Transmisión de Datos en Internet

Indice

1. Introducción

La gran rapidez con la que Internet se ha expandido y popularizado en los últimos años ha supuesto una revolución muy importante en el mundo de las comunicaciones, llegando a causar cambios en muchos aspectos de la sociedad. Lo que se conoce hoy como Internet es en realidad un conjunto de redes independientes (de área local y área extensa) que se encuentran conectadas entre si, permitiendo el intercambio de datos y constituyendo por lo tanto una red mundial que resulta el medio idóneo para el intercambio de información, distribución de datos de todo tipo e interacción personal con otras personas.

2. Transmisión De Datos En Internet

Una red de ordenadores permite conectar a los ordenadores que la forman con la finalidad de compartir información, como documentos o bases de datos, o recursos físicos, como impresoras o unidades de disco. Las redes suelen clasificarse según su extensión en:

- LAN (Local Area Network): Son las redes de área local. La extensión de este tipo de redes suele estar restringida a una sala edificio, aunque también podría utilizarse para conectar dos más edificios próximos.

- WAN (Wide Area Network): Son redes que cubren un espacio muy amplio, conectando a ordenadores de una cuidad o un país completo. Para ello se utilizan las líneas de teléfono y otros medios de transmisión más sofisticados, como pueden ser las microondas. La velocidad de transmisión suele ser inferior que en las redes locales.

Varias redes pueden conectarse entre S formando una red lógica de área mayor. Para que la transmisión entre todas ellas sea posible se emplean los routers, que son los sistemas que conectando físicamente varias redes se encargan de dirigir la información por el camino adecuado. Cuando las redes que se conectan son de diferente tipo y con protocolos distintos se hace necesario el uso de los gateways, los cuales además de encaminar la información también son capaces de convertir los datos de un protocolo a otro. Generalmente los términos router y gateway se emplean indistintamente para referirse de forma general a los sistemas encargados del encaminamiento de datos en Internet.

Lo que se conoce como Internet es en realidad una red de redes, la interconexión de otras redes independientes de manera que puedan compartir información entre ellas a lo largo de todo el planeta. Para ello es necesario el uso de un protocolo de comunicaciones común. El protocolo que proporciona la compatibilidad necesaria para la comunicación en Internet es el TCP/IP.

Los protocolos de comunicaciones definen las normas que posibilitan que se establezca una comunicación entre varios equipos o dispositivos, ya que estos equipos pueden ser diferentes entre S.

Un interfaz, sin embargo, es el encargado de la conexión física entre los equipos, definiendo las normas para las características eléctricas y mecánicas de la conexión.

Exceptuando a los routers cualquier ordenador conectado a Internet y, por tanto, capaz de compartir información con otro ordenador se conoce con el nombre de host (anfitrión). Un host debe identificarse de alguna manera que lo distinga de los demás para poder recibir o enviar datos. Para ello todos los ordenadores conectados a Internet disponen de una dirección única y exclusiva. Esta dirección, conocida como dirección de Internet o dirección IP, es un número de 32 bit que generalmente se representa en cuatro grupos de 8 bit cada uno separados por puntos y en base decimal (esto es así en la versión número 4 del protocolo IP, pero no en la 6). Un ejemplo de dirección IP es el siguiente: 205.198.48.1.

3. El modelo OSI.

El modelo OSI (Open System Interconection) es utilizado por prácticamente la totalidad de las redes del mundo. Este modelo fue creado por el ISO (Organización Internacional de Normalización), y consiste en siete niveles o capas donde cada una de ellas define las funciones que deben proporcionar los protocolos con el propósito de intercambiar información entre varios sistemas. Esta clasificación permite que cada protocolo se desarrolle con una finalidad determinada, lo cual simplifica el proceso de desarrollo e implementación. Cada nivel depende de los que están por debajo de él, y a su vez proporciona alguna funcionalidad a los niveles superiores.

Los siete niveles del modelo OSI son los siguientes:

Aplicación	El nivel de aplicación es el destino final de los datos donde se proporcionan los servicios al usuario.
Presentación	Se convierten e interpretan los datos que se utilizarán en el nivel de aplicación.
Sesión	Encargado de ciertos aspectos de la comunicación como el control de los tiempos.
Transporte	Transporta la información de una manera fiable para que llegue correctamente a su destino.
Red	Nivel encargado de encaminar los datos hacia su destino eligiendo la ruta más efectiva.
Enlace	Enlace de datos. Controla el flujo de los mismos, la sincronización y los errores que puedan producirse.
Físico	Se encarga de los aspectos físicos de la conexión, tales como el medio de transmisión o el hardware.

4. Arquitectura cliente-servidor.

La arquitectura cliente-servidor es una forma específica de diseño de aplicaciones, aunque también se conoce con este nombre a los ordenadores en los que se estas aplicaciones son ejecutadas. Por un lado, el cliente es el ordenador que se encarga de efectuar una petición o solicitar un servicio. El cliente no posee control sobre los recursos, sino que es el servidor el encargado de manejarlos. Por otro lado, el ordenador remoto que actúa como servidor evalúa la petición del cliente y decide aceptarla o rechazarla consecuentemente. Una vez que el servidor acepta el pedido la información requerida es suministrada al cliente que efectuó la petición, siendo este último el responsable de proporcionar los datos al usuario con el formato adecuado. Finalmente debemos precisar que cliente y servidor no tienen

que estar necesariamente en ordenadores separados, sino que pueden ser programas diferentes que se ejecuten en el mismo ordenador.

5. El protocolo TCP/IP.

TCP/IP es el protocolo común utilizado por todos los ordenadores conectados a Internet, de manera que éstos puedan comunicarse entre S. Hay que tener en cuenta que en Internet se encuentran conectados ordenadores de clases muy diferentes y con hardware y software incompatibles en muchos casos, además de todos los medios y formas posibles de conexión. Aquí se encuentra una de las grandes ventajas del TCP/IP, pues este protocolo se encargará de que la comunicación entre todos sea posible. TCP/IP es compatible con cualquier sistema operativo y con cualquier tipo de hardware.

