1

Teoría de las Comunicaciones

INDICE

1. CONCEPTOS BÁSICOS

El siglo XX ha estado dominado por la tecnología de la información, es decir, todos los aspectos relacionados con la recolección, procesamiento y distribución de la información. En este siglo, hemos visto el nacimiento de la radio y la televisión, la extensión por todo el planeta de las redes telefónicas, el nacimiento y la expansión de los ordenadores, así como la puesta en órbita de satélites de comunicaciones.

A medida que se acerca el fin de siglo, estas áreas han ido convergiendo, y las fronteras entre captura, transporte, almacenamiento y procesamiento de la información, son cada vez más tenues. El crecimiento de la demanda de estos servicios es exponencial. A medida que aumenta la capacidad para recoger, procesar y distribuir la información, las exigencias de procesamientos más sofisticados crecen con mayor rapidez.

En estos últimos años se ha producido una drástica reducción en los costes de los equipos informáticos. Simultáneamente, el campo de las comunicaciones ha experimentado también una importante reducción de costes, así como unas mejoras técnicas substanciales. El resultado de esta evolución ha sido la aparición de redes de ordenadores como una solución más barata, fiable y flexible para muchas situaciones prácticas, y a la vez ha abierto la puerta a nuevas aplicaciones impensables anteriormente.

El viejo modelo de un ordenador para satisfacer todas las necesidades de cálculo, ha sido reemplazado por otro en el que un número grande de ordenadores autónomos pero interconectados realizan todo el trabajo. Este tipo de sistemas se conocen como ***redes de ordenadores***. Para aclarar más el concepto, entenderemos por *interconexión* de dos o más ordenadores, aquella situación en la que éstos sean capaces de intercambiar información. La forma física de lograr esta situación no implica la utilización de hilos de cobre, sino que puede realizarse con otras tecnologías como fibra óptica, microondas o conexiones vías satélite. Con *autónomos* queremos excluir aquellos casos donde existe una clara relación maestro/esclavo. Si un ordenador puede forzosamente arrancar, parar o controlar a otro, éstos no se considerarán autónomos. Llamaremos ***sistema distribuido*** a una red de ordenadores en la que la existencia de múltiples equipos autónomos es transparente para el usuario, es decir, no le resulta visible la división del sistema en distintos equipos.

1.1. Uso de las Redes de Ordenadores

1.1.1. Redes de Ordenadores para las Compañías

Muchas compañías y organismos oficiales disponen de un gran número de ordenadores. A menudo, éstos están dispersos por un edificio, una región, por todo el país o incluso en distintas partes del mundo. En estos casos, suele ser habitual que unos ordenadores necesiten acceder a información o recursos disponibles en otro. La mejor forma de lograr resolver este problema es interconectarlos para formar una red.

El objetivo fundamental de la red es pues compartir recursos, es decir, que todos los programas, equipos auxiliares (como impresoras, unidades de cinta, ...) y fundamentalmente los datos, estén disponibles para todos los demás equipos que integran la red independientemente de su ubicación física y de la del usuario.

Un segundo objetivo es proporcionar una mayor fiabilidad como consecuencia de la existencia de varias fuentes alternativas para un mismo recurso. Por ejemplo, todos los datos pueden estar duplicados en varias máquinas, de modo que si una no está disponible pueda recurrirse a otra. Este aspecto es particularmente importante en casos como las redes bancarias o el control del tráfico aéreo. En general, en cualquier caso en el que un sistema deba seguir funcionando a pesar de un fallo.

Otro objetivo básico de las redes de ordenadores desde el punto de vista empresarial es la reducción de costes. En general, la relación prestaciones / precio es más favorable en los ordenadores pequeños que en los grandes supercomputadores. Esta no linealidad hace que un sistema basado en una red de ordenadores personales conectados a uno o varios servidores de ficheros sea más atractiva que otra basada en un número menor de mainframes. A estas arquitecturas se las denomina ***cliente - servidor***.

Una ventaja adicional de las redes de ordenadores desde el punto de vista del coste, es la mayor flexibilidad que ofrece frente a una solución basada en un gran ordenador único. En una red es muy fácil añadir nuevos equipos para ir adaptándola a cargas de trabajo crecientes, mientras que sustituir un gran ordenador es costoso y suele suponer la parada del sistema durante días.

También hay que destacar que una red de ordenadores ofrece a la empresa una forma rápida y sencilla de comunicación entre sus empleados. Por ejemplo, usando la red de ordenadores es fácil que personas que están separadas escriban un informe juntas.

1.1.2. Redes de Ordenadores para el Público General

Si en el apartado anterior se comentaban las principales ventajas técnicas y económicas que una red de ordenadores ofrece a una organización (empresarial o

gubernamental), esta sección está dedicada a estudiar los atractivos que puede ofrecer a los usuarios particulares.

Al comienzo de los años 90 comenzaron a introducirse en los hogares las redes de ordenadores, ofreciendo servicios a clientes particulares. Esta situación ofrecía tres nuevas alternativas:

1. Acceso a información remota.
2. Una nueva forma de comunicación personal.
3. Nuevas formas de entretenimiento.

La primera opción puede permitirnos consultar el estado de nuestras cuentas bancarias, consultar el catálogo de una biblioteca o realizar compras. Dentro de esta categoría podemos incluir el acceso a sistemas de información como el ***World Wide Web*** que contiene información sobre arte, negocios, política, salud, deportes, hobbies, ...

Dentro del segundo apartado, el correo electrónico es una realidad y cada vez su uso se hace más generalizado. Inicialmente era una forma de enviar ficheros de texto entre distintas máquinas. Hoy en día podemos enviar, texto, gráficos, documentos, programas, ... e incluso todo mezclado en un mismo mensaje. Además, la posibilidad de enviar a través de la red, sonido e imágenes, cada vez con mayor calidad y menores retardos, podrían terminar destronando al teléfono.

Desde el punto de vista del entretenimiento, una de las aplicaciones que más interés despierta es el llamado cine a la carta. Otra alternativa ya disponibles es la participación en juegos a través de la red.

1.1.3. Aspectos Sociales

La difusión de las redes de ordenadores ha introducido problemas legales, éticos y sociales. Una aplicación de red de uso generalizado son los grupos de noticias (news), en los que personas con intereses similares, intercambian ideas. Mientras los temas son

técnicos o hobbies como la jardinería o el submarinismo, no existe ningún problema. Éste aparece cuando un grupo se dedica a discutir sobre temas políticos, religiosos o sobre sexo. Algunas opiniones expuestas pueden resultar ofensivas para la mayoría de la sociedad, e incluso rozar la ilegalidad. Son recientes casos como la propaganda de ETA desde un servidor en Suiza, o la pornografía infantil.

Algunos mantienen que debe adoptarse una postura de "vive y deja vivir", mientras que otras voces se alzan a favor de una regulación de los contenidos que pueden difundirse. Conviene tener en cuenta que para el acceso a esta información no se realiza ninguna comprobación sobre el usuario. Por ejemplo, no se puede asegurar que a un menor de edad se le impida el acceso a páginas Web o grupos de noticias relacionados con la pornografía.

Por otra parte, culpar directamente a los proveedores del servicio sería como acusar al servicio postal o a la compañía de teléfonos de permitir la comunicación entre traficantes de droga o terroristas. Además, el establecimiento de un control (o censura) podría entrar en conflicto con la libertad de expresión. En cualquier caso, la solución de este tipo de problemas no es en absoluto trivial. El vacío legal provocado por el rápido desarrollo de esta tecnología deberá ser cubierto de alguna forma, sin lesionar los derechos de los usuarios actuales del servicio.

1.2. Funciones de los Sistemas de Comunicación

Hemos visto las ventajas que nos ofrece una red de ordenadores, pero antes de continuar profundizando en el tema, indicaremos qué funciones son exigibles al sistema de comunicación que permite construir dicha red.

Por sistema de comunicación entenderemos el conjunto de hardware y software que permite la comunicación entre estaciones. Las estaciones están interconectadas mediante caminos (o enlaces) que permiten el envío y/o recepción de la información. En líneas generales, el sistema de comunicación debe permitir:

- Identificar las estaciones que conforman la red.
- Fragmentación y reconstrucción de los mensajes intercambiados.
- Compactación de mensajes.
- Establecimiento de conexiones y multiplexación/demultiplexación de canales.
- Control de errores durante la comunicación.
- Manejo de congestiones y control del flujo de la información.
- Sincronización.
- Establecimiento de distintos niveles de prioridad.

1.2.1. Nombres y Direcciones

En la comunicación, la identificación de las partes que intervienen es fundamental. No sólo hay que saber qué nos están diciendo sino que hay que saber quién lo dice. En general, distinguiremos entre ***nombres***, ***direcciones*** y ***rutas***, a la hora de identificar una estación, o de forma más general, un recurso.

Mediante el nombre identificaremos el recurso al que queremos acceder. Su dirección nos indicará en qué punto de la red se encuentra, y la ruta nos definirá el camino óptimo a seguir para llegar al recurso. La función que se optimiza puede ser el coste de la comunicación, la fiabilidad, el tiempo, o una ponderación de varios de estos criterios.

1.2.2. Fragmentación y Reconstrucción de Mensajes

Resulta evidente que la longitud de la información que se desea enviar o se va a recibir no tiene que coincidir necesariamente con el tamaño del paquete que realmente circula por la red. En ese caso, el mensaje original debe se fragmentado en trozos más pequeños para su envío a través del canal de comunicación. Esta situación obliga a que la estación receptora sea capaz de identificar los diferentes bloques y reensamblarlos con el fin de obtener la información original.

1.2.3. Compactación

En ocasiones, para aumentar la eficiencia de un canal, pueden enviarse en un mismo paquete varios bloques pequeños de información. Es obligación del sistema de comunicación hacer esta tarea de forma transparente al usuario.

1.2.4. Establecimiento de Conexiones y Multiplexación

Para poder establecer un comunicación que involucre varios mensajes, es necesario establecer una sesión o una conexión. La sesión mantiene información sobre el estado de las comunicaciones para permitir la recuperación de la misma tras un error, o bien para ordenar la secuencia de mensajes. En este sentido, una conexión puede verse como un flujo de mensajes entre dos estaciones.

Por otra parte, puede ocurrir que una estación tenga un único canal de comunicación, pero quiera mantener simultáneamente varias sesiones abiertas. Esto obliga a que las distintas sesiones existentes compartan el canal mediante su multiplexación. También puede ocurrir lo contrario, es decir, que una sesión desee emplear varios canales disponibles en una máquina con el fin de aumentar la capacidad de la conexión. Esta multiplexación / demultiplexación del canal exige un control adicional sobre el flujo de mensajes.

1.2.5. Control de Errores

En la comunicación es importante disponer de canales fiables, es decir, libres de errores. Esto incluye tres aspectos fundamentales: detección, corrección y recuperación de errores. Las principales causas de error son el ruido en la línea de transmisión, el deterioro de la información en algún nodo intermedio o la pérdida de paquetes. Así pues, deben detectarse:

- Deterioros en la información (errores a nivel de bit)
- Pérdidas de mensajes
- Duplicación de mensajes
- Mensajes fuera de secuencia.

La detección de errores de bits se logra añadiendo información redundante, por ejemplo usando bits de paridad. Los errores de secuencia se detectan añadiendo a los mensajes identificadores de secuencia únicos. En general, cuando se detecta un error, la solución suele ser la petición de retransmisión del paquete o paquete afectados.

1.2.6. Congestión y Control de Flujo

Un sistema de comunicación puede sufrir los mismos problemas de congestión que las carreteras. Esto es debido a que un gran número de usuarios comparten un número limitado de recursos. Si en un momento dado hay una gran demanda de dicho recurso, éste puede llegar a saturarse y no ser capaz de atender todas las peticiones que recibe. Estamos ante una congestión.

Los ***mecanismos de control de congestiones*** son los medios de que dispone la red para evitar un bloqueo de la misma a medida que aumenta el tráfico de información. Los ***mecanismos de control de flujo***, permiten regular el intercambio de información entre dos entidades de forma que una no envíe más información de la que la otra es capaz de procesar.

1.2.7. Sincronización

Para que pueda existir comunicación entre dos entidades, es necesario que exista una sincronización a distintos niveles:

- **Nivel de bit**: El receptor debe conocer o ser capaz de determinar el comienzo y duración de cada elemento de señal para poder leerla de forma correcta.
- **Nivel de byte**: Muchos sistemas intercambian información en forma de caracteres de 8-bits (byte), aunque varios bytes pueden empaquetarse en un único mensaje para su transmisión. Por ello, el receptor debe ser capaz de distinguir el comienzo y final de cada byte dentro del paquete.
- **Nivel de bloque**: Es necesario determinar el inicio y final de un bloque de bytes. La información contenida en un bloque suele tener un significado u

otro en función de su posición. Es habitual que los bytes iniciales actúen como cabecera y contengan información que permite al protocolo de la capa, controlar la comunicación.

- **Nivel de acceso al medio de comunicación**: En el caso de acceder a un medio de comunicación con estructura de bus, es importante asegurar que sólo un usuario tiene acceso al medio en un instante determinado.
- **Nivel de protocolo**: Dos entidades pares que se comunican y que mantienen información sobre el estado de la comunicación, deben estar sincronizadas al comienzo de la misma o tras un error grave de la comunicación, para poder recuperarla.
- **Nivel de proceso**: Este tipo de sincronización es necesaria para acceder a un recurso compartido como por ejemplo datos comunes almacenados en un disco.

1.2.8. Prioridad

Con el fin de establecer jerarquías a la hora de competir por el acceso a un recurso, pueden establecerse distintos niveles de prioridad para los mensajes. En general, mensajes de alta prioridad sufrirán retardos menores. Un uso típico es la transmisión de alarmas en aplicaciones de control, indicar la parada de una aplicación, o el uso de mensajes de control de comunicación.

2. ASPECTOS HARDWARE Y SOFTWARE DE LAS REDES DE COMPUTADORES

2.1. Hardware de Red

No existe una clasificación en la que estén incluidos todos los tipos de red existentes, sin embargo, sí hay dos claros rasgos que las distinguen: tipo de tecnología y alcance de la red

2.1.1. Tipos de Tecnología

Básicamente hablando, hay dos tipos de tecnologías de transmisión: ***redes broadcast*** o de difusión y ***redes punto a punto***.

En las redes broadcast hay un único canal de comunicación compartido por todas las máquinas de la red. Las máquinas envían mensajes cortos, denominados generalmente tramas, y que son recibidos por todas las demás estaciones. Dentro de la trama suele haber un campo que indica el origen y otro con la especificación del destino, que identifican a la estación que originó la trama y a la que lo debe recibir.

Cuando una máquina recibe una trama, comprueba si la dirección del destino coincide con la suya propia, en cuyo caso la trama será procesada. Si la trama no iba dirigida a la estación, será ignorada. Este tipo de canales también permiten la posibilidad de dirigir una trama a todas las estaciones de la red mediante la utilización de un código de dirección especial. Esta operación se denomina ***mensaje broadcast***. También es posible enviar tramas a grupos de estaciones, lo que se conoce como ***mensaje multicast***. Cada máquina puede pertenecer a uno o varios grupos.

La otra alternativa son las redes punto a punto. En este caso la red se forma mediante múltiples conexiones punto a punto entre pares de máquinas. Para que un mensaje llegue a su destino, puede tener que pasar por uno o varios nodos intermedios. Habitualmente, existe más de un camino, cada uno con su longitud, precio, etc.. Por ello, los ***algoritmos de encaminamiento*** (o ***routing***) resultan vitales. Como norma general (por supuesto con sus excepciones), las redes pequeñas que se extienden en un área geográfica limitada suelen ser redes broadcast, frente a las redes más extensas que suelen ser redes punto a punto. Dentro de este tipo de redes podemos considerar dos clases. Las redes de conmutación de circuitos y las de conmutación de paquetes, también conocidas como redes de almacenamiento y reenvío (store and fordward). En las primeras, al establecer la comunicación, los canales físicos que unen ambos extremos quedan reservados para uso exclusivo hasta que la conexión se libera. En el caso de redes de reenvío, cada nodo intermedio recibe mensajes en forma de paquetes de datos y los almacena hasta que los reenvía hacia su destino final o a otro nodo intermedio.

Broadcast	***Punto a Punto***
Fundamentalmente empleada en redes locales.	Fundamentalmente empleada en redes de largo alcance.
El software es más simple puesto que no necesita emplear algoritmos de routing y el control de errores es extremo a extremo.	Los algoritmos de routing pueden llegar a ser muy complejos. Se necesitan dos niveles de control de errores: entre nodos intermedios y entre extremos.
Para que la estación reciba el mensaje, debe reconocer su dirección en el campo de destino.	La información se recibe y, una vez leído el mensaje, se procesa si va dirigido a la estación, o se reenvía si tiene un destino diferente.
Un único medio de transmisión debe soportar todos los mensajes de la red, por lo que son necesarias líneas de alta velocidad (>1 Mbps).	Varias líneas de comunicación pueden funcionar en paralelo, por lo que pueden usarse líneas de baja velocidad (2-50 kbps).
Los principales retrasos son debidos a las esperas para ganar el acceso al medio.	Los principales retardos son debidos a la retransmisión del mensaje entre varios nodos intermedios.
El medio de transmisión puede ser totalmente pasivo y por ello más fiable.	El medio de transmisión incluye nodos intermedios por lo que es menos fiable.
Se necesitaría duplicar las líneas en caso de que se quiera asegurar la funcionalidad ante fallos.	La redundancia es inherente siempre que el número de conexiones de cada nodo sea mayor que dos.
Los costes de cableado de la red son menores. Sólo es necesario una tarjeta de interfaz por estación.	Los costes de cableado son superiores, y la estación requiere al menos dos tarjetas de interfaces.

2.1.2. Tipos de Red por su Extensión

Redes de Area Local (LAN)

En general, una LAN es una red privada cuya extensión está limitada en el espacio: un edificio, un campus o en general una extensión inferior a unos cuantos kilómetros. Su aplicación más extendida es la interconexión de ordenadores personales y estaciones de trabajo en oficinas y fábricas para compartir recursos e intercambiar datos y aplicaciones.

Una LAN suele distinguirse por tres características:

1. Tamaño
2. Tecnología de transmisión
3. Topología

Las LAN están limitadas en el espacio, eso implica que para un determinado medio de transmisión es posible saber el tiempo máximo de transmisión. Este dato permite el uso de ciertos diseños y simplifica la administración.

En cuanto al medio de transmisión, suelen emplear enlaces que consisten en un único cable al que se conectan todas las máquinas que componen la red. Se alcanzan velocidades de entre 10 y 100 Mbps, con retardos muy bajos.

Las topologías más típicas son las conexiones en bus, anillo o estrella.

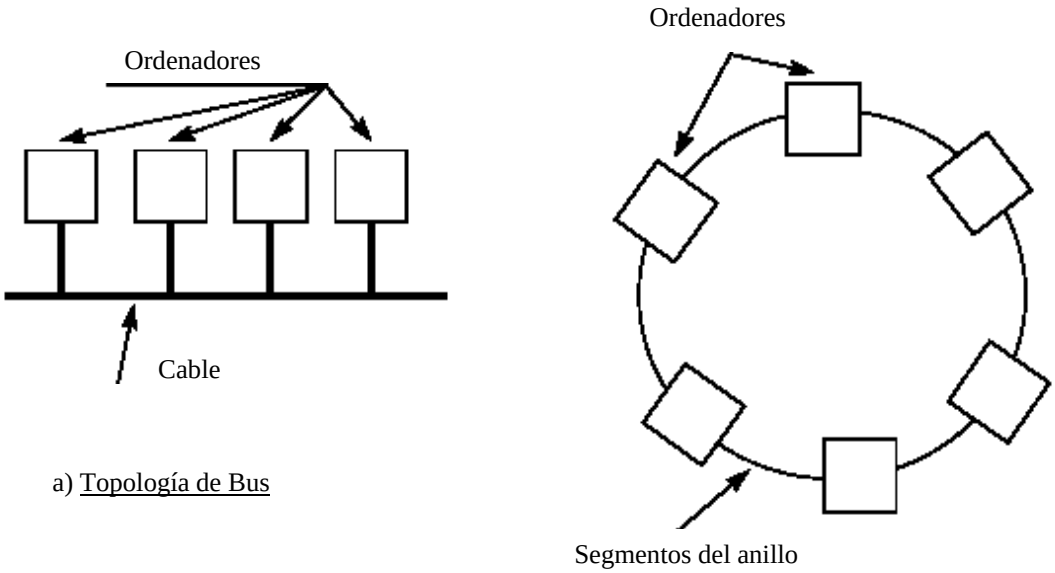


Fig. 1: Tipos de redes de área local

Redes de Area Extendida (WAN)

Una WAN se caracteriza por ocupar una gran área geográfica (hasta un continente entero). Contiene una serie de ordenadores en los que corren las aplicaciones de los usuarios (también conocidos como hosts), que se conectan mediante lo que se viene en llamar subred.

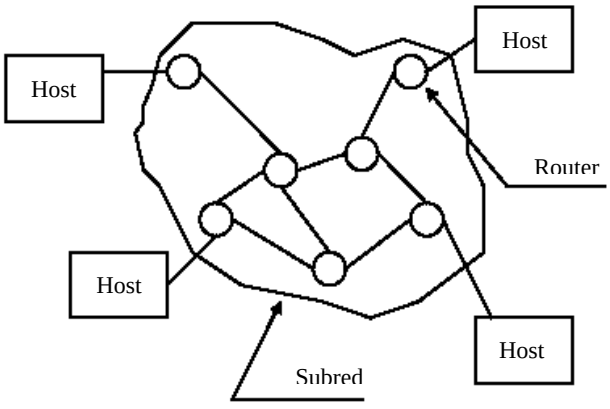


Fig. 2: Organización de una red WAN

El trabajo de la subred consiste en llevar los mensajes de un host a otro. Al separar las partes estrictamente relacionadas con la transmisión de datos, de los aspectos relacionados con la aplicación, el diseño se simplifica enormemente. En general, la subred está formada por líneas de transmisión y nodos de conmutación de paquetes.

Redes de largo alcance (WAN)	***Redes de área local (LAN)***
Distancias de hasta miles de Kilómetros.	Distancias inferiores a unos pocos kilómetros.
Velocidades típicas inferiores a 10 Mbps.	Velocidades típicas superiores a 10 Mbps.
Protocolos complejos.	Protocolos simples.
Interconecta sistemas de ordenadores independientes.	Interconecta ordenadores que cooperan, habitualmente formando un sistema distribuido.
Suelen ser públicas y administrada por empresas u organismos nacionales.	Suelen ser privadas y administradas por sus propietarios.
Habitualmente usa circuitos de la red telefónica para sus conexiones.	Suele emplear comunicaciones digitales sobre cables propios.
Tasas de error altas (1 bit erróneo entre cada 10^5 bits transmitidos).	Tasas de error bajas (1 bit erróneo entre cada 10^9 bits transmitidos).
Suele emplear enlaces punto a punto.	Suele emplear redes broadcast.
Suele emplear estructura de interconexión parcial o de estrella.	Las topologías habituales son bus o anillo.

Las líneas de transmisión, también llamadas circuitos o canales, se encargan de mover la información de una máquina a otra. Los nodos de conmutación de paquetes o routers son ordenadores especializados que se emplean para conectar dos o más líneas

de transmisión. Cuando llegan datos por una línea de entrada, el router selecciona el canal de salida más adecuado para enviar el mensaje hacia su destino. Habitualmente los hosts son redes LAN con un router.

En las redes WAN, la subred contiene numerosos cables o líneas de teléfono que interconectan pares de nodos. Si dos routers no conectados directamente desean intercambiar información deben hacerlo a través de nodos intermedios.

2.2. Topologías de Red

La topología de la red define la estructura de las conexiones entre estaciones. El tipo de topología influye en:

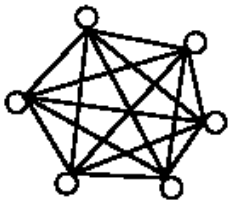
- El coste de ampliación de la red.
- La facilidad para reconfigurar la red.
- La fiabilidad, es decir, el grado de dependencia de un único componente de la red.
- La complejidad del software.
- El rendimiento
- La posibilidad de enviar mensajes broadcast

Además de la diferenciación entre redes broadcast y punto a punto comentada en la sección 2.1, las redes pueden adoptar distintas configuraciones físicas, que se mezclan a medida que aumentamos la extensión geográfica.

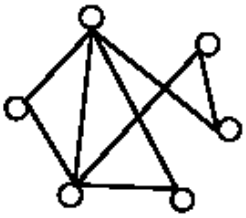
2.2.1. Conexión total

Entre cada par de estaciones de la red existe un canal punto a punto dedicado. Los diferentes canales pueden funcionar simultáneamente de forma que la cantidad de información que puede distribuir es alta, y los retrasos son pequeños. El software es sencillo puesto que no son necesarios algoritmos de routing.

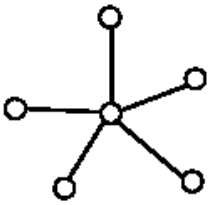
El nivel de fiabilidad es muy alto dado que pueden fijarse varios caminos alternativos si un enlace falla. El principal inconveniente es el coste. Una red de este tipo con **n** estaciones tiene **(n-1)n/2** enlaces bidireccionales y cada estación necesita **(n-1)** tarjetas de interfaz con la red, una por enlace. El coste y la dificultad de añadir un nuevo nodo es evidente. El envío de un mensaje broadcast exige que se envíe a través de cada enlace. Su uso suele estar restringido a redes pequeñas en las que la redundancia es vital.



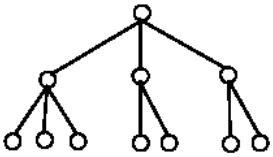
Interconexión total



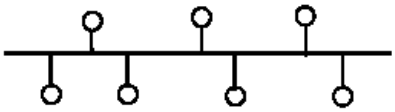
Interconexión parcial



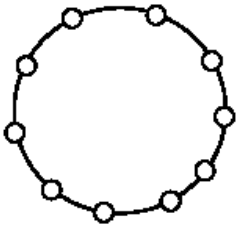
Interconexión en estrella



Interconexión en árbol



Interconexión en bus



Interconexión en anillo

Fig. 3: Topologías de red

2.2.2. Conexión parcial

Es una red en la que existen enlaces punto a punto entre pares de estaciones, pero no todos los posibles pares están conectados. Para algunas comunicaciones será obligatorio el uso de nodos intermedios. Si cada estación tiene por lo menos un par de canales disponibles, la fiabilidad del conjunto resulta alta y con un coste relativamente bajo. La selección y dimensionado de los enlaces puede hacerse de acuerdo al tráfico previsible entre nodos.

Los retardos pueden ser relativamente altos en función del número de enlaces y del tráfico existente, así como del origen y del destino. Es necesario incluir algoritmos de routing. El envío de mensajes broadcast no es fácil de implementar.

2.2.3. Conexión en estrella

Todas las estaciones se conectan mediante un único enlace a un nodo central. Esto facilita la expansión de la red al ser barato y resultar fácil de configurar. Por contra, el funcionamiento del nodo central, que puede ser activo o pasivo, resulta crítico y los retrasos aumentan al tener que circular todos los mensajes a través de dicho nodo. El riesgo de fallo es pues elevado.

2.2.4. Conexión en árbol o jerárquica

Es una extensión de la red en estrella pudiendo considerase como un conjunto de estrellas cuyos nodos centrales se conectan a otro, de ahí que sus propiedades sean semejantes a las de la conexión en estrella. Suele usarse en sistemas de control puesto que refleja de forma natural la jerarquización de los diferentes niveles de control: desde la planificación general hasta el regulador de cada máquina individual. Sin embargo, un fallo puede aislar una rama de la red.

2.2.5. Bus serie

Básicamente, en un bus se envía un mensaje broadcast a todas las estaciones dado que el medio de transmisión es compartido por todas las estaciones. Para evitar que varias estaciones accedan a la vez al canal, es necesario incorporar un mecanismo de acceso y detección de colisiones.

El coste de instalación es bajo, y resulta muy fácil añadir estaciones nuevas. El software de comunicaciones no necesita incluir algoritmos de routing. El medio de transmisión puede ser totalmente pasivo y por tanto, básicamente fiable. Todo esto hace que la conexión en bus resulte muy atractiva para su uso en redes de área local.

Para aumentar la fiabilidad, puede duplicarse el bus, mientras que la longitud puede aumentarse mediante el uso de repetidores para evitar la atenuación de la señal. También hay que tener en cuenta que el bus debe ofrecer una gran capacidad para absorber el flujo de datos generados por todas las estaciones.

2.2.6. Conexión en anillo

Cada estación está unida a su vecina por un enlace unidireccional y la comunicación sigue ese camino hasta completar el lazo. También es necesario disponer de un mecanismo de acceso al medio.

Es sencillo incorporar nuevas estaciones al anillo aunque el tamaño de éste no puede crecer indefinidamente. El software es sencillo al no necesitar algoritmos de encaminamiento. Los retrasos suelen ser pequeños. Pero, el fallo de un enlace provoca el fallo de todo el anillo.

2.3. Técnicas de Conmutación en Redes

De las distintas topologías en red se deduce que no siempre va a existir un enlace físico directo entre dos estaciones. En este caso la red debe establecer las “conexiones”

necesarias para proporcionar un camino físico o lógico entre las estaciones. Existen dos modelos básicos: conmutación de circuitos y conmutación de paquetes.

El modelo de conmutación de circuitos se basa en las líneas telefónicas, de forma que las estaciones intermedias que intervienen en la comunicación conectan circuitos de entrada y salida hasta establecer un canal físico entre los extremos. Los nodos intermedios sólo intervienen en la creación y eliminación del circuito. La subred no necesita proporcionar ningún procesamiento o almacenamiento de los datos que transmite.

La conmutación de paquetes sigue una filosofía completamente distinta. Es habitual que un mensaje largo se subdivida en otros más pequeños (de entre 100 y 2000 bytes) para su transmisión a través de la red. En el caso de una red de conmutación de paquetes, estos paquetes se multiplexan por los distintos canales de comunicación de un nodo para su envío hacia el destino final. Si el destino no está disponible, el mensaje se descarta. La diferencia fundamental es, que en este caso, no existe un canal físico único y constante durante toda la comunicación sino que, en función de la ocupación de la red en cada momento, los distintos paquetes irán por caminos físicos distintos hacia su destino.

2.4. Arquitectura de Red

Para la interconexión de sistemas abiertos se construyen arquitecturas de red. Se define un ***sistema abierto*** como: "un sistema capaz de interconectarse con otros de acuerdo con unas normas establecidas". La ***interconexión de sistemas abiertos*** se ocupará del intercambio de información entre sistemas abiertos y su objetivo será la definición de un conjunto de normas que permitan a dichos sistemas cooperar entre sí.

El diseño de una red de ordenadores es un problema suficientemente complejo como para que deba estructurarse si quiere ser resuelto con éxito. Como en otros aspectos de la computación, la técnica empleada es la división en ***capas*** o ***niveles***. Estas capas están jerarquizadas y dividen el problema en partes más sencillas. Cada capa

añade nuevas características a partir de los servicios que proporciona la capa inmediatamente inferior.

Una capa se implementa mediante un cierto número de ***entidades*** que llevan a cabo las funciones asignadas a la capa y equivalen a procesos software o dispositivos hardware inteligentes. Entidades pertenecientes a capas equivalentes en dos equipos diferentes de llaman ***entidades homólogas*** (***peers***).

