



Universidad
de Huelva

DA
iESI

TERCER CURSO. REDES DE COMPUTADORES

Escuela Politécnica Superior
Universidad de Huelva

Departamento de Ing. Electrónica,
Sistemas Informáticos y Automática

TEMA 1

Fundamentos de comunicaciones y redes de datos

Manuel Sánchez Raya
Versión 0.1
1 de Octubre de 2003

ÍNDICE

1.- Introducción.	3
2.- Conceptos de comunicaciones.	4
2.1.- Redes en las empresas.	5
2.2.- Redes de ordenadores para el público en general.	5
3.- Funciones de los sistemas de comunicación.	7
3.1.- Nombres y direcciones.	7
3.2.- Fragmentación y reconstrucción de mensajes.	7
3.3.- Compactación.	8
3.4.- Establecimiento de conexiones y Multiplexación.	8
3.5.- Control de errores.	8
3.6.- Congestión y control de flujo.	8
3.7.- Sincronización.	9
3.8.- Gestión de Prioridades.	9
4.- Topología de las Redes de Comunicación.	10
4.1.- Tipos de Red por el tipo de tecnología.	10
4.2.- Tipos de Red por su extensión.	11
5.- Arquitectura de Red.	14
5.1.- Unidades de información transmitidas en la comunicación.	15
5.2.- El modelo de referencia OSI de ISO.	16
5.3.- El modelo de referencia TCP/IP.	24
5.4.- Comparación entre los modelos OSI y TCP/IP.	25
6.- Breve historia de las redes de comunicación.	26
6.1.- Arpanet.	26
6.2.- Nfsnet.	28
6.3.- Usenet.	28
6.4.- El nacimiento de Internet.	29
6.5.- Novell Netware.	30
7.- Estándares y Agencias de Normalización.	31
7.1.- Organizaciones de Estándarización en Comunicaciones.	32
7.2.- Agencias de Normalización Internacionales.	32
7.3.- Normas sobre Internet.	33

BIBLIOGRAFÍA:

Apuntes de Redes de Comunicaciones. Universidad de Oviedo.

Stallings, W.; "Comunicaciones y Redes de Computadores". 6ª Edición; Prentice-Hall; 2000 (681.324 STA com)

1.- Introducción.

El siglo XX ha estado dominado por la tecnología de la información, es decir, todos los aspectos relacionados con la recolección, procesamiento 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.

En 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 vía 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.

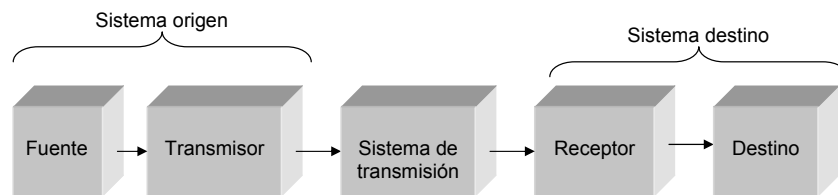
Llamaremos **sistema distribuido** a una red de ordenadores en la que la existencia de múltiples equipos autónomos es transparente para el usuario, es decir, no le resulta visible la división del sistema en distintos equipos.

2.- Conceptos de comunicaciones.

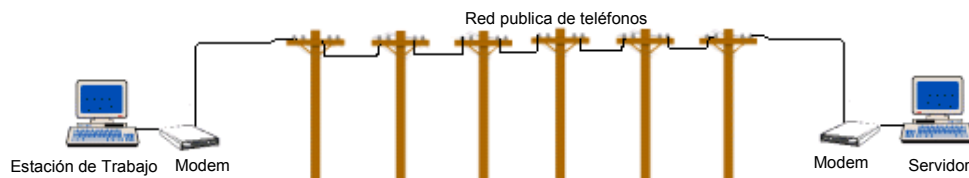
Desde un punto de vista de las comunicaciones, podríamos hacer las siguientes definiciones:

- **Dato:** Una entidad que transporta información.
- **Señal:** Es una codificación eléctrica o electromagnética de los datos.
- **Señalización:** Es el acto de propagar la señal a través de un medio adecuado.
- **Transmisión:** Se define como la comunicación de los datos mediante la propagación y el procesamiento de señales.
- **Comunicación entre computadores:** Se denomina así al intercambio de información entre ellos.
- **Redes de computadores:** Se denomina así al conjunto de ordenadores que se comunican entre sí mediante una red de comunicaciones.

