

Utilización de los Protocolos TCP/IP en el Desarrollo de un Servidor WEB para Windows 95

Judith Fernández Lizama
Alumna de la Maestría en Informática de UPIICSA-IPN.
Héctor Samuel García Salas
Profesor e Investigador del CIDETEC-IPN

En este trabajo se presenta el esquema de desarrollo de un servidor WEB para el sistema operativo Windows 95. Se incluye un análisis detallado de los protocolos TCP/IP (Protocolo de Control de Transmisión/Protocolo Internet) como base para las comunicaciones del servidor a través de Internet. Se expone la división por niveles, el direccionamiento en Internet, las técnicas de enrutado y los protocolos más importantes que utiliza el servidor Web: IP, TCP y HTTP. El manejo de las comunicaciones TCP/IP para Windows se basa en la biblioteca Winsock, por lo que se presenta el concepto de socket y las clases sockets más importantes que utiliza el servidor.

Introducción

Es ampliamente conocido que por la red Internet circula una gran cantidad de información de toda índole. Una de las principales herramientas que permiten administrar las comunicaciones y la circulación de la información en Internet es el servidor WEB. El WEB es un sistema de transferencia de información con base en hipertextos usado popularmente en Internet, con comunicaciones de da-

tos que operan de acuerdo a un modelo Cliente/Servidor.

Los usuarios acceden al servidor mediante alguna aplicación que entienda el lenguaje de programación del documento. El cliente es el programa encargado de hacer una petición de documentos y, una vez que la información es recibida, la despliega al usuario. El servidor es el programa que da respuesta a la solicitud mediante el envío del documento requerido.

guaje común que utilizan dos computadoras para intercambiar información. De esta manera, se define la comunicación en forma precisa. Cada una de las máquinas enlazadas a Internet maneja el protocolo de comunicaciones TCP/IP (que abarca a su vez una familia de protocolos). Esto hace posible que se puedan comunicar redes tan diversas.

Internet puede definirse como una red virtual formada por redes físicas interconectadas por enrutadores. El

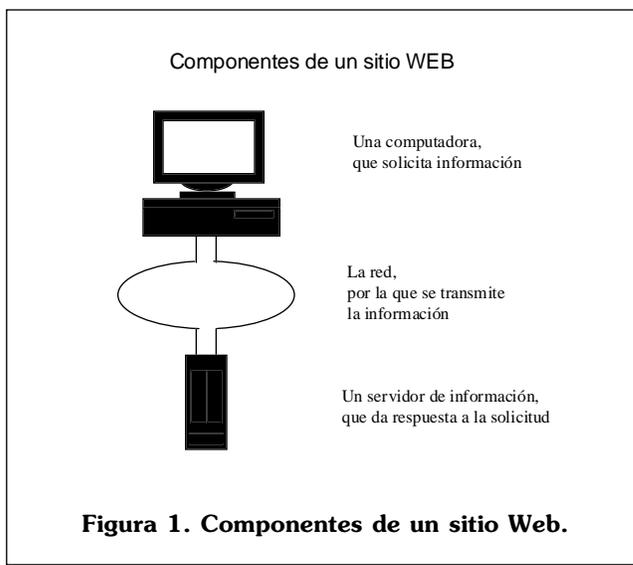


Figura 1. Componentes de un sitio Web.

Desarrollo

Internet y TCP/IP

Un protocolo de comunicaciones es un acuerdo que especifica un len-

direccionamiento ayuda a la programación TCP/IP a esconder los detalles físicos de la red y hace parecer a Internet como una unidad simple y uniforme. Para que Internet sea un sistema de comunicación universal, se necesita un método globalmente aceptado para identificar a todas las computadoras que se conectan a la red (hosts). TCP/IP tiene un esquema análogo al direccionamiento físico de red en el que a cada host se le asigna un número entero de 32 bits llamado dirección Internet o dirección IP, que se usa en todas las comunicaciones a ese host.

Las direcciones IP contienen dos campos (identificador de red; identificador de host) por lo que se pueden usar tanto para hacer referencia a un host individual como a una red. Los enrutadores, para pasar los paquetes de una red a otra, usan la porción del identificador de red, no la del host.

Para asignar a las máquinas nombres fáciles de recordar, TCP/IP utiliza el Sistema de Nombres de Dominio (DNS). Mediante servidores de nombres jerárquicos se mapean estos nombres a las direcciones IP.

División en capas

Los módulos de software del protocolo se dividen verticalmente en capas. Cada capa se responsabiliza con el manejo de una parte del problema, toma decisiones sobre la corrección del mensaje y elige acciones apropiadas basadas en el tipo o la dirección de destino. TCP/IP presenta 4 capas, que son:

1. **Aplicación:** Al mayor nivel, los usuarios invocan a los programas de aplicación que permiten el acceso a los servicios disponibles en una red TCP/IP.
2. **Transporte:** Su tarea principal es proporcionar la comunicación de una aplicación a otra. Esta comunicación se llama frecuentemente end-to-end.
3. **Red:** Maneja la comunicación de una máquina a otra. Acepta una solicitud para enviar un paquete desde la capa de transporte, junto con la dirección destino.
4. **Físico:** Es el nivel más bajo de TCP/IP. Transmite la información a través de una red específica.

