multimodo)	2000 m	No permite el uso de hubs
100BaseCX	100Mbps	2 pares trenzado (STP)	25 m	Estrella (Hub o Switch)
1000BaseT	1000Mbps	4 pares trenzado (categoría 5e ó UTP6)	100 m	Estrella. Full Duplex (switch)
1000BaseSX	1000Mbps	Fibra óptica (multimodo)	550 m	Estrella. Full Duplex (switch)
1000BaseLX	1000Mbps	Fibra óptica (monomodo)	5000 m	Estrella. Full Duplex (switch)
10GBaseSR	10Gbps	Fibra óptica (multimodo)	500 m	Estrella. Full Duplex (switch)
10GBaseLX4	10Gbps	Fibra óptica (multimodo)	500 m	Estrella. Full Duplex (switch)
10GBaseT	10 Gbps	Par Trenzado (UTP6)	55 m	Estrella. Full Duplex (switch)

Tabla 4-1. Tecnologías de red y principales características dependiendo del cable utilizado.

Una primera clasificación de los medios de comunicación puede hacerse atendiendo a si se trata de medios guiados o no guiados:

- **Medios guiados:** son aquellos en los cuales la información está confinada en un cable determinado.
- **Medios no guiados:** son los que no confinan las señales mediante ningún tipo de cable, sino que las señales se propagan libremente a través del medio.

4.2 Medios guiados

Como se ha indicado anteriormente, los medios guiados son aquéllos en los que el canal por el que se transmiten las señales son medios físicos, es decir, por medio de un cable. En este tipo de medios se emplea un conductor de un dispositivo a otro, limitando la propagación de la señal al interior del conductor. A lo largo de este capítulo se van a estudiar los medios guiados que más comúnmente se emplean en las redes de computadores actuales. Estos medios serán los siguientes: par trenzado, cable coaxial y fibra óptica.

4.2.1 Par trenzado

Características generales

Este se trata de uno de los medios de transmisión más empleados en las redes de área local actuales. En su configuración básica está constituido por 2 cables de cobre entrecruzados en forma de espiral recubiertos por un aislante. La señal se envía por ambos cables de forma balanceada y en banda base. Cuando se transmiten señales balanceadas cada par de hilos transporta una señal independiente del resto de pares. La señal que transporta un par se mide como la diferencia de potencial entre los dos hilos que constituye el par. La disposición de los cables de manera trenzada provoca que sea más inmune al ruido y a la atenuación, de hecho a mayor número de torsiones por cm (lo que se suele denominar como longitud de trenzado) se consigue un aumento de la robustez del medio a las interferencias electromagnéticas, obteniendo una mayor calidad en la transmisión.

Normalmente un cable de par trenzado está formado por un grupo de pares trenzados, habitualmente cuatro, recubiertos de una envoltura protectora (a veces se agrupan muchos pares trenzados en un mismo cable para la transmisión de distintas líneas de comunicación a largas distancias, véase la Figura 4-2.a). Cada uno de estos pares se identifica mediante un color, siendo los colores asignados y las agrupaciones de los pares de la siguiente forma (ver Figura 4-1):

- Par 1: Blanco-Azul/Azul
- Par 2: Blanco-Naranja/Naranja
- Par 3: Blanco-Verde/Verde
- Par 4: Blanco-Marrón/Marrón

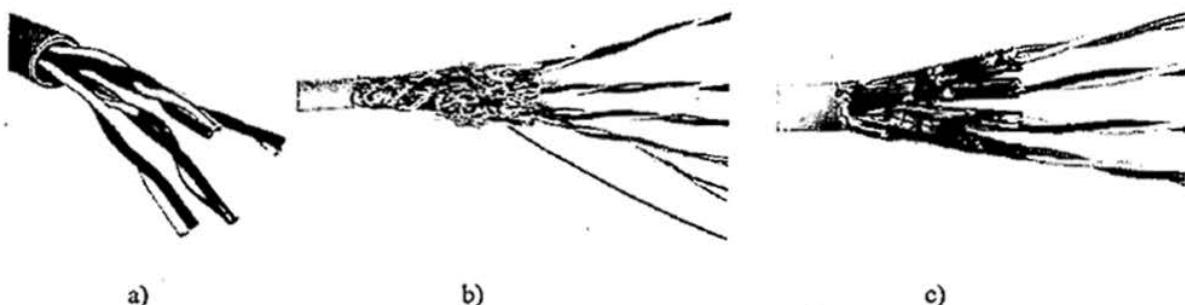


Figura 4-1. Cable de par trenzado. a) No apantallado UTP. b) Apantallado STP. c) Apantallado S/STP

Aplicaciones

El par trenzado es empleado indistintamente para la transmisión de señales analógicas y digitales. Así, es el medio más extendido tanto para redes de telefonía como en comunicación dentro de edificios en LANs. En cuanto a esta última aplicación de los pares trenzados, la velocidad clásica que se solía alcanzar era de 10 Mbps. No obstante existen redes LAN de pares trenzados de 10 Gbps, aunque este último tipo de redes se encuentran limitadas en cuanto al número de posibles dispositivos a conectar así como en la extensión geográfica de la red.

Su éxito se ha debido a que es mucho menos costoso que cualquiera de los otros tipos de medios de transmisión descritos en este capítulo y, además, es el más fácil de manejar. Sin embargo, comparando este medio con las fibras ópticas o los cables coaxiales, se puede afirmar que el par trenzado permite distancias menores, menor ancho de banda y menor velocidad de transmisión.

Tipos de par trenzado

Existen básicamente dos tipos de pares trenzados: apantallados y sin apantallar:

- UTP Unshielded Twisted Pair (Par trenzado no apantallado). Este es el cable más empleado en telefonía y LANs debido a su bajo precio y su fácil manipulación. A pesar del aumento de robustez que proporciona el trenzado del cable, los pares trenzados no apantallados se pueden ver afectados por interferencias electromagnéticas externas como pueden ser interferencias causadas por pares cercanos. El estándar para conectores de cable UTP es el RJ-45.
- STP Shielded Twisted Pair (Par trenzado apantallado). En este caso el par trenzado se encuentra recubierto por una malla metálica, lo que garantiza un aislamiento a interferencias electromagnéticas externas (véase Figura 4-1b). Existe una variante, denominada S/STP en la que se encuentra apantallado cada par de cables además de un apantallamiento general de todos los cables (véase Figura 4-1c). Este tipo de cables, aunque más costoso que el UTP, proporciona mejores prestaciones a velocidades de transmisión mayores. La protección del STP, para que sea más eficaz, requiere una configuración de interconexión con tierra (dotada de continuidad hasta el terminal), es por ello por lo que con el STP se suele utilizar conectores RJ49 con carcasa metálica.

Los principales tipos de cable UTP empleados en LANs actuales se recopila en la EIA (Asociación de Industrias Electrónicas) Commercial Building Wiring Standard 568. El estándar EIA 568 utiliza UTP para crear estándares que se apliquen a todo tipo de espacios y situaciones de cableado, garantizando de esta manera productos homogéneos al público. Estos estándares incluyen las siguientes categorías de cables UTP dependiendo del número de pares que tenga el cable, del

número de vueltas por metro que posea su trenzado y de los materiales utilizados (si no se indica lo contrario la longitud máxima del cable será de 100 metros):

- Categoría 1: Cable de teléfono tradicional con un ancho de banda de 1 MHz (transmisión de voz pero no de datos). Este es el cable telefónico instalado habitualmente antes de 1983, aunque ya no es utilizado.
- Categoría 2: Permite transmisión de datos hasta un máximo de 4 Mbps en un ancho de banda de 1 MHz. Este tipo de cable contiene 4 pares trenzados y es empleado en redes como la Token Ring IEEE 802.5 a 4 Mbps.
- Categoría 3: Admite una velocidad máxima de hasta 16 Mbps y presenta un ancho de banda de 16 MHz. Este tipo de cable contiene 4 pares trenzados y es el cable que habitualmente se instala para la transmisión de voz. Es empleado en la red IEEE 802.3 10BaseT Ethernet a 10 Mbps. Presenta una longitud de trenzado aproximada de 7,5 a 10 cm. A pesar de las limitaciones anteriores existe un estándar 100BaseT4 que permite alcanzar los 100 Mbps empleando este tipo de cables y haciendo uso de múltiples pares.
- Categoría 4: Permite transmisión de hasta 20 Mbps con un ancho de banda máximo de 20 MHz. Este tipo de cable contiene 4 pares de hilos de cobre trenzados y es empleado en redes como Token Ring IEEE 802.5 a 16 Mbps.
- Categoría 5: Puede transmitir datos hasta una velocidad máxima de 100 Mbps presentando un ancho de banda de 100 MHz. Este tipo de cable contiene 4 pares de hilos de cobre trenzados. Este es el cable empleado en tecnologías como FastEthernet 100BaseTX, 100BaseVGAnyLan y transferencia modo asíncrono (ATM) a 155 Mbps. Presenta una longitud de trenzado aproximada de 0.6 a 0.85 cm.
- Categoría 5e (CAT 5 enhanced "mejorada"): Presenta un ancho de banda de 100 MHz y contiene 4 pares de hilos de cobre trenzados. Este cable es utilizado en la red Gigabit Ethernet 1000BaseT.
- Categoría 6: Presenta un ancho de banda de 250 MHz. Soporta velocidades de hasta 1000 Mbps con la red IEEE 1000BaseT. Véase en la Figura 4-2.b como el trenzado en esta categoría es mucho mayor que en la Categoría 3 (Figura 4-2.a).
- Categoría 6e: Dispone de un ancho de banda máximo de 500 MHz y puede alcanzar una velocidad de transmisión máxima de 10Gbps en la tecnología IEEE 10GBaseT.
- Categoría 7: Presenta un ancho de banda de 600 MHz y puede alcanzar velocidades de transmisión superiores a 10Gbs.



Figura 4-2. Cable de par trenzado. a) Agrupación de 50 pares trenzados UTP-3. b) Cable UTP-6.

4.2.2 Cable coaxial

Características generales

Los cables coaxiales están constituidos por un conductor interno, por el que se envía el voltaje correspondiente a la información a enviar, rodeado por un material aislante. Alrededor del material aislante se dispone una malla conductora externa que sirve de referencia y que también está recubierta de un material aislante. En la Figura 4-3a se muestran dos ejemplos de cable coaxial en los que se observa el núcleo y la malla conductora que la recubre. En la Figura 4-3b se ha representado el conector BNC que se emplea habitualmente en cables coaxiales para datos.

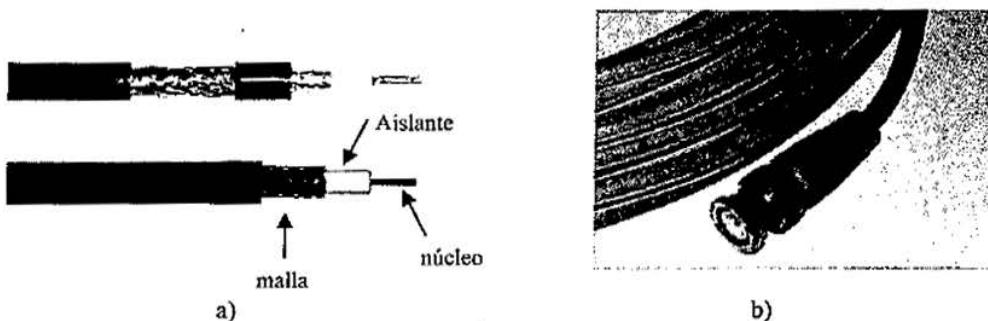


Figura 4-3. Cable coaxial.

Aplicaciones

Frente al par trenzado, los cables coaxiales presentan una mayor robustez al ruido empleándose, por ejemplo, en entornos industriales con excesivo ruido electromagnético. Los cables coaxiales se utilizan para transmitir tanto señales analógicas como digitales. Su uso es muy común en televisión, aunque también es posible encontrarlos en redes de área local y telefonía a gran distancia. Tradicionalmente, el coaxial ha sido un elemento fundamental en la red de telefonía a larga distancia. En la actualidad viene sustituyéndose progresivamente por la fibra óptica, las microondas terrestres y las comunicaciones vía satélite debido, entre otros motivos, a que requiere menos cantidad de repetidores. Entre las ventajas de la utilización de este tipo de cables cabe citar que se logra una alta inmunidad a ruido, interferencias, etc. Además, posee un gran ancho de banda (alrededor de 500 MHz) y es posible usarlo para transmisión de datos a largas distancias. Sin embargo, su principal inconveniente que ha provocado que haya sido sustituido en LANs gradualmente por pares trenzados ha sido su poca flexibilidad y dificultad de manipulación.

Tipos de cables coaxiales

Hay cables coaxiales con distintas impedancias dedicados a diferentes tipos de transmisión. Cabe mencionar los siguientes tipos:

- **Cable coaxial 50 Ω.** Es empleado en transmisión en banda base en LANs utilizando Manchester y Manchester diferencial. Este tipo de medio ha sido remplazado por el par trenzado ya que permite una detección de errores más sencilla. Básicamente, existen dos tipos de cable coaxial de 50 Ω: coaxial grueso y coaxial fino. El cable coaxial grueso era empleado en las redes Ethernet 10Base5, con una velocidad máxima de 10 Mbps a 500 m. Este tipo de redes están prácticamente en desuso en la actualidad. El cable coaxial fino se emplea en las redes Ethernet 10Base2, que permite una velocidad máxima de transmisión de 10 Mbps a una distancia de 200 m. En la Figura 4-4a se muestra el conector en BNC-T

empleado para unir los distintos segmentos que conforma el bus que pasa por cada uno de los equipos integrantes de la LAN. En la Figura 4-4b se muestra una tarjeta de red Ethernet con conexión BNC para cables coaxiales. Hoy en día es difícil encontrar este tipo de tarjetas.

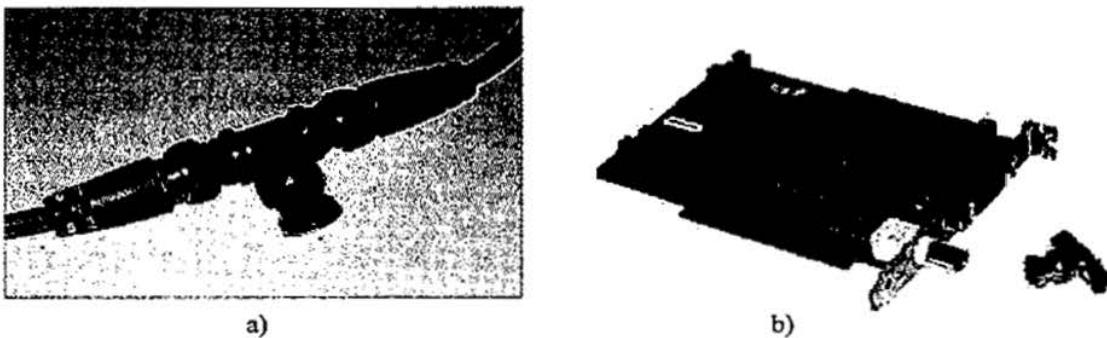


Figura 4-4. a) Conector BNC-T. b) Tarjeta de red Ethernet con conexión para cable coaxial fino.

- Cable coaxial $75\ \Omega$. Utilizado para transmisión en banda modulada en televisión analógica por cable empleando multiplexión por división en frecuencia de múltiples canales. Se pueden emplear para transmitir datos digitales, pero modulando para poder convertir la señal que circula por el medio a analógica.

4.2.3 Fibra óptica

Características generales

Si en los medios de comunicación anteriores la información se transmitía en forma de tensión o corriente eléctrica, en las fibras ópticas las señales se transmiten en forma de luz. Las fibras ópticas están constituidas por un núcleo de cristal de silicio por el que se envía un haz de naturaleza óptica que codifica la información. El núcleo está rodeado por un recubrimiento que puede ser otro cristal o plástico con propiedades ópticas distintas al núcleo (véase Figura 4-5a). La separación entre el núcleo y el revestimiento actúa como un reflector, confinando así el haz de luz dentro del núcleo. Este recubrimiento a su vez suele disponer también de una protección adicional con el objetivo de evitar que le afecte factores externos como la humedad. Con todo, las fibras ópticas son significativamente más finas que los medios de transmisión anteriores y con un menor peso. Los principales tipos de conectores empleados son los que se muestran en la Figura 4-5b. El conector FC se usa en la transmisión de datos y en las telecomunicaciones. El FDDI, se usa para redes de fibra óptica de datos. El LC y MT-Array se utilizan en transmisiones de alta densidad de datos. El SC y SC-Dúplex se utilizan para la transmisión de datos y, finalmente, el ST o BFOC se usa en redes de edificios y en sistemas de seguridad.

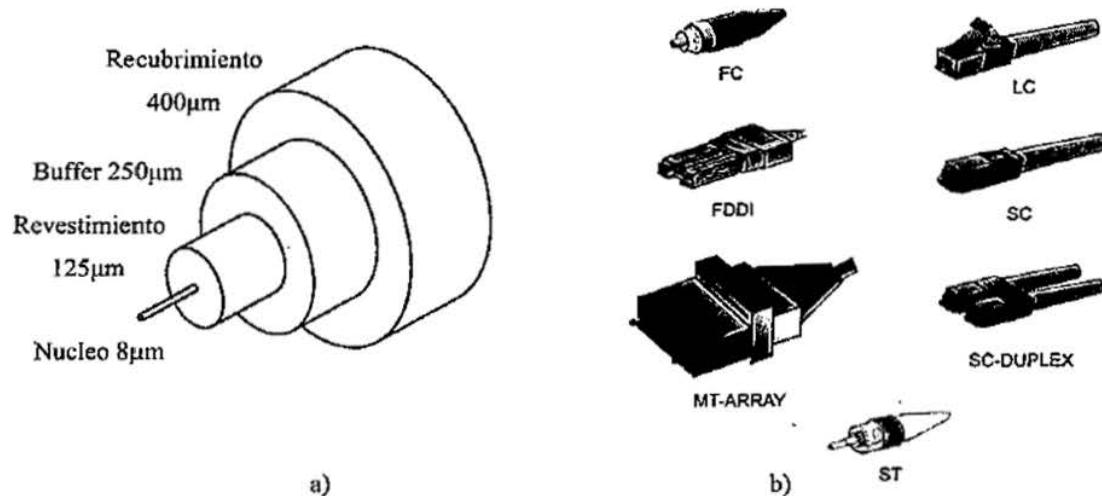


Figura 4-5. a) Composición de la fibra óptica. b) Principales tipos de conectores.

Aplicaciones

Las fibras ópticas constituyen actualmente el medio físico más ampliamente utilizado para la transmisión de datos a largas distancias, siendo cada vez más utilizado en las redes de telefonía (son capaces de albergar hasta 100.000 canales de voz). También es un medio muy extendido como red de área metropolitana enlazando centrales telefónicas o cabeceras de cable-módem dentro del área metropolitana y sin la necesidad de emplear repetidores. También se encuentran importantes aplicaciones de las fibras ópticas dentro del ámbito de las LAN. Se han desarrollado estándares de red basados en fibra óptica que permiten velocidades desde 100 Mbps hasta 10 Gbps (véase Tabla 4-1), las cuales a su vez permiten interconectar miles de estaciones.

Una de las principales ventajas de la fibra óptica es que es inmune al ruido electromagnético y presenta una atenuación significativamente menor que en los cables coaxiales y pares trenzados. Además, el ancho de banda y, por lo tanto, la velocidad de transmisión, son considerablemente mayores en las fibras ópticas. Las fibras ópticas tienen un ancho de banda de alrededor de 1 THz, de manera que este ancho de banda excede ampliamente al de los cables de cobre. Se ha determinado experimentalmente que es posible alcanzar velocidades de transmisión de varios cientos de Gbps para decenas de kilómetros de distancia. Sin embargo, sus principales desventajas radican en el elevado coste de instalación (requieren de transmisores y receptores ópticos adecuados), en la dificultad de realizar empalmes y montaje de los conectores, la fragilidad de las fibras a torsiones excesivas así como la dificultad de detectar posibles averías o roturas en la fibra.

Tipos de fibras ópticas

Las fibras ópticas se basan en el principio de reflexión total para guiar ondas de luz. El fenómeno de reflexión total se da en cualquier medio transparente que tenga un índice de refracción mayor que el medio que lo contenga. En las fibras ópticas los rayos de luz que inciden con un determinado ángulo en el núcleo se reflejan y se propagan dentro del núcleo de la fibra. Para otros ángulos de incidencia en la fibra, los haces de luz son absorbidos por el material que forma el revestimiento. Los principales tipos de fibra óptica que se emplean en redes de computadores son los siguientes:

- Fibra óptica multimodo de índice salto o índice discreto (Figura 4-6.a). En este tipo de fibra óptica se transmiten varios haces de luz simultáneamente, es decir, hay multitud de ángulos de incidencia para los que se da reflexión total. A este conjunto de ángulos se le

denomina modos. Cada rayo de luz describirá un camino desde el origen al destino, esto provoca que los pulsos de luz que se transmiten se dispersen con el tiempo, lo que provoca que se limite la velocidad a la que los datos puedan ser correctamente recibidos. La principal ventaja de este tipo de fibras óptica es que es la más barata, sin embargo, presenta un menor ancho de banda.

- Fibra óptica multimodo de índice gradual (Figura 4-6.b). En este caso también se transmiten varios modos simultáneamente, sin embargo, se tiende a igualar la velocidad de todos ellos de manera que todos los haces de luz llegan aproximadamente en el mismo instante al extremo. Para conseguir esto, el núcleo presenta un índice de refracción superior en la parte central y dicho índice va disminuyendo progresivamente en la periferia. Esto provoca que la luz periférica llegue al receptor al mismo tiempo que los rayos axiales del núcleo. Este tipo de fibra óptica es más cara que la anterior pero presenta un mayor ancho de banda, por lo que se pueden alcanzar mayores velocidades. Este tipo de fibras es habitualmente empleado en LANs.
- Fibra óptica monomodo (Figura 4-6.c). Partiendo de una fibra multimodo de índice salto, si se reduce el radio del núcleo a dimensiones del orden de magnitud de la longitud de onda un solo modo podrá pasar. En este tipo de fibra óptica se transmite un único haz de luz a lo largo del eje de la fibra, evitando de esta manera los problemas mencionados para las fibras multimodo.

En las redes de fibra óptica se emplean dos tipos de fuentes de luz: diodos LED (Light Emitting Diodes), menos costoso, y diodos ILD (Injection Laser Diode) con un principio de funcionamiento semejante al láser, más eficaces y permiten alcanzar velocidades de transmisión mayores. La diferencia entre diodos LED e ILD es el ancho espectral y la potencia lumínica. El ILD proporciona potencias de miliwattios frente a los microwattios del LED. Además, en los ILD su ancho espectral es menor que en los LED, proporcionando un haz de luz más coherente y enfocado.

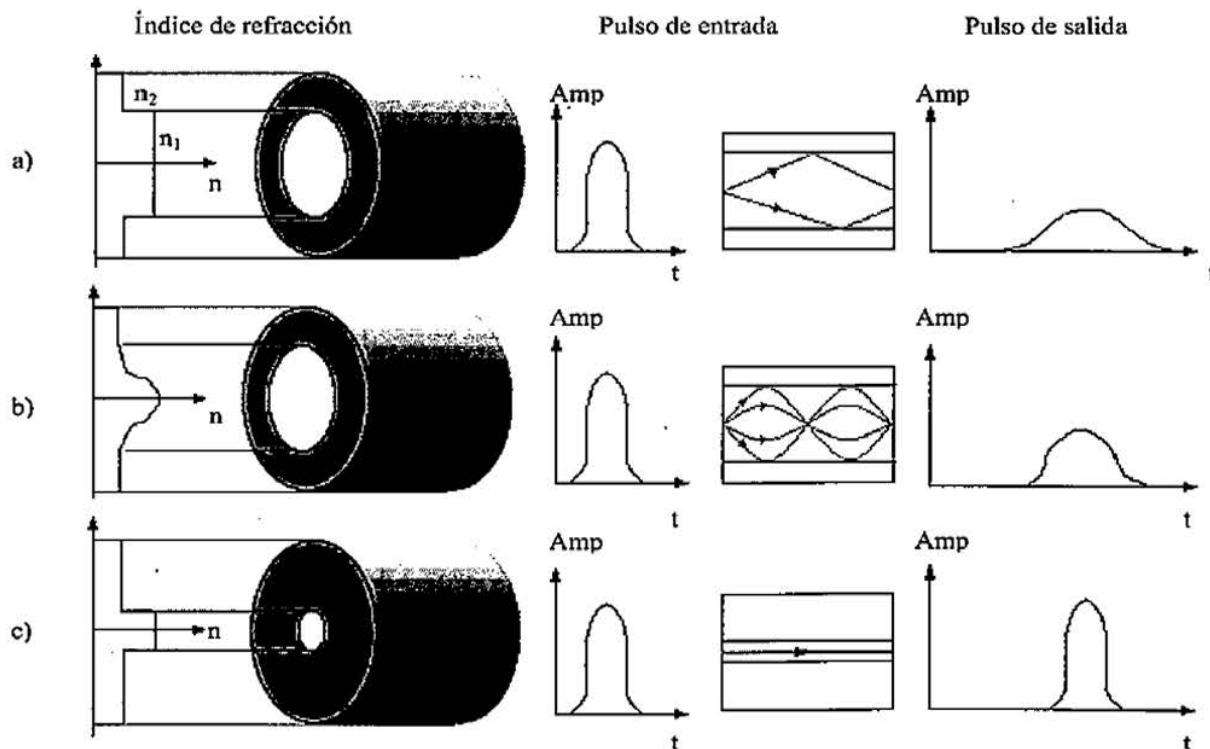


Figura 4-6. Principales tipos de fibras ópticas. a) Fibra óptica multimodo de índice salto. b) Fibra óptica multimodo de índice gradual. c) Fibra óptica monomodo.

4.3 Medios no guiados

Una de las áreas que más ha evolucionado en los últimos años en el mundo de las telecomunicaciones es la transmisión inalámbrica. Un sistema inalámbrico es aquel que permite la transmisión de cualquier tipo de información (audio, vídeo, datos) desde cualquier lugar y en cualquier momento, con posibilidad de transmitir en tiempo real de ser necesario. Entre las ventajas de un sistema inalámbrico sobre uno cableado podemos mencionar:

- Movilidad, la cual apoya la productividad y la efectividad con que se presta el servicio.
- Aunque los costos iniciales son mayores que los que supondría un sistema cableado, a lo largo del tiempo los gastos de operación pueden ser significativamente menores.
- Menor tiempo de instalación y puesta en marcha del sistema. La instalación es más sencilla debido a que no necesitan repartir el cableado por los ordenadores que constituyen la red.
- Existe completa flexibilidad en cuanto a la configuración del sistema. Se pueden tener diversas topologías para satisfacer los requerimientos de aplicaciones e instalaciones específicas.

Habitualmente, para soportar este tipo de comunicaciones se realiza la transmisión de datos en forma de ondas electromagnéticas de radio o microondas por el aire. Estas ondas pueden emplearse para codificar información digital sin más que modificar alguno de sus parámetros físicos. Estas ondas electromagnéticas engloban todo el espectro de radiación de energía del universo. Dentro de este espectro de radiofrecuencia destacan tres tipos de ondas electromagnéticas usadas en telecomunicaciones: ondas de radio, microondas e infrarrojos.

Dependiendo de la frecuencia de la onda electromagnética se tiene un modelo de propagación:

- Propagación superficial. Las ondas se propagan a lo largo de la superficie terrestre. Las ondas siguen la curvatura de la tierra por la cual la señal es capaz de alcanzar grandes distancias antes de que la señal sea absorbida por la tierra. La señal puede sortear edificios y montañas. Es la propagación que se produce para las ondas que se encuentran en la banda VLF y LF. Esto se debe a que las frecuencias altas se ven menos afectadas por los fenómenos atmosféricos (no requiere de antenas grandes para tener una transmisión efectiva de gran directividad), lo que provoca la confiabilidad de que la información llegue al receptor. Este tipo de propagación se utiliza para la televisión y radio FM.
- Propagación troposférica. La capa troposférica se encuentra entre los 11 Km y los 16 Km. En esta capa se forman las nubes y la temperatura desciende rápidamente debido a la altura. Cuando se produce la inversión del gradiente de temperatura se generan los denominados canales de ionización, los cuales son ideales para que las ondas puedan viajar. Este tipo de propagación se emplea ocasionalmente en VHF.
- Propagación ionosférica. La ionosfera se encuentra entre 40 Km y 320 Km y está formada por aire altamente ionizado por la radiación solar. Cuando esta capa se encuentra eléctricamente cargada hace que la señal comience a cambiar en un cierto ángulo, esto lo hace sucesivamente hasta que se realiza una reflexión total y la señal regresa a la tierra. Se trata de uno de los modos de propagación más importantes y permite conectar puntos a una distancia de 4000 Km. Este tipo de propagación se emplea en HF.
- Propagación en línea vista. Para realizar este tipo de propagación es necesario que exista una línea de vista entre el transmisor y el receptor. En este tipo de comunicación se utilizan frecuencias por encima de los 50 MHz, las correspondientes a VHF, UHF y canales con frecuencia superior.

- Propagación espacial. En este caso se emplean satélites como reflector y presentan frecuencias de UHF y superiores. Existen dos tipos de satélites:
 - Satélites geoestacionarios pasivos. Estos satélites se encuentran en órbita alrededor de la Tierra y sirven como espejos, reflejando la onda de radio y regresándola a la Tierra.
 - Satélites geoestacionarios activos. Estos satélites funcionan igual que los geoestacionarios pasivos, pero la diferencia es que estos reciben la señal y la amplifican para mandarla de regreso a la Tierra.

4.3.1 Ondas de radio

Las ondas de radio son omnidireccionales, es decir, se propagan en todas las direcciones. Esto hace que este tipo de ondas no requieran antenas parabólicas para su envío/recepción. El rango de frecuencias comprendido entre 30 MHz y 1 GHz es muy adecuado para la difusión simultánea de información a varios destinos (las ondas comprendidas en este rango son denominadas ondas de radio y cubren la banda VHF y parte de UHF). A diferencia de la región de las microondas, las ondas de radio son menos sensibles a la atenuación producida por la lluvia.

Las redes LAN inalámbricas están normalizadas en el estándar 802.11 del IEEE, y sus dispositivos son certificados como WiFi (wireless fidelity). Las ondas más extendidas para su empleo como LAN inalámbricas son aquellas que no requieren licencia (ondas de radio: banda de 915 MHz, microondas: bandas de 2,4 GHz y 5 GHz).

4.3.2 Microondas

Se distinguen dos tipos de microondas dependiendo de la forma de realizar la transmisión: microondas terrestres y microondas por satélite.

Microondas terrestres de larga distancia

En los sistemas de microondas terrestres se emplea una propagación en línea vista en la que se dispone de una antena parabólica que debe estar perfectamente orientada con la antena receptora. Para conseguir transmisiones a larga distancia se concatenan distintos enlaces punto a punto entre antenas situadas en torres adyacentes hasta cubrir la distancia deseada.

Este tipo de microondas se emplean en servicios de telecomunicación para transmitir televisión y voz a largas distancias. La banda de frecuencias está comprendida entre 1 y 40 GHz. Cuanto mayor sea la frecuencia mayor es el ancho de banda potencial para los canales que se definen y, por lo tanto, mayor es la velocidad de transmisión que es posible alcanzar. En este tipo de ondas la atenuación aumenta con la lluvia y masas de agua, incluidos los seres vivos (p. ej. zonas muy pobladas), siendo este efecto especialmente significativo para frecuencias por encima de 10 GHz. La banda de 12 GHz es empleada para la transmisión por cable. Frecuencias superiores tienen mayor atenuación por lo que son empleadas para la transmisión a cortas distancias.

Microondas por satélite

Un satélite de comunicaciones es esencialmente una estación que retransmite microondas. Se usa como enlace entre dos o más receptores/transmisores terrestres denominados estaciones base. El satélite recibe la señal en una banda de frecuencia, la amplifica y posteriormente lo retransmite en otra

banda. La utilización de satélites tiene distintas aplicaciones como la difusión de televisión, transmisión telefónica y datos a larga distancia o creación de redes privadas.

El rango de frecuencias óptimo para la transmisión vía satélite está en el intervalo comprendido entre 1 y 10 GHz. Por debajo de 1GHz el ruido producido por causas naturales es apreciable y por encima, como se ha indicado anteriormente, la señal se ve afectada por la meteorología y existe mayor atenuación. La mayoría de satélites proporcionan la banda de 5,925-6,425 GHz para la transmisión de estaciones terrestres hacia el satélite y entre 3,7 GHz y 4,2 GHz para la comunicación de satélite a estación terrestre, existiendo un retardo de propagación aproximado de un cuarto de segundo para una transmisión que vaya desde una estación terrestre hasta otra pasando por un satélite.

WiFi

Actualmente existen distintos estándares IEEE que emplean microondas en LANs. Los estándares más empleados son el IEEE 802.11b y el IEEE 802.11g con una velocidad de 11 Mbps y 54 Mbps respectivamente. Ambos emplean la banda de 2,4 GHz que está disponible universalmente.

También existe el estándar IEEE 802.11a, conocido como WiFi 5, que opera en la banda de 5 GHz y es empleado fundamentalmente en EEUU. La banda de 5 GHz ha sido recientemente habilitada en Europa y, además no existen otras tecnologías (Bluetooth, microondas) que la estén utilizando, por lo tanto existen muy pocas interferencias. Su alcance es algo menor que el de los estándares que trabajan a 2.4 GHz (aproximadamente un 10%), debido a que la frecuencia es mayor (a mayor frecuencia, menor alcance).

Por último, destacar la normativa 802.11n (actualmente sigue en borrador) que permite una velocidad de hasta 500Mbps con una distancia de operatividad óptima de 50 metros en entornos cerrados. Además, es compatible con las redes 802.11a, 802.11b y 802.11g, lo que significa que funciona en las dos bandas actuales, 2.4GHz y 5GHz.

Existen otras tecnologías inalámbricas como Bluetooth que también funcionan a una frecuencia de 2.4 GHz, por lo que puede presentar interferencias con WiFi. Debido a esto, en la versión 1.2 del estándar Bluetooth por ejemplo se actualizó su especificación para reducir en la medida de lo posible las interferencias con la utilización simultánea de ambas tecnologías.

4.3.3 Infrarrojos

Las comunicaciones mediante infrarrojos se llevan a cabo mediante transmisores/receptores que modulan luz infrarroja. Las distancias alcanzadas son pequeñas (varios metros como mucho). Son empleados en los mandos a distancia de muchos dispositivos que envían datos en banda base. Uno de sus principales inconvenientes es el hecho de que en el espectro infrarrojo se producen muchas interferencias. Actualmente, para la transmisión de datos por infrarrojos, existen normas que alcanzan velocidades de hasta 4 Mbps.

Capítulo 5. DISEÑO DEL NIVEL DE ENLACE Y CONTROL DE ERRORES

En este capítulo se presentan las funciones del nivel de enlace de la arquitectura de red OSI. Dentro de estas funciones cabe mencionar algunas como las de crear y delimitar tramas, identificar estaciones, control de flujo, etc. Una de las principales funciones es la de ofrecer al nivel de red un enlace libre de errores. En este capítulo se describirán también distintos métodos para que un receptor sea capaz de determinar si una trama es correcta o si algún bit se ha alterado durante la transmisión. También se mostrarán algoritmos para la corrección de dichos errores.



CAPÍTULO 5

5-1 Servicios del nivel de enlace

5-2 Funciones del nivel de enlace

5-3 Iniciación y terminación de la comunicación

5-4 Delimitación de tramas

- Delimitación temporal
- Delimitación por número de caracteres
- Delimitación por caracteres especiales
- Delimitación por bits especiales
- Delimitación por códigos de línea especiales

5-5 Direccionamiento de tramas

5-6 Detección y corrección de errores en tramas

- Comprobación de paridad
- Comprobación de redundancia cíclica
- Corrección de errores empleando códigos de bloque
- Códigos Reed-Solomon

5-7 Medición de la tasa de errores

OBJETIVOS

- Describir las principales funciones que ofrece el nivel de enlace y detallar los tipos de servicios que ofrece el nivel de enlace al nivel de red
- Explicar en qué consiste las funciones de iniciación y terminación de la comunicación realizadas por el nivel de enlace
- Describir el formato de trama habitual empleado por las tecnologías de red a nivel de enlace
- Indicar distintas formas de delimitar tramas a nivel de enlace
- Detallar métodos para la detección de errores: paridad y códigos de redundancia cíclica
- Describir como es posible realizar la corrección de errores empleando códigos de bloque: Código de Hamming y códigos de Reed-Solomon
- Exponer conceptos empleados para la medición de la tasa de errores en un medio físico.

TERMINOLOGÍA – PALABRAS CLAVE

- | | | |
|-----------------------------------|----------------------------------|-------------------------|
| • Nivel de enlace | • Delimitación de trama | • Códigos Hamming |
| • Control de flujo | • Direccionamiento | • BER |
| • Servicios orientados a conexión | • Códigos de redundancia cíclica | • Tasa de errores |
| • Servicios con reconocimiento | • Paridad | • Detección de errores |
| • Trama | • Códigos de Bloque | • Corrección de errores |
| | | • Reed- Solomon |

5.1 Servicios del nivel de enlace

El objetivo fundamental del nivel de enlace del modelo OSI es el de ofrecer un enlace lógico sobre el medio físico libre de errores al nivel de red. La capa de enlace está en un nivel superior a la capa física. Por lo tanto, la capa de enlace de datos es responsable de la transferencia fiable de información a través de un circuito de transmisión de datos. A este nivel la transmisión de datos se realiza mediante tramas que son las unidades de información con sentido lógico para el intercambio de datos en la capa de enlace (ver Figura 5-1).

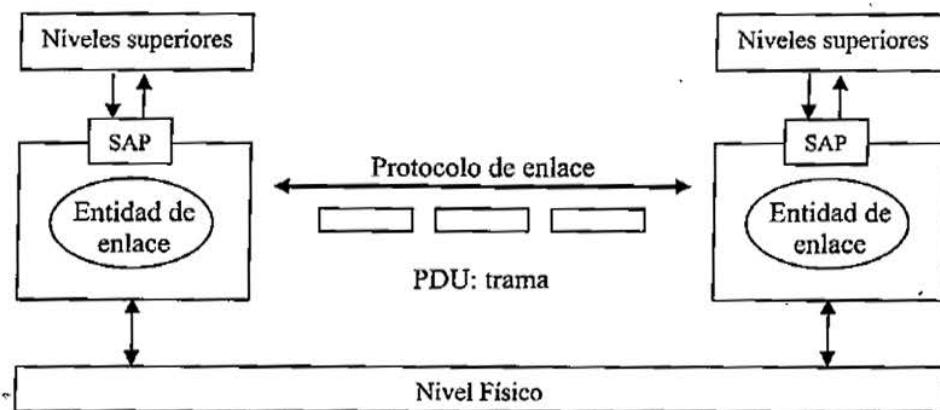


Figura 5-1. Nivel de enlace del modelo OSI

Según el modelo OSI, la capa o nivel de enlace se encarga de los servicios de trasferencia de datos seguros a través del enlace físico. Por lo tanto, envía bloques de datos denominados tramas, llevando a cabo la sincronización, el control de errores y de flujo necesario. La capa de enlace pretende conseguir una comunicación eficiente, es decir, aprovechar al máximo el medio de transmisión. Además, esa comunicación debe realizarse de manera fiable, es decir, debe garantizarse que los datos lleguen correctamente al destino. En esta comunicación pueden surgir algunos problemas como son posibles errores en el medio de transmisión o, por ejemplo, que el emisor realice el envío de datos a una velocidad mayor a la que es capaz de procesar el receptor. El nivel de enlace debe proporcionar protocolos que consigan realizar una comunicación eficiente y fiable y que sean capaces de resolver los problemas como los anteriores.

Como se ha indicado anteriormente, el nivel de enlace de datos detecta y corrige los errores que ocurren en el nivel físico, proporcionando un medio libre de errores de transmisión al nivel de red. Para ello, trocea los datos del nivel de red en tramas de datos, los transmite secuencialmente, y procesa las tramas reconocidas, ofreciendo una comunicación fiable y eficiente entre dos máquinas adyacentes. Como se describirá a lo largo de este capítulo, el nivel de enlace implementará algoritmos de detección y/o corrección de errores que permiten detectar si se ha producido un error en la transmisión de algunos de los bits que constituyen la trama y, en ese caso, corregir los errores si es posible.

Para evitar el segundo problema indicado anteriormente, es decir, que el emisor trabaje a una velocidad mayor a la que es capaz de procesar al receptor, se implementan técnicas denominadas de control de flujo, con las cuales se sincroniza el envío y confirmación de datos entre emisor y receptor. Los circuitos de comunicación tienen una velocidad de transmisión de datos finita y hay un retardo de propagación diferente de cero, es decir, existe un determinado tiempo que transcurre desde que un bit se envía hasta que se recibe. Estas limitaciones, asociadas con la velocidad finita de las máquinas para

procesar los datos, tienen implicaciones muy importantes en la eficiencia de la transferencia de datos. El nivel de enlace se diseña para paliar estos problemas y que la comunicación sea fiable y eficiente a través de una serie de algoritmos que los resuelven.