Un ***protocolo*** es el conjunto de reglas (semánticas y sintácticas) que gobiernan la comunicación entre entidades de un misma capa. Es decir el protocolo de la capa N intercambia información con su homóloga en la máquina destino, de cara a proporcionar los servicios asignados a esa capa. Para ello, hará uso de los servicios que proporciona la capa anterior.

a) El sistema de interconexión está formado por un conjunto de *entidades* situadas en diferentes *capas*.
b) Las *entidades* de una determinada *capa* N cooperan entre sí de acuerdo con un determinado *protocolo* N.
c) Las *entidades* de una *capa* N utilizan los *servicios* N-1 proporcionados por las *entidades* de las *capas* inferiores, mediante un *acceso* a ellos. La estructura de estas *capas* es desconocida para la *capa* N, la cual, sólo tiene en cuenta los *servicios* proporcionados por lo que se ha denominado *bloque* N-1.
d) Las *entidades* de una *capa* N realizan unas determinadas *funciones* N, utilizando los *servicios*.

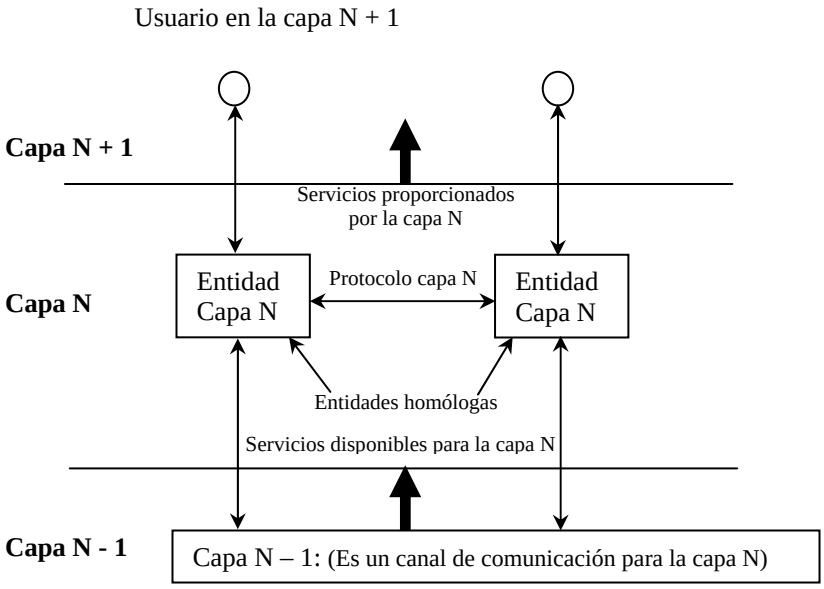


Fig. 4: Relación entre capas, protocolos y servicios

Una capa, la N, proporcionará a la capa inmediatamente superior, la N+1, una serie de servicios. Para ello puede usar los servicios ofrecidos por la capa N-1. Por ejemplo, a partir de un enlace físico con errores, se podría construir un enlace lógico libre de errores. Una capa puede ofrecer más de una clase de servicio. Conviene dejar claro que:

- No todas las funciones que realiza una capa deber ser vistas como servicio por la capa siguiente
- La especificación del servicio no detalla la forma en que éste está implementado. De esta forma, un servicio puede ofrecerse con distintas calidades en función de la implementación.

Se entiende por ***arquitectura de red*** el conjunto de capas y protocolos de capas que constituyen el sistema de comunicaciones.

Según ISO, el modelo que hemos definido, es válido para configuraciones simples como sería el caso de una línea punto a punto dedicada. Pero para cubrir

configuraciones más complejas, como es el caso de interconexiones a través de una red pública de transmisión de datos, se elaboró otro modelo en el que se ha permitido el encadenamiento entre bloques o capas.

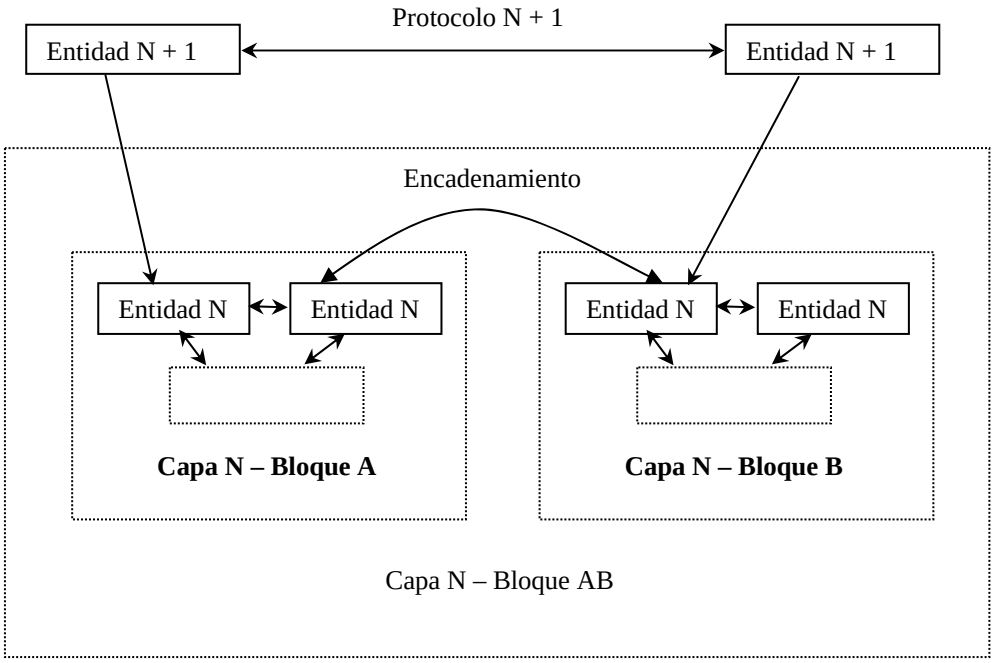


Fig. 5: Encadenamiento entre capas

2.5. Unidades de Información Transmitidas en la Comunicación

Para acceder a un servicio, se recurre a un ***SAP*** (Punto de Acceso al Servicio), que puede entenderse como el punto en el que interactúan dos capas contiguas de una misma estación. Puede haber más de un SAP entre dos capas, en función de la calidad de servicio que se requiera. Cada SAP tiene una dirección que lo identifica. En UNIX, los SAP son los puertos o sockets y su dirección, el número del puerto o socket.

Para permitir la comunicación entre dos capas, debe existir un conjunto de reglas que definan la interfaz. Así, la interfaz incluirá aspectos físicos (conectores, niveles eléctricos) y/o lógicos (estructuras de datos, temporización, etc.) que permitan la

interconexión de las capas. Cada máquina puede tener sus propias interfaces entre capas, sin que esto afecte a la comunicación entre capas equivalentes.

En una comunicación típica, la capa N+1 pasa una ***IDU*** (Unidad de Datos de la Interfaz) a través de un SAP a la capa N dentro de la misma máquina. Una IDU está compuesta por una información de control de la interfaz (***ICI***) y una parte de datos o ***SDU*** (Unidad de Datos del Servicio). La SDU es la información para la que se requiere el servicio, mientras que la ICI es la información que necesita la interfaz para proporcionar el servicio en la forma deseada.

Mientras la ICI puede variar de una máquina a otra, la SDU permanece invariable. La SDU de la capa N junto con la cabecera y la cola que forman la información de control del protocolo (PCI), integran la llamada ***N-PDU*** (Unidad de Datos del Protocolo) de la capa N. Si la información no se fragmenta, la información de la SDU de la capa N coincide con los datos de la PDU de la capa N. Si por el contrario, la información es fragmentada, se formarán varias PDU de capa N. Estas trozos deberán ser reensamblados en el destino para obtener la SDU.

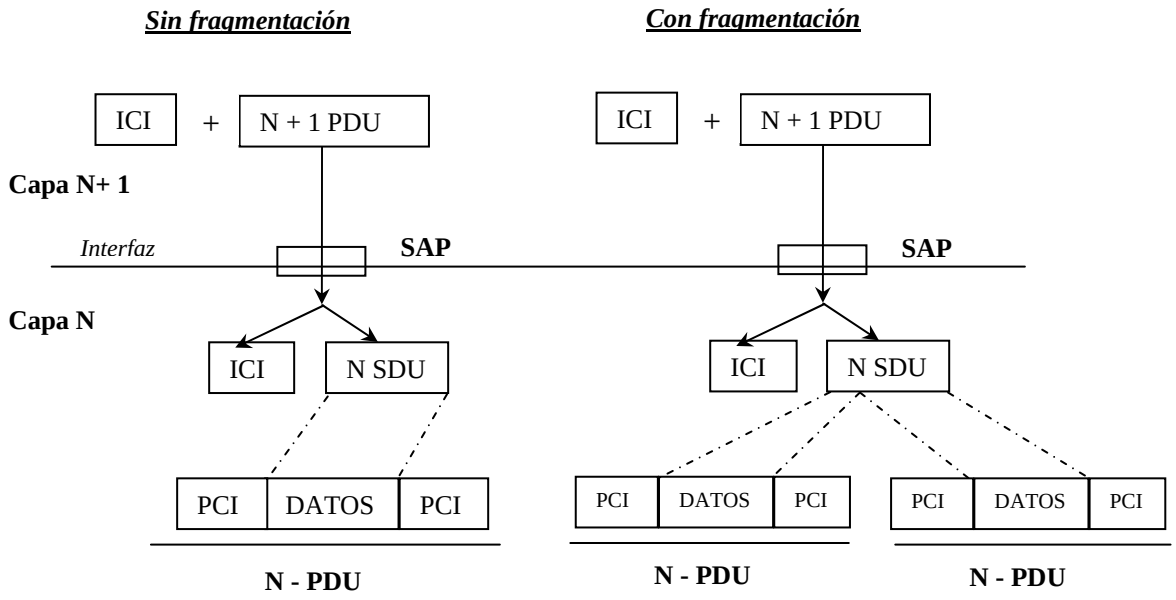


Fig. 6: Uso de los puntos de acceso al servicio

Un servicio ofrecido por una capa, puede mapearse directamente sobre un servicio de la capa inferior, o bien, la capa puede disponer de un protocolo que le permita mejorar el servicio que ofrece la capa inferior (por ejemplo corrección de errores). En cualquier caso, el usuario de los servicios de una capa, debe ver a ésta como una caja negra.

2.6. Clasificación de los Servicios de Comunicación

Los servicios ofrecidos por las distintas capas de una arquitectura de red pueden clasificarse en dos categorías: servicios sin conexión o servicios orientados a conexión.

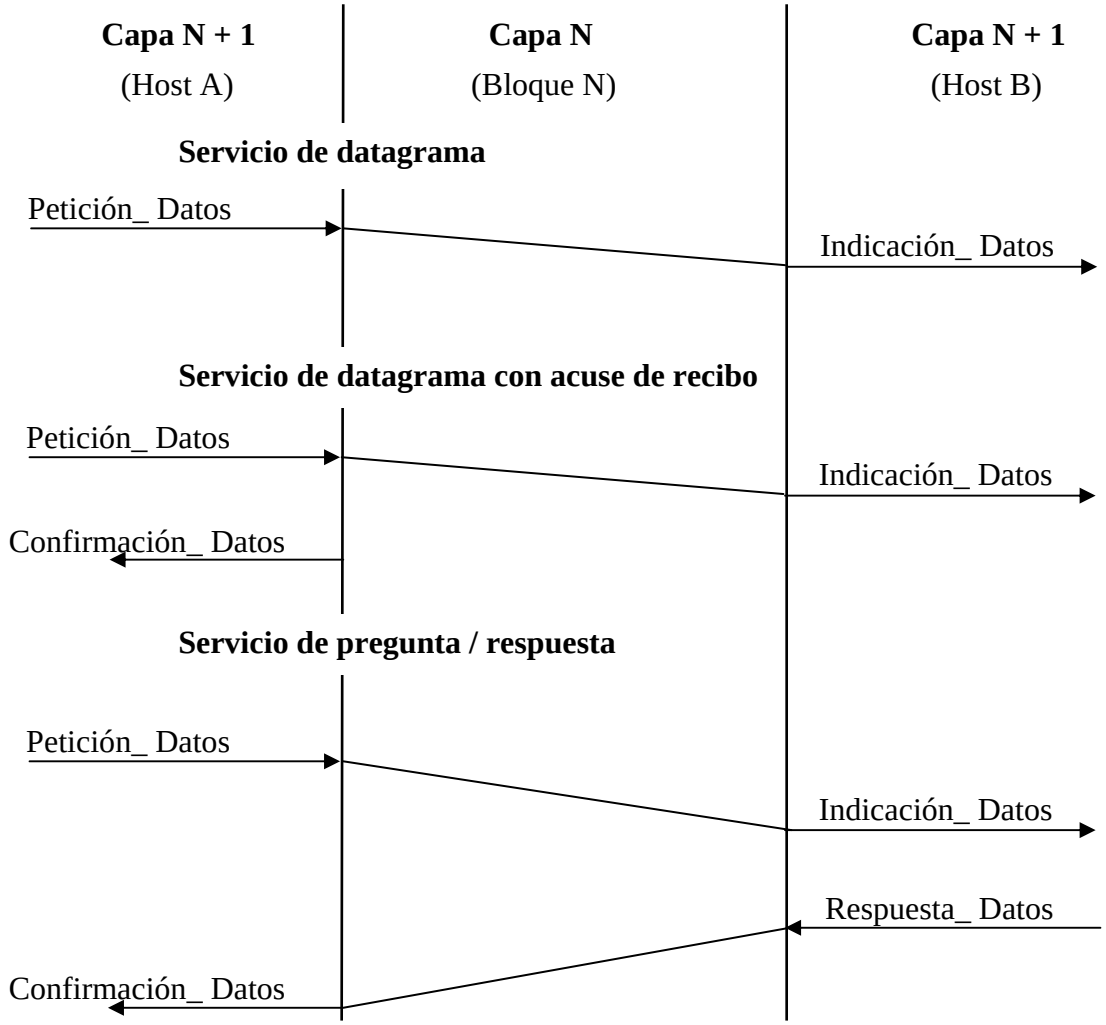


Fig. 7: Posibles servicios sin conexión en la Capa N

2.6.1. Servicios sin conexión

Un servicio sin conexión es más fácil de implementar ya que la capa tiene poco que hacer para mejorar el servicio ofrecido por la inferior. La capa no tiene que preocuparse de resolver problemas ocasionados por la pérdida de mensajes, su duplicación, o de la llegada de mensajes fuera de secuencia.

El manejo de un mensaje es totalmente independiente de mensajes anteriores y/o posteriores. Cada mensaje puede tener un destino diferente, y por tanto cada uno debe llevar la indicación completa de su destino.

El tipo de servicio sin conexión más simple es el ***datagrama***. El receptor no responde al mensaje. Este servicio es denominado en ocasiones "envía y reza", ya que el usuario depende de la fiabilidad de la red y no recibe confirmación de si el mensaje ha sido o no, recibido. Un servicio de datagramas suele ofrecer la posibilidad de transmisiones broadcast y multicast.

Una mejora evidente sobre el servicio de datagramas es el ***datagrama con acuse de recibo***. En este caso, la capa que ofrece el servicio proporciona una respuesta para cada mensaje enviado a través de la misma. La confirmación del mensaje no contiene datos del receptor puesto que el acuse de recibo es generado por la propia capa. Aunque el usuario que envía la información sabe si el mensaje llegó o no, el servicio no evita la pérdida, duplicación o salida de secuencia del mensaje.

Un servicio sin conexión alternativo es el ***servicio de pregunta / respuesta*** (request /reply), en el que el usuario que envía la información emplea un datagrama simple y espera a que el receptor le envíe la contestación.

2.6.2. Servicios orientados a conexión

El servicio se modeló basándose en el sistema telefónico. En este caso, el usuario del servicio establece una conexión con el destinatario, la usa y después la libera. El aspecto fundamental es que una vez establecida la conexión, ésta es similar a un tubo: el

que envía introduce objetos por un extremo y le receptor los recoge, en el mismo orden, por el otro extremo.

La dirección completa del destinatario debe ser conocida para el establecimiento de la comunicación. Después, un identificador de la conexión sirve para identificar al usuario remoto durante la transferencia de datos. Las fase de establecimiento puede usarse para negociar la calidad del servicio o cualquier otra opción disponible.

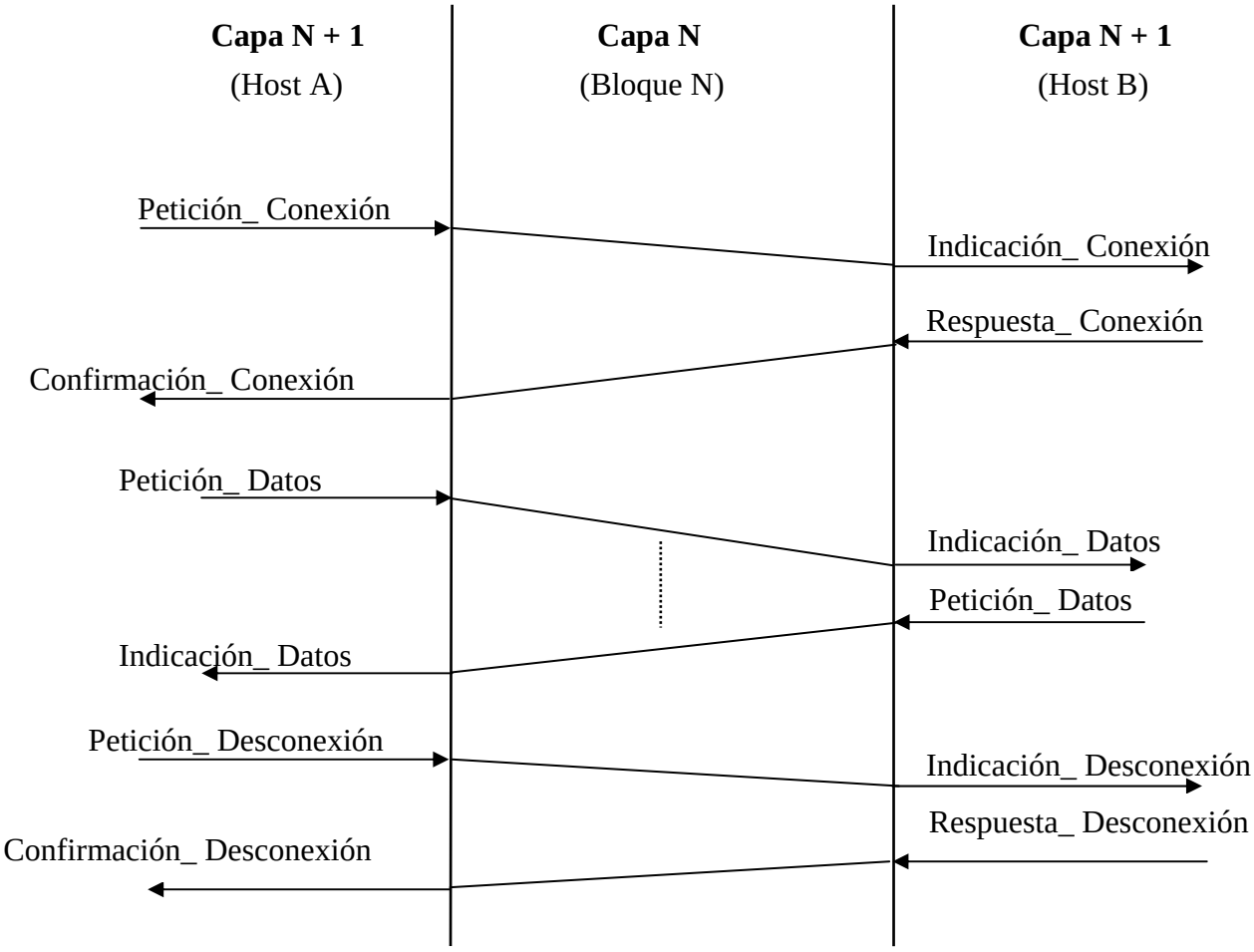


Fig. 8: Servicio orientado a conexión en la Capa N

La forma más común de los servicios orientados a conexión son las conexiones libres de errores. Tras el establecimiento de la conexión, cada usuario puede enviar diferentes mensajes que son enviados al otro en la misma secuencia. Cualquier error que no pueda recuperar la capa, automáticamente se transmite a ambos extremos como una

pérdida de conexión. Debido a la similitud del funcionamiento con el caso de la conmutación de circuitos, este servicio se denomina en ocasiones ***servicio de circuito virtual***. Este tipo de conexión es adecuado para aplicaciones como la transferencia de ficheros o sesiones de terminal remoto. Debido a lo costoso del establecimiento del circuito, suele ser habitual que se limite el número máximo de conexiones simultáneas que se pueden establecer.

3. MODELOS DE REFERENCIA

3.1. El modelo de referencia OSI de ISO

El modelo OSI (Open Systems Interconnection) de ISO (International Standards Organization) fue una propuesta para la standarización de las redes de ordenadores. Este modelo tiene siete capas, diseñadas con arreglo a los siguientes principios:

1. Una capa se creará en situaciones en las que se requiera un nivel diferente de abstracción.
2. Cada capa deberá realizar una función bien definida.
3. La función que realiza cada capa deberá seleccionarse teniendo en cuenta la minimización del flujo de información a través de las interfaces.

4. El número de capas será suficientemente grande como para que funciones diferentes no estén en la misma capa, y suficientemente pequeño para que la arquitectura no sea difícil de manejar.

El modelo OSI por sí mismo, no es una arquitectura de red puesto que no especifica el protocolo que debe usarse en cada capa.

3.1.1. Capa física

La capa física se ocupa de la transmisión de bits a través de un canal de comunicación. Debe asegurar que cuando un extremo envía un bit con valor 1, sea recibido como tal en el otro extremo. Los problemas de diseño a considerar aquí son los aspectos mecánico, eléctrico, de interfaz y el medio de transmisión física.

3.1.2. Capa de enlace

Su principal tarea consiste en proporcionar una línea sin errores a partir de un medio de transmisión cualquiera. Esta capa debe crear y reconocer los límites de las tramas; además debe resolver los problemas creados por el deterioro, pérdida o duplicidad de tramas. La capa de enlace ofrece distintos servicios a la capa de red, cada uno con distinta calidad y precio.

También deberá incluir algún mecanismo de regulación del tráfico que permita evitar que un emisor muy rápido sature a un receptor muy lento.

3.1.3. Capa de red

La capa de red se ocupa del control de la operación de la subred. Un punto vital de su diseño, es la decisión sobre cómo encaminar los paquetes del origen al destino. El encaminamiento puede basarse en unas tablas estáticas o bien determinarse dinámicamente en función del tráfico de red. También debe detectar y corregir problemas de congestión de tráfico.

En ocasiones también incluye funciones de contabilidad para el cobro de los servicios de subred. La capa de red también debe resolver los problemas de comunicación entre distintas redes.

3.1.4. Capa de transporte

La principal función es aceptar los datos de la capa de sesión, dividirlos si es necesario y pasarlos a la capa de red. Además debe asegurar que todos lleguen correctamente al otro extremo. Este trabajo debe hacerse de forma eficiente para aislar la capa de sesión de cambios en el hardware.

Lo habitual es establecer una conexión de red distinta para cada conexión de transporte solicitada por la capa de sesión. Si la conexión de transporte necesita un gran caudal, ésta podría crear múltiples conexiones de red. Por otra parte, si el mantenimiento de una conexión de red es costoso podría multiplexar varias conexiones de transporte sobre la misma conexión de red.

La capa de transporte determina qué tipo de servicio debe dar a la capa de sesión. El tipo de conexión más habitual es el punto a punto libre de errores. La capa de transporte es la primera capa extremo a extremo dentro de la jerarquía. Debe preocuparse del establecimiento y liberación de conexiones, así como proporcionar mecanismos de control de flujo y de congestiones.

3.1.5. Capa de sesión

Una capa de sesión permite que los usuarios de diferentes máquinas puedan establecer sesiones entre ellos. Un servicio de la capa de sesión es gestionar el control de diálogo. Puede permitir que el tráfico vaya en las dos direcciones simultáneamente, o bien alternativamente, en cuyo caso determinará qué estación tiene el turno.

Otro servicio asociado a la capa de sesión es la administración del testigo. También debe encargarse de la sincronización. Esto implica la inserción de puntos de

verificación en el flujo de datos, en los que puede retomarse la conversación en caso de fallo.

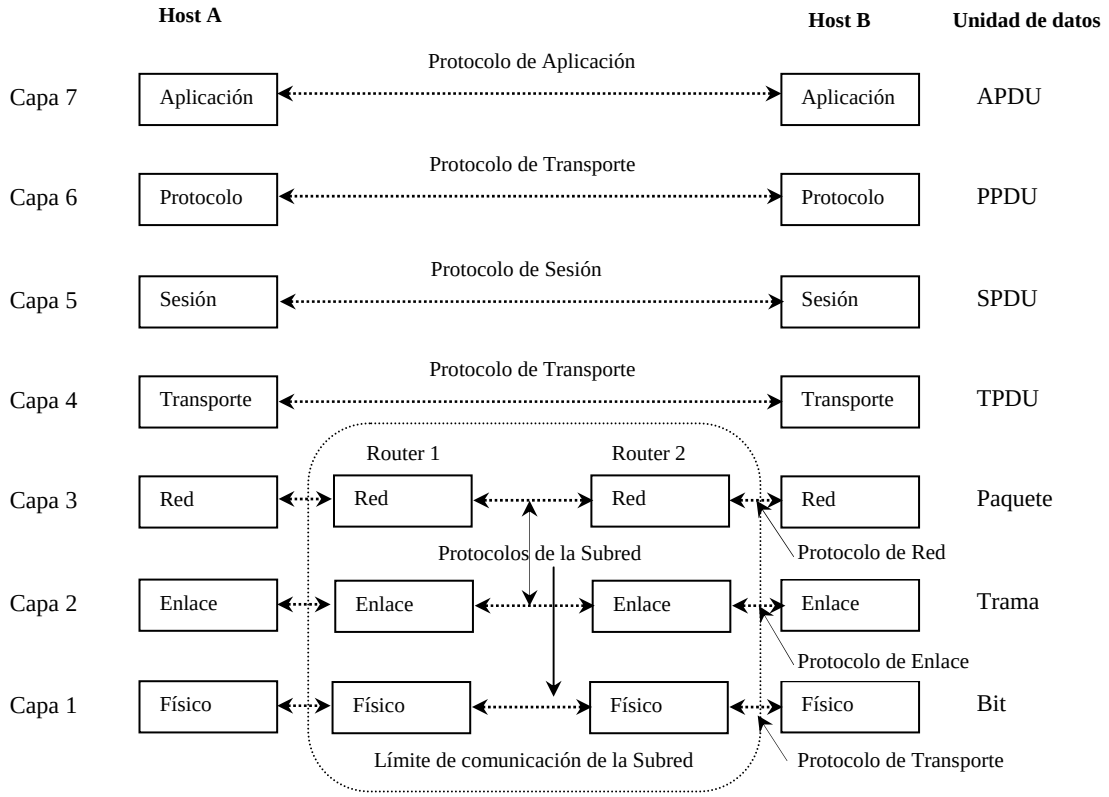


Fig. 9: Modelo de referencia OSI

3.1.6. Capa de presentación

La capa de presentación se ocupa de los aspectos de sintaxis y semántica de la información que transmite. También puede ocuparse de la compresión y encriptación de los datos intercambiados.

3.1.7. Capa de aplicación

Contiene una cantidad de protocolos usados frecuentemente, como por ejemplo ofrecer servicios de terminal virtual, transferencia de archivos, correo electrónico, ejecución remota de procesos, etc.

3.1.8. Transmisión de datos en el modelo OSI

Una vez vistas las distintas capas que especifica el modelo de referencia OSI, conviene estudiar la forma en que se produce una comunicación. Supongamos que el proceso emisor tiene una información que enviar, para ello, entregará los datos a la capa de aplicación. La capa de aplicación añade a la información que recibe, una cabecera (que puede ser nula) que permite a la capa seguir el protocolo que tenga definido. El conjunto formado por los datos originales y la cabecera de aplicación es entregado a la capa de presentación.

La capa de presentación transforma este bloque de distintas formas, en función del servicio pedido, y añade una nueva cabecera, la correspondiente a la capa de presentación. El nuevo conjunto de datos es entregado a la capa inmediatamente inferior, la capa de sesión. Es importante destacar que la capa de presentación no distingue qué parte de los datos que recibió corresponden a la cabecera de la capa de aplicación y qué parte son los datos del usuario.

Es importante hacer notar que, en una o varias de las capas, el conjunto de datos que recibe la capa N de la N+1 puede ser fragmentado en bloques más pequeños para su entrega a la capa N-1. En ese caso, cada bloque recibirá su propia cabecera y además la capa que realiza la fragmentación deberá ser la encargada (en la máquina receptora) de reensamblar los bloques hasta formar el conjunto inicial de datos, y entregarlos a la capa superior.

El proceso se repite hasta llegar a la capa física, momento en el cual los datos son enviados a través del canal físico disponible hacia la máquina de destino. La capa física de la estación receptora recibirá el conjunto de bits del mensaje y comenzará el proceso inverso. Capa a capa deberá ir eliminando las distintas cabeceras y transmitiendo el resultado hacia las capas superiores hasta llegar al proceso receptor.

Evidentemente, el objeto de añadir y eliminar las cabeceras no es tener algo que hacer, sino que las cabeceras permiten a cada capa suministrar el servicio que le fue requerido por la capa superior de acuerdo al protocolo establecido para la capa. De esta

manera, la comunicación funciona como si cada capa se comunicase directamente con su homóloga en la máquina de destino a través de un canal lógico proporcionado por el resto de capas en ambas máquinas.

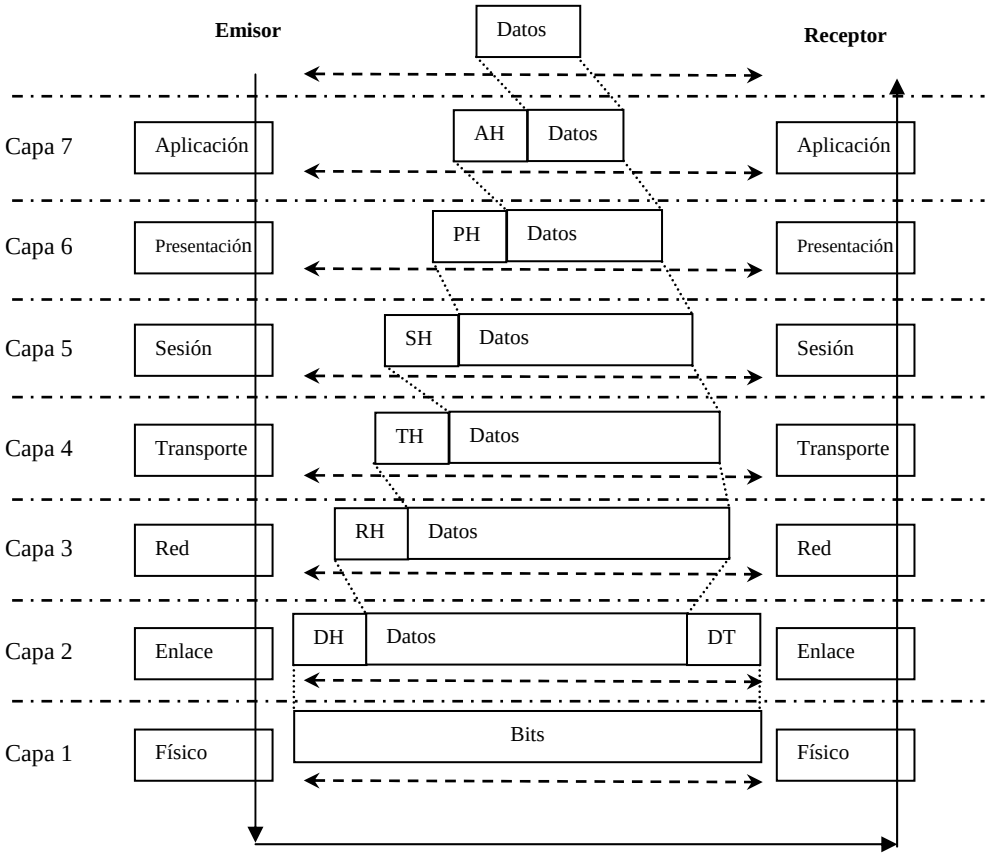


Fig. 10: Transmisión de datos en un modelo de capas

Aunque la idea puede parecer rebuscada, es similar a lo que sucede en la comunicación entre personas. Inicialmente tenemos una idea que queremos comunicar a nuestro contertulio. Esa idea se entrega a la zona del cerebro que se encarga del lenguaje. A su vez, el área del lenguaje se encargará de generar los impulsos nerviosos necesarios para hacer vibrar nuestras cuerdas vocales. Esta vibración se transformará en un sonido recogido por el oído de nuestro interlocutor. Los impulsos nerviosos generados por su oído serán enviados al cerebro que los transformará en palabras, y de ellas extraeremos la idea.