TCP/IP no es un único protocolo, sino que es en realidad lo que se conoce con este nombre es un conjunto de protocolos que cubren los distintos niveles del modelo OSI. Los dos protocolos más importantes son el TCP (Transmission Control Protocol) y el IP (Internet Protocol), que son los que dan nombre al conjunto. En Internet se diferencian cuatro niveles o capas en las que se agrupan los protocolos, y que se relacionan con los niveles OSI de la siguiente manera:

- Aplicación: Se corresponde con los niveles OSI de aplicación, presentación y sesión. Aquí se incluyen protocolos destinados a proporcionar servicios, tales como correo electrónico (SMTP), transferencia de ficheros (FTP), conexión remota (TELNET) y otros más recientes como el protocolo HTTP (Hypertext Transfer Protocol).

- Transporte: Coincide con el nivel de transporte del modelo OSI. Los protocolos de este nivel, tales como TCP y UDP, se encargan de manejar los datos y proporcionar la fiabilidad necesaria en el transporte de los mismos.

- ternet: Es el nivel de red del modelo OSI. Incluye al protocolo IP, que se encarga de enviar los paquetes de información a sus destinos correspondientes. Es utilizado con esta finalidad por los protocolos del nivel de transporte.

- Enlace: Los niveles OSI correspondientes son el de enlace y el nivel físico. Los protocolos que pertenecen a este nivel son los encargados de la transmisión a través del medio físico al que se encuentra conectado cada host, como puede ser una línea punto a punto o una red Ethernet.

El TCP/IP necesita funcionar sobre algún tipo de red o de medio físico que proporcione sus propios protocolos para el nivel de enlace de Internet. Por este motivo hay que tener en cuenta que los protocolos utilizados en este nivel pueden ser muy diversos y no forman parte del conjunto TCP/IP. Sin embargo, esto no debe ser problemático puesto que una de las funciones y ventajas principales del TCP/IP es proporcionar una abstracción del medio de forma que sea posible el intercambio de información entre medios diferentes y tecnologías que inicialmente son incompatibles.

Para transmitir información a través de TCP/IP, ésta debe ser dividida en unidades de menor tamaño. Esto proporciona grandes ventajas en el manejo de los datos que se transfieren y, por otro lado, esto es algo común en cualquier protocolo de comunicaciones. En TCP/IP cada una de estas unidades de información recibe el nombre de "datagrama" (datagram), y son conjuntos de datos que se envían como mensajes independientes.

6. TCP (Transmission Control Protocol).

El protocolo de control de transmisión (TCP) pertenece al nivel de transporte, siendo el encargado de dividir el mensaje original en datagramas de menor tamaño, y por lo tanto, mucho más manejables. Los datagramas serán dirigidos a través del protocolo IP de forma individual. El protocolo TCP se encarga además de añadir cierta información necesaria a cada uno de los datagramas. Esta información se añade al inicio de los datos que componen el datagrama en forma de cabecera.

La cabecera de un datagrama contiene al menos 160 bit que se encuentran repartidos en varios campos con diferente significado. Cuando la información se divide en datagramas para ser enviados, el orden en que éstos lleguen a su destino no tiene que ser el correcto. Cada uno de ellos puede llegar en cualquier momento y con cualquier orden, e incluso puede que algunos no lleguen a su destino o lleguen con información errónea. Para evitar todos estos problemas el TCP numera los datagramas antes de ser enviados, de manera que sea posible volver a unirlos en el orden adecuado. Esto permite también

solicitar de nuevo el envío de los datagramas individuales que no hayan llegado o que contengan errores, sin que sea necesario volver a enviar el mensaje completo.

Formato de la cabecera TCP.

Puerto origen

Puerto destino

Número de secuencia

Señales de confirmación

Tamaño

Reservado

Bits de control

Window

Checksum

Puntero a datos urgentes

En cualquier caso el tamaño de la cabecera debe ser múltiplo de 32 bits, por lo que puede ser necesario añadir un campo de tamaño variable y que contenga ceros al final para conseguir este objetivo cuando se incluyen algunas opciones. El campo de tamaño contiene la longitud total de la cabecera TCP expresada en el número de palabras de 32 bits que ocupa. Esto permite determinar el lugar donde comienzan los datos.

Dos campos incluidos en la cabecera y que son de especial importancia son los números de puerto de origen y puerto de destino. Los puertos proporcionan una manera de distinguir entre las distintas transferencias, ya que un mismo ordenador puede estar utilizando varios servicios o transferencias simultáneamente, e incluso puede que por medio de usuarios distintos. El puerto de origen contendrá un número cualquiera que sirva para realizar esta distinción. Además, el programa cliente que realiza la petición también se debe conocer el número de puerto en el que se encuentra el servidor adecuado. Mientras que el programa del usuario utiliza números prácticamente aleatorios, el servidor deber tener asignado un número estándar para que pueda ser utilizado por el cliente. (Por ejemplo, en el caso de la transferencia de ficheros FTP el número oficial es el 21). Cuando es el servidor el que envía los datos, los números de puertos de origen y destino se intercambian.

En la transmisión de datos a través del protocolo TCP la fiabilidad es un factor muy importante. Para poder detectar los errores y pérdida de información en los datagramas, es necesario que el cliente envíe de nuevo al servidor unas señales de confirmación una vez que se ha recibido y comprobado la información satisfactoriamente. Estas señales se incluyen en el campo apropiado de la cabecera del datagrama (Acknowledgment Number), que tiene un tamaño de 32 bit. Si el servidor no obtiene la señal de confirmación adecuada transcurrido un período de tiempo razonable, el datagrama completo se volverá a enviar. Por razones de eficiencia los datagramas se envían continuamente sin esperar la confirmación, haciéndose necesaria la numeración de los mismos para que puedan ser ensamblados en el orden correcto.

También puede ocurrir que la información del datagrama llegue con errores a su destino. Para poder detectar cuando sucede esto se incluye en la cabecera un campo de 16 bit, el cual contiene un valor calculado a partir de la información del datagrama completo (checksum). En el otro extremo el

receptor vuelve a calcular este valor, comprobando que es el mismo que el suministrado en la cabecera. Si el valor es distinto significaría que el datagrama es incorrecto, ya que en la cabecera o en la parte de datos del mismo hay algún error.

La forma en que TCP numera los datagramas es contando los bytes de datos que contiene cada uno de ellos y añadiendo esta información al campo correspondiente de la cabecera del datagrama siguiente. De esta manera el primero empezará por cero, el segundo contendrá un número que será igual al tamaño en bytes de la parte de datos del datagrama anterior, el tercero con la suma de los dos anteriores, y así sucesivamente. Por ejemplo, para un tamaño fijo de 500 bytes de datos en cada datagrama, la numeración sería la siguiente: 0 para el primero, 500 para el segundo, 1000 para el tercero, etc.