Como resumen, cabe mencionar que el servicio principal de la capa de enlace es la transferencia de datos de la capa de red de la máquina origen a la capa de red de la máquina destino. El nivel de enlace proporciona al nivel de red diferentes servicios de acuerdo a la calidad que se quiere conseguir en la transmisión:

- *Servicios sin conexión y sin reconocimiento.* En este tipo de servicio no se establece conexión previa y las tramas se envían a la máquina destino sin acuse de recibo, es decir, no se confirma la recepción correcta de los datos. Es válido para aplicaciones en tiempo real (Se emplea en medios físicos donde los retardos perjudican más que los errores). Además, este tipo de servicios es adecuado en líneas de comunicación con tasa de errores bajas como LAN cableadas.
- *Servicios sin conexión y con reconocimiento.* En este caso tampoco se establece conexión previa, pero ahora cada trama enviada por el emisor es notificada y confirmada por el destino empleando tramas de asentimiento ACKs y, por lo tanto, se llevará a cabo un reenvío de tramas ante errores. Es apropiada para líneas con errores y evita la sobrecarga en niveles superiores. Este tipo de servicio es ofrecido por el nivel de enlace en redes inalámbricas ya que se busca comunicación fiable.
- *Servicios con conexión y con reconocimiento.* Las tramas no se envían independientemente, sino que son numeradas y antes del envío de las mismas es necesaria una fase de establecimiento de la conexión en la que se negocian los parámetros de comunicación entre emisor y receptor. Una vez establecida la conexión se envía la información y finalmente, tras ser recibida y confirmada, se cierra la conexión. Se usan tramas numeradas y cada trama se recibe una única vez. Estas serán confirmadas por el receptor y si hay errores se producirá el reenvío de los datos erróneos. Este tipo de servicios es utilizado en comunicaciones donde prima la fiabilidad como en redes inalámbricas.

5.2 Funciones del nivel de enlace

En el apartado anterior se han ido enunciando las principales características del nivel de enlace. Entre todas ellas la más importante es garantizar que los datos llegan sin errores y ordenados al receptor sea cual sea el medio físico empleado. Las funciones del nivel de enlace que permiten llevar a cabo los servicios indicados en el apartado anterior al nivel de red son las siguientes:

- *Iniciación y terminación de la comunicación.* El inicio, mantenimiento y la conclusión del envío de datos requiere una cantidad razonable de coordinación y cooperación entre las estaciones emisora y receptora. Se necesitan pues una serie de procedimientos para gestionar este intercambio. Mientras que la capa física proporciona solamente un servicio bruto de flujo de datos, la capa de enlace de datos intenta hacer el enlace físico seguro y proporciona medios para activar, mantener y desactivar el enlace.
- *Delimitación de tramas.* Los datos se envían en bloques que se denominan tramas (en inglés se emplea el término "frames"). El comienzo y el final de cada trama debe identificarse. Puesto que la capa física se encarga de la transmisión de cadenas de bits no estructuradas, una de las funciones del nivel de enlace será agrupar estos bits en tramas en las que se incluirán, además de los bits de información, campos de delimitación de la trama para saber cuál es el inicio y el final de la misma.

- *Direccionamiento de tramas.* En diversos tipos de redes como pueden ser las multipunto y las de difusión es necesario algún tipo de identificación de las estaciones que conforman la red. El direccionamiento de tramas permite identificar la estación origen y destino de una comunicación.
- *Control de errores.* Todos los errores en los bits generados por el sistema de transmisión se deberían corregir o, al menos, detectar. Entre las principales funciones que realizan el nivel de enlace están la detección y corrección de dichos errores. Para ello se implementan mecanismos para detectar cuándo ha habido errores en una trama y posteriormente corregir los mismos si es posible o interesa hacerlo. Hay que considerar que en ocasiones es más costosa la corrección de errores que realizar el reenvío de la trama errónea.
- *Control de flujo.* La estación emisora no debe enviar datos a una velocidad mayor que la estación receptora pueda absorber. Para ello, el nivel de enlace se encarga también del control de flujo, es decir, de la regulación del ritmo de envío de tramas desde el trasmisor con el fin de no congestionar a los receptores lentos.
- *Gestión y coordinación de la comunicación.* Se encarga de establecer quien tiene la responsabilidad en la gestión del enlace, es decir, si existe una estación que controla el enlace o si esa gestión es distribuida. Dependiendo del tipo de gestión del enlace, las entidades de nivel de enlace se encargarán de determinar como se coordinan las distintas estaciones de la red para realizar el envío de los datos a los respectivos destinos. Esto es especialmente importante en redes de difusión donde se requiere controlar qué equipo puede acceder al medio en cada instante.

En los siguientes apartados se realizará una descripción detallada de cada una de las funciones del nivel de enlace enumeradas en este apartado.

5.3 Iniciación y terminación de la comunicación

La función de *iniciación* comprende los procesos necesarios para activar el enlace e implica el intercambio de tramas de control con el fin de establecer la disponibilidad de las estaciones para transmitir y recibir información. También permite sincronizar numeraciones de tramas.

Las funciones de terminación se emplean para liberar los recursos ocupados después de la recepción/envío de la última trama. Para lograr el intercambio de información entre las estaciones a través del enlace se pueden considerar las fases siguientes:

- *Conexión del circuito.* Consiste en la obtención de un circuito físico que conecta las estaciones. En el caso de líneas conmutadas comprenden todos los procesos necesarios para el establecimiento del circuito (marcación, conmutación, etc.).
- *Establecimiento del enlace lógico.* Comprende los procesos necesarios para poder iniciar la transferencia de datos a través de un circuito físico ya establecido. La estación que quiera iniciar la transferencia deberá comprobar si la otra está preparada para recibir, si se disponen de los recursos necesarios, etc. Durante esta fase la información intercambiada por el enlace es información de control generada en ambas estaciones.
- *Transferencia de datos.* Comprende todos los procesos necesarios para lograr la transferencia de mensajes de una estación a otra.
- *Terminación.* Durante esta fase tiene lugar la liberación de los recursos asociados a la transferencia de información (buffers en las estaciones, etc.).
- *Desconexión del circuito.* Comprende los procesos de liberación del circuito. Las fases de conexión y desconexión del circuito son competencia del nivel físico, en el resto de fases interviene el nivel de enlace.

5.4 Delimitación de tramas

Como se ha indicado anteriormente la unidad de información con la que se trabaja a nivel de enlace es la trama. Una trama no es más que una agrupación de bits en una unidad de información mayor. Una función del nivel de enlace consiste en identificar el inicio y fin de la trama de datos. En general, las tramas presentan la siguiente estructura:

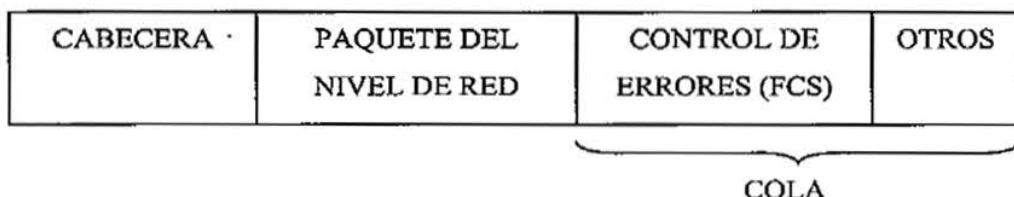


Figura 5-2. Estructura de una trama

- Los distintos campos representados en la Figura 5-2 son los siguientes:
 - Cabecera.* La cabecera identifica el comienzo de la transmisión, por lo tanto, delimita el inicio de la trama. En el campo de cabecera también se especifican las direcciones de las estaciones origen y destino. Asimismo, puede especificarse otra información de control como un código identificador de la trama, el tipo de trama (si es de datos, confirmación, u otro tipo de trama de control) o longitud de la trama. Cuando se requiere sincronización en la cabecera puede especificarse una secuencia especial de bits que es empleada por el receptor para sincronizarse en tiempo con el emisor.
 - Paquete de nivel de red.* En este campo se encapsula el paquete generado a nivel de red. En muchas ocasiones este campo presenta un tamaño máximo por lo que si el paquete a nivel de red sobrepasa este tamaño máximo será necesario realizar la fragmentación del paquete original. Puede ocurrir que el nivel de enlace envíe tramas propias de control (por ejemplo, ACK) que no tienen encapsuladas información del nivel de red.
 - Cola.* El emisor tiene la responsabilidad de agrupar la información en tramas que serán retransmitidas posteriormente a través del nivel físico como una secuencia de bits. El receptor recibe los bits individuales procedentes del nivel físico y ha de agruparlos en tramas, indicando donde empieza y acaba cada uno. El inicio de la trama se especifica en la cabecera y el final de la misma en la cola. Dentro de la cola se encuentra el siguiente campo:
 - Control de errores (FCS, "Frame Control Sequence". Secuencia de verificación de trama).* Este campo es utilizado para verificar si la transmisión se ha llevado a cabo correctamente, es decir, los bits que recibe el receptor no se han alterado respecto a los enviados por el emisor.

En terminología OSI, la cabecera, y la cola forman la PCI (Información de Control de Protocolo), el paquete de red es la SDU (Unidad de Datos de Usuario) y todo el conjunto de la trama es la PDU (Unidad de Datos de Protocolo).

Por lo tanto, una de las principales funciones del nivel de enlace es delimitar el inicio y final de las tramas y para ello puede emplearse algunas de las siguientes aproximaciones que serán descritas en los siguientes apartados:

- Delimitación temporal.
- Delimitación por numeración de caracteres.
- Delimitación por caracteres especiales

- Delimitación por bits especiales.
- Delimitación por códigos de línea especiales.

5.4.1 Delimitación temporal

Una posibilidad de delimitar una trama es lo que se denomina dejar “silencios” entre tramas, es decir, dejar de transmitir entre trama y trama. De esta forma, cuando se detecta un silencio será porque ha finalizado la transmisión de la trama.

Otra manera de realizar una delimitación temporal es suponer que todas las tramas presentan el mismo tamaño. En este caso no es necesario emplear cabeceras adicionales con el objetivo de la delimitación. El tamaño se asigna en función de una estimación de tiempo que tarda la trama en llegar al receptor. Este método es sensible a los retardos lo que provoca que se introduzcan errores. Este método se emplea en los entramados de TDM.

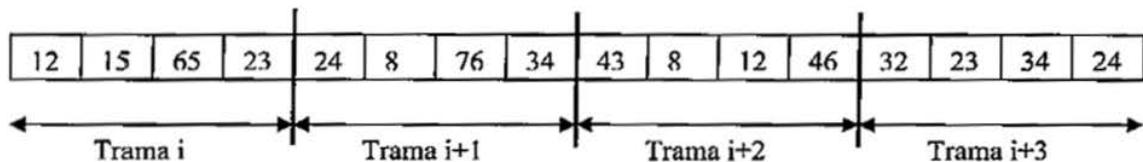


Figura 5-3. Delimitación temporal de tramas

5.4.2 Delimitación por número de caracteres

En este caso sí que es necesario emplear una cabecera para delimitar la trama. En concreto, en la cabecera se indica el número de bytes que posee la trama. Este método se emplea en protocolos muy simples y es sensible al ruido, ya que si hay variaciones en los datos de la cabecera debido al ruido se pierde la sincronización.

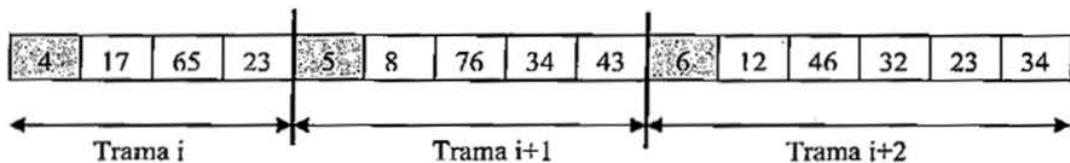


Figura 5-4. Delimitación por número de caracteres

5.4.3 Delimitación por caracteres especiales

Para realizar esta delimitación la trama a nivel de enlace debe presentar una cabecera y una cola. Tanto la cabecera como la cola almacenan caracteres especiales que no se emplean en los datos

transmitidos. Estos caracteres especiales suelen ser caracteres ASCII con código menor que 32 ya que se suponen que están reservados y no se emplean en la transmisión de datos alfabéticos (p. ej. caracteres ASCII SYN, DLE, STX, ACK, NACK, etc). Por lo tanto, este método se emplea sobretodo cuando los datos transmitidos son caracteres de texto. Además, si algún carácter especial se usa como dato, se usa un carácter de relleno que no es más que el carácter duplicado.

En la Figura 5-5 se muestra un ejemplo de trama delimitada por caracteres especiales. Primero se establece un carácter especial para advertir sobre caracteres especiales, lo que se conoce como secuencia de escape. En ASCII se usa el carácter DLE (Data Link Escape). Luego se emplea la secuencia de caracteres DLE, STX (start of transmission) para delimitar la trama. Los datos transmitidos son los caracteres A, B, C, D, F, G, DLE, C, F, H. Como dentro de los datos aparece el carácter DLE que en principio está reservado para definir secuencias especiales, éste carácter también se precede con otro DLE en la trama. El final de la trama se delimita con la secuencia de caracteres DLE, ETX (end of transmission). También es muy empleado como delimitador el código 7E en hexadecimal que se corresponde en binario con 01111110.

DLE	STX	A	B	C	D	F	G	DLE	DLE	C	F	H	DLE	ETX
-----	-----	---	---	---	---	---	---	-----	-----	---	---	---	-----	-----

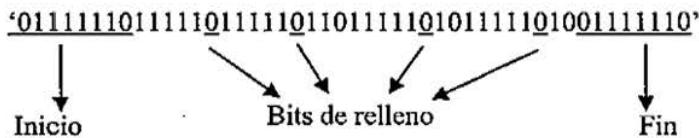
Figura 5-5. Delimitación por caracteres especiales.

5.4.4 Delimitación por bits especiales

Al igual que en el caso anterior, ahora también se van a emplear cabeceras y colas para la delimitación. La cabecera y la cola almacenan una secuencia de bits única que no se repetirá en los datos de la trama para indicar cuál es el inicio y final de la trama. Esta secuencia de bits o delimitadores es una combinación especial de bits llamada flag que indica el comienzo y final de la trama. El procedimiento es similar a la delimitación de tramas por caracteres especiales pero en este caso a nivel de bit.

Para evitar que la secuencia de bits especial pueda aparecer como datos se usan ‘bits de relleno’. Esto se observa en el siguiente ejemplo:

Supóngase que se desea transmitir los siguientes bits de datos: ‘1111111111101111101111110’. Además como delimitador de inicio y fin se va a considerar la secuencia de bits 01111110. Por lo tanto, la trama final que se enviaría, es decir, los datos junto con los delimitadores será:



Se observa que se introduce un ‘0’ en los datos cada vez que aparece la secuencia 011111. Cuando esta trama llegue al receptor comprobará cual es el valor del siguiente bit cada vez que recibe la secuencia 011111. Si ese bit es ‘1’ interpretará que es el fin o inicio de trama. Mientras que si es ‘0’ elimina el bit ya que se tratará de un bit de relleno.

5.4.5 Delimitación por códigos de línea especiales

En este caso se delimita el inicio y final de la trama empleando codificaciones distintas a las empleadas para codificar el resto de bits de la trama. En la Figura 5-6 se muestra una trama de datos cuyos bits están codificados empleando Manchester. En este ejemplo los bits de inicio y fin presentan una codificación distinta al resto de bits ya que no presentan una transición sino que el bit de inicio se codifica con un estado alto y el bit de final con un estado bajo.



Figura 5-6. Delimitación por caracteres especiales.

5.5 Direccionamiento de tramas

El direccionamiento de las tramas a nivel de enlace es empleado para identificar el origen y el destino de la trama de datos y permite identificar las estaciones o extremos de la comunicación que intercambian la información.

Los métodos de direccionamiento pueden clasificarse en primer lugar en función de la forma de especificar el origen y el destino en:

- *Implicito.* No es necesario especificar las direcciones de las estaciones origen y destino. Este tipo de direccionamiento se emplea en conexiones punto a punto en las que los datos enviados por el origen llegan al destino sin necesidad de especificar el destino ya que únicamente existe una posible estación destino.
- *Explícito.* En este caso sí que hay que indicar cual es la dirección origen y la destino en la trama. Cada equipo conectado a la red presenta una dirección única. Este tipo de direccionamiento se emplea en redes de difusión y en redes multipunto donde existen varias estaciones en la misma red que pueden ser el destino.

Asimismo, en función de cómo están conectadas las estaciones al canal los métodos de direccionamiento pueden clasificarse en:

- *Preselección.* Existe un dispositivo central que se encarga de seleccionar secuencialmente cada una de las estaciones de destino accesibles desde una estación.
- *Maestro único.* En este tipo de direccionamiento existe una estación central denominada master que se encarga de gestionar las comunicaciones. Sólo el equipo master tiene iniciativa de enviar datos. El resto de estaciones se le denominan esclavos y se encargan únicamente de responder a las peticiones del master.
- *Maestro múltiple.* Este caso es similar al anterior pero pueden existir varios equipos master en la red.

Además del direccionamiento individual o unicast de cada máquina suelen emplearse dos tipos de direcciones especiales:

- *Direcciones de broadcast.* Una trama enviada a la dirección de broadcast está destinada a todos los equipos de la red.
- *Direcciones de multicast.* Una trama enviada a la dirección de multicast está destinada a un subconjunto de los equipos que componen la red.

5.6 Detección y corrección de errores en tramas

Cualquier medio de transmisión se ve sometido a ruidos e interferencias externas que pueden causar que, al reducirse la relación señal-ruido, se produzcan errores en la transmisión. En general, se dice que se ha producido un error cuando el valor de un bit enviado por el receptor es alterado durante la transmisión. Básicamente, es posible distinguir dos tipos de errores:

- *Errores aislados.* Cuando, por ejemplo, el emisor envía una trama y un bit de la trama que se correspondía a un '1' binario es recibido por el receptor como un '0' (error aislado en un bit sin llegar a afectar a los bits vecinos). En resumen, los bits erróneos son independientes de los demás.
- *Errores a ráfagas.* Se dice que se ha producido un error en una ráfaga de longitud B cuando se recibe una secuencia de B bits en la que el primero, el último y cualquier número de bits intermedios son erróneos. Es lo que ocurre habitualmente cuando la señal con los datos se deteriora. En resumen, los bits erróneos se encuentran secuencialmente en el tiempo, además de estar agrupados.

Los protocolos de nivel de enlace deben ser capaces de asegurar la transmisión sin errores producidos por ruido o atenuaciones del medio físico. Los errores aislados son más fáciles de detectar y corregir, sin embargo, las ráfagas son más difíciles de tratar y también más frecuentes. Así, en cuanto a tratamiento de errores realizado por los protocolos de nivel de enlace se distingue dos posibilidades:

- *Detección de errores.* Consiste en únicamente detectar si los bits recibidos son los que realmente quiso enviar el emisor. Normalmente se asocia la detección a alguna técnica de reenvío de tramas, de forma que si el receptor determina que se ha producido algún error, se encarga de notificar esta situación al emisor y este último deberá reenviar la trama de datos que había llegado de forma incorrecta al receptor. Este tipo de aproximaciones se emplean cuando el tiempo de reenvío es inferior al tiempo de chequeo y corrección, o cuando la corrección es demasiado compleja.
- *Corrección de errores.* Consiste en, no solo detectar si se ha producido algún error o no, sino también determinar cuáles son los bits de la trama que han llegado alterados al receptor. Una vez detectados, el propio receptor modifica los datos erróneos de forma que no es necesario que el emisor realice una retransmisión de la trama original. Se emplea detección y corrección de errores simultáneamente en los medios físicos donde el reenvío produce retardos elevados. Hay que tener en cuenta que la detección/corrección de errores de ráfagas de bits puede ser compleja y supone un tiempo de proceso a considerar.

Por lo tanto, el proceso de control de errores se lleva a cabo de forma combinada tanto en el emisor como en el receptor. Así, el emisor añade en la trama información para el control de errores. Normalmente, este campo se denomina FCS (Frame Control Sequence). Esta información llegará al receptor el cual se encarga de interpretar la información recibida con el fin de detectar errores y corregirlos si fuera posible.

Se van a describir dos métodos para la detección de errores: *comprobación de paridad* y *comprobación de redundancia cíclica*. Por último, para la corrección de errores se describirá la

utilización de otros dos métodos *códigos Hamming* y *códigos Reed-Solomon*. Mientras que los códigos Hamming sólo se pueden emplear para corregir errores aleatorios, los códigos Reed-Solomon pueden además corregir los errores de ráfaga.

5.6.1 Comprobación de paridad

Este método constituye la forma más sencilla de detectar errores y básicamente consiste en añadir un único bit al final de la trama de datos. El valor de este bit se determina de tal forma que el bloque de bits resultante tenga un número par de unos (paridad par) o un número impar de unos (paridad impar).

Supóngase que el emisor debe realizar la transmisión de la siguiente trama de datos: 1011. Si se emplea paridad par, el bloque de datos que se enviaría finalmente será: 10111. Se observa que en este caso se añade un bit '1' para obtener un número par de unos en el bloque final de bits obtenido. Sin embargo, si se empleara paridad impar el bloque final sería: 10110. Considerese ahora que se envía el bloque de datos anterior empleando paridad par. En este caso le llegará al receptor el bloque de bits 10111 (si no se ha producido ningún error durante la transmisión). El receptor volverá a calcular la paridad y si el bit de paridad calculado teóricamente coincide con el bit de paridad introducido por el emisor significará que no se ha producido ningún error.

Este tipo de detección de errores presenta la ventaja de que no introduce excesiva información de control adicional para la detección y es muy sencillo de implementar. Sin embargo, este método no detectará el error si dos o cualquier número par de bits se alteran debido a algún error en la transmisión, por ejemplo un error de ráfagas.

Existen extensiones de la técnica de paridad que emplean varios bits de paridad asociados a diferentes bloques de bits de los datos.

5.6.2 Comprobación de redundancia cíclica (CRC)

Los códigos de redundancia cíclica son una secuencia de bits que se colocan en el campo control de errores de la trama (FCS) y constituyen una de las formas más habituales y potente de detección de errores. Se basa en la división binaria en módulo 2 e inclusión de un resto en el campo control de errores de la trama de datos. En la Figura 5-7 se muestra la aritmética básica para la aplicación del método de generación y comprobación de errores mediante CRC. La trama enviada es teóricamente la diferencia entre los datos, D, y el resto, R, obtenido al dividir D entre un divisor conocido G. Ese divisor es conocido tanto por el emisor como el receptor. Cuando la trama llegue al receptor, éste dividirá la trama entre ese divisor conocido y si el resto de la división es nulo supondrá que no ha habido errores.

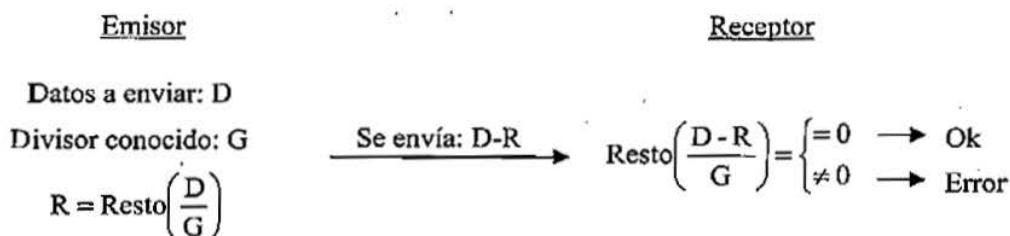


Figura 5-7. Aritmética para el cálculo y comprobación de CRC.

Sin embargo, a diferencia del esquema representado en la Figura 5-7 los datos realmente transmitidos no son números sino un conjunto de bits que pueden representarse en forma de polinomios de manera que los bits constituyan los coeficientes del polinomio. Por ejemplo, la secuencia de bits 10010110 puede representarse con el polinomio $1x^7+0x^6+0x^5+1x^4+0x^3+1x^2+1x^1+0x^0 = x^7+x^4+x^2+x$.

Considerando los datos representados en forma de polinomio y que el divisor, G, habitualmente denominado polinomio generador, es el polinomio $G(x)$ de grado r , el esquema de la Figura 5-7 sería el que se muestra en la Figura 5-8.

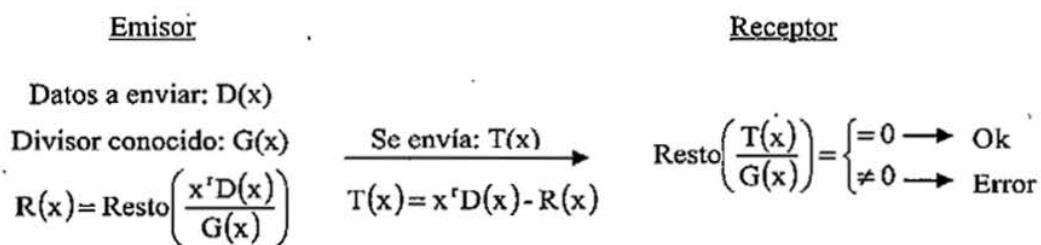


Figura 5-8. Proceso de cálculo y comprobación de CRC considerando los datos representados en forma de polinomio.

Se observa como los datos a enviar $D(x)$ son multiplicados por x^r . Esto equivale a poner r ceros tras el valor binario de $D(x)$. Por ejemplo, considerando los datos a enviar 10010110, entonces $D(x) = x^7+x^4+x^2+x$. Si $r=2$, $x^r D(x) = x^9+x^6+x^4+x^3$. Además, la resta $D(x)-R(x)$ se hace en módulo 2, sin acarreo. Así, a la hora de calcular $T(x)$ no se mezcla $D(x)$ y $R(x)$, y los datos y el resto irán por separado (véase Figura 5-9.a).

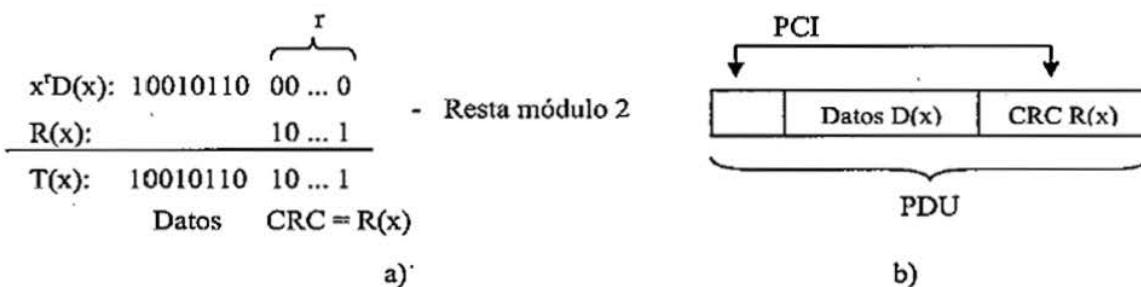


Figura 5-9. a) Combinación de datos, $D(x)$ y $R(x)$. b) Trama de datos

En la trama de datos (PDU de nivel de enlace), el emisor colocará los bits $D(x)$ como información útil, y el resto $R(x)$ como un campo de PCI de nivel de enlace. Es común denominar al campo $R(x)$ como "CRC" (véase Figura 5-9.b).

El proceso seguido para la generación del CRC y la posterior comprobación de errores en el receptor se puede sintetizar en los siguientes pasos:

- En el emisor:
 - Para calcular el resto es necesario disponer de un divisor, o polinomio generador $G(x)$.
 - A los datos, $D(x)$, que se quieren transmitir se le añaden r ceros por la derecha mediante el producto $x^r D(x)$.

- Se divide $x^r D(x)$ entre el divisor, $G(x)$ para determinar el resto, $R(x)$, donde r es el grado del polinomio generador. $R(x)$ será de grado $r-1$.
 - El resto de la división $R(x)$ es el CRC.
 - Se calcula $T(x)$. Para ello se calcula la siguiente diferencia: $x^r D(x) - R(x)$. Se envía $T(x)$.
- En el receptor:
 - Se divide $T(x)$ entre el polinomio generador, $G(x)$, y se calcula el nuevo resto, $R'(x)$.
 - Si resto, $R'(x)$, es cero la trama no presenta errores. Si el resto, $R'(x)$, no es cero será porque la trama presenta errores.

Hay que tener en cuenta que en la división binaria en módulo 2 no hay acarreos, y una cifra se puede dividir entre otra simplemente cuando tienen igual número de bits sin contar los ceros a la izquierda. Por ejemplo, 100 o 110 son divisibles entre 101, mientras que 011 o 010 no lo son.

Se observa que es necesario que tanto el emisor como el receptor conozcan el valor del polinomio generador, $G(x)$. Existen algunos valores estándar para $G(x)$ como son los siguientes:

- CRC-12. Se utiliza para la transmisión de secuencias de caracteres de 6 bits y generará un CRC de 12. El valor de $G(x)$ en este caso es $x^{12} + x^{11} + x^3 + x^2 + x + 1$.
- CRC-16 y CRC-CCITT. Ambos son habituales para los caracteres de 8 bits son empleados en Estados Unidos y en Europa respectivamente y sus valores son $x^{16} + x^{15} + x^2 + 1$ y $x^{16} + x^{12} + x^5 + 1$.
- CRC-32. Es empleado en algunas normas de transmisión síncrona sobre enlaces punto a punto y su valor es $x^{32} + x^{26} + x^{23} + x^{22} + x^{16} + x^{12} + x^{11} + x^{10} + x^8 + x^7 + x^5 + x^4 + x^2 + x + 1$.

En general, empleando el método de comprobación de código de redundancia es posible detectar los siguientes errores:

- Se detectan todos los errores de un bit si el polinomio generador $G(x)$ tiene más de un término distinto de cero.
- Se detectan todos los errores dobles, siempre que $G(x)$ tenga al menos un factor con tres términos.
- Se detecta cualquier número impar de errores, siempre que $G(x)$ contenga el factor $x+1$.
- Se detecta cualquier ráfaga de errores con longitud menor o igual a la longitud del CRC.

Ejercicio 5.1

Calcular la trama de datos a enviar $T(x)$ si el emisor quiere enviar la cadena de datos '10100111' usando como polinomio generador $G(x) = x^2 + x + 1$. A continuación, comprobar si se han producido errores en los datos recibidos por el receptor.

Paso 1. Calcular $x^r D(x)$ para ello habrá que añadir tantos ceros a la derecha de los datos a enviar como grado r tiene el polinomio generador $G(x)$ ($r=2$).

$$G(x) = x^2 + x + 1 \rightarrow '111'. \quad (r=2)$$

$$D(x) = x^7 + x^5 + x^2 + x + 1 \rightarrow '10100111'$$

$$x^r D(x) = x^9 + x^7 + x^4 + x^3 + x^2 \rightarrow '1010011100'$$

Paso 2. En este paso hay que calcular el resto $R(x)$. Para ello, se divide $x^r D(x)$ entre $G(x)$ usando división módulo 2. En este caso el resto obtenido es '11' (grado 1).

$$\begin{array}{r}
 1\ 0\ 1\ 0\ 0\ 1\ 1\ 1\ 0\ 0 \\
 \underline{1\ 1\ 1} \\
 0\ 1\ 0\ 0 \\
 \underline{1\ 1\ 1} \\
 0\ 1\ 1\ 0 \\
 \underline{1\ 1\ 1} \\
 0\ 0\ 1\ 1\ 1 \\
 \underline{1\ 1\ 1} \\
 0\ 0\ 0\ 1\ 0\ 0 \\
 \underline{1\ 1\ 1} \\
 0\ 1\ 1
 \end{array}$$

Paso 3. Calcular $T(x)$. Para ello, a la cadena resultante del paso 1 se le substrae el resto $R(x)$ obtenido de calcular la división del paso 2. Como se mostró en la Figura 5-9 esto equivale a combinar los datos y el resto, $R(x) = CRC$. Por lo tanto, en este caso $T(x) = x^9+x^7+x^4+x^3+x^2+x+1 \rightarrow '101001110011'$.

Paso 4. En este caso se va a describir la comprobación que realiza el receptor con el objetivo de determinar si la trama presenta o no errores a partir de los datos y CRC recibido. Para realizar dicha comprobación se divide $T(x)$ entre $G(x)$ y se observa el valor del resto. En este caso el resto es nulo, por lo que no se habrá producido ningún error al enviar los datos.

$$\begin{array}{r}
 1\ 0\ 1\ 0\ 0\ 1\ 1\ 1\ 1 \\
 \underline{1\ 1\ 1} \\
 0\ 1\ 0\ 0 \\
 \underline{1\ 1\ 1} \\
 0\ 1\ 1\ 0 \\
 \underline{1\ 1\ 1} \\
 0\ 0\ 1\ 1\ 1 \\
 \underline{1\ 1\ 1} \\
 0\ 0\ 0\ 1\ 1\ 1 \\
 \underline{1\ 1\ 1} \\
 0\ 0\ 0
 \end{array}$$

5.6.3 Corrección de errores empleando códigos de bloque

La mayoría de los protocolos de nivel de enlace implementan alguna técnica de detección de errores como la mostrada en el apartado anterior basada en códigos de redundancia cíclica. Sin embargo, existen aplicaciones en las cuales puede ser más conveniente realizar una corrección de los datos recibidos que solicitar al emisor que vuelva a enviar las tramas erróneas. Supóngase, por ejemplo, la transmisión de información entre satélites. En este caso, el tiempo que tarda una trama en propagarse desde el emisor hasta el receptor puede ser muy elevado. Además, debido a las características del medio de transmisión, la tasa de errores también puede ser bastante elevada, lo que provocará un gran número de retransmisiones. Una alternativa sería que el receptor haga uso del conjunto de bits recibidos, no solo para detectar la existencia o no de errores, sino también para corregirlos. En este apartado se va a describir una aproximación basada en códigos de bloque que permitirá al receptor detectar la existencia de errores y poder reconstruir la trama que originalmente había enviado el emisor.

En las técnicas de corrección de errores basadas en códigos de bloque no se emplea un campo FCS (como se hacía en la detección basada en CRC) sino que la información adicional necesaria para la detección y corrección se intercala entre los bits de datos que el emisor desea hacer llegar al receptor. Cuando la trama de datos llega al receptor, éste se encarga de separar los bits de datos de la información adicional intercalada para enviar sólo al nivel de red la parte de datos. Para realizar la comprobación y corrección de errores empleando códigos de bloque es necesario definir los siguientes conceptos previos:

- *Palabras de datos.* A partir de un conjunto de k bits es posible obtener un total de 2^k palabras de datos. Estos van a ser los bloques de datos que en principio desea enviar el emisor.
- *Palabras-código.* Como se ha indicado anteriormente, para la corrección de errores se intercalará un conjunto de bits (un total de r bits redundantes) a las palabras de datos. Por lo tanto, se enviarán finalmente $n=k+r$ bits y a partir de estos bits es posible generar un total de 2^n palabras código. Las palabras-código serán los datos que finalmente se enviarán al medio.
- *Distancia Hamming entre dos palabras código.* La distancia Hamming entre dos palabras-código se define como el número de bits en que difieren las dos palabras. Por ejemplo, la distancia de Hamming $d(v_1, v_2)$ entre las palabras $v_1=011101$, $v_2=010011$ será de 3.
- *Distancia Hamming de un código.* Dado un conjunto de palabras código correctas, se define la distancia Hamming de un código como la menor de todas las distancias Hamming entre las palabras-código.

Cuando se emplea corrección por códigos de bloque, en lugar de transmitir cada bloque de k bits (palabras de datos), se asigna cada secuencia de entrada a una única palabra-código de n bits. En la Figura 5-10 se muestran los pasos seguidos en la corrección, que se pueden resumir en los siguientes:

- En el emisor, a través de un proceso de codificación, se asigna una palabra-código a la palabra de datos que se desea enviar. Por lo tanto, cada palabra de datos de k bits tendrá asociada una palabra-código de n bits.
- Cuando el receptor recibe una palabra-código sin errores, es decir, idéntica a la palabra-código original, el decodificador generará el bloque de datos correspondiente, una vez eliminados los bits añadidos.
- Para ciertos patrones de error, es posible que el decodificador detecte y corrija esos errores. Así, aunque el bloque de datos recibido difiera de la palabra-código transmitida por el emisor, el decodificador será capaz de extraer la palabra de datos original.
- Para otros patrones de error el decodificador no será capaz de realizar la corrección pero podrá detectar los errores. En este caso, el receptor notifica esta situación al nivel de red.

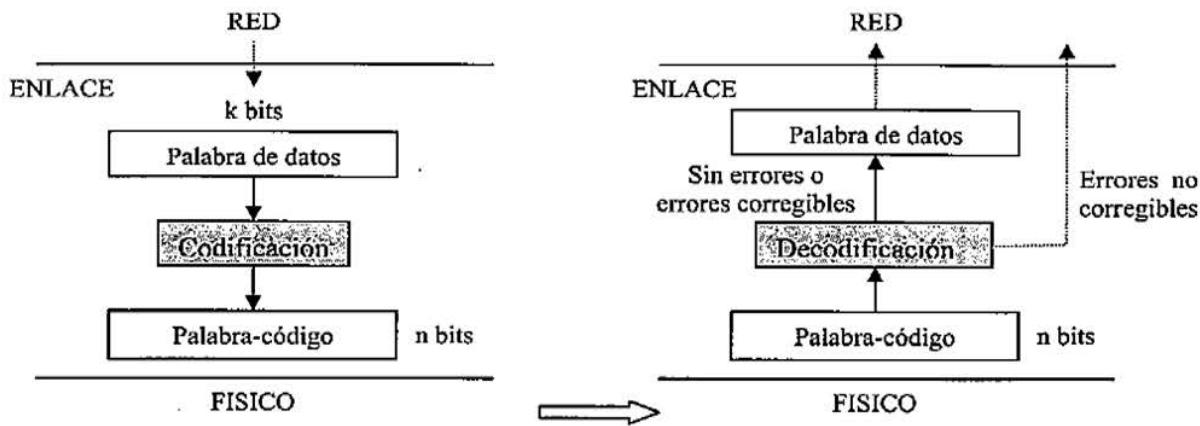


Figura 5-10. Corrección de errores empleando palabras código.

Por ejemplo, para $k=2$ y $n=5$ se pueden realizar las siguientes asignaciones:

Bloque de datos	Palabra-código
00	00000
01	00111
10	11001
11	11110

Si el receptor recibe el patrón de bits 00100, ésta no será una palabra-código válida, por lo que el receptor detecta un error. A partir de la palabra-código errónea recibida no es posible asegurar qué bloque de datos fue originalmente enviado ya que el ruido puede haber alterado 1, 2 o más bits. Sin embargo, para convertir la palabra-código 00100 en 00000 (palabra-código válida) solo se necesita alterar un bit. Además, para convertirla en cualquier otra palabra se necesitará alterar más bits. Por lo tanto, puede suponerse que el bloque de datos originalmente enviado es 00. De esta manera, si se recibe una palabra código inválida, entonces se selecciona la palabra código válida a menor distancia Hamming. Por lo tanto, tanto el emisor como el receptor deben conocer todas las palabras código y su correspondencia con palabras de datos. Estas palabras-código dependerán del tipo de paridad que se escoga.

Supóngase ahora que se recibe la palabra 01010. Esta palabra-código tiene como menor distancia a cualquier palabra-código 2. Además, se encuentra a la misma distancia de 00000 y 11110 (distancia = 2 como se ha indicado anteriormente). En este caso el receptor no tendría forma de elegir entre las dos alternativas y no se podría corregir el error. Así, este código puede corregir errores simples pero no errores dobles. Dicho de otra forma, la distancia del código es 3. Por tanto, un error en un bit dará lugar a una palabra-código inválida que está a distancia 1 de la palabra-código válida, pero al menos a distancia 2 de cualquiera de las otras palabras código.

A partir de las definiciones anteriores se puede afirmar que, dado un código binario de distancia Hamming d , es posible detectar errores de $d-1$ bits y corregir errores de $(d-1)/2$ bits.

Un caso específico de detección y corrección de errores es el Algoritmo de Hamming. El algoritmo de Hamming permite realizar la corrección de errores en un bit y para ello deberá cumplirse que $k+r+1 \leq 2^r$. Esta relación permite obtener cuantos bits redundantes, r , es necesario introducir para una palabra de datos con un total de k bits. A continuación se muestran los pasos que se han de seguir para la generación de palabras código que permitan realizar la posterior corrección de errores en 1 bit:

1. Determinar la cantidad de bits redundantes, r , que son necesarios. Para ello se hace uso de la relación $k+r+1 \leq 2^r$.
2. Numerar los bits de la palabra-código de izquierda a derecha, en orden ascendente, desde $i=1 \dots n$.
3. Todos los bits cuya posición es potencia de dos se utilizan como bits redundantes (posiciones 1, 2, 4, 8, 16, 32, 64, etc.).
4. Los bits del resto de posiciones son utilizados como bits de datos (posiciones 3, 5, 6, 7, 9, 10, 11, 12, 13, 14, 15, 17, etc.).
5. El valor de los bits redundantes se va a determinar como bits de paridad. Cada bit de paridad se obtiene calculando la paridad de alguno de los bits de datos. La posición del bit de paridad determina la secuencia de los bits que alternativamente comprueba y salta, tal y como se explica a continuación:
 - Posición 1: salta 1, comprueba 1, salta 1, comprueba 1, etc.
 - Posición 2: comprueba 1, salta 2, comprueba 2, salta 2, comprueba 2, etc.
 - Posición 4: comprueba 3, salta 4, comprueba 4, salta 4, comprueba 4, etc.
 - Posición 8: comprueba 7, salta 8, comprueba 8, salta 8, comprueba 8, etc.
 - Posición 16: comprueba 15, salta 16, comprueba 16, salta 16, comprueba 16, etc.
 - Y así sucesivamente.