El proceso de la comunicación es similar si el área del lenguaje decide enviar la información al área encargada de la escritura. En este caso, el área del lenguaje estará pidiendo un servicio diferente a la capa inferior: escribir en lugar de hablar. Además, el medio físico empleado será distinto, papel en lugar del aire. En cualquier caso nosotros sólo somos conscientes de que enviamos o recibimos un pensamiento.

3.1.9. Ejemplos

A modo de ejemplo en las páginas siguientes se muestra cómo dos sistemas abiertos interconectados realizan el intercambio de información. Se ha supuesto una red formada por dos dominios constituidos por redes locales y unidos a través de una red pública de transmisión de datos. Dentro de cada red, local o pública, las interfaces de cada nodo están identificadas mediante una dirección física (que en el caso de la red pública puede ser un número de abonado) impuesta por el propio hardware de red y que normalmente el usuario no puede modificar.

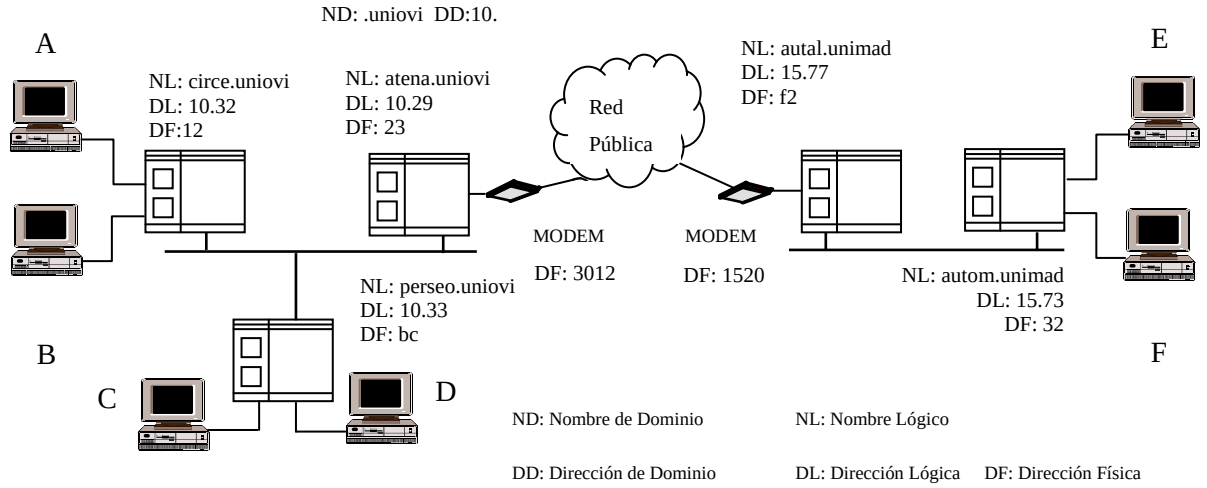


Fig. 11: Ejemplo de estructura de red

Por razones mnemónicas, a los nodos y dominios se les dan nombres que permitan recordar su denominación en la red fácilmente. Estos nombres están asociados a direcciones lógicas, que son las que realmente utiliza el sistema de comunicaciones para

identificar cada nodo y dominio. Por lo general, el nombre o dirección de un nodo se compone de la identificación del dominio donde se encuentra, junto con su identificación individual dentro de ese dominio. Las identificaciones lógicas son asignadas por los usuarios a los nodos, generalmente bajo la supervisión de un administrador de la red.

Cuando se transmite un mensaje, pasa de la capa 7 a la 1 del sistema emisor, y cada capa añade su propia cabecera o trata el mensaje de alguna forma. Las tramas que constituyen el mensaje se transmiten sobre el medio hasta el sistema receptor, en el que pasan de la capa 1 a la 7, eliminándose las cabeceras y reconstituyéndose el mensaje. Cuando las funciones de una capa en particular no son necesarias, se emplea una capa nula.

En el primer ejemplo, el mensaje va destinado a un nodo que se encuentra en la misma red física que el nodo emisor. Por ello, las funciones de encadenamiento entre entidades no son necesarias y la capa de red y la distinción entre direcciones lógicas y físicas pierden sentido al no ser necesario para realizar el encaminamiento.

El mensaje es adquirido por la capa de aplicación, que se implementaría como el software necesario para recoger el mensaje del teclado del usuario del terminal B de “circe” y enviarlo por la red. Una vez obtenido el mensaje, la aplicación lo entregaría al módulo o programa que implementa la capa de presentación, que adecuará el mensaje a la sintaxis de la red. En este caso se ha ejemplificado como una traducción a idioma de la red, que podríamos suponer que es el inglés. En la realidad, la capa de presentación adecua estructuras de datos, representaciones de datos enteros, de coma fija, de coma flotante, comprime, encripta, etc., a unas estructuras estándar para el sistema de comunicaciones.

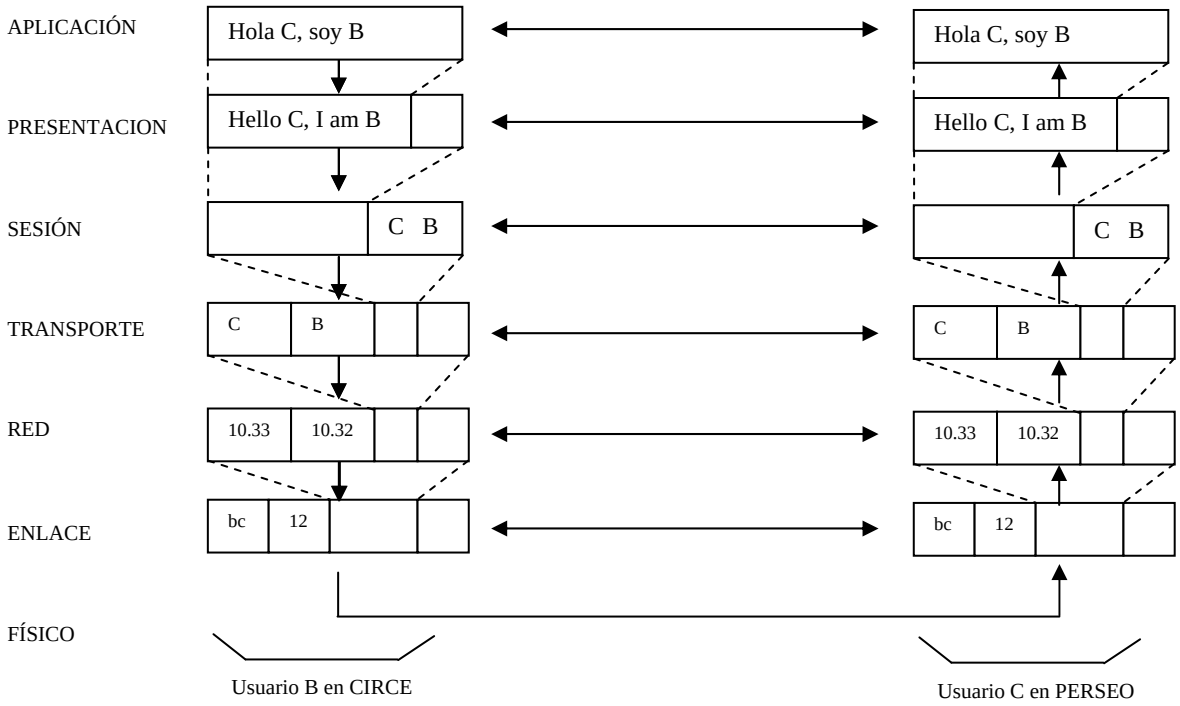


Fig. 12: Ejemplo de comunicación dentro de una subred

La capa de sesión mantiene la sesión de trabajo de cada usuario dentro de un mismo nodo, identificando a cada usuario para diferenciar su sesión de la de los demás. Todas estas sesiones convenientemente identificadas (generalmente mediante la identificación tanto del origen, B, como del destinatario, C) se multiplexan en la capa de transporte que transfiere a la capa de red los datos destinados a cada nodo (correspondientes a una o varias sesiones) dando su identificación lógica en la red (10.33 como destino y 10.32 como origen).

Cuando el nodo destinatario se encuentra en la misma red, esta capa simplemente entrega a la de enlace los datos a enviar con la identificación de la interfaz física (bc) que corresponde al destinatario. La trama de datos creada por la capa de enlace es convertida en señales eléctricas (en este caso) que se propagan por el medio de transmisión.

Una vez captadas las señales por la interfaz física del destinatario, se convierten de nuevo en una trama. La capa de enlace se encarga de determinar si está dirigida al

nodo en el que se encuentra mediante la comprobación de la dirección física que viene en la trama. Si es así, la acepta y la entrega para ser procesada por la capa de red, si no, la rechaza.

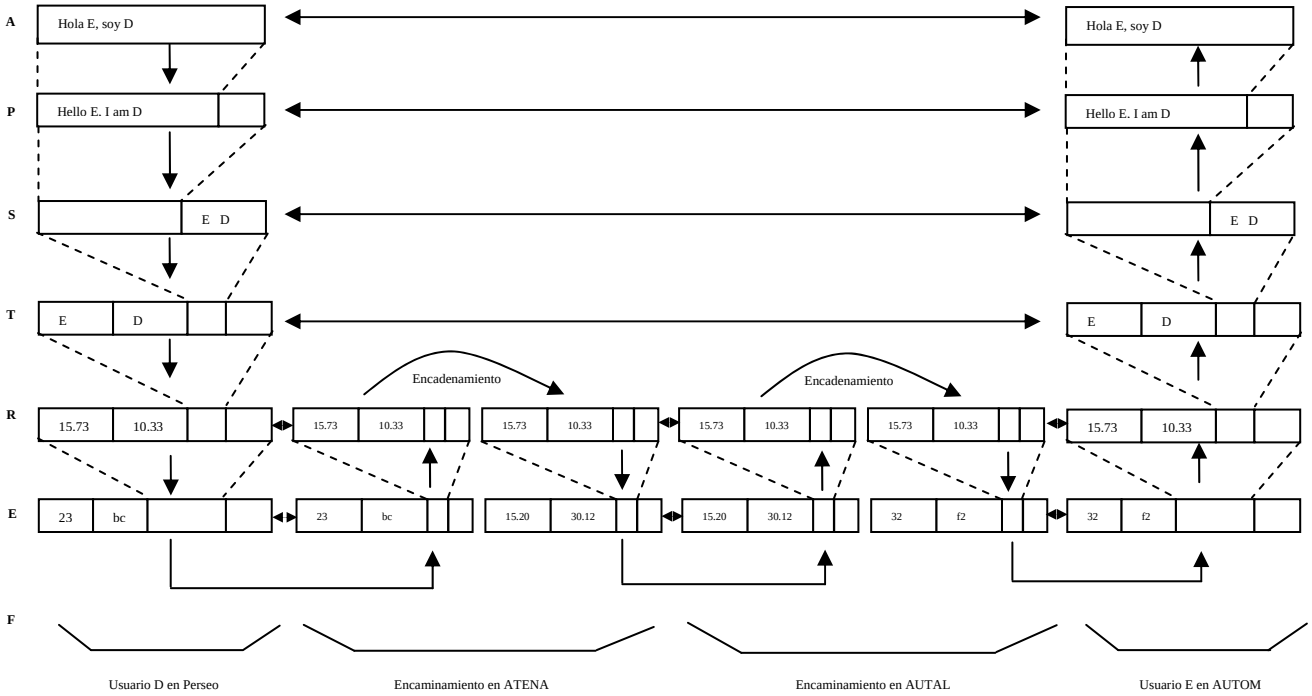


Fig. 13: Ejemplo de comunicación entre subredes

A: Aplicación **P**: Presentación **S**: Sesión

T: Transporte **R**: Red **E**: Enlace **F**: Físico

La capa de red comprueba la dirección lógica de destino, y si es la suya entrega los datos a la de transporte. Esta última identifica los datos que vienen para las distintas sesiones y los demultiplexa entre ellas (en este caso la sesión del usuario C). La capa de sesión elabora sus datos para el mantenimiento de la misma y pasa el mensaje aún en la forma de representación de la red, a la capa de presentación. Esta lo descomprime, desencripta y/o adecua su representación a la utilizada en el nodo destinatario (que no tiene por qué ser la misma que la del nodo de origen). Finalmente la aplicación

correspondiente hará aparecer el mensaje en la pantalla del terminal del usuario destinatario.

En el segundo ejemplo la transmisión se realiza entre dos nodos localizados en dominios diferentes, “perseo.uniovi” y “autom.unimad”. Esto obliga a la información a pasar por nodos intermedios en su camino entre el usuario D, origen de la transmisión, y el destinatario E.

En principio todo el proceso es igual al anterior hasta que la información llega a la capa de red, encargada precisamente del encaminamiento entre subredes. Esta capa se encuentra con el problema de que si entrega la información a la capa de enlace indicando como destinatario la dirección física de “autom.unimad” (32), nadie en su subred atenderá esa trama de datos. Sin embargo, sí conoce la dirección física en su red del nodo que le sirve de enlace con nodos de otros dominios, “atena.uniovi” (23) y a esa dirección física dirige la trama.

La trama es aceptada por la capa de enlace de “atena.uniovi” pues está dirigida a su dirección física. Pero cuando los datos llegan a la capa de red, éste detecta que la dirección lógica del destinatario no es la suya. Sin embargo, “atena.uniovi” está preparada para estas situaciones ya que se encarga del encaminamiento del tráfico que va y viene desde fuera de la subred local. Dispone de dos interfaces de comunicación con características y sintaxis de dirección diferentes, y de unas tablas de encaminamiento que le permiten saber, en función de la dirección lógica del destinatario, a qué red y a qué dirección física ha de dirigir la información. En este caso, decide pasar a la capa de red implementada para la red pública, los datos, y ésta los destina a través de la capa de enlace hacia la dirección física “1520” que corresponde al nodo que realiza funciones similares en la red “unimad”.

La información se transmite a través de la red pública con señales eléctricas de características muy distintas a las de la red local, y son aceptadas por la capa de enlace de “autal.unimad”. Su capa de red detecta también una dirección lógica de destino distinta a la suya para realizar a continuación un proceso similar al de “atena”. Ahora los datos pasan de nuevo a unas capas relacionadas con la red local “unimad” (que

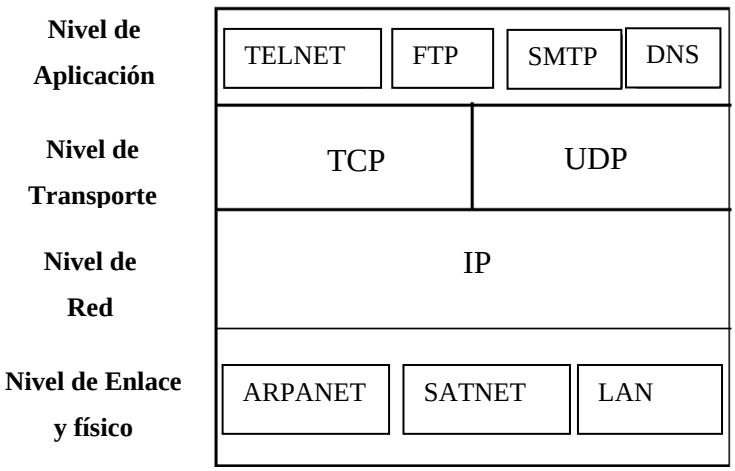
puede ser un estándar diferente a la red que se utiliza en “uniovi”) y son dirigidos, ahora sí, a la dirección física del destinatario, “autom.unimad” (32). El proceso hasta llegar a la pantalla del usuario del terminal E, es el ya descrito en el ejemplo anterior.

3.2. El modelo de referencia TCP/IP

Este modelo es el usado por ARPANET, el abuelo de las redes de ordenadores.

3.2.1. La capa Internet

Por diversas razones, en el caso de ARPANET, se eligió una red basada en conmutación de paquetes sobre un servicio de red sin conexión. Esta capa de red es la capa Internet. Su función es permitir que los host inserten paquetes en cualquier red, y que éstos viajen independientemente hacia su destino (que quizá sea una red distinta). Incluso pueden llegar en distinto orden del que fueron enviados, en cuyo caso, es obligación de las capas superiores reordenarlos si fuese preciso.



La capa internet define un tipo oficial de paquete y un protocolo llamado IP (internet protocol). La principal obligación de la capa es distribuir los paquetes hacia su destino, por ello, su función es el encaminamiento de los mensajes y evitar atascos,

aunque sus mecanismos de control de congestiones son bastantes limitados. Equivale a la capa de red del modelo OSI.

3.2.2. Capa de transporte

Es la siguiente capa en el modelo TCP/IP. Está diseñada para permitir el diálogo entre entidades homólogas extremo a extremo, al igual que la capa de transporte de modelo OSI. Utiliza dos protocolos: TCP (Transmision Control Protocol) y UDP (User Datagram Protocol). El primero es un protocolo orientado a conexión, libre de errores, que permite enviar bloques de bytes de una máquina a otra por un canal libre de errores. TCP también administra el control de flujo. El protocolo UDP es un protocolo sin conexión, basado en datagramas simples. Se pensó para aquellos casos en los que la capa de sesión necesitase un canal lógico distinto del que proporciona TCP.

3.2.3. Capa de aplicación

El modelo TCP/IP no tiene las capas de presentación ni de sesión. La experiencia ha demostrado que esta aproximación es la correcta. Esta capa contiene todos los protocolos de alto nivel como por ejemplo: TELNET (terminal remoto), FTP (transferencia de ficheros), SMTP (correo electrónico), DNS (servidor de nombres), etc. Más recientemente se le han añadido otros protocolos como NNTP (news) y HTTP.

La capa de enlace entre el host y la red no está definida en TCP/IP. En realidad sólo especifica que el host debe estar unido a la red a través de algún protocolo que permita el envío de paquetes IP.

3.3 Comparación de los modelos OSI y TCP/IP

El modelo OSI y el TCP/IP tienen muchas cosas en común. Ambos se basan en la idea de una pila de protocolos independientes. Además, la funcionalidad de las capas es bastante similar. Por ejemplo, en ambos modelos, las capas hasta la de transporte deben proporcionar un servicio de transporte extremo a extremo independiente de la red, a procesos que desean comunicarse. En ambos casos, las capas que están por encima de la

capa de transporte son usuarios de los servicios que ésta proporciona, orientados a la aplicación.

Aún así, también poseen muchas diferencias. El modelo OSI tiene tres conceptos básicos: servicios, interfaces y protocolos. Probablemente, la principal contribución del modelo OSI es hacer explícita la distinción entre estos conceptos. Cada capa realiza unos servicios para la capa superior. La definición de los servicios indica qué es lo que hace la capa, no cómo es el acceso de las capas superiores o cómo funcionan las mismas.

La interfaz de una capa indica cómo acceder a los servicios que ofrece, pero tampoco dice nada sobre cómo funciona interiormente. Finalmente, el protocolo de la capa es un problema exclusivo de la misma. Sólo debe ser capaz de asegurar que la capa proporciona correctamente sus servicios. Su modificación no debería afectar al software de las demás capas.

En su origen, el modelo TCP/IP no hizo esta distinción, aunque con el tiempo se ha adecuado a estos propuestos por el modelo OSI. Como consecuencia, los protocolos del modelo OSI están mejor escondidos que en el modelo TCP/IP. El modelo OSI se planteó antes de definir los protocolos de cada capa, por ello el modelo no se desvió en favor de ningún protocolo en particular. El principal inconveniente es que los diseñadores del modelo no tenían mucha experiencia y por ello no sabían muy bien en qué capa incluir cada servicio.

Por ejemplo, la capa de enlace estaba pensada para redes punto a punto. Cuando aparecieron las redes broadcast hubo que insertar una subcapa para acomodarlas. Cuando se comenzaron a diseñar sistemas basados en OSI con los protocolos que existían, se dieron cuenta que no encajaban con los servicios requeridos de la capa. Los miembros del comité ISO pensaban que cada país tendría una red, controlada por el gobierno y adecuada al modelo OSI. El problema es que las cosas no evolucionaron así.

Con TCP/IP sucedió lo inverso: primero se definieron los protocolos y el modelo resultó ser una descripción de los mismos. Evidentemente, los protocolos se ajustan al

modelo, pero el modelo no se ajusta a ningún otro conjunto de protocolos, por lo que no es útil para describir redes que no sean de tipo TCP/IP.

Otra diferencia está en el tipo de conexión. El modelo OSI soporta servicios sin conexión y orientados a conexión en la capa de red, pero la capa de transporte sólo acepta servicios orientados a conexión. El modelo TCP/IP sólo soporta servicio de datagramas en la capa de red, pero admite ambas formas de servicio en la capa de transporte, con lo que el usuario puede elegir. Esto es importante para aplicaciones basadas en un protocolo simple de pregunta / respuesta.

3.4. Crítica del modelo OSI

Cabe preguntarse la razón por la que un standard con interesantes aportaciones teóricas y capaz de describir cualquier red no se ha impuesto. El tiempo ha dejado claras cuatro razones:

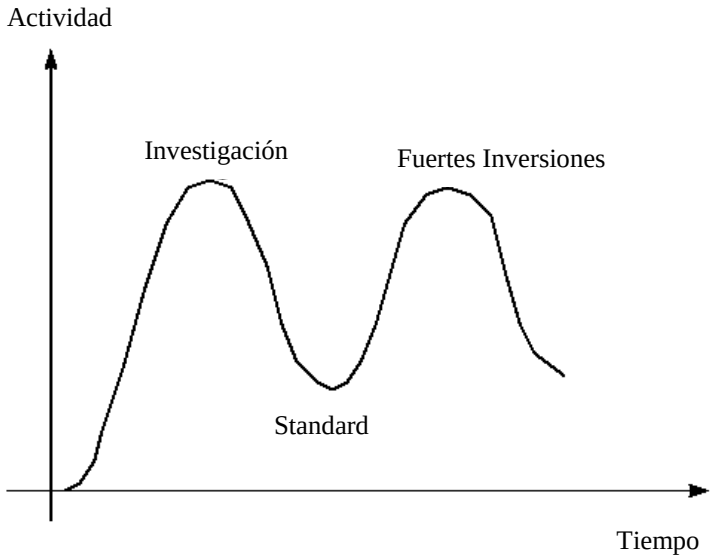
- Mala elección del momento.
- Mala tecnología.
- Malas implementaciones.
- Malas políticas.

Según David Clark, del MIT, para que la definición de un standard resulte exitosa, debe producirse lo que él llama ***apocalipsis de los dos elefantes***. La figura muestra la evolución en el tiempo de la actividad que provoca un nuevo tema. Al comienzo, existe una intensa actividad investigadora que se refleja en artículos, congresos y reuniones de grupos de trabajo. Después de un tiempo de mantenerse la actividad investigadora, las compañías descubren el tema e invierten fuertes sumas de dinero para lograr su aplicación comercial.

Es muy importante que las normas se escriban durante la parte intermedia, localizada entre los dos “elefantes”. Si éstas se escribiesen antes de culminar la fase investigadora, las normas resultantes podrían reflejar lagunas en el conocimiento del tema. Por contra, si se espera tanto como para que las compañías hayan efectuado

grandes inversiones, es posible que prefieran ignorar las recomendaciones del standard para no perder su posición en el mercado. Lo que pasó con la normalización propuesta por el modelo OSI, es que el intervalo entre los “elefantes” fue muy pequeño en comparación con el tiempo empleado en desarrollar la norma, por lo que ésta quedó colapsada entre ambos.

Por otra parte, aunque puede parecer evidente que el número de capas del modelo, así como su contenido, es la única alternativa disponible. Sin embargo, no está claro que esto sea así; de hecho, la propuesta británica era de cinco capas. Muchas aplicaciones no necesitan los servicios ofrecidos por las capas de sesión y presentación. Además la capa de presentación está prácticamente vacía de contenidos. Por contra, otras capas como la física o la de enlace, debieron ser subdivididas debido a la gran cantidad de funciones que debían soportar.



Teoría de los elefantes de David Clark

Aunque no es oficial, una posible razón para adoptar el modelo de siete capas fue que IBM disponía de una arquitectura de red de 7 capas (SNA). IBM dominaba de tal manera el mercado, que todos estaban convencidos de que hubiese usado su poder para imponer SNA frente al standard, pudiéndolo después modificar a su voluntad. Por ello, se pensó en hacer un modelo a la medida de SNA.

En el modelo OSI algunas funciones tales como direccionamiento, control de flujo y detección de errores están duplicadas en cada capa, lo que resulta redundante e ineficiente. Además, aunque muchas LAN trabajaban usando servicios y protocolos sin conexión, el standard original no incluía esta posibilidad, que fue añadida mediante extensiones de la norma. Aspectos importantes como los de administración fueron excluidos del modelo.

En cualquier caso, y dada la complejidad del modelo y de los protocolos, las primeras implementaciones resultaron excesivamente grandes, incontroladas y lentas. Se asoció OSI a poca calidad, y aunque los productos fueron mejorando, la idea no cambió.

Por el contrario una de las primeras implementaciones de TCP/IP era parte del Unix de Berkeley y su calidad resultó bastante alta. Por si esto fuera poco, era software de libre distribución. En estas condiciones, es fácil entender que su utilización se generalizase, lo que llevó a nuevas mejoras y de ahí a un número de usuarios aún mayor.

TCP/IP nació dentro los ambientes universitarios, mientras que se esperaba que OSI fuese un producto elaborado por los ministerios de telecomunicaciones europeos, la Comunidad Europea y el gobierno de E.E.U.U. Evidentemente, la idea de imponer desde la burocracia una tecnología inferior a la disponible no funcionó.

3.5. Crítica del modelo TCP/IP

El modelo y los protocolos de TCP/IP no es general y por tanto, describe mal cualquier otro conjunto de protocolos distintos de TCP/IP. Por ejemplo, describir la arquitectura SNA resultaría imposible.

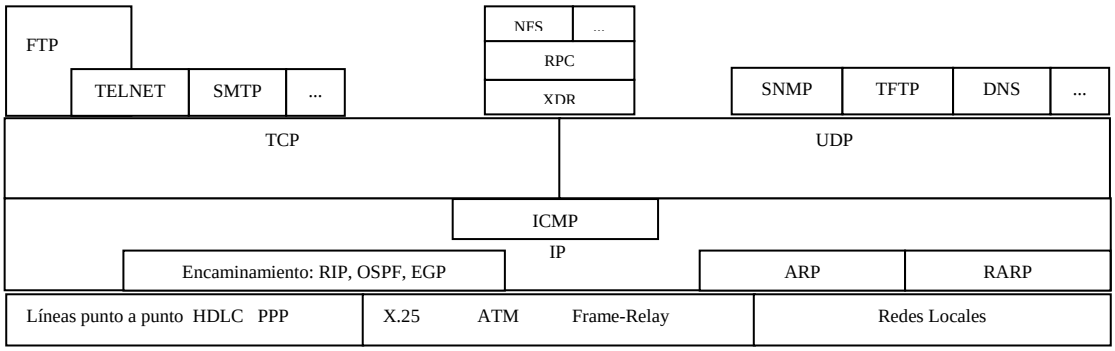
Por otra parte, la conexión a red (capas de enlace y física) no es una capa en el sentido normal del término. Es más bien, una interfaz, se indica que debe permitir el envío de tramas IP. Además, no se establece ninguna diferencia entre la capa física y la

de enlace. Aunque TCP/IP tiene unos protocolos bien pensados y bien implementados, muchos protocolos de la capa de aplicación se hicieron sobre la marcha, pero su rápida difusión popularizó su uso, con lo que resultan difíciles de sustituir.

4. EL PROTOCOLO TCP/IP

4.1. Una Familia de Protocolos

La denominación TCP/IP recoge la descripción de una serie de protocolos, la topología y la arquitectura que sirven de base para una red de área extensa (WAN) como es el caso de Internet. Entre los protocolos descritos bajo esa denominación se encuentran el IP (Internet Protocol) y el TCP (Transmision Control Protocol) junto con varios más. Todos ellos sirven de soporte a un conjunto de aplicaciones y servicios de aplicación, muy conocidos por su utilización en la red Internet. La descripción de todos los elementos que forman parte de la arquitectura TCP/IP y la mayor parte de las aplicaciones que hacen uso de ella, se encuentran recogidos como estándares "de facto" en los RFCs (Request For Comments). Se trata de documentos manejados por la comunidad de Internet donde se incluyen los protocolos y estándares de la red Internet, las propuestas de estándar, documentos puramente informativos, etc.



4.2. Direcciones IP y Encaminamiento mediante Routers

Para hacer un sistema de comunicación universal, se necesita un método de identificar computadores aceptado globalmente. Cada computador tendrá su propio identificador, conocido como dirección IP o dirección Internet.

Puede pensarse en la red Internet como cualquier otra red física. La diferencia está en que la red Internet es una estructura virtual implementada enteramente en "software". Por tanto, los diseñadores fueron libres de escoger los tamaños y formatos de los paquetes, las direcciones, las técnicas de distribución de paquetes, etcétera.

Para las direcciones, se escogió un sistema análogo al direccionamiento en redes físicas, en el cual a cada "host" se le asigna un número entero de 32 bits como identificador, llamado dirección internet. Estos enteros están cuidadosamente escogidos para hacer el proceso de encaminamiento o "routing" eficiente. Las direcciones internet codifican la identificación de la red a la que el "host" se encuentra conectado, así como la identificación de ese "host" dentro de la red. Por tanto, todos los computadores conectados a una misma red tienen en su número de dirección una serie de bits comunes (evidentemente, los bits de identificación de red).

Cada dirección internet es un par de identificadores (redid, hostid), donde ***redid*** identifica una red y ***hostid*** identifica a un computador dentro de esa red. En la práctica,

hay tres clases distintas de direcciones (clases A, B y C), como se muestra en las figuras.

Clase A

0	redid (7 bits)	hostid (24 bits)

Clase B

1	0	redid (14 bits)	hostid (16 bits)

Clase C

1	1	0	redid (21 bits)	hostid (8 bits)

Dada una dirección IP, se puede determinar su clase a partir de los tres bits de orden alto, siendo sólo necesario dos bits para distinguir entre las clases primarias. Las direcciones de clase A se usan para computadores en redes que tienen más de 2^{16} estaciones o “hosts” (esto es, 65.636), utilizando 7 bits para ***redid*** y 24 bits para ***hostid***. Las direcciones de clase B se usan para redes de tamaño intermedio, que tienen entre 2^8 (esto es, 256) y 2^{16} "hosts", localizando 14 bits en ***redid*** y 16 bits en ***hostid***. Finalmente, las redes de clase C, que tienen menos de 2^8 "hosts", utilizan 21 bits para ***redid*** y solamente 8 bits para ***hostid***.