Un sistema de transmisión de datos está formado por una fuente de los datos, generalmente un computador que extrae esos datos del usuario a través de un teclado o de disco, un transmisor que es el encargado de adecuar la información al sistema de transmisión empleado, realizando para ello la generación de la señal que se va a transmitir.

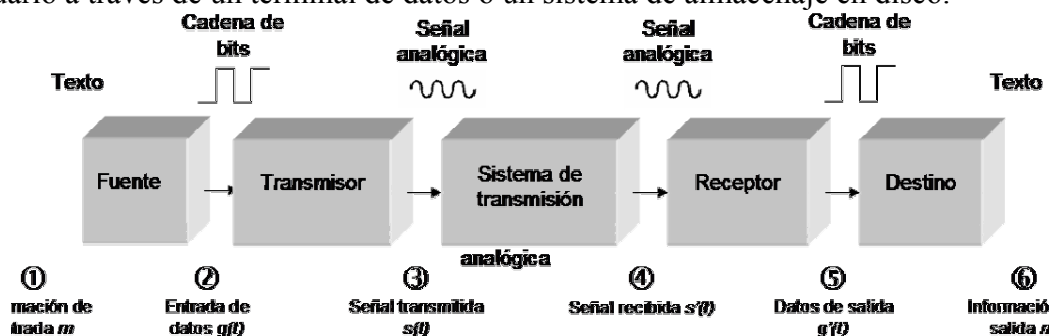


(a) Diagrama general de bloques



(b) Ejemplo

El receptor realiza el proceso inverso, extrayendo del medio o sistema de transmisión la señal distorsionada o degradada previsiblemente proveniente del transmisor y convirtiéndola mediante tareas de decisión en datos digitales. Estos datos digitales se deben interpretar en el Destino, por el receptor final del mensaje, que puede ser el usuario a través de un terminal de datos o un sistema de almacenaje en disco.



2.1.- Redes en las empresas.

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 el 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 substituir 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.

2.2.- Redes de ordenadores para el público en 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 desbancando 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 disponible es la participación en juegos a través de la red.

3.- 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 que 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.

3.1.- 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 que 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.

3.2.- 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 ser 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.

3.3.- 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 transparente al usuario.

3.4.- Establecimiento de conexiones y Multiplexación.

Para poder establecer una 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.

3.5.- 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.

3.6.- 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.

3.7.- 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 bloques:** Es necesario determinar el inicio y final de un bloque de bytes. La información contenida 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.

3.8.- Gestión de Prioridades.

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.

4.- Topología de las Redes de Comunicación.

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

4.1.- Tipos de Red por el tipo 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 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 de 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 forward). 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.

DIFUSIÓN	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. 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.

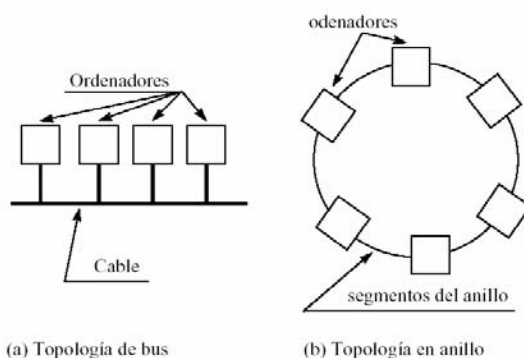
4.2.- Tipos de Red por su extensión.

Redes de Área 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 cuando 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.

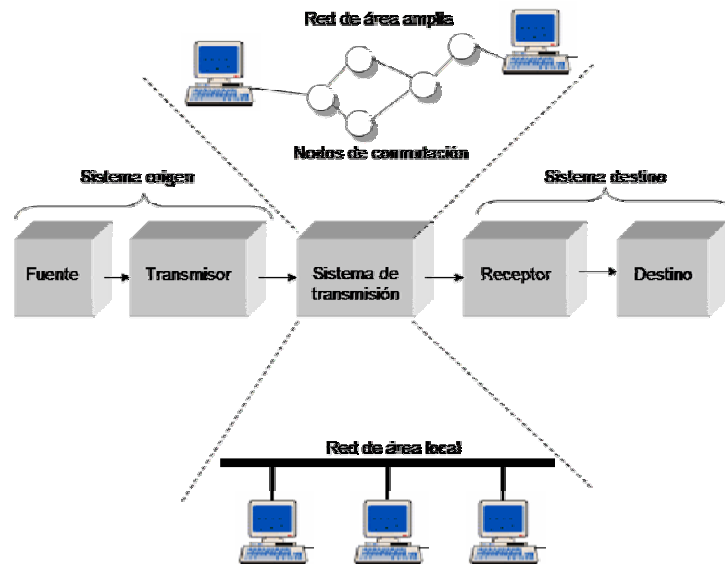
Redes de Área Extensa (WAN)