En otras palabras, el bit de paridad de la posición 2^k comprueba los bits en las posiciones que tengan al bit k en su representación binaria.

Ejercicio 5.2

Dada la palabra de datos de 7 bits "0110110" determinar la palabra-código asociada teniendo en cuenta que se emplea paridad par.

Paso 1. Determinar el conjunto de bits necesarios. Para ello, hay que aplicar la relación $k+r+1 \leq 2^r$. En este caso $k=7$, por lo que $r+8 \leq 2^r$. Si $r=1 \rightarrow 9 > 2^1$, Si $r=2 \rightarrow 10 > 2^2$, Si $r=3 \rightarrow 11 > 2^3$, Si $r=4 \rightarrow 12 \leq 2^4$ (para $r=4$ se cumple la relación $k+r+1 \leq 2^r$). Por lo tanto, es necesario introducir 4 bits redundantes.

Pasos 2, 3 y 4. En la Figura 5-11 se muestra una tabla en la que se representa la palabra de datos original y se ha identificado la posición de los bits de datos, d_i , y la de los bits redundantes, r_i .

Paso 5. Se determinan los valores de los bits redundantes como bits de paridad. En la tabla de la Figura 5-11 se ha introducido una fila para cada bit redundante en la que se observa, no solo el valor de la paridad, sino también qué bits se utiliza para su cálculo. En este caso todos los bits redundantes son nulos.

Por lo tanto, la nueva palabra código será 00001100110. En la tabla de la Figura 5-12 se describe el proceso de detección de errores cuando la trama llega al bit sin alteración en ninguno de sus bits y en la Figura 5-13 la palabra código que llega al receptor tiene el último bit alterado.

	r_1	r_2	d_1	r_3	d_2	d_3	d_4	r_4	d_5	d_6	d_7
Palabra de datos:			0		1	1	0		1	1	0
r_1	0		0		1		0		1		0
r_2		0	0				1	0		1	0
r_3				0	1	1	0				
r_4									0	1	1
Palabra-código:	0	0	0	0	1	1	0	0	1	1	0

Figura 5-11. Ejemplo de cálculo de palabra-código.

	r_1	r_2	d_1	r_3	d_2	d_3	d_4	r_4	d_5	d_6	d_7	Prueba paridad	Bit de paridad	
Palabra-código recibida	0	0	0	0	1	1	0	0	1	1	0	0		
r_1	0		0		1		0		1		0	Ok	0	
r_2		0	0			1	0			1	0	Ok	0	
r_3				0	1	1	0					Ok	0	
r_4									0	1	1	0	Ok	0

Figura 5-12. Comprobación de errores. La palabra código llega sin errores.

	r_1	r_2	d_1	r_3	d_2	d_3	d_4	r_4	d_5	d_6	d_7	Prueba paridad	Bit de paridad	
Palabra-código recibida:	0	0	0	0	1	1	0	0	1	1	1	0		
r_1	0		0		1		0		1		1	Error	1	
r_2		0	0			1	0			1	1	Error	1	
r_3				0	1	1	0					Ok	0	
r_4									0	1	1	1	Error	1

Figura 5-13. Comprobación de errores. La palabra código llega con errores.

Hasta ahora se ha descrito como generar las palabras código. Sin embargo, también es necesario definir la manera en que el receptor, una vez detectado el error, es capaz de determinar cuál es el bit que ha llegado alterado. En el caso de la Figura 5-13 ese bit es el undécimo (el número once empezando a contar los bits de izquierda a derecha). Para ello, se hace uso de un contador c que se inicializa a 0. Si se analiza la tabla de la Figura 5-13 la paridad que se debe obtener tras la llegada del mensaje sin errores debe ser siempre 0 (por cada fila), pero en el momento en que ocurre un error esta paridad cambia a 1. Se observa que en la fila en que el cambio no afectó la paridad es cero y llega sin

errores. Por lo tanto, si un determinado bit redundante tiene la paridad errónea (se comprueba que debería ser 1), se incrementa la variable c en el valor de la posición del bit redundante (nótese que los bits se habían numerado de izquierda a derecha comenzando por 1). Si, una vez verificados todos los bits redundantes $c=0$, entonces no habrá error. Si $c\neq0$, el valor de c indicará cual es la posición del bit redundante.

Para el caso mostrado en la Figura 5-13 los bits redundantes con paridad errónea eran r_1 , r_2 y r_4 que se corresponden con las posiciones 1, 2 y 8. Por lo tanto $c=1+2+8=11$. Por lo tanto, como se había indicado en un principio, el bit erróneo es el undécimo.

5.6.4 Los códigos Reed-Solomon

Los códigos Reed-Solomon (RS) son un sistema de corrección de errores que consiste en dada una secuencia de bits datos llamada bloque de información, dividir éste en bloques y a cada uno de ellos añadirle un cierto número de símbolos de redundancia. Estos símbolos adicionales contienen la información necesaria y suficiente para poder localizar la posición de los símbolos de datos erróneos y corregirlos.

Para realizar la comprobación y corrección de errores empleando códigos Reed-Solomon necesario definir los siguientes conceptos:

- *Símbolo de un código Reed-Solomon.* Un símbolo es un conjunto de m bits.
- *Borrado:* Símbolo erróneo cuya localización se conoce.
- *Error de símbolo:* Cuando al menos 1 bit del símbolo es erróneo.
- *Código Reed-Solomon.* Se puede describir un código por la tupla $RS(n,k)$ donde k es el número de símbolos del bloque de información y n es el tamaño del mensaje completo una vez añadidos los símbolos de redundancia.
- *Palabras Reed-Solomon.* Son cada uno de los bloques individuales de información más los símbolos de redundancia añadidos a cada bloque. Así una palabra estará formada por $n=k+r$ símbolos y donde $r=n-k$ siendo r la cantidad de símbolos redundantes añadidos.

Los códigos Reed-Solomon permiten corregir hasta un máximo de $t=r/2$ símbolos, donde t es como se denota a los símbolos erróneos cuya localización se desconoce. Además, Reed-Solomon es capaz de corregir hasta $2t$ borrados.

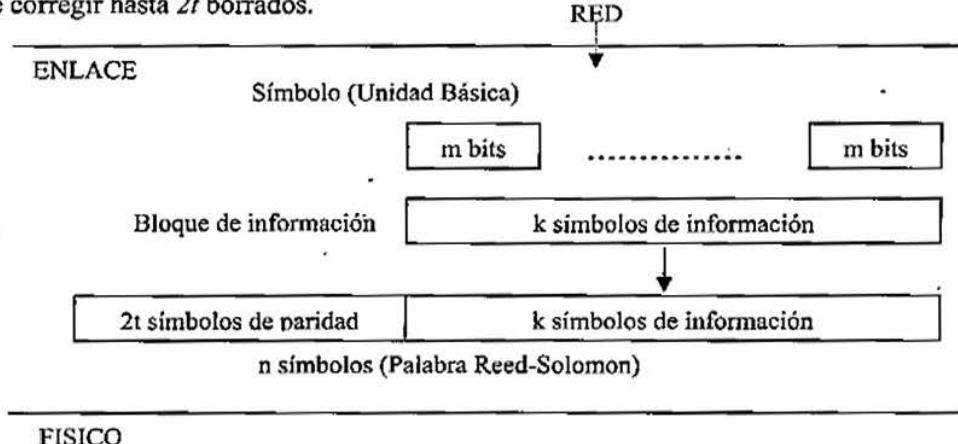


Figura 5-14. Ejemplo de formación de una palabra de código de RS.

Ejercicio 5.3

Supóngase que un MODEM-ADSL recibe en el nivel de enlace 255 símbolos, y se sabe que este dispositivo maneja un código Reed-Solomon RS(255, 239), ¿cuántos bits es capaz de corregir sabiendo que cada símbolo está formado por 1 byte?

Si RS(255,239) entonces $n=255$ y $k=239$

$$r=n-k=255-239=16 \text{ símbolos}$$

$$t=r/2=16/2=8 \text{ símbolos}$$

Sí 1 símbolo son 8 bits, entonces 8 símbolos son 8 bytes.

Por lo tanto, se puede deducir que se pueden corregir hasta 64 bits de paridad.

A partir del ejemplo anterior se puede deducir que la corrección con códigos Reed-Solomon depende del tamaño del símbolo empleado. Si en el caso anterior, el símbolo estuviera constituido por 1 sólo bit, entonces se tendría que sólo podría corregirse errores en 16 bits cada uno de los cuales estaría en un símbolo (bloque unitario de datos) distinto. Esto determinaría que un código del tipo RS(255, 239) en el peor de los casos permitiría corregir únicamente 16 bits, ya que la unidad más pequeña para 1 símbolo es 1 bit y éste es indivisible.

Además, el ejemplo anterior también ilustra el comportamiento de los códigos Reed-Solomon para detectar errores aleatorios, por ejemplo algún error en algún símbolo. Ahora con otro ejemplo se va a comprobar si los códigos RS se presentan tan eficaces para corregir los errores en ráfaga.

Ejercicio 5.4

Supóngase que un MODEM-ADSL recibe en el nivel de enlace 255 símbolos, y se sabe que este dispositivo maneja un código Reed-Solomon RS(255, 251), ¿cuántos bits es capaz de corregir sabiendo que cada símbolo está formado por 1 byte y que se produce un error en ráfaga que afecta a 3 símbolos?

Si RS(255,251) entonces $n=255$ y $k=251$

$$r=n-k=255-251=4 \text{ símbolos}$$

$$t=r/2=4/2=2 \text{ símbolos se pueden corregir}$$

$2 < 3$ entonces no queda corregido el error de ráfaga

Como se ha observado en este ejemplo, no es posible la corrección de errores cuando se produce un error de ráfaga lo suficientemente grande como para afectar a más de los t símbolos que puede corregir el código. Sin embargo, se puede hacer que la corrección de errores con códigos Reed-Solomon mejore combinándolo con técnicas de *interleaved*.

La técnica de *interleaved* consiste en realizar un mezclado de los símbolos de las distintas palabras RS que se van a enviar en una transmisión y de esta manera obligar a que los errores que se produzcan parezcan aleatorios en vez de formar parte de una ráfaga. En definitiva, el sistema de transmisión, si emplea técnica de interleaved, en vez de transmitir cada palabra original RSi como la mostrada en la figura 4, crea una nueva palabra RS_n de tamaño idéntico a cada RSi pero en este caso formada por símbolos de éstas entremezclados. Esto tiene la ventaja de que cuando ocurre una ráfaga de errores y las palabras RSi son reordenadas en el receptor, los errores quedan repartidos entre las distintas palabras RSi. De este modo, los errores parecerán aleatorios y en ningún momento quedarán localizados en la misma palabra RSi y aumentará la probabilidad de que los errores sean corregibles.

Los códigos RS son los códigos de error más empleados en aplicaciones comerciales, así destaca su uso en la grabación de CDs, DVDs, Blue-ray, en la transmisión de datos en tecnologías como DSL y WiMAX, o en servicios de broadcast como DVB.

5.7 Medición de la tasa de errores

En el capítulo 4 se describieron los medios de transmisión que se emplean habitualmente en redes de computadores. Independientemente de las características de los medios de transmisión, éstos sufren una determinada atenuación dependiendo de la frecuencia de transmisión y de la longitud del medio. Además, en cualquier medio existirá un determinado ruido que puede provocar errores que harán que uno o varios bits se vean alterados durante el proceso de transmisión.

Esa tasa de error, que a menudo también se denomina como BER (Bit Error Rate) define la probabilidad de que un bit recibido sea erróneo y depende de diversos factores. Dentro de estos factores cabe mencionar algunos como el tipo del medio físico empleado, el entorno del medio, la velocidad de transmisión, la calidad del servicio del medio físico o el horario en que se realiza la comunicación. El BER puede calcularse como la relación entre el número de errores y el número de bits totales transmitidos; y presenta valores típicos comprendidos entre 10^{-3} y 10^{-6} .

Asociado a los errores que puede dar lugar la transmisión de tramas a nivel de enlace se pueden definir las siguientes probabilidades:

- P_b : Probabilidad de que un bit recibido sea erróneo (BER).
- P_1 : Probabilidad de que una trama llegue sin errores.
- P_2 : Probabilidad de que, utilizando un algoritmo para la detección de errores, una trama llegue con uno o más errores no detectados.
- P_3 : Probabilidad de que, utilizando un algoritmo para la detección de errores, una trama llegue con uno o más errores detectados y sin errores no detectados.

Supóngase que P_3 es cero, es decir, no se aplica ningún método para la detección de errores y que además, todos los bits transmitidos (un total de N bits) tienen una probabilidad de error igual a P_b . En este caso las probabilidades P_1 y P_2 serán respectivamente:

$$P_1 = (1-P_b)^N \quad (5.1)$$

$$P_2 = 1 - P_1 \quad (5.2)$$

Por lo tanto, se puede concluir que:

- La probabilidad de que una trama llegue sin ningún bit erróneo disminuye al aumentar la probabilidad de que un bit sea erróneo.
- La probabilidad de que una trama llegue sin errores disminuye al aumentar la longitud de la misma.

Ejercicio 5.5

Calcular la probabilidad de que una trama de 4 bits llegue al receptor sin errores si se sabe que la probabilidad de que se produzca un error en un bit es de 0,4.

$$P_b = 0,4$$

$$P_1 = (1-0,4)^N = (1-0,4)^4 = 0,129$$

Capítulo 6. CONTROL DE FLUJO EN EL NIVEL DE ENLACE

Una de las funciones implementadas habitualmente por los protocolos de nivel de enlace es el control de flujo de la transmisión. Este tipo de protocolos permiten garantizar que un equipo receptor no llegue a verse colapsado por recibir una gran cantidad de información procedente de otro equipo emisor. A lo largo de este capítulo se describirá el protocolo de parada y espera así como el de ventana deslizante y se mostrará cómo pueden utilizarse para controlar el flujo de datos que se intercambia en una comunicación.

CAPÍTULO 6

6-1 Introducción

6-2 Protocolo de parada y espera

- Parada y espera en ausencia de errores
- Parada y espera considerando que se producen errores
- Parada y espera con confirmaciones negativas
- Protocolo bilateral de parada y espera
- Análisis de las prestaciones del protocolo de parada y espera

6-3 Protocolo de ventana deslizante

- Características del protocolo de ventana deslizante
- Análisis de las prestaciones del protocolo de ventana deslizante

OBJETIVOS

- Dar a conocer el objetivo de realizar un control de flujo en las arquitecturas de red
- Mostrar el funcionamiento del protocolo de parada espera con y sin errores
- Determinar la eficiencia del protocolo de parada y espera
- Describir el funcionamiento del protocolo de ventana deslizante
- Determinar la eficiencia del protocolo de ventana deslizante

TERMINOLOGÍA – PALABRAS CLAVE

- | | | |
|--------------------|----------------------|-------------------------|
| • Nivel de enlace | • Ventana deslizante | • Piggyback |
| • Control de flujo | • Asentimientos | • Cálculo de eficiencia |
| • Parada y espera | • ACK/NACK | • Repetición selectiva |

6.1 Introducción

Como se indicó en el capítulo anterior, una de las funciones que habitualmente implementa el nivel de enlace de las arquitecturas de red es el control del flujo de la transmisión de datos entre el emisor y el receptor. El control de flujo se emplea para evitar que el receptor sea saturado por el envío de datos desde un emisor. Cualquier dispositivo de recepción tiene una cantidad limitada de memoria para almacenar datos y una velocidad limitada para procesarlos. El control de flujo permitirá restringir la cantidad de datos que el emisor envía al receptor. Esto permitirá que la memoria reservada en el receptor no llegue a desbordarse mientras éste se encuentra procesando datos que le habían llegado con anterioridad.

A lo largo de este capítulo se describirán 2 métodos para el control de flujo implementados por el nivel de enlace:

- Parada y espera.
- Ventana deslizante.

6.2 Protocolo de parada y espera

6.2.1 Parada y espera en ausencia de errores

En este método, el emisor espera un reconocimiento después de cada trama que se envía. Sólo se podrá enviar la siguiente trama cuando se recibe el reconocimiento. Para estudiar el funcionamiento de las técnicas de control de flujo se hará uso de un esquema como el que se muestra en la Figura 6-1. En este tipo de esquemas se representa el tiempo con una línea vertical, la inclinación de la flecha se debe al retardo de tiempo que pasa hasta que el otro equipo recibe la trama correspondiente. Con las flechas se indica el tránsito de tramas por el medio de comunicación, de forma que una flecha representará que una única trama circula por el medio físico que une las dos estaciones. El sentido de la flecha indica si la trama es enviada por el emisor o recibida por éste. Además, a menudo se identificarán las tramas de datos como Datos 0, Datos 1, etc. y las tramas de confirmación como ACK (del inglés Acknowledgment).

La principal ventaja del protocolo de parada y espera es su facilidad de implementación. Sin embargo, no aprovecha el canal lo suficiente (en el Apartado 6.2.4 se realiza un análisis de las prestaciones de este protocolo). En la Figura 6-1 se muestra el funcionamiento del protocolo de parada y espera. En este caso se supone que todas las tramas que se transmiten se reciben con éxito y además llegan en el mismo orden en el que se enviaron. Se observa como, una vez que la estación destino recibe una trama de datos, indica al emisor su deseo de aceptar otra trama mediante el envío de una confirmación de la trama que acaba de recibir. Además, la estación emisora queda a la espera de recibir la confirmación antes de proceder a realizar el envío de una nueva.

Asimismo, cabe indicar que este protocolo puede funcionar sobre un nivel físico que sólo soporta semiduplex.

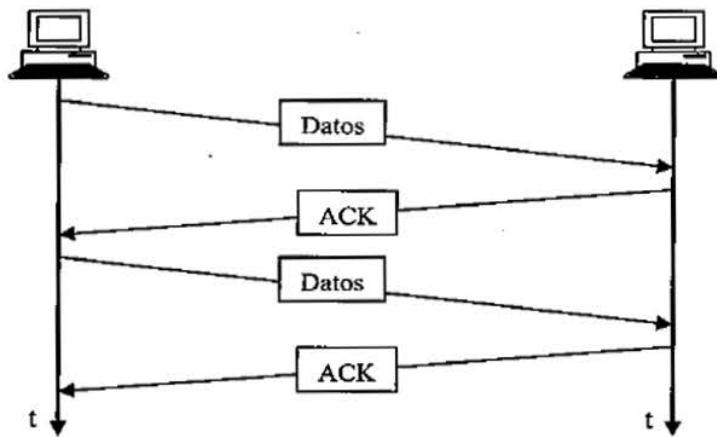


Figura 6-1. Protocolo de parada y espera.

6.2.2 Parada y espera considerando que se producen errores

Un caso particular es la implementación del protocolo de parada y espera considerando la posibilidad que puedan producirse errores en las transmisiones y que estos errores puedan provocar la pérdida de tramas. Para ello, cada vez que el emisor envía una trama, le asigna un identificador, inicia un temporizador y queda a la espera de la llegada del asentimiento (ACK). Cuando expira el temporizador, si no ha llegado el ACK se procede al reenvío de la trama. El temporizador expirará cuando por algún motivo el ACK enviado por el receptor se pierde o bien éste no llegue a enviarlo. Esto último puede deberse a que el paquete de datos sea recibido con errores o bien no llegue al receptor. Si la trama ACK llega al emisor antes de que expire el temporizador, este se para.

En la Figura 6-2 se muestra un ejemplo de aplicación del protocolo de parada y espera con numeración de tramas en el que la trama de datos "Datos 0" se pierde antes de llegar al receptor. En este caso, cuando expira el temporizador la trama de datos vuelve a ser enviada.

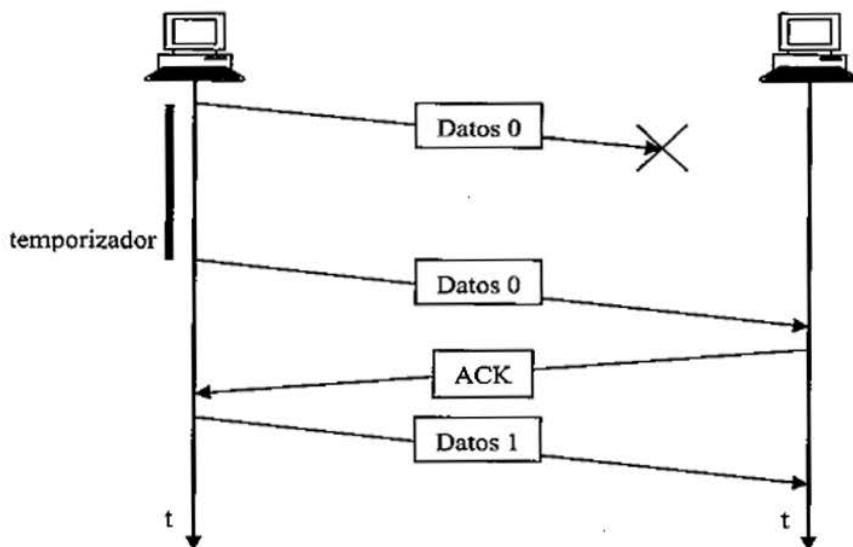


Figura 6-2. Protocolo de parada y espera con errores (pérdida de trama de datos).

En la Figura 6-3 se describe el caso en el que se pierde la confirmación de la trama "Datos 0". Igual que en el caso anterior expiraría el temporizador y se volvería a enviar la trama "Datos 0". Sin embargo, como el receptor ya la había recibido previamente, esta trama le llega duplicada. Como se ha realizado una numeración previa de las tramas el receptor es capaz de identificar que la trama le ha llegado duplicada, por lo que la descarta y vuelve a enviar la confirmación.

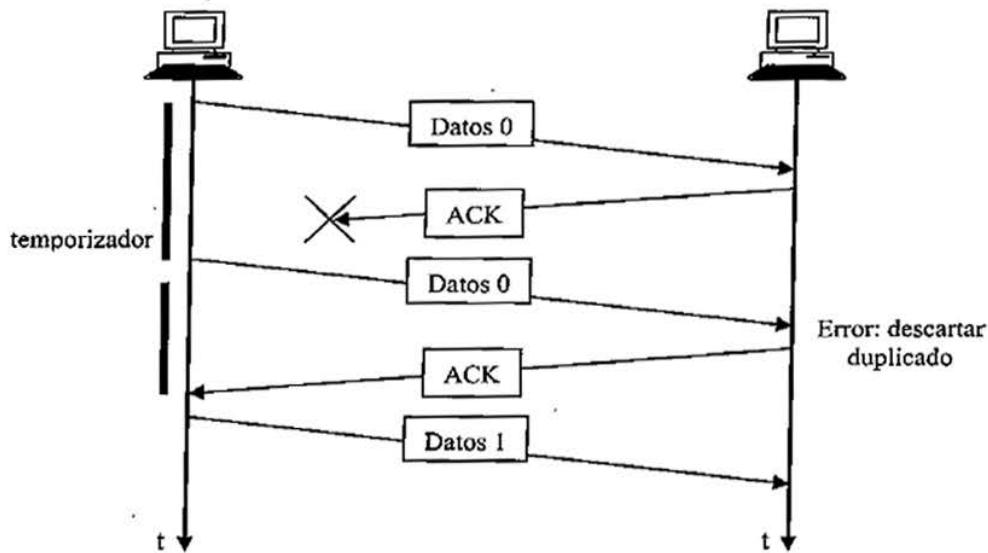


Figura 6-3. Protocolo de parada y espera con errores (tramas duplicadas por pérdida de ACK).

Un problema adicional es el que se muestra en la Figura 6-4, que considera el uso de una línea duplex. Se observa que el retraso en la llegada de la confirmación de "Datos 0" provoca, no solo la generación de tramas duplicadas, sino también la pérdida de sincronización. El retraso de la primera confirmación de Datos 0 provoca que se interpreten las confirmaciones de las tramas de datos de forma incorrecta. En este ejemplo, la segunda confirmación de Datos 0 es interpretada por el emisor como la confirmación de Datos 1. Así, la confirmación de Datos 1 será interpretada por el emisor como la de Datos 2 y así sucesivamente. La solución a este problema consiste en numerar también las confirmaciones (véase Figura 6-5).

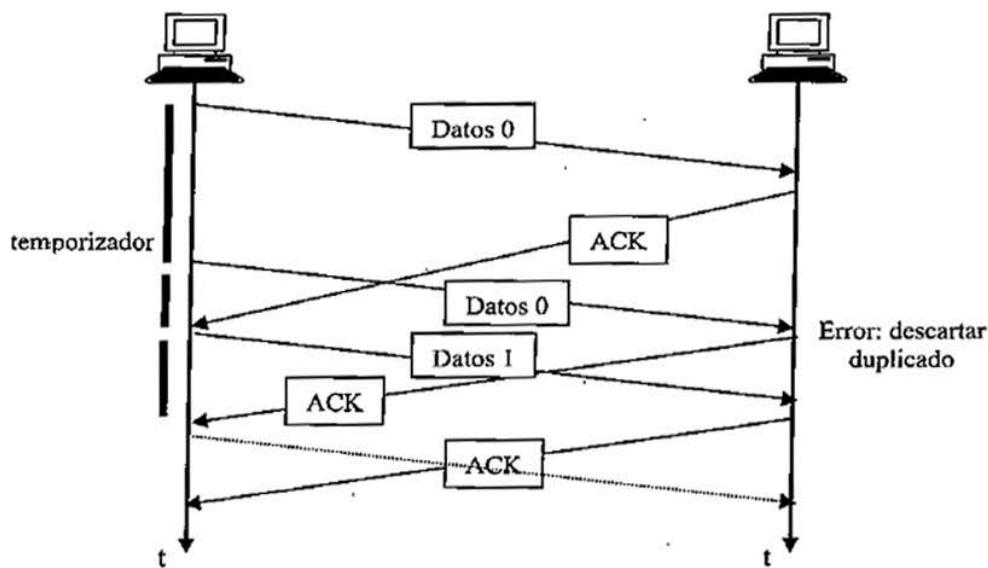


Figura 6-4. Protocolo de parada y espera con errores (tramas duplicadas y pérdida de sincronización).

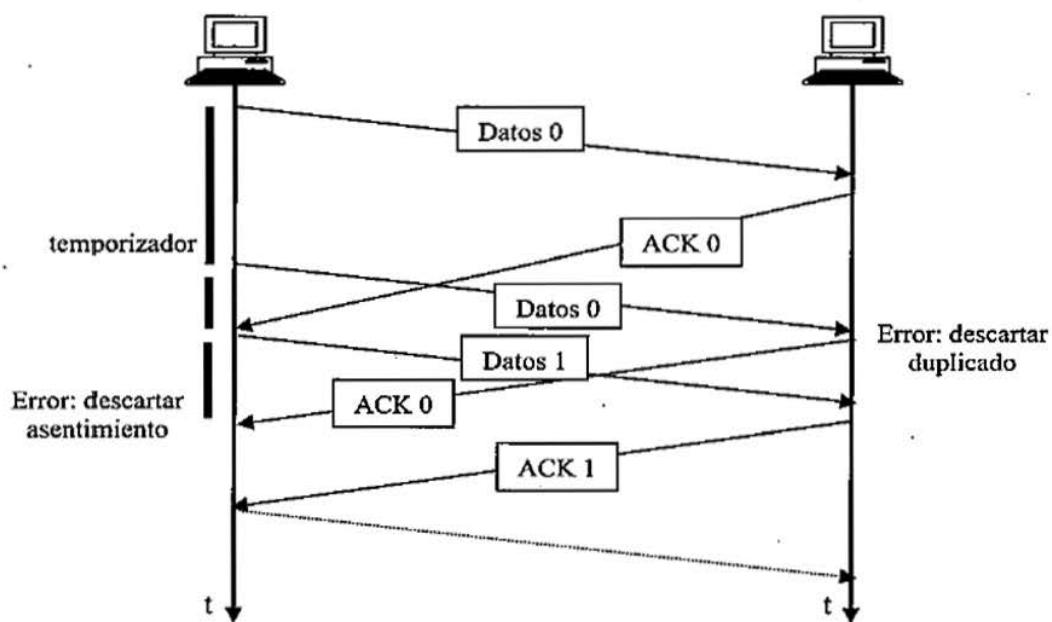


Figura 6-5. Protocolo de parada y espera con errores. Numeración de tramas y asentimientos.

6.2.3 Parada y espera con confirmaciones negativas

Como se ha visto antes, con el protocolo de parada y espera que usa temporizador y ACKs, tanto si se pierde o altera la trama de datos como su ACK, el emisor debe esperar el tiempo del temporizador antes de reenviar. También se puede dar el caso de que el receptor detecte errores en la trama aunque esta le llegue y bastará con que el receptor no confirme para que el emisor le envíe de nuevo la trama. Pero este tendrá que esperar a que venza el temporizador. Se puede agilizar el proceso haciendo que el receptor envíe una trama especial NACK (negative ACK) cuando reciba errónea, y así el emisor que recibe un NACK reenvía los datos sin tener que esperar al temporizador (Figura 6-6). La pérdida del NACK se trata como la pérdida del ACK.

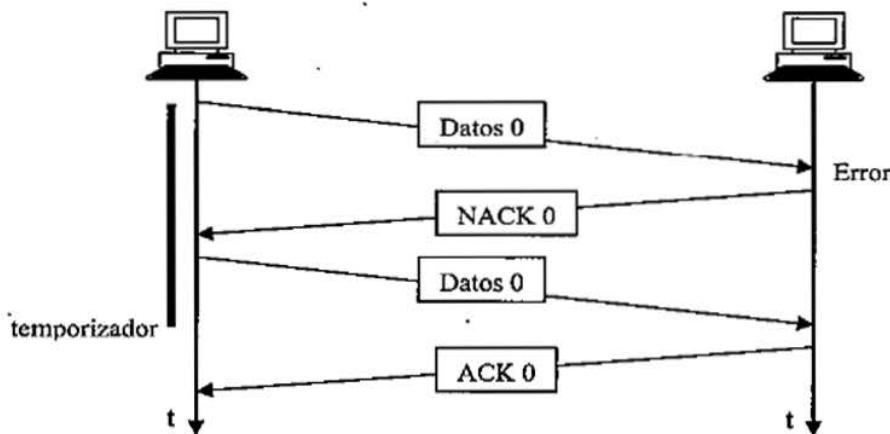


Figura 6-6. Protocolo de parada y espera con confirmaciones negativas.

6.2.4 Protocolo bilateral de parada y espera

El protocolo bilateral de parada y espera permite extender el protocolo de parada y espera a nivel de enlace duplex. Para ello se considera que las dos estaciones son emisoras y receptoras de tramas de datos. A su vez, son emisoras y receptoras de confirmaciones. En este caso se aprovecha la trama de datos a enviar para confirmar la trama de datos inmediatamente recibida (véase Figura 6-7). Este procedimiento se denomina de incorporación de confirmación (piggybacking en inglés). Por tanto, si una estación tiene para enviar una confirmación además de datos, lo hará conjuntamente utilizando una sola trama, ahorrando así capacidad del canal. Por supuesto, si una estación tiene que enviar una confirmación pero no tiene datos, se enviará una trama de confirmación convencional ACK. Si la estación tiene datos para enviar pero nada que confirmar, deberá repetir el último número de secuencia de confirmación enviado con anterioridad. Esto se debe a que en la trama de datos se prevé un campo para el número de confirmación y, por tanto, habrá que especificar algún valor en el mismo. Cuando una estación reciba una confirmación repetida, simplemente la ignorará.

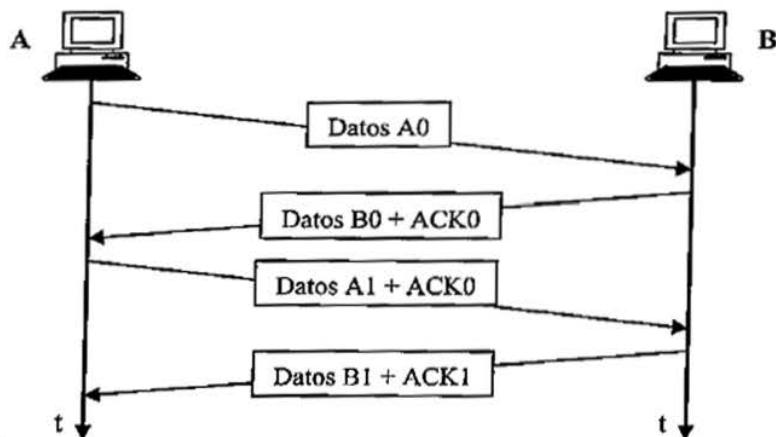


Figura 6-7. Protocolo bilateral de parada y espera.

6.2.5 Análisis de las prestaciones del protocolo de parada y espera

Independientemente del tipo de protocolo de parada y espera que se haya implementado se dice que se hace control de flujo porque el destino puede parar el flujo de los datos sin más que retener las confirmaciones. En general, el procedimiento es eficaz y funciona adecuadamente cuando los datos a transmitir se envían empleando pocas tramas de gran tamaño.

A continuación se va a determinar la eficiencia de una línea punto a punto half-duplex donde se usa el esquema de parada y espera unilateral. Para ello, se va a considerar que la estación emisora desea enviar un total de datos T que los fragmenta en un conjunto de n tramas. Además, el emisor envía las tramas de datos numeradas como Datos 0, Datos 1, Datos 2, etc. y la estación receptora confirma esas tramas como ACK0, ACK1, ACK2, etc. Así, el tiempo total que necesitaría el equipo emisor para transmitir los T datos sería nT_i , siendo T_i el tiempo necesario para transmitir la trama de

datos Datos i y recibir su confirmación. Así, el tiempo total para enviar los datos de una trama se puede definir como:

$$T_i = t_{tx} + t_{prop} + t_{procT} + t_{conf} + t_{prop} + t_{procC} \quad (6.1)$$

donde:

t_{tx} = tiempo de duración de la trama.

t_{prop} : tiempo de propagación de la trama/confirmación.

t_{procT} : tiempo de procesamiento de la trama.

t_{conf} : tiempo de duración de la confirmación.

t_{procC} : tiempo de procesamiento de la confirmación.

Si los equipos emisor y receptor son suficientemente rápidos t_{procT} , t_{conf} y t_{procC} se pueden considerar despreciables respecto al resto de los términos, por lo que la Ecuación (6.1) quedará:

$$T_i = t_{tx} + 2t_{prop} \rightarrow T = (t_{tx} + 2t_{prop})n \quad (6.2)$$

De todo el tiempo T_i definido en la Ecuación (6.2) únicamente se destina a transmitir la trama un tiempo igual a t_{tx} , por lo que la utilización de la línea o eficiencia será:

$$e = \frac{t_{tx}}{t_{tx} + 2t_{prop}} \quad (6.3)$$

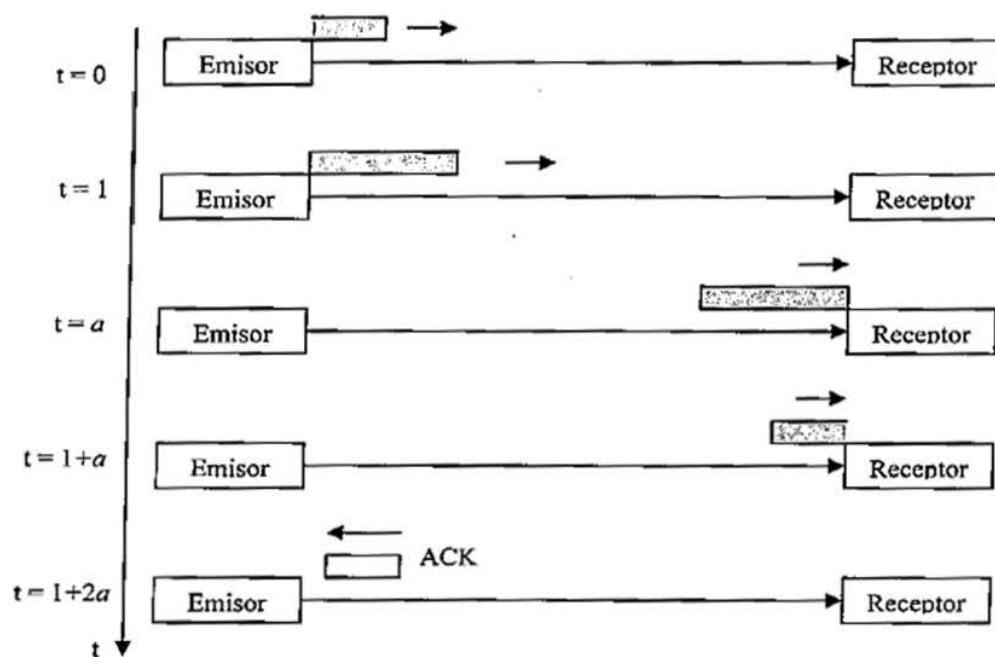
Si se define $\alpha = t_{prop}/t_{tx}$ la expresión (6.3) queda como:

$$e = \frac{1}{1 + 2\alpha} \quad (6.4)$$

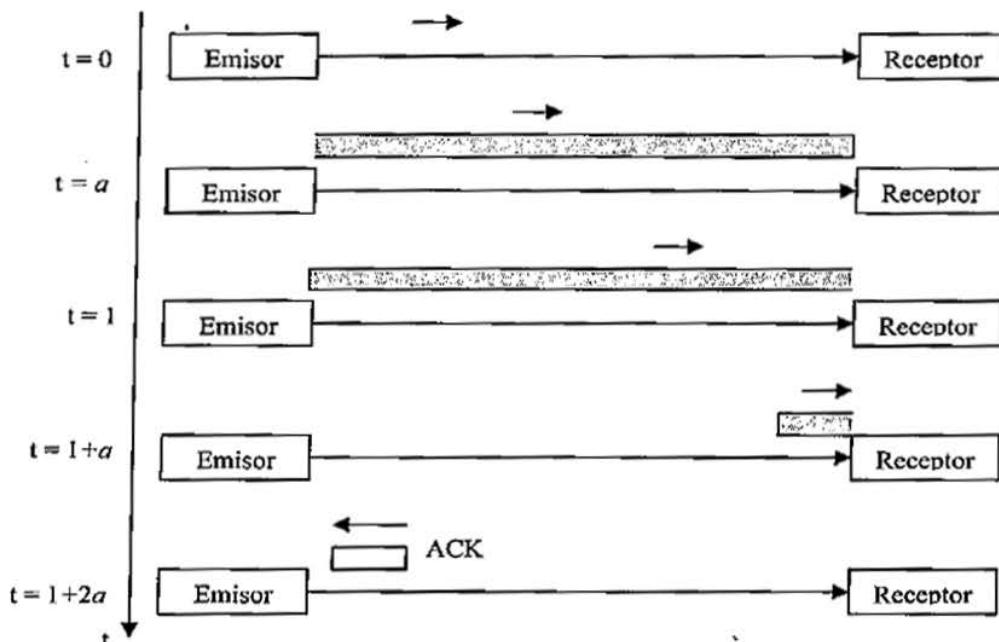
Por otro lado, el tiempo de propagación, t_{prop} , es igual a la distancia del enlace, d, dividida por la velocidad de propagación, V. Para transmisiones no guiadas, a través del aire o el espacio, V es la velocidad de la luz (aproximadamente igual a 3×10^8 m/s). Para transmisiones guiadas V es aproximadamente igual a 0,67 veces la velocidad de la luz (depende del medio de transmisión). Además, el tiempo de transmisión, t_{tx} , es igual a la longitud de la trama en bits, L, dividida por la velocidad de transmisión, R, por lo tanto:

$$\alpha = \frac{t_{prop}}{t_{tx}} = \frac{\frac{d}{V}}{\frac{L}{R}} = \frac{Rd}{VL} \quad (6.5)$$

Cuando α es menor que 1, el tiempo de propagación es menor que el de transmisión. En este caso, la trama es lo suficientemente grande para que los primeros bits de la misma lleguen al destino antes de que el origen haya concluido la transmisión de dicha trama. Cuando α es mayor que 1, el tiempo de propagación es mayor que el de transmisión. En este caso, el emisor completa la transmisión de toda la trama antes que el primer bit de la misma llegue al receptor. Es decir, para velocidades de transmisión (R) y/o distancias grandes (d) se tiene valores grandes de α y la utilización e es baja. En la Figura 6-8 se muestra el comportamiento de parada y espera con $\alpha > 1$. En esta figura se representa el proceso de transmisión de la trama de datos (rectángulo sombreado) y al final se muestra la transmisión de la trama de confirmación correspondiente. Se considera que el tiempo de transmisión es 1, por lo que el tiempo de propagación será igual a α . En esta figura puede observarse que la línea está infrautilizada.

Figura 6-8. Protocolo de parada y espera. $\alpha > 1$.

Por último, en la Figura 6-9 se describe el proceso de transmisión de una trama y su confirmación considerando $\alpha < 1$. En este caso se mejora la utilización del canal, sin embargo, también se utiliza de manera ineficiente. En general, el control de flujo mediante parada y espera da lugar a una utilización ineficiente del canal cuando se emplean velocidades de transmisión muy altas entre estaciones muy distantes. Con el objetivo de mejorar el aprovechamiento del canal en el Apartado 6.3 se describirá el protocolo de ventana deslizante.