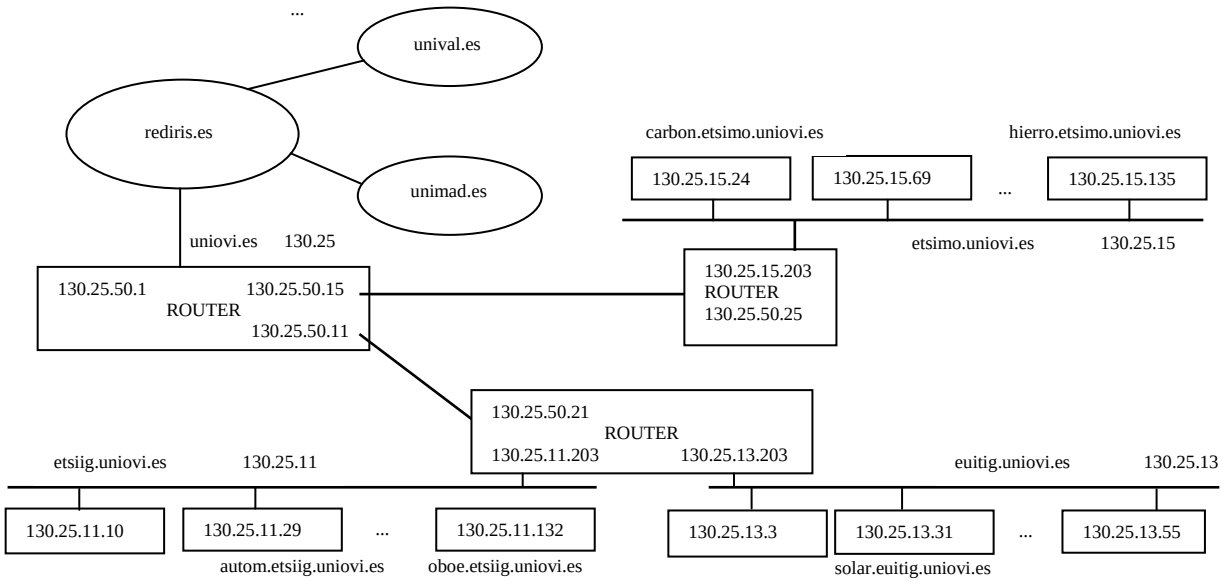
En cada red de clase A, B o C el administrador de la misma puede hacer uso de parte de los bits correspondiente al ***hostid*** para a su vez denominar distintas subredes dentro de su organización.

Los "routers" basan sus decisiones de encaminamiento en el ***redid*** del nodo destino, es decir en la red a la que van destinados los datos. Los "routers", por tanto, deciden dónde van a mandar los datos basándose en la red destino y no en el computador específico al que van destinados los datos, lo que disminuye las necesidades de memoria de los "routers" a medida que aumenta el número de estaciones conectadas a la red Internet. Sin embargo, siguiendo el criterio explicado hasta ahora no sería posible, por ejemplo, dar una dirección a un "router" que esté unido a dos redes,

puesto que el ***redid*** de las dos redes no es el mismo. La solución a esto es asignar a este tipo de máquinas conectadas a más de una red, varias direcciones, una por cada red a que estén conectadas.

Para que al usuario le resulte más cómodo recordar la identificación de los distintos "hosts" y dominios, se les dan nombres o alias. En algunos casos un nombre hace referencia a más de una dirección IP (generalmente varias direcciones de dominios) y en otras varios nombres (alias) hacen referencia a una misma dirección IP de "host" o de dominio. Esto obliga a que exista una base de datos en el "host" que relacione nombres lógicos con direcciones IP.

Generalmente, esta base de datos es un fichero de tipo texto que sólo contiene unos pocos de los nombres existentes dentro de la red Internet. Sería inviable que cada "host" tuviese una base de datos completa y actualizada con todos los nombres y direcciones IP. Si un "host" no tiene en su propia base de datos la identificación de otro "host", consulta a un servidor DNS (***Domain Name Server***). Se trata de un "host" que mantiene la base de datos para uno o varios dominios y que da servicio de nombres a los ordenadores de ese dominio, que a su vez deben conocer cuál o cuáles son los servidores DNS que tienen más cercanos. Si el servidor DNS no contiene la identificación del "host" que le han solicitado, consulta a su vez a otros servidores DNS de la red Internet hasta encontrarlo.



4.2.1. Direcciones de red y broadcast

Se ha dicho que una de las ventajas de las direcciones IP es que codifican información sobre la red, con lo que se simplifica el encaminamiento. Otra ventaja es que una dirección IP puede referirse tanto a redes como a "hosts". Por convenio, *hostid* 0 no se asigna nunca a un computador, si no que una dirección IP con *hostid* igual a 0, se usa para referirse a la propia red.

Otra ventaja importante del direccionamiento ***IP*** es que soporta la dirección *broadcast*, que se refiere a todos los "hosts" conectados a la red. De acuerdo con el convenio, cualquier *hostid* con todos los bits valiendo 1 se reserva para *broadcast*. En muchas tecnologías de red (por ejemplo en la red Ethernet), la transmisión *broadcast* es tan eficiente como una transmisión normal; en otras redes se admiten las direcciones *broadcast*, aunque suponen un retraso considerable; otras redes no las admiten en absoluto. Por tanto, la existencia de direcciones *broadcast* no garantiza la eficacia de este tipo de transmisión.

De igual manera que un campo en la dirección IP con todos los bits 1 significaba "todos", el software IP interpreta un campo con todos los bits valiendo 0 como "éste". Así, una dirección con *hostid* igual a 0, se refiere a "ese" computador, y una dirección con *redid* igual a 0 se refiere a "esa" red. El uso de direcciones con *redid* igual a 0 es especialmente importante en aquellos casos en que un "host" quiere comunicarse sobre una red, pero todavía no sabe su dirección . El "host" utiliza temporalmente un *redid* igual a 0, y los otros computadores conectados a la red, interpretan esa dirección significando "esta" red. En la mayoría de los casos, las respuestas tendrán una dirección con el identificador de red especificado, permitiendo al "host" almacenarlo para futuros usos.

4.2.2. Notación decimal con puntos

Las direcciones IP se representan como cuatro enteros decimales separados por puntos, donde cada entero da el valor de un octeto de la dirección. Así, por ejemplo, la dirección de 32 bits 10000000 00001010 00000010 00011110 se escribe 128.10.2.30

Todas las direcciones IP en Internet son asignadas por una autoridad central, el *Centro de Información de Redes* (NIC,*Network Information Center*), localizado en los Estados Unidos de América. Esta autoridad central sólo asigna la parte de la dirección que identifica a la red, delegando la asignación de las direcciones de los "hosts" a la organización que ha formulado la petición de conexión a la red Internet.

4.2.3. Orden de Byte en la red

Estandarizar el orden de byte en la representación de enteros es especialmente importante, pues los paquetes TCP/IP llevan números que especifican información tal como la dirección destino, o la longitud del paquete. Estos valores deben ser perfectamente entendidos tanto por el que envía como por el que recibe. El TCP/IP resuelve el problema definiendo un orden de byte estándar para la red, que debe ser usado por todas las máquinas en los campos binarios de los paquetes. Cada "host" debe convertir los datos de su propia representación interna al orden de byte de la red antes de enviar un paquete. Lo mismo tiene que hacer el destinatario cuando reciba un paquete. Naturalmente, el campo de datos en el paquete está exento de este estándar; cada usuario es libre de elegir el formato de datos.

El estándar TCP/IP especifica que los enteros son enviados con el byte más significativo delante; es decir, si se consideran sucesivos bytes en un paquete cuando va de una máquina a otra, los enteros tienen el byte alto más cerca del principio del paquete y el byte bajo más cerca del final.

4.3. El Protocolo ARP

Se ha dicho anteriormente que una dirección IP era un número de 32 bits que se asignaba a las máquinas conectadas a la red para su identificación, y que esa identificación era suficiente para enviar y recibir paquetes. Sin embargo, las máquinas conectadas a una red física pueden comunicarse sólo si conocen sus respectivas direcciones físicas. Por tanto, cuando un "host" quiere comunicarse con otro, necesita

"mapear" la dirección IP del destinatario en su correspondiente dirección física, si éste se encuentra en la misma red física. Si el destinatario no está en la misma red física, deberá "mapear" la dirección del encaminador o "router" que le permita enviar los datagramas IP fuera de la subred en la que se encuentra. De esto se encarga el protocolo ARP.

La idea del ARP(*Address Resolution Protocol*) es sencilla. Cuando un "host" A desea comunicarse con otro B, del que conoce su dirección IP (I_B), pero no su dirección física (F_B), envía un paquete especial con dirección destino *broadcast*, pidiendo al "host" que tiene como dirección IP, I_B, que responda con su dirección física, F_B. Todos los "hosts" conectados a la red reciben el paquete ARP, pero sólo el "host" B que es, al que iba dirigida la pregunta, responde con otro paquete ARP, enviando su dirección física.

Así, una vez completado el intercambio de información mediante el protocolo ARP, el "host" conoce la dirección física del otro con el que quiere comunicarse, de manera que puede enviarle sucesivos paquetes a él directamente. La experiencia demuestra que merece la pena mantener una tabla dinámica en la memoria volátil con las direcciones IP y las correspondientes direcciones físicas de los computadores con los que se ha establecido comunicación más recientemente, ya que, normalmente, la comunicación requiere el envío de varios paquetes. Así, cuando un "host" quiere comunicarse con otro, antes de enviar un paquete ARP, busca en la memoria para ver si tiene su dirección física, con lo que se reducen costes de comunicación. Sin embargo, las entradas de la tabla se eliminan si no son utilizadas durante un cierto tiempo o cuando el computador se apaga. Se evitan así problemas de comunicación si alguno de los ordenadores registrados ha cambiado de dirección física por avería de su interfase de comunicaciones o cualquier otra eventualidad.

Cuando un paquete ARP viaja por la red de una máquina a otra, lo hace encapsulado en una trama del nivel de enlace, como en el ejemplo de la figura.

Preámbulo	Destino	Origen	Tipo	Mensaje ARP tratado como datos	CRC

Mensaje ARP encapsulado en una trama Ethernet

Para identificar que la trama está llevando un paquete ARP, el computador fuente, asigna un valor especial al campo de tipo en la cabecera del paquete. Cuando la trama llega al destino, el "host" examina el campo tipo para determinar qué contiene esa trama. En el caso de la red Ethernet los paquetes ARP tienen un campo de tipo de valor 0806h (valor en notación hexadecimal).

Al contrario que la mayoría de los protocolos, los datos en paquetes ARP no tienen un formato de encabezamiento fijo. El mensaje está diseñado para ser válido con una variedad de tecnologías de transmisión y de protocolos. El ejemplo de la figura muestra el mensaje ARP de 28 octetos usado para las redes Ethernet (en las que la dirección física tiene una longitud de 48 bits, 6 octetos) y protocolo de red IP (con dirección lógica de 32bits, 4 octetos).

<table>
<tr><td colspan="2">HARDWARE</td><td>PROTOCOLO</td></tr>
<tr><td>HLON</td><td>PLON</td><td>OPERACIÓN</td></tr>
<tr><td colspan="3">DF ORIGEN (octetos 0-3)</td></tr>
<tr><td colspan="2">DF ORIGEN (octetos 4-5)</td><td>DL ORIGEN (octetos 0-1)</td></tr>
<tr><td colspan="2">DL ORIGEN (octetos 2-3)</td><td>DF DESTINO (octetos 0-1)</td></tr>
<tr><td colspan="3">DF DESTINO (octetos 2-5)</td></tr>
<tr><td colspan="3">DL DESTINO (octetos 0-4)</td></tr>
</table>

Formato del paquete ARP usado para redes Ethernet

A continuación se explica el significado de cada uno de los campos del paquete:

- HARDWARE: Especifica el tipo de interfaz hardware para el que el computador fuente solicita la respuesta; el valor es 1 para la red Ethernet.
- PROTOCOLO: Contiene el número del protocolo al que corresponden las direcciones lógicas.
- HLON y PLON: Especifican, respectivamente, las longitudes de la dirección física y de la dirección de protocolo.
- OPERACION: Especifica el tipo de operación que realiza el mensaje: vale 1 para una petición ARP, 2 para una respuesta ARP, 3 para una petición RARP y 4 para una respuesta RARP. (El protocolo RARP se explicará a continuación).
- DF ORIGEN: Contiene la dirección física del "host" que realiza la petición ARP.
- DL ORIGEN: Contiene la dirección lógica (o de protocolo) del "host" que realiza la petición ARP.
- DF DESTINO: Es un campo vacío en una petición ARP y contiene la dirección física del "host" al que va destinada la petición en la respuesta.
- DL DESTINO: Contiene la dirección lógica del "host" al que va destinado la petición ARP.

4.4. El Protocolo RARP

Las máquinas sin acceso a un tipo de almacenamiento secundario (por ejemplo, máquinas sin disco) conectadas a una red, no tienen posibilidad de guardar su dirección IP cuando se apagan. Por tanto, cuando estas máquinas se arrancan, tienen que usar la red para contactar con un servidor que les indique su dirección IP. El protocolo que usan las máquinas sin disco para obtener su dirección es el RARP (*Reverse Address Resolution Protocol*). Este protocolo está adaptado del ARP y usa su mismo formato de paquete. Al igual que el mensaje ARP, el mensaje RARP se transmite de una máquina a las otras encapsulado en una trama física. En el caso de redes Ethernet el campo de tipo de trama correspondiente a éste protocolo es el 0835h.

El uso del protocolo RARP es similar al del ARP: el "host" que quiere conocer su dirección IP envía un mensaje de petición RARP *broadcast*, con su dirección física en

el campo de dirección física de destino (DF DESTINO). Todas las máquinas conectadas a la red reciben el mensaje, pero sólo aquellas autorizadas procesan la petición y contestan. Estas máquinas se conocen con el nombre de servidores RARP. Los servidores contestan rellenando el campo de dirección de lógica de destino (DL DESTINO), y cambiando el tipo de operación a "respuesta RARP".

Hay que tener en cuenta que la comunicación entre "hosts" buscando su dirección IP y servidores RARP tiene que hacerse usando solamente la red física a que están conectados ambos, pues la única identificación que tienen los "hosts" es su dirección física para esa red concreta.

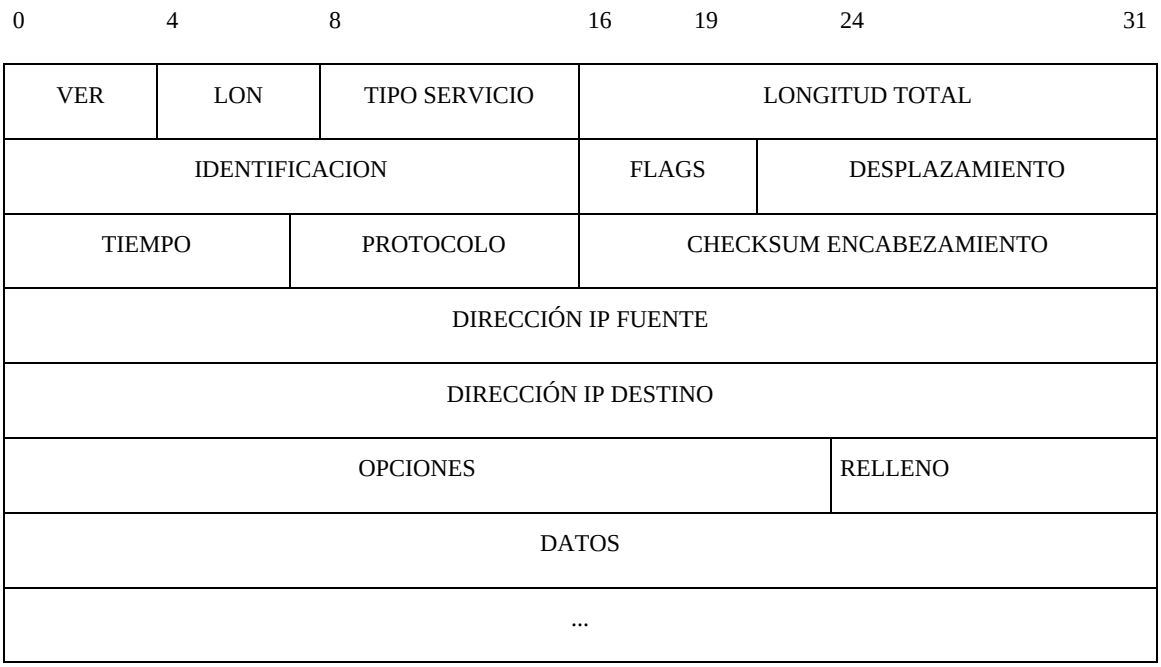
4.5. El Protocolo IP

El protocolo IP (*Internet Protocol*) define la unidad básica de transmisión de datos, y el formato exacto de todos los datos cuando viajan por una red TCP/IP. Además, el protocolo IP incluye una serie de reglas que especifican cómo procesar los paquetes y cómo manejar los errores. El protocolo IP se basa en la idea de que los datos se transmiten con un mecanismo no fiable y sin conexión. El decir un mecanismo no fiable, se refiere a que un paquete puede perderse, duplicarse o enviarse a otro destino del deseado. El mecanismo es sin conexión porque cada paquete se trata independientemente de los otros. Paquetes de una secuencia enviados de una máquina a otra pueden ir por distintos caminos, a la vez que unos pueden alcanzar su destino mientras que otros no. El protocolo incluye también la idea del encaminamiento (o *routing*) de paquetes.

4.5.1. El datagrama IP

El datagrama es la unidad básica de transmisión de datos en la red Internet. El datagrama, al igual que las tramas en las redes físicas, se divide en encabezamiento y campo de datos. El encabezamiento contiene las direcciones IP de la fuente y el destino.

La longitud máxima de un datagrama es de 65.536 octetos (64 Kbytes). Sin embargo, para viajar de una máquina a otra, los datagramas lo hacen en el campo de datos de tramas de enlace, por lo tanto los datagramas de longitud excesiva deben ser divididos en fragmentos que "quepan" en las tramas. Cada fragmento perteneciente a un mismo datagrama tiene el mismo número de identificación que el datagrama original, con lo que es posible su reconstrucción.



Preámbulo	Destino	Origen	Tipo	Datagrama IP (o fragmento) tratado como datos	CRC

Datagrama IP encapsulado en una trama Ethernet

- VER: Este campo de 4 bits se usa para especificar la versión del protocolo IP, y se usa para que destino, fuente y "routers" entre ellos, están de acuerdo en el tipo de datagrama. Si los estándares cambian, las máquinas rechazarán datagramas con versión de protocolo distinta de las de éllas.
- LON: Este campo de 4 bits especifica la longitud del encabezamiento del datagrama medido en palabras de 32 bits. El encabezamiento más común, sin incluir opciones, tiene este campo con valor igual a 5.

- TIPO DE SERVICIO: Este campo de 8 bits especifica cómo debe ser manejado el datagrama, y se divide en los 5 campos que muestra la figura:

PRIORIDAD (3 bits)	D (1 bit)	T (1 bit)	R (1 bit)	NO USADO

PRIORIDAD: Estos tres bits especifican la prioridad del datagrama, con valores en el rango de 0 (prioridad normal) a 7 (datagramas de control de la red). Bits D, T y R especifican el tipo de servicio que el datagrama solicita. Cuando están a 1, el bit D solicita bajo retraso, el bit T alta velocidad de transmisión de la información y el bit R solicita alta fiabilidad. Por supuesto, la red no puede garantizar el tipo de servicio requerido si, por ejemplo, el camino al destino no tiene esas propiedades. Estos bits se usan por los "routers" cuando pueden escoger entre varios posibles caminos para un datagrama; eligiendo el que mejor se adapte a los servicios solicitados.

- LONGITUD TOTAL: Este campo da la longitud total del fragmento del datagrama, incluido el encabezamiento, medida en octetos. El tamaño del campo de datos puede calcularse a partir de este campo y del campo LON.
- IDENTIFICACIÓN: Este campo, junto con los de FLAGS y DESPLAZAMIENTO se utilizan para el control de la fragmentación. Su propósito es permitir al "host" destino unir todos los fragmentos que pertenecen a un mismo datagrama, y que le pueden llegar fuera de orden. A medida que los fragmentos llegan, el "host" destino utiliza el campo IDENTIFICACION, junto con la dirección fuente para identificar a qué datagrama pertenece ese fragmento. Los computadores normalmente generan un campo IDENTIFICACION único, incrementando un contador cada vez que forman un datagrama.
- FLAGS: Los dos bits más bajos de los tres bits del campo FLAGS controlan la fragmentación. El primero especifica si el datagrama puede ser fragmentado (si está a 1, no puede serlo). El bit más bajo de FLAGS indica, si está a 1, que este fragmento no es el último del datagrama. Este bit es necesario pues el campo

LONGITUD TOTAL del encabezamiento se refiere a la longitud del fragmento, y no a la longitud total del datagrama.

- DESPLAZAMIENTO: Especifica el desplazamiento del fragmento en el datagrama original, medido en unidades de 8 octetos, empezando con desplazamiento 0. Para ensamblar un datagrama, el "host" destino debe obtener todos los fragmentos; desde el de desplazamiento cero hasta el de desplazamiento más alto. Si uno o más fragmentos se pierden, el datagrama entero debe ser descartado.
- TIEMPO: Especifica el tiempo máximo que se le permite al datagrama permanecer en la red IP; así se evita que datagramas perdidos viajen por la red indefinidamente. Es difícil estimar el tiempo exacto porque los "routers" normalmente no saben el tiempo que requieren, para la transmisión, las redes físicas. Para simplificar, "host" y "routers" suponen que cada red utiliza una unidad de tiempo en la transmisión; así deben decrementar el valor de este campo en uno cada vez que procesen un encabezamiento de datagrama. Si el valor del campo llega a cero el datagrama se destruye y se devuelve un mensaje de error ICMP.
- PROTOCOLO: Especifica el tipo de protocolo de alto nivel que soporta los datos que lleva el datagrama. Los valores de este campo para distintos protocolos los asigna una autoridad central (el NIC). Algunos de ellos aparecen en la siguiente tabla:

Protocolo	**Valor**
ICMP	1
TCP	6
EGP	8
UDP	17

- CHECKSUM DEL ENCABEZAMIENTO: Asegura que el encabezamiento no tiene errores. El "checksum" se forma tratando el encabezamiento como una secuencia de enteros de 16 bits. Se suma con aritmética de complemento a uno, el complemento a uno de todos ellos. A efectos de calcular el "checksum" se

supone que el campo CHECKSUM DEL ENCABEZAMIENTO tiene valor cero. Como sólo se chequean errores en el encabezamiento, los protocolos de nivel superior deberán añadir otro tipo de comprobación para detectar errores en los datos.

- DIRECCINES IP FUENTE Y DESTINO: Contienen las direcciones internet de 32 bits de los "hosts" fuente y destino del datagrama respectivamente.
- OPCIONES: No es un campo necesario en todos los datagramas. Se incluyen normalmente para chequear o depurar la red. La longitud de las opciones varía dependiendo de cuáles de éstas se seleccionan. Algunas opciones son de longitud un octeto, mientras que otras son de longitud variable. Cada opción consiste en un octeto de código de opción, un octeto de longitud y una serie de octetos para la opción. El código de opción se divide en tres campos, como aparece en la figura.

COPIA (1 bit)	CLASE DE OPCION (2 bits)	NUMERO DE OPCION (5 bits)

Cuando el bit COPIA está a 1, especifica que la opción sólo debe ser copiada al primer fragmento, y no a los demás. Los campos CLASE DE OPCION y NUMERO DE OPCION, especifican la clase general de la opción y dan la opción específica dentro de esa clase. En la tabla siguiente se muestra la asignación de las clases.

Clase de opción	**Significado**
0	Control de datagrama o red
1	Reservado para futuro uso
2	Depuración y medida
3	Reservado para futuro uso

La tabla siguiente muestra las posibles opciones que pueden acompañar a un datagrama y da su clase y número de opción. La mayoría de ellas se usan para propósitos de control.

Clase de opción	Número de opción	Longitud	Descripción
0	0	1	Fin de lista de opción. Usado si las opciones no acaban al final del datagrama.
0	1	1	No operación
0	2	11	Restricciones de seguridad y manejo
0	3	variable	Encaminamiento fuente impreciso. Usado para encaminar un datagrama a lo largo de un camino fijado.
0	7	variable	Grabar ruta. Usado para localizar el camino seguido.
0	8	4	Secuencia identificadora. Usado para llevar una secuencia SATNET identificadora
0	9	variable	Encaminamiento fuente estricto. Usado para encaminar un datagrama por una vía determinada.
2	4	variable	Tiempos internet. Usado para grabar tiempos durante la ruta.

- RELLENO: Representa octetos conteniendo ceros, que son necesarios para que el encabezamiento del datagrama sea un múltiplo exacto de 32, ya se había visto que el campo de longitud del encabezamiento se especificaba en unidades de palabras de 32 bits.
- DATOS: Es la zona de datos del datagrama.

4.6. La nueva versión del Protocolo IP: IPv6

En un futuro próximo, la actual versión del protocolo IP (la versión cuatro, IPv4) será sustituida por una nueva versión, la seis, con el denominado IPv6 o IPng (IP new generation). El principal motivo es la ampliación del campo de direcciones IP que pasará ahora de 32 a 128 bits. Con 2^{32} direcciones, es decir, aproximadamente 4000 millones, debería ser suficiente, pero la ineficaz distribución de direcciones en subredes hacen que se desaprovechen la mayor parte de ellas.

Las mejoras del IPv6 incluyen los siguientes aspectos:

- Espacio de direcciones ampliado: 128 bits.
- Mecanismo de opciones mejorado: Las opciones van en cabeceras opcionales separadas a partir de la principal. Hay hasta ocho tipos diferentes que sólo se incluyen y/o procesan cuando son necesarias.
- Direcciones de autoconfiguración: Permiten la asignación dinámica de direcciones.
- Mayor flexibilidad en el direccionamiento.
- Facilidad de asignación de recursos al tráfico de alta prioridad: Desaparece el campo de TIPO DE SERVICIO y se incluye uno de PRIORIDAD que clasifica el tráfico en función de sus necesidades especiales como el vídeo en tiempo real.
- Capacidades de seguridad: Incluyen la autentificación y la privacidad de los datos.

La cabecera del IPv6 pasa a tener 40 en lugar de 60 bytes, sin embargo es más simple y el número de campos que incluye es menor.

4.7. Protocolo ICMP: Mensajes de Error y Control

Se ha visto que el protocolo IP proporciona un servicio no fiable y sin conexión; y que los mensajes viajan de "router" en "router" hasta alcanzar el nodo destino. El sistema funciona bien si todas las máquinas trabajan adecuadamente y los "routers" están de acuerdo en los encaminamientos. En caso contrario, ocurren errores. Para permitir a las máquinas en la red Internet informar sobre errores o circunstancias inesperadas, está el protocolo ICMP (*Internet Control Message Protocol)*, que es considerado como una parte del protocolo IP.

Los mensajes ICMP viajan en la porción de datos de los datagramas IP, como se muestra en la figura:

Encabezamiento IP	Mensaje ICMP tratado como datos

Los datagramas que llevan mensajes ICMP, se encaminan como los demás, por lo que pueden producirse errores. El protocolo dice que, en este caso, se produce una excepción, y especifica que no se deben generar mensajes ICMP sobre errores resultantes de datagramas llevando mensajes ICMP. Los mensajes ICMP facilitan también el control de congestión.

Aunque cada tipo de mensaje ICMP tiene su propio formato, todos empiezan con los mismos tres campos: Un entero de 8 bits indicando TIPO. Un campo de 8 bits, (CODIGO) dando más información sobre el tipo de mensaje y un campo de 16 bits con el "checksum" (se usa el mismo algoritmo que para los datagramas IP, pero incluye sólo el mensaje ICMP). Además, los mensajes ICMP incluyen el encabezamiento del datagrama IP que causó el problema, así como los primeros 64 bits de datos, para ayudar a determinar qué protocolo y qué programa de aplicación causaron el problema. El campo TIPO define el tipo del mensaje ICMP y el formato del resto del paquete. Los tipos son:

Campo TIPO	**Tipo de mensaje ICMP**
0	Respuesta de eco
3	Destino inalcanzable
4	Disminución de flujo de la fuente
5	Redireccionar (cambiar la ruta)
8	Petición de eco
11	Tiempo excedido por el datagrama
12	Problema de parámetro en un datagrama
13	Petición de grabar tiempos
14	Respuesta de grabar tiempos
15	Petición de información
16	Respuesta de información
17	Petición de máscara de direcciones
18	Respuesta de máscara de direcciones

4.8. El Protocolo UDP

Cuando un datagrama llega a su destino se ha de determinar a cuál de las aplicaciones existentes en la máquina, se ha de hacer llegar la información contenida en él. Esto no lo hace el protocolo IP directamente, sino que es misión del Protocolo de Transporte. El protocolo UDP proporciona la funcionalidad necesaria para identificar al destinatario final de un datagrama de una manera simple. Como no proporciona ningún mecanismo para el acuse de recibo, secuenciación de mensajes, ni control de flujo, proporciona el mismo servicio no fiable que el protocolo IP. Los mensajes UDP pueden perderse o llegar fuera de secuencia, y además, los paquetes pueden llegar más rápido de lo que es receptor es capaz de procesar.

El protocolo UDP permite a varios programas de aplicación en una misma máquina comunicarse simultáneamente y demultiplexa el tráfico que se recibe entre las distintas aplicaciones. El protocolo UDP incorpora los denominados puertos, que identifican el último destino dentro de una máquina, el programa de aplicación que está haciendo uso de ese puerto. Cada puerto tiene asociado un entero para identificarlo. Dado que el número de puerto es único dentro de una máquina, el destino último queda perfectamente determinado por la dirección internet del "host" y el número de puerto UDP en ese "host".

4.8.1. Formato del mensaje UDP

Los mensajes UDP se denominan *datagramas de usuario* y viajan en la porción de datos de los datagramas IP, como se muestra en la figura:

Encabezamiento IP	Mensaje UDP tratado como datos

El protocolo UDP entrega al servicio IP el segmento que contiene los datos, que es el que se transmite realmente, y una *pseudo-cabecera* que no se trasmite, pero que permite al protocolo IP completar los datos del o los datagramas que va a generar. El

formato del segmento y de la pseudo-cabecera son los que se muestran en la figura siguiente.

0 16 31

PUERTO FUENTE	PUERTO DESTINO
LONGITUD	CHECKSUM
DATOS	
...	

DIRECCION IP FUENTE		
DIRECCION IP DESTINO		
CERO	PROTOCOLO	LONGITUD UDP

A continuación se describe el significado de los distintos campos del segmento.

- PUERTOS FUENTE Y DESTINO: Estos dos campos contienen los números de los puertos UDP que identifican los programas de aplicación en las máquinas fuente y destino.
- LONGITUD: Es el número total de octetos que forman el datagrama UDP incluida la cabecera.
- CHECKSUM: Para calcular el "checksum", la máquina fuente antepone la pseudo-cabecera al datagrama y añade al final de los datos bytes, conteniendo ceros hasta conseguir una longitud del segmento múltiplo de 16 bits. El "checksum" se calcula una vez hechos los cambios según el siguiente algoritmo: se considera el segmento formado por enteros de 16 bits y se suman todos los complementos a uno de esos enteros utilizando la aritmética de complemento a uno. A efectos de hacer los cálculos se supone que ese campo tiene valor cero. Los ceros añadidos para rellenar, así como la pseudo-cabecera no se cuentan en la longitud del datagrama y no son transmitidos. El “checksum” es opcional. Si

no se utiliza, su contenido es cero. La razón de usar la pseudo-encabecera es permitir a la máquina destino comprobar que el datagrama ha alcanzado su destino correcto, ya que incluye la dirección internet del "host" destino, así como el número de puerto UDP de la conexión. La máquina destino puede conseguir la información usada en la pseudo-cabecera a partir del datagrama IP que transporta al datagrama UDP.