Existe otro factor más a tener en cuenta durante la transmisión de información, y es la potencia y velocidad con que cada uno de los ordenadores puede procesar los datos que le son enviados. Si esto no se tuviera en cuenta, el ordenador de más potencia podría enviar la información demasiado rápido al receptor, de manera que éste no pueda procesarla. Este inconveniente se soluciona mediante un campo de 16 bit (Window) en la cabecera TCP, en el cual se introduce un valor indicando la cantidad de información que el receptor está preparado para procesar. Si el valor llega a cero será necesario que el emisor se detenga. A medida que la información es procesada este valor aumenta indicando disponibilidad para continuar la recepción de datos.

Protocolos alternativos a TCP.

TCP es el protocolo más utilizado para el nivel de transporte en Internet, pero además de éste existen otros protocolos que pueden ser más convenientes en determinadas ocasiones. Tal es el caso de UDP y ICMP.

7. UDP (User Datagram Protocol)

El protocolo de datagramas de usuario (UDP) puede ser la alternativa al TCP en algunos casos en los que no sea necesario el gran nivel de complejidad proporcionado por el TCP. Puesto que UDP no admite numeración de los datagramas, éste protocolo se utiliza principalmente cuando el orden en que se reciben los mismos no es un factor fundamental, o también cuando se quiere enviar información de poco tamaño que cabe en un único datagrama.

Cuando se utiliza UDP la garantía de que un paquete llegue a su destino es mucho menor que con TCP debido a que no se utilizan las señales de confirmación. Por todas estas características la cabecera del UDP es bastante menor en tamaño que la de TCP. Esta simplificación resulta en una mayor eficiencia en determinadas ocasiones.
Un ejemplo típico de una situación en la que se utiliza el UDP es cuando se pretende conectar con un ordenador de la red, utilizando para ello el nombre del sistema. Este nombre tendrá que ser convertido a la dirección IP que le corresponde y, por tanto, tendrá que ser enviado a algún servidor que posea la base de datos necesaria para efectuar la conversión. En este caso es mucho más conveniente el uso de UDP.

8. ICMP (Internet Control Message Protocol)

El protocolo de mensajes de control de Internet (ICMP) es de características similares al UDP, pero con un formato aún más simple. Su utilidad no está en el transporte de datos "de usuario", sino en los mensajes de error y de control necesarios para los sistemas de la red.

9. IP (Internet Protocol)

El IP es un protocolo que pertenece al nivel de red, por lo tanto, es utilizado por los protocolos del nivel de transporte como TCP para encaminar los datos hacia su destino. IP tiene únicamente la misión de encaminar el datagrama, sin comprobar la integridad de la información que contiene. Para ello se utiliza una nueva cabecera que se antepone al datagrama que se está tratando. Suponiendo que el protocolo TCP ha sido el encargado de manejar el datagrama antes de pasarlo al IP, la estructura del mensaje una vez tratado quedaría así:

(20 byte)	(20 byte)

La cabecera IP tiene un tamaño de 160 bit y está formada por varios campos de distinto significado. Estos campos son:

- Versión: Número de versión del protocolo IP utilizado. Tendrá que tener el valor 4. Tamaño: 4 bit.
- Longitud de la cabecera: (Internet Header Length, IHL) Especifica la longitud de la cabecera expresada en el número de grupos de 32 bit que contiene. Tamaño: 4 bit.
- Tipo de servicio: El tipo o calidad de servicio se utiliza para indicar la prioridad o importancia de los datos que se envían, lo que condicionará la forma en que éstos serán tratados durante la transmisión. Tamaño: 8 bit.
- Longitud total: Es la longitud en bytes del datagrama completo, incluyendo la cabecera y los datos. Como este campo utiliza 16 bit, el tamaño máximo del datagrama no podrá superar los 65.535 bytes, aunque en la práctica este valor será mucho más pequeño. Tamaño: 16 bit.
- Identificación: Valor de identificación que se utiliza para facilitar el ensamblaje de los fragmentos del datagrama. Tamaño: 16 bit.
- Flags: Indicadores utilizados en la fragmentación. Tamaño: 3 bit.
- Fragmentación: Contiene un valor (offset) para poder ensamblar los datagramas que se hayan fragmentado. Está expresado en número de grupos de 8 bytes (64 bit), comenzando con el valor cero para el primer fragmento. Tamaño: 16 bit.
- Límite de existencia: Contiene un número que disminuye cada vez que el paquete pasa por un sistema. Si este número llega a cero, el paquete será descartado. Esto es necesario por razones de seguridad para evitar un bucle infinito, ya que aunque es bastante improbable que esto suceda en una red correctamente diseñada, no debe descuidarse esta posibilidad. Tamaño: 8 bit.
- Protocolo: El número utilizado en este campo sirve para indicar a qué protocolo pertenece el datagrama que se encuentra a continuación de la cabecera IP, de manera que pueda ser tratado correctamente cuando llegue a su destino. Tamaño: 8 bit.
- Comprobación: El campo de comprobación (checksum) es necesario para verificar que los datos contenidos en la cabecera IP son correctos. Por razones de eficiencia este campo no puede utilizarse para comprobar los datos incluidos a continuación, sino que estos datos de usuario se comprobarán posteriormente a partir del campo de comprobación de la cabecera siguiente, y que corresponde al nivel de transporte. Este campo debe calcularse de nuevo cuando cambia alguna opción de la cabecera, como puede ser el límite de existencia. Tamaño: 16 bit.
- Dirección de origen: Contiene la dirección del host que envía el paquete. Tamaño: 32 bit.
- Dirección de destino**:** Esta dirección es la del host que recibirá la información. Los routers o gateways intermedios deben conocerla para dirigir correctamente el paquete. Tamaño: 32 bit.

10. La dirección de Internet.

El protocolo IP identifica a cada ordenador que se encuentre conectado a la red mediante su correspondiente dirección. Esta dirección es un número de 32 bit que debe ser único para cada host, y normalmente suele representarse como cuatro cifras de 8 bit separadas por puntos.