Figura 6-9. Protocolo de parada y espera. $\alpha < 1$.

6.3 Protocolo de ventana deslizante

6.3.1 Características del protocolo de ventana deslizante

El poco aprovechamiento del canal del protocolo de parada y espera se debe fundamentalmente a que sólo puede haber en circulación una trama. En el protocolo de ventana deslizante se mejora este aspecto permitiendo que se transmitan varias tramas al mismo tiempo sobre el enlace. Haciendo uso del protocolo de ventana deslizante el emisor puede enviar varias tramas antes de necesitar un reconocimiento. Se emplea numeración para la identificación tanto de las tramas de datos como de las tramas de confirmación. Para realizar la numeración se emplea un total de n bits por lo que el rango de numeración será de 0 a 2^{n-1} .

Para el funcionamiento del protocolo de ventana deslizante, el receptor hace uso de una memoria temporal de suficiente tamaño para almacenar W tramas. Considerando este tamaño en la memoria del receptor, el emisor podrá enviar hasta un total de W tramas sin necesidad de recibir confirmación. Como se ha indicado anteriormente, cada trama de datos va numerada con un identificador o número de secuencia. El receptor envía en las confirmaciones el número de secuencia de la siguiente trama que se espera recibir. Con esto quedarán confirmadas todas las tramas anteriores.

Una vez descrito el funcionamiento básico, a continuación se va a mostrar el comportamiento del algoritmo desde el punto de vista del emisor y del receptor.

Desde el punto de vista del emisor, se emplea una ventana que almacena las tramas que pueden ser enviadas y aquellas de las que todavía que no se ha recibido confirmación. En la Figura 6-10 se muestra una forma de representar la ventana del emisor. El rectángulo sombreado representa las tramas que se pueden transmitir. Cada vez que se envíe una trama, la ventana sombreada reducirá su tamaño en una unidad por el límite inferior (la ventana se cierra). Además, cada vez que se reciba una trama de confirmación la ventana se abrirá y aumentará su tamaño por el límite superior. Las tramas comprendidas entre la barra vertical y la ventana son aquellas que ya han sido enviadas pero que están pendientes de confirmación. El emisor debe almacenar estas tramas en una memoria temporal por si hubiera que retransmitirlas.

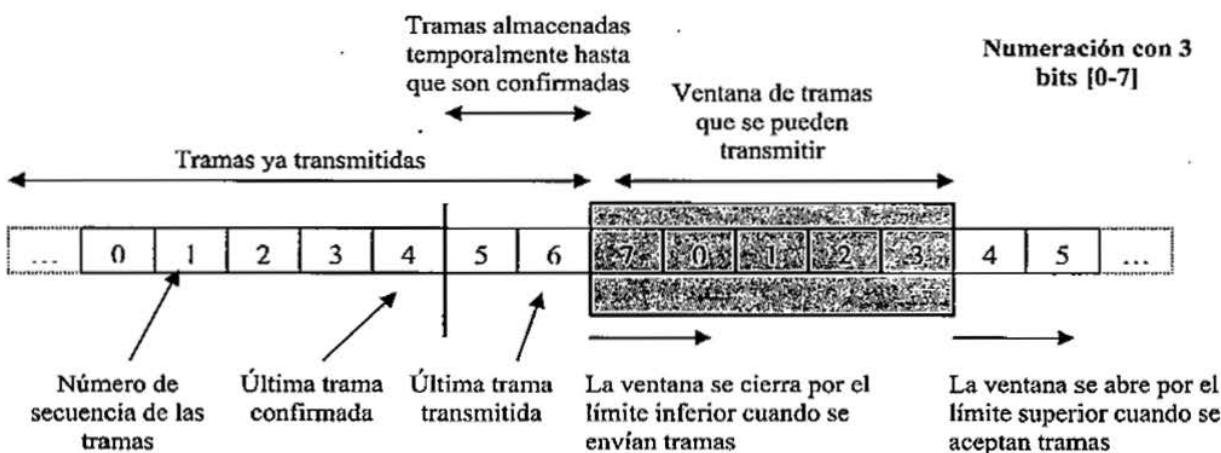


Figura 6-10. Ventana del emisor.

Desde el punto de vista del receptor, se emplea una ventana que indica las tramas que se espera recibir y para las que se enviará confirmación (véase la Figura 6-11).

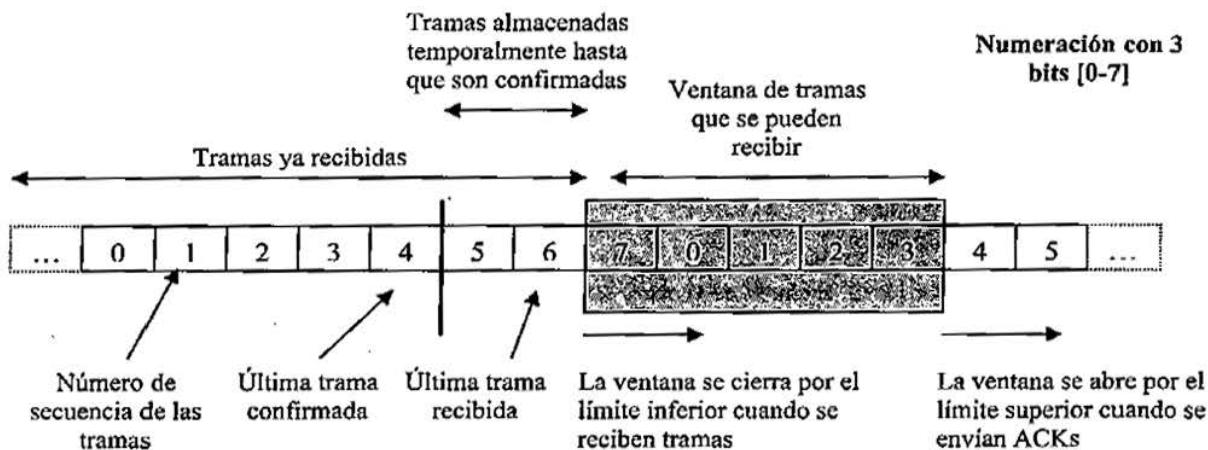
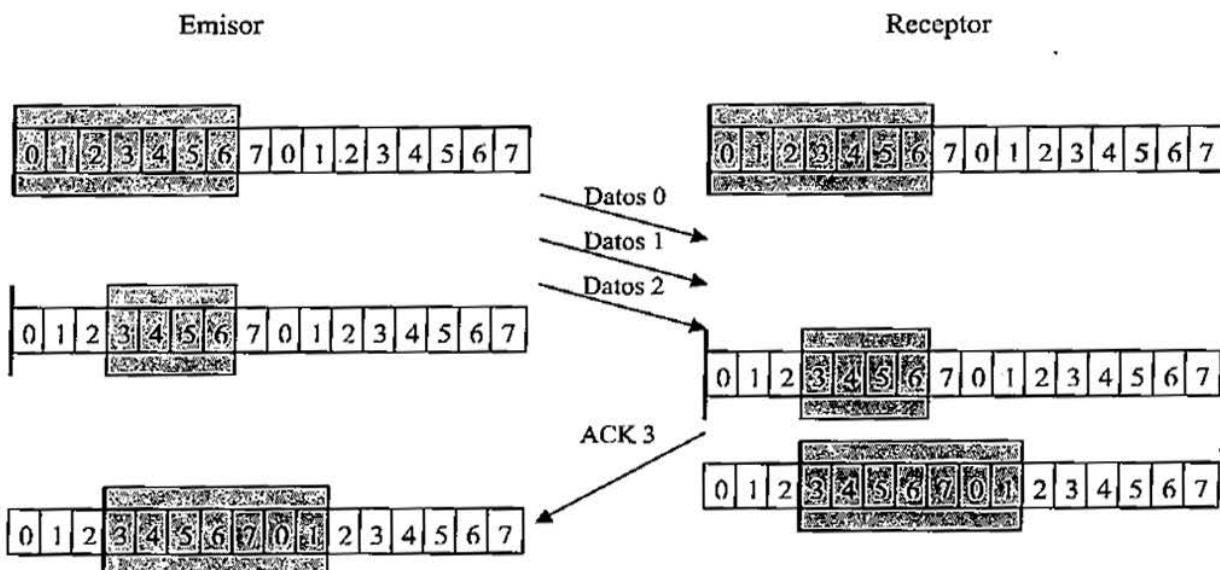


Figura 6-11. Ventana de receptor.

Ejercicio 5.1

Control de flujo con ventana deslizante considerando que no hay errores.

A continuación se va a mostrar una traza del protocolo de ventana deslizante sin errores considerando numeración de 3 bits (las ventanas permiten almacenar un máximo de 7 tramas). Inicialmente el emisor puede transmitir las tramas de 0-6 y el receptor puede recibir las tramas de 0-6. Cuando el emisor transmite las tramas de datos 0-2 la ventana del emisor se cierra tres unidades presentando únicamente un tamaño de 4 tramas. Se mantiene una copia de las tres primeras tramas a la espera de su confirmación. Cuando el receptor confirma con ACK3, está indicando que le ha llegado correctamente hasta la trama 2 por lo que a continuación la ventana de emisor vuelve a presentar su tamaño original. En la traza también puede observarse que el receptor recibe las tramas de datos 4-6 pero no las confirma. Esto provoca que la ventana del receptor se reduzca de tamaño y que estas tramas queden almacenadas en una memoria temporal del receptor hasta que finalmente el nivel superior en el receptor lea esas tramas y se confirmen.



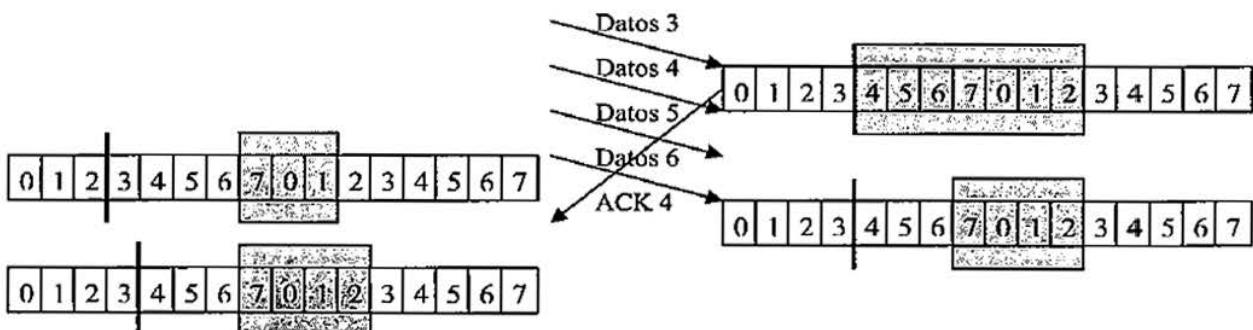


Figura 6-12. Ejemplo de ventana deslizante con numeración de 3 bits.

Empleando un protocolo de ventana deslizante se puede controlar el flujo de datos ya que el receptor sólo acepta las W tramas siguientes a la última que ha confirmado (siendo W el tamaño de la ventana del receptor). La mayoría de los protocolos permiten que una estación pueda interrumpir totalmente la transmisión de tramas desde el otro extremo mediante el envío de un mensaje de tipo RNR (Receiver Not Ready: receptor no preparado), con el que se confirman las tramas anteriores pero se prohíbe la transmisión de tramas adicionales. Por lo tanto, si el receptor envía la trama RNR 6 querrá informar que se han recibido correctamente todas las tramas de datos hasta la 5 pero no se aceptarán más. En cuanto el receptor vuelve a transmitir una trama ACK querrá decir que vuelve a aceptar más tramas de datos. A menudo, en los protocolos de ventana deslizante las tramas de confirmación también son denominadas como mensajes de tipo RR (Receiver Ready: receptor preparado).

Al igual que se describió en el apartado 6.2.3 para el protocolo de parada y espera, también se puede realizar la comunicación bidireccional a nivel de enlace empleando ventana deslizante. En este caso, cada estación deberá mantener dos ventanas (una para transmitir y otra para recibir). Para ello, se emplea la misma técnica que la descrita en el apartado 6.2.3, es decir, cada trama de datos incluye un campo en el que se indica el número de secuencia de la trama más un campo adicional con el número de secuencia que se confirma.

Según los tamaños de la ventana del emisor, We , y la ventana del receptor, Wr , se definen las siguientes variantes del protocolo de ventana deslizante:

- $We = 1$, $Wr = \text{cualquiera}$. El comportamiento del protocolo sería el mismo que el protocolo de parada y espera bidireccional. El emisor envía las tramas de datos de una en una.
- $We > 1$, $Wr = 1$. Se trata del protocolo de ventana deslizante de envío continuo con repetición no selectiva. El emisor puede enviar un cierto número de tramas antes de recibir sus confirmaciones. Así, se aprovecha mucho mejor el medio que en parada y espera.
- $We > 1$, $Wr > 1$. Mejora el caso anterior para aumentar la eficiencia cuando se produzcan errores. Si no se produjeron errores el aprovechamiento del medio sería similar al caso anterior. Si el receptor recibe una trama errónea la descarta y no envía su ACK, pero acepta otras posteriores que estén en la ventana del receptor. El receptor confirma las tramas tras un error con el ACK de la recibida correctamente antes del error. A continuación las tramas que va recibiendo el receptor las guarda en su lista hasta que tiene la primera y las pasa al nivel de red ordenadas. Seguidamente ya puede confirmar todas las tramas con un ACK de la última recibida correctamente. Con esto quedarán confirmadas todas las anteriores.

Además, para que un protocolo de ventana deslizante con envío continuo ($We > 1$) proporcione una buena eficiencia se debe disponer de un medio físico duplex, aunque la transmisión de enlace sea semi-duplex, ya que sólo así se pueden enviar simultáneamente los datos y sus ACKs.

En el capítulo 8 se describirán algunos ejemplos de protocolos estandarizados que emplean ventana deslizante para realizar el control de flujo de la transmisión.

6.3.2 Análisis de las prestaciones del protocolo de ventana deslizante

El control de flujo mediante ventana deslizante es mucho más eficiente que el de parada y espera debido a su mejor aprovechamiento del canal. El grado de utilización del canal depende directamente del tamaño de la ventana W , así como del valor de α definido en el Apartado 6.2.4. Al igual que se hizo en la Figura 6-8 se va a considerar el tiempo de transmisión de la trama igual a la unidad, por lo que el tiempo de propagación será igual a α .

Supóngase en primer lugar que $W \geq 2\alpha+1$. La confirmación de la primera trama enviada llega antes de que ésta agote su ventana. Por tanto, la estación emisora puede transmitir continuamente sin necesidad de detener la transmisión para quedarse a la espera de confirmaciones. Por lo tanto, la eficiencia será de 1.

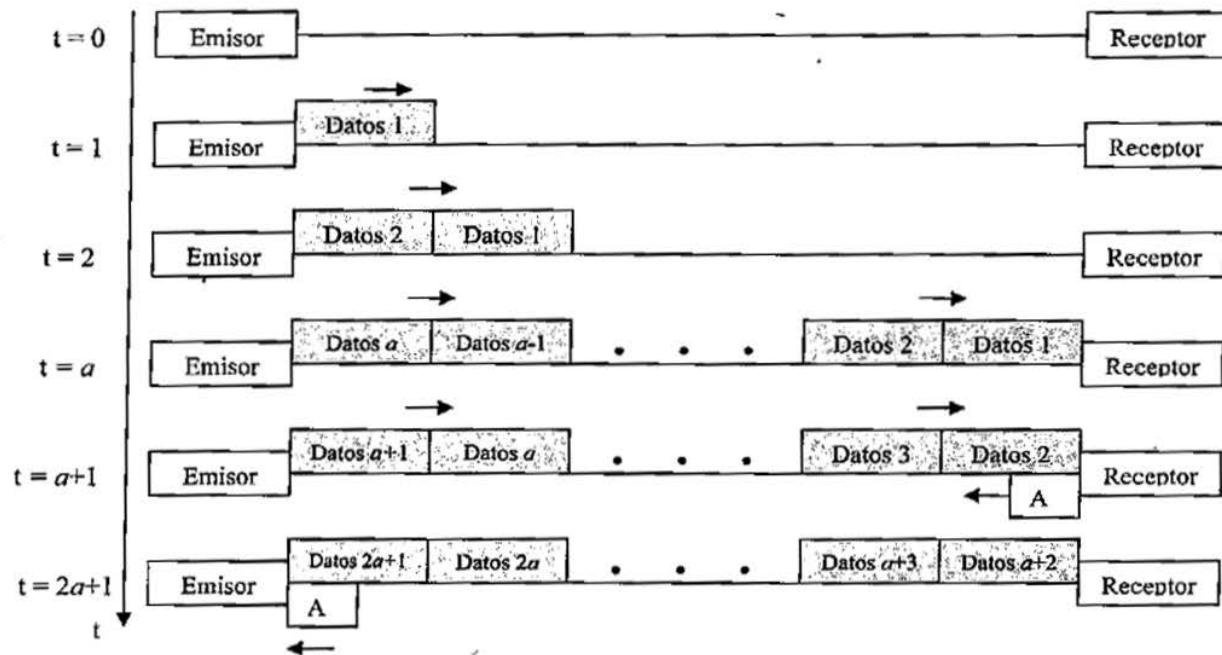
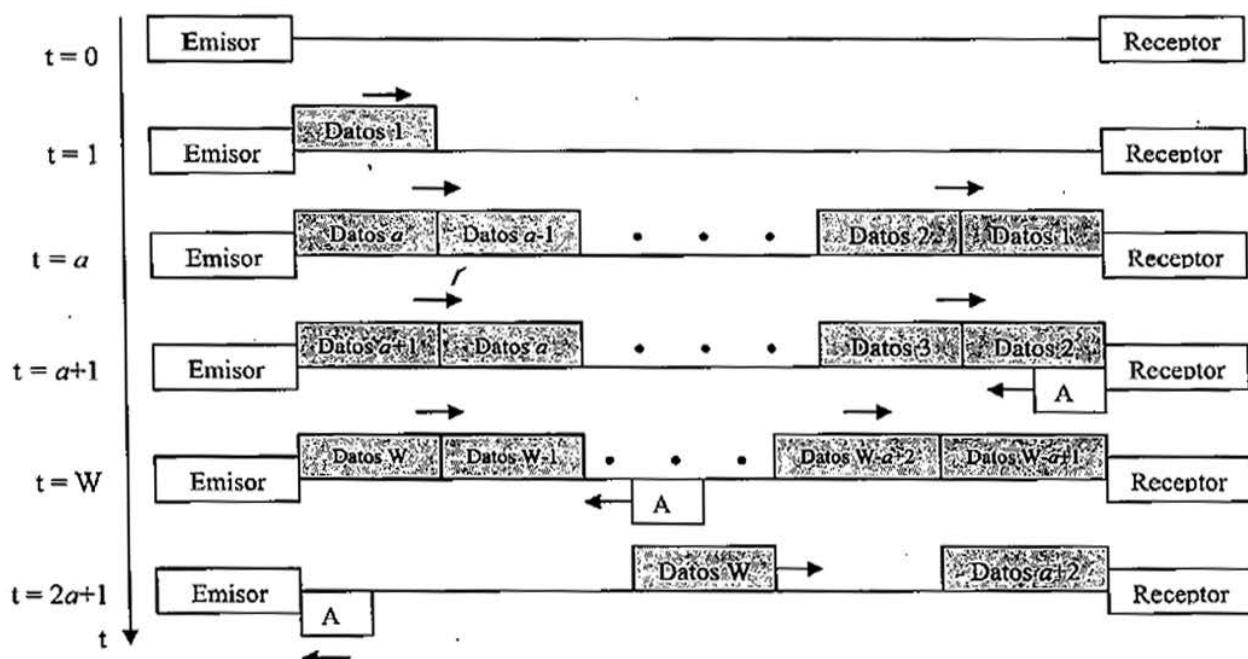


Figura 6-13. Ventana deslizante. $W \geq 2\alpha+1$.

En el caso de que $W < 2\alpha+1$ la estación transmisora no podrá enviar más tramas de datos en el momento $t=2\alpha+1$ ya que se llena la ventana del emisor (véase Figura 6-14). Por lo tanto, la utilización de la línea es W unidades de tiempo por cada periodo $2\alpha+1$, o lo que es lo mismo la eficiencia será:

$$e = \frac{W}{2\alpha+1} \quad (6.6)$$

Figura 6-14. Ventana deslizante. $W < 2a+1$.Ejercicio 5.2

Dos equipos de datos (DTEs) implementan un protocolo de parada y espera para la transferencia de ficheros. Las tramas que se emplean para enviar bloques de datos son de 200 bits, incluyendo datos, cabeceras, colas y demás bits de control. Si la velocidad de transmisión del enlace de datos en ambos sentidos es de 64Kbits/s, y la velocidad de propagación de las señales por el medio físico es de $2 \cdot 10^8$ m/s, se pide:

- Calcular la máxima distancia de separación entre los ETD para conseguir una eficiencia en el protocolo, sin errores, de al menos el 95%.
- ¿Cuál sería el valor del temporizador de retransmisión de tramas más adecuado en las condiciones del apartado anterior, para el caso en que se produzcan errores de perdida de trama?
- Proponer un nuevo mecanismo de control de flujo distinto al parada y espera (de entre los descritos en este capítulo) que permita una eficiencia del 100% cuando la distancia entre DTEs sea de 1000Km. Calcular el parámetro más significativo que permite determinar el intercambio de tramas para este mecanismo de control de flujo.

Apartado a:

Si el tiempo de confirmación ACK y el tiempo de proceso es despreciable, entonces el tiempo del envío de una trama y su confirmación, viene dado por:

$$T_i = t_{tx} + t_{prop} + t_{procT} + t_{conf} + t_{prop} + t_{procC} = t_{tx} + 2t_{prop}$$

$$t_{procT} = t_{conf} = t_{procC} = 0$$

Por otro lado, el tiempo de propagación y el tiempo de transmisión de la trama puede calcularse según los datos aportados por el problema de la siguiente manera (nótese que la distancia a calcular se ha representado como d):

$$t_{prop} = \frac{d}{2 \cdot 10^8}$$

$$t_{tx} = \frac{200\text{ bits}}{64 \cdot 10^3 \text{ bits/s}} = 3.125\text{ ms}$$

Según se ha determinado en la Ecuación (6.3), la eficiencia puede calcularse de la siguiente manera:

$$e = \frac{t_{tx}}{t_{tx} + 2t_{prop}} = \frac{3.125}{3.125 + \frac{2d}{2 \cdot 10^8}} \geq 0.95$$

Despejando de la expresión anterior la distancia, d, puede determinarse que la distancia máxima de separación entre los DTEs es de 16.447,3 m.

Apartado b:

A partir de la distancia máxima obtenida en el apartado anterior se puede calcular el tiempo total necesario para el envío de una trama de la siguiente manera:

$$T_i = t_{tx} + 2t_{prop} = 3.125 + \frac{2 \cdot 16447,3}{2 \cdot 10^8} = 3,2895\text{ msec}$$

Por lo tanto, transcurrido este tiempo debería haberse recibido la confirmación de la trama de datos. De forma que esta debería ser la duración del temporizador ya que si no se ha recibido la confirmación será porque ha ocurrido un error en la transmisión.

Apartado c:

El protocolo que permite una eficiencia es el de ventana deslizante considerando que el tamaño de ventana, W, es mayor o igual que $2\alpha+1$ en la transmisión, siendo α la relación entre el tiempo de propagación y el tiempo de transmisión de la trama.

El tiempo de propagación para este caso será:

$$t_{prop} = \frac{1000 \cdot 10^3}{2 \cdot 10^8} = 5\text{ msec}$$

Además, para conseguir una eficiencia del 100% es necesario que:

$$W \geq 2\alpha+1 = 2 \frac{t_{prop}}{t_{tx}} + 1 = \frac{2 \cdot 5}{3.125} + 1 = 4,2$$

Por lo tanto, el tamaño de la ventana empleada deberá ser mayor o igual que 5 para conseguir una eficiencia del 100%.

Capítulo 7. PROTOCOLOS ESTANDARIZADOS DEL NIVEL DE ENLACE

En muchas ocasiones se dispone de un único medio de comunicación que es compartido por distintos equipos y que es empleado para la transferencia de información entre todos ellos. Por ello, en este capítulo se mostrarán las técnicas que permiten compartir un medio físico y gestionar el acceso a éste para intercambio de información. Además, se describirá el funcionamiento del protocolo HDLC para regular el intercambio de información y controlar errores y el protocolo PPP como ejemplo de un protocolo punto a punto que se emplea habitualmente tanto para conexiones por línea alquilada entre nodos encaminadores como conexiones commutadas entre equipos terminales de datos.

CAPÍTULO 7

7-1 Técnicas de compartición del medio

- Técnicas de contienda
- Selección distribuida

7-2 Técnicas de control del enlace: Protocolo HDLC

- Tipos de estaciones y enlaces
- Modos de funcionamiento
- Estructura de la trama
- Tipos de tramas
- Repertorio básico de instrucciones
- Control del enlace
- Configuración del medio de transmisión y transferencia de datos
- Transferencia de datos con rechazo

7-3 Protocolo PPP

- Estructura de la trama
- Tipos de tramas
- Funcionamiento de PPP

OBJETIVOS

- Clasificar y describir las técnicas que permiten a varios equipos emplear un mismo medio para transferir información: contienda, reserva y selección
- Exponer y comentar las técnicas de compartición más empleadas con algunas de las tecnologías de uso más habitual en redes LAN
- Presentar y describir el funcionamiento de un protocolo de nivel de enlace que permite controlar el enlace y que además hace uso de técnicas de ventana deslizante para regular el flujo de información: HDLC
- Presentar y describir el funcionamiento del protocolo de nivel de enlace más empleado en conexiones punto a punto: PPP

TERMINOLOGÍA – PALABRAS CLAVE

- | | | |
|------------------------|-----------------------|------------------------|
| • Técnica de Contienda | • Estación primaria | • Trama de información |
| • Técnica de Reserva | • Estación secundaria | • Trama de supervisión |
| • Técnica de Selección | • Estación combinada | • Trama no numerada |
| • ALOHA | • Modo balanceado | • LCP |
| • CSMA | • Mono asíncrono | • CHAP/PAP |
| • CSMA/CD | • Modo normal | • IPCP/NCP |
| • HDLC | • PPP/PPPoE/PPPoA | • CCP/MPPE |

7.1. Técnicas de compartición del medio

En muchas ocasiones se dispone de un único medio de comunicación que es compartido por distintos equipos y que es empleado para la transferencia de información entre todos ellos. Fundamentalmente se adopta un medio compartido con el objetivo de ahorrar costes de material y hacer más sencillo el mantenimiento e instalación. Véase, por ejemplo, los casos mostrados en la Figura 7-1 donde se ha representado una red basada en bus y otra en anillo. En ambos casos existe un único medio de transmisión que es compartido por cuatro equipos y al cual pueden acceder indistintamente para comunicar información al resto de equipos. Por lo tanto, es necesario que se implemente algún tipo de protocolo de control de acceso para coordinar los equipos con el objetivo de decidir cual de ellos envía información en un instante de tiempo determinado sin que las tramas enviadas "colisionen" con las enviadas por otro equipo. En caso de que pueda haber colisiones, el protocolo debe tratar de resolverla. De realizar esta coordinación para el acceso al medio físico compartido se encarga el nivel de enlace de las LAN mediante las denominadas técnicas de compartición del medio (MAC).

El término colisión se emplea para indicar la situación en la que una trama de datos interfiere en el medio de transmisión con la que simultáneamente ha enviado otro equipo. Esto se produce cuando dos equipos transmiten información casi al mismo tiempo. En este caso, las dos tramas pueden interferir entre sí. Como se verá hay técnicas de compartición que evitan las colisiones, otras que tratan de evitarlas y otras que admiten la existencia de colisiones y se preocupan de cómo resolverlas cuando se producen.

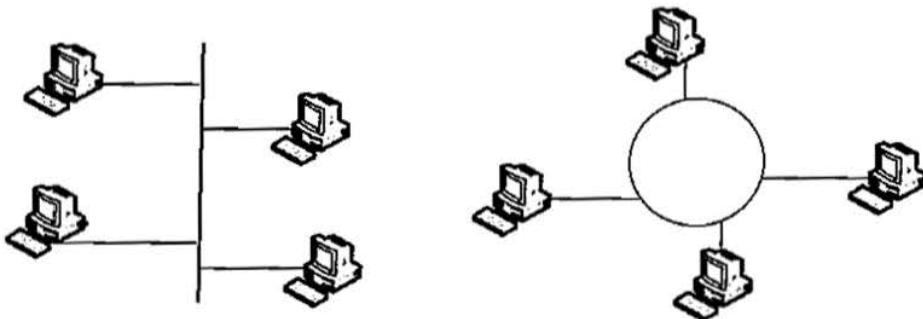


Figura 7-1. Medios compartidos.

Las técnicas de compartición del medio pueden clasificarse en:

- *Contienda o acceso múltiple.* Son técnicas sencillas en general aunque también las menos eficientes. Se denominan técnicas de contienda debido a que los equipos que pretenden acceder al medio para realizar la transmisión de los datos "compiten" por el medio antes de realizar el envío de los datos. Una de las desventajas de este método es que pueden producirse colisiones en los datos transmitidos por los distintos equipos.
- *Reserva.* El acceso al medio se realiza mediante dos pasos: reserva y envío de datos. Durante la fase de reserva el equipo que quiere acceder al medio solicita la reserva del medio de transmisión. Una vez realizada la reserva y concedida ésta, podrá realizarse la transmisión de datos sin correr el riesgo de que los datos enviados puedan colisionar con los transmitidos por otro equipo. Aunque se evitan las colisiones al transmitir los datos,

aún pueden producirse colisiones en la petición de reserva. Sin embargo, las peticiones de reserva generalmente son más cortas y, por lo tanto, se reduce la probabilidad de que llegue a producirse una colisión. Existen dos tipos de acceso al medio por reserva:

- *Centralizada.* En la que existe un equipo especial en la red que gestiona y atiende las reservas.
- *Distribuida.* Si las reservas se gestionan entre todos los equipos que conforman la red.
- *Selección.* En esta técnica se aplica un determinado algoritmo que va seleccionando los equipos que pueden acceder al medio. Cuando el equipo es seleccionado podrá realizar el envío de los datos. La principal ventaja de este método es que no se producirán colisiones ya que en un momento dado solo podrá realizar envío de datos el equipo que se encuentre seleccionado. Al igual que el acceso por reserva , el acceso al medio por selección puede ser:
 - *Centralizada.* Hay un equipo que va concediendo los turnos al resto de equipos para realizar el envío de datos.
 - *Distribuida.* El turno es gestionado entre todos los equipos.

El método de acceso por selección distribuída es empleado por redes en anillo tipo IEEE 802.5 (Token Ring), redes ópticas FDDI y diversas redes industriales. Mientras que los métodos basados en contienda son empleados por las redes tipo Ethernet IEEE 802.3. Las redes 802.11 (WiFi) usan una variante mejorada de las técnicas de contienda que busca prevenir las colisiones. En el siguiente apartado se describe con más detalle los distintos tipos y funcionamiento de los métodos de acceso al medio típicamente utilizados.

7.1.1. Técnicas de contienda

Las técnicas de contienda pueden clasificarse en:

- *Transmisión sorda (ALOHA).* Los equipos que componen la red no tienen en cuenta si hay información en el medio antes de realizar el envío de sus datos. Es sencilla, pero no se usa hoy en día por su bajo rendimiento.
- *Transmisión con escucha (CSMA, carrier sense multiple access, Técnica de acceso múltiple con detección de portadora).* Los equipos solo enviarán datos al medio si determinan que no hay información previa enviada por otro equipo. Empleando este método sólo se producirán colisiones cuando dos o más equipos envíen datos a la vez (ya que ambos determinarían simultáneamente que el medio de comunicación está libre).

A continuación se pasan a describir las distintas técnicas de contienda basadas en la clasificación anterior.

ALOHA puro y ranurado

Empleando esta técnica de acceso al medio una determinada estación puede transmitir una trama siempre que lo necesite. A continuación, la estación pasa a escuchar el medio durante un tiempo igual al máximo retardo de propagación posible de ida y vuelta a través de la red más un pequeño incremento de tiempo. Se considera que todo ha ido bien si durante este intervalo de escucha, a la estación le llega la confirmación procedente del destino. En caso contrario retransmitirá la trama. La estación desistirá si no recibe una confirmación después de varias retransmisiones.

ALOHA es una técnica muy sencilla de implementar, sin embargo, las colisiones aumentan rápidamente cuando aumenta la carga de la red. Con el objetivo de mejorar el comportamiento del

método se implementó el denominado ALOHA ranurado. En ALOHA ranurado el canal se divide en ranuras de tiempo discreto, considerando ranuras uniformes de duración igual al tiempo de transmisión de una trama. De esta manera, la transmisión solo se permite en los instantes de tiempo que coincidan con el comienzo de una ranura. Así, las tramas que se solapen lo harán completamente, lo que incrementa la utilización máxima del sistema desde el 18% en ALOHA puro al 37% en ALOHA ranurado.

Habitualmente ocurre que el tiempo de propagación de una trama es notablemente inferior al tiempo requerido para su transmisión. De esta forma, cuando una estación transmita una trama, el resto de estaciones lo sabrá casi inmediatamente. Así, si las estaciones que componen una red pueden determinar que hay otra que ya está transmitiendo se evitarán que se produzcan colisiones a no ser que dos equipos decidan transmitir datos casi simultáneamente. Esto llevó a la definición de las técnicas de control del acceso al medio denominadas CSMA.

CSMA no persistente

En CSMA si una estación determina que el medio está ocupado deberá esperar. El tiempo de espera se establece en una cantidad de tiempo razonable en el cual debería llegar la confirmación del equipo destinatario, teniendo en consideración el retardo de propagación máximo, del trayecto de ida y vuelta y el hecho de que la estación que confirma debe competir también por conseguir el medio para responder. Si no llega la confirmación, la estación supone que se ha producido una colisión. En este caso esperará un tiempo aleatorio y volverá a comprobar si el medio está libre antes de realizar el reenvío de los datos.

Por lo tanto, en CSMA un equipo antes de enviar datos escucha el medio para determinar si está o no ocupado. Sólo cuando está libre puede transmitir y si está ocupado deberá esperar un tiempo prudencial antes de realizar el reenvío de los datos.

En general, la utilización máxima que se puede conseguir en CSMA puede superar con mucho la de ALOHA ranurado. La utilización máxima depende de la longitud de la trama y del tiempo de propagación. Cuanto mayor sea la longitud de las tramas o cuanto menor sea el tiempo de propagación, mayor será la utilización.

En CSMA se precisa de un algoritmo que determine qué debe hacer una estación si encuentra el medio ocupado. Un posible algoritmo es el implementado por CSMA no persistente en el que una estación que desea transmitir escuchará el medio y procederá según las siguientes reglas (véase Figura 7-2):

1. Si el medio se encuentra libre transmite. Si está ocupado se aplica el paso 2.
2. Si el medio se encuentra ocupado espera una cierta cantidad de tiempo obtenido de una distribución de probabilidad (retardo de transmisión) y repite el paso 1.

El uso de retardos aleatorios reduce la probabilidad de las colisiones. Esto es debido a que si dos estaciones detectan simultáneamente que se ha producido una colisión y ambas esperan el mismo tiempo para realizar la retransmisión de las tramas, se volverá a producir la colisión.

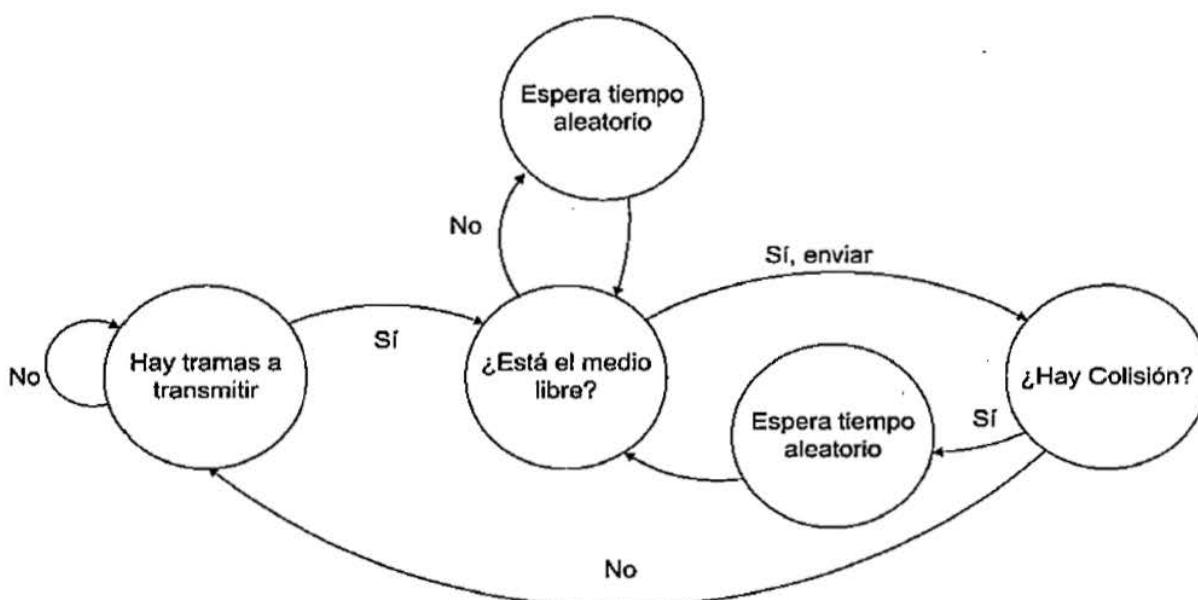


Figura 7-2. CSMA no persistente.

CSMA 1-persistente

Uno de los principales problemas de CSMA no persistente es que se desaprovecha la capacidad debido a que el medio permanecerá libre justo tras el fin de una transmisión, incluso si una o más estaciones se encuentran listas para transmitir. Para aprovechar estos intervalos de tiempo se puede utilizar el protocolo CSMA 1-persistente. En este caso un equipo que desea transmitir escuchará el medio y actuará de acuerdo siguiendo los siguientes pasos (véase Figura 7-3):

1. Si el medio se encuentra libre transmite. Si está ocupado se aplica el paso 2.
2. Si el medio está ocupado, continúa escuchando hasta que el canal se detecte libre, momento en el cual se transmite inmediatamente.

El principal problema de este algoritmo es que si dos estaciones desean transmitir se garantizará que se producirá la colisión entre sus tramas.

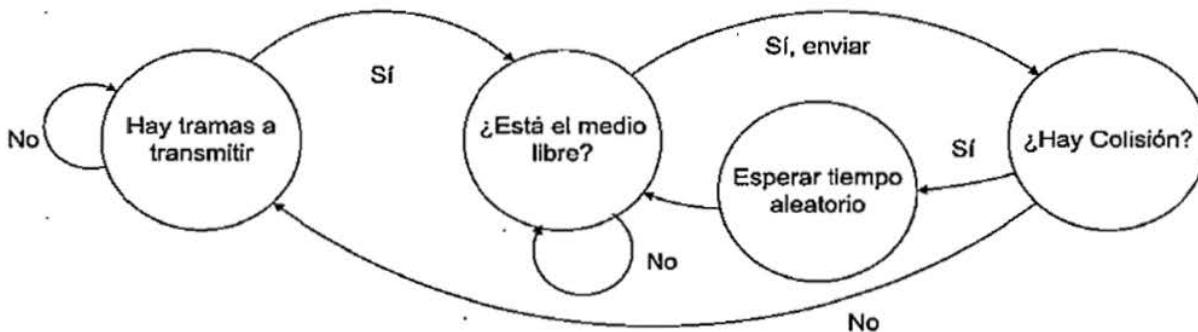


Figura 7-3. CSMA 1-persistente.

CSMA p-persistente

Esta técnica representa un compromiso entre reducir el número de colisiones (como en el caso de no persistente) y reducir el tiempo de desocupación del canal (como en el 1-persistente). Las reglas a aplicar en este caso son las siguientes (ver Figura 7-4):

1. Si el medio se encuentra libre, entonces se transmite con una probabilidad p y se espera una unidad de tiempo con una probabilidad $(1-p)$. La unidad de tiempo es generalmente igual al retardo máximo de propagación.
2. Si el medio está ocupado, se continúa escuchando hasta que se detecte libre y se repite el paso 1.
3. Si la transmisión se ha retardado una unidad de tiempo, se repite el paso 1.