- DATOS: Representa los datos del datagrama UDP.

El significado de los campos de la pseudo-cabecera es el siguiente:

- DIRECCION IP FUENTE: Es la dirección internet del "host" que envía el segmento.
- DIRECCION IP DESTINO: Es la dirección internet del "host" al que va dirigido el segmento.
- CERO: Campo que contiene el valor cero.
- PROTOCOLO: Especifica protocolo UDP (es decir, 17).
- LONGITUD UDP: Este campo especifica la longitud total del datagrama UDP.

4.8.2. Números reservados para puertos UDP

El protocolo UDP combina una adjudicación de números de puertos UDP dinámica y estática, usando una serie de asignaciones de puertos conocida para un conjunto de programas que se utilizan comúnmente (por ejemplo, correo electrónico). Sin embargo, la mayoría de los números de puerto están disponibles para que el sistema operativo los utilice a medida que los programas de aplicación los necesiten.

4.9. El Protocolo TCP

Al más bajo nivel, la comunicación de computadores proporciona un servicio de distribución no fiable de paquetes. Cuando la red física falla, los paquetes pueden perderse, llegar con errores o duplicados.

Al nivel más alto, los programas de aplicación necesitan, frecuentemente, enviar grandes volúmenes de datos. Utilizar el sistema de distribución no fiable anteriormente mencionado, requiere que los programadores construyan algoritmos de detección y recuperación de errores en los programas de aplicación. Estos algoritmos son muy complejos, y por tanto pocos programadores tienen los conocimientos necesarios para diseñarlos. Como consecuencia, se desarrolló el protocolo de transporte TCP, que permite la transmisión fiable de datos mediante el establecimiento y liberación de conexiones, de manera que los datos que llegan a los programas de aplicación lo hagan sin errores, pudiendo ser utilizados directamente, sin necesidad de escribir algoritmos de detección y recuperación de errores.

Para conseguir esto, la mayoría de los protocolos utilizan una técnica conocida como "acuse de recibo positivo con retransmisión". Esta técnica requiere que la máquina destino envíe a la fuente un mensaje de acuse de recibo cada vez que recibe datos. La fuente mantiene un registro de paquetes enviados, y espera por un acuse de recibo antes de enviar el siguiente paquete. Asimismo, la fuente inicia un temporizador cada vez que envía un paquete y retransmite el paquete si el temporizador expira antes de recibir el acuse de recibo.

El problema final surge cuando la red física duplica paquetes (por ejemplo, si la red tiene un gran retraso pueden ocurrir retransmisiones prematuras). Para solucionar esto, los mensajes de acuse de recibo tienen números de secuencia, con lo que la fuente puede asociar correctamente acuses de recibo con paquetes.

4.9.1. La ventana deslizante del protocolo TCP

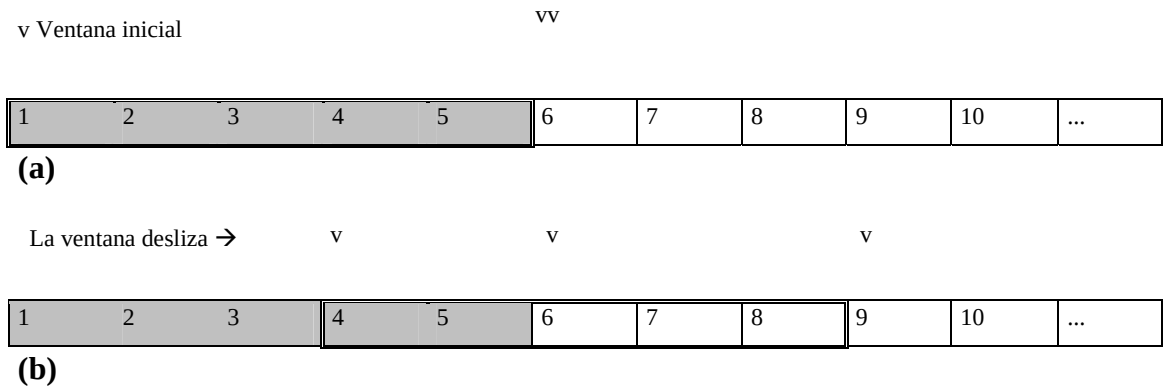
Un protocolo de acuse de recibo simple desperdicia una gran cantidad de ancho de banda de la red puesto que debe retrasar el envío de un nuevo paquete hasta recibir un acuse de recibo para el paquete previo.

La técnica de la *ventana deslizante* utiliza una forma más compleja del algoritmo de acuse de recibo y retransmisión que la anteriormente descrita. La forma de ver este algoritmo es pensar en una secuencia de paquetes que deben ser transmitidos, situar una

ventana sobre la secuencia y transmitir los paquetes dentro de la ventana. Cuando se recibe un acuse de recibo para el primer paquete de la ventana, ésta "desliza" una posición y se puede enviar el siguiente paquete, pues ya está dentro de la ventana.

El protocolo TCP utiliza un mecanismo de ventana deslizante para resolver dos problemas importantes: transmisión eficiente y control de flujo extremo a extremo, permitiendo a la máquina destino ordenar a la fuente restringir el envío de datos hasta que tenga suficiente espacio para tratar más datos. Sin embargo, el protocolo TCP ve a la corriente de datos como una secuencia de octetos o bytes que divide en *segmentos* para la transmisión. Generalmente, cada segmento viaja a través de la red Internet en un solo datagrama IP.

Así, el mecanismo de ventana deslizante TCP opera a nivel de byte, no al de paquete. Los bytes en la secuencia de datos se numeran secuencialmente y el "host" fuente mantiene tres punteros asociados a cada conexión que define una ventana deslizante. El primer puntero apunta al comienzo de la ventana, separando los bytes que han sido enviados y confirmados de los bytes que se han de enviar. El segundo apunta al final de la ventana, definiendo el byte más alto en la secuencia que puede ser enviado sin que llegue un acuse de recibo. El tercer puntero, dentro de la ventana, separa los bytes que han sido enviados de los bytes que no lo han sido. En la siguiente figura se muestra un ejemplo.



En la figura (a) se muestra la técnica de la ventana deslizante con 5 bytes en la ventana ya enviados y no confirmados. En la figura (b), los bytes hasta el 3 han sido

enviados y confirmados; los bytes 4 y 5 han sido enviados pero no confirmados; los bytes del 6 al 8 no han sido enviados, pero lo serán sin esperar a recibir acuse de recibo alguno, y los bytes 9 y más altos no se enviarán hasta que la ventana no se mueva.

Los bytes no confirmados son los que han sido enviados pero no han tenido acuse de recibo. El número de bytes no confirmados en un momento dado está limitado al tamaño de

la ventana.

Se ha dicho que la fuente mantiene una ventana para cada conexión y que el destino mantiene otra ventana similar. Sin embargo, como la comunicación TCP es "full-duplex", dos transferencias de datos ocurren simultáneamente en cada conexión, una en cada sentido. Las transmisiones son independientes, puesto que en cualquier momento los datos pueden ir en cualquiera de las dos direcciones, o en ambas. Por tanto, el software TCP mantiene dos ventanas para cada conexión en cada máquina; una desliza sobre la secuencia de bytes a enviar y la otra sobre los bytes que se van recibiendo.

4.9.2. Control de Flujo

El método de ventana deslizante del protocolo TCP permite que el tamaño de la ventana varíe. Cada acuse de recibo que especifica cuántos bytes han sido recibidos, contiene un "informe de ventana" que especifica cuántos bytes adicionales de datos está dispuesto a aceptar el "host" destino. Si este informe indica un tamaño de la ventana mayor que el actual tamaño de la ventana de la máquina fuente, ésta lo incrementa y comienza a mandar más bytes. Si el informe de ventana especifica un tamaño menor, la fuente disminuye el tamaño de su ventana y no enviará datos más allá del límite de la misma. El software TCP no contradice informes de ventana previos y nunca reduce la ventana a posiciones anteriores a las aceptadas anteriormente.

La ventaja de tener un tamaño de ventana variable es que proporciona control de flujo así como transferencia fiable. Si la capacidad de aceptar datos de la máquina destino disminuye, envía un informe de ventana menor. En el caso extremo, el destino

puede informar con un tamaño de ventana cero para detener la transmisión. Más tarde, cuando haya espacio disponible, el destino envía un informe de ventana mayor que cero para comenzar la transmisión de nuevo.

Con el método de las ventanas deslizantes de tamaño variable, el protocolo TCP soluciona el problema de transferencia fiable y control de flujo extremo a extremo; sinembargo, no soluciona el control de congestión en la red Internet, pues ésta tiene conectadas máquinas intermedias de distintas velocidades y tamaños que se comunican a través de redes de distintas características. La misión de ese control corresponde lógicamente al Protocolo de Red IP, que para solucionarlo utiliza el mecanismo menos preciso de los mensajes ICMP de disminución de flujo de la fuente.

4.9.3. Puertos TCP

El protocolo TCP está encima del IP en el esquema de capas de la red Internet. TCP permite a varios programas de aplicación en una misma máquina comunicarse simultáneamente y demultiplexar el tráfico TCP que se recibe entre las distintas aplicaciones. El protocolo TCP incorpora los denominados puertos, que identifican el último destino dentro de una máquina, el programa de aplicación que está haciendo uso de ese puerto. Cada puerto tiene asociado un entero para identificarlo. Dado que el número de puerto es único dentro de una máquina, el destino último para el tráfico TCP queda perfectamente determinado por la dirección internet del "host" y el número de puerto TCP en ese "host".

4.9.4. Formato del segmento TCP

Una vez explicados los mecanismos de trabajo del TCP, puede verse el formato del segmento TCP, que es la unidad de transferencia de datos con este protocolo. Los segmentos se intercambian para establecer y liberar conexiones, transferir datos, enviar acuses de recibo e informar sobre tamaños de ventana. Un acuse de recibo que vaya de una máquina A a otra B puede viajar en el mismo segmento que lleve datos de la máquina A a la máquina B (*piggybacking*).

Los segmentos TCP viajan en la porción de datos de los datagramas IP, como se muestra en la figura:

Encabezamiento IP	Mensaje TCP tratado como datos

El protocolo TCP entrega al servicio IP el segmento que contiene los datos, que es el que se transmite realmente, y una *pseudo-cabecera* que no se trasmite, pero que permite al protocolo IP completar los datos del o los datagramas que va a generar. El formato del segmento y de la pseudo-cabecera son los que se muestran en la figura siguiente.

0 16 31

PUERTO FUENTE			PUERTO DESTINO
NUMERO DE SECUENCIA			
NUMERO DE ACUSE DE RECIBO			
DESP.	RES.	CODIGO	VENTANA
CHECKSUM			PUNTERO URGENTE
OPCIONES			RELLENO
DATOS			
...			

DIRECCION IP FUENTE		
DIRECCION IP DESTINO		
CERO	PROTOCOLO	LONGITUD TCP

A continuación se describe el significado de los distintos campos del segmento.

- PUERTOS FUENTE Y DESTINO: Estos dos campos contienen los números de los puertos TCP que identifican los programas de aplicación en las máquinas fuente y destino
- NÚMERO DE SECUENCIA: Este campo identifica la posición, en la secuencia de bytes de la máquina fuente, del primer byte de datos del segmento.
- NÚMERO DE ACUSE DE RECIBO: Identifica la posición del byte más alto que la máquina fuente ha recibido, referido al número de secuencia de byte de la máquina a la que va destinado el segmento.
- DESPLAZAMIENTO (DESP.): Este campo de 4 bits contiene un entero que especifica el desplazamiento de la porción de datos del segmento. Es decir, indica la longitud de la cabecera en unidades de 32 bits. Se necesita porque la longitud del campo OPCIONES varía en longitud, según las opciones elegidas.
- RESERVADO (RES.): Este campo (6 bits) está reservado para su uso en el futuro.
- CÓDIGO: Es un campo de 6 bits que determina el propósito y contenido del segmento. Los seis bits explican cómo hay que interpretar otros campos en el encabezamiento de acuerdo con la tabla que se muestra a continuación.

Bit (de izquierda a derecha)	**Significado**
URG	El campo puntero urgente es válido
ACK	El campo acuse de recibo es válido
PSH	El segmento requiere un “push”
RST	“Resetear” la conexión
SYN	Sincronizar números de secuencia
FIN	La fuente ha alcanzado el fin de su secuencia de bytes

- VENTANA: Este campo contiene el "informe de ventana" explicado, es decir, el número de bytes que la máquina está dispuesta a aceptar. Este campo se incluye

tanto en los segmentos que llevan datos como en los que sólo llevan un acuse de recibo.

- CHECKSUM: Este campo contiene un entero de 16 bits usado para verificar la integridad del encabezamiento y los datos del segmento. Para calcular el "checksum", la máquina fuente antepone la pseudo-cabecera al segmento, y añade al final de los datos bytes conteniendo ceros hasta conseguir una longitud del segmento múltiplo de 16 bits. El checksum" se calcula una vez hechos los cambios según el siguiente algoritmo: se considera el segmento formado por enteros de 16 bits y se suman todos los complementos a uno de esos enteros utilizando la aritmética de complemento a uno. A efectos de hacer los cálculos se supone que ese campo tiene valor cero. Los ceros añadidos para rellenar, así como la pseudo-cabecera no se cuentan en la longitud del segmento y no son transmitidos. La razón de usar la pseudo-encabecera es permitir a la máquina destino comprobar que el segmento ha alcanzado su destino correcto, ya que incluye la dirección internet del "host" destino, así como el número de puerto TCP de la conexión. La máquina destino puede conseguir la información usada en la pseudo-cabecera a partir del datagrama IP que lleva el segmento.
- PUNTERO URGENTE: Cuando el bit URG está a 1, el campo PUNTERO URGENTE especifica la posición en la secuencia de bytes en la que los datos urgentes acaban. Los datos urgentes deben ser distribuidos lo más rápidamente posible. El protocolo especifica que cuando se encuentran datos urgentes, el software TCP que los recibe debe informar al programa de aplicación asociado a la conexión para que se ponga en modo "urgente". Los detalles de cómo el software TCP debe informar al programa de aplicación dependen del sistema operativo que se esté usando. Los datos urgentes contienen mensajes en vez de datos normales; por ejemplo, señales de interrupción del teclado.
- OPCIONES: El software TCP usa este campo para comunicarse con el software TCP al otro lado de la conexión. En particular, una máquina puede comunicar a otra el máximo tamaño de segmento que está dispuesta a aceptar. Esto es muy importante cuando un computador pequeño va a recibir datos de uno grande. Escoger un tamaño adecuado de segmento es difícil, puesto que el rendimiento de la transmisión empeora para segmentos excesivamente grandes (pues hay que

fragmentarlos en varios datagramas) o excesivamente pequeños (puesto que el tanto por ciento de encabezamientos es muy grande).

- RELLENO: Representa octetos conteniendo ceros, que son necesarios para que el encabezamiento sea un múltiplo exacto de 32, ya se había visto que el campo de DESPLAZAMIENTO indica la longitud del encabezamiento en unidades de 32 bits.
- DATOS: Representa los datos del segmento.

El significado de los campos de la pseudo-cabecera es el siguiente:

- DIRECCION IP FUENTE: Es la dirección internet del "host" que envía el segmento.
- DIRECCION IP DESTINO: Es la dirección internet del "host" al que va dirigido el segmento.
- CERO: Campo que contiene el valor cero.
- PROTOCOLO: Especifica protocolo TCP (es decir, 6).
- LONGITUD TCP: Este campo especifica la longitud total del segmento TCP.

4.9.5. Acuses de recibo y retransmisiones

Puesto que el software TCP envía los datos en segmentos de longitud variable, los acuses de recibo se refieren a la posición en la secuencia de bytes y no a paquetes o segmentos. Cada acuse de recibo especifica la posición inmediatamente superior al byte más alto recibido.

Cada vez que envía un segmento, el software TCP inicia un temporizador y espera por un acuse de recibo. Si éste no llega antes de que el temporizador expire, se supone que el segmento se perdió y, en consecuencia, se retransmite. Esto lo hacen la mayoría de los protocolos. La diferencia del protocolo TCP es que está pensado para usarse en la red Internet, en la que el camino entre dos máquinas puede ser desde una simple red de alta velocidad, hasta un camino a través de muchas redes y nodos intermedios. Por tanto, es imposible conocer *a priori* lo rápido que los acuses de recibo van a llegar

desde la fuente. Además, el retraso depende del tráfico, por lo que éste varía mucho de unos instantes a otros. El protocolo TCP se acomoda a los cambios en los retrasos de la red mediante un algoritmo adaptativo. Para conocer los datos necesarios para el algoritmo adaptativo, el software TCP graba la hora a la que envía el segmento y la hora a la que recibe el acuse de recibo para los datos del segmento. Con estos dos datos el computador calcula el llamado tiempo de retardo. Cada vez que el computador calcula un nuevo tiempo de retardo modifica su noción de tiempo medio de retardo para la conexión. Para calcular este tiempo medio se usa una media ponderada entre el último tiempo de retardo calculado y el tiempo medio anterior; y así el tiempo medio de retardo varía lentamente.

4.9.6. Establecimiento y liberación de una conexión TCP

Para establecer una conexión, el protocolo TCP utiliza el *three-way handshake.* El primer segmento transmitido se puede identificar puesto que tiene el bit SYN a 1 en el campo CODIGO. El segundo segmento, respuesta al anterior, tiene los bits SYN y ACK a 1, indicando que reconoce el primer SYN y continúa el proceso de conexión. El último caso es simplemente un acuse de recibo para informar al destino de que ambas máquinas están de acuerdo en que la conexión ha sido establecida. Normalmente, el TCP en una máquina espera pasivamente por una conexión y la otra la inicia; sin embargo, el *three-way handshake* está diseñado para que la conexión se abra incluso si las dos partes la intentan iniciar simultáneamente. Una vez que la conexión está abierta, los datos pueden ir en ambas direcciones.

Este intercambio de tres mensajes es necesario para una correcta sincronización entre los dos lados de la conexión. Esto es debido a que, dado que se trabaja en una red no fiable, el software TCP debe usar retransmisiones para las peticiones de conexión. El problema surge cuando se reciben peticiones retransmitidas de conexión cuando ésta está estableciéndose, o cuando la conexión ya ha sido abierta, usada y terminada. El *three-way handshake* junto con la regla de que TCP ignora peticiones de retransmisión una vez que la conexión está establecida, soluciona el problema.

El *three-way handshake* sirve además para que los dos lados de la conexión se pongan de acuerdo en los números de secuencia iniciales. Estos números de secuencia son transmitidos y confirmados durante el establecimiento de la conexión. Los números de secuencia no tienen que empezar en 1 obligatoriamente. De hecho, no deberían. Al empezar, una máquina (por ejemplo la A) pasa su número de secuencia inicial, *X*, en el primer paquete, que tiene el bit SYN a 1. La segunda máquina, B, recibe el SYN, graba el número de secuencia inicial de A y responde indicando su número de secuencia inicial, *Y*, así como un número de acuse de recibo que especifica que B está esperando el byte *X+1*. En el último paquete, la máquina A reconoce el número de secuencia inicial de B enviando como número acuse de recibo *Y+1*.

Para entender el proceso de libreración de una conexión hay que tener en cuenta que la comunicación es "full-dúplex", y que se transfieren dos secuencias de bytes independientes, una en cada dirección. Cuando un programa de aplicación indica al software TCP que no tiene más datos que transmitir, éste libera la conexión en esa dirección. Para liberar su mitad de conexión, el software TCP envía el último paquete de datos con el bit FIN a 1. En el otro extremo, el software TCP reconoce el FIN e indica al programa de aplicación que no se van a recibir más datos (por ejemplo, usando el mecanismo de final de fichero del sistema operativo). El bit FIN, al igual que se había visto con el SYN, ocupa una posición en la secuencia de bytes de la cabecera.

Una vez que la conexión se ha liberado en una dirección, no se aceptan más datos en esa dirección. Mientras tanto, los datos pueden seguir transmitiéndose en la dirección opuesta hasta que ésta también se libere. Cuando ambas direcciones han sido liberadas, la conexión se borra y los recursos que ésta utiliza (buffers, etc.) se liberan.

Un programa de aplicación libera una conexión cuando deja de usarla. Por tanto, liberar la conexión es algo normal. Sin embargo, a veces, aparecen condiciones anormales que obligan a una aplicación a abandonar la conexión. El protocolo TCP proporciona un mecanismo de "reset" para estas desconexiones anormales. Un lado de la conexión inicia el "reset" enviando un segmento con el bit RST a 1. El otro lado de la conexión responde abortando la misma. También informa al programa de aplicación que el "reset" ha ocurrido. El "reseteo" de una conexión significa que la transferencia en

ambas direcciones cesa inmediatamente, y los recursos de la misma (buffers,etc.) se liberan.

4.9.7. Envío forzado de datos

El protocolo TCP permite dividir la secuencia de bytes a transmitir en segmentos. La ventaja es la eficiencia. Se pueden acumular suficientes bytes en un "buffer" para crear segmentos razonablemente largos, con lo que se reduce el alto porcentaje de encabezamientos de los segmentos cortos.

Aunque el uso de "buffers" incrementa la densidad de transmisión de la red, puede no ser conveniente para algunos tipos de aplicación; por ejemplo, en el caso de una conexión TCP utilizada para enviar caracteres desde un terminal interactivo a una máquina remota. El usuario espera una respuesta instantánea a cada pulsación de una tecla. Si el software TCP "bufferea" los datos, la respuesta se puede retrasar.

Para satisfacer a usuarios interactivos, el TCP proporciona una operación de *push*, que un programa de aplicación puede utilizar para obligar al software TCP a transmitir los bytes de la secuencia sin que se llene el "buffer". Además, esta operación hace que el bit PSH del campo CODIGO esté a 1, con lo que los datos serán entregados al programa de aplicación en la máquina destino inmediatamente. Por lo tanto, cuando se envían datos desde una terminal interactiva, se usa la función de *push* después de cada pulsación de tecla.

Además de la función de *push*, el protocolo TCP proporciona una utilidad de *puntero urgente* que permite a la fuente informar al destino que los datos de la secuencia, hasta los que señala este puntero, son urgentes, y, por tanto, deben ser procesados lo más rápidamente posible. Por ejemplo, en una conexión TCP usada para intercambiar datos entre un terminal interactivo y un computador, los caracteres que paran y arrancan la salida por pantalla (por ejemplo control-s y control-q) deberían ser considerados urgentes, ya que el destino debe procesarlos inmediatamente.

4.9.8. Números reservados para puertos TCP

El protocolo TCP combina una adjudicación de números de puertos TCP dinámica y estática, usando una serie de asignaciones de puertos conocida para un conjunto de programas que se utilizan comúnmente (por ejemplo, correo electrónico). Sin embargo, la mayoría de los números de puerto están disponibles para que el sistema operativo los utilice a medida que los programas de aplicación los necesiten.

5. EL PROTOCOLO TELNET

TELNET permite a un usuario establecer una conexión TCP a un servidor de "login" en un lugar alejado. Aunque TELNET no es tan sofisticado como algunos protocolos de terminal remota, está disponible a lo largo de casi toda la red Internet.

TELNET ofrece tres servicios básicos. En primer lugar, define un terminal virtual que proporciona una interfase con sistemas remotos. En segundo, incluye un mecanismo que permite al cliente y servidor negociar opciones y proporciona una serie de opciones estándar. Por último, TELNET trata a ambos lados de la conexión simétricamente. Así, en vez de forzar a un lado a conectarse a un terminal de usuario, permite a ambos lados de la conexión ser un programa.

Cuando un usuario invoca TELNET, un programa de aplicación en la máquina se convierte en cliente y contacta con un servidor en uno de los puertos TCP reservados, estableciendo una conexión sobre la que se comunicarán. El cliente acepta pulsaciones

de tecla del terminal del usuario y las envía al servidor, a la vez que acepta caracteres que envía el servidor y los imprime en la pantalla del terminal.

6. TRANSFERENCIA DE FICHEROS

6.1 El Protocolo FTP

El protocolo FTP (*File Transfer Protocol*) permite, a usuarios autorizados, entrar en un sistema remoto, identificarse y listar directorios remotos, copiar ficheros desde ó a la máquina remota y ejecutar algunos comandos remotos. Además, FTP maneja varios formatos de ficheros y puede hacer conversiones entre las representaciones más utilizadas (por ejemplo, entre EBCDIC y ASCII). FTP puede ser usado por usuarios interactivos así como por programas.

El protocolo FTP permite al usuario acceder a varias máquinas en una misma sesión. Mantiene dos conexiones TCP independientes para control y transferencia de datos. Usa el protocolo TELNET para el control de la conexión.

La implementación del protocolo FTP depende del sistema operativo usado, aunque casi todas siguen el mismo patrón. En el lado del servidor, un proceso de aplicación, S, corre esperando por una conexión en el puerto asignado para TCP. Cuando un cliente abre una conexión con ese puerto, el proceso S lanza un nuevo proceso de control, N, para manejar la conexión, y vuelve a esperar por otro cliente. El proceso N se comunica con el cliente por medio de la llamada "conexión de control" y cuando recibe una petición de transferencia, inicia otro proceso adicional, D, que abre otra conexión con el cliente que solamente se usa para la transmisión de datos. Una vez que la transferencia de datos ha concluido, el proceso D cierra la conexión y termina. El cliente vuelve a su interactividad con el proceso N y puede solicitar otra transferencia de datos.

6.2. El Protocolo TFTP

El protocolo TFTP (*Trivial File Transfer Protocol*) proporciona un servicio barato y poco sofisticado de transferencia de ficheros.

Al contrario que FTP, el protocolo TFTP no utiliza un servicio fiable de transmisión. No utiliza el protocolo TCP, sino que se basa en el protocolo UDP (*User Data Protocol)* que corre por encima del IP, pero no es tan complejo como el TCP. TFTP utiliza temporización y retransmisión para asegurar que los datos llagan a su destino. La máquina fuente transmite un fichero en bloques de tamaño fijo (512 bytes) y espera por un acuse de recibo para cada bloque antes de enviar el siguiente. La máquina destino reconoce cada bloque que le llega.

Aunque TFTP contiene poco más que lo mínimo necesario para la transmisión, soporta varios tipos de formato de ficheros.

7. CORREO ELECTRÓNICO

Es una de las aplicaciones más comúnmente usadas en la red Internet ya que ofrece una forma sencilla de transmitir información. El correo electrónico es distinto a las otras aplicaciones de la red, pues no es necesario esperar a que la máquina remota reciba el mensaje para continuar trabajando. Cada vez que se quiere enviar un mensaje, el sistema crea una copia del mismo junto con la identificación del destino y se realiza una transferencia en modo "background". Posteriormente, el proceso de envío de paquetes tratará de entregar el mismo, contactando con el servidor de mensajes en la máquina destino.

Las ventajas de usar el correo electrónico en la red Internet es que ésta proporciona un servicio universal y además fiable, ya que, al usar una comunicación extremo a extremo, se garantiza que el mensaje permanece en la máquina fuente hasta que ha sido copiado satisfactoriamente en la destino.

El estándar que se utiliza para el envío de mensajes es el SMTP (*Simple Mail Transfer Protocol*). Este protocolo se centra en cómo el sistema de distribución de mensajes pasa los datos a través de una unión de una máquina con la otra. Inicialmente, el cliente establece una conexión TCP con el servidor y se intercambian una serie de comandos para establecer la unión. Una vez la conexión está establecida, el cliente puede enviar uno o más mensajes, terminar la conexión o pedir al servidor intercambiar los papeles de emisor y receptor para que los mensajes puedan fluir en la dirección contraria. El receptor debe reconocer cada mensaje y puede abortar la conexión entera o la transferencia del mensaje actual.

El protocolo de Oficina Postal (POP, actualmente POP3), define el diálogo entre un servidor de correo POP y la aplicación de correo electrónico en el computador del usuario. Esto es necesario cuando el usuario no lee el correo en la propia máquina donde se encuentra el buzón de su correo. Al recibir los mensajes, el servidor de correo los almacena en buzones privados para cada usuario. POP permite que un Agente de Usuario (UA) acceda al buzón, descargue todos los mensajes pendientes y después los borra del servidor. De forma similar funciona IMAP (Internet Message Protocol), sólo que este protocolo permite al usuario mantener sus mensajes en el servidor donde están los buzones y clasificarlos en carpetas, sin necesidad de hacer copias en su computador local.

8. EL PROTOCOLO HTTP

El Protocolo de Transferencia de Hipertexto (HTTP) es la base de la existencia del World Wide Web (WWW). El servicio HTTP en un host permite que usuarios a distancia puedan acceder a los ficheros que almacena si éstos conocen su dirección exacta. El protocolo HTTP define un sistema de direcciones basado en Localizaciones Uniformes de Recursos (URL). El URL de un recurso indica el protocolo o servicio que se emplea para ser accedido, la dirección del host donde se encuentra el recurso, y la ubicación del recurso dentro del host.

La información hypertexto se almacena en formato HTML. Se refiere a cada fichero como “página”. El “Browser” es el programa de usuario que conecta con el servidor mediante HTTP e interpreta la página HTML antes de mostrarla al usuario.

9. REDES DE AREA EXTENSA (WAN)

9.1. Establecimiento de Enlaces punto a punto

Para el establecimiento de comunicaciones entre sistemas informáticos a larga distancia se suelen utilizar tecnologías de redes conmutadas, bien de circuitos, o bien de paquetes. En ocasiones los paquetes de la red conmutada se denominan tramas o celdas debido a su reducido tamaño. Estas redes se construyen mediante enlaces punto a punto.

La tecnología de conmutación, además de las distancias entre sistemas, suele diferenciar las redes de área extensa de las redes locales, que son fundamentalmente de tipo "broadcast".

Los enlaces a larga distancia necesarios para comunicaciones punto a punto o la estructura de una red de área extensa, los puede conseguir el usuario de varias formas diferentes en función de sus necesidades y coste.

9.1.1. Conexiones temporales a través de RTB o RDSI

Cuando los intercambios de datos se realizan de forma muy esporádica, no resulta rentable mantener una comunicación permanente entre los sistemas informáticos. La alternativa es el establecimiento de una conexión temporal por medio de una línea telefónica convencional de la Red Telefónica Básica (RTB) o de una línea digital de la Red Digital de Servicios Integrados (RDSI).

En el primer caso es necesario el uso de un módem telefónico en ambos extremos de la línea, pudiendo alcanzar velocidades de transmisión de la información de hasta 33,6 Kbps. o 56 Kbps. dependiendo de la calidad de las líneas telefónicas. La facturación se realiza en función de la duración de la llamada telefónica que se establece con la marcación del número de abonado telefónico del destinatario, del horario en que se realiza y del tipo de ésta (local, provincial, nacional, etc.), a parte de los costes fijos mensuales o bimestrales por disposición y mantenimiento de la línea.

RDSI se presenta para el usuario en dos modalidades de acceso: *básico* y primario. La primera se suele utilizar para establecer conexiones temporales de voz y datos. Los costes para una llamada a través de la RDSI son similares, la forma de tarifación es la misma y solo varía la cuantía de los costes fijos. Mediante un "módem RDSI", que no es un módem analógico, si no que simplemente codifica los datos para su transmisión a través de la línea digital, el usuario dispone en la modalidad más económica de un canal digital full-duplex a 64 Kbps. La velocidad de transmisión de la información se puede incrementar mediante la modificación del contrato con la compañía suministradora a múltiplos de 64 Kbps. con el aumento de costes correspondientes. El establecimiento de la comunicación se realiza de la misma forma que con una línea telefónica convencional ya que el disponer de una línea RDSI permite sustituir a la anterior si se dispone de un teléfono digital RDSI, de tal manera que el usuario utiliza igualmente un número de abonado telefónico convencional. Además, a

través de la línea RDSI se pueden utilizar simultáneamente el teléfono y la conexión de datos.