La dirección de Internet (IP Address) se utiliza para identificar tanto al ordenador en concreto como la red a la que pertenece, de manera que sea posible distinguir a los ordenadores que se encuentran conectados a una misma red. Con este propósito, y teniendo en cuenta que en Internet se encuentran conectadas redes de tamaños muy diversos, se establecieron tres clases diferentes de direcciones, las cuales se representan mediante tres rangos de valores:

- Clase A: Son las que en su primer byte tienen un valor comprendido entre 1 y 126, incluyendo ambos valores. Estas direcciones utilizan únicamente este primer byte para identificar la red, quedando los otros tres bytes disponibles para cada uno de los hosts que pertenezcan a esta misma red. Esto significa que podrán existir más de dieciséis millones de ordenadores en cada una de las redes de esta clase. Este tipo de direcciones es usado por redes muy extensas, pero hay que tener en cuenta que sólo puede haber 126 redes de este tamaño. ARPAnet es una de ellas, existiendo además algunas grandes redes comerciales, aunque son pocas las organizaciones que obtienen una dirección de "clase A". Lo normal para las grandes organizaciones es que utilicen una o varias redes de "clase B".

- Clase B: Estas direcciones utilizan en su primer byte un valor comprendido entre 128 y 191, incluyendo ambos. En este caso el identificador de la red se obtiene de los dos primeros bytes de la dirección, teniendo que ser un valor entre 128.1 y 191.254 (no es posible utilizar los valores 0 y 255 por tener un significado especial). Los dos últimos bytes de la dirección constituyen el identificador del host permitiendo, por consiguiente, un número máximo de 64516 ordenadores en la misma red. Este tipo de direcciones tendría que ser suficiente para la gran mayoría de las organizaciones grandes. En caso de que el número de ordenadores que se necesita conectar fuese mayor, sería posible obtener más de una dirección de "clase B", evitando de esta forma el uso de una de "clase A".

- Clase C: En este caso el valor del primer byte tendrá que estar comprendido entre 192 y 223, incluyendo ambos valores. Este tercer tipo de direcciones utiliza los tres primeros bytes para el número de la red, con un rango desde 192.1.1 hasta 223.254.254. De esta manera queda libre un byte para el host, lo que permite que se conecten un máximo de 254 ordenadores en cada red. Estas direcciones permiten un menor número de host que las anteriores, aunque son las más numerosas pudiendo existir un gran número redes de este tipo (más de dos millones).

Tabla de direcciones IP de Internet.

Clase	Primer byte	Identificación de red	Identificación de hosts	Número de redes	Número de hosts
A	1 .. 126	1 byte	3 byte	126	16.387.064
B	128 .. 191	2 byte	2 byte	16.256	64.516
C	192 .. 223	3 byte	1 byte	2.064.512	254

En la clasificación de direcciones anterior se puede notar que ciertos números no se usan. Algunos de ellos se encuentran reservados para un posible uso futuro, como es el caso de las direcciones cuyo primer byte sea superior a 223 (clases D y E, que aún no están definidas), mientras que el valor 127 en el primer byte se utiliza en algunos sistemas para propósitos especiales. También es importante notar que los valores 0 y 255 en cualquier byte de la dirección no pueden usarse normalmente por tener otros propósitos específicos. El número 0 está reservado para las máquinas que no conocen su dirección, pudiendo utilizarse tanto en la identificación de red para máquinas que aún no conocen el número de red a la que se encuentran conectadas, en la identificación de host para máquinas que aún no conocen su número de host dentro de la red, o en ambos casos.

El número 255 tiene también un significado especial, puesto que se reserva para el broadcast. El broadcast es necesario cuando se pretende hacer que un mensaje sea visible para todos los sistemas conectados a la misma red. Esto puede ser útil si se necesita enviar el mismo datagrama a un número determinado de sistemas, resultando más eficiente que enviar la misma información solicitada de manera individual a cada uno. Otra situación para el uso de broadcast es cuando se quiere convertir el nombre por dominio de un ordenador a su correspondiente número IP y no se conoce la dirección del servidor de nombres de dominio más cercano.

Lo usual es que cuando se quiere hacer uso del broadcast se utilice una dirección compuesta por el identificador normal de la red y por el número 255 (todo unos en binario) en cada byte que identifique al host. Sin embargo, por conveniencia también se permite el uso del número 255.255.255.255 con la misma finalidad, de forma que resulte más simple referirse a todos los sistemas de la red.

El broadcast es una característica que se encuentra implementada de formas diferentes dependiendo del medio utilizado, y por lo tanto, no siempre se encuentra disponible. En ARPAnet y en las líneas punto a punto no es posible enviar broadcast, pero sí que es posible hacerlo en las redes Ethernet, donde se supone que todos los ordenadores prestarán atención a este tipo de mensajes.

En el caso de algunas organizaciones extensas puede surgir la necesidad de dividir la red en otras redes más pequeñas (subnets). Como ejemplo podemos suponer una red de clase B que, naturalmente, tiene asignado como identificador de red un número de dos bytes. En este caso sería posible utilizar el tercer byte para indicar en qué red Ethernet se encuentra un host en concreto. Esta división no tendrá ningún significado para cualquier otro ordenador que esté conectado a una red perteneciente a otra organización, puesto que el tercer byte no será comprobado ni tratado de forma especial. Sin embargo, en el interior de esta red existirá una división y será necesario disponer de un software de red especialmente diseñado para ello. De esta forma queda oculta la organización interior de la red, siendo mucho más cómodo el acceso que si se tratara de varias direcciones de clase C independientes.

11. Niveles físico y de enlace: Ethernet.

Los protocolos que pertenecen al nivel de enlace o interfaz de red de Internet (niveles físico y de enlace en el modelo OSI) deben añadir más información a los datos provenientes de IP para que la transmisión pueda realizarse. Es el caso, por ejemplo, de las redes Ethernet, de uso muy extendido actualmente. Este tipo de redes utiliza su propio sistema de direcciones, junto con una nueva cabecera para los datos.

Las redes locales Ethernet son posiblemente la tecnología que domina en Internet. Este tipo de redes fue desarrollado por Xerox durante los años 70, y entre sus características podemos destacar su alto nivel de rendimiento, la utilización de cable coaxial para la transmisión, una velocidad de 10Mbit/seg. y CSMA/CD como técnica de acceso.

Ethernet es un medio en el que todos los ordenadores pueden acceder a cada uno de los paquetes que se envían, aunque un ordenador sólo tendrá que prestar atención a aquellos que van dirigidos a él mismo.