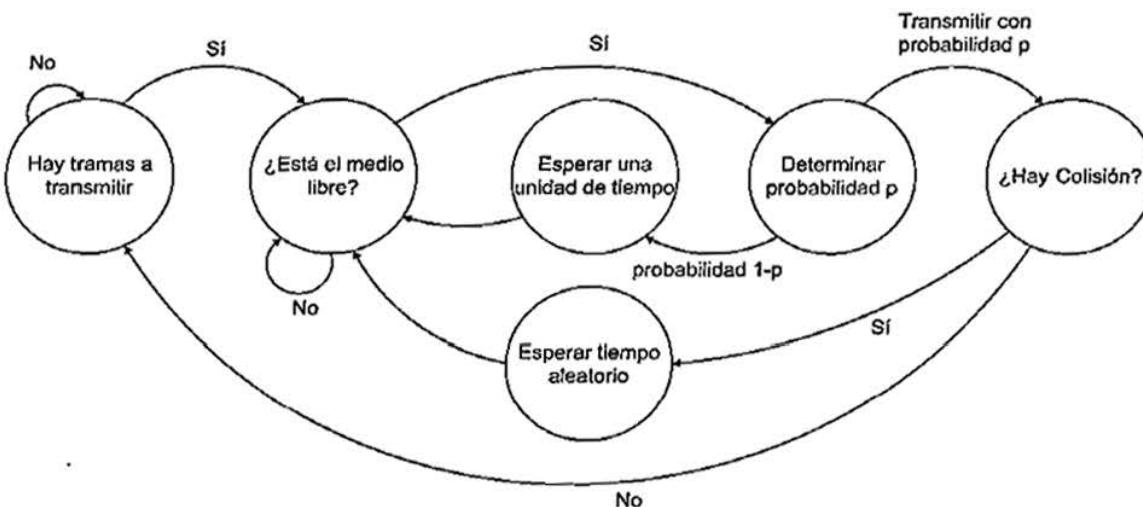


Figura 7-4. CSMA p-persistente.

CSMA/CD

En CSMA, una estación que deseé transmitir escuchará primero el medio para determinar si existe alguna otra transmisión en curso. Sin embargo, pueden producirse colisiones si dos estaciones comienzan a transmitir casi al mismo tiempo. Independientemente del método de CSMA (no persistente, 1-persistente o p-persistente) cuando se producen colisiones las estaciones esperan una determinada cantidad de tiempo antes de retransmitir. Por lo tanto, el medio permanece inutilizable mientras dure la transmisión de ambas tramas dañadas. En el caso de que la longitud de las tramas sea elevada comparada con el tiempo de propagación, la cantidad de tiempo desaprovechado puede ser importante. Este desaprovechamiento puede reducirse si una estación continúa escuchando el medio mientras dura la transmisión. Esta característica es empleada por CSMA/CD (acceso múltiple con detección de portadora y detección de colisiones (carrier sense multiple access with collision detection)). En este caso para el acceso al medio se siguen las siguientes reglas (ver Figura 7-5):

1. Si el medio se encuentra libre transmite. Si está ocupado se aplica el paso 2.
2. Si el medio se encuentra ocupado, continúa escuchando hasta que el canal se libere, en cuyo caso transmite inmediatamente.



3. Si se detecta una colisión durante la transmisión, se transmite una pequeña señal de interferencia para asegurarse de que todas las estaciones constaten la colisión. A continuación, se deja de transmitir.
 4. Tras la emisión de la señal de interferencia, la estación espera una cantidad aleatoria de tiempo conocida como espera, intentando transmitir de nuevo a continuación (volviendo al paso 1).

Este método de acceso es el empleado por las redes Ethernet IEEE 802.3. Cuando se implementa este método de acceso al medio las tramas a nivel de enlace deben presentar un tamaño mínimo como para permitir la detección de la colisión antes de que finalice la transmisión.

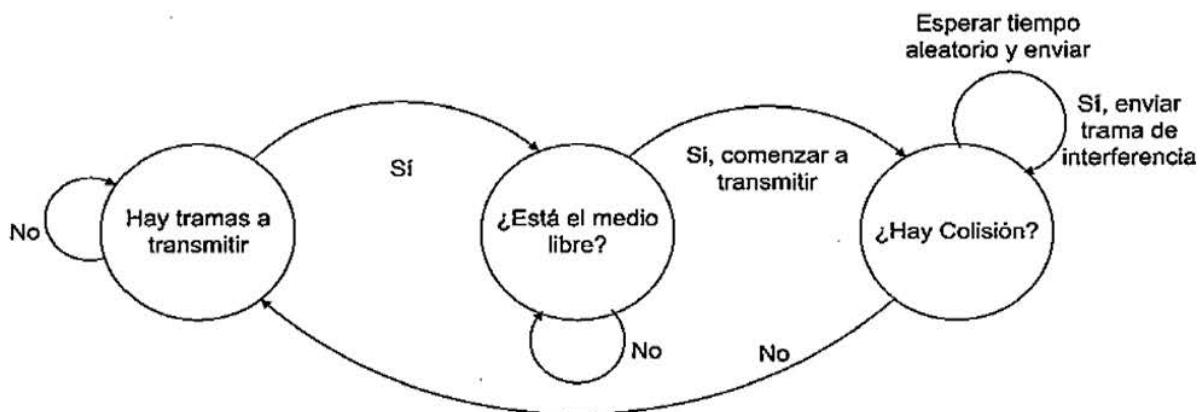


Figura 7-5. CSMA/CD.

CSMA/CA

Las redes inalámbricas IEEE 802.11 emplean para el acceso al medio una modificación de CSMA/CD denominada acceso múltiple con detección de portadora y prevención de colisiones CSMA/CA. Con este método se evita la aparición de colisiones debido a la dificultad de su detección en el medio inalámbrico. Para ello se define un algoritmo con varios pasos previos a la transmisión. Empleando este mecanismo cada equipo anuncia opcionalmente su intención de transmitir antes de hacerlo para evitar colisiones entre los paquetes de datos. De esta forma, el resto de equipos de la red sabrán cuando hay colisiones y en lugar de transmitir la trama en cuanto el medio está libre, se espera un tiempo aleatorio adicional corto y solamente si, tras ese corto intervalo el medio sigue libre, se procede a la transmisión reduciendo la probabilidad de colisiones en el canal.

7.1.2. Selección distribuída

El método de control de acceso al medio por selección distribuida es el empleado típicamente por redes en anillo tipo IEEE 802.5. En estas redes el control de acceso al medio se realiza por paso de testigo. El testigo es una trama especial que circula por el anillo y que indica el turno para transmitir datos. El testigo se pasa de un DTE a otro dentro de una topología en anillo y hasta que un DTE no recibe el testigo no puede enviar datos. De esta forma, cuando un DTE recibe el testigo transmite datos a otro DTE destino y “retiene” el testigo durante el tiempo que dura la transmisión de los datos. Cuando ha finalizado la transmisión de los datos reenvía el testigo al siguiente equipo del anillo habilitándole de esta manera para poder realizar el envío de datos.

7.2. Técnicas de control del enlace: Protocolo HDLC

En el capítulo anterior se presentó el protocolo de ventana deslizante como uno de los protocolos más usados para regular el flujo de información entre dos dispositivos. En este capítulo se proporcionan los conocimientos básicos, relacionados con algunos de los protocolos estándar del nivel de enlace, que hacen uso del protocolo de ventana deslizante para regular el intercambio de información y controlar errores. En concreto, se presenta el estándar HDLC (High Level Data Link Control), su estructura de trama, sus modos de funcionamiento y su repertorio básico de instrucciones. El protocolo HDLC está regulado por el estándar ISO 3309 e ISO 4335 y se caracteriza por ser un protocolo de nivel de enlace orientado a bit. Este protocolo es de vital importancia porque ha servido como base para la implementación de otros muchos protocolos. Entre los que cabe destacar LAPB: Link Access Procedure Balanced para redes X.25, LAPD: Link Access Protocol D-Channel en RDSI, LLC: Logical Link Control (IEEE 802.2) para redes LAN/MAN y también usado en ADSL, ATM, etc. o PPP: Point to Point, muy extendido en conexiones entre pares de equipos). El protocolo HDLC tiene sus orígenes en el protocolo SDLC (Synchronous Data Link Control) de IBM.

7.2.1. Tipos de estaciones y enlaces

El protocolo HDLC define tres tipos de DTEs (Equipos terminales de datos). Los DTEs pueden funcionar como estaciones primarias, secundarias o combinadas.

- *Primarias*: Una estación se dice que funciona como primaria cuando es la responsable de controlar el enlace de datos, por ejemplo, creando o liberando éste.
- *Secundaria*: Una estación se denomina secundaria cuando funciona siempre bajo el control de una estación primaria y sólo se puede comunicar con primarias.
- *Combinada o Balanceada*: Las estaciones combinadas son todas aquellas que pueden funcionar como estaciones primarias y secundarias. Dependiendo de la situación del enlace funcionara como primaria o en su defecto como secundaria.

Las tramas HDLC que envía una estación funcionando como primaria se denominan *órdenes* y las tramas que envía una estación secundaria se denominan *respuestas*. Por lo tanto, una estación combinada puede enviar tramas orden o tramas respuesta.

Además, de acuerdo a la combinación de los tipos de DTE que determinan un enlace de datos, se establecen dos tipos de configuraciones de enlace de datos.

- *No Balanceada*: Es un enlace de datos formado por una estación primaria y una o más estaciones secundarias sobre un canal con transmisión dúplex o semidúplex.
- *Balanceada*: Es un enlace de datos formado por dos estaciones combinadas sobre un canal con transmisión dúplex o semidúplex.

7.2.2. Modos de funcionamiento

El protocolo HDLC define tres modos de funcionamiento que determinan el modo en cómo se realiza el intercambio de datos entre DTEs.

- *Modo normal*: El modo normal o NRM (Normal Response Mode) es siempre utilizado en configuraciones no balanceadas. En este tipo de configuraciones, la estación que

funciona como primaria se encarga de iniciar la transmisión y la estación secundaria únicamente puede transmitir respuestas a dichas ordenes.

- *Modo asincrono*: El modo asincrono o ARM (Asynchronous Response Mode) es empleado también en configuración no balanceada. La diferencia con respecto a NRM es que con este modo de funcionamiento, la secundaria puede iniciar la transmisión sin tener permiso de la primaria.
- *Modo asincrono balanceado*: El modo asincrono balanceado o ABM (Asynchronous Balanced Mode) es empleado con configuraciones balanceadas del enlace. En este caso, cualquiera de las dos estaciones combinadas puede iniciar la transmisión sin solicitar permiso a la otra.

7.2.3. Estructura de la trama

Al igual que ocurre en otros protocolos del nivel de enlace (Ethernet, PPP, etc), el intercambio de información entre dos DTEs se realiza organizando los bloques de datos en tramas. El protocolo HDLC para realizar el control de errores y el control de flujo emplea siempre un mismo formato de tramas (Figura 7.6) numerado con 3 bits, lo que permite numerar las tramas entre 0 (cuando el valor de los bits es 000) y 7 (cuando es 111). El intercambio de datos en HDLC se realiza siempre usando una transmisión síncrona.

DELIMITADOR (8 bits)	DIRECCION (8 ó k-8 bits)	CONTROL (8 ó 16 bits)	INFORMACION (Variable)	FCS (8 ó 32 bits)	DELIMITADOR (8 bits)
-------------------------	-----------------------------	--------------------------	---------------------------	----------------------	-------------------------

Figura 7-6: Formato de trama HDCL.

El formato de una trama HDLC se organiza de la siguiente manera:

- *Delimitador*: Es el campo que se sitúa al principio y al final de la trama y que determina dónde empieza y termina el bloque de datos que se envía. Al ser el protocolo HDLC un protocolo orientado a bit, emplea delimitación por bits especiales. Así, tanto el delimitador de inicio como delimitador de final emplean la secuencia '0111110', y se emplea el mecanismo de bits de relleno para evitar que la secuencia especial de delimitación aparezca en los bits de datos (véase Apartado 5.4.4). La longitud del campo delimitador es fija y de 8 bits.
- *Dirección*: Identifica a la estación secundaria cuando la trama representa una orden enviada por una estación primaria. También, puede identificar a una estación secundaria cuando la trama representa una respuesta enviada hacia una estación primaria. En el primer caso, el campo identifica a una estación secundaria que funciona como receptor, y en el segundo caso a una estación secundaria que funciona como emisor. El campo dirección tiene una longitud cuyo valor corresponde con múltiplos de 8 bits.
- *Control*: Identifica el tipo de trama que se intercambian los dispositivos. Según el tipo de trama que se intercambien dos estaciones, el campo control será distinto. Así, en el campo control, los dos primeros bits se emplean para identificar el tipo de trama, y el resto de bits se suelen dividir en varios subcampos que se organizan de modo distinto dependiendo del tipo de trama. La longitud de este campo suele ser de 1 o 2 bytes.
- *Datos o Información*: Contiene la secuencia de bits de datos que quieren intercambiar los dos dispositivos. La longitud de este campo es variable. Este campo no está

presente en cualquier tipo de trama, por lo tanto, su longitud puede ir desde 0 bits hasta cualquier múltiplo entero de 8 bits.

- *FCS o campo de detección de errores:* Almacena la secuencia de bits de CRC necesaria para detectar errores en el campo de datos o información. Para generar la secuencia de ceros y unos que determina el CRC se emplean polinomios generadores del tipo CRC-CCITT de 16 o 32 bits.

7.2.4. Tipos de tramas

El protocolo HDLC emplea tres tipos de tramas. Estas tramas se distinguen unas de otras en función de los dos primeros bits del campo control de la trama. Así en función de que estos dos bits se pueden clasificar las tramas como:

- *Información:* Las tramas de información se emplean para el intercambio de datos del nivel de red entre dispositivos (Figura 7.7a).
- *Supervisión:* Estas tramas se emplean para incorporar mecanismos adicionales que permitan supervisar el control de flujo mediante ventana deslizante y control de errores (Figura 7.7b).
- *No-Numerada:* Estas tramas permiten incorporar otras funciones para controlar el enlace de datos (Figura 7.7c).

a)	0 (1 bit)	N(S) (3 bits)	P ó F (1 bit)	N(R) (3 bits)
b)	10 (2 bit)	S (2 bit)	P ó F (1 bit)	N(R) (3 bits)
c)	11 (2 bit)	M (2 bit)	P ó F (1 bit)	M (3 bits)

Figura 7-7: Formato del campo control de 8 bits. a) Tramas de Información. b) Tramas de Supervisión.
c) Tramas No-Numeradas

El campo control en las tramas de Información está formado por 3 subcampos. Un subcampo N(S) que determina el número de secuencia enviada, un subcampo N(R) que indica el número de secuencia que se espera recibir, confirmando así hasta la secuencia N(R)-1, y un subcampo de sondeo (P/F), que permite identificar la última trama asociada a una secuencia de paquetes del nivel de red fragmentados. Este bit P/F, cuando la trama HDLC indica una orden se denomina bit P, y se fija a 1 para solicitar (sondear) una respuesta a la entidad HDLC par situada en el otro extremo de la comunicación. Sin embargo, cuando la trama HDLC indica una respuesta, se denomina F, y su valor se pone a 1 para identificar a la trama devuelta tras una recepción de una orden.

El campo control en las tramas de Supervisión está formado por 3 subcampos. Un código S que determina el tipo de trama de supervisión, un subcampo N(R) que indica el número de secuencia que se espera recibir y un subcampo de sondeo (P/F), que al igual que en la trama de Información, permite identificar la última trama asociada a una secuencia de paquetes del nivel de red fragmentados. Los códigos que puede adoptar S son los mostrados en la figura 7-8.

El campo control en las tramas No-Numeradas, también, está formado por 3 subcampos. Dos subcampos M que indican el tipo de tramas No-Numeradas y un subcampo de sondeo (P/F), que al igual que en las otras tramas, permite identificar la última trama asociada a una secuencia de paquetes

del nivel de red fragmentados. Los códigos que puede adoptar M, también, se muestran en la figura 7-8.

7.2.5. Repertorio básico de instrucciones

El protocolo HDLC tiene un conjunto de órdenes y respuestas que constituyen el conjunto de todo el repertorio de tramas HDLC. En la tabla de la figura 7.8 se han definido las órdenes y respuestas básicas organizadas en función del tipo de trama que se emplea para comunicarlas. Los distintos tipos de estaciones, enlaces, modos de funcionamiento y tipos de tramas permiten definir diferentes variantes del protocolo HDLC, como se verá a lo largo de los siguientes apartados.

Tipo trama	Instrucción	Orden/Respuesta	Descripción
Información Supervisión	I	O/R	Información
	RR	O/R	Receptor preparado
	RNR	O/R	Receptor no preparado
	REJ	O/R	Rechazo
	SREJ	O/R	Rechazo selectivo
No numeradas	SNRM	O	Fijar modo normal no balanceado
	SARM	O	Fijar modo asíncrono no balanceado
	SABM	O	Fijar modo balanceado asíncrono
	DISC	O	Desconectar
	UA	R	Confirmación no numerada

Tipo trama	Instrucción	Campos 'S' o 'M'	Descripción
Supervisión	RR	00	Se ha recibido correctamente hasta la trama indicada en N(R)-1 y se indica en N(R) la siguiente trama que se espera recibir.
	RNR	10	Se ha recibido correctamente hasta la trama N(R)-1, pero no se pueden recibir más tramas
	REJ	01	Se solicita el reenvío de todas las tramas a partir de la indicada en N(R), incluida ésta.
	SREJ	11	Se solicita el reenvío de la trama indicada en N(R)
No numeradas	SNRM	00 001	Fijar modo normal no balanceado
	SARM	11 000	Fijar modo asíncrono no balanceado
	SABM	11 100	Fijar modo balanceado asíncrono
	DISC	00 010	Desconectar
	UA	00 110	Confirmación no numerada

Figura 7-8: Repertorio simplificado de Instrucciones.

7.2.6. Control del enlace en Modo Normal No Balanceado

En un modo de funcionamiento Normal no balanceado (NRM), una de las estaciones se tiene que definir como primaria y la otra como secundaria. De modo, que si se configura la estación A como primaria, será la única que podrá enviar órdenes lo que imposibilita a la estación B para establecer y liberar el enlace de comunicación. Como se muestra en el ejemplo de la figura 7.9, sólo la estación A se encargará de establecer la comunicación, enviando la orden SNRM mediante una trama no numerada. Y la estación B confirmará ese establecimiento con la respuesta UA mediante una trama no numerada.

Posteriormente, si la estación A tiene que enviar datos a la estación B procederá como se indica en la figura 7.9, enviando tantas tramas como permita la ventana deslizante del emisor (We). Y la estación B, podrá recibir tantas tramas sin enviar asentimiento como permita la ventana del receptor (Wr). Para el caso que nos ocupa $We=Wr=2$. Los datos en este caso se encapsularán siempre en tramas de información, I. Y la confirmación o asentimiento de dichos datos se realizará, por parte de B, con tramas de supervisión RR. En este ejemplo, la primera trama de Información indica que $N(S)=0$ y $N(R)=0$ (se envío la trama con número de secuencia 0 y se espera recibir la 0), la segunda trama de Información indica que $N(S)=1$ y $N(R)=0$ (se envío la trama con número de secuencia 1 y se espera recibir la 0), y el flag P/F se ha puesto a 1. En ese mismo ejemplo, la trama de supervisión tiene el valor S a 00, indicando que es una trama RR cuyo campo $N(R)=2$ (se espera recibir la trama con número de secuencia 2) y cuyo campo P/F adopta el valor 0.

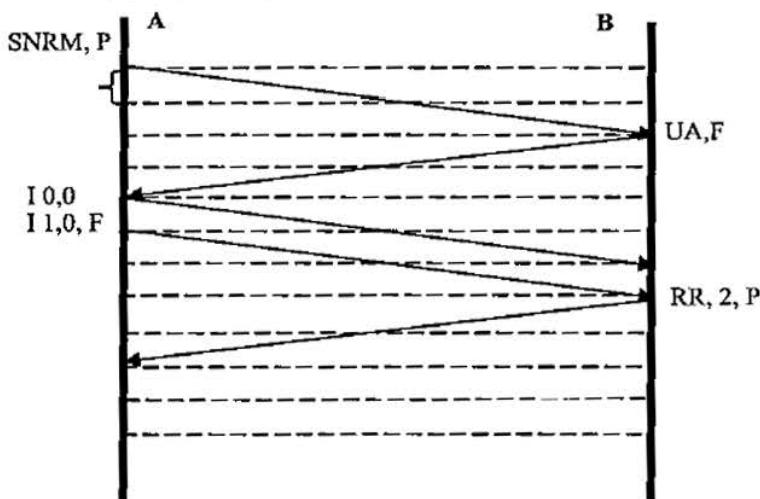


Figura 7-9: Transmisión de datos en Modo Normal no Balanceado.

Supóngase ahora que el modo de funcionamiento y el tipo de estaciones son las mismas que se indicaban en el ejemplo anterior, sin embargo la que necesita enviar datos es la estación B. Es decir, la estación contraria a la que controla el enlace. En este caso (Figura 7.10), una vez el enlace ha sido establecido, la estación A informa mediante una trama de supervisión RR que está preparada para recibir datos de la estación B. Sino ésta no sabrá cuando la estación A está preparada para recibirllos. Posteriormente, la estación B ya está preparada para enviar datos mediante tramas de información, I. Y la confirmación o asentimiento de dichos datos se realizará, de nuevo, con tramas de supervisión RR.

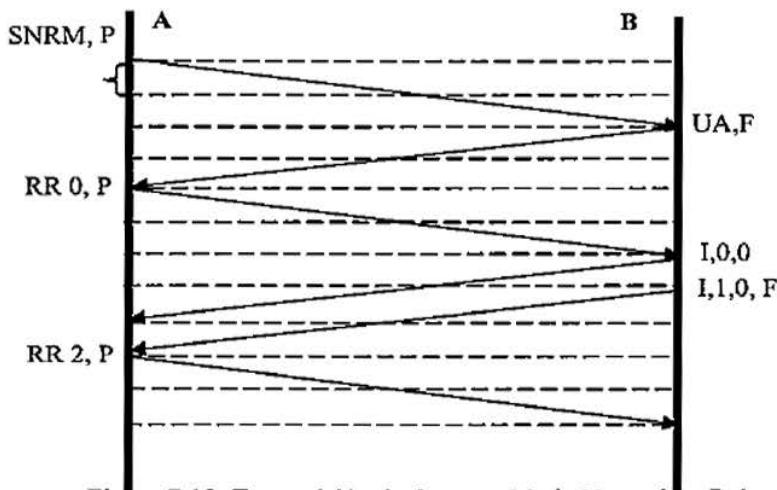


Figura 7-10: Transmisión de datos en Modo Normal no Balanceado.

7.2.7. Control del enlace en Modo Asíncrono No Balanceado

En un modo de funcionamiento Asíncrono no Balanceado (ARM), al igual que en el NRM, una de las estaciones se tiene que definir como primaria y la otra como secundaria. De modo que, será la estación A como primaria, la única que podrá enviar órdenes para establecer y liberar el enlace de datos. Para llevar a cabo el control del enlace, la estación A se encargará de establecer la comunicación enviando la orden SARM mediante una trama no numerada. La estación B confirmará ese establecimiento con la respuesta UA mediante una trama no numerada. Sin embargo, a diferencia del modo NRM, en este caso la estación B que es secundaria puede iniciar la transmisión de datos sin necesidad de esperar un permiso explícito de la estación primaria, A.

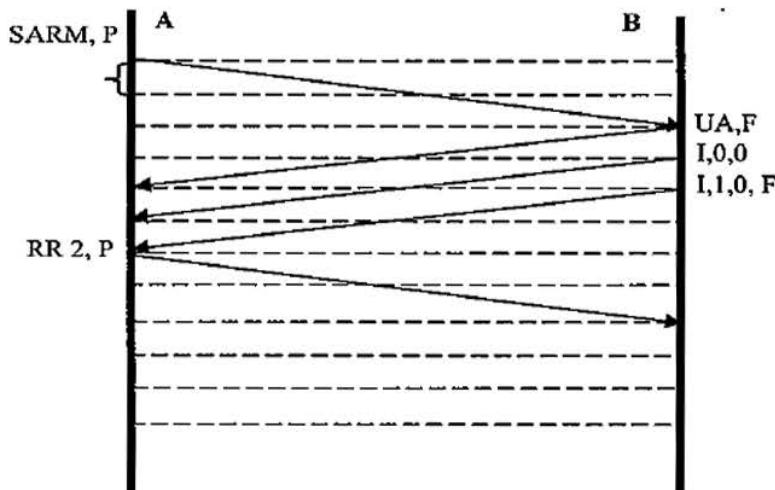


Figura 7-11: Transmisión de datos en Modo Asíncrono no Balanceado.

7.2.8. Control del enlace en Modo Asíncrono Balanceado

En el modo de funcionamiento Asíncrono Balanceado (ABM), las dos estaciones que configuran el enlace son balanceadas. Por lo tanto, ambas estaciones, A y B, pueden controlar el enlace e iniciar la transmisión de datos, sin pedir permiso explícito a la otra estación. Por ejemplo, en la figura 7.12, es la estación A la que controla el enlace y comienza la transmisión de datos. Así, comienza creando un enlace enviando la orden SABM mediante una trama no numerada. Y la estación B confirmará ese establecimiento con la respuesta UA mediante una trama no numerada. Posteriormente, A comienza el envío de datos, I, mediante tramas de información y B los confirma con tramas de supervisión RR.

En la figura 7.13, es la estación B la que controla el enlace y comienza la transmisión de datos. Así, comienza creando un enlace enviando la orden SABM mediante una trama no numerada. La estación A confirmará ese establecimiento con la respuesta UA mediante una trama no numerada. Como se observa en la figura se realiza el mismo intercambio de tramas que en el caso anterior, pero tomando como origen B y como destino A.

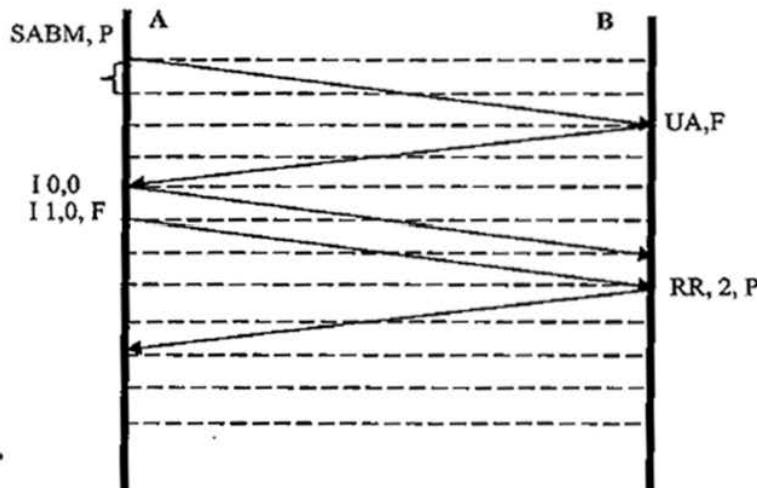


Figura 7-12: Transmisión de datos en Modo Asíncrono Balanceado.

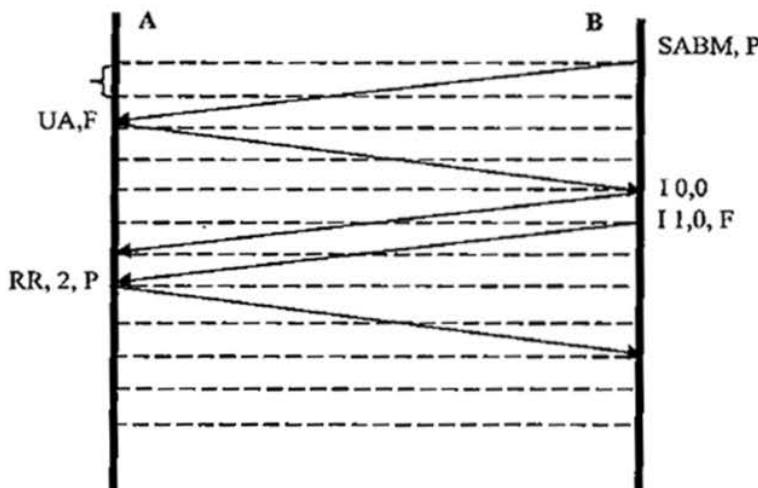


Figura 7-13: Transmisión de datos en Modo Asíncrono Balanceado.

Y en la figura 7.13, es la estación B la que controla el enlace y comienza la transmisión de datos. Así, comienza creando un enlace enviando la orden SABM mediante una trama no numerada. Y la estación A confirmará ese establecimiento con la respuesta UA mediante una trama no numerada. Y se realiza el mismo intercambio de tramas que en el caso anterior, pero tomando como origen B y como destino A.

7.2.9. Configuración del medio de Transmisión y transferencia de datos

El medio de transmisión puede estar configurado en modo *simplex*, *semiduplex* (half-duplex) o *dúplex* (full-duplex). Las líneas dúplex permiten enviar varias tramas en ambos sentidos de la transmisión. Esto implica que si una estación transmite datos, la otra estación no tiene que esperar a que se libere la línea para confirmar esos datos. En la figura siguiente se puede observar un intercambio de datos en un medio de transmisión dúplex. La estación A envía 3 tramas de datos y la estación B va confirmando una a una cada una de las tramas de información, I, recibidas utilizando tramas RR. En ningún caso necesita esperar a que la estación A libere la línea.

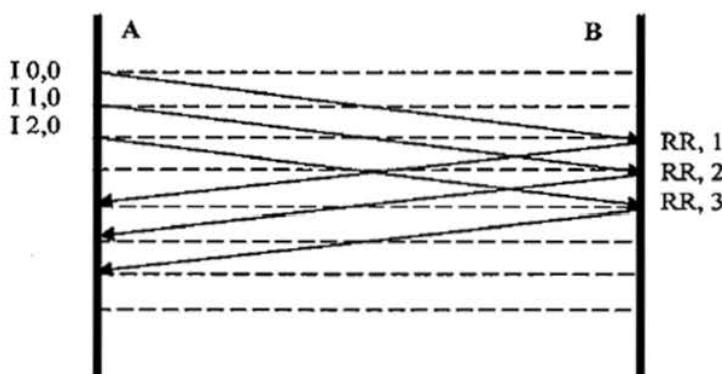


Figura 7-14: Transmisión dúplex.

Sin embargo, si el medio empleado es un medio de transmisión semiduplex. La estación que confirma los datos, no puede confirmar dichos datos hasta que el medio de transmisión queda libre. Así, en la figura 7.15, se observa como la estación A envía una ventana entera de tramas de datos, antes de ceder el turno de transmisión a la estación B para que confirme las tramas enviadas.

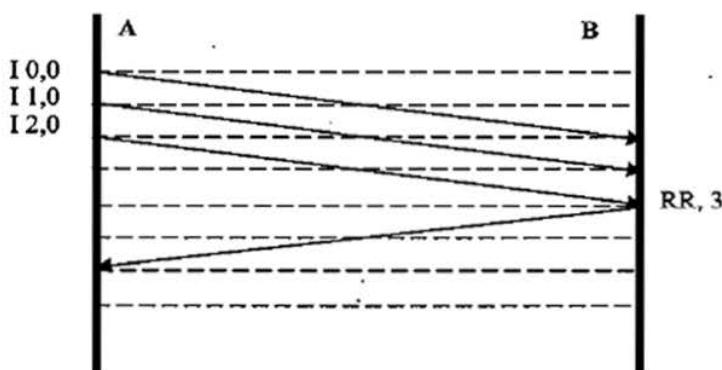


Figura 7-15: Transmisión semiduplex con $W_t = W_c$.

En los ejemplos anteriormente mostrados, se puede deducir fácilmente que en muchos casos el tamaño de la ventana y la configuración del modo de transmisión de la línea pueden ser incompatibles. Así, por ejemplo si la estación A tiene una ventana de emisor, $We=3$, y la ventana de B para recepción es de $Wr=1$, nunca podrá emplearse una línea semiduplex. Ya que el receptor, B, necesitaría confirmar cada trama I que recibe con un RR, justo en el momento de recibirla. Sin embargo, al estar el medio ocupado por las otras tramas de datos, no puede confirmarlas. Es decir, en el ejemplo 7.15, tras recibir la trama I,0; B tendría que confirmarla con un RR,1. Sin embargo, no puede enviar esa trama de supervisión porque la línea de transmisión está ocupada aún con las tramas I1 e I2. Hasta que B no reciba I2 y la línea quede liberada, no podría enviar el RR,1. Si el protocolo HDLC esperara a que llegase la trama I2 y la línea quedase liberada para enviar RR,1 necesitaría poder ir almacenando las tramas recibidas hasta su posterior confirmación. En cuyo caso, el sistema de ventana estaría funcionando como un ventana deslizante selectivo, y puesto que se definió el emisor como $Wr=1$ estaríamos concurriendo en una incongruencia.

Por lo tanto, la situación anterior, HDLC la resuelve descartando los datos cuando la $Wr < We$, y la línea es semiduplex. Así, el diagrama de transmisión correcto queda contemplado en la figura 7.16.

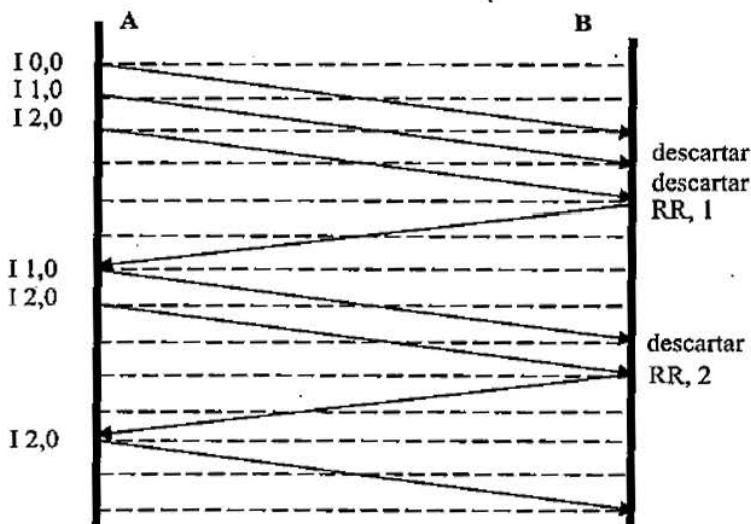


Figura 7-16: Transmisión semiduplex con $Wr < We$.

Una vez establecida la inicialización del enlace, es decir se haya realizado la conexión y esta haya sido aceptada (ver Apartados 7.2.6, 7.2.7 y 7.2.8) es posible comenzar a enviar datos mediante el empleo de tramas de información, I., empezando por el número de secuencia 0. Los campos N(R) y N(S) contendrán los números de secuencia necesarios para llevar a cabo el control de flujo y de errores.

En los ejemplos vistos en las figuras anteriores, siempre se han usado tramas de supervisión del tipo RR para confirmar las tramas recibidas. La trama RR (campo S=00) mediante el campo N(R) indica el número de secuencia de la siguiente trama que espera recibir.

7.2.10. Transferencia de datos con rechazos

Hasta ahora, siempre se ha supuesto que las tramas que intercambian emisor y receptor no tienen errores. Pero, ¿cómo se comportaría el protocolo HDLC en caso de recibir tramas con errores? Para ver este caso se plantean dos situaciones bien distintas. En la primera de ellas, se supone que emisor y receptor funcionan con ventanas de distinto tamaño y además, se sabe que $We > Wr$. En el

segundo caso, se supone que el emisor y receptor trabajan con una ventana deslizante de igual tamaño, $We=Wr$.

En la figura 7.17 se muestra un medio de transmisión configurado en modo dúplex, y por lo tanto, se pueden enviar tramas en ambos sentidos de la transmisión. En el ejemplo, la estación A envía 3 tramas de datos: I,0,0; I,1,0 e I,2,0; y además la estación B detecta errores en la segunda de estas tramas I,1,0. Si en el control de flujo la ventana del emisor, A, es $We=3$ y la ventana del receptor, B, es $Wr=1$, entonces el protocolo HDLC está haciendo uso de una ventana en modo no selectivo. De ahí, que el receptor B no pueda almacenar las tramas que recibe correctamente, tras la trama de datos que contiene errores. Además el receptor B requiere confirmar todas y cada una de las tramas que le van llegando procedentes de A, una a una.

Por lo tanto, la estación B, confirma la primera trama que le llega correctamente y cuando le llega la segunda trama, es decir la trama con errores I,1,0; la descarta y envía una trama de supervisión de tipo rechazo, REJ, para solicitar el reenvío de tramas a partir de la trama que ha fallado. Esto implica que puesto que la trama que falla es I,1,0, la trama de rechazo no selectivo se haya numerado como REJ,1. Y además, la imposibilidad de aceptar tramas después del rechazo, puesto que $Wr=1$, obliga al receptor B a descartar todas las tramas que llegan a continuación.

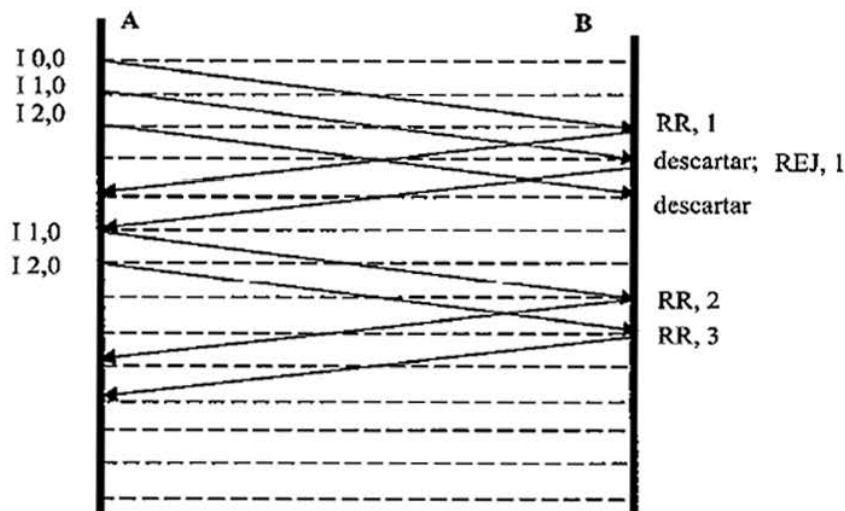


Figura 7-17: Transmisión con rechazo no selectivo.

En la figura 7.18 se muestra un medio de transmisión configurado en modo semidúplex y, por lo tanto, no se pueden enviar simultáneamente tramas en ambos sentidos de la transmisión. En el ejemplo, la estación A envía 3 tramas de datos: I,0,0; I,1,0 e I,2,0; y además la estación B detecta errores en la segunda de estas tramas I,1,0. Si en el control de flujo la ventana del emisor, A, es $We=3$ y la ventana del receptor, B, también es $Wr=3$, entonces el protocolo HDLC está haciendo uso de una ventana en modo selectivo. Consecuentemente, el receptor B puede almacenar las tramas que recibe correctamente, tras la trama de datos que contiene errores. Además, el receptor B no requiere confirmar todas y cada una de las tramas que le van llegando procedentes de A, una a una, ya que puede emplear una única trama para confirmar todas las tramas enviadas en la ventana.

Por este motivo, la estación B, no confirma la primera trama sino que espera confirmar las tres tramas enviadas por A. Sin embargo, como la segunda trama, es decir la trama I,1,0 llega con errores; es descartada por B, sin necesidad de descartar las que llegan tras ésta. Así, la estación B, envía una trama de supervisión de tipo rechazo selectivo, SREJ para solicitar el reenvío de la trama que ha fallado y el resto de tramas I,0,0 e I,2,0 se almacenan temporalmente en un buffer a la espera de su

confirmación. Esto implica que puesto que la trama que falla es I,1,0, la trama de rechazo selectivo se haya numerado como SREJ,1.

A diferencia de los ejemplos vistos en el apartado 7.2.9, en este apartado, además de emplearse tramas de supervisión del tipo RR para confirmar las tramas recibidas, se han empleado tramas de supervisión REJ (campo S=01) y SREJ (campo S=11) para solicitar el reenvío de datos erróneos.

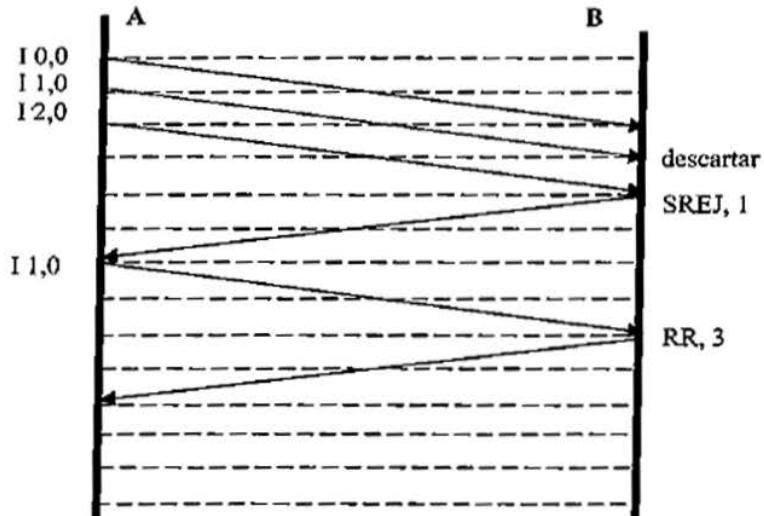
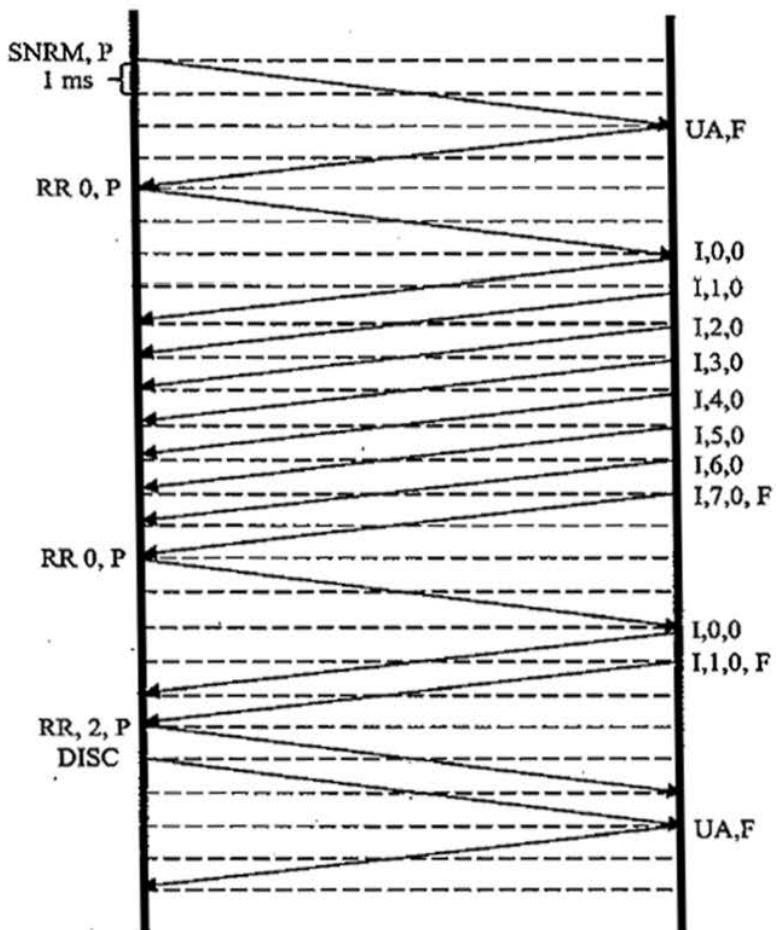


Figura 7-18: Transmisión con rechazo selectivo.