El acceso primario no permite conexiones temporales, por ello su descripción se incluye en otro apartado.

En cualquiera de los dos casos, el uso de la RTB o la RDSI es para el usuario un simple medio de transmisión, con distintas velocidades posibles, que le permiten la conexión a otros usuarios utilizando el protocolo que en cada caso sea preciso.

9.1.2. Accesos permanentes con ADSL o con Cablemódem

Para el acceso a una red de datos como Internet, muchos usuarios desean poder disponer de una conexión permanente, barata, con un ancho de banda aceptable y que no impida el uso independiente del teléfono. Este tipo de acceso es posible mediante un *Módem ADSL* a través de una línea telefónica convencional o mediante un *Cablemódem* a través de una red de televisión por cable. Se utilizan fundamentalmente para acceder a los servicios de los proveedores de acceso a Internet (ISP).

Sin embargo, no se suelen utilizar estos accesos para el establecimiento de enlaces punto a punto a larga distancia o de redes de área extensa privadas, salvo que se establezca un sistema de Red Privada Virtual.

El acceso por ADSL proporciona al usuario un canal privado con el ISP independiente en velocidad y uso del acceso del resto de usuarios del ISP. Sin embargo, con el módem de cable todos los usuarios de una zona geográfica comparten el medio de transmisión hacia el ISP (constituido de un sistema de cableado coaxial con amplificadores de señal). La velocidad de acceso de cada usuario depende del número de usuarios de la zona conectados en cada momento y el uso que estén haciendo de la red. Además la privacidad de cada conexión podría verse comprometida.

Además el acceso ADSL puede ser conducido hacia otros servicios, a parte del de un ISP. Por ejemplo, hacia la red de la empresa en la que trabaja el usuario, proporcionando así al usuario un acceso punto a punto con la red de su empresa.

9.1.3. Alquiler de circuitos permanentes

Cuando la frecuencia del intercambio de datos aconseja al usuario mantener una conexión permanente punto a punto, una de las opciones es el alquiler de una línea de transmisión, a una compañía que disponga de ellas. No sólo las compañías telefónicas disponen de estas líneas, sino que hay otros tipos de compañías que tienen facilidades para el tendido de ellas. Este es el caso de empresas eléctricas, de distribución de gas, ferrocarriles, radiodifusión, etc., que normalmente realizan tendidos para la transmisión de datos paralelos a sus instalaciones para su propio uso o el alquiler a terceros.

Las velocidades de transmisión de la información y los dispositivos de interfaz para la conexión a esas líneas dependen de la tecnología y características de las mismas. La facturación suele ser fija e independiente del volumen de datos transmitidos. Permite al usuario disponer de un canal de capacidad fija para su uso exclusivo, lo que puede dar lugar a un desaprovechamiento del mismo en periodos de baja actividad.

9.1.4. Alquiler de circuitos virtuales permanentes o temporales

Los operadores de transmisión de datos suelen ofrecer servicios avanzados de red mediante sus redes de conmutación X.25, Frame-Relay, ATM, etc. Esto permite ofrecer al usuario un ancho de banda mínimo para sus conexiones temporales o permanentes y la posibilidad de aumentar ese ancho de banda si lo necesita y la carga de la red en esos momentos lo permite. El aprovechamiento de las líneas de transmisión es mayor, ya que al ser compartidas por múltiples usuarios, los periodos de inactividad se reducen al mínimo.

La velocidad de la línea de transmisión que une al usuario a la red suele ser bastante superior al ancho de banda contratado, ya que este suele tener un mayor peso en el coste de la conexión. Los parámetros de facturación pueden ser tremendamente

complejos, añadiendo en general a los costes fijos de conexión, velocidad de la línea de transmisión y ancho de banda contratado, costes variables en función del volumen de tráfico, horario de acceso, exceso sobre el ancho de banda contratado, etc.

9.1.5. Red Privada Virtual (VPN)

Una alternativa más económica que la contratación de circuitos virtuales, es la utilización de una red pública de conmutación de paquetes, como Internet, para establecer conexiones punto a punto de larga distancia de la empresa.

En el transporte de los datos pueden intervenir múltiples nodos pertenecientes a diferentes empresas y organizaciones públicas o privadas. Por ello, hay que poner especial interés en mantener su privacidad en su viajes a través de la red mediante el uso de técnicas de codificación. Estas técnicas son también recomendables en el caso de los ejemplos descritos en los apartados anteriores, especialmente, cuando no exista confianza en el medio de transmisión o las compañías que le dan soporte.

El acceso de cada extremo a la red Internet se hará a través de un proveedor de servicio (ISP), hasta el que se llega mediante alguno de los métodos habituales: RTB, RDSI, ADSL, Cablemódem, Frame-Relay, punto a punto, etc. Desde el ISP el acceso a la red Internet se realiza a través de alguna organización conectada a la misma normalmente mediante conexiones Frame-Relay o ATM. Es responsabilidad del ISP que estas últimas tengan el ancho de banda adecuado para dar un servicio de calidad a todos sus clientes, conectados temporal o permanentemente a la red Internet.

El elemento final que conecta cada sede a la línea de transmisión que va hacia el ISP, suele ser un encaminador (router) con capacidad para establecer una *Red Privada Virtual*, es decir, una conexión punto a punto con el encaminador del otro extremo, a través del que viajan todos los datos que se intercambian entre ambas sedes. Los datos viajan normalmente codificados y obviamente se pueden conectar más de dos sedes, manteniendo varias conexiones punto a punto.

De esta manera, los nodos de ambos extremos del enlace tienen la sensación de pertenecer a la misma red a la vez que ésta resulta prácticamente invisible para el resto del mundo. Los inconvenientes son la imposibilidad de garantizar un ancho de banda mínimo e incluso el servicio de los enlaces punto a punto, sometidos a los avatares del estado de las conexiones y el tráfico de la red Internet.

9.2. Circuitos de transmisión para redes de área extensa

Como se ha visto en el apartado anterior, existirán distintas tecnologías de transmisión física de los datos punto a punto que permiten el establecimiento temporal o permanente de conexiones para enlaces simples entre dos sistemas o la organización de la estructura de una red de área extensa.

9.2.1. Líneas de telefonía analógica

El modo más elemental para establecer un enlace temporal a larga distancia es utilizando una línea telefónica de la RTB. Originalmente se trataba del establecimiento de un enlace por conmutación de circuitos mediante la marcación del número del abonado de destino, pero en la actualidad las tecnologías de conmutación, sobre todo entre centrales de la RTB han cambiado mucho. Sin embargo, no ha cambiado tanto la línea que une al abonado con la central más próxima de la RTB. Se trata de un par de cables trenzados que debido a los sistemas de repetidores analógicos que se emplean para amplificar la señal que porta la voz del usuario, apenas tiene un ancho de banda de 4 KHz en el mejor de los casos, estando generalmente limitado entre 300 y 3000 Hz.

El no poder bajar de 300 Hz impide que señales digitales en banda base, que establecen generalmente valores de tensión continua en la codificación, puedan viajar por estas líneas. Por lo tanto se ha de utilizar un equipo modulador-demodulador, *módem*, para enviar y recibir señales a través de ellas. La mejora en la tecnología de estos equipos con la utilización de varias frecuencias portadoras sobre las que se modula la información digital combinando la modulación de fase y de amplitud, permite la transmisión de información full-duplex a velocidades de hasta 33,6 y 56 Kbps según la

calidad de la línea, en un medio que apenas permite transmitir 600 elementos de señal por segundo, baudios, sobre una frecuencia portadora.

Si la distancia del abonado a la central telefónica más próxima no es muy grande (unos 5 km.) y la central telefónica está suficientemente modernizada, el usuario podrá optar por soluciones RDSI o ADSL, ya que para esas distancias no son necesarios los amplificadores de banda en la línea de cobre. Para distancias mayores, se precisará algún sistema de amplificación específico que normalmente no es rentable para las compañías que ofertan el servicio.

ADSL (*Asymmetric Digital Subscriber Line o Línea de Abonado Digital Asimétrica*) consiste en transmitir conjuntamente voz y datos modulados a distintas frecuencias sobre la línea telefónica convencional. Ambas transmisiones se separan en la recepción por medio de un filtro (o *splitter*) colocado en ambos extremos de la línea telefónica. El filtro separa las frecuencias correspondientes a la voz (o telefonía convencional) de las frecuencias sobre las que se modulan los datos digitales. Así, cuando se está utilizando el módem ADSL (un módem especial para este tipo de tecnología), se tiene la línea de teléfono disponible para realizar simultáneamente llamadas de voz. Además la conexión del usuario a través del módem ADSL puede mantenerse las 24 h del día.

La comunicación que se establece mediante ADSL es asimétrica, ya que la velocidad en bits por segundo a la que se transmite la información al usuario es mucho mayor que la que se utiliza en sentido contrario. En la actual red telefónica española, la velocidad de entrada puede llegar hasta 2 Mb/s y la de salida hasta 300Kb/s, de ahí que sea considerada una tecnología de banda ancha,frente a la actual capacidad de los módem convencionales de enviar y recibir a 56 kb/s.

9.2.2. Líneas RDSI

Las compañías telefónicas han optado por las ventajas de la comunicación digital y han ido cambiando sus tradicionales circuitos analógicos para la interconexión de centrales telefónicas por sistemas digitales en los que las señales analógicas se

digitalizan para su posterior transmisión. Esto facilita en alguna manera la transmisión de datos digitales, que ya no necesitan ser modulados sobre portadoras analógicas.

La *Modulación por Codificación de Pulsos* (PCM) consiste en muestrear una señal analógica cada 125 μs. Esta frecuencia de muestreo se considera suficiente para reproducir sin problemas la voz humana en una línea telefónica. El valor de cada muestra se codifica en 7 u 8 bits (es decir, en valores de 0 a 127 ó de 0 a 255) que se envían con esa misma cadencia por la línea en banda base. Esto quiere decir que la velocidad de transmisión de información necesaria para un canal de voz en PCM será 8 bits / 125 μs = 64 Kbps. Precisamente los canales B de las líneas RDSI están diseñados para ser el soporte de conexiones telefónicas de voz con la ayuda de teléfonos digitales que muestrean y reconstruyen la voz de los interlocutores.

La forma más básica en la que se suelen utilizar los canales PCM es la que presentan las líneas RDSI. Las modalidades de acceso RDSI son en principio dos. El *acceso básico* y el *acceso primario*.

El ***acceso básico*** consiste en dos canales full-duplex a 64 Kbps, denominados de tipo B y uno a 16 kbps, denominado de tipo D, que suman en total 144 kbps. De todas formas, el usuario puede optar por un uso restringido del acceso básico a un coste más económico. Aún siendo tres canales independientes, se multiplexan sobre un único par de hilos para la transmisión y otro para la recepción que trabajan en realidad a 192 kbps debido a la necesidad de añadir bits para la sincronización y la compensación de niveles de continua en la señal.

El ***acceso primario*** no permite conexiones temporales, si no que se trata de un acceso punto a punto permanente hacia otro usuario o hacia un servicio de red de área extensa. El acceso primario se adapta de forma diferente a las características de la estructura de la red telefónica en Norte América y en Europa. En Norte América la velocidad de acceso es 1,544 Mbps que, a parte de los canales utilizados para señalización o control, se organiza generalmente para el usuario en 23 canales B y un canal D a 64 Kbps multiplexados en el tiempo sobre un único medio físico. En Europa la velocidad es 2,048 Mbps organizados en 30 canales B y un canal D.

9.3. Servicios de red de área extensa

Con el objeto de mejorar el servicio que da al cliente una red de transmisión de datos, se han desarrollado distintos servicios de red para redes de área extensa. Uno de los más utilizados ha sido la red X.25 que en España se denominó red IBERPAC-X.25. Se trata de una red de conmutación de paquetes que ofrece un servicio de red fiable con conexión sobre enlaces, en principio, poco fiables.

La mejora de las tecnologías de transmisión, ha hecho que muchas de las funciones del protocolo de X.25 orientadas a mejorar la fiabilidad del enlace, resulten innecesarias, eincluso perjudiciales por hacer más lentas las comunicaciones. Por ello aparece un sistema de conmutación de tramas denominado Frame Relay cuyo protocolo más simplificado se basa en la mejor calidad de las nuevas líneas digitales y, además, realiza la conmutación a nivel de enlace en lugar de a nivel de red como ocurre con X.25. Frame Relay consigue superar la velocidad de X.25 en al menos un orden de magnitud.

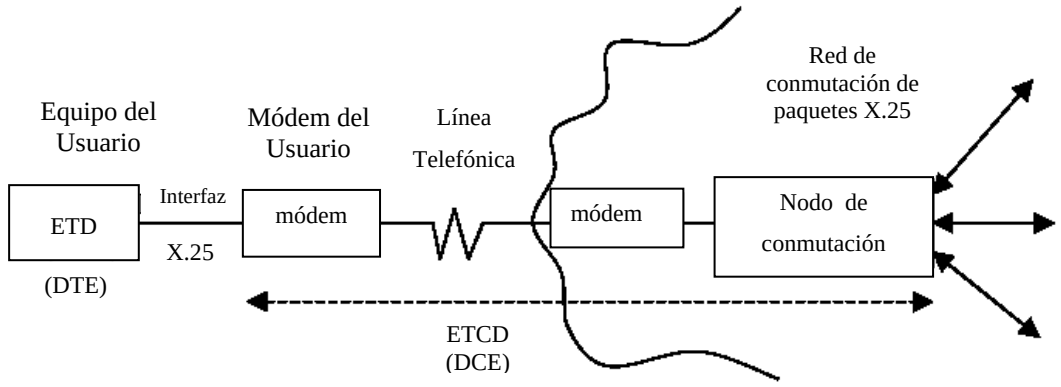
Posteriormente se ha desarrollado el Modo de Transferencia Asíncrono (ATM, *Asinchronous Transfer Mode*) donde se realiza conmutación de celdas. Conceptualmente es similar a Frame Relay ya que las celdas ATM son básicamente tramas de pequeño tamaño. La ventaja de ATM está en su funcionalidad que permite alcanzar velocidades varios ordenes de magnitud superiores a Frame Relay.

9.3.1. X.25

Aunque X.25 se refiere a una especificación que define la interfaz entre un usuario y una red de conmutación de paquetes como se verá más adelante, bajo esta denominación se conocen las propias redes de conmutación de paquetes que dan soporte a esos usuarios. El desarrollo de la red X.25 en España la llevó a cabo desde 1983 Telefónica, denominándola IBERPAC-X.25. Fue la sucesora de IBERPAC-RSAN que desde 1971 era la red de conmutación de paquetes existente fundamentalmente para los

clientes bancarios y que requería del desarrollo de aplicaciones propias bastante costosas. El desarrollo original de la red se basó en nodos de conmutación multiprocesador Intel 8086, con velocidades de acceso de los usuarios que variaban entre los 50 bps y los 64 kbps. Los nodos de conmutación estaban unidos entre sí mediante líneas a 16 y 64 kbps. Tanto los nodos de conmutación como la velocidad de las líneas de acceso e interconexión se fueron mejorando con el tiempo, de hecho, también la recomendación X.25 sufrió revisiones desde su primera definición por el CCITT (Comité Consultivo Internacional Telegráfico y Telefónico) en 1976, en los años 80, 84, 88, 92 y 93.

La recomendación X.25 especifica las funciones de tres capas que coinciden con las tres inferiores del modelo OSI: Física, Enlace y Red, esta última de nominada en X.25 capa de Paquete. En ella se define el intercambio de información entre un Equipo Terminal de Datos (ETD o en inglés DTE) y el Equipo Terminal del Circuito de Datos (ETCD o en inglés DCE) de la red de conmutación de paquetes más próximo al usuario. Hay que advertir que en X.25 se entiende como ETCD al nodo de la red de conmutación de paquetes más próximo al usuario, no exclusivamente al módem que utiliza este, aunque para el usuario, su interfaz con la red puede estar perfectamente representada por su módem.



9.3.1.1. Capa Física

La capa física en la interfaz está definida por el protocolo X.21 que fue pensado para un enlace completamente digital entre el ETD y el nodo de conmutación más próximo. Sin embargo la falta de disponibilidad de estos enlaces hizo que se generalizara el uso de X.21bis que recoge el uso de módem síncrono según las recomendaciones V.24 y V.28 que son en realidad equivalentes a la norma RS-232 de la EIA. En resumen, la interfaz física más generalizada para el acceso a una red X.25 es una conexión serie síncrona RS-232 entre el equipo del usuario y su módem.

9.3.1.2. Capa de Enlace

La capa de enlace se encarga de la transmisión fiable de datos a través del enlace físico mediante la transmisión de secuencias de tramas. La capa de enlace estándar es el LAPB (*Protocolo Equilibrado de acceso al Enlace*) que es un subproducto del protocolo HDLC y utiliza su formato de trama, aunque no todas las funciones que éste tiene. Las tramas transportan paquetes X.25 generados por la capa superior entre el usuario y el nodo de la red de conmutación al que está conectado o entre dos nodos de la red.

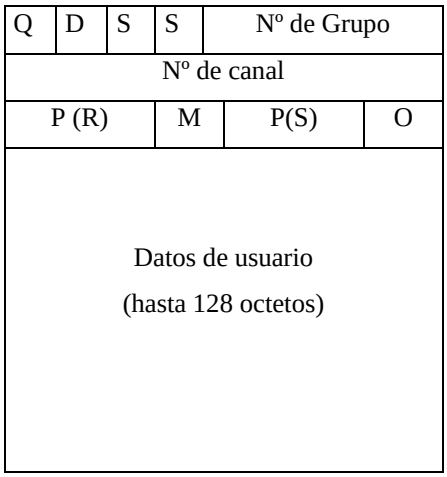
De todas formas, LAPB proporciona un enlace con conexión libre de errores y de duplicados sobre la línea que une al usuario con el primer nodo de conmutación de la red. Si los enlaces punto a punto entre cada pareja de nodos de conmutación se realizan también con LAPB, el control de flujo y de errores se hará también en cada uno de esos enlaces. Esto implica, por ejemplo, que la retransmisión de los datos que van en una trama errónea de un enlace se hará sólo en ese enlace, y no desde el punto de origen de los datos.

9.3.1.3. Capa de Paquete (Red)

Esta capa se basa en el establecimiento de circuitos virtuales que pueden ser *temporales* o *permanentes*. Los primeros se establecen mediante una *llamada virtual* que provoca la actualización de las tablas de encaminamiento de los nodos de la red

para dar servicio a ese circuito virtual temporal y se liberan una vez finalizado el intercambio de datos borrando las entradas de ese circuito en las tablas de encaminamiento. Los circuitos virtuales permanentes transportan datos como los anteriores pero no es necesario el proceso de llamada y liberación ya que son fijados mediante la configuración de los nodos de la red por tiempo indeterminado.

Existen básicamente seis formatos diferentes de paquetes X.25 para tres funciones distintas: transporte de datos, control de las llamadas virtuales (establecimiento del circuito virtual) y control de flujo y reinicio. Para cada función existen dos modelos de paquete diferente, en función de que se utilicen para numerar los paquetes números de secuencia de 3 bit o de 7 bits, en cuyo caso la cabecera del paquete es más grande. Como ejemplo se representa el paquete de datos con números de secuencia de 3 bits:



El bit Q es generalmente “0” en los paquetes X.25, se emplea para distinguirlos de los paquetes procedentes de un Equipo Empaquetador-Desempaquetador de Paquetes que se describirá más adelante. El bit D es un “1” si el paquete requiere acuse de recibo. Los bits SS indican la extensión de los campos de numero de secuencia, “01” significa 3 bits y “10” significa 7 bits. Los campos Nº de grupo (4bits) y Nº de canal (8 bits) definen en conjunto, el identificador del circuito virtual al que pertenece el paquete. P(R) y P(S) son campos de 3 bits para los números de secuencia de paquetes recibidos y enviados respectivamente. El bit M indicará si llegarán más datos y el bit O es “0” para indicar que se trata de un paquete de datos y un “1” para otros tipos de paquetes.

9.3.1.4. Acceso de terminales en modo carácter

En el desarrollo de la red X.25 hubo de especificarse un modo de acceso para terminales (ETD) que trabajaban en modo carácter ya que el acceso definido en principio es sólo para terminales en modo paquete, es decir, que son capaces de construir paquetes X.25. Determinados dispositivos como terminales financieras, terminales de videotexto (denominado Ibertext en España), datáfonos y otros como podría ser un PC con un módem asíncrono precisan de un dispositivo intermedio en el acceso a la red que empaquete los datos en el formato X.25 para su transporte por la red.

Este dispositivo se denomina *Desensamblador-Ensamblador de Paquetes* (DEP, o en inglés *Packet Assembly - Disassembly*, PAD). Su funcionalidad esta definida por la recomendación X.3 del CCITT, su diálogo con los ETD en modo carácter por la X.28 y con los ETD en modo paquete por la X.29.

9.3.1.5. Interconexión de redes

Cuando el transporte de los datos es hacia un ETD que está en otra red X.25, por ejemplo en la de otro país, el transporte se hace mediante un sistema de interconexión de redes X.25. El protocolo utilizado en este tipo de interconexiones se denomina generalmente *Acceso Internacional X.25* y está recogido en la especificación X.75 del CCITT.

Resumen de algunas de las recomendaciones X del CCITT	
X.25	Protocolo entre el ETD y el ETCD de una red de conmutación de paquetes
X.21	Circuitos físicos digitales para la interfaz ETD-ETCD
X.21 bis	Circuitos físicos para la interfaz con un módem síncrono
X.3	Funcionalidad del DEP
X.28	Protocolo entre el DEP y el ETD en modo carácter
X.29	Protocolo entre el DEP y el ETD en modo paquete
X.75	Protocolo de acceso internacional X.25

Los paquetes X.25, aún siendo una unidad de datos a nivel de red, pueden ser utilizados para transportar unidades de datos de otros protocolos de red. De hecho, X.25 fue uno de los primeros soportes para la conexión de encaminadores (routers) de la red Internet. Los datagramas IP que intercambian estos encaminadores se fragmentan si es necesario y transportan en paquetes X.25 hasta su destino, donde son reconstruidos los datagramas IP. A este proceso se le conoce como IP-Tunneling y es bastante ineficiente, por lo que sólo se recurre a él cuando no existe otra alternativa más eficaz.

9.3.2. Frame Relay

La especificación técnica de Frame Relay contempla su utilización para velocidades de acceso por encima de 2 Mbps. Se ha diseñado además para eliminar gran parte del coste de proceso que supone X.25 para la red y el usuario sobre todo al utilizar altas velocidades sobre líneas de alta calidad. Se podrían indicar como principales diferencias:

- La multiplexación y conmutación de conexiones lógicas tiene lugar en la capa de enlace en lugar de la capa de red, eliminando una capa completa de procesamiento.
- No hay control de flujo ni de errores en los enlaces individuales entre nodos. Este control se realiza extremo a extremo (entre la máquina origen y la de destino) y es responsabilidad de capas superiores.

Dentro de Frame Relay algunos autores distinguen entre los conceptos conmutación de tramas y retransmisión de tramas. En el primer caso la red realizaría los procedimientos de control de errores y de flujo y en el segundo, no. Como quiera que el segundo caso es más habitual y, como se comentaba antes, estas tareas se dejan a capas superiores, la traducción habitual de Frame Relay es *Retransmisión de Tramas*.

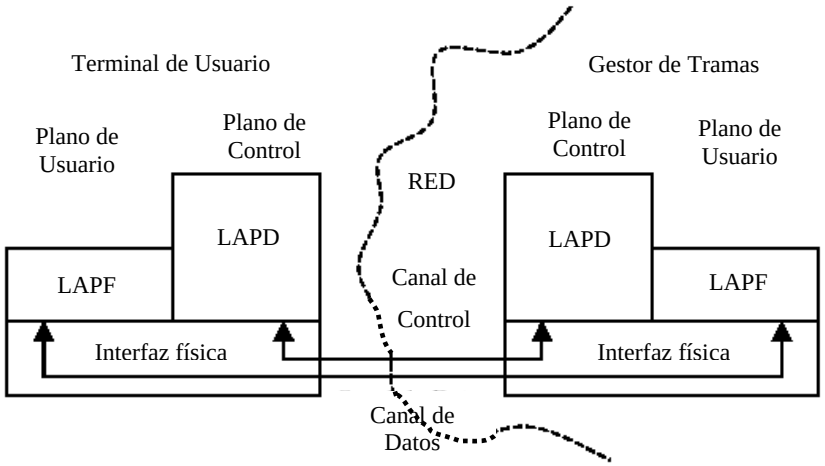
La retransmisión de tramas ofrece un servicio de enlace orientado a conexión con las siguientes propiedades:

- Se preserva el orden de la transferencia entre el origen y el destino aunque las tramas no van numeradas.
- Existe una pequeña probabilidad de pérdida de tramas.

Las conexiones pueden ser *temporales* o *permanentes*. Las primeras se establecen mediante un sistema de *control de llamada* que provoca la actualización de las tablas de conexión de los nodos de la red. Las conexiones permanentes se establecen cuando se acepta la conexión del usuario a la red y son fijadas mediante la configuración de los nodos de la red por tiempo indeterminado.

9.3.2.1. Protocolos de enlace en Frame Relay

En la arquitectura de Frame Relay existen varios protocolos de enlace para la realización de distintas funciones. Se habla de un **Plano de Control** mediante el que se realiza el establecimiento y liberación de conexiones lógicas. Estas funciones se llevan a cabo transmitiendo tramas del protocolo LAPD, otro derivado del HDLC (suele ser el método habitual de acceso a las redes Frame Relay). A través de este canal el terminal del usuario se comunica con el *gestor de tramas* (nodo de conmutación de la red Frame Relay) más próximo.



Una vez establecida la conexión entra en funcionamiento el **Plano de Usuario**, encargado del intercambio de los datos del usuario a través de una conexión por la que sólo viajan tramas de datos, evitándose así la sobrecarga de los mensajes de control. En el caso de una conexión a través de RDSI los datos viajan por canales de tipo B o H para los que se establece y libera la conexión, desde un canal D. Este intercambio de datos de usuario se realiza mediante tramas de protocolo LAPF (que en este caso es una versión de LAPD). LAPF está formado por dos partes que se denominan *LAPF control* y *LAPF core.* En Frame Relay sólo se utiliza el núcleo (LAPF core) lo cual quiere decir que no se realiza el control de errores ni de flujo.

1	2 a 4	Variable	2	1
Flag	Dirección	Datos	CRC	Flag

La trama es muy semejante a la trama HDLC, salvo que no existe campo de control. El campo de dirección tiene en principio dos octetos pudiendo ampliarse a 3 ó 4, de tal manera que contiene un DLCI (Identificador de Conexión de Enlace de Datos) de 10, 17 ó 24 bits según las necesidades. Los bits denominados EA de este campo permiten determinar la longitud del mismo. Otros bits de este campo permiten realizar el control de congestión en la red Frame Relay y se describirá su función más adelante. El formato del campo de dirección con dos octetos es el siguiente:

8	7	6	5	4	3	2	1
DLCI superior						C/R	EA 0
DLCI inferior				FECN	BECN	DE	EA 1

donde:

C/R	Bit de orden/respuesta
EA	Bit de ampliación del campo direcciones
FECN	Notificación de congestión explícita hacia delante
BECN	Notificación de congestión explícita hacia atrás
DE	Bit de rechazo de trama

9.3.2.2. Función de los gestores de tramas

El funcionamiento de la red Frame Relay se basa en la funcionalidad de *los gestores de tramas*, que son los nodos de conmutación de la red. Estos encaminan las tramas descritas anteriormente en función de su DLCI. Los gestores de tramas mantienen una tabla de conexión basada en DLCI que hace corresponder tramas de entrada por un canal, con su salida por otro, traduciendo adecuadamente su DLCI.

Por otro lado la conexión lógica con DLCI=0 de cada canal está reservada para control y es la que permite el control de llamadas para el establecimiento y liberación de las conexiones lógicas y, por lo tanto, la actualización de las tablas de conexión.

Cada vez que llega una trama, el gestor comprueba el CRC para detectar la existencia de posibles errores en la misma. Si se detecta un error, la trama es simplemente descartada. La recuperación de errores será misión de los usuarios finales.

9.3.2.3. Control de la congestión

El objetivo del control de congestión en Frame Relay es limitar la longitud de las colas en los gestores de tramas para evitar un colapso en el rendimiento de la red. Por lo general, esta congestión se produce cuando la carga ofrecida a la red es superior a su capacidad de proceso, lo que suele dar lugar a una congestión general de la red. También se puede producir de forma puntual en uno o varios gestores de tramas si la carga de trabajo de éstos es superior a su capacidad, con el consiguiente aumento del tamaño de sus colas de entrada y salida.

Los objetivos del control de congestión en Frame Relay, son los siguientes:

- Minimizar del rechazo de tramas.
- Generación mínima de tráfico de red suplementario.
- Sencillez de implementación y reducido coste para el usuario y la red

- Mantenimiento con alta calidad y mínima varianza, de una calidad de servicio adecuada.
- Distribución de los recursos de red entre los usuarios finales minimizando la posibilidad de que un usuario final pueda monopolizar los recursos de la red a expensas de otros usuarios finales.

Los procedimientos para controlar las congestiones en Frame Relay son tres: la e*strategia de rechazo*, *prevención de congestión* y *recuperación de congestión*.

9.3.2.3.1. Estrategia de rechazo

Es la respuesta más básica a una congestión severa en la cual la red se ve forzada a rechazar tramas. Este rechazo se hace de forma selectiva empezando por las tramas que tienen alterado el bit de rechazo, DE, del campo de dirección. La manipulación de este bit la hace el gestor de tramas que recibe una trama procedente del terminal del usuario.

El usuario contrata con la red para cada conexión *una tasa de información contratada,* CIR. La suma de las CIR de las conexiones del usuario debe ser siempre inferior o igual a la velocidad de la línea de acceso del usuario a la red. Por ejemplo, si el usuario tiene contratada la posibilidad de hacer dos conexiones simultáneas con una CIR de 16 kbps cada una, su línea de acceso debería tener una velocidad superior a 32 kbps. Si la velocidad de esta línea de acceso es por ejemplo 64 kbps el usuario no estará limitado por la CIR, es decir, podrá enviar hasta 64 kbps entre sus dos conexiones. Sin embargo, aquellas tramas que supongan que el usuario supera su CIR en un determinado intervalo de tiempo T, en alguna de las conexiones, serán marcadas por el primer gestor de tramas en su bit DE. En caso de congestión severa serán las primeras en ser descartadas por la red.

Para las conexiones permanentes la CIR se fija cuando el usuario se conecta a la red, y para las temporales, la determinación de la CIR forma parte del protocolo de control de llamada.

Lo lógico es que los recursos de un gestor de tramas tengan capacidad suficiente para la suma de las CIR de todas las conexiones de todos los sistemas de usuario finales conectados a él. De todas formas, esto no asegura que no haya rechazos antes de alcanzar la CIR, dando un servicio inferior a la CIR en caso de congestión extrema.

También se impone al usuario una tasa de transmisión máxima (inferior normalmente a la velocidad de acceso) por encima de la cual las tramas que supongan que esta tasa se supera en un intervalo de tiempo T son descartadas inmediatamente por el primer gestor de tramas.

Hay que destacar que estas funciones de marcado del bit DE y rechazo de las tramas de una conexión de usuario sólo las realiza el gestor de tramas que las recibe directamente del usuario. El resto de gestores de la red no modificarán el bit DE, ni controlarán la tasa máxima del usuario.