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.

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 ÁREA EXTENSA (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.

5.- Arquitectura de Red.

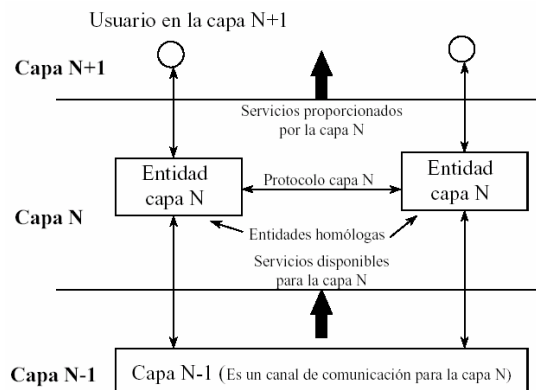
Para la interconexión de sistemas abiertos se construyen arquitecturas de red. Se define un **sistema abierto** como: "Un sistema capaz de interconectarse con otros de acuerdo con unas normas establecidas". La **Interconexión de Sistemas Abiertos**: "se ocupará del intercambio de información entre sistemas 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, sintácticas y de temporización) que gobiernan la comunicación entre entidades de una 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.

- El sistema de interconexión está formado por un conjunto de *entidades* situadas en diferentes *capas*.
- Las *entidades* de una determinada *capa* N cooperan entre sí de acuerdo con un determinado *protocolo* N.
- 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.
- Las *entidades* de una *capa* N realizan unas determinadas *funciones* N, utilizando los *servicios*.



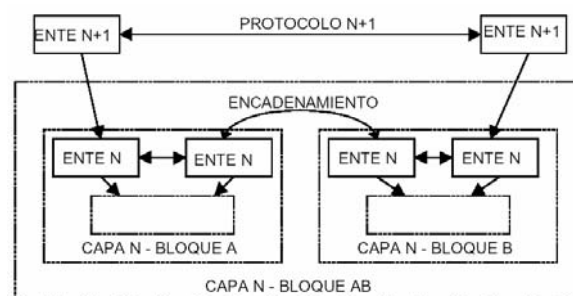
Una capa, la N, proporcionará a la capa inmediatamente superior, la N+1, una serie de servicios. Para ello puede usar los servicios ofrecidos por la capa N-1. Por ejemplo, a partir de un enlace físico con errores, se podría construir un enlace lógico libre de errores.

Una capa puede ofrecer más de una clase de servicio. Conviene dejar claro que:

- No todas las funciones que realiza una capa deben 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, formando una estructura modular.

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.



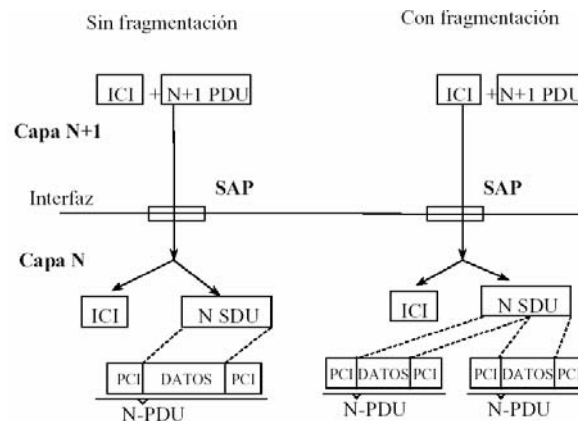
5.1.- Unidades de información transmitidas en la comunicación.

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 definirá 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. Estos trozos deberán ser reensamblados en el destino para obtener la SDU.



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.

5.2.- 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 tomando 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 si mismo, no es una arquitectura de red puesto que no especifica el protocolo que debe usarse en cada capa.