Ejemplo 7.1*Control de Flujo sin errores en HDLC.*

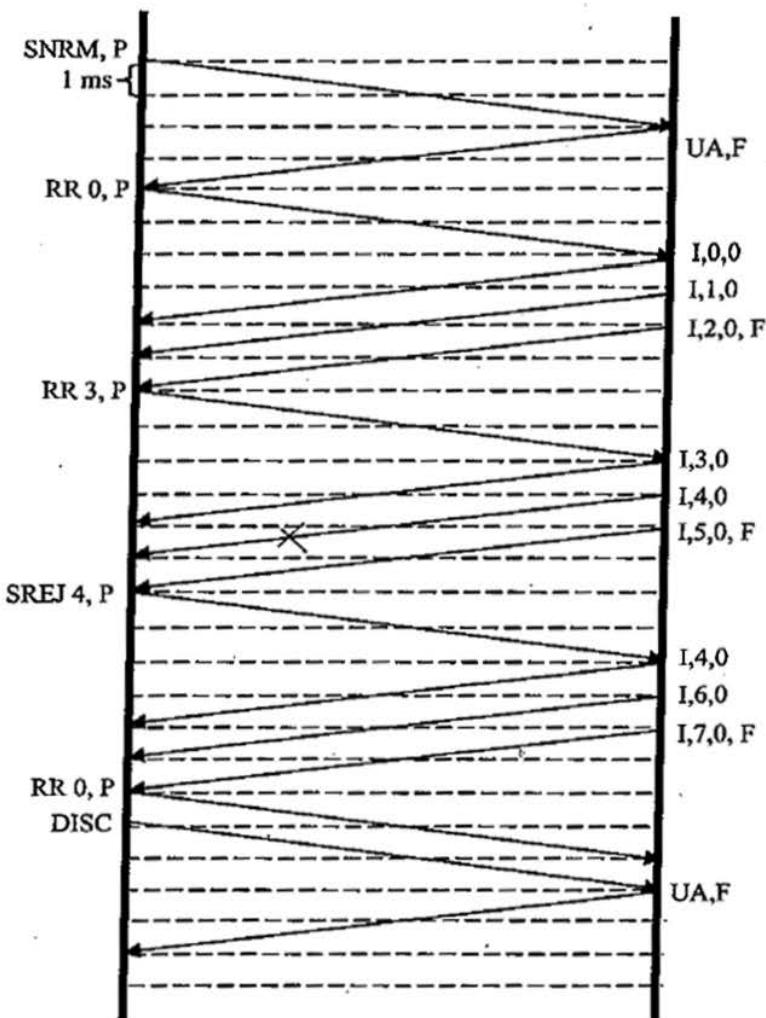
Se dispone de dos estaciones, A y B, separadas una distancia de 10Km, conectadas mediante una línea dúplex, con una capacidad de 1Mbps y una velocidad de propagación de 300m/s. Para comunicarse, ambas estaciones, emplean un protocolo de nivel de enlace del tipo HDLC en modo Normal no Balanceado, empleando ventana deslizante de un máximo de 8 tramas en el Emisor y de 8 tramas en el receptor, así como tramas de tamaño de 1000bits. Si se desea enviar un fichero de 10000bits desde la estación B hasta la A, siendo A la estación primaria. Calcula el diagrama de transmisión de tramas, indicando la instrucción según el tipo de trama, así como la numeración N(R), N(S) y M, en el caso de ser necesario, para el caso en que el canal esté libre de errores de transmisión. Además, en el diagrama deberán aparecer las fases de establecimiento, transferencia de datos e información y liberación del enlace. (Se aconseja emplear el repertorio básico de instrucciones que se facilita en la figura 7.8)



Ejemplo 7.2

Control de Flujo con errores en HDLC.

Se dispone de dos estaciones, A y B, separadas una distancia de 15m, conectadas mediante una línea semidíplex, con una capacidad de 100Mbps y un tiempo de propagación de 2ms. Para comunicarse, ambas estaciones, emplean un protocolo de nivel de enlace del tipo HDLC en modo Normal No Balanceado con temporizador de 7ms. Se sabe que Emisor y Receptor emplean ventana deslizante de un máximo de 3 tramas, y que el tamaño de una trama de datos es de 1000bits. Si se desea enviar un fichero de 8000bits desde la estación B hasta la A, y A es la estación primaria. Calcula el diagrama de transmisión de tramas, indicando la instrucción según el tipo de trama, así como la numeración N(R), N(S) y M, en el caso de ser necesario, para el caso en que se produzca un error en la quinta trama enviada. Además, en el diagrama deberán aparecer las fases de establecimiento, transferencia de datos e información y liberación del enlace. (Se aconseja emplear el repertorio básico de instrucciones que se facilita en la figura 7.8) ¿Cuánto tiempo transcurre desde que la máquina B envía la primera trama hasta que envía la última?



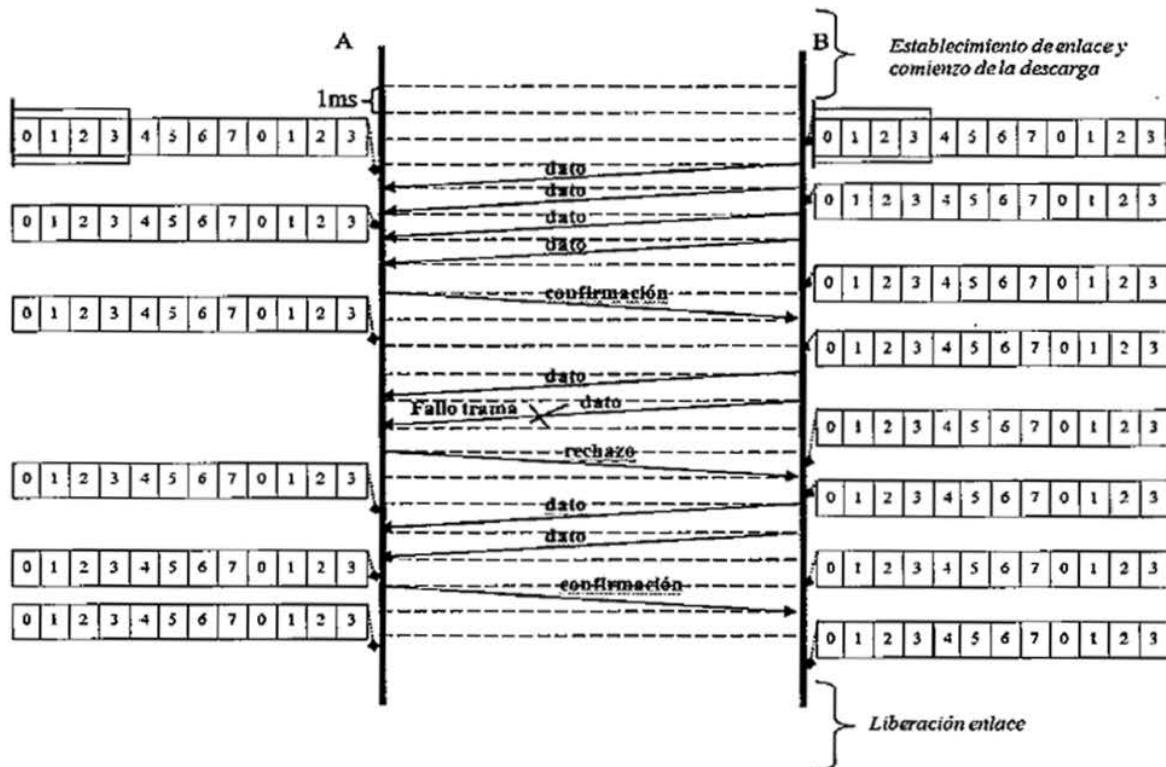
El tiempo transcurrido entre que envía la primera trama I,0,0 y envía la última I,7,0 es de 14 ms.

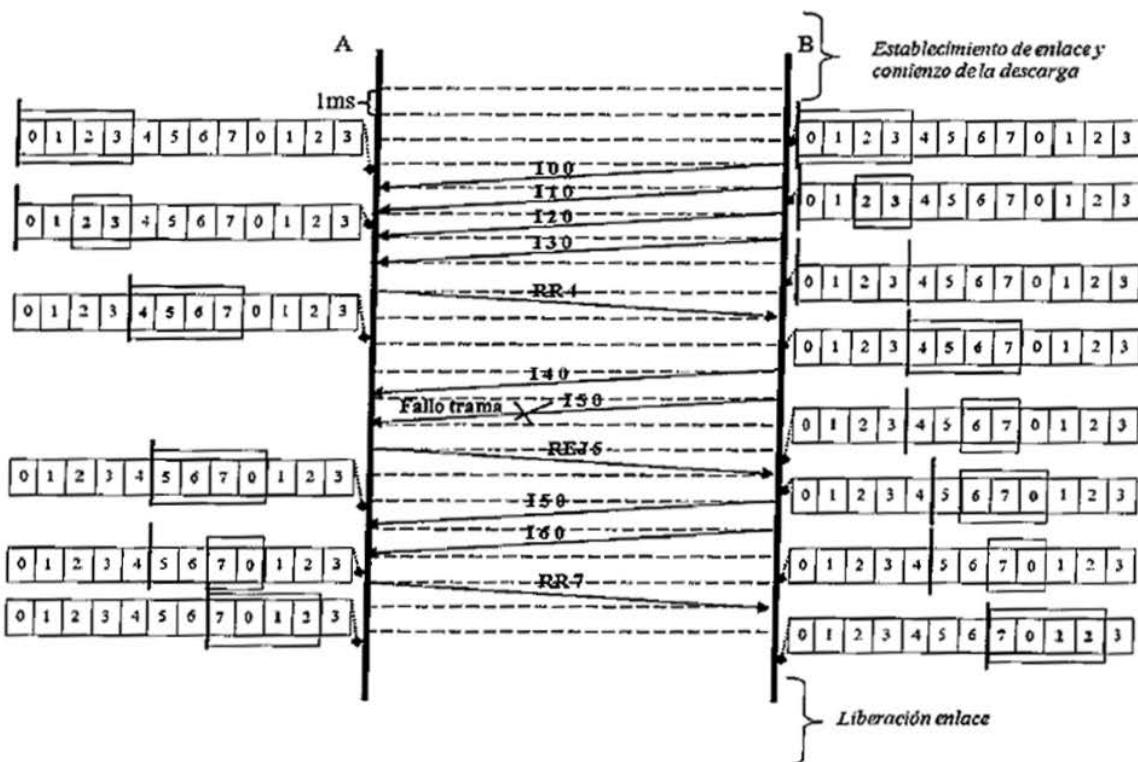
Ejemplo 7.3

Control de la ventana deslizante en HDLC.

Dado el diagrama de transmisión de tramas de la figura, en el cual se refleja el intercambio de información entre dos estaciones, A y B, separadas una distancia de 15m, conectadas mediante una línea semidúplex, con una capacidad de 10Mbps y un tiempo de propagación de 1ms. Para comunicarse, ambas estaciones emplean un protocolo de nivel de enlace del tipo HDLC en modo Asíncrono Balanceado con temporizador de 7ms. Se sabe que Emisor y Receptor emplean ventana deslizante de un máximo de 4 tramas, y que el tamaño de una trama de datos es de 10Kbytes. Si se desea enviar un fichero de 70Kbytes desde la estación B hasta la A. Además, se sabe que el receptor puede utilizar tramas de confirmación para indicar al emisor la recepción correcta y de rechazo simple no selectivo para indicar la recepción incorrecta de una trama.

- Identifica los tipos de tramas HDLC, identificando el subtipo en función del campo de control, que como mínimo se necesitaría utilizar en el nivel de enlace para transmitir el fichero y realizar las funciones de establecimiento, liberación del enlace.
- Completa el siguiente diagrama correspondiente a las tramas de intercambio de datos durante la descarga del fichero, representando fielmente el estado de la ventana deslizante en el emisor (B) y en el receptor (A).





7.3. Protocolo PPP

El protocolo PPP es quizás el protocolo del nivel de enlace más empleado en cualquier tipo de conexiones punto a punto. Este protocolo se emplea tanto en conexiones por línea alquilada entre dos routers como en accesos ADSL o telefónicos entre el equipo de usuario (DTE o router del usuario) y el router del ISP (Internet Service Provider). Algunas de estas situaciones son:

- Cuando se quiere establecer una conexión entre un particular y un proveedor de servicios a Internet (ISP), a través del empleo de un módem telefónico.
- Cuando se quiere establecer una conexión entre un particular y un ISP de cablemodem o xDSL, en cuyo caso, el protocolo PPP se encapsula sobre Ethernet formando lo que se conoce como protocolo PPPoE (Point to Point over Ethernet).
- Cuando se quiere establecer una conexión entre un particular y un ISP de cablemodem o xDSL empleando un modo de transferencia asíncrona (ATM). En este caso, el protocolo PPP se encapsula sobre ATM formando lo que se conoce como protocolo PPPoA (Point to Point over ATM).
- Cuando se accede a recursos de una red privada remota mediante protocolos de VPN (Virtual Private Network).

Además, del protocolo PPP, existen otros protocolos de nivel de enlace para conexiones punto a punto, como el protocolo SLIP (Serial Line Internet Protocol). No obstante, aunque el protocolo SLIP es un protocolo que está específicamente diseñado para TCP/IP, la tendencia actual es el empleo de PPP. Las diferencias entre SLIP y PPP radican en que SLIP necesita configurar algunos parámetros de forma estática. Así, SLIP necesita conocer la dirección IP de los dos extremos de la comunicación, además de necesitar conocer la unidad máxima de transmisión del nivel de enlace (MTU), por lo tanto se hace difícil su uso cuando alguna de estas direcciones IP se asignan dinámicamente.

El protocolo PPP fue diseñado y desarrollado por IETF (Internet Engineering Task Force) en 1993 para solventar estas deficiencias, y proporcionar un estándar internacional para las conexiones punto a punto. El protocolo PPP define su estándar en el documento RFC 1661, el PPPoA está definido en RFC 2364 y el PPPoE en la RFC 2516.

Por lo tanto, el protocolo PPP proporciona un método estándar multiprotocolo, independiente del protocolo del nivel de red empleado, ya que soporta los más habituales (IP, IPX y Appletalk) sobre enlaces simples punto a punto entre dos máquinas. En estos casos, la comunicación es bidireccional (full-dúplex) y los paquetes se entregan en orden.

PPP utiliza el mismo control de enlace de datos de alto nivel que HDLC como base para encapsular datagramas multiprotocolo y llevar a cabo la detección de errores. Además, hace uso de otros protocolos auxiliares como son:

- *LCP* (Link Control Protocol) para establecer la configuración del enlace de datos.
- *IPCP* (IP Control Protocol) y *NCP* (Network Control Protocol) para el control, establecimiento y configuración del protocolo IP en el caso de IPCP o de los distintos protocolos del nivel de red que permite encapsular en el caso de NCP.
- *CHAP* (Challenge Handshake Authentication Protocol) y *MS-CHAP* (Microsoft CHAP) que permiten el control de acceso y la autenticación de usuarios.
- *CCP* (Compression Control Protocol) y *MPPC* (Microsoft Point-to-Point Compression Protocol) que son usados para establecer y configurar la compresión de datos sobre PPP.
- *MPPE* (Microsoft Point-to-Point Encryption) que está diseñado para transmitir datagramas encriptados sobre enlaces PPP.

7.3.1. Estructura de la trama

La estructura de una trama PPP es la misma que la de una trama HDLC, puesto que PPP toma muchas cosas en común de éste último protocolo. El protocolo PPP emplea el mismo formato de trama comentado en la Figura 7.6. No obstante, algunos de los campos que en HDLC podían adoptar distintos valores, en PPP adquieren valores fijos (Figura 7.19).

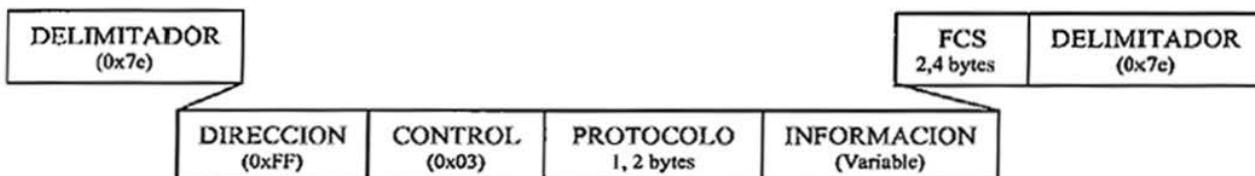


Figura 7-19: Formato de trama PPP.

El formato de una trama PPP se organiza de la siguiente manera:

- *Delimitador*: Todas las tramas comienzan y terminan con el byte '01111110' (0x7e).
- *Dirección*: El campo dirección suele tener asignado el valor broadcast '11111111' (0xFF).
- *Control*: El campo control tiene como valor predeterminado '00000011' (0x03). Este valor indica una trama no numerada. Por lo general, PPP no proporciona numeración

de tramas de datos y de asentimientos, aunque es posible numerar tramas y asentimientos cuando el ambiente en el que se transmite es ruidoso y se requiere.

- **Protocolo:** El campo protocolo es de 1 o 2 bytes y su valor identifica el contenido del datagrama almacenado en el campo información. Para que el valor del campo sea válido, el bit menos significativo del byte menos significativo debe ser '1', y el bit menos significativo del byte más significativo debe ser '0'. En caso de no ser así, se considerará el paquete almacenado en el campo Información como irreconocible.
 - **Información:** Contiene la secuencia de bits de datos correspondiente a un nivel superior cuyo protocolo se especifica en el campo Protocolo. Esta es la información que intercambian los dos dispositivos. La longitud es variable y puede tener un valor mínimo de 0 bytes y un valor máximo determinado por la MTU (por defecto 1500 bytes). A la información se le puede agregar un relleno ('padding') con un número arbitrario de bytes para alcanzar el valor de MTU.
 - **FCS:** Es un campo de chequeo de errores o verificación. Este campo almacena una suma de comprobación y normalmente ocupa 2 bytes, aunque puede llegar a ocupar 4 bytes.

7.3.2. Tipos de tramas

El protocolo PPP encapsula varios tipos de paquetes en función del valor que adopte el campo *Protocolo*. Los identificadores del campo protocolo están especificados en la norma RFC 1661. Los más importantes son:

- *IP*: El campo toma el valor (0x0021).
 - *LCP*: El campo toma el valor (0xc021).
 - *PAP*: El campo toma el valor (0xc023).
 - *CHAP*: El campo toma el valor (0xc223).
 - *IPCP*: El campo toma el valor (0x8021).
 - *CCP*: El campo toma el valor (0x80FD).

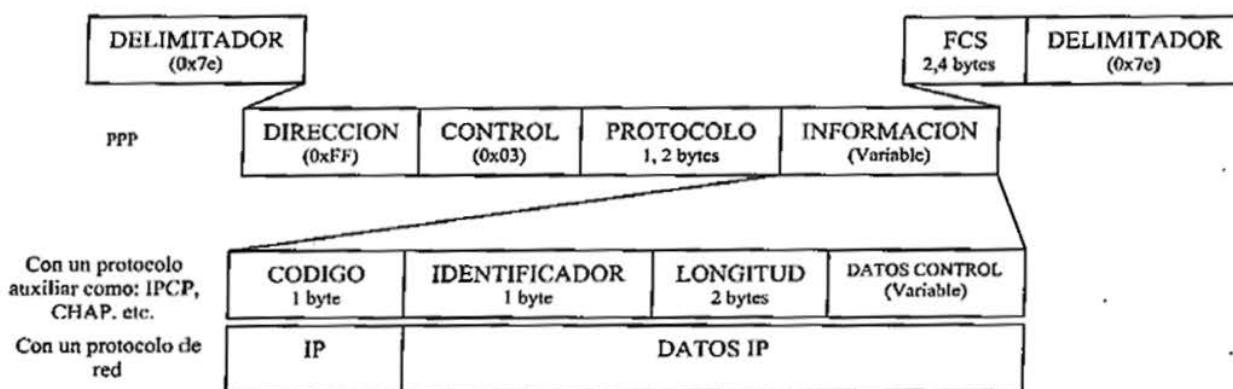


Figura 7-20: Formato de tipos de paquetes para PPP.

Cualquier tipo de paquete PPP tiene el formato que se indica en la Figura 7.20. En función del valor del campo *protocolo* (NCP, CHAP, etc.), el campo de *Información* de PPP puede transportar bits adicionales de: *código*, *identificador* (tipo) y *longitud* (cuando el protocolo es LCP, NCP, IPCP, CHAP, etc.) o un datagrama de red (cuando el protocolo es IP, IPX, etc.). El campo *código* se emplea para determinar la clase de paquete NCP, clase de paquete CHAP, etc. Esto es así porque dentro de los distintos protocolos que puede encapsular PPP, hay diversas clases de paquetes para cada uno de ellos (por ejemplo, ACK, NACK, Configure Request, Reject, etc.). Y los otros dos campos permiten definir el número de paquete (*identificador*) y la longitud que ocupa el paquete (*longitud*).

Paquetes LCP

Los paquetes LCP (Link Control Protocol) se emplean para negociar los parámetros del nivel de enlace en el inicio del establecimiento de la conexión. Hay 3 tipos de paquetes LCP:

- Paquetes de establecimiento que son empleados para establecer y configurar el enlace.
- Paquetes de finalización que son empleados para cerrar la conexión.
- Paquetes de testeo que son empleados para mantenimiento del enlace y chequeo de errores.

Paquetes NCP o IPCP

Los paquetes NCP (Network Control Protocol) se emplean para negociar opcionalmente parámetros de configuración para poder encapsular multiprotocolos, como por ejemplo la asignación dinámica de dirección de red. Sin embargo, los paquetes IPCP se emplean para establecer y configurar el protocolo de red IP sobre PPP.

Así, para establecer la comunicación sobre un enlace punto a punto mediante el empleo de PPP, cada extremo del enlace envía primero paquetes LCP para configurar y chequear el enlace de datos. Una vez que el enlace queda establecido, los extremos intercambian mediante el protocolo PPP paquetes NCP encapsulados para elegir y configurar uno o más protocolos de red elegidos (en el caso de que se vaya emplear paquetes IP, el protocolo auxiliar empleado será IPCP en vez de NCP). Despues, los datagramas de cada protocolo del nivel de red podrán ser enviados a través del enlace. Finalmente, el enlace permanecerá activo hasta que una serie de paquetes NCP (o IPCP según caso) y/o paquetes LCP cierren la conexión, o hasta que ocurra un evento externo como por ejemplo un temporizador expirado debido a una detección de inactividad.

Más formalmente, los paquetes NCP o IPCP (RFC 1332) que se intercambian son los siguientes:

- El router del ISP manda un NCP o IPCP con código 1, ‘Configure Request’, indicando su dirección IP y solicitando la del usuario.
- El cliente responde con un NCP o IPCP con código 2, ‘Configure ACK’.
- El cliente envía un NCP o IPCP con código 1, dándole su IP actual (0.0.0.0 si no está aun configurada) para que router del ISP verifique si es válida o le asigne otra en caso de no serlo.
- El router del ISP responde con un NCP o IPCP con código 3, ‘Configure NACK’, negando esa IP y ofreciéndole una dirección válida.
- El cliente solicita con un NCP o IPCP con código 1 conectarse con esa IP ofrecida.
- El router del ISP responde con un NCP o IPCP con código 2 aceptando la IP ofrecida.

Paquetes PAP

Los paquetes PAP (Password Authentication Protocol) se emplean para autenticar un usuario que intenta acceder a un servidor de acceso remoto o a un proveedor de servicios a través de una línea alquilada. Los paquetes PAP se emplean sobre PPP para transmitir contraseñas escritas mediante códigos ASCII y además éstas están sin cifrar.

Paquetes CHAP

Los paquetes CHAP (Challenge Handshake Authentication Protocol), también, se emplean para autenticar un usuario que intenta acceder a un servidor de acceso remoto o a un proveedor de servicios a través de una línea alquilada. La diferencia con PAP es que aquí sí se produce encriptación de la información.

CHAP permite autenticar al usuario de la siguiente manera: Primero, se establece el enlace, y posteriormente, el servidor de acceso manda un mensaje 'Challenge' CHAP para que el usuario se identifique. A continuación, el usuario responde con un valor codificado almacenado en otro paquete CHAP 'Response'. Finalmente, el servidor verifica la respuesta, comprobando el valor codificado. Si este valor coincide con el esperado por el servidor, entonces se informará al cliente de la autenticación, mediante un mensaje CHAP 'Success'. Si el valor no coincide, entonces se cerrará la conexión y se informará con un mensaje CHAP 'Failure'. El mensaje de autenticación que solicita el servidor se realiza y repite de manera aleatoria, durante la conexión.

Los dos extremos del enlace pueden intercambiar distintos paquetes CHAP, los cuales se definen a través del campo *Código* (Figura 7.20). Así, paquetes CHAP con el código a 1 indican que se envía un mensaje 'Challenge'. El mensaje 'Response' del usuario es un paquete CHAP con código 2 que indica el identificador de trama, el password y el mensaje de 'challenge' todo ello concatenado y codificado con un algoritmo de cifrado (por ejemplo MD5), y a esta cadena se le añade, también, el login sin cifrar. El servidor responde con un CHAP 'Success' código 3 si es la respuesta esperada y con CHAP 'Failure' código 4 no es la esperada.

En la actualidad existen variantes del protocolo CHAP, un claro ejemplo de ello es el caso de MS-CHAP (Microsoft CHAP) que es usado por sistemas MS-Windows (Microsoft Windows).

7.3.3. Funcionamiento de PPP

Para describir el funcionamiento general del protocolo PPP (Figura 7.21) y cómo hace uso de los distintos paquetes que encapsula (PAP, CHAP, LCP, NCP, etc.) se ha optado ilustrarlo con un ejemplo de conexión a un router de un proveedor de servicios de Internet (ISP), por parte de un DTE de usuario que emplea un modem analógico de 56Kbps, que es un caso sencillo.

El funcionamiento de PPP se puede dividir en varias fases o etapas:

- *Establecimiento de la conexión:* Inicialmente, el usuario llama al router del proveedor de servicios, empleando el modem analógico. Así establece el circuito físico. Una vez, se ha establecido el circuito conmutado, el usuario hace llegar al router del proveedor de servicios un paquete LCP para negociar los parámetros del enlace de datos que empleará PPP en el intercambio de datos. Haciendo uso de paquetes LCP se negocia también el método de autenticación que se empleará (PAP, CHAP o no se requiere).

- *Autentificación:* Si durante la fase de establecimiento se ha negociado PAP, no se requiere autentificación cifrada. Sin embargo, si por el contrario, se ha negociado CHAP se requerirá autentificación cifrada de una contraseña.
- *Configuración de la red:* A continuación, mediante paquetes NCP o IPCP el protocolo PPP negocia los parámetros de la capa de red. Así, se configura el protocolo de red que se usa de entre todos los que se pueden encapsular sobre PPP. Por ejemplo, para un protocolo de nivel de red IP, se negocia mediante IPCP la asignación dinámica de una IP valida. A partir de este momento, el DTE de usuario ya es una máquina de Internet y puede enviar y recibir paquetes.
- *Negociar otros parámetros:* En esta fase es cuando se realiza la elección de otras opciones como puede ser la compresión de datos y/o la encriptación. Para ello se hace uso de algunos de los protocolos auxiliares del tipo CCP, MPPC, MPPE, etc.
- *Transmisión de datos:* Es en esta fase cuando se mandan y reciben datagramas de nivel red encapsulados en tramas PPP. Durante esa fase pueden aparecer paquetes LCP encapsulados en tramas PPP para comprobar aleatoriamente si la línea sigue activa o ha quedado inactiva durante algún periodo de tiempo.
- *Finalización:* Finalmente, cuando el intercambio de datos haya finalizado. El DTE del usuario terminará la comunicación, liberando los parámetros del nivel de red con NCP. Posteriormente, enviará paquetes LCP para liberar el enlace de datos. Y ya para concluir, cuando se cuelgue el teléfono mediante una orden del modem, se habrá liberado el circuito que había quedado definido en la capa física.

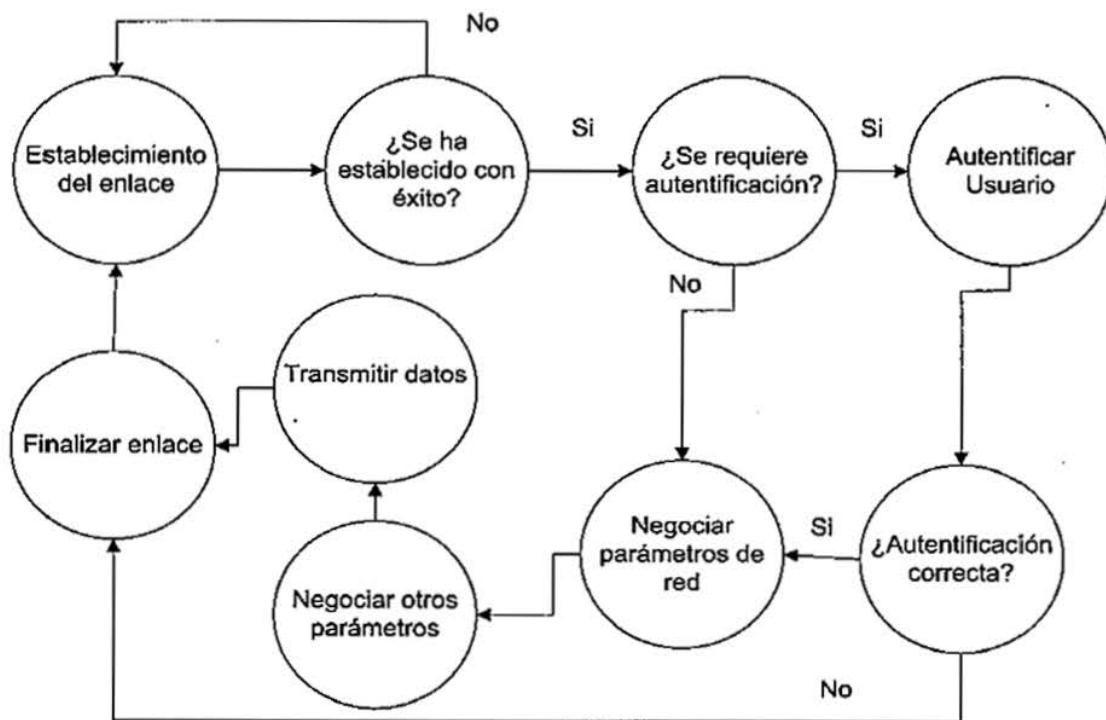


Figura 7-21: Funcionamiento general de PPP.

Capítulo 8. DISEÑO DEL NIVEL DE RED

A lo largo de este capítulo se van a describir las funciones básicas de la capa de red del modelo OSI, destacando dentro de éstas las funciones de encaminamiento. Para ello, tras describir las funciones o servicios que proporciona el nivel de red al nivel de transporte, se describirán los métodos de commutación de circuitos y de paquetes. Finalmente, se presentarán los tipos de encaminamiento empleados en commutación de paquetes por datagramas: inundación, encaminamiento estático y adaptativo.

CAPÍTULO 8

- 8-1 Nivel de Red
- 8-2 Servicios de la capa de red
- 8-3 Métodos de conmutación
- 8-4 Algoritmos de encaminamiento
 - Encaminamiento fijo o estático
 - Encaminamiento por inundación
 - Encaminamiento adaptativo o dinámico

OBJETIVOS

- Comentar las funciones básicas de la capa de red de una arquitectura de comunicaciones
- Mostrar el tipo de servicios que la capa de red puede ofrecer a la capa de transporte en una arquitectura de comunicaciones
- Explicar el funcionamiento interno de una red desde un punto de vista de métodos de conmutación para determinar el camino entre origen y destino
- Exponer y comentar los algoritmos que permiten determinar el modo en cómo los datos se transmiten entre dispositivos de interconexión intermedios

TERMINOLOGÍA – PALABRAS CLAVE

- | | | |
|--------------------------------------|-----------------------|-----------------------------|
| • Paquete/PDU | • Datagramas | • Encaminamiento adaptativo |
| • Servicios orientados a conexión | • Número de secuencia | |
| • Servicios no orientados a conexión | • Contador de salto | |
| • Conmutación de circuitos | • Camino | |
| • Conmutación de paquetes | • Inundación | |
| • Circuitos virtuales | • Encaminamiento | |

8.1 Nivel de red

Este capítulo está destinado a describir las principales funciones del nivel de red de la arquitectura OSI. A nivel de red la unidad de datos de protocolo (PDU) empleada suele denominarse paquete. El principal objetivo del nivel de red del modelo OSI es el realizar el correcto encaminamiento de los paquetes de información que van dirigidos desde el equipo origen al destino. Para realizar estas funciones de encaminamiento suelen emplearse los denominados routers o encaminadores que son capaces de interconectar entre sí redes con distintas topologías y niveles de enlace. Este tipo de dispositivos, implementan los niveles físico y enlace para cada interfaz de red, además de un nivel de red común a todos los interfaces. Véase, por ejemplo, la Figura 8-1 en la que se ha hecho uso de un router para interconectar 3 redes: una con topología en estrella, una con topología en bus y otra en anillo. La función fundamental del router es la de encaminar los paquetes de información procedentes de un determinado nodo origen de manera que se calcule el mejor camino para llegar a un nodo destino. Para determinar dicho camino pueden emplearse distintos criterios como puede ser el cálculo del camino que proporcione un menor número de saltos entre origen y destino (que pase por una menor cantidad de routers), cálculo del camino con el cual se alcance el destino con menor congestión, etc.

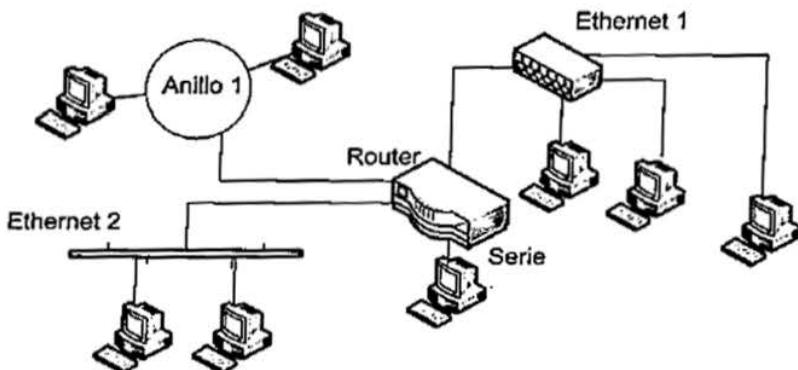


Figura 8-1. Router interconectando redes con distinta topología.

A lo largo de este capítulo se van a describir las funciones básicas de la capa de red del modelo OSI, destacando dentro de éstas las funciones de encaminamiento. Para ello, tras describir las funciones o servicios que proporciona el nivel de red al nivel de transporte, se describirán los métodos de conmutación. En los métodos de conmutación basados en circuitos virtuales el camino que seguirán los paquetes desde el origen al destino será el mismo durante toda la conexión. Sin embargo, empleando datagramas el encaminamiento puede variar, es decir, cada paquete puede seguir un camino distinto desde el origen al destino. Es por esto por lo que el capítulo finalizará realizando una clasificación de los tipos básicos de encaminamiento de paquetes empleando distintas estrategias y criterios. Será en el capítulo 9 cuando se describan algoritmos concretos que emplean protocolos de red reales para realizar el encaminamiento.

8.2 Servicios de la capa de red

Los servicios de red proporcionados a la capa superior deben ser capaces de aislar a la capa de transporte de tareas de direccionamiento. Como se ha indicado anteriormente esta es la principal función del nivel de red de manera que en el nivel superior se indicará el destino del paquete de información y a nivel de red se irá encaminando dicho paquete a través de los nodos intermedios hasta alcanzar el nodo destinatario. Además, la capa de red también aislará a las capas superiores de los diferentes esquemas de red que pueden haber a lo largo del camino del paquete desde el origen al destino (el nivel de enlace puede ser diferente y esto no afectará al nivel de transporte). Otro aspecto del que será transparente la capa de transporte y superiores es el coste en la conexión, es decir, número de encaminadores necesarios, congestión, etc. De esta forma, los servicios que la capa de red ofrece a la capa de transporte se caracterizan por las siguientes propiedades:

- La capa de red ofrece servicios que son independientes del nivel de enlace en el que se encuentre el equipo correspondiente. Es decir, ofrece a la capa de transporte una serie de servicios independientes de si se emplea commutación de circuitos, commutación de paquetes, tecnologías LAN, comunicaciones móviles, etc.
- Todos los aspectos relativos al encaminamiento de los paquetes desde el nodo origen al destino son considerados en la capa de red, por lo tanto, la capa de transporte debe ser transparente a la cantidad, tipo y topología de las subredes presentes.
- El nivel de red debe proporcionar un sistema de identificación de cada equipo de la red independientemente del tipo de red en el que se encuentre.

Además, los servicios ofrecidos por el nivel de red pueden clasificarse en servicios orientados a conexión y servicios no orientados a conexión. En el primer caso, la complejidad de la comunicación recae en los DTEs origen y destino, mientras que en el segundo la complejidad de la comunicación recae en los nodos intermedios o encaminadores de la subred. Las características de cada uno de estos dos tipos de servicios son las siguientes:

- Servicios orientados a conexión. En este caso sí que es necesario antes de enviar los datos establecer una conexión. Habitualmente, al establecer la conexión se negocian parámetros tales como el coste, la calidad del servicio y parámetros del control de flujo: ventanas y modo selectivo, por ejemplo. Como resultado de ese establecimiento de conexión se obtendrá un valor identificativo de la conexión. A diferencia de los servicios no orientados a conexión ahora sí que se garantiza el orden de los datos transmitidos. La transmisión de información se realiza de forma secuencial y bidireccional. Además, se realiza control de flujo para evitar la saturación en el equipo destinatario de la información. Al igual que es necesario el establecimiento de conexión, también se requiere finalizarla. Por lo tanto, una vez realizada la transmisión de datos será necesario realizar la liberación de la conexión liberando también el identificador para poder ser empleado en sucesivas conexiones. Este es el tipo de servicio empleado en redes para telefonía. En general, este servicio es empleado cuando se exige una alta fiabilidad.
- Servicios no orientados a conexión. Se denominan así ya que no es necesario establecer ningún tipo de conexión entre el nodo origen y el destino antes de realizar el envío de los datos. Además, los paquetes pueden circular desordenados por la red por lo que el equipo destino deberá realizar la ordenación de todos los paquetes antes de enviarlos a la capa superior. Esta aproximación es la empleada por el nivel de red IP dentro de la arquitectura TCP/IP. En general, este servicio es empleado cuando se pretende una mayor rapidez frente a una menor fiabilidad.

8.3 Métodos de conmutación

Independientemente del tipo de servicio proporcionado por la capa de red, el funcionamiento interno de la subred puede ser mediante conmutación de circuitos o conmutación de paquetes (véase apartado 1.2.3):

- Conmutación de circuitos. Se crea una conexión física directa entre dos dispositivos, estableciéndose un circuito dedicado entre las estaciones para la comunicación (ver Figura 8-2). Es el método empleado por las Redes Telefónicas Conmutadas originales o clásicas para la transmisión de señales de voz y se caracteriza por:
 - En este caso se emplea un servicio orientado a conexión por lo que previo a la transmisión de datos es necesario realizar un establecimiento de la conexión. Una vez establecida se realiza la transmisión de datos. Finalizada la transmisión de datos también se requiere realizar la liberación de la conexión
 - Emplea caminos físicos fijos origen-destino. Estos caminos son establecidos tras la conexión.
 - Habitualmente, el establecimiento de conexión introduce un retardo de tiempo elevado.
 - Empleando conmutación de circuitos es frecuente que los nodos intermedios no realicen comprobaciones de errores.
 - Como se ha indicado anteriormente este tipo de conmutación era empleada para la transmisión con flujo constante de datos (p. ej. transmisión de voz) hasta la llegada de las transmisiones digitales PCM. En la actualidad se emplean circuitos virtuales principalmente.
 - Opera a nivel físico.