Por otro lado también tiene mucha importancia el intervalo de tiempo T en el que se computa la CIR. Si T fuera muy grande, por ejemplo varios minutos, probablemente el usuario no la superaría nunca ya que la información suele viajar a ráfagas con intervalos largos de inactividad entre ellas. Sin embargo, la red podría resentirse frente a esas ráfagas de tráfico intenso. Si T es muy pequeño, la red marcará muchas tramas como rechazables y se defenderá mejor de la congestión. Pero el usuario puede verse perjudicado y obligado a no enviar tramas muy seguidas para evitar superar su CIR en el tiempo T, a pesar de que su media de transmisión de información a lo largo del día sea muy inferior a esa CIR.

9.3.2.3.2. Prevención de congestión

Cuando un gestor de tramas de la red aprecia que existe congestión, comienza a marcar los bits BECN y FECN de las tramas. Los demás gestores no deben desactivar estos bits cuando la trama pasa por ellos. Estos bits son mensajes para los usuarios finales que reciben las tramas.

- **Notificación de congestión explícita hacia atrás (BECN):** notifica al usuario final que, las tramas que transmite en sentido contrario a la recibida pueden encontrar recursos congestionados. El usuario debe poner en marcha procedimientos para evitar la congestión. Estos procedimientos suelen consistir en reducir la velocidad a la que el usuario transmite las tramas hasta que desaparezca la señal.

- **Notificación de congestión explícita hacia delante (FECN):** notifica al usuario final que, la recibida ha encontrado recursos congestionados y que el tráfico que venga en ese sentido puede también sufrir la congestión. El usuario debe poner en marcha procedimientos para evitar la congestión que, en este caso, son más complejos ya que deberá comunicar la situación al usuario del otro extremo de la conexión para que reduzca la velocidad a la que transmite las tramas hasta que desaparezca la señal.

9.3.2.3.3. Recuperación de congestión

Los niveles superiores del usuario final pueden detectar de forma implícita que existe congestión en la red cuando detecta que ésta ha descartado una trama, bien por no haber llegado el acuse de recibo (del último mensaje enviado en la trama) desde el otro extremo, o bien, por haber sido rechazado éste por el otro extremo, por no tener el número de secuencia esperado.

En este caso, para recuperarse de la congestión, los niveles superiores del usuario final hacen uso de su control de flujo. Por ejemplo, reduciendo la ventana deslizante del control de flujo. Una vez recuperada la congestión, cuando no se detecten pérdidas de tramas, se volvería a recuperar poco a poco el tamaño de la ventana original.

10. OPERADORES DE CABLE

Antes de introducirnos en el mundo del cable, es necesario conocer una serie de conceptos que nos ayudarán a entender las ventajas e inconvenientes de la transmisión por cable, su comparación con otros medios de transmisión,...

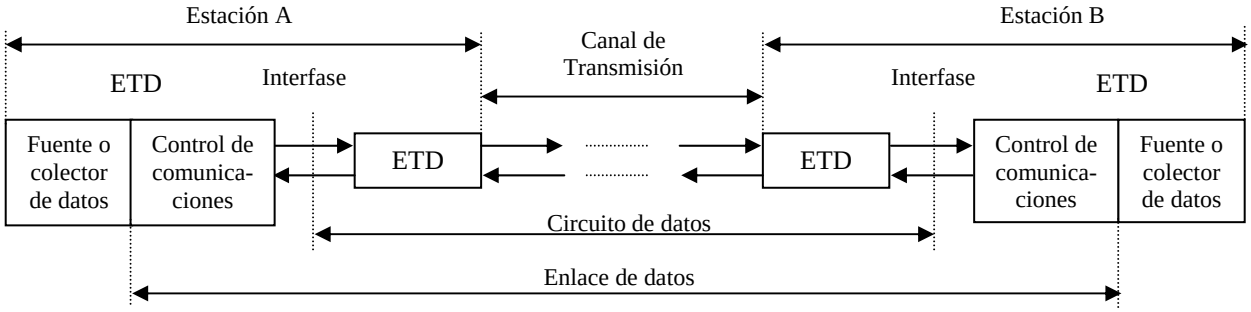
10.1. Circuitos de Transmisión de datos

En la estructura física de un circuito para la transmisión de datos se pueden distinguir los siguientes elementos:

- **Equipo terminal de datos (ETD):** Es la fuente o destino de los datos y puede ser más o menos inteligente, desde un equipo de fax a un computador. En él se encuentran la *fuente o colector de datos*, en el caso de un computador sería la aplicación que genera o recibe datos, y el *control de comunicaciones*, que se podría asemejar al software (driver) que controla el dispositivo de

comunicaciones (módem o interfaz de red). Se le denomina en numerosas ocasiones *DTE* por las siglas inglesas Data Terminating Equipment.

- **Equipo terminal del circuito de datos (ETCD):** Es el dispositivo encargado de convertir las señales que llegan por el *canal de transmisión* en otras legibles por el *ETD* y viceversa, formando junto con él, la *Estación*, *Nodo* o *Host* de comunicaciones. En el caso de que el *ETD* sea un computador, se trata de un dispositivo de comunicaciones (módem o interfaz de red). Se le denomina también *DCE* por las siglas inglesas Data Circuit-terminating Equipment.
- **Canal de transmisión:** Es el conjunto de medios de transmisión que unen los dos ETCD. También se le denomina *línea de transmisión* por que en muchas ocasiones el canal es algún tipo de sistema cableado, pero no siempre es así.
- **Circuito de datos:** Es el conjunto que forman el *canal de transmisión* y los *ETCD*.
- **Enlace de datos:** Es el conjunto que forman el *circuito de datos* junto con el *control de comunicaciones* de los *ETD.*



Circuito para la transmisión de datos

10.1.1. Características del canal de transmisión

Para describir la calidad de un determinado canal o medio de transmisión se utilizan algunos conceptos que describen sus características. Los más importantes son los siguientes:

Atenuación del canal

Es la relación entre la potencia de la señal a la entrada del canal de transmisión y la potencia que tiene esta señal a la salida del canal expresada en decibelios [dB]. Como esta atenuación suele ser proporcional a la longitud del canal, se indica la atenuación por cada 100 m o cada km de canal de transmisión.

Ancho de banda

Describe el rango de valores de frecuencia que pueden tener las señales a transmitir a través de ese medio de transmisión, ya sean de carácter eléctrico o electromagnético.

Generalmente, un canal presenta distintas atenuaciones a señales de distintas frecuencias, haciendo que determinadas frecuencias apenas se propaguen por el canal. Esto hace también que el ancho de banda se reduzca a medida que aumenta la longitud del canal, debido a que también aumenta la atenuación de la señal. Por ello, en ocasiones se expresa el ancho de banda del canal por cada 100 m o cada km de longitud.

Una línea telefónica, por ejemplo, tiene un ancho de banda que va desde los 300 Hz a los 3400 Hz, y cubre con ello la parte fundamental de las frecuencias que puede generar la voz humana. El límite no viene impuesto por el tipo de cable utilizado que tiene un ancho de banda mucho mayor, sino por los amplificadores que se insertan en la línea para contrarrestar el fenómeno de atenuación de la misma y poder transmitir la señal a larga distancia.

Velocidad de transmisión

Es el número de elementos de señal o cambios de condición por segundo. Se mide en *baudios*, que son el número de *elementos de señal* que se transmiten por segundo.

Un *elemento de señal* es un determinado valor de amplitud, fase, frecuencia, etc. o combinación de ellos, que codifica un determinado valor binario, de uno o varios bits.

Un *elemento de señal* puede tener N estados diferentes, que codificarán cada uno un valor binario diferente.

Capacidad del canal

También se la denomina *velocidad de transmisión de la información*. Es la velocidad máxima a la que se puede transmitir información sin errores, expresada en bits por segundo. Por lo tanto, la *capacidad del canal* será la velocidad en baudios máxima admisible por el canal, multiplicada por el número de bits que codifica cada *elemento de señal*.

10.2. Medios de Transmisión

El propósito de la capa física consiste en transportar el flujo original de bits de una máquina a otra. Hay varios medios de transmisión sobre los que se puede llevar a cabo este propósito. A continuación se mencionan algunos de ellos:

10.2.1. Par trenzado

El medio de transmisión más antiguo es el par trenzado, que aún es muy usado hoy en día. Consiste en dos hilos de cobre aislados, de 1 mm de espesor aproximadamente. Los conductores se trenzan en forma helicoidal para reducir la interferencia eléctrica con respecto a los pares cercanos que se encuentran a su alrededor. Dos cables paralelos constituyen una antena simple, mientras que si se trenzan no.

Su aplicación más común es el sistema telefónico. Con estos cables se pueden recorrer varios kilómetros sin tener que amplificar las señales, aunque sí son necesarios repetidores para distancias más largas. Cuando hay muchos pares trenzados en paralelo, recorriendo una distancia considerable, éstos se agrupan y se cubren con una malla

protectora. Los pares dentro de estos grupos podrían sufrir interferencias mutuas si no estuviesen trenzados.

Los pares trenzados pueden usarse para transmisión analógica o digital, y su ancho de banda depende del trenzado del cable y de la distancia que recorre. En muchos casos, pueden obtenerse transmisiones de varios Mbits por segundo sobre distancias de pocos kilómetros.

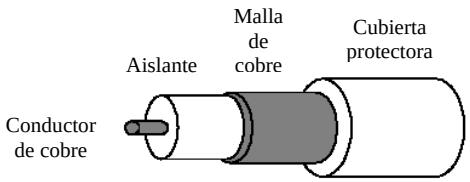
Debido a su buen comportamiento y bajo coste, están ampliamente difundidos.

Las capacidades típicas que se suelen alcanzar son: 100 Mbps sobre 100 metros, 2 Mbps sobre 1500 metros y 60 kbps sobre líneas telefónicas. Las tasas de error están entorno a 1 bit entre cada millón.

10.2.2. Cable coaxial

El cable coaxial es otro medio típico de transmisión. Hay dos tipos de cable coaxial, el cable coaxial de 50 Ω, que se usa en la transmisión digital y el cable coaxial de 75 Ω que se emplea para la transmisión analógica. El cable de 50 Ω también se conoce como cable coaxial de banda base, mientras que el 75 Ω se denomina cable coaxial de banda ancha.

El cable coaxial consta de un alambre de cobre en su parte central o núcleo. Este se encuentra rodeado por un material aislante. A su vez, el material aislante está recubierto por un conductor que suele presentarse como una malla trenzada. Por último, dicha malla está recubierta por una capa de plástico protector. De este diseño en forma de capas concéntricas es de donde se deriva el nombre.



El cable coaxial produce una buena combinación de un gran ancho de banda con una alta inmunidad al ruido. El ancho de banda que puede alcanzarse depende de la longitud del cable y del tipo, pudiendo ser de hasta 450 MHz. Así, un cable de 50 Ω y de 1 km de longitud permite obtener velocidades de hasta 10 Mbps en banda base y hasta 150 Mbps en transmisiones en banda ancha sobre cables de 75 Ω. Por otro lado, la señal eléctrica se propaga, según el tipo cable, a una velocidad que varía entre el 66% y el 80% de la velocidad de la luz. La atenuación de los cables varía entre los 20 y los 60 dB/100 m a 400 MHz.

Cable coaxial de banda base

En las redes locales se suele usar el cable coaxial como bus de comunicación sobre el que se transmiten señales en banda base. El bus de cable coaxial ha de tener en cada extremo una resistencia con la impedancia característica del cable (p.ej. 50 Ω) para evitar reflexiones en los mismos de la señal eléctrica que producirían interferencias e impedirían la comunicación. Ocasionalmente, se utilizan en conexiones punto a punto sin necesidad del uso de terminadores.

Existen dos formas de conectar ordenadores a un bus de cable coaxial: uso de conectores T o uso de conectores tipo vampiro. En el primer caso, hay que cortar el cable en dos partes e insertar una unión T, que vuelve a reconectar el cable y además proporciona una tercera conexión hacia el ordenador. El segundo tipo de conector consiste en hacer un orificio en el cable, de un diámetro y profundidad muy precisos, que atraviesa el cable hasta el núcleo. En el orificio se atornilla un conector especial que lleva a cabo la misma función de la unión en T, pero sin la necesidad de cortar el cable en dos.

El hecho de incluir una unión en T implica realizar un corte en el cable y por tanto desconectar temporalmente la red. Para una red con un gran nivel de utilización, detenerla cada vez que se conecta un nuevo equipo puede ser un gran inconveniente. Además, cuantos más conectores haya en el cable, más probabilidad existe de que alguna conexión sea defectuosa y ocasione problemas de vez en cuando.

Los conectores tipo vampiro no ofrecen este problema, pero son más difíciles de instalar. Si el orificio es muy profundo puede llegar a romper el núcleo provocando falsos contactos. Por otra parte, si no es suficiente profundo, pueden provocarse falsos contactos debido al aislante. Además, los cables en este tipo de conexión son más gruesos y por tanto más caros.

Cable coaxial de banda ancha

Este cableado se utiliza comúnmente para el envío de la señal de televisión por cable. El término banda ancha proviene del medio telefónico, y se refiere a frecuencias mayores a 4 kHz.

Utilizan la tecnología patrón para envío de señales de televisión por cable y por ello pueden llegarse a alcanzar hasta 450 MHz de ancho de banda para longitudes de hasta 100 m. Un cable típico de 300 MHz puede, por lo general, mantener velocidades de hasta 150 Mbps.

Es habitual que los sistemas de banda ancha se dividan en varios canales, por ejemplo en canales de 6 MHz para el envío de señal de televisión. Cada canal puede emplearse de forma independiente, por lo que en un mismo cable pueden coexistir señales de vídeo, voz y datos.

Una diferencia clave entre los sistemas de banda base y los de banda ancha es que los últimos necesitan amplificadores que repitan la señal en forma periódica. Estos amplificadores sólo pueden transmitir señales en una dirección de manera que un ordenador que dé salida a un bloque de información sólo puede alcanzar a otros ordenadores que estén “aguas abajo”. Hay dos formas de solucionar este problema: uso de cable dual y uso de canales distintos.

En los sistemas de cable dual, se tienden dos cables idénticos paralelos. Para transmitir información, el ordenador emplea uno de ellos, que envía el mensaje hacia el

repetidor central (en la cabeza de la red). Una vez que el mensaje alcanza dicho repetidor se reenvía por el otro cable para que todos los ordenadores puedan leerlo.

El otro sistema consiste en aplicar diferentes frecuencias para las señales que entran y salen de un ordenador, sobre un cable sencillo. La banda de baja frecuencia se emplea para enviar información hacia el repetidor central, para que éste la reenvíe hacia los ordenadores por la banda de mayor frecuencia. En el *sistema de asignación baja* el tráfico de llegada al repetidor usa una frecuencia de entre 5 y 30 MHz, mientras que el de salida usa una banda entre 40 y 300 MHz. En el *sistema de asignación alta*, el tráfico entrante va entre 5 y 116 MHz, mientras que el de salida va entre 168 y 300 MHz. La adopción de estas técnicas se debe en parte a la fiabilidad y bajo coste del hardware empleado.

Un sistema de banda ancha puede usarse de diferentes maneras. Por ejemplo, se puede asignar un canal para su uso exclusivo por un par de ordenadores, mientras que los demás deben competir por el uso de un canal temporal mientras dure la comunicación.

La instalación del sistema de banda base es simple y económica y emplea interfaces baratas. Ofrece un sólo canal digital con velocidades de unos 10 Mbps para distancias de 1 km. Son muy empleados para el diseño de redes locales.

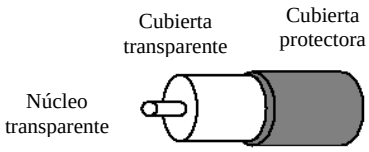
La instalación del sistema de banda ancha requiere por lo general personal especializado. Además es necesario realizar un mantenimiento del sistema para asegurar que todos los repetidores están correctamente sintonizados. Por otra parte, un fallo en el repetidor central llevaría a la desconexión del sistema. Este resulta en general, más costoso. Sin embargo, ofrece el uso de varios canales, aunque se limitan a unos 3 Mbps cada uno, y permite la transmisión simultánea de datos, voz y señales de televisión.

10.2.3. Fibras ópticas

Los avances en el campo de la tecnología óptica han hecho posible la transmisión de información mediante pulsos de luz. Un pulso de luz puede utilizarse para indicar un

bit de valor 1, y su ausencia un bit de valor cero. La luz visible tiene una frecuencia de alrededor de 10^8 MHz, por lo que el ancho de banda de un sistema de este tipo tiene un potencial enorme.

Un sistema de transmisión óptica tiene 3 componentes: el medio de transmisión, la fuente de luz y el detector. El medio de transmisión es una fibra ultradelgada de vidrio o silicio fundido. También existen fibras fabricadas con polímeros plásticos de calidad inferior a las de vidrio. La fuente de luz puede ser un LED o un diodo láser; cualquiera de los dos emite luz cuando se le aplica una corriente eléctrica. El detector es un fotodiodo que genera un pulso eléctrico en el momento en el que recibe un rayo de luz. La transmisión de datos que se obtiene es unidireccional.



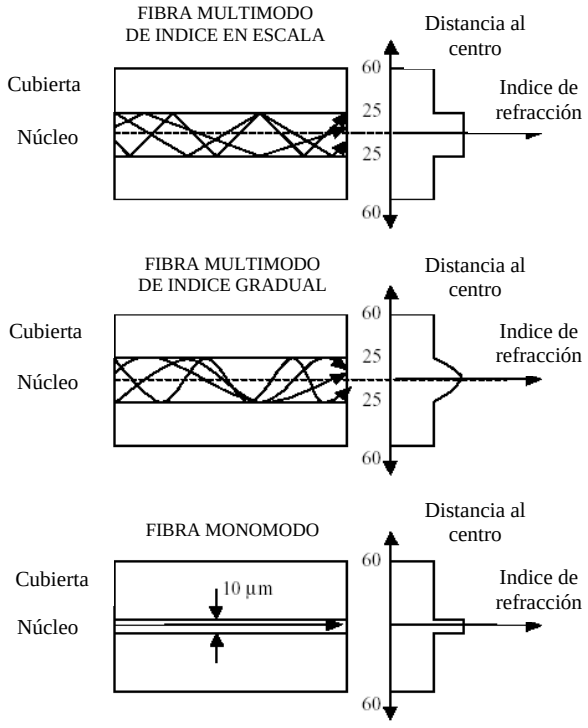
El sistema se basa en el principio físico de la refracción. Cuando un rayo de luz pasa de un medio a otro, el rayo se refracta en la frontera entre ambos medios. En general, el ángulo de refracción depende de las propiedades de los medios en contacto, en particular de sus índices de refracción. Si el ángulo de incidencia se encuentra por encima de un determinado valor crítico, la luz se refleja y no sale del medio.

La fibra óptica esta compuesta por dos medios transparentes de distinto índice de refracción, un núcleo y un revestimiento que lo envuelve. Finalmente se cubre el conjunto con una cubierta opaca. Así, los rayos que incidan por encima del ángulo crítico van a quedar atrapados dentro del núcleo de la fibra, y pueden propagarse a lo largo de varios kilómetros sin apenas tener pérdidas.

Dado que cualquier rayo de luz incidente, por encima del ángulo crítico, se reflejará internamente, existirá una gran cantidad de rayos diferentes rebotando a distintos ángulos. A esta situación se la conoce como **fibra multimodo**. Si el índice de

refracción es uniforme en todo el núcleo, la fibra se denomina de **índice de escala** y los haces rebotarán bruscamente en el punto de contacto del núcleo con el revestimiento, que tiene un índice de refracción diferente. Si el índice de refracción del núcleo varía gradualmente, aumentando poco a poco hacia el centro del mismo, la fibra se denomina de **índice gradual** y los haces de luz son conducidos de forma más suave hacia el interior de la fibra, sin que reboten bruscamente reduciendo así las pérdidas en la propagación del haz.

Si el diámetros se reducen hasta que sea semejante al valor de la longitud de onda de la luz, la fibra actúa como una guía de ondas, y la luz se propaga en línea recta sin rebotar, produciendo así una **fibra monomodo**. Estas fibras necesitan diodos láser para su excitación, se asegura una mayor eficiencia y pueden usarse en distancias muy largas.



La **apertura numérica** de una fibra óptica es el parámetro que define el ángulo crítico para que la luz se propague a través de la fibra óptica. Este parámetro está íntimamente relacionado con los diámetros del núcleo y el revestimiento. Cuanto más

grandes sean éstos, mayor es la **apertura numérica** y más fácil resultará el acoplamiento de dos segmentos de fibra óptica o de ésta con los dispositivos emisor y receptor. Sin embargo, crecerán a la vez, las pérdidas en la propagación de la luz.

Los enlaces de fibra óptica se están usando para la sustitución de enlaces telefónicos de larga distancia. Hasta ahora, se usaba cable coaxial de banda ancha. También se usan para el montaje de redes LAN, aunque requieren una tecnología más compleja que el cable coaxial. El problema fundamental es que la realización de conexiones intermedias es complicada y supone una importante pérdida de luz.

Una red en forma de anillo es una solución al problema ya que es en realidad una colección de enlaces punto a punto. La interfaz que existe en cada ordenador permite el paso del flujo de los pulsos de luz al siguiente enlace y como unión en T por medio de la cual el ordenador envía y acepta mensajes.

Hay dos tipos de interfaz. Uno es de tipo pasivo. Está formado por dos conectores fusionados con la fibra principal, uno tiene un LED en su extremo (para transmisión) y el otro tiene un fotodiodo (para recepción). La conexión es completamente pasiva y por tanto muy fiable.

El otro tipo de interfaz es el receptor activo. La luz incidente se convierte en señal eléctrica y se regenera a su máximo valor, retransmitiéndose de nuevo como luz. Como en cada enlace se regenera la señal, cada línea puede tener varios kilómetros de longitud. En cambio en un anillo pasivo, se pierde luz en cada enlace por lo que está limitado el número de estaciones y la longitud total del anillo.

Entre las principales ventajas cabe destacar las siguientes:

a) Mayor velocidad de propagación de la señal. La señal luminosa se propaga a la velocidad de la luz.
b) Mayor capacidad de transmisión. En la actualidad se pueden hacer transmisiones de hasta 1 Gbps en distancias de 1 km.
c) Inmunidad ante interferencias electromagnéticas.

d) Menor atenuación. 5 a 20 dB/km a 400 Mhz.
e) Mayor ancho de banda.
f) Tasas de error menores. 1 error por cada 10^9 bits frente a 1 por cada 10^6 en los cables eléctricos.
g) No hay riesgos de cortocircuitos o daños de origen eléctrico.
h) Peso mucho menor.
i) Menor diámetro y más flexibles lo que facilita su instalación.
j) Es más difícil realizar escuchas sobre una fibra óptica que sobre un cable eléctrico.
k) Se pueden emplear varios canales empleando longitudes de onda diferentes simultáneamente sobre la misma fibra.
l) Tiene mayor resistencia a los ambientes corrosivos que los cables eléctricos.
m) Las materias primas para su fabricación son abundantes.
n) Su vida media es mucho más larga que la de un cable eléctrico.

Sin embargo también presentan inconvenientes. Por un lado, las fibras ópticas son inherentemente unidireccionales y el coste de las interfaces es mucho mayor que en el caso eléctrico. Por otro lado, la unión de fibras ópticas es complicada y todavía más su derivación. Uno de los elementos más costosos de una instalación de fibra óptica es la incorporación de las férulas de conexión en los extremos de las fibras. Las férulas suelen ser complejas y de laboriosa instalación. De la delicada y correcta instalación de estas férulas, depende el correcto alineamiento entre los extremos de las dos fibras que se vayan a conectar o del extremo de la fibra con los dispositivos emisor o receptor. Si el alineamiento no es correcto, la limitada apertura numérica de una fibra puede impedir total o parcialmente la propagación de la señal luminosa.

10.3. El cable. Definición. Ventajas. Inconvenientes.

Definiciones

<< El cable es el sistema de telecomunicación que nos trae la televisión, el teléfono, la conexión a Internet de Banda Ancha (principalmente por cablemódem) y

otros servicios a través de la fibra óptica. Originariamente, se utilizó un cable de TV, para transmitir canales sin necesidad de utilizar antenas y con decodificador. Posteriormente, se le fue añadiendo la característica de bidireccionalidad, de forma que se pudiera ofrecer telefonía e Internet, empleando principalmente para ello la fibra óptica.>>

<< Un **cable módem** es un dispositivo que permite conectar el PC a una línea local de TV por cable a aproximadamente 1,5 Mbps. Esta tasa de datos excede con mucho la de los módems telefónicos de 28,8 y 56 Kbps actualmente prevalecientes, y los hasta 128 Kbps de RDSI y es más o menos la tasa de transferencia de datos disponible para los suscriptores del servicio telefónico de Línea Digital del Suscriptor (Digital Subscriber Line, DSL). Un módem cable puede ampliarse o integrarse a una caja "set top" que convierte nuestro televisor en un canal de Internet. Para conectarse al PC, la línea de cable debe dividirse de modo que parte de ella vaya al televisor y la otra parte al módem cable y al PC .>>

<< En sentido estricto, se llama **cable módem** al dispositivo por el cual se establece la comunicación entre Internet y la Computadora Personal del domicilio u oficina y que, si bien encierra diferencias sustanciales con los módem tradicionales, cumple su misma función. Luego, por asociación, se denomina así al servicio de acceso a Internet a través de estos dispositivos y de la infraestructura de la red de cable. El acceso a Internet a través de cable comparado con los accesos actuales dial-up es una nueva *categoría de servicio*. Es como comparar una bicicleta con un auto, y la percepción de los clientes es exactamente ésa: no pueden volver a los sistemas dial-up después de haber probado el cable. En términos de velocidad, la red de cable módem supera en hasta 240 veces a los módem telefónicos más rápidos. En teoría, permite alcanzar los 36 Megabits por segundo al recibir datos de la Red, aunque la velocidad real dependerá de otros factores. Así, la transferencia de un archivo de 10 MB tardaría idealmente 24 minutos con un módem de 56 Kbps (velocidad estándar en conexión dial-up), ofrecida por cualquier proveedor de Internet en la actualidad, mientras que con un cable módem de 10 Mbps de ancho de banda tardaría tan sólo 8 segundos. En términos sencillos, el cable módem permite el acceso a Internet a través de la red de TV por cable (CATV). Con una doble conexión -una con la compañía proveedora y otra con el PC-,

la mayoría de estos dispositivos son externos a la computadora y se vinculan con ella por una placa de red tipo Ethernet. La mayoría de las redes de CATV actuales son de tipo híbrido, pues combinan fibra óptica, en el tramo que va desde la central del cable (headend) hasta los distintos nodos, y cable coaxial desde el mismo nodo hasta el hogar de cada abonado. Para el cable módem de doble vía, la ida y vuelta de datos se realiza en la red de cable en un flujo de datos que puede ser, a su vez, *simétrico* o *asimétrico*, según el ancho de banda asignado >> para el envío y recepción de datos, sea igual o diferente, respectivamente.

<< Un cable módem es realmente más comparable a una placa de interfaz de red, (network interface card, NIC) que a un módem de ordenador. Todos los módems cable conectados a la línea coaxial de una compañía de TV por cable se comunican con un Sistema de Terminación de Módem Cable (Cable Modem Termination System, CMTS) en las instalaciones de la compañía local de cable. Todos los cable módems puede recibir y enviar señales a los CMTS únicamente, no a otros cable módems en la línea.>>

<<El ancho de banda real para el servicio de Internet por medio de una línea de cable de TV es de hasta 27 Mbps en el camino de bajada hacia el suscriptor, con un ancho de banda de aproximadamente 2,5 Mbps para respuestas interactivas en dirección opuesta. Pero, como probablemente el proveedor local no estará conectado a Internet en una línea más rápida que 1,5 Mpbs, la tasa más verosímil de transferencia de datos estará cerca a los 1,5 Mpbs.>>

<<Además de la mayor velocidad de transferencia de datos, una ventaja de Internet por cable sobre la que se provee por teléfono es que se trata de una conexión continua. Esta nueva tecnología le permite al usuario estar conectado a Internet apenas enciende su computador. No necesita utilizar una línea telefónica debido a que la conexión es por cable y no por el teléfono.>>

<<Otra de las ventajas es su velocidad: el aparato es cinco veces más rápido que el módem convencional.>>

(VULNERABILIDADES DE ACCESO Y NAVEGACIÓN EN INTERNET POR CABLE MÓDEM © Jorge Machado Lima-Perú)

<<Nadie puede negar que el acceso a Internet a través del Cable Módem es una conveniente oportunidad para usuarios domésticos y hasta para pequeños negocios. Sin embargo, ¿cuáles son los riesgos de seguridad para este tipo de conexión? La conectividad y transmisión de datos en Internet ha llegado a convertirse en el deseo más anhelado de todos los usuarios de computadoras. Obtener 256 kbps o más, por una razonablemente pequeña cantidad de dinero es "algo demasiado bueno para poder desperdiciarlo".

A causa de ello muchísimos particulares y pequeñas organizaciones han empezado a tomar la decisión de instalar conexiones de Cable Módem y que, en el Perú, la empresa que provee este servicio la denomina CableNet, aunque por ahora limitada a una velocidad máxima de velocidad 128 kbps. La empresa no informa sobre el valor de Overbooking, vale decir, la cantidad máxima de usuarios conectados a una misma sub-net.

De acuerdo a definiciones precisas, los cable módems son conexiones de banda ancha, mientras que el ADSL (Asymmetric Digital Subscriber Line) debería ser llamado más exactamente una conexión de "banda de base" ya que no es compartida. A pesar de ello y en la práctica, banda ancha significa "velocidad" y ese concepto es aplicado a ambos tipos de conexiones.

La Banda Ancha ofrece muchos beneficios, pero la amplia difusión de su uso no significa del todo una muy buena noticia. ¿La razón? El Cable es inseguro y la mayoría de sus usuarios están desinformados o desprevenidos contra estos graves riesgos. Sin embargo, existen algunas soluciones y contra ofensivas para contrarrestar o combatir estas amenazas y proteger su información y privacidad.

La más desapercibida vulnerabilidad de este tipo de banda ancha es la naturaleza del uso común y compartido de las conexiones por cable. Y es que, por ejemplo, todos los usuarios en una área local, tal como un edificio de apartamentos o los vecinos de un

mismo barrio, comparten la misma sub-red cuando se conectan al Cable Módem. Cualquier otro usuario que comparta ese mismo tipo de conexiones, provisto de las herramientas de software adecuadas o quizás con conocimientos de "hacking" tiene la capacidad potencial de ocasionar ataques o incursiones en otros sistemas compartidos.

Aún siendo el caso de que el usuario integre una sub-red conectada a otras de distritos cercanos, su sistema siempre será vulnerable, debido a que otros sistemas comparten potencialmente su misma conexión de red.

Una situación bastante perjudicial resulta del hecho que, en el caso que uno o más usuarios estén descargando archivos de gran extensión, como los populares MP3, ocasionará que los demás participantes de una misma sub-red vean reducidas dramáticamente sus velocidades de acceso, navegación y descarga de archivos. Otra vulnerabilidad de la Banda Ancha por Cable Módem es que siempre está activada, esto significa que una vez que un cable módem está en línea, su red está siempre conectada a Internet. Cuando se usa la conexión por marcación directa, al desconectarse, el usuario ya no participa por completo en Internet. En cambio con la banda ancha del cable, un sistema está constantemente sujeto a riesgo de ataques, los cuales pueden suceder durante las 24 horas del día, 7 días de la semana.

Por si no fuese suficiente el estar permanentemente conectado en una sub-red compartida, los sistemas de Cable Módem tienen direcciones IP, ya sea estáticas o son las mismas, también en forma compartida. Este factor no necesariamente ocasiona ataques dirigidos a su sistema, pero sí los facilita. El caso contrario es el acceso por marcación directa, ya que cada vez que un usuario se conecta, su Proveedor de Servicios de Internet (ISP) le asigna automáticamente una dirección IP diferente.