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.

Capa de enlace

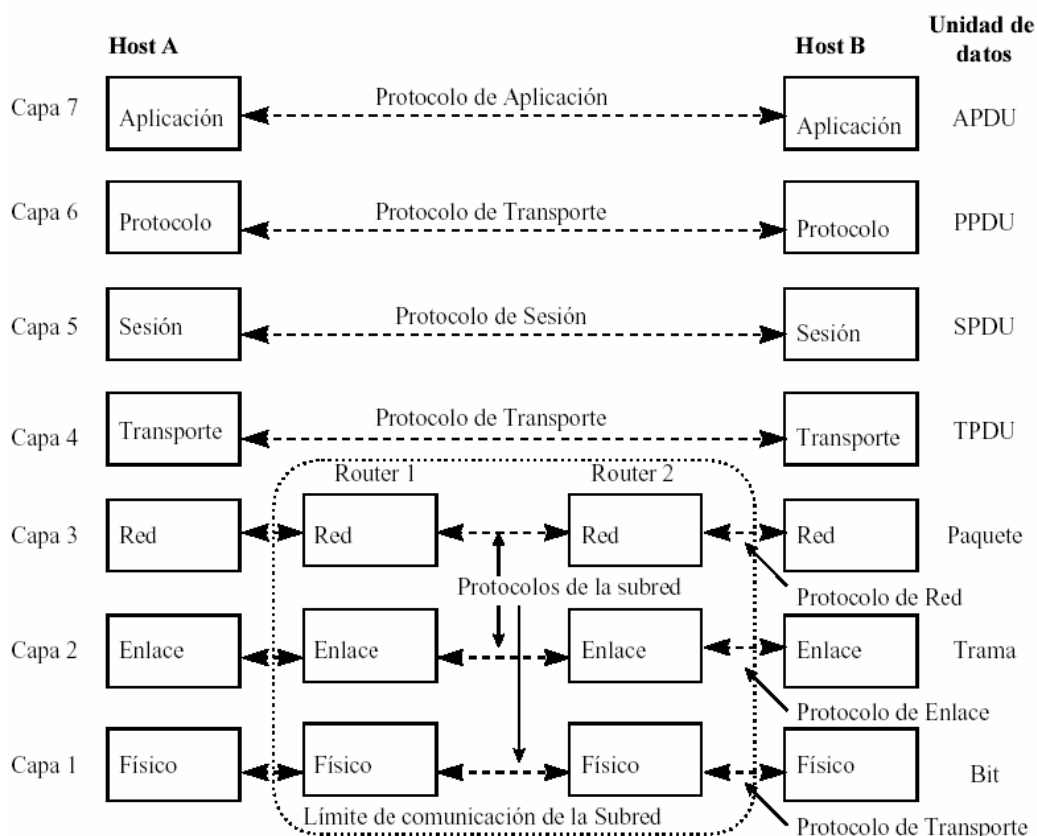
Su principal tarea consiste en proporcionar una línea sin errores a partir de un medio de transmisión cualquiera. Esta capa debe crear y reconocer los límites de las tramas. Además debe resolver los problemas creados por el deterioro, pérdida o duplicidad de tramas. La capa de enlace ofrece distintos 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.

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 como 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.



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.

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 que 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.

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.

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.