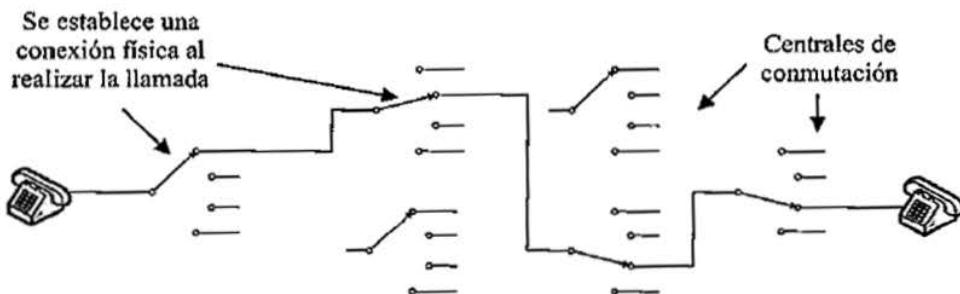


Figura 8-2. Conmutación de circuitos.

- Conmutación de paquetes. Es el método de conmutación empleado para la transmisión de datos y se caracteriza por:
 - La información se envía como fragmentos más pequeños denominados paquetes con un determinado formato compuesto por una cabecera con información de control y datos procedentes del nivel de transporte.
 - Los nodos intermedios pueden realizar comprobación de errores.

- Los nodos de la red reenvían los paquetes de datos atendiendo a un determinado criterio. Atendiendo a la forma en la que se realiza el reenvío existen dos variantes: commutación de circuitos y datagramas.

Como se ha indicado anteriormente, los métodos de commutación basados en circuitos virtuales y datagramas son dos variantes de commutación de paquetes que se caracterizan por:

- Circuitos virtuales. Presenta un comportamiento similar a la Red Telefónica Conmutada de circuitos conmutada y se caracteriza por las siguientes propiedades:
 - Está orientada a conexión de forma que en una primera fase, durante la conexión, se establece el camino que seguirán los paquetes de información desde el origen hasta el destino (Figura 8-3a). Una vez finalizada la conexión se realiza la transmisión de datos (Figura 8-3b). Esta transmisión finaliza con la desconexión del canal.
 - Se establece un camino fijo para la transmisión de datos, pero no exclusivo. Por lo tanto, el mismo camino (o un determinado enlace perteneciente al camino) puede ser utilizado por varias conexiones.
 - Generalmente los circuitos virtuales presentan una calidad alta en la transmisión, es decir, controlan los errores de datos.
 - Este tipo de circuitos es adecuado para la transmisión tanto de datos normales como señales analógicas digitalizadas (p. ej. voz o vídeo). Se trata, por tanto, de la tendencia actual ya que permite la transmisión de datos multimedia que presentan características muy variadas.
 - Los circuitos virtuales pueden ser permanentes o no. Un circuito virtual permanente es aquel que se mantiene fijo en sucesivas conexiones. Sin embargo, si el circuito virtual no es permanente, dado un mismo origen y un mismo destino, si se establecen varias conexiones los caminos obtenidos para interconectarlos puede variar. El primer caso es útil para conectar permanentemente redes remotas.
 - Opera a nivel de enlace o de red, según la tecnología.

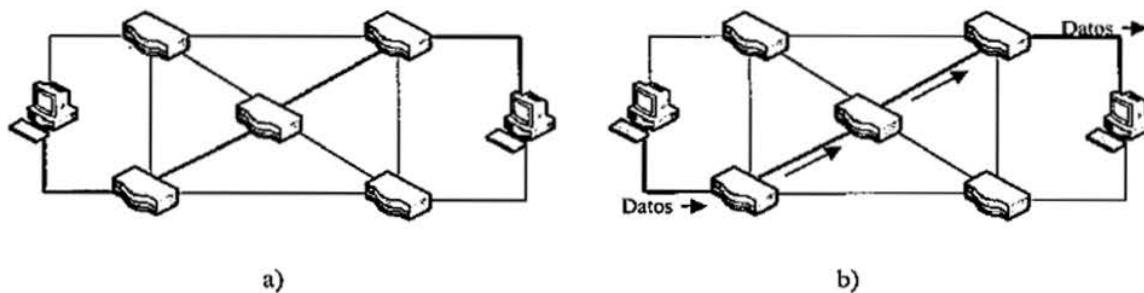


Figura 8-3. a) Establecimiento de conexión en circuitos virtuales. b) Envío de datos empleando circuitos virtuales.

- Datagramas. Se trata del método de commutación empleado en TCP/IP y se caracteriza por:
 - No hay conexión ni caminos fijos para conectar un origen con un destino. Cada paquete es independiente y en dicho paquete se especifica la dirección del equipo origen que envió el paquete y la dirección destino del mismo. De esta manera, cada nodo intermedio encamina el paquete atendiendo a un determinado criterio.

En la Figura 8-4 se muestra un ejemplo en el que el equipo A quiere enviar las tramas 1, 2 y 3 al B. En esta figura se ha indicado el camino seguido por cada trama hasta llegar al destino. Se observa como cada una de las tramas puede ser enviada por cada encaminador por un enlace distinto.

- Cada paquete sigue su camino en la red, según el estado de ésta y pueden llegar desordenados. Por lo tanto, el equipo destino deberá ser capaz de ordenar los datos.
- Se trata de un método tolerante a fallos ya que si falla un enlace sólo se pierde un paquete. El resto de paquetes podrán reenviarse por un camino alternativo.
- El método de conmutación basado en datagramas está totalmente orientado a datos informáticos y no a señales digitales. Fue el método original para transmitir datos a larga distancia.

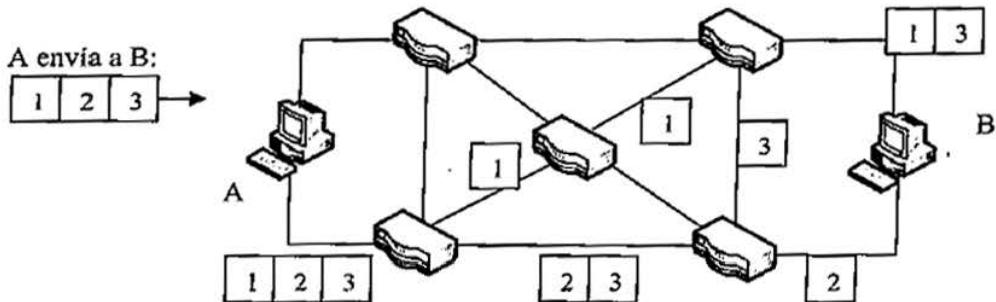


Figura 8-4. Transmisión de datos mediante datagramas.

Actualmente las infraestructuras WAN son principalmente circuitos virtuales ya que permiten enviar voz y vídeo (digital) y datos. Además, existen diversas tecnologías que permiten la transmisión de circuitos virtuales de datagramas sobre ellos, y es común enviar datagramas IP sobre redes de circuitos virtuales. IPv6 se ha diseñado pensando también en la utilización de circuitos virtuales.

8.4 Algoritmos de encaminamiento

Los algoritmos de encaminamiento son implementados en el nivel de red para determinar el camino que han de seguir los paquetes para alcanzar su destino. Por lo tanto, estos algoritmos son ejecutados en los nodos intermedios de una red para establecer el enlace de datos por el cual se retransmitirá el paquete atendiendo a la dirección origen y destino del mismo. La forma en la que se ejecutan estos algoritmos depende de si se emplean circuitos virtuales o datagramas:

- Circuitos virtuales. La decisión de encaminamiento se toma sólo una vez en cada establecimiento de conexión entre el origen y destino. De forma que en este establecimiento se determina el camino fijo que se empleará para comunicar los dos extremos de la comunicación.
- Datagramas. La decisión de encaminamiento se tomará por cada nodo intermedio. De manera que es responsabilidad del nivel de red de cada uno de estos nodos el reenviar cada paquete por el enlace más adecuado atendiendo al criterio que se considere oportuno.

Uno de los motivos por los que se tiende a circuitos virtuales a largas distancias es porque facilitan el encaminamiento una vez establecidas las conexiones. Así, los routers de circuitos virtuales procesan más rápido el reenvío de bloques de datos que los routers de datagramas.

Los principales tipos de encaminamiento empleando datagramas pueden clasificarse en encaminamiento fijo, inundación y encaminamiento adaptativo o dinámico.

8.4.1 Encaminamiento fijo o estático

Empleando este tipo de encaminamiento cada vez que se vaya a transmitir datos desde un determinado origen hasta un destino los paquetes seguirán el mismo camino. Habitualmente, en cada nodo se definen unas tablas de encaminamiento de forma que indican por qué enlace de datos se transmitirá el paquete dependiendo del origen y el destino del mismo. Presenta una baja tolerancia a fallos ya que si cae un determinado enlace de datos el nodo no podrá escoger otra alternativa para realizar la transmisión del paquete. Se trata de un método sencillo y rápido ya que en los nodos intermedios no se ejecutan complejos algoritmos de encaminamiento.

8.4.2 Encaminamiento por inundación

Haciendo uso de este tipo de encaminamiento cada paquete de datos entrante en un nodo intermedio se enviará por todas las rutas o enlaces del nodo, excepto por el de llegada. Este método garantiza que al menos uno de los paquetes llegará al destino por el camino óptimo. Sin embargo, se generan un gran número de paquetes duplicados lo cual puede provocar que la red se congestione. Para tratar de evitar esta congestión pueden emplearse mecanismos que limiten la cantidad de paquetes generados:

- Contador de salto. Se trata de un valor almacenado en la cabecera y que va siendo disminuido por cada nodo por el que pasa el paquete. En el momento que un determinado nodo detecte que el valor es 0 se encargará de eliminarlo.
- Número de secuencia. En este caso cada nodo fuente asigna un valor distinto a cada paquete denominado número de secuencia. Los nodos intermedios generan una lista de secuencias de paquetes enviados por cada nodo, de forma que si aparece un paquete con un número de secuencia que ya está en la lista, se descarta.

En la Figura 8-5 se muestra un ejemplo de inundación. En el esquema de la Figura 8-5a se representa el esquema de la red compuesto por 4 equipos interconectados a través de 4 encaminadores. Se desea establecer el camino para que A envíe información a B. Para ello, tras enviar la información de A a R1, este encaminador reenvía la información a sus encaminadores adyacentes, es decir, R2 y R4. Estos a su vez la reenviarán a sus adyacentes y así sucesivamente se alcanzará el destino. El esquema de cómo se realiza la difusión se ha representado en la Figura 8-5b. Se ha resaltado con un cuadro el momento en el que aparece en la ruta el encaminador R1, que ya se había visitado anteriormente. Tal y como se ha indicado anteriormente, el encaminador verá que su número de secuencia ya se encuentra en la ruta y, por lo tanto, deberá descartar el paquete.

Este tipo de encaminamiento no es empleado en la práctica en general, pero sí que es utilizado por determinados protocolos de nivel de red para enviar mensajes urgentes o de control de la red. Por ejemplo: actualizaciones de rutas en encaminamiento adaptativo, sincronización de tiempos, etc.

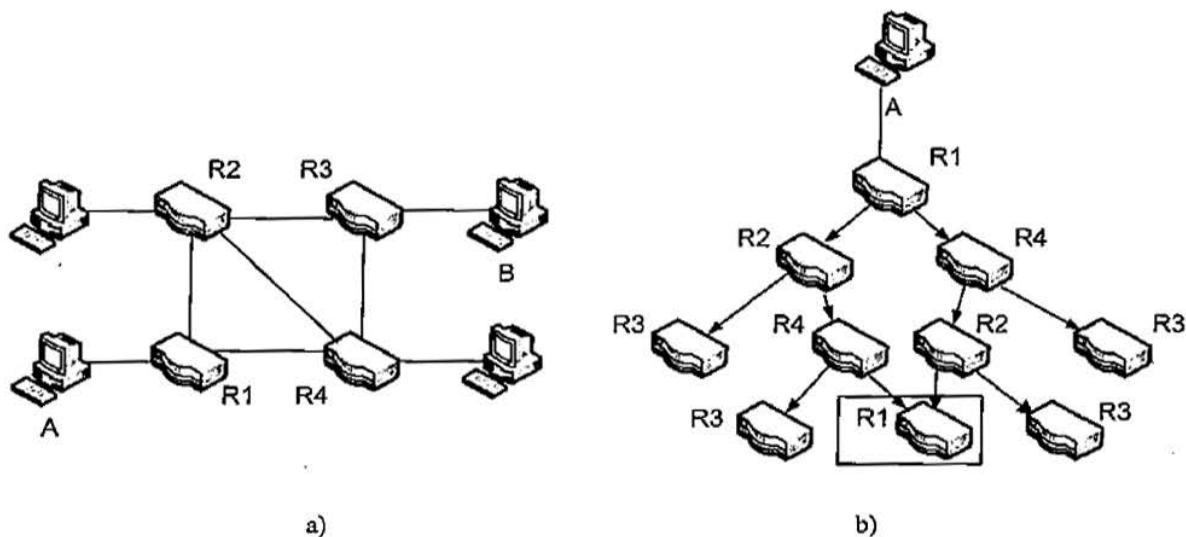


Figura 8-5. Ejemplo de inundación.

8.4.3 Encaminamiento adaptativo o dinámico

En este caso los caminos generados no son fijos sino que se adaptan a las condiciones de cambio de la subred. Cada nodo intermedio encaminará los paquetes por un enlace u otro atendiendo al estado de la red; por lo que es necesario que dichos nodos presenten una mayor capacidad de proceso. Este método de encaminamiento puede clasificarse en:

- Encaminamiento adaptativo centralizado. En la red existirá un nodo intermedio que se denominará centro de control del encaminamiento. Este nodo calcula las rutas óptimas entre los nodos que constituyen la red y las transmitirá a cada uno de ellos.
- Encaminamiento adaptativo aislado. Las decisiones de encaminamiento sólo se basan en información propia de cada nodo.
- Encaminamiento adaptativo distribuido. Empleando este tipo de encaminamiento la elección de la ruta del paquete al destino se realiza obteniendo la ruta de menos coste. La elección del coste de una ruta puede ser según uno o varios parámetros. Los parámetros típicos son: retardos (congestión), velocidad o ancho de banda y número de saltos (número de routers en el camino).

Capítulo 9. ENCAMINAMIENTO DEL NIVEL DE RED

En la práctica existen multitud de protocolos estándar de encaminamiento. En este capítulo se mostrará una clasificación de protocolos de encaminamiento en función de: su ámbito de uso y modo de calcular la ruta óptima. Posteriormente, se presentarán dos algoritmos básicos para el cálculo de la ruta óptima y una breve descripción de las métricas que se emplean para determinar cuál es ésta. Para concluir el capítulo se describirá el funcionamiento de dos protocolos estandarizados de encaminamiento en redes LAN: OSPF y RIP.

CAPÍTULO 9

- 9-1 Introducción
- 9-2 Tipos de protocolos de encaminamiento
- 9-3 Algoritmos de cálculo de ruta
 - Algoritmo de Dijkstra
 - Algoritmo de Bellman-Ford
- 9-4 Protocolo OSPF
 - Acciones OSPF para construir encaminamiento
 - OSPF encaminando a través de múltiples áreas
 - Tipos de paquetes OSPF
- 9-5 Protocolo RIP
 - Características de RIP
 - Acciones RIP para construir encaminamiento
- 9-6 Protocolos EGP: BGP

OBJETIVOS

- Clasificar los protocolos de encaminamiento
- Detallar las métricas más usadas que emplean los algoritmos de encaminamiento para determinar la ruta óptima
- Describir la casuística de dos algoritmos básicos de cálculo de ruta
- Exponer y comentar el funcionamiento de dos protocolos estandarizados para encaminamiento en redes LAN: OSPF y RIP
- Dar a conocer el protocolo BGP para encaminamiento para redes MAN/WAN

TERMINOLOGÍA – PALABRAS CLAVE

- | | | |
|----------------|-----------------------|--------------------|
| • IGP/EGP | • Vector de distancia | • LSA/LSD |
| • IGRP/EGRP | • Estado de enlace | • LSR/LSU/DD |
| • RIP | • Protocolo híbrido | • TOS |
| • OSPF | • MTU | • NSSA/ABR |
| • Dijkstra | • Coste/Carga/Retardo | • Nodo Backbone |
| • Bellman-Ford | • Número de satos | • Nodo Frontera |
| • Métrica | • Nodo encaminador | • Nodo interior |
| • Grafo | • Árbol de nodos | • Área/Stub |
| • Multicasting | • Tabla de rutas | • Request/Response |
| • BGP | • ISP | • Gateway |

9.1 Introducción

Los protocolos de encaminamiento se utilizan para actualizar las tablas de encaminamiento que se almacenan en los dispositivos de una red. Las tablas de encaminamiento son un conjunto de entradas donde se define cómo hacer llegar paquetes desde un nodo origen a un nodo destino. Ya se vio en el capítulo 8 que estas tablas de encaminamiento pueden ser estáticas o dinámicas.

Las tablas de encaminamiento estáticas son aquellas cuya información es introducida manualmente por un administrador de red que se encarga de modificarlas y actualizar sus entradas. Generalmente, este tipo de tablas se emplean en redes pequeñas que no suelen cambiar frecuentemente. Por el contrario, las tablas de encaminamiento dinámicas se actualizan periódicamente usando lo que se conocen como protocolos de encaminamiento. De este modo, cada vez que se produce un cambio en la red, por ejemplo el apagado o fallo de un nodo encaminador, el fallo de un enlace, etc., los protocolos actualizan las tablas de forma automática permitiendo el enrutamiento evitando esos enlaces o nodos encaminadores.

En la práctica existen multitud de protocolos estándar de encaminamiento que emplean algoritmos de encaminamiento basados en criterios de mínimo coste como los que se van a comentar en el apartado 9.2. En general, todos los protocolos de encaminamiento cumplen las mismas funciones, que consisten en aprender y determinar cuál es el mejor camino para ir de un nodo origen a un nodo destino.

Los algoritmos de encaminamiento usados en conmutación de paquetes y redes de tipo INTERNET basan sus decisiones en criterios de mínimo coste (apartado 9.3).

9.2 Tipos de protocolos de encaminamiento

En concreto, es común tipificar los protocolos de encaminamiento en dos grandes grupos en función de su ámbito de uso. Así, los protocolos se clasifican como:

- *Protocolos de 'gateway' interior* (IGP: Interior Gateway Protocol). Son protocolos de encaminamiento que se usan para intercambiar información de enrutamiento dentro de un mismo sistema de red autónomo. Un sistema autónomo, no es más que un conjunto de redes administradas bajo un mismo dominio (es decir, por un mismo administrador). Por lo tanto, este tipo de protocolos de encaminamiento se suelen emplear en redes LAN. Los protocolos estandarizados IGP que más se emplean son RIP y OSPF.
- *Protocolos de 'gateway' exterior* (EGP: Exterior Gateway Protocol). Son protocolos de encaminamiento que se usan para intercambiar información de enrutamiento entre distintos sistemas de red autónomos. Es decir, generalmente, este tipo de protocolos de encaminamiento se emplean en redes WAN. El protocolo EGP más conocido es BGP.

También, existen otros protocolos de encaminamiento IGP y EGP que no se van a tratar en este capítulo por tratarse de protocolos no estandarizados y propietarios de compañías como CISCO, siendo este el caso de IGRP y EIGRP. En la figura 9.1 se puede observar un ejemplo de ámbito de uso de protocolos IGP y EGP.

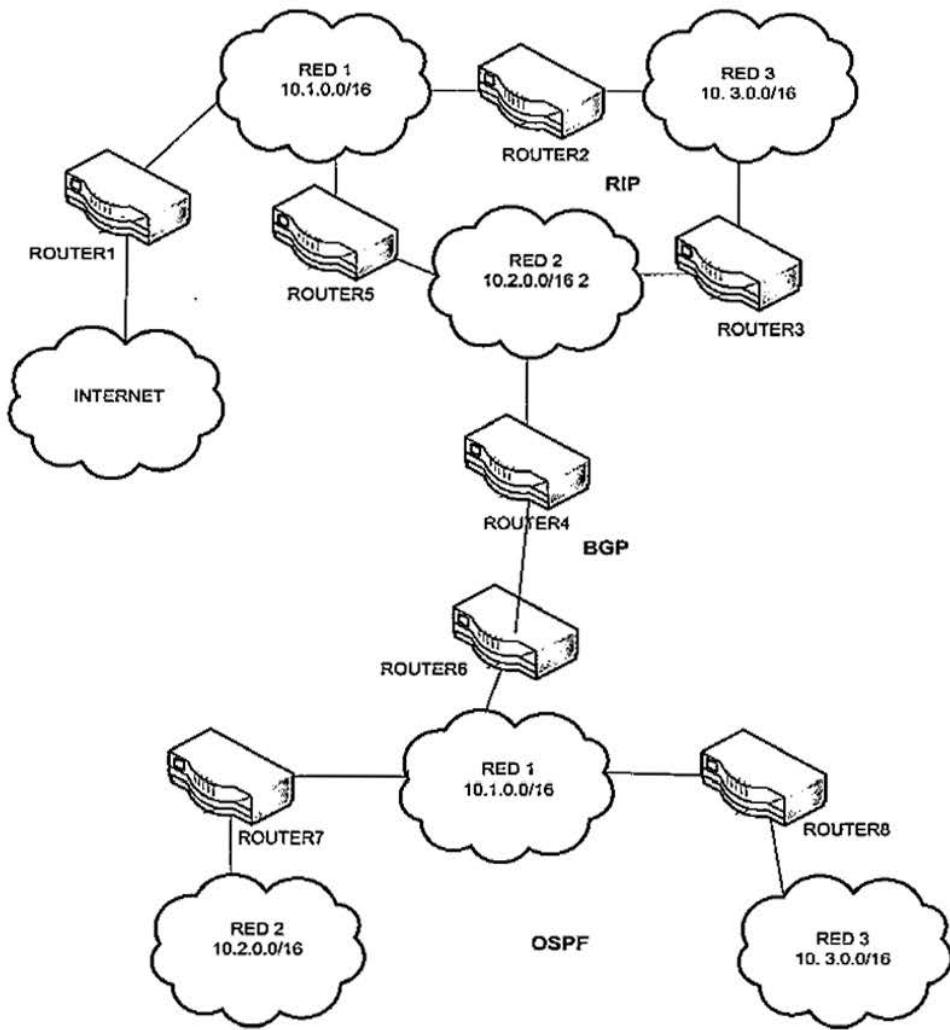


Figura 9-1. Esquema de Sistemas Autónomos empleando protocolos de distintas clases.

Además, existe otra forma de catalogar a los protocolos de encaminamiento en función de cómo se calcule la ruta óptima. Estos son tres:

- *Protocolos basados en vector de distancia.* Éste tipo de protocolos permiten calcular la distancia para llegar a un nodo destino situado en cualquier red, y la dirección necesaria para llegar hasta ese nodo destino. En estos protocolos de encaminamiento se denomina distancia a una medida de longitud, y vector significa una dirección. Existen gran diversidad de protocolos de encaminamiento estándar que están basados en vector de distancia. Algunos de los más representativos son los protocolos RIP e IGRP. En este tipo de protocolos las rutas calculadas por unos routers pueden ser empleadas por otros, ya que los distintos routers de una topología se intercambian periódicamente la información de rutas que disponen. Además, en este tipo de protocolos, todos y cada uno de ellos pueden emplear métodos distintos para calcular la ruta óptima. Uno de los algoritmos de cálculo de ruta más representativo basado en vector de distancia es el algoritmo de Bellman-Ford.

- *Protocolos basados en estado del enlace.* Este otro tipo de protocolo se basa en un conocimiento exacto de la topología de red sobre la que se quiere encaminar. Así, se crean tablas de encaminamiento basándose en información de la topología, a partir de paquetes denominados de estado de enlace que se intercambian todos los routers que forman la red para describir y determinar el estado de los distintos enlaces. La ventaja de estos protocolos es que cada router calcula de manera independiente la ruta óptima hacia un nodo destino. Esto permite que la ruta calculada por un router no depende de la calculada por otro. Un ejemplo de este tipo de protocolo es el protocolo OSPF. Y el algoritmo de cálculo de ruta más común es el algoritmo de Dijkstra.
- *Protocolos híbridos.* Protocolos de encaminamiento que combinan aspectos de protocolos de vector de distancia y de estado del enlace. Un ejemplo de este tipo de protocolos es el protocolo EIGRP de CISCO.

Cada uno de estos protocolos puede emplear algoritmos distintos para calcular la ruta óptima. Los algoritmos de encaminamiento más representativos son los algoritmos de Bellman-Ford y de Dijkstra. Ambos se comentarán en el apartado 9.3.

9.3 Algoritmos de cálculo de ruta

Los algoritmos que permiten calcular la ruta óptima para los distintos protocolos de encaminamiento son muy diversos, y su diversidad viene determinada por la métrica. La métrica no es más que un valor numérico, es decir un peso, que cada dispositivo de enrutamiento asigna a todas y cada una de las posibles rutas calculadas. Por lo general, este peso determina cómo de buena es cada una de las posibles rutas calculadas. La de mejor métrica será la ruta escogida como óptima para ir desde un nodo origen a uno nodo destino.

Las métricas se pueden calcular a partir de uno o varios parámetros. Así, algunos de los parámetros que estos algoritmos suelen considerar para el cálculo de la ruta óptima, son:

- *Número de saltos.* Número de 'routers' o encaminadores por los que pasará un paquete.
- *Ancho de banda:* Capacidad de datos del enlace. Es decir, velocidad de transmisión máxima que soporta el enlace.
- *Carga:* Número de paquetes que están gestionando un router o que está soportando un enlace en un determinado instante de tiempo.
- *Coste:* Valor arbitrario, que suele basarse en un valor ponderado o no, obtenido a partir del ancho de banda del enlace, el coste económico o cualquier otra medida y que es asignado por el administrador del sistema autónomo.
- *Tasa de errores:* Número de bits erróneos que se producen en cada enlace.
- *MTU:* Longitud máxima de datos de la trama de nivel de enlace que puede ser aceptada por todos los enlaces de la ruta.
- *Retardo:* Hace referencia al tiempo que un paquete de datos es almacenado en la cola de salida de un encaminador antes de ser reenviado. Es común considerar la suma de los retardos para un determinado camino. Normalmente este parámetro estará relacionado con los anchos de banda de los enlaces pero también puede depender de las estrategias de planificación y reenvío de paquetes empleados en los nodos encaminadores.



9.3.1 Algoritmo de Dijkstra

El algoritmo de Dijkstra es un algoritmo de cálculo de ruta basado en estado del enlace. Para ilustrar cómo el algoritmo de Dijkstra calcula la ruta óptima, se va a suponer que el estado del enlace viene determinado por una métrica. De ser así, cada enlace debería tener asociado un valor numérico. Generalmente, este valor es inversamente proporcional a la capacidad del enlace y proporcional a la carga de éste o una combinación ponderada de ambos. Por lo tanto, una ruta vendrá determinada por la suma de todas las métricas de todos los enlaces por los que se pase. Y la ruta óptima, será aquella que menor métrica calculada tenga.

Además, es importante indicar que el coste de un enlace puede ser distinto en cada uno de los dos sentidos de circulación de los paquetes. Esto es así, porque el coste depende precisamente de que se consideren parámetros como el retardo y la carga que no tienen por qué ser iguales en ambos sentidos de la comunicación, principalmente porque estos parámetros dependen de las características de los interfaces de salida de los nodos encaminadores.

Dada una red, el algoritmo de Dijkstra permite, por lo tanto, calcular la ruta óptima entre un nodo origen y todos y cada uno de los demás nodos que constituyen la red. El algoritmo consta de varias etapas o pasos que pueden resumirse en la figura 9.2.

Paso 1: <i>Inicio:</i> $T=\{s\}$ y $L(n)=w(s,n)$, con $n \neq s$.
Paso 2: <i>Nodo Siguiente:</i>
Encontrar el nodo $x \in N$, $L(x)=\min L(j)$ con $x, j \notin T$.
Añadir x a T : $T=\{s, \dots, x\}$
Paso 3: <i>Actualizar el camino de coste mínimo:</i>
Calcular $L(n)=\min \{L(n), L(x)+w(x,n)\}$ con ' n ' no $\in T$.
Paso 4: <i>Finalizar</i> cuando todos los nodos han sido añadidos a T .

donde:

- ' n ' \sqcup al conjunto de nodos de la red.
- ' s ': es el nodo origen
- ' T ': es la lista de nodos recorridos durante la ejecución del algoritmo
- ' $w(i,j)$ ': es el coste del enlace directo del nodo ' i ' al nodo ' j '.
- ' $L(n)$ ': es el coste en curso desde el nodo ' s ' al nodo ' n '.
- Caminos recorridos desde ' s ' al resto de los nodos de N

Figura 9-2. Etapas del algoritmo de Dijkstra.

Las variables que maneja el algoritmo (Figura 9.2) son las siguientes: como N se define el conjunto de nodos de la red o conjunto de routers, s representa el nodo inicial origen a partir del cual se quiere calcular la ruta, T es la lista de nodos hasta los cuales se ha calculado la ruta óptima, $w(i,j)$ define el coste del enlace entre el nodo i y el nodo j , y $L(n)$ determina el coste mínimo (es decir el coste de la ruta óptima) desde el nodo s al nodo n . Además, es importante indicar que si los nodos i y j no se encuentran directamente conectados, entonces los costes que se asocian durante la ejecución del algoritmo son infinitos, es decir $w(i,j)=\infty$, y si se encuentran directamente conectados el coste será $w(i,j)\geq 0$. El algoritmo consiste en ir añadiendo a una lista T aquellos nodos que formen parte del camino más corto para ir desde el nodo origen hasta el nodo destino, y que no se encuentren ya incluidos en la lista (es decir que no se haya pasado por ellos). Cada vez que se añade un nodo nuevo a T se está incorporando un nuevo router a la ruta. Además, $L(n)$ irá actualizándose con nuevo coste tras haber introducido un nuevo nodo.

Ejemplo 9.1

Se pretende hacer uso de algoritmo de Dijkstra para generar el camino de mínimo coste entre el nodo 1 y el resto de nodos de una red LAN, representada en la figura 9.3; de modo que la información obtenida pueda usarse como criterio para construir las tablas de encaminamiento estático para cada nodo de la red. Para ello, completa la siguiente tabla indicando la ruta y el coste mínimo de la ruta para cada iteración i.

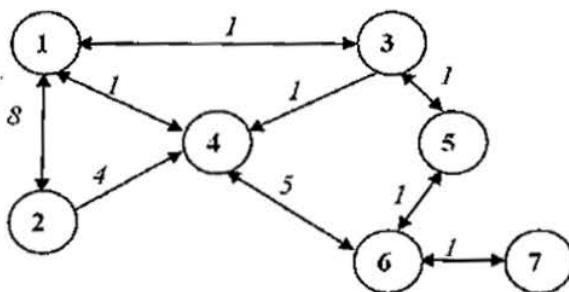


Figura 9-3. Topología de red para el cálculo de rutas por Dijkstra.

i	T	L(2)	Ruta	L(3)	Ruta	L(4)	Ruta	L(5)	Ruta
1	1	8	1-2	1	1-3	1	1-4	∞	-
2	1,3	8	1-2	1	1-3	1	1-4	2	1-3-5
3	1,3,4	5	1-4-2	1	1-3	1	1-4	2	1-3-5
4	1,3,4,5	5	1-4-2	1	1-3	1	1-4	2	1-3-5
5	1,3,4,5,6	5	1-4-2	1	1-3	1	1-4	2	1-3-5
6	1,3,4,5,6,7	5	1-4-2	1	1-3	1	1-4	2	1-3-5
7	1,2,3,4,5,6,7	5	1-4-2	1	1-3	1	1-4	2	1-3-5

L(6)	Ruta	L(7)	Ruta
∞	-	∞	-
∞	-	∞	-
6	1-4-6	∞	-
3	1-3-5-6	∞	-
3	1-3-5-6	4	1-3-5-6-7
3	1-3-5-6	4	1-3-5-6-7
3	1-3-5-6	4	1-3-5-6-7

9.3.2 Algoritmo de Bellman-Ford

El algoritmo de Bellman-Ford es un algoritmo de cálculo de ruta basado en el camino más corto. En el caso de Bellman-Ford, el cálculo de la ruta óptima se reduce a calcular el camino más corto entre un nodo origen y un nodo destino. Por lo tanto, en este caso no se suele hablar de coste asociado a un enlace, ya que normalmente la métrica que se asocia a un enlace es uno. Al final, el camino más corto tendrá como coste la suma de todas las métricas de cada uno de los enlaces que, por lo general, no será otra que el número de enlaces por los que se ha pasado.

El algoritmo comienza calculando el camino más corto entre un nodo y todos los demás con la única exigencia de que los caminos calculados contengan al menos un enlace. Posteriormente, se repite la misma iteración pero suponiendo que la condición que se exige sea de al menos dos enlaces, y a continuación se volverá a repetir suponiendo tres enlaces, y así sucesivamente. Los pasos básicos para calcular el camino más corto se van a ilustrar sobre el ejemplo de la figura 9.4:

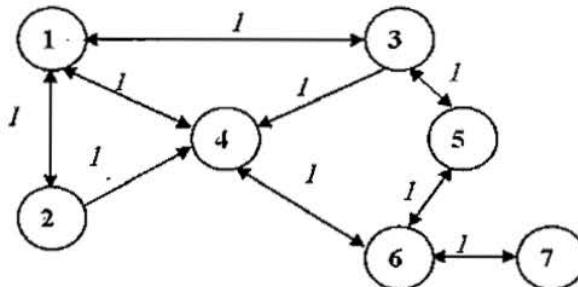


Figura 9-4. Topología de red para el cálculo de rutas por Bellman-Ford.

Paso 1: Inicialmente, cada nodo de la red se inicializa con su una tabla de información que contiene como ir de ese nodo origen a todos los demás nodos de la red. Así, para el ejemplo de la figura 9.5 se muestra la información de inicialización para cada uno de los nodos. En cada nodo, se inicializa el coste del camino a métrica 0 si el nodo destino es el mismo que el origen y se inicializa a métrica, ∞ , si no hay un camino directo entre el nodo origen y el nodo destino. Además, el símbolo '-' significa que el nodo destino está conectado directamente o es el origen mismo. El resto de caminos entre el nodo origen y el nodo destino tendrán como coste 1. En la información de las tablas se almacena cuál es el nodo destino (Des) y cuál es el nodo siguiente para alcanzar ese destino (Sig).

Des	Sig	M.												
1	-	0	1	-	1	1	-	1	1	-	1	1	-	1
2	-	1	2	-	0	2	-	∞	2	-	1	2	-	∞
3	-	1	3	-	∞	3	-	0	3	-	1	3	-	1
4	-	1	4	-	1	4	-	1	4	-	0	4	-	∞
5	-	∞	5	-	∞	5	-	1	5	-	0	5	-	0
6	-	∞	6	-	∞	6	-	∞	6	-	1	6	-	1
7	-	∞												

Nodo '1'			Nodo '2'			Nodo '3'			Nodo '4'			Nodo '5'		
Des	Sig	M.												
1	-	∞	2	-	∞	3	-	∞	4	-	∞	5	-	∞
2	-	∞	3	-	∞	4	-	∞	5	-	∞	6	-	∞
3	-	∞	4	-	∞	5	-	∞	6	-	∞	7	-	∞
4	-	∞	5	-	∞	6	-	∞	7	-	∞	1	-	∞
5	-	∞	6	-	∞	7	-	∞	1	-	∞	2	-	∞
6	-	∞	7	-	∞	1	-	∞	2	-	∞	3	-	∞
7	-	∞	1	-	∞	2	-	∞	3	-	∞	4	-	∞

Des	Sig	M.	Des	Sig	M.
1	-	∞	1	-	∞
2	-	∞	2	-	∞
3	-	∞	3	-	∞
4	-	1	4	-	∞
5	-	1	5	-	∞
6	-	0	6	-	1
7	-	1	7	-	0

Nodo '6'			Nodo '7'		
Des	Sig	M.	Des	Sig	M.
1	-	∞	1	-	∞
2	-	∞	2	-	∞
3	-	∞	3	-	∞
4	-	1	4	-	∞
5	-	1	5	-	∞
6	-	0	6	-	1
7	-	1	7	-	0

Figura 9-5. Inicialización de las tablas para encaminamiento basado en vector de distancia y Actualización tabla de encaminamiento Nodo '1'.

Paso 2: Se actualizan las tablas de información de rutas calculando cual es la ruta más corta para ir de un nodo origen a un nodo destino considerando dos enlaces (Figura 9-6). Veamos, por ejemplo, cuál es la ruta más corta para ir desde el nodo '1' de la Figura 9.6, a un nodo destino predeterminado empleando dos enlaces. En primer lugar, el nodo '1' consultará en sus tablas cuales son los enlaces adyacentes, que en este caso serían '2', '3' y '4'. A continuación se consultan en las tablas de '2', '3' y '4', cuáles son sus respectivos nodos adyacentes (de 2: '1' y '4', de 3: '1', '4' y '5' y de 4: '1', '2', '3' y '6'). Por lo tanto, se puede determinar que el nodo '1' puede alcanzar a los nodos '5' y '6' a través de dos enlaces. Así observando las tablas de información se observa que la ruta del nodo '1' hasta el nodo '3' empleando dos enlaces tiene métrica 2. Lo mismo ocurre para ir del nodo '1' a los nodos '2' y '4'. Por lo tanto, como en estos casos la métrica obtenida al recorrer dos enlaces es mayor que la que se tenía almacenada en la tabla de la figura 9.5, la ruta y la métrica no se actualiza y se mantienen las del paso 1. Por otro lado, para ir desde el nodo '1' al nodo '5' y '6' con dos enlaces, también, se emplea una métrica de 2 y, sin embargo en este caso si se actualizan las rutas almacenadas en la tabla de la figura 9.5, ya que la métrica obtenida es menor que la almacenada inicialmente. En un principio, para ir del nodo '1' a los nodos '5' y '6' se tenía una métrica ∞ .

Paso 3: Este proceso se repite iterativamente tantas hasta que se alcanza el nodo más alejado.

The figure shows the iterative update of the routing table for Node 1. It starts with an initial table for Node 1 (labeled 'Nodo 1') and four other nodes (Nodo 2, Nodo 3, Nodo 4) which are not yet modified. The 'Nodo 1' table has columns 'Des' (Destination), 'Sig' (Signature), and 'M.' (Metric). The 'Nodo 2', 'Nodo 3', and 'Nodo 4' tables also have these columns. An arrow points from the final 'Tabla modificada de Nodo 1' back to the original 'Nodo 1' table, indicating that the changes made to the other nodes' tables are reflected back to the original table.

Des	Sig	M.
1	-	0
2	-	1
3	-	1
4	-	1
5	∞	
6	∞	
7	∞	

Des	Sig	M.
1	-	1
2	-	0
3	-	∞
4	-	1
5	∞	
6	∞	
7	∞	

Des	Sig	M.
1	-	1
2	-	∞
3	-	0
4	-	1
5	-	1
6	-	∞
7	-	∞

Des	Sig	M.
1	-	1
2	-	1
3	-	1
4	-	1
5	-	∞
6	-	1
7	-	∞

Des	Sig	M.
1	-	0
2	-	1
3	-	1
4	-	1
5	3	2
6	4	2
7	6	∞

Figura 9-6. Actualización de la tabla de rutas del nodo '1'.

En la figura 9.6, se representa el proceso de actualización. En un instante determinado la tabla del Nodo '1', se actualiza con la información que le proporcionan los nodos '3' y '4'. Un ejemplo práctico de este algoritmo es el protocolo estándar RIP que emplea este algoritmo para calcular las rutas y actualizarlas.

9.4 Protocolo OSPF

El protocolo OSPF (Open Shortest Path First) o protocolo abierto del primer camino más corto, data de finales de los años 80. Este protocolo fue diseñado como protocolo de encaminamiento para solventar algunos de los problemas que el protocolo RIP no contemplaba. Actualmente la versión del protocolo OSPF que se emplea es la versión 2 cuando se usa IPv4, u OSPF versión 3 para IPv6. En cualquier caso, OSPF es el protocolo de encaminamiento más empleado en grandes redes LAN. Además, OSPF permite usar el algoritmo MD5 para firmar los datos y comprobar la veracidad de un mensaje antes de actualizar y crear nuevas rutas. También, es importante destacar, que a diferencia de RIP, OSPF no requiere protocolos de nivel de transporte para encapsular los datos de encaminamiento, y que éste opera directamente sobre IP (RIP encapsula sobre paquetes de transporte UDP). La

normativa de IETF que define la implementación del protocolo OSPF está publicada en las normas RFC 2328 y RFC 2740, para las versiones 2 y 3 respectivamente. Estas versiones datan de finales de los 90.