La técnica de acceso CSMA/CD (Carrier Sense and Multiple Access with Collition Detection) permite a que todos los dispositivos puedan comunicarse en el mismo medio, aunque sólo puede existir un único emisor en cada instante. De esta manera todos los sistemas pueden ser receptores de forma simultánea, pero la información tiene que ser transmitida por turnos. Si varios dispositivos intentan transmitir en el mismo instante la colisión es detectada, de forma que cada uno de ellos volverá a intentar la transmisión transcurrido un pequeño intervalo de tiempo aleatorio.

Es importante notar que las direcciones utilizadas por Ethernet no guardan ninguna relación con las direcciones de Internet. Así como las direcciones IP de Internet son asignadas por el usuario, las direcciones Ethernet se asignan "de fábrica". Esta es la razón por la que se utilizan 48 bit en las direcciones, ya que de esta manera se obtiene un número lo suficientemente elevado de direcciones como para asegurar que no sea necesario repetir los valores.

En una red Ethernet los paquetes son transportados de un ordenador a otro de manera que son visibles para todos, siendo necesario un procedimiento para identificar los paquetes que pertenecen a cada ordenador. Cuando el paquete es recibido en el otro extremo, la cabecera y el checksum se retiran, se comprueba que los datos corresponden a un mensaje IP, y este mensaje se pasa al protocolo IP para que sea procesado.

El tamaño máximo para un paquete de datos varía de unas redes a otras. En el caso de Ethernet el tamaño puede ser de 1500 bytes, para otras redes puede ser menor o bastante mayor en el caso de redes muy rápidas. Aquí surge otro problema, pues normalmente los paquetes de tamaño mayor resultan más eficientes para transmitir grandes cantidades de información. Sin embargo, se debe tener en cuenta que las redes del receptor y el emisor pueden ser muy distintas. Por este motivo el protocolo TCP está preparado para negociar el tamaño máximo de los datagramas que serán enviados durante el resto de la conexión. Pero así el problema no queda completamente resuelto porque hasta que los paquetes lleguen a su destino es muy probable que tengan que atravesar otras redes intermedias, las cuales puede que no sean capaces de soportar el tamaño de los paquetes que se está

enviando. Se hace necesario entonces dividir el paquete original en otros más pequeños para que puedan ser manejados: Esto se conoce como fragmentación (fragmentation).

La fragmentación es posible gracias a determinados campos que el protocolo IP introduce en su cabecera. Estos campos de fragmentación se usan cuando ha sido necesario dividir el paquete enviado originalmente, de manera que éste pueda ser reconstruido por el host receptor a través del protocolo TCP/IP. Este último proceso de reconstrucción de los paquetes se conoce como "reensamblaje" (reassembly).

ARP (Address Resolution Protocol).

El Protocolo de Resolución de Direcciones (ARP) es necesario debido a que las direcciones Ethernet y las direcciones IP son dos números distintos y que no guardan ninguna relación. Así, cuando pretendemos dirigirnos a un host a través de su dirección de Internet se necesita convertir ésta a la correspondiente dirección Ethernet.

ARP es el protocolo encargado de realizar las conversiones de dirección correspondientes a cada host. Para ello cada sistema cuenta con una tabla con la dirección IP y la dirección Ethernet de algunos de los otros sistemas de la misma red. Sin embargo, también puede ocurrir que el ordenador de destino no se encuentre en la tabla de direcciones, teniendo entonces que obtenerla por otros medios.

Con la finalidad de obtener una dirección Ethernet destino que no se encuentra en la tabla de conversiones se utiliza el mensaje ARP de petición. Este mensaje es enviado como broadcast, es decir, que estará disponible para que el resto de los sistemas de la red lo examinen, y el cual contiene una solicitud de la dirección final de un sistema a partir de su dirección IP. Cuando el ordenador con el que se quiere comunicar analiza este mensaje comprueba que la dirección IP corresponde a la suya y envía de regreso el mensaje ARP de respuesta, el cual contendrá la dirección Ethernet que se estaba buscando. El ordenador que solicitó la información recibirá entonces el mensaje de respuesta y añadirá la dirección a su propia tabla de conversiones para futuras referencias.

El mensaje de petición ARP contiene las direcciones IP y Ethernet del host que solicita la información, además de la dirección IP del host de destino. Estos mensajes son aprovechados en algunas ocasiones también por otros sistemas de la red para actualizar sus tablas, ya que el mensaje es enviado en forma de broadcast. El ordenador de destino, una vez que ha completado el mensaje inicial con su propia dirección Ethernet, envía la respuesta directamente al host que solicitó la información.

12. Routing.

Ya se ha expuesto anteriormente la forma en que los datagramas pasan de un ordenador de la red a otro mediante el protocolo IP, sin embargo en esta sección se comenta con más detalle el proceso que permite que la información llegue hasta su destino final. Esto se conoce con el nombre de routing.

Las tareas de routing son implementadas por el protocolo IP sin que los protocolos de un nivel superior tales como TCP o UDP tengan constancia de ello. Cuando se quiere enviar información por Internet a un ordenador, el protocolo IP comprueba si el ordenador de destino se encuentra en la misma red local que el ordenador origen. Si es así, se enviará el correspondiente datagrama de forma directa: la cabecera IP contendrá el valor de la dirección Internet del ordenador destino, y la cabecera Ethernet contendrá el valor de la dirección de la red Ethernet que corresponde a este mismo ordenador.

Cuando se pretende enviar información a un ordenador remoto que está situado en una red local diferente al ordenador de origen, el proceso resulta más complicado. Esto se conoce como routing indirecto, y es el caso que se presenta más frecuentemente cuando se envía información en Internet. La figura 1 muestra un ejemplo en el que dos redes locales que utilizan la tecnología de Internet se enlazan para intercambiar información, creando una red lógica de mayor tamaño gracias a la funcionalidad del protocolo IP.

En Internet existen un elevado número de redes independientes conectadas entre sí mediante el uso de los routers. Un ordenador puede actuar como un router si se conecta a varias redes al mismo tiempo, disponiendo por lo tanto de más de una interfaz de red así como de varias direcciones IP y Ethernet (tantas como redes a las que se encuentre conectado). El router, por supuesto, puede enviar y recibir información de los hosts de todas las redes a las que está conectado, y siempre será de forma directa. Continuando con el ejemplo anterior, el host A puede comunicarse de forma directa con el host

B, así como los hosts A y B pueden enviar o recibir información del router. En ambos casos se trata de routing directo, pues el ordenador que actúa como router está conectado a la red 'alfa' de la misma manera que los ordenadores A y B, teniendo una dirección IP propia asignada que lo identifica dentro de esta misma red. La situación es la misma para la red 'omega' donde el router es identificado a través de una segunda dirección IP que corresponde con esta red.