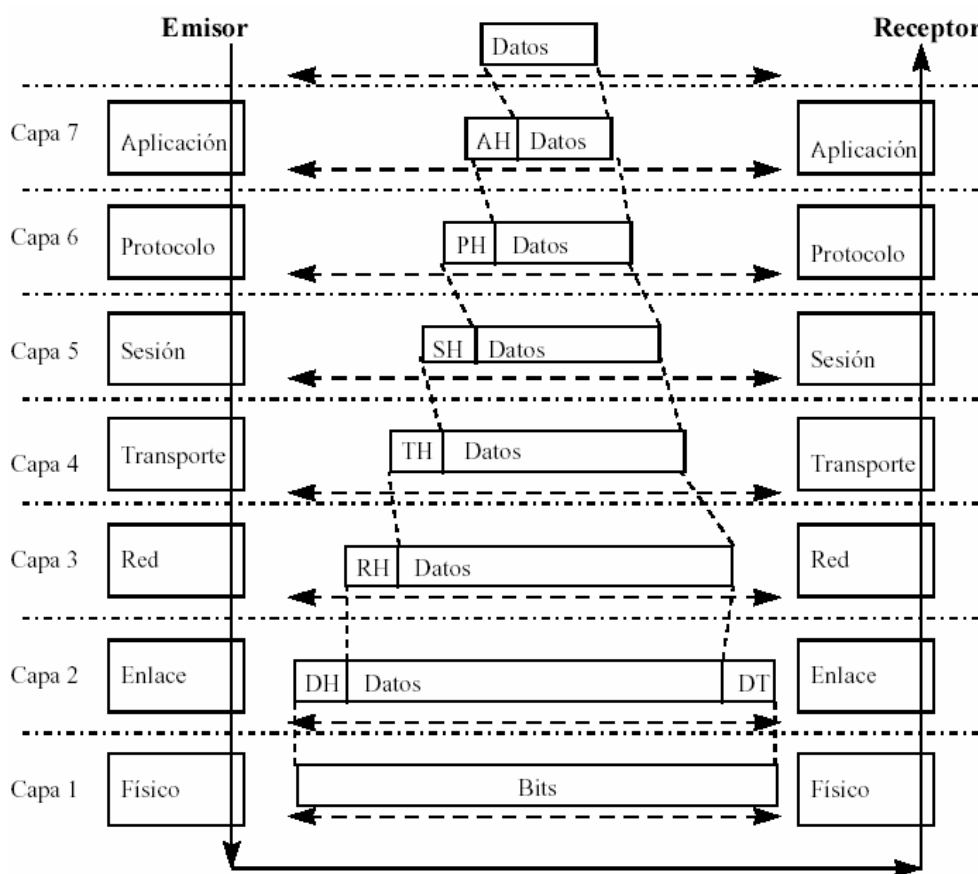
Transmisión de datos en el modelo OSI

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 que parte de los datos que recibió corresponden a la cabecera de la capa de aplicación y que 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 pueden ser fragmentados 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

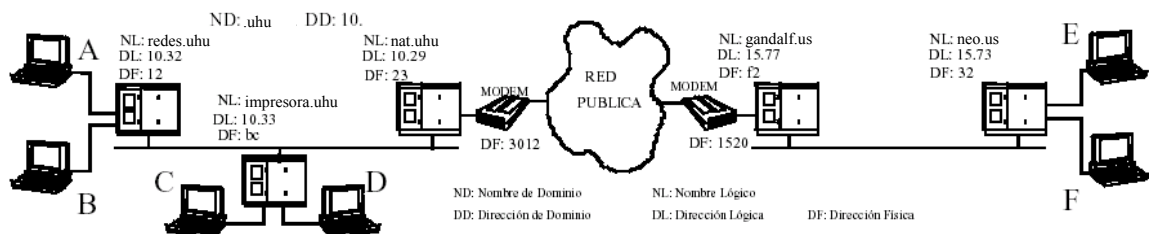
homóloga en la máquina de destino a través de un canal lógico proporcionado por el resto de capas en ambas máquinas.

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 es entrega a la zona del cerebro 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.

Ejemplo

A modo de ejemplo en las páginas siguientes se muestra como 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.



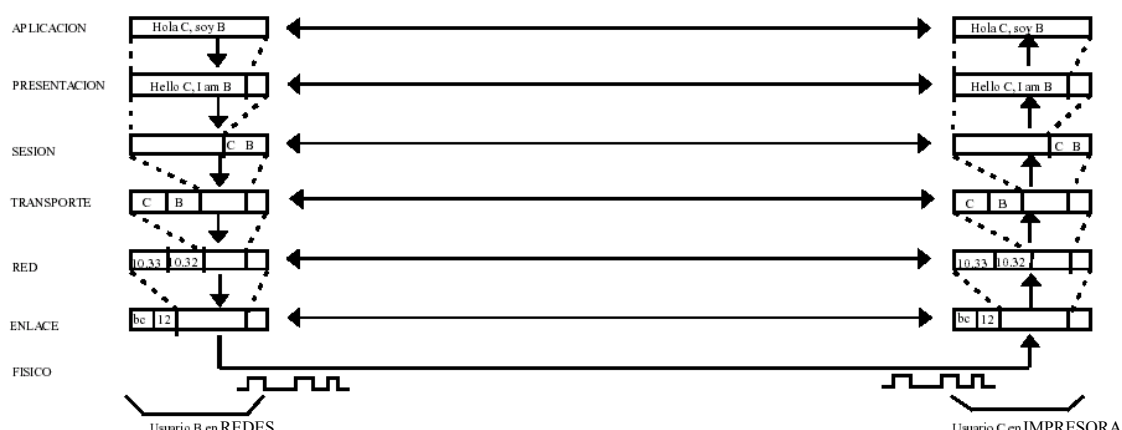
Por razones nemotécnicas a los nodos y dominios se les dan nombres que permitan recordar su denominación en la red fácilmente. Estos nombres están asociados a direcciones lógicas, que son las que realmente utiliza el sistema de comunicaciones para identificar cada nodo y dominio. Por lo general el nombre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