OSPF es un protocolo basado en estado del enlace. La principal diferencia entre OSPF y RIP viene determinada por cómo ambos protocolos comunican actualizaciones de rutas. Mientras que RIP permite a un nodo encaminador enviar sus propias rutas únicamente a sus nodos vecinos adyacentes, el protocolo OSPF permite a un nodo enviar publicaciones del estado del enlace, a cualquier nodo perteneciente a la misma jerarquía (ver más adelante el concepto de área). Así, el protocolo OSPF permite que cada nodo encaminador disponga de una base de datos que refleja la topología de la red, expresada en formato de grafo dirigido. En ambos casos se emplean paquetes multidifusión de IP para propagar las rutas entre nodos encaminadores. En OSPF la dirección IP de multidifusión o ‘multicast’ es 224.0.0.5.

OSPF, divide un sistema autónomo (red con un mismo administrador) en unidades jerárquicas de menor tamaño para realizar un encaminamiento más eficiente (Figura 9.7). Estas unidades se conocen con el nombre de áreas. Un *área* no es más que un subconjunto de nodos encaminadores y estaciones conectados dentro de un mismo sistema autónomo. Por lo tanto, un área es una red más pequeña que constituye un subconjunto de aquella a partir de la cual partió la división. Los cambios que se produzcan en la topología de un área no afectan al rendimiento de otras áreas, puesto que OSPF gestiona el encaminamiento de cada área de forma independiente.

Para asignar la métrica o coste de los enlaces a cada nodo encaminador, se pueden emplear diferentes criterios: mínimo retardo, máxima productividad, etc. Además, es frecuente construir varias tablas de encaminamiento para un mismo nodo encaminador empleando cada uno de esos diferentes criterios. Por lo general, es frecuente usar como métrica de un enlace un valor numérico que se calcula en base al ancho de banda del interfaz (Figura 9.8).

9.4.1 Acciones OSPF para construir encaminamiento

En el encaminamiento basado en estado del enlace, se necesita de un conjunto de acciones que permitan asegurar que cada nodo encaminador tenga una tabla de encaminamiento que disponga de información del nodo de menor coste a partir del cual se alcanza al resto de nodos. Estas acciones se pueden resumir en:

- *Crear estados de los enlaces* a los que conecta cada nodo (Figura 9.8). A partir de estos estados cada nodo construye paquetes LSA (Link-State Advertisement) o paquetes de estado de sus enlaces. Estos paquetes se emplean para que los nodos se intercambien la información de los enlaces a los que están conectados. El paquete de estado de enlace, LSA, suele transportar gran cantidad de información. Por lo general:
 - La identidad de cada nodo y la lista de enlaces a los que está conectado para poder definir la topología de la red del dominio.
 - El número de secuencia y la edad para distinguir los nuevos paquetes de los viejos.

Además, los paquetes LSA sólo se crean cuando se da alguna de estas situaciones:

- Cambios en la topología de red del dominio. Por lo tanto, se requiere que cada nodo encaminador se adapte a la nueva topología. En este caso los paquetes LSA los generan los nodos que conectan enlaces nuevos o enlaces que han sido modificados. Esos paquetes se enviarán, después, a nodos adyacentes.
- Siempre de forma periódica. Cada intervalo de tiempo comprendido entre 60 minutos y 2 horas, dependiendo de la configuración, se crean paquetes LSA que

permitan actualizar las entradas de los encaminadores. De este modo se elimina información vieja que pueda quedar almacenada.

- *Enviar paquetes LSA a cada nodo adyacente mediante el método de inundación* por cada interfaz de salida, generalmente cada 10 segundos. Cada nodo que recibirá paquetes LSA, los compara con la copia que pueda tener almacenada, descartando el que sea más viejo de los dos (número de secuencia más pequeño). El más nuevo lo almacena y envía copia de nuevo por cada interfaz por el método de inundación. Una vez se hayan enviado los paquetes LSA, cada nodo tendrá una copia de toda la topología completa, esta copia constituye una base de datos de los enlaces de la misma (LSD, Link-State Database).
- *Formar un árbol del camino más corto para cada nodo* (Figura 9.9). Puesto que la topología completa almacenada en cada nodo a partir de los LSA (Figura 9.8) no es suficiente para determinar cuál es el camino más corto, el algoritmo de Dijkstra permite determinar el camino mínimo entre un nodo encaminador (nodo raíz en el árbol) y el resto de nodos encaminadores (nodos hojas en el árbol). El árbol representa la ejecución del algoritmo de Dijkstra.
- *Calcular la tabla de encaminamiento basada en el árbol de camino más corto*. Los nodos encaminadores que tienen activado el protocolo OSPF almacenan información sobre el estado de los enlaces y ejecutar el algoritmo de Dijkstra para calcular la ruta más corta de un nodo a cualquier otro. Así a partir del árbol de Dijkstra para cada nodo (Figura 9.9), se construye una tabla de encaminamiento empleando la ruta de camino más corto para ir de un nodo a otro.

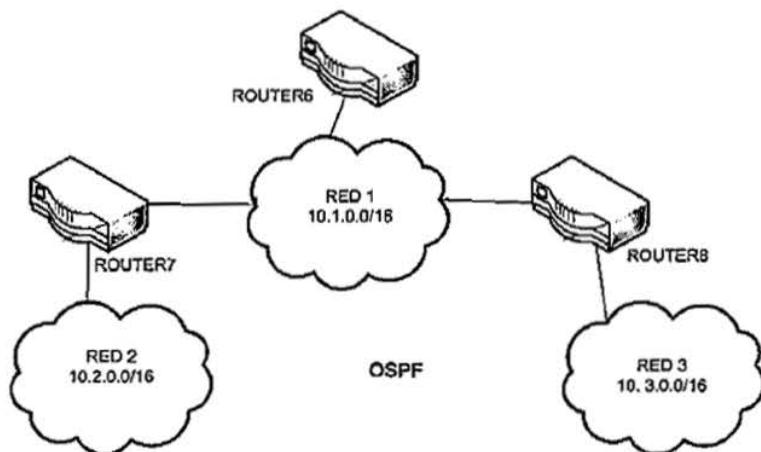


Figura 9-7. Ejemplo de una red como sistema autónomo funcionando con protocolo OSPF.

Para determinar cómo se fija el coste de los enlaces o métrica, OSPF dispone de cinco servicios que determinan múltiples formas de calcular la métrica. Estos se clasifican como:

- *TOS 0 (Type of Service)*: Métrica asociada por el administrador de la red.
- *TOS 2*: Métrica empleada si se asigna coste económico al uso de la red.
- *TOS 4*: Métrica basada en el historial reciente de fallos en la red o tasa de paquetes erróneos.
- *TOS 8*: Métrica basada en la capacidad del enlace.
- *TOS 16*: Métrica basada en el retardo de propagación y retardo de los buffers de interfaces.

En la práctica, OSPF construye y mantiene cinco grafos distintos para la red en la que está ejecutándose, así como las correspondientes tablas de encaminamiento para cada grafo. OSPF se ayuda del campo TOS que incorporan los datagramas IP que encapsulan los paquetes. Según el valor de este campo, cada nodo encaminador consulta la tabla de encaminamiento apropiada para encaminar el datagrama. En caso de que el campo TOS no se incluya entonces OSPF usa por defecto la métrica *TOS 0*.

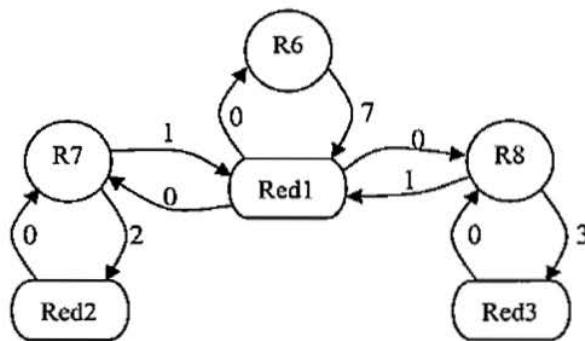


Figura 9-8. Grafo dirigido construido por OSPF para la red de la figura anterior.

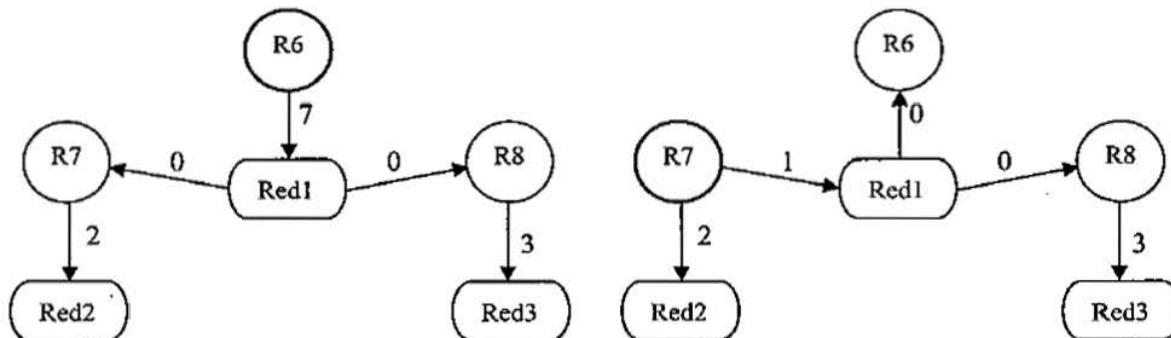


Figura 9-9. Árbol de expansión obtenido al aplicar el algoritmo de Dijkstra con $s=R6$ y $R7$ respectivamente.

9.4.2 OSPF encaminando a través de múltiples áreas

Cuando un sistema autónomo se jerarquiza y se divide en múltiples áreas para facilitar el encaminamiento por OSPF, cada área puede ser de uno de estos tres tipos o categorías:

- **Backbone:** Constituye el núcleo central de una red encaminada con OSPF. Este área se caracteriza porque mantiene una conexión física o lógica con todas y cada una de las demás áreas en las que se ha jerarquizado la red.
- **Stub:** Hace referencia a todas aquellas áreas que no pueden recibir rutas procedentes de redes externas. Se consideran redes **externas** todas aquellas que son gestionadas por protocolos de encaminamiento distintos a OSPF y están fuera de la jerarquía de red.
- **Not-so-stubby (NSSA):** Son áreas stub que pueden importar rutas externas y enviarlas a un área de backbone, pero que sin embargo no pueden exportar rutas externas procedentes del backbone hacia otras áreas.

Además, se pueden diferenciar los siguientes tipos de nodos encaminadores:

- Nodos encaminadores Interiores: Son aquellos en los que todos sus interfaces pertenecen a una misma área de la jerarquía, y además la base de datos de estados de enlace (LSD) es idéntica.
- Nodos encaminadores de Backbone: Son aquellos en los que al menos un interface pertenece al área de Backbone, también conocida como área cero. Estos nodos tienen como objetivo calcular la mejor ruta para todos los nodos encaminadores a partir de la información que les facilitan los nodos encaminadores de tipo frontera. Una vez calculada la mejor ruta, el nodo backbone la propaga de regreso a los nodos frontera los cuales la divultan a sus respectivas áreas.
- Nodos encaminadores de Frontera (ABR: Area Border Router): Son los encaminadores que tienen interfaces conectadas a diferentes áreas. Estos encaminadores necesitan la base de datos de cada área y deben calcular, para cada área por separado, la obtención de rutas óptimas.

9.4.3 Tipos de paquetes OSPF

OSPF es un protocolo de encaminamiento cuya información se encapsula sobre datagramas IP. Existen cinco tipos de mensajes básicos en OSPF, algunos de los cuales ya se han comentado en el apartado 9.4.1:

- *Hello*: Permiten que un nodo encaminador detecte cambios en el estado de nodos encaminadores adyacentes o en los enlaces que los unen con estos. Cuando en los nodos encaminadores se activa OSPF se envían mensajes de este tipo.
- *DD (Database description)* o *LSD*: Descripción de la base de datos. Estos mensajes se envían cuando se activa un enlace entre dos nodos encaminadores. Se emplea para sincronizar la información topológica que poseen ambos.
- *LSR (Link State Request)*: Solicitud de información. Mensajes que un nodo encaminador emplea para solicitar información topológica a otro encaminador adyacente.
- *LSU (Link State Update)*: Actualización de información. Es el mensaje respuesta a un LSR y pueden contener uno o varios paquetes LSA. La confirmación de actualización se realiza con un mensaje *Link State Acknowledgment*.
- *LSA (Link State Advertisement)*: Existen cinco tipos distintos de mensajes LSA: mensajes *router links*, *network links*, dos mensajes del tipo *summary link*, uno para notificar el coste de alcanzar una subred situada fuera de un área y otro para notificar el coste de alcanzar un encaminador de frontera. Finalmente, se dispone de un mensaje *AS external link* el cual es generado por encaminadores de frontera y transmitidos por inundación al resto.

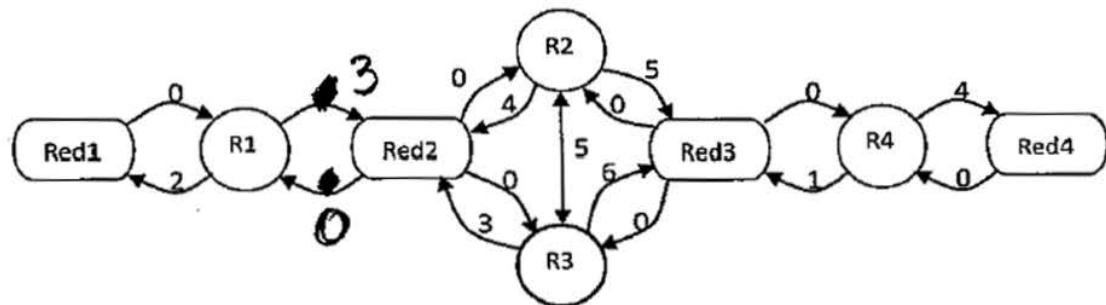
Ejemplo 9.2

Actualización de tablas de encaminamiento con protocolo OSPF: Dada la siguiente información LSD con costes asociados a los enlaces mostrada en la figura 9.10, se pide construir el árbol que determina los caminos óptimos para el nodo encaminador R3.

Destino	Origen							
	Routers				Redes			
Router/Red	R1	R2	R3	R4	Red1	Red2	Red3	Red4
R1					0	0		
R2			5			0	0	
R3		5				0	0	
R4							0	0
Red1	2							
Red2	3	4	3					
Red3		5	6	1				
Red4				4				

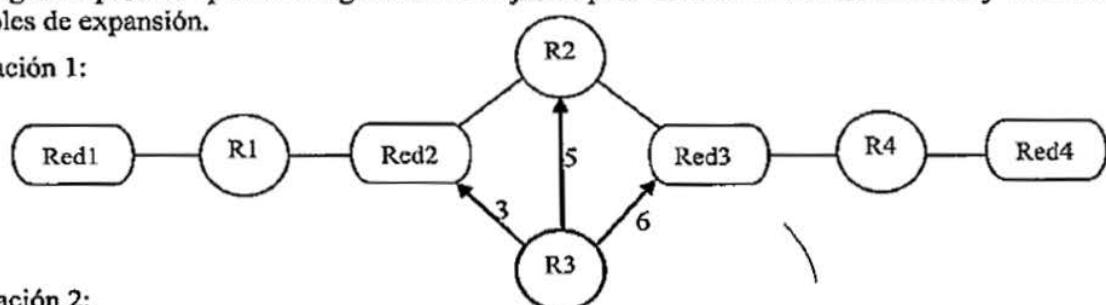
Figura 9-10. Información de enlaces de redes y nodos encaminadores (LSD: Link State Database).

El primer paso es construir el grafo dirigido.

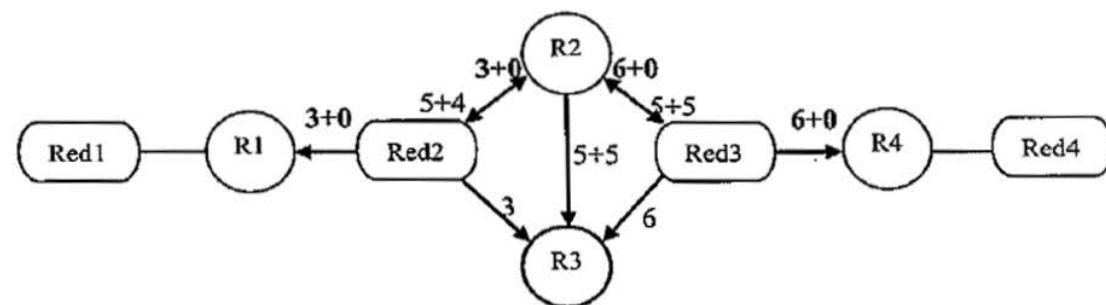


El segundo paso es aplicar el algoritmo de Dijkstra para obtener el camino mínimo y así calcular los árboles de expansión.

Iteración 1:



Iteración 2:



Iteración 2: Para construir el segundo nivel, se buscan todos los dispositivos del primer nivel que pueden ser alcanzados directamente. Entonces, se añade al coste de alcanzar cada dispositivo en el primer nivel, el coste de alcanzarlo en el segundo nivel. Así, se tiene:

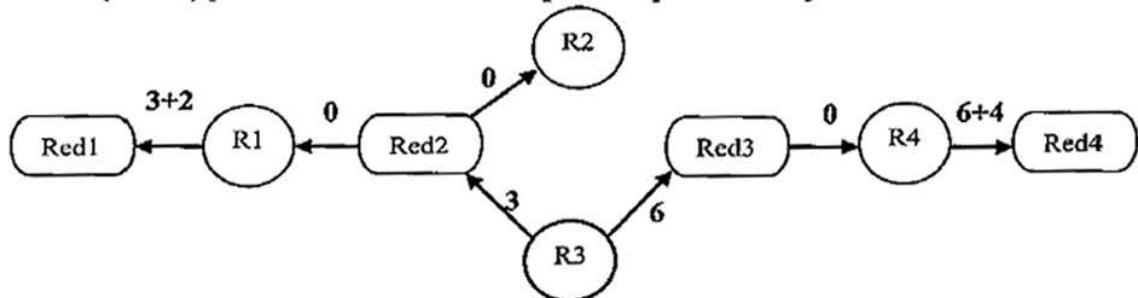
R2: si el coste de alcanzar R2 desde el nodo R3 es 5 en el primer nivel, entonces se considera que R2 parte de un coste inicial de 5 al que hay que añadir el coste de que R2 alcance el router R3 ($5+5=10$), el coste de que alcance la Red2 ($5+4=9$) y el coste de que alcance la Red3 ($5+5=10$).

Red2: si el coste de alcanzar la Red2 desde el nodo R3 es 3 en el primer nivel, entonces se considera que desde la Red2 se parte de un coste inicial de 3 al que hay que añadir el coste de que desde la Red2 se alcance el router R1 ($3+0=3$), el coste de que se alcance el router R2 ($3+0=3$) y el coste de que se alcance el router R3 ($3+0=3$).

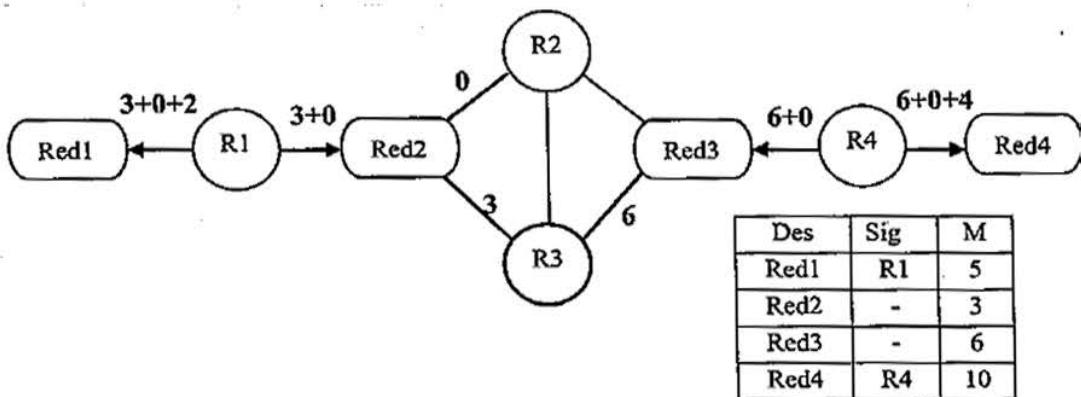
Red3: si el coste de alcanzar la Red3 desde el nodo R3 es 6 en el primer nivel, entonces se considera que desde la Red3 se parte de un coste inicial de 6 al que hay que añadir el coste de que desde la Red3 se alcance el router R2 ($6+0=6$), el coste de que se alcance el router R3 ($6+0=6$) y el coste de que se alcance el router R4 ($6+0=6$).

En tal caso, nos encontramos con algunas rutas o caminos que no tienen sentido puesto que forman bucles. Por tal motivo, no hay que considerar las rutas de alcanzar R3 desde cualquier nodo o red puesto que en la Iteración 1 se partió como nodo inicial de R3. Además, tampoco hay que considerar todas aquellas rutas o caminos que sean más largas que las calculadas en la iteración anterior (iteración 1). Así, el camino a través del R2 para alcanzar la Red3 tiene un coste de 10 y es superior al camino de ir de R3 a Red3 directamente, como se calculó en la iteración anterior (iteración 1). Por lo tanto, tras la iteración 2 las nuevas rutas de coste mínimo calculadas tienen un coste de (3+0) para ir a R1 desde R3 pasando por la Red2, un coste de (6+0) para ir a R4 desde R3 pasando por Red3 y un coste de (3+0) para ir a R2 puesto que tiene menor coste ir a través de la Red2 que ir directamente. Esto implica que en la iteración 2 además de añadir dos nuevas rutas se ha actualizado la ruta para ir al nodo R2 minimizando el coste.

Iteración 3: Repitiendo el procedimiento anterior pero considerando los caminos y costes de la iteración 2 y el coste de añadir un nuevo enlace, se obtiene que las rutas de coste mínimo calculadas tienen un coste de (3+0+2) para ir a la Red1 desde R3 pasando por la Red2 y R1, y un coste de (6+0+4) para ir a la Red4 desde R3 pasando por la Red3 y R4.



Una vez calculado el camino mínimo con Dijkstra se puede obtener el árbol de expansión y una tabla de encaminamiento para R3 a partir del árbol de extensión como los mostrados en la siguiente figura:



9.5 Protocolo RIP

El protocolo RIP, igual que el protocolo OSPF, permite que los nodos encaminadores mantengan actualizadas sus tablas de encaminamiento gracias a mensajes de difusión usando el protocolo UDP y el puerto 520. Como se comentó anteriormente, el protocolo RIP emplea un algoritmo de vector de distancia cuya métrica se basa en el número de saltos desde un nodo origen a un nodo destino. El número de saltos se entiende como el número de nodos encaminadores que tiene que atravesar un paquete de datos hasta alcanzar un nodo destino, sin contar el origen e incluyendo el destino. También, se puede entender como el número de redes por las que debe pasar un paquete. Así, RIP asocia esta métrica a cada una de las entradas de las tablas de encaminamiento de cada nodo encaminador. A partir de esa métrica, el protocolo decide qué entrada de la tabla de encaminamiento se debe aplicar para alcanzar un nodo destino.

Con RIP el máximo número de saltos se sitúa en 15, y por ello es utilizado en redes con dimensiones reducidas en cuanto a número de nodos encaminadores. De hecho, una métrica de 16 indica el valor infinito. La descripción del protocolo RIP (Routing Information Protocol) está publicado en las normas RFC 1058 y RFC 2453, para las versiones 1 y 2 respectivamente.

9.5.1 Características de RIP

En el encaminamiento mediante el protocolo RIP en su versión 2 caben destacar las siguientes características:

- Capacidad de identificar rutas externas.** Permite propagar información sobre rutas establecidas con otros protocolos de encaminamiento (como EGP o BGP) sin alterarlas. Su objetivo principal es separar rutas “internas” de la red donde funciona RIP de rutas “externas” a esa red.
- Máscaras de subred.** Permite trabajar con rutas de subredes. El gran problema que tenía RIP 1 era no disponer de esta característica, aunque su necesidad es evidente.
- Dirección del siguiente salto.** En cada entrada de ruta de un mensaje RIP se puede especificar, además del número de saltos para llegar a la dirección IP destino (como se hace en RIP 1), la dirección IP del siguiente nodo encaminador al que pueden ser enviados los paquetes, en vez de utilizar el nodo encaminador que genera el mensaje. Permite la optimización del encaminamiento en la red y se evitan futuras redirecciones.

- *Autenticación.* Aporta mecanismos para que un nodo encaminador solo acepte mensajes RIP determinados, con el objetivo de aumentar la seguridad de acceso gracias a una firma MD5 incluida en los mensajes, de forma similar a OSPF. Se evita así que cualquier equipo de una red pueda enviar paquetes RIP a un nodo encaminador para confundirlo.
- *Multicasting.* Los paquetes RIP 2 se envían a una dirección IP específica; la dirección IP de multidifusión 224.0.0.9 (de la clase especial de *multicast* de IP: 224.0.0.0/28). Solo los nodos encaminadores con RIP 2 activo hacen caso de lo recibido por esa dirección, esto es, funciona como si de un *broadcast* selectivo se tratase.

9.5.2 Acciones RIP para construir encaminamiento

En el encaminamiento RIP, existe una serie de acciones que permitan asegurar que cada nodo encaminador disponga de una tabla de encaminamiento lo más actualizada posible sobre la métrica de cada camino para ir de un nodo encaminador a otro. Estas acciones se realizan a través del intercambio de paquetes UDP que encapsulan diferentes mensajes RIP. Las acciones básicas son dos:

- *Enviar petición o solicitud RIP.* Cuando se inicia el proceso de actualización de rutas tras la activación del protocolo RIP en una red, todos y cada uno de los nodos encaminadores comienzan a enviar mensajes '*Request*' que difunden por todos sus interfaces activos. Con estos mensajes se solicita la información de encaminamiento que conocen los nodos encaminadores adyacentes.
- *Enviar respuesta RIP.* Los nodos encaminadores que reciben un mensaje de solicitud de rutas envían mensajes RIP '*Response*' con información de la información de rutas que conocen. Además, en cada nodo encaminador funcionando con RIP, cada cierto tiempo (típicamente 30 segundos), una parte o la totalidad de la tabla de encaminamiento es enviada a los nodos encaminadores adyacentes a través de la dirección multicast 224.0.0.9.
- *Calcular la tabla de encaminamiento.* Cuando un nodo encaminador, que acepta mensajes RIP recibe un RIP '*Response*', examina las entradas de rutas que contiene éste para así comprobar si debe actualizar su tabla de encaminamiento. Si en el mensaje RIP Response recibido aparece una ruta referente a un destino que no se conocía previamente, o a una entrada dinámica que ya conoce pero cuya métrica indica que se podría alcanzar el destino con menos saltos, entonces el nodo encaminador receptor procede a actualizar su tabla de encaminamiento. Para ello añade o modifica la entrada, colocando como puerta de enlace la dirección IP del nodo encaminador que envió el mensaje RIP Response o la del siguiente salto si se especifica en la entrada recibida.

RIP dispone de un temporizador asociado a cada ruta dinámica en la tabla. Un sistema con RIP que encuentra una ruta no actualizada desde hace cierto tiempo (3 minutos) procede a marcarla para su destrucción con el valor infinito (16). La eliminación permanente se retrasa 60 segundos más para asegurarse de que esta acción ha sido notificada al resto de la red con el tiempo suficiente.

En la figura 9.11 se muestra un ejemplo de actualización de tablas de encaminamiento. En dicha figura se muestra la topología de una red LAN cuyos nodos encaminadores (routers) tienen activo el protocolo de encaminamiento dinámico RIP. En la misma figura se muestra la tabla de encaminamiento con las rutas estáticas que tienen asignadas cada nodo encaminador antes de cualquier actualización dinámica. Tras el comienzo del proceso de actualización RIP cada uno de los nodos encaminadores comienza a mandar paquetes RIP Request, y cada vez que un nodo encaminador recibe un paquete RIP Request, procedente de otro nodo encaminador, envía paquetes RIP Response a todos sus nodos adyacentes. Si en un instante determinado, el router1 recibe un paquete RIP Response procedente del router5, será necesario que el router1 calcule su nueva tabla de encaminamiento a partir de la información de rutas que le ha facilitado el router5. Para ello, el router1 compara todas sus

entradas de rutas con las facilitadas por el router5. En caso de que el router5 le haya proporcionado una ruta nueva que inicialmente el router1 no la tuviese almacenada, la incorporará a su tabla de encaminamiento, incrementando en 1 el número de saltos (M: métrica) de la entrada que incorpora y actualizando el nodo encaminador de acceso (Sig: Nodo siguiente) sustituyendo el existente por la dirección del interfaz del router5 que ha enviado el RIP Response al router1. Por otro lado, si el router5 facilita al router1 una ruta ya existente cuya métrica es inferior en 2 o más saltos, el router1 también actualiza la entrada para esa ruta, de modo similar al comentado anteriormente, incrementando el número de saltos en 1 y modificando el nodo encaminador de acceso a la ruta. Nótese como para las redes conectadas directamente se emplea el valor 1 de métrica, indicando que todos los equipos de su red estarán a distancia 1.

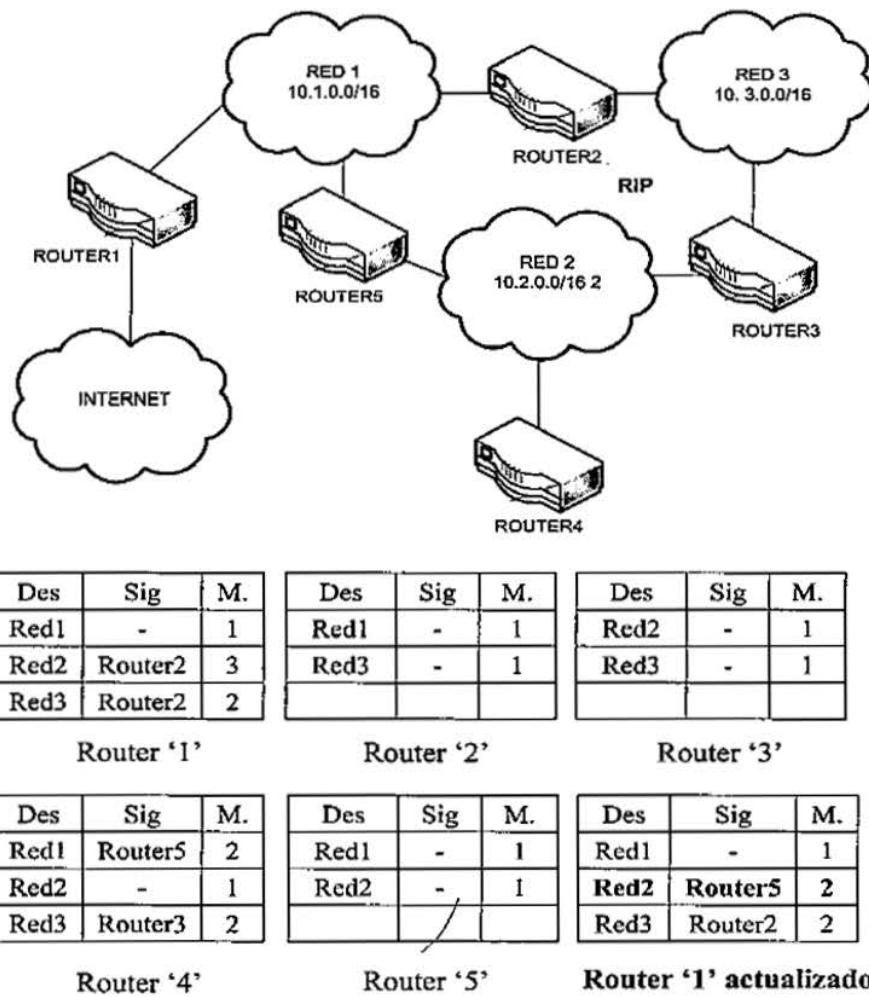
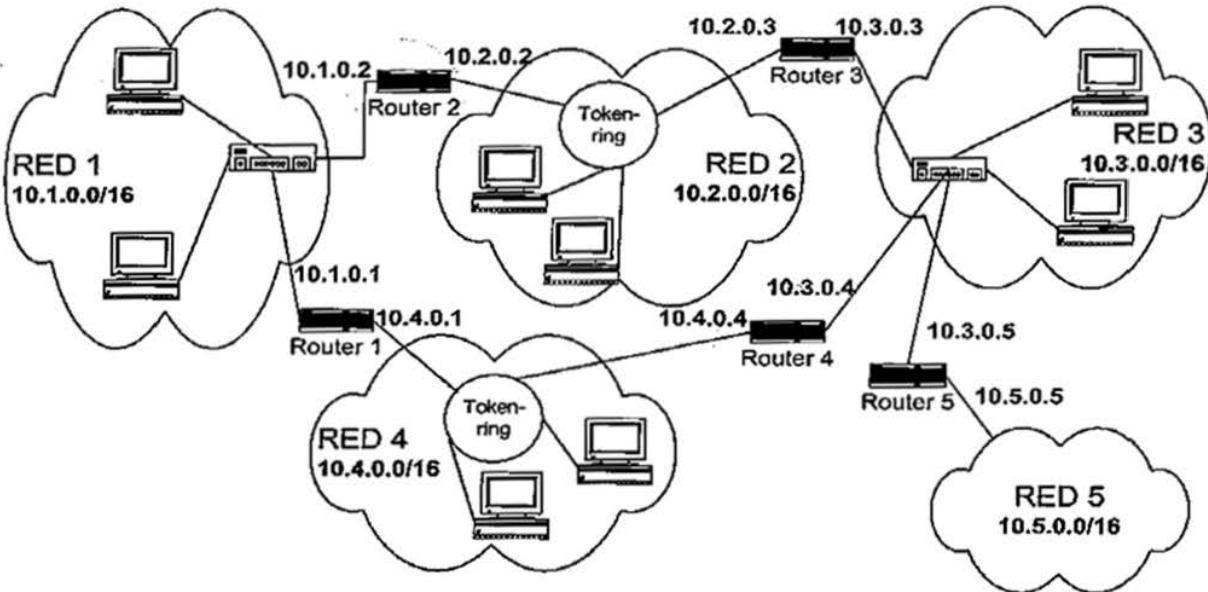


Figura 9-11. Topología de red y tablas de encaminamiento de la red con actualización RIP.

Ejemplo 9.3

Actualización de tablas de encaminamiento con protocolo RIP: Dada la topología de red mostrada en la figura 9.12 y las tablas de encaminamiento que se facilitan para cada uno de los nodos encaminadores, si se conoce que se usa el protocolo de encaminamiento RIP, se pide actualizar las tablas de encaminamiento de cada router e indicar cómo quedarán estas después de la actualización.

Se sabe que el orden de ejecución del protocolo RIP es: Router1, Router2, Router3, Router4 y Router5.



Des	Sig	M.
Red1	-	1
Red4	-	1

Router '1'

Des	Sig	M.
Red1	-	1
Red2	-	1

Router '2'

Des	Sig	M.
Red2	-	1
Red3	-	1

Router '3'

Des	Sig	M.
Red3	-	1
Red4	-	1

Router '4'

Des	Sig	M.
Red3	-	1
Red5	-	1

Router '5'

Figura 9-12. Topología de red y tablas de encaminamiento.

Des	Sig	M.
Red1	-	1
Red4	-	1
Red2	10.1.0.2	2
Red3	10.4.0.4	2

Router '1'

Des	Sig	M.
Red1	-	1
Red2	-	1
Red4	10.1.0.1	2
Red3	10.2.0.3	2

Router '2'

Des	Sig	M.
Red2	-	1
Red3	-	1
Red1	10.2.0.2	2
Red4	10.3.0.4	2

Router '3'

Des	Sig	M.
Red3	-	1
Red4	-	1
Red1	10.4.0.1	2
Red2	10.3.0.3	2

Router '4'

Des	Sig	M.
Red3	-	1
Red5	-	1
Red1	10.3.0.3	3
Red2	10.3.0.3	2
Red4	10.3.0.4	2

Router '5'

9.6 Protocolos EGP: BGP

El protocolo BGP (Border Gateway Protocol) o protocolo de pasarela de frontera es un protocolo que se diseñó para encaminamiento entre diferentes sistemas autónomos. Este protocolo es empleado por los proveedores de servicio de Internet, ISPs. BGP no se basa en el estado del enlace sino en el vector de distancia y fue creado a finales de los 80. Actualmente, la única versión en uso del protocolo BGP es la versión 4. Las normas que determinan el funcionamiento de esta versión están publicadas en RFC 4271. Esta versión funciona en Internet desde mediados de los años 90.

BGP se diseñó para permitir la cooperación en el intercambio de información de encaminamiento entre nodos encamionadores de distintos sistemas autónomos (ver apartado 9.2). Los nodos encamionadores que emplean protocolos del tipo EGP se suelen llamar nodos pasarelas (*gateway*). Estos nodos realizan tres funciones básicas:

- *Adquisición de vecinos.* Este procedimiento se emplea para crear una relación de vecindad y se realiza cuando dos nodos encamionadores que comparten la misma subred física y pertenecen a sistemas autónomos distintos requieren intercambiar información de encaminamiento con cierta regularidad. El procedimiento para llevar a cabo esta función consiste en que un nodo encamionador ofrece al otro la posibilidad de intercambiar información mediante un mensaje denominado *open*, en cuyo caso el otro decide aceptarla con un mensaje *keepalive* o rechazarla.
- *Detección de vecinos alcanzables.* Con esta función, el protocolo BGP genera un procedimiento para mantener la relación de vecindad anteriormente establecida. Durante esta etapa se intercambian periódicamente mensajes *keepalive*, entre un nodo encamionador y uno de sus vecinos, para determinar de qué ambos nodos siguen existiendo y mantienen relación de vecindad.
- *Detección de redes alcanzables.* Cada nodo pasarela mantiene una base de datos con las subredes que puede alcanzar y la ruta completa de cómo hacerlo. Siempre que se producen modificaciones en la base de datos, es necesario que el nodo pasarela lo notifique a los demás nodos encamionadores que implementan el protocolo BGP. Para realizar las notificaciones se emplean mensajes *update*.

Al igual que RIP pero a diferencia de OSPF, los mensajes del protocolo BGP se encapsulan en paquetes del nivel de transporte. En este caso la encapsulación se lleva a cabo dentro de paquetes TCP y no UDP como ocurría en RIP. Esto permite realizar un intercambio de información de un modo confiable, sobre redes que no tienen porqué ser LAN.

Bibliografía

Las URLs proporcionadas eran válidas a fecha de edición del libro. Cuando se han empleado referencias a URLs se ha hecho lo posible porque éstas sean URLs estables cuyas direcciones no cambien. Nuevos materiales, correcciones de erratas y/o actualizaciones estarán disponibles a través del Repositorio Institucional de la Universidad de Alicante (<http://rua.ua.es/>) y/o OpenCourseWare de la misma universidad (<http://ocw.ua.es/>).

Todas las URLs de las RFCs que se encuentran a continuación pueden estar actualizadas, o bien resultar obsoletas debido a la entrada en vigor de nuevas RFCs. Será labor del lector acceder a los sitios recomendados en la bibliografía en busca de las RFCs más actualizadas posibles. Las URLs para búsquedas de RFCs recomendadas son:

- RFCs del ISI: Information Science Institute: <http://www.rfc-editor.org/rfcsearch.html>
- RFCs del grupo RFCs en Español: <http://www.rfc-es.org/>
- RFCs en Internet: <http://www.faqs.org/rfcs/>

Otras fuentes bibliográficas empleadas en la redacción del libro y que permiten realizar aclaraciones conceptuales al lector son, por orden alfabético, las siguientes:

- [Barcia, 2005] N. Barcia, C. Fernandez, S. Frutos, G. Gómez y otros, "Redes de computadores y arquitecturas de comunicaciones. Supuestos prácticos". Prentice-Hall, Madrid, 2005.
- [Beasley, 2008] J. S. Beasley, "Networking". 2º Edición. Pearson Education, Michigan, 2008.
- [Berná, 2002] J.A. Berná, M. Pérez, L.M. Crespo, "Redes de Computadores para Ingenieros en Informática". Publicaciones Universidad de Alicante, Alicante, 2002.
- [Cisco, 2008a] "Academia de Networking de Cisco Systems: Guía del primer año CCNA 1 y 2". 3º Edición. Cisco Press, Madrid, 2008.
- [Cisco, 2008b] "Academia de Networking de Cisco Systems: Guía del segundo año CCNA 3 y 4". 3º Edición. Cisco Press, Madrid, 2008.
- [Forouzan, 2007] B. Forouzan, "Transmisión de datos y redes de comunicaciones". 4º Edición. Mc-Graw Hill, Madrid, 2007.
- [Kurose, 2004] J. F. Kurose, K.W. Ross, "Redes de Computadores: Un Enfoque Descendente Basado en Internet". 2º Edición. Pearson Education, Madrid, 2004.
- [Magaña, 2003] E. Magaña, E. Izkue, M. Prieto, J. Villadangos, "Comunicaciones y Redes de Computadores. Problemas y ejercicios resueltos". Prentice-Hall, Madrid, 2003.
- [Stallings, 2004] W. Stallings, "Comunicaciones y Redes de Computadores". 7º Edición. Pearson Education, Madrid, 2004.
- [Tanenbaum, 2003] A.S. Tanenbaum, "Redes de Computadoras". 4º Edición. Pearson Education, Mexico, 2003.
- [Torres, 2001] F. Torres, F.A. Candelas, S.T. Puente, "Sistemas para la Transmisión de Datos". 2º Edición. Publicaciones Universidad de Alicante, Alicante, 2001.