Si sólo fuésemos a enviar información de manera directa dentro de una misma red no sería necesario el uso del protocolo TCP/IP, siendo el mismo especialmente indicado cuando se desea una comunicación con otras redes. En este caso los datagramas tendrán que ser encaminados a través del router para llegar a su destino. La forma de hacer esto es a través del protocolo IP, el cual decide si la información puede enviarse directamente o si por el contrario debe utilizarse el método indirecto a través de un router. Tomamos de nuevo el ejemplo de la figura 1: Suponemos que el host B de la red 'alfa' necesita comunicarse con el host X situado en la red 'omega'. Una vez que se ha determinado que el destino no se encuentra en la misma red, envía el datagrama IP hacia el router correspondiente. Como este router y el ordenador que envía la información se encuentran conectados a la misma red, se trata por tanto de routing directo, ya comentado anteriormente, y por consiguiente sólo será necesario determinar la dirección Ethernet del router mediante empleo del protocolo ARP. El paquete enviado incluirá la dirección del router como dirección Ethernet de destino, pero sin embargo, la dirección de destino IP corresponderá al ordenador final al que va dirigido el paquete, el host X en el ejemplo. El router recibe el paquete y a través del protocolo IP comprueba que la dirección de Internet de destino no corresponde con ninguna de las asignadas como suyas, procediendo entonces a determinar la localización de la 'omega', en la que se entrega el paquete al ordenador de destino.

Hasta este punto se ha supuesto que sólo existe un único router, pero es bastante probable que una red con conexión a Internet posea múltiples enlaces con otras redes, y por lo tanto más de un router. Entonces... ¿cómo determina el protocolo IP el sistema correcto al que debe dirigirse? Para resolver este problema cada ordenador utiliza una tabla donde se relaciona cada una de las redes existentes con el router que debe usarse para tener acceso. Debe tenerse en cuenta que los routers indicados en estas tablas pueden no estar conectados directamente a las redes con las que están relacionados, sino que lo que se indica es el mejor camino para acceder a cada una de ellas. Por esta razón, cuando un router recibe un paquete que debe ser encaminado, busca en su propia tabla de redes la entrada correspondiente a la red para, una vez encontrada, entregarlo al ordenador de destino. Es importante notar que en el caso de que el router no tenga conexión directa a la misma red que el ordenador de destino, la búsqueda en su tabla de redes dará como resultado la dirección de un nuevo router al que dirigir el paquete, y así continuará el proceso sucesivamente hasta encontrar el destino final.

La figura 2 muestra la estructura de los protocolos para cada ordenador de Internet que se encuentre conectado a una red Ethernet. Para un ordenador con más de un interfaz de red en el esquema aparecerían todas las Ethernet con sus correspondientes protocolos ARP, pero en cualquier caso sería un único protocolo IP el que se utilice, aunque éste disponga de varias direcciones asignadas.

A causa de la extensión de Internet, es normal que un paquete atraviese numerosas redes (pueden ser decenas) hasta llegar a su destino. La ruta que tiene que recorrer un paquete en su viaje a través de la red no está determinada inicialmente, sino que es el resultado de la consulta en las tablas de direcciones individuales de los ordenadores intermedios.

La creación y mantenimiento de la tabla de redes para routing es un proceso complejo que debe ser realizado por el administrador de la red. Aquí hay que tener en cuenta que la enorme extensión de Internet supone una gran dificultad para conseguir que sean correctas todas las entradas de la tabla, además de que esta tabla puede llegar a tener un tamaño considerable. La utilización de routers por defecto mejora la situación al permitir que sean estos los que guarden el registro de la red sin que los ordenadores individuales tengan que ocuparse en ello, pero estos routers sí que deberían tener una tabla completa. Para facilitar el mantenimiento de la tabla existen algunos protocolos para routing que permiten que un router o gateway cualquiera pueda encontrar por sí mismo la localización de otros routers o gateways y guardar la información acerca del mejor camino para acceder a cada red.

Lógicamente el proceso real de routing sobre Internet suele ser mucho más complejo que el expuesto aquí, principalmente por el uso de redes y tecnologías muy distintas e incompatibles. Esto obliga a que se realicen conversiones en el formato de los paquetes para que puedan pasar a través de

medios diferentes, pero en cualquier caso el protocolo IP proporciona una transmisión transparente para los protocolos de nivel superior y las aplicaciones de red.

13. Sistema de nombres por dominio.

El sistema de nombres por dominio (DNS, Domain Name System) es una forma alternativa de identificar a una máquina conectada a Internet. La dirección IP resulta difícil de memorizar, siendo su uso más adecuado para los ordenadores. El sistema de nombres por dominio es el utilizado normalmente por las personas para referirse a un ordenador en la red, ya que además puede proporcionar una idea del propósito o la localización del mismo.

El nombre por dominio de un ordenador se representa de forma jerárquica con varios nombres separados por puntos (generalmente 3 ó 4, aunque no hay límite). Típicamente el nombre situado a la izquierda identifica al host, el siguiente es el subdominio al que pertenece este host, y a la derecha estará el dominio de mayor nivel que contiene a los otros subdominios:

nombre_ordenador.subdominio.dominio_principal

Aunque esta situación es la más común, el nombre por dominio es bastante flexible, permitiendo no sólo la identificación de hosts sino que también puede utilizarse para referirse a determinados servicios proporcionados por un ordenador o para identificar a un usuario dentro del mismo sistema. Es el caso de la dirección de correo electrónico, donde el nombre por dominio adquiere gran importancia puesto que el número IP no es suficiente para identificar al usuario dentro de un ordenador.

Para que una máquina pueda establecer conexión con otra es necesario que conozca su número IP, por lo tanto, el nombre por dominio debe ser convertido a su correspondiente dirección a través de la correspondiente base de datos. En los inicios de Internet esta base de datos era pequeña de manera que cada sistema podía tener su propia lista con los nombres y las direcciones de los otros ordenadores de la red, pero actualmente esto sería impensable. Con esta finalidad se utilizan los servidores de nombres por dominio (DNS servers).