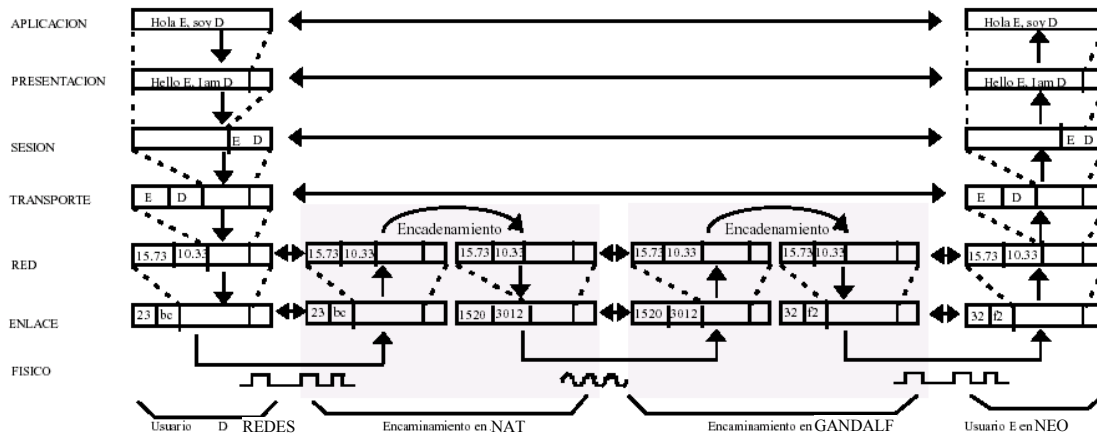
El mensaje es adquirido por la capa de aplicación, que se implementaría como el software necesario para recoger el mensaje del teclado del usuario del terminal B de “redes” y enviarlo por la red. Una vez obtenido el mensaje la aplicación lo entregaría al módulo o programa que implementa la capa de presentación, 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.



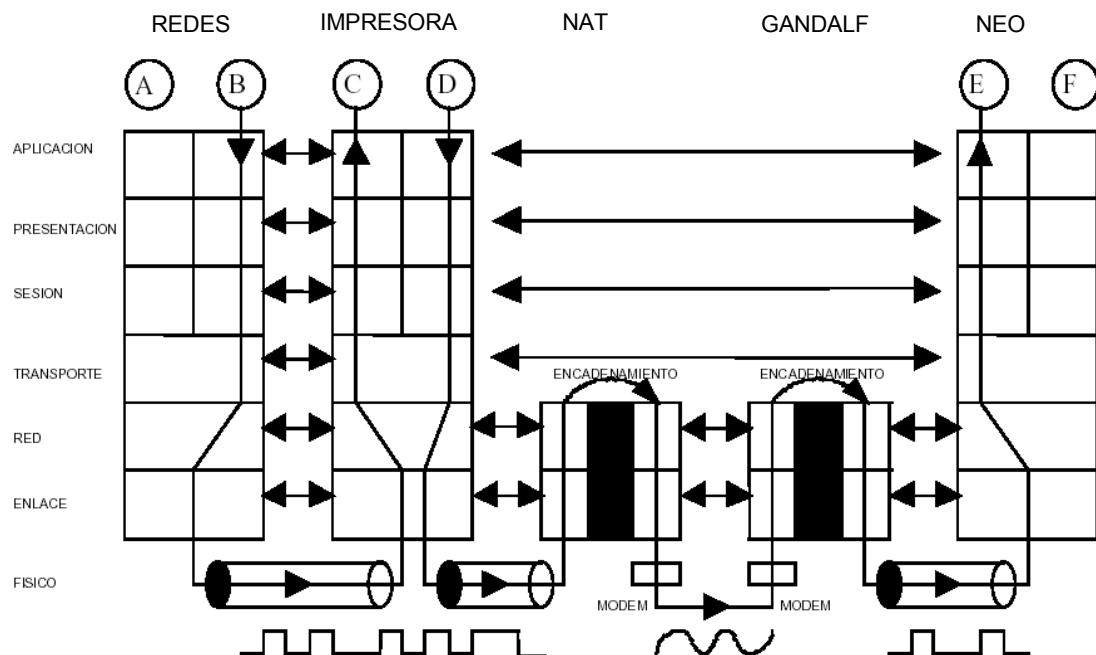
La capa de sesión mantiene la sesión de trabajo de cada usuario dentro de un mismo nodo, identificando a cada usuario para diferenciar su sesión de la de los demás. Todas estas sesiones convenientemente identificadas (generalmente mediante la identificación tanto del origen, B, como del destinatario, 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, sino la rechaza.