Una final pero significativa vulnerabilidad es la velocidad que hace atractiva la banda ancha para los usuarios. Una vez que la red es compartida con la conexión de alta velocidad, permite que el intruso pueda enviar rápidamente archivos infectados, troyanos o con el uso de herramientas de hackers, introducirse de forma no autorizada a un sistema y poder descargar archivos, passwords y documentos importantes.

No está demás mencionar que la banda ancha de alta velocidad también sufre de serias caídas o interrupciones de conexión. Conjuntamente con los aspectos fundamentales de seguridad, tales como rastreo de virus y accesos de passwords seguros, el usuario debe implementar algunas medidas de seguridad primordiales, antes de que piense en tomar la decisión de elegir un servicio de acceso a Internet vía Cable Módem.

Una forma común para que los intrusos obtengan acceso a un sistema es a través de los recursos compartidos, los mismos que son aplicados dentro de una Red Local. Sin embargo la libertad de compartir estos recursos entre receptores internos que sean confiables, no debería ser ofrecido a desconocidos o posiblemente entidades externas sospechosas, particularmente en un entorno de banda ancha cuando su dirección IP rara vez es modificada.

Para protegerse de los ataques directos, búsqueda de puertos abiertos y otras vulnerabilidades se requieren 2 sistemas importantes:

- *Un buen software antivirus con soporte técnico 24/7 y actualizaciones diarias.*
- *Un Firewall personal.>>*

Como siempre, un consejo de un conejo:

La conectividad de banda ancha le ofrece acceso y navegación a altas velocidades a un costo razonablemente bajo. Pero el costo fundamental de ignorar los riesgos de seguridad inherentes en este tipo de conectividad, puede costar muy caro. Obtener ventaja de este recurso, requiere indagar por información, alguna preparación e inversión adicional, antes de estar seguro de tomar la mejor decisión.

CÓMO COMPETIR LIBREMENTE CON EL CABLE-MÓDEM, (Joel Bendersky Gerente General Inter.Net Chile)

<< El negocio de proveer conectividad a Internet no deja de ser complicado. Hay que adaptarse rápidamente a los requerimientos de los usuarios, que en este mercado, quizás más que en ningún otro, cambian casi a la velocidad de la luz. Lo que está claro es, que lo que las personas quieren hoy es banda ancha, al mejor precio posible y con las mejores condiciones técnicas.

¿Cómo se puede favorecer el logro de ese objetivo? Aumentando la competencia entre proveedores, y de esa manera obligando a cada cuál a dar el mejor servicio que le sea posible. Sin embargo, hoy esa situación se ve afectada por un hecho puntual, que impide la sana y libre competencia entre los distintos actores de este sector.

Tres son las tecnologías a través de las cuales, se ofrece Internet de Banda Ancha al usuario final hoy en el mundo: el xDSL, cuya modalidad más conocida es el ADSL, que usa los pares telefónicos de cobre; el Cable Coaxial, a través del cual se entregan los servicios de televisión pagada; y el WLL (Wireless Local Loop), tecnología inalámbrica fija, que igualmente sirve para ofrecer servicios de telefonía.

Sin embargo, en Chile hoy eso significa que existen dos tipos de proveedores: los tradicionales ISPs, que ofrecemos todos servicios de ADSL; el WLL, que ofrece solamente una compañía, pero cuya salida al mercado es aún incipiente; y las compañías de televisión por cable.

Esta situación, que a ojos del público en general quizás parece normal, no lo es y de hecho afecta al mercado Internet en Chile en general. Lo que sucede es que, hoy, los operadores de cable cuentan con redes propias, con una infraestructura que usa cada uno para ofrecer, en forma exclusiva, ese servicio, sin que los ISPs tengamos ninguna posibilidad de entrar al negocio de la conectividad a través del cable.

El tema es bastante delicado, ya que lo que estamos enfrentando hoy es, una suerte de "duopolio" o "cartel", donde dos compañías acaparan el cien por ciento de la capacidad de oferta de un determinado servicio.

Esta no es una situación nueva en el país. Hace un tiempo atrás pasaba algo similar en el tema de las telecomunicaciones, específicamente en el área de telefonía, asunto en el cual la Subtel se vio obligada a intervenir para terminar con esa situación. Hoy, la compañía, dueña de la propiedad de la mayoría de las redes telefónicas del país, subarrienda su infraestructura para que el resto de los proveedores ofrezcan tanto Internet como servicios de telefonía.

Compartir las redes con el resto de los proveedores es algo esencial, porque permite la competencia justa también en cuanto a variedad de servicios. Un proveedor que abarca varios servicios tiene la posibilidad de subvencionar alguno, ofreciendo así tarifas muy por debajo de la competitividad promedio del mercado.

Lo que hoy está claro es, que los dos operadores de cable con cobertura nacional que ofrecen conectividad entre sus servicios, ofrecen un servicio de banda ancha que es ostensiblemente más barato que cualquier plan ADSL del resto de los ISPs. Si a eso se suma que el servicio es ofrecido dentro de un paquete que incluye telefonía y televisión, y que por su contrato se restan costos al usuario de los otros servicios contratados (como el cargo fijo en el caso del teléfono), estamos frente a una evidente desigualdad de condiciones a la hora de competir.

Más allá de las características técnicas o de las ventajas de una tecnología sobre otra, de lo que se trata en realidad es de una correcta regulación del mercado, donde todos los actores puedan competir en igualdad de condiciones, y donde el usuario pueda decidir entre un servicio u otro, basándose sólo en factores como costo y calidad de atención al cliente.

De eso es lo que se trata finalmente. Porque un mercado sano, donde exista la libre competencia, beneficia tanto a los participantes en él como al usuario final. En ese sentido, alguien tendría que tomar las riendas del asunto, tal como se hizo en el pasado

con tecnologías dominantes, mientras al resto de los ISPs no nos queda más que seguir impulsando el negocio como lo hemos hecho hasta el momento.>>

10.3.1. CABLE MODEM <versus> ADSL

	Cable	ADSL
Canal hasta la central telefónica	Compartido para todos los usuarios de varios bloques de edificios.	Compartido para cada 10 usuarios.
Cableado adicional en el edificio	Necesario.	Sólo cableado interno. Aprovecha el cableado ya existente
Cobertura	Parcial. Concentrada en zonas urbanas.	80% del territorio nacional
Velocidad independiente del número de usuarios	No	Sí (garantizando el 10% de la velocidad contratada)
Seguridad	**Baja**, al compartir un mismo cable todos los vecinos pertenecientes a un área	**Alta**, al disponer de un cable independiente y exclusivo hasta la central
Velocidad		
Descendente hacia el PC	256 Kbps-1Mbps	256 Kbps-2 Mbps
Ascendente hacia Internet	128 Kbps-512 Kbps	128 Kbps-300 Kbps

10.3.2. AOC. Agrupación de Operadores de Cable

A continuación se listan los operadores de cables en España:

- ONO (Cantabria, zona de Levante y Andalucía atlántica)
- Retecal (Castilla, León)
- Able (Aragón).
- Canariastelecom
- Euskaltel (País Vasco).

- Madritel
- Menta (Cataluña).
- Mundo-r (Galicia).
- Reterioja.
- Retena (Navarra).
- Supercable (Andalucía, salvo las demarcaciones atlánticas).
- Telecable (Asturias).
- Telefónica Cable (Todas las demarcaciones).

La Agrupación de Operadores de Cable (AOC) tienen como accionistas a Endesa y Unión Fenosa, socios de Retevisión. Forman parte de la misma Telecable, Retecal, Supercable, Grupo Gallego de Cable-MundoR, Retena y Reterioja, Aragón de Cable-Able, CTC-Menta, Madritel y Cabletelca-Canarias Telecom.

APENDICES

1. HISTORIA DE LAS REDES DE COMPUTADORES. NORMALIZACION

1.1. ARPANET

ARPANET (Red de la Agencia de Proyectos de Investigación Avanzada) es la creación de ARPA, que es la Agencia de Proyectos de Investigación Avanzada del Departamento de Defensa de EEUU. Su programa, iniciado en los últimos años de la década de los 60, comenzó por estimular la investigación en temas relacionados con redes de ordenadores, mediante la canalización de recursos a los departamentos de ciencias de la computación de varias Universidades de Estados Unidos, así como a algunas compañías privadas. Esta investigación produjo una red experimental de cuatro nodos, que se dio a conocer públicamente en diciembre de 1969. Desde entonces, creció en forma substancial, hasta llegar a tener varios centenares de hosts, cubriendo casi la mitad de la Tierra. En 1983, una vez demostrada su capacidad para establecer un servicio fiable de comunicaciones, ARPA cedió la administración de la red a la DCA

(Defense Communications Agency), para que la utilizase como una red operacional. Lo primero que hizo la DCA fue separar la parte militar en una subred separada, llamada MILNET, con fuertes restricciones para su acceso desde otras redes externas. En 1990, fue sustituida por otras redes que ella misma había creado, de forma que fue cerrada y desmantelada, aunque MILNET sigue operativa.

A comienzos de los años 60, Paul Baran había sugerido la idea de la conmutación de paquetes frente a la conmutación de circuitos propia de las líneas telefónicas. ARPA decidió que esta novedosa solución debía ser la base para las comunicaciones entre los ordenadores militares dado que resultaba más segura en caso de ataque, pues la destrucción de un nodo de comunicaciones no implicaría la interrupción automática de las mismas. Por ello, la red que se desarrollara debía ser una red de conmutación de paquetes, formada por una subred y unos host que la utilizan.

La subred estaba formada por una serie de minicomputadores llamados IMP (Interfaz Message Processors) conectados entre sí por líneas de transmisión de datos. Para mayor seguridad, cada IMP debía estar conectado al menos a otros dos, de esta forma si alguna línea o algún IMP resultaba destruido, los mensajes continuarían circulando por caminos alternativos.

Cada nodo de la red consistiría en un IMP y un host, en la misma habitación y conectados por un cable que permitiese comunicaciones fiables a alta velocidad. Un host podría enviar mensajes a un IMP de hasta 8063 bits. El IMP lo fragmentaría en trozos menores de 1008 bits y los enviaría de forma independiente hacia su destino. Cada paquete debía ser recibido entero antes de que un nodo intermedio lo reenviase hacia el destino final.

ARPA seleccionó a BBN, una empresa de Massachusetts, para que construyera la subred en diciembre de 1968. BBN eligió un modelo modificado de los DDP-316 de Honeywell, con 12K palabras de 16-bits como memoria principal para utilizarlos como IMP. Los IMP no tenían discos, ya que las partes móviles se consideraban poco fiables. Los IMP estaban conectados entre sí por líneas alquiladas de 56 Kbps.

El software se dividió en dos partes: el host y la subred. El software de la subred incluía los protocolos de comunicación entre dos IMP consecutivos y entre IMP origen - IMP destino. El software del host se encargaba de las comunicaciones host - IMP, host - host, y el software de aplicación.

Para resolver el problema del software del host, ARPA convocó un encuentro entre investigadores, la mayor parte estudiantes de graduado. Los estudiantes esperaban encontrar a algún experto en redes para que les explicase el diseño de las mismas y de su software, para después asignar a cada uno una parte del trabajo. La realidad es que no hubo ningún experto, y ellos mismos tuvieron que hacer todo el trabajo.

Sin embargo, una primera red experimental comenzó a funcionar a finales de 1969 con cuatro nodos: UCLA, UCSB, SRI y UTAH. Se eligieron estas cuatro universidades por el número de contratos que ya tenían con ARPA, y además porque sus ordenadores de proceso eran totalmente incompatibles entre sí. La red creció rápidamente y se añadieron más IMP. En menos de tres años estaba extendida por todo Estados Unidos.

Posteriormente, el software de los IMP se modificó para permitir la conexión de terminales a los IMP, sin necesidad de un host intermedio. A este tipo de IMP se les denominó TIP (Terminal Interfaz Processor). También se permitió la conexión de varios hosts a un mismo IMP para ahorrar dinero, la conexión de un host a varios IMP para aumentar la seguridad y la separación entre host e IMP.

Para favorecer la difusión de ARPANET, ARPA también financió la investigación sobre redes vía satélite y redes vía radio. Llegado este punto, se concluyó que los protocolos de que se disponían no eran los más adecuados para enfrentarse a redes heterogéneas. Como consecuencia se buscaron nuevos protocolos, lo que culminó con la propuesta en 1974 de TCP/IP por parte de Cerf y Kahn. TCP/IP estaba específicamente concebido para la comunicación entre diversos tipos de redes. Esto favoreció que nuevas redes se incorporasen a ARPANET.

Para facilitar la difusión de estos protocolos ARPA financió a BBN y la Universidad de California en Berkeley para que los integrasen el Unix de Berkeley. Se

crearon así los sockets, como interfaz del sistema con la red, y escribieron muchas aplicaciones, utilidades y programas de administración para facilitar su uso.

El momento fue el idóneo, coincidió con la compra de nuevos VAX en muchas universidades y redes locales para interconectarlos, pero no tenían el software. La aparicióndel Unix BSD 4.2 fue providencial, y su uso se generalizó rápidamente. Es más con TCP/IP era fácil conectar la LAN a ARPANET. La expansión de la red hizo necesario crear un nuevo protocolo para organizar las máquinas en dominios y mapear los nombres de las máquinas con sus direcciones IP. El nuevo protocolo fue DNS (Domain Naming System).

1.2. NSFNET

A finales de los 70, NSF (la Fundación Nacional para la Ciencia de Estados Unidos) se fijó en el enorme impacto que ARPANET estaba teniendo sobre la investigación universitaria, permitiendo que investigadores de todo el país compartiesen datos y colaborasen en proyectos de investigación. Sin embargo, para conectarse a ARPANET, la universidad debía tener algún contrato de investigación con el Departamento de Defensa. Esta dificultad para el acceso a ARPANET llevó a NSF a crear una red virtual, llamada CSNET (Red de Ciencias de la Computación) entorno a una máquina de BBN que tenía líneas módem y conexiones a ARPANET. Usando CSNET, los investigadores podían llamar y dejar correo electrónico para que otros los leyesen más tarde. Era simple, pero funcionaba.

Hacia 1984 NSF comenzó el diseño de una red de alta velocidad que sucediese a ARPANET, y estuviese abierta a todos los grupos de investigación universitarios. Para comenzar, NSF estableció una red base que conectase sus seis centros de supercomputación. El software sobre el que corrían las comunicaciones fue TCP/IP desde el comienzo.

NSF financió la creación de diversas redes regionales conectadas a NSFNET y constituyó la base para intercomunicar universidades, centros de investigación,

bibliotecas y museos. NSFNET tenía también conexiones con ARPANET. El éxito fue inmediato.

A medida que la red fue creciendo, NSF se dio cuenta de que no podría seguir financiando el servicio para siempre. Además, existían empresas que deseaban conectarse a NSFNET pero lo tenían prohibido debido las restricciones impuestas por NSF. De esta forma, NSF animó a MERIT, MCI e IBM a formar una corporación sin ánimo de lucro, ANS, como paso intermedio hacia la comercialización de la red. En 1990, ANS se hizo cargo de NSFNET y actualizó los enlaces de 1.5 Mbps a 45 Mbps formando ANSNET.

En 1991, el Congreso de Estados Unidos autorizó la financiación de NREN, el sucesor de NSFNET para la investigación, para su funcionamiento a velocidades de Gigabits. El objetivo es tener una red nacional a 3 Gbps antes del próximo siglo. Es un prototipo de la pretendida superautopista de la información.

1.3. USENET

Cuando apareció el Unix por primera vez, y se utilizó ampliamente en los laboratorios Bell, los investigadores descubrieron que necesitaban una forma de copiar archivos de un sistema Unix a otro. Para resolver este problema, escribieron el ***uucp*** (Unix to Unix Copy). A medida que los sistemas Unix adquirieron módems de llamada automática, fue posible copiar archivos entre máquinas distantes, mediante el programa uucp, de forma automática. Vino el surgimiento de redes informales, en las que una máquina central con un marcador telefónico automático se encargaba de llamar a un grupo de máquinas, durante la noche, para acceder y transferir archivos y correo electrónico entre ellas. Dos máquinas que tuviesen módems, pero sin llamada automática, podían comunicarse al hacer que la máquina central llamara a la primera, cargase los archivos y correo pendientes, y luego llamase al destino para descargarlos.

Estas redes crecieron muy rápido debido a que todo lo que se necesitaba para que uno se uniera a la red, era el sistema UNIX con un modem, algo que prácticamente cualquier departamento de ciencias de la computación tenía. Estas redes, se unieron para

formar una sola red que se denominó UUCP, constituida por aproximadamente 10.000 máquinas y un millón de usuarios.

La rama europea correspondiente se denominó EUNET y disponía de una estructura más organizada. Cada país europeo tenía una sola máquina de entrada operada por un único administrador. Los administradores mantienen un contacto permanente para administrar el tráfico de la red. Todo el tráfico internacional circula entre los puntos de entrada de los diferentes países. La conexión con Estados Unidos se hacía a través de un enlace entre Amsterdam y Virginia. También existían ramas en Japón, Corea, Australia y otros países.

El único servicio que ésta red ofrecía es el correo electrónico, pero una red similar llamada USENET, que se creó entre las universidades de Duke y Carolina del Norte, ofrecía un servicio de noticias. En la práctica todas las máquinas de EUNET y UUNET disponen de ambos servicios, por ello, se suele utilizar el nombre de USENET para referirse a todas ellas.

En el servicio de ***news***, se establecen infinidad de grupos de noticias a los que puede subscribirse cualquier usuario. Algunos grupos son de tipo técnico, aunque otros están relacionados con hobbies, política, ... Cada usuario puede poner mensajes en los grupos a los que está subscrito y leer los enviados por los demás. Estos mensajes se copian mediante uucp y se distribuyen a todas las máquinas que actúan como servidores.

1.4. INTERNET

El número de redes, máquinas y usuarios conectados a ARPANET creció rápidamente después de que TCP/IP se convirtiese en el protocolo “oficial”. Cuando NSFNET y ARPANET se interconectaron, el crecimiento se hizo exponencial. Hacia mediados de los 80, se comenzó a ver todo este conjunto de redes y subredes como la Internet, aunque no hubo ningún acto oficial que inmortalizase el momento.

El crecimiento ha seguido siendo exponencial, y hacia 1990 Internet contaba ya con 3000 redes y 200.000 ordenadores conectados. En 1992, se llegó al millón de hosts. En 1994 se estimó que el número de hosts se duplicaba cada año. El pegamento que une todas estas redes es el modelo de referencia TCP/IP junto con sus protocolos.

Pero, ¿qué significa estar en Internet?. Podemos considerar que una máquina está en Internet si corre los protocolos del modelo TCP/IP, tiene una dirección IP, y la capacidad de enviar paquetes IP a otras máquinas que tienen las mismas características. El concepto queda oscurecido por el hecho de que muchos ordenadores personales tienen la capacidad de conectarse a servicios de Internet a través de un intermediario mediante el uso del modem.

Con la expansión sufrida, no es posible administrar la red con el estilo informal con que se hacía. En 1992, se fundó la Internet Society para promover el uso de Internet e incluso poder hacerse cargo de su administración.

Las cuatro aplicaciones básicas de Internet son:

1. **Correo Electrónico**
2. **Servicio de noticias**
3. **Login remoto:** Telnet, Rlogin, ...
4. **Transferencia de ficheros**

Hasta comienzos de los 90, Internet era usada fundamentalmente por las universidades, organismos gubernamentales y algunas compañías con fuertes departamentos de investigación. La aparición de una nueva aplicación, el World Wide Web lo cambió todo y atrajo a millones de usuarios. Esta aplicación desarrollada en el CERN, no cambiaba los servicios básicos, sino que simplemente facilitaba su uso sin más que usar el ratón.

1.5. NOVELL NETWARE

Es la red local para ordenadores personales más extendida del mundo. Se diseñó para su uso en compañías que sustituían sistemas basados en mainframes por grupos de ordenadores personales. Cada usuario posee un PC que hace las veces de cliente de otros más potentes, que actúan como servidores de ficheros, de bases de datos, ofrecen colas de impresión, etc.

NetWare usa una arquitectura de red propia, basada en el antiguo sistema XNS de Xerox. Esta arquitectura, anterior a OSI, es más parecida a TCP/IP. De hecho, consta de 5 capas, con funciones similares a los de TCP/IP, pero el conjunto de protocolos es distinto. Las capas física y de enlace se pueden elegir de entre varios estándares como Ethernet, TokenRing o ARCnet. Sobre ellos, define un nivel de red en el que usa el protocolo IPX que proporciona un servicio sin conexión no fiable. Su funcionalidad es muy similar a IP.

Sobre IPX, en la capa de transporte se dispone de un protocolo orientado a conexión y libre de errores, llamado NCP y que es el núcleo fundamental de NetWare. Hay otro protocolo que sólo proporciona servicios de datagramas, que es el SPX. Otra posible opción es el uso de TCP. Cada aplicación de la capa superior (transferencia de ficheros, anuncio de servidor, correo, ...) puede elegir el servicio de transporte que desea utilizar.

1.6. Normalización de las Redes de Ordenadores

Existen muchos fabricantes y suministradores de redes de ordenadores, cada uno con sus propias ideas sobre como deben funcionar las comunicaciones entre ordenadores. Por ejemplo, IBM tenía más de una docena de protocolos propios. Esta situación hacía que fuese difícil construir redes de ordenadores si éstos pertenecían a distintos fabricantes. El caos generado por esta situación dio lugar a la exigencia de que se estableciesen normas.

El objeto de la normalización no solo era facilitar la interconexión de equipos diferentes, sino lograr un incremento del mercado para aquellos productos que se

acogiesen a la norma, lo que conduciría a una economía de escala que permitiría la reducción de costes y con ello un mercado aún mayor.

Las normas se dividen en dos categorías que pueden definirse como: de facto y de jure. Las normas ***De Facto***, son aquellas que se han establecido sin ningún planteamiento formal. Por ejemplo, las normas IBM-PC y sus sucesoras son normas de hecho porque docenas de fabricantes decidieron copiar fielmente las máquinas que IBM sacó al mercado.

Por el contrario, las normas ***De Jure*** (de derecho), son normas formales, adoptadas por un organismo que se encarga de su normalización. Las autoridades internacionales encargadas de la normalización se dividen, por lo general, en dos clases: la establecida por convenio entre gobiernos nacionales, y la establecida voluntariamente sin un tratado entre organizaciones. En el área de normas de redes de ordenadores, existen dos organizaciones principales, de cada uno de los dos tipos.

1.6.1. Quién es quién en el mundo de las comunicaciones

El status legal de las compañías telefónicas en el mundo varía considerablemente de un país a otro. En un extremo está Estados Unidos que tiene unas 1500 compañías distintas, todas ellas privadas. Antes de su fragmentación en 1984, AT&T era la mayor de estas compañías, prestando servicio al 80 % de la población de Estados Unidos y cubriendo más de la mitad de su área geográfica. Las demás compañías daban servicio al resto de usuarios, principalmente en áreas rurales. En el otro extremo, están los países en los que el gobierno detenta un monopolio sobre las comunicaciones, como suele suceder en muchos países europeos.

Es clara la necesidad de que los servicios de comunicación sean compatibles a escala mundial, para asegurar que la gente (y los ordenadores) de un país pueden comunicarse con los de otro país diferente. Esta coordinación la ofrece una agencia de las Naciones Unidas llamada, ***ITU*** (Unión Internacional de Telecomunicaciones). La ITU tiene tres órganos principales, dos de ellos se ocupan sobre todo de la difusión

internacional de radio y el otro está fundamentalmente relacionado con sistemas telefónicos y de comunicaciones de datos.

A este último grupo se le conoce como ***CCITT*** (Comité Consultivo Internacional Telegráfico y Telefónico). El CCITT tiene cinco clases de miembros:

- Miembros A, que son las compañías telefónicas nacionales, o los ministerios de telecomunicaciones.
- Miembros B, que son los reconocidos como administraciones privadas (por ejemplo AT&T).
- Miembros C, que son las organizaciones científicas e industriales.
- Miembros D, que corresponden a otras organizaciones internacionales.
- Miembros E, que corresponden a aquellas organizaciones cuya misión fundamental está en otro campo, pero que están interesadas en el trabajo de la CCITT.

De esta clasificación, sólo los miembros de tipo A tienen derecho a voto.

La tarea del CCITT consiste en promover las recomendaciones técnicas sobre aspectos telefónicos, telegráficos e interfaces de comunicación de datos. Esta labor ha producido normas que tienen un reconocimiento internacional como por ejemplo la norma V.24 (EIA RS-232 en Estados Unidos), y la norma X.25 que especifica la interfaz entre un ordenador y una red de ordenadores (conmutación de paquetes).

1.6.2. Quién es quién en el mundo de las normas

Las normas internacionales son producidas por la ISO (International Standards Organization), que es una organización voluntaria, fuera de tratados y fundada en 1946, cuyos miembros son las organizaciones nacionales de normalización correspondientes a los 89 países miembros, y otros 85 organismos.

La ISO emite normas en una gama amplia de temas, que van desde las tuercas y los tornillos, hasta los recubrimientos de los postes telefónicos. La ISO tiene casi 200

comités técnicos (TC), cuyo orden de numeración se base en el momento de su creación, ocupándose cada uno de ellos de un tema específico. Por ejemplo, TC1 está relacionado con temas relativos a tuercas y tornillos, mientras que el TC 97 está relacionado con ordenadores y procesamiento de información. Cada uno de los TC tiene subcomités (SC), los cuales se dividen a su vez en grupos de trabajo (WG).

Los WG, constituidos por unos 100.000 voluntarios distribuidos en todo el mundo, son los que realizan el trabajo. Varios de estos "voluntarios" son por lo general asignados por las propias compañías, representantes de gobiernos nacionales o expertos provenientes del mundo académico.

La ISO y el CCITT algunas veces cooperan (de hecho, ISO es un miembro de clase D del CCITT), con respecto a la emisión de normas sobre telecomunicaciones, con objeto de evitar el absurdo de dos normas internacionales oficiales, mutuamente incompatibles.

El procedimiento que utiliza la ISO para el establecimiento de normas, está diseñado para conseguir el mayor consenso posible. El proceso comienza cuando alguna de las organizaciones nacionales considera necesario el establecimiento de una norma internacional. Entonces, se forma un grupo de trabajo que llega a plantear una propuesta de trabajo (DP). Una vez que se genera la DP se hace circular entre todos los miembros, los cuales cuentan con seis meses, a partir de ese momento, para plantear sus comentarios y críticas. Si una mayoría significativa aprueba la propuesta, se produce un documento revisado, denominado DIS (Anteproyecto de Norma Internacional), el cual se hace circular nuevamente con objeto de tener más comentarios y realizar una votación al respecto. Con base en los resultados de esta vocación, se prepara, aprueba y publica el testo final de la IS (norma internacional). En algunas de las áreas, en donde existe una gran polémica, la DP o DIS probablemente tenga que pasar por varias versiones, en su planteamiento, antes de adquirir el número de votos necesarios para su aprobación. El proceso completo puede llevar varios años.

Existen otros organismos que también establecen normas a distintos niveles. Por ejemplo NIST (National Institute of Standards and Technology) de Estados Unidos se

encarga de establecer normas de obligado cumplimiento para las adquisiciones que realiza el gobierno de Estados Unidos, con excepción de las que realiza directamente el ministerio de Defensa, que tiene sus propias normas (normas MIL).

Otro participante importante en el mundo de las normas es el IEEE, que es la organización profesional más grande del mundo. Esta institución, además de publicar numerosas revistas y programar un número muy importante de conferencias anuales, ha establecido un grupo dedicado al desarrollo de normas en el área de ingeniería eléctrica y computación. La norma 802 del IEEE, para una red de área local, es la norma clave para el desarrollo de las LAN. Posteriormente, fue adoptada por la ISO como base para la norma ISO 8802.

1.6.3. Quién es quién en los standard de Internet

Internet tiene sus propios mecanismos de standarización, diferentes de los del CCITT y la ISO. De forma sencilla, podemos decir que los participantes en los encuentros de ITU o de la ISO llevan trajes. Las personas que llegan a las reuniones para standarización de Internet llevan vaqueros o uniformes militares.

ITU-T e ISO están pobladas por funcionarios y representantes de las grandes empresas que han hecho de la standarización su trabajo. Por el contrario, la gente relacionada con Internet buscan un acuerdo para que las cosas funcionen, pero sin que sea un fin en sí mismo.

Cuando se creo ARPANET, el departamento de defensa creó una comité informal para su desarrollo. En 1983, el comité se renombró y se denominó ***IAB*** (Internet Activities Board). Recibió una serie de encargos adicionales cuyo objetivo básico era lograr que los investigadores involucrados en ARPANET e Internet avanzan en la misma dirección. Posteriormente, el acrónimo “IAB” se cambió por Internet Architecture Board.

Cada uno de los diez miembros del IAB encabeza un grupo de trabajo (task force) sobre algún aspecto de especial relevancia. El IAB tiene varias reuniones al año para

discutir resultados y comunicarlos al ministerio de Defensa y el NSF. Cuando se necesita un standard, el IAB elabora el nuevo standard y lo distribuye para que se elaboren distintas implementaciones. Las comunicaciones se realizan en forma de ***RFC*** (Request For Comments). Las RFC se encuentran disponibles a través de la red y pueden ser consultadas por cualquiera. Su numeración sigue un estricto orden cronológico y en la actualidad es de unas 2000.

Con la difusión de Internet, esta forma de trabajo no era efectiva. En 1989, el IAB se reorganizó de nuevo. Los investigadores formaron el IRTF (Internet Research Task Force), y al IETF (Internet Engineering Task Force), ambos dependientes del IAB. El IAB se amplió para incluir representantes de otras organizaciones. El IRTF se debe hacer cargo de la investigación a largo plazo, mientras que el IETF debe resolver los problemas técnicos a corto plazo.

BIBLIOGRAFIA

[TANEMBAUM, 96]

Tanenbaum, A.S. (1996).

Computer Networks. (Third Edition).

Prentice-Hall.

[HALSALL, 95]

Halsall, F. (1995).

Data Communications, Computer Networks and Open Systems.

Addison-Wesley.

[FREER, 88]

Freer, J. (1988).

Introducción a la tecnología y diseño de Sistemas de Comunicaciones y Redes de Ordenadores.

Anaya Multimedia.

[STALLINGS, 97]

Stallings, W. (1997).

Comunicaciones y redes de computadores, 5ª ed.

Prentice Hall Iberia.

[COMER, 88]

Comer, D.E. (1988).

Internetworking With TCP/IP. Principles, Protocols, and Architecture.

Prentice-Hall.

[COMER, 91]

Comer, D.E. (1991).

Internetworking With TCP/IP. Vol II: Design, Implementation and Internals.

Prentice-Hall.

[TELEFÓNICA, 89]

Telefónica. (1989).

Manual estractado de operación, Red IBERPAC. Protocolo X.25

Telefónica/Formación.

[TELEFÓNICA, 90]

Telefónica. (1990).

Estructura y funcionamiento de la Red IBERPAC.

Telefónica/Formación.

[CORRALES, 95]

Corrales, J.A. ; Ojea, G. (1995)

La red corporativa de la Universidad de Oviedo.

Jornadas Técnicas de RedIris 1995.

[ALONSO, 95]

Alonso, J. M. (1995).

Protocolos de comunicaciones para sistemas abiertos.

Addison-Wesley Iberoamericana.

Apuntes de Comunicaciones I y II de la escuela de Informática de Sevilla

Por Adrián Ramírez.

[ACADEMIA DE NETWORKING DE CISCO SYSTEMS: GUIA DEL PRIMER AÑO (CCNA 1 Y 2) 3/E]

Editado por Cisco System

[INTERNET]: Artículos de interés sobre el Cable-módem. Sus ventajas, inconvenientes y comparativas con otros medios de transmisión.