Los servidores de nombres por dominio son sistemas que contienen bases de datos con el nombre y la dirección de otros sistemas en la red de una forma encadenada o jerárquica.

Para comprender mejor el proceso supongamos que un usuario suministra el nombre por dominio de un sistema en la red a su ordenador local, realizándose el siguiente proceso:

- El ordenador local entra en contacto con el servidor de nombres que tiene asignado, esperando obtener la dirección que corresponde al nombre que ha suministrado el usuario.
- El servidor de nombres local puede conocer la dirección que se está solicitando, entregándosela al ordenador que realizó la petición.
- Si el servidor de nombres local no conoce la dirección, ésta se solicitará al servidor de nombres que esté en el dominio más apropiado. Si éste tampoco tiene la dirección, llamará al siguiente servidor DNS, y así sucesivamente.
- Cuando el servidor DNS local ha conseguido la dirección, ésta se entrega al ordenador que realizó la petición.
- Si el nombre por dominio no se ha podido obtener, se enviará de regreso el correspondiente mensaje de error.

14. Servicios de Internet: el nivel de aplicación.

Los diferentes servicios a los que podemos tener acceso en Internet son proporcionados por los protocolos que pertenecen al nivel de aplicación. Estos protocolos forman parte del TCP/IP y deben aportar entre otras cosas una forma normalizada para interpretar la información, ya que todas las máquinas no utilizan los mismos juegos de caracteres ni los mismos estándares. Los protocolos de los otros niveles sólo se encargan de la transmisión de información como un bloque de bits, sin definir las normas que indiquen la manera en que tienen que interpretarse esos bits. Los protocolos del nivel de

aplicación están destinados a tareas específicas, algunos de los cuales se consideran como tradicionales de Internet por utilizarse desde los inicios de la red, como son por ejemplo:

- Transferencia de ficheros (File Transfer).
- Correo electrónico (e-mail).
- Conexión remota (remote login).

15. Transferencia de ficheros.

El protocolo FTP (File Transfer Protocol) se incluye como parte del TCP/IP, siendo éste el protocolo de nivel de aplicación destinado a proporcionar el servicio de transferencia de ficheros en Internet. El FTP depende del protocolo TCP para las funciones de transporte, y guarda alguna relación con TELNET (protocolo para la conexión remota).

El protocolo FTP permite acceder a algún servidor que disponga de este servicio y realizar tareas como moverse a través de su estructura de directorios, ver y descargar ficheros al ordenador local, enviar ficheros al servidor o copiar archivos directamente de un servidor a otro de la red. Lógicamente y por motivos de seguridad se hace necesario contar con el permiso previo para poder realizar todas estas operaciones. El servidor FTP pedirá el nombre de usuario y clave de acceso al iniciar la sesión (login), que debe ser suministrado correctamente para utilizar el servicio.

La manera de utilizar FTP es por medio de una serie de comandos, los cuales suelen variar dependiendo del sistema en que se esté ejecutando el programa, pero básicamente con la misma funcionalidad. Existen aplicaciones de FTP para prácticamente todos los sistemas operativos más utilizados, aunque hay que tener en cuenta que los protocolos TCP/IP están generalmente muy relacionados con sistemas UNIX. Por este motivo y, ya que la forma en que son listados los ficheros de cada directorio depende del sistema operativo del servidor, es muy frecuente que esta información se muestre con el formato propio del UNIX. También hay que mencionar que en algunos sistemas se han desarrollado clientes de FTP que cuentan con un interfaz gráfico de usuario, lo que facilita notablemente su utilización, aunque en algunos casos se pierde algo de funcionalidad.

Existe una forma muy utilizada para acceder a fuentes de archivos de carácter público por medio de FTP. Es el acceso FTP anónimo, mediante el cual se pueden copiar ficheros de los hosts que lo permitan, actuando estos host como enormes almacenes de información y de todo tipo de ficheros para uso público. Generalmente el acceso anónimo tendrá algunas limitaciones en los permisos, siendo normal en estos casos que no se permita realizar acciones tales como añadir ficheros o modificar los existentes. Para tener acceso anónimo a un servidor de FTP hay que identificarse con la palabra "anonymous" como el nombre de usuario, tras lo cual se pedirá el password o clave correspondiente. Normalmente se aceptará cualquier cadena de caracteres como clave de usuario, pero lo usual es que aquí se indique la dirección de correo electrónico propia, o bien la palabra "guest". Utilizar la dirección de correo electrónico como clave de acceso es una regla de cortesía que permite a los operadores y administradores hacerse una idea de los usuarios que están interesados en el servicio, aunque en algunos lugares puede que se solicite esta información rechazando el uso de la palabra "guest".

E FTP proporciona dos modos de transferencia de ficheros: ASCII y binario. El modo de transferencia ASCII se utiliza cuando se quiere transmitir archivos de texto, ya que cada sistema puede utilizar un formato distinto para la representación de texto. En este caso se realiza una conversión en el formato del fichero original, de manera que el fichero recibido pueda utilizarse normalmente. El modo de transferencia binario se debe utilizar en cualquier otro caso, es decir, cuando el fichero que vamos a recibir contiene datos que no son texto. Aquí no se debe realizar ninguna conversión porque quedarían inservibles los datos del fichero.

16. Conexión remota.

El protocolo diseñado para proporcionar el servicio de conexión remota (remote login) recibe el nombre de TELNET, el cual forma parte del conjunto de protocolos TCP/IP y depende del protocolo TCP para el nivel de transporte.
El protocolo TELNET es un emulador de terminal que permite acceder a los recursos y ejecutar los programas de un ordenador remoto en la red, de la misma forma que si se tratara de un terminal real directamente conectado al sistema remoto. Una vez establecida la conexión el usuario podrá iniciar la

sesión con su clave de acceso. De la misma manera que ocurre con el protocolo FTP, existen servidores que permiten un acceso libre cuando se especifica "anonymous" como nombre de usuario.

Es posible ejecutar una aplicación cliente TELNET desde cualquier sistema operativo, pero hay que tener en cuenta que los servidores suelen ser sistemas VMS o UNIX por lo que, a diferencia del protocolo FTP para transferencia de ficheros donde se utilizan ciertos comandos propios de esta aplicación, los comandos y sintaxis que se utilice en TELNET deben ser los del sistema operativo del servidor. El sistema local que utiliza el usuario se convierte en un terminal "no inteligente" donde todos los caracteres pulsados y las acciones que se realicen se envían al host remoto, el cual devuelve el resultado de su trabajo. Para facilitar un poco la tarea a los usuarios, en algunos casos se encuentran desarrollados menús con las distintas opciones que se ofrecen.