La capa de red comprueba la dirección lógica de destino, y si es la suya entrega los datos a la de transporte. Esta 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 en mensaje aún en la forma de representación de la red a la capa de presentación. Esta lo descomprime, descrypta 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, "redes.uhu" y "neo.us". 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.

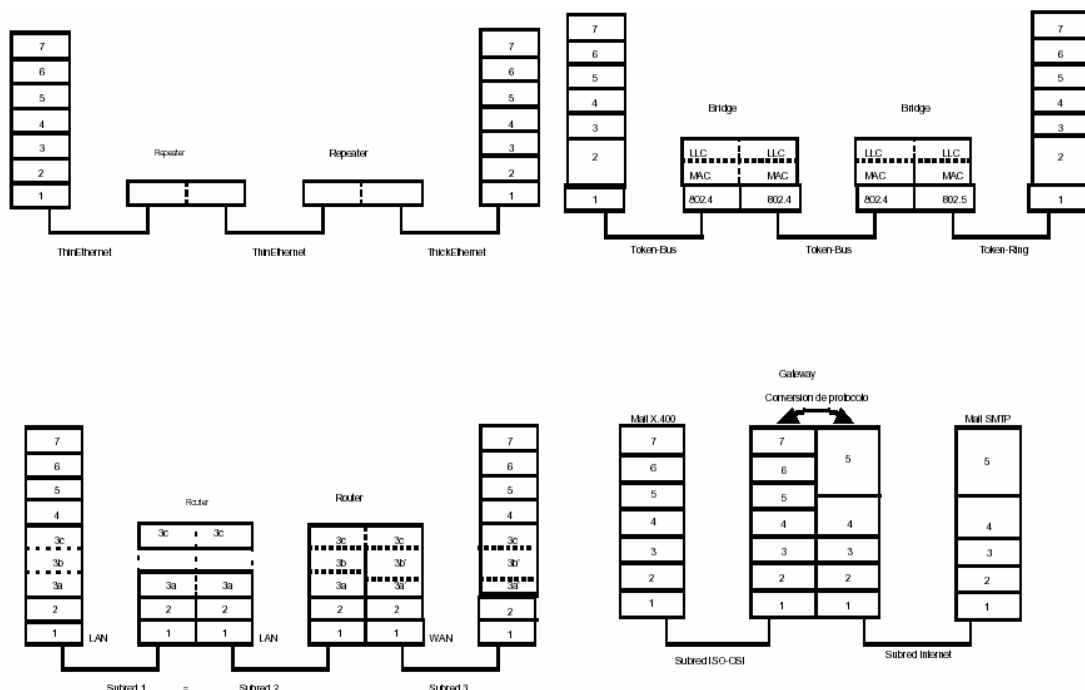
En principio todo el proceso es igual al anterior hasta que la información llega a la capa de red, encargada precisamente del encaminamiento entre subredes. Esta capa se encuentra con el problema de que si entrega la información a la capa de enlace indicando como destinatario la dirección física de "neo.us" (32), nadie en su subred

atenderá esa trama de datos. Sin embargo si conoce la dirección física en su red del nodo que le sirve de enlace con nodos de de otros dominios, “nat.uhu” (23) y a esa dirección física dirige la trama.

La trama es aceptada por la capa de enlace de “nat.uhu” pues está dirigida su dirección física. Pero cuando los datos llegan a la capa de red este detecta que la dirección lógica del destinatario no es la suya. Sin embargo, “nat.uhu” 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 que red y a que 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 “us”.

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 “gandalf.us”. Su capa de red detecta también una dirección lógica de destino distinta a la suya para realizar a continuación un proceso similar al de “nat”. Ahora los datos pasan de nuevo a unas capas relacionadas con la red local “us” (que puede ser un estándar diferente a la red que se utiliza en “uhu”) y son dirigidos, ahora si, a la dirección física del destinatario, “neo.us” (32). El proceso hasta llegar a la pantalla del usuario del terminal E es el ya descrito en el ejemplo anterior.

En la tercera figura se muestran los caminos seguidos por la información a través de las capas y se pueden observar los fenómenos de multiplexación sobre la capa de transporte y que esta es la primera de las capas que mantiene un diálogo extremo a extremo en la comunicación entre subredes. La cuarta figura muestra distintos dispositivos en los que se producen encadenamientos en distintas capas.

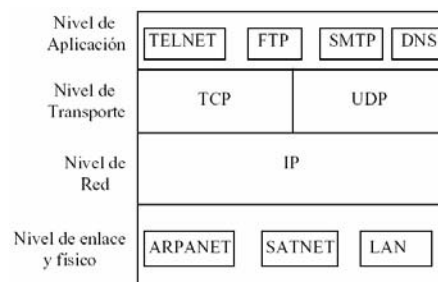


5.3.- El modelo de referencia TCP/IP.

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

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 estos 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 bastante limitados. Equivale a la capa de red del modelo OSI.

La 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 (Transmission 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 penso para aquellos casos en los que la capa de sesión necesitase un canal lógico distinto del que proporciona TCP.

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.

5.4.- Comparación entre 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 como funcionan las mismas.

La interfaz de una capa indica cómo acceder a los servicios que ofrece, pero tampoco dice nada sobre como 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.

6.- Breve historia de las redes de comunicación.

6.1.- 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 en 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 aparición de 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).

6.2.- Nfsnet.

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.

6.3.- 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 era 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 suscrito y leer los enviados por los demás. Estos mensajes se copian mediante uucp y se distribuyen a todas las máquinas que actúan como servidores.

6.4.- El nacimiento de 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 ejecuta 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 (news).**
3. **Login remoto:** Telnet, ssh, rlogin...
4. **Transferencia de ficheros (ftp).**

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, consistía en un programa para interfaz gráfica y se denominó Mosaic. En sí no cambiaba los servicios básicos, sino que simplemente facilitaba su uso sin más que usar el ratón. Luego llegó la lucha de los Navegadores de Internet, Internet Explorer, Netscape, Mozilla, Opera, etc, pero eso ya es otra historia...

6.5.- 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.

7.- Estándares y Agencias de Normalización.

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 planeamiento 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.

Las normalizaciones tienen las siguientes ventajas:

- Las normalizaciones aseguran un gran mercado. Se estimula la producción masiva y en algunos casos la utilización de alta y muy alta escala de integración lo que reduce mucho los costos.
- Un estándar permite que productos de diferentes suministradores se comuniquen entre sí, dotando al comprador de mayor flexibilidad en la selección y uso de los equipamientos.

Pero por otro lado tienen las siguientes desventajas:

- Los estándares tienden a congelar la tecnología. Mientras un estándar se desarrolla, se revisa y se adopta, se habrán desarrollado otras técnicas más eficaces.
- Hay varios estándares para la misma función. Recientemente, las organizaciones dedicadas a desarrollar estándares han comenzado a cooperar más estrechamente para que esto no suceda.

7.1.- Organizaciones de Estándarización en 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, **UIT** (Unión Internacional de Telecomunicaciones). La UIT 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 **UIT-T** o **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.

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).

7.2.- Agencias de Normalización Internacionales.

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 votación, se prepara, aprueba y publica el texto 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.

7.3.- Normas sobre Internet.

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

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

Cuando se creó ARPANET, el departamento de defensa creó un 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 avancen 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 al 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.