Los programas clientes de TELNET deben ser capaces de emular los terminales en modo texto más utilizados para asegurarse la compatibilidad con otros sistemas, lo que incluye una emulación del teclado. El terminal más extendido es el VT100, el cual proporciona compatibilidad con la mayoría de los sistemas, aunque puede ser aconsejable que el programa cliente soporte emulación de otro tipo de terminales.

17. Correo electrónico.

El servicio de correo electrónico se proporciona a través del protocolo SMTP (Simple Mail Transfer Protocol), y permite enviar mensajes a otros usuarios de la red. A través de estos mensajes no sólo se puede intercambiar texto, sino también archivos binarios de cualquier tipo.

Generalmente los mensajes de correo electrónico no se envían directamente a los ordenadores personales de cada usuario, puesto que en estos casos puede ocurrir que esté apagado o que no esté ejecutando la aplicación de correo electrónico. Para evitar este problema se utiliza un ordenador más grande como almacén de los mensajes recibidos, el cual actúa como servidor de correo electrónico permanentemente. Los mensajes permanecerán en este sistema hasta que el usuario los transfiera a su propio ordenador para leerlos de forma local.

18. El acceso a Internet.

El acceso a Internet es proporcionado por cualquier proveedor que disponga de esta posibilidad, para lo cual se hace completamente necesario el protocolo TCP/IP. El número IP que dispondrá como dirección el ordenador del usuario final es suministrado por el proveedor (puede ser distinto cada vez que se establezca una conexión) y será una dirección válida de Internet.

19. Otras fuentes de información:

Todos los protocolos y estándares que se consolidan como propios de Internet han de ser organizados y dirigidos de alguna manera. Esta es la misión principal del IETF (Internet Engineering Task Force), que es una gran comunidad de carácter abierto formada por diseñadores de redes, operadores, usuarios, etc. Todos los protocolos agrupados normalmente bajo el nombre TCP/IP son estándares de Internet cuyo desarrollo depende del IETF. Las actividades que realiza el IETF se dividen en distintos grupos, llamados Working Groups (WG) con finalidades específicas, los cuales se clasifican en distintas áreas comunes (Aplicaciones, seguridad, estandarización, servicios de transporte, etc.). El IESG (Internet Engineering Steering Group) se encarga de coordinar y dirigir al IETF por medio de los directores de área, que controlan las actividades número de los Working Groups que se encuentren dentro de cada área.

Las tareas de coordinación de los números asignados a los distintos protocolos de Internet están a cargo de IANA (Internet Assigned Numbers Authority). Los protocolos definidos por el IETF y su grupo de dirección correspondiente IESG contienen ciertos valores tales como: direcciones de Internet, números de protocolos y de puertos, nombres por dominio, etc. La funcionalidad de IANA está en que todos estos parámetros deben ser únicos, y por tanto, debe existir un registro que controle los valores que se encuentran asignados.

Request for Comments.

Los documentos denominados Request for Comments (RFC) contienen información de gran interés acerca de Internet. Existen miles de estos documentos con información sobre cualquier aspecto relacionado con la red. Los RFC comenzaron a funcionar sobre el año 1969 como un medio informal de

intercambio de ideas entre la comunidad investigadores de temas concernientes a las redes. Estos documentos se distribuían inicialmente de forma impresa por correo convencional hasta que la transferencia de ficheros a través de FTP (File Transfer Protocol) se comenzó a utilizar. Con el paso del tiempo los RFC se han convertido en una manera más oficial de presentar los protocolos de Internet, aunque aún se crean algunos de estos documentos con carácter únicamente informativo.

Los RFC se utilizan actualmente para fines de investigación y desarrollo de Internet por el Network Working Group, y en ellos se documentan los protocolos y estándares ya existentes, o bien las propuestas de nuevos protocolos o nuevas versiones de los actuales esperándose que se conviertan en un estándar. A cada RFC se le asigna un número siempre distinto para poder identificarlo, incluso cuando un RFC ya existente se modifica o actualiza se obtendrá un nuevo documento con su propio número exclusivo. Por este motivo y como las revisiones se producen continuamente se hace necesario el uso de un índice en el que se puede encontrar el número correspondiente a la última revisión de un determinado documento.

Cualquiera que lo desee puede elaborar un texto para que sea editado y publicado como un nuevo RFC por medio de una persona que actúa como editor (consultar RFC 2200 para más información). Sin embargo, si lo que pretende documentar en un nuevo RFC es un protocolo estándar o la propuesta correspondiente para ello, primero se debe notificar al IESG (Internet Engineering Steering Group).

Para que un protocolo de Internet se convierta en un estándar debe pasar por una serie de estados o niveles. El nivel de proposición de protocolo es asignado cuando un protocolo tiene posibilidades de convertirse en un estándar en el futuro, siendo recomendables algunas pruebas y revisiones hasta que el IESG considere su avance. Después del nivel de proposición el protocolo puede pasar a considerarse como un "borrador" (draft standard). Esto sólo ocurrirá cuando hayan transcurrido al menos 6 meses desde el nivel anterior, permitiendo de esta manera que la comunidad de Internet evalúe y considere el proceso de estandarización. Durante otros 4 meses el protocolo permanecerá en este nivel mientras se hacen pruebas y se analizan los comentarios recibidos con la posibilidad de efectuar algún cambio. Finalmente, el protocolo puede llegar a convertirse en un estándar oficial de Internet a través del IESG cuando su funcionalidad ha quedado suficientemente demostrada.

El carácter abierto con que se trata a esta información sobre los aspectos de diseño de la red permite que Internet evolucione y se desarrolle de una manera rápida y eficaz. Cualquiera puede tener acceso a todos los RFC creados desde el comienzo, los cuales se conservan como información de consulta y registro.

Trabajo enviado y realizado por:
Gilda Isabel Valera Guerrera
isabelvalera55@hotmail.com
Matrícula: 97-3840
Universidad Tecnológica de Santiago
UTESA
Recinto Santo Domingo de Guzmán
Profesor:
Saidan Batista
Materia:
Sistema Operativo II