Protocolo Internet (IP)

El servidor WEB utiliza como protocolo de red al IP. El servicio de IP se define como entrega de paquetes sin conexión, del mayor esfuerzo y no confiable (la entrega no se garantiza). Los paquetes pueden perderse, duplicarse, demorarse o entregarse fuera de orden, y el servicio no detecta estas condiciones ni lo informa al que envía ni al que recibe.

Se caracteriza como protocolo sin conexión porque trata de forma independiente a cada paquete; así, una secuencia de paquetes puede viajar de una máquina a otra por diferentes caminos, unos perderse y otros no. Se llama también del mayor esfuerzo, porque la red no elimina paquetes arbitrariamente, sino sólo cuando se acaban los recursos o existe alguna falla.

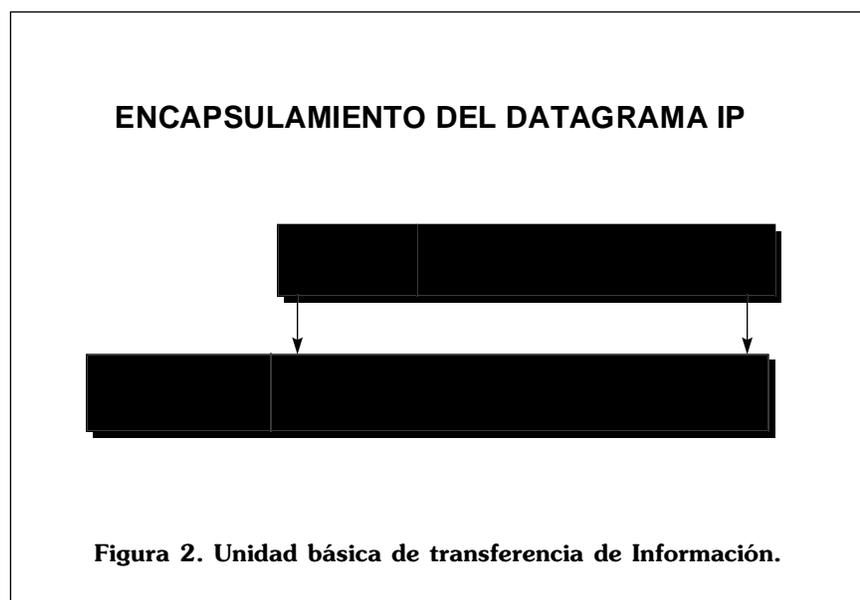
IP define la unidad básica de transferencia de información: el datagrama, que se encapsula en la trama física de la red (ver **figura 2**). Además, realiza la función de enrutado seleccionando un camino a través del cual se entregue la información e incluye un conjunto de reglas que

caracterizan la entrega como, por ejemplo, las condiciones bajo las cuales se eliminan los paquetes.

En un sistema de entrega de paquetes, el enrutado es el proceso de elegir un camino a través del cual mandar los paquetes, y un enrutador se refiere a cualquier computadora que realice esta elección. El algoritmo de enrutado IP debe seleccionar cómo mandar un datagrama a través de múltiples redes físicas.

Para enviar un datagrama desde su origen a su destino, el IP necesita sólo decidir cuál nodo recibirá próximamente el datagrama. Cada instancia del IP mantiene un directorio llamado tabla de rutas, usualmente colocado en memoria. Cada entrada de la tabla tiene dos valores: un destino y un próximo salto a ese destino (próximo enrutador en el camino hacia el destino).

Cuando un datagrama se enruta, se compara con el contenido de la tabla para determinar el próximo salto, hasta que llegue a su destino.



Protocolo de Control de Transporte (TCP)

Como protocolo de transporte, el servidor WEB utiliza al TCP. El TCP define una entrega de cadenas de información confiables. Aunque forma parte del TCP/IP, TCP es un protocolo independiente que puede utilizarse con cualquier otro sistema de entrega.

TCP usa números de puertos de protocolo para identificar el último destino en una máquina. Utiliza un concepto algo complejo: la abstracción de la conexión, en la que los objetos a identificar no son puertos sino circuitos virtuales de conexión que se identifican por un par de puntos finales. Los puntos finales son pares de números enteros que definen la dirección del host y del puerto TCP dentro del host. Como TCP identifica una conexión por un par de puntos finales, un puerto TCP dado puede establecer múltiples conexiones en una misma máquina. TCP es un protocolo orientado a conexión que requiere del acuerdo de los dos puntos finales para establecer la comunicación. Su unidad básica de información es el segmento, que se encapsula dentro de un datagrama IP.

Protocolo de Transferencia de Hipertextos (HTTP)

Es un protocolo del nivel de aplicaciones de TCP/IP, con la ligereza y velocidad necesarias para un sistema de información de hipertexto distribuido, como es el WEB, que ha utilizado HTTP desde 1990. Dentro del WEB, numerosos recursos de información están disponibles al usuario y existe una disciplina de referencia entre los diversos recursos, por lo que es posible seguir rápidamente las referencias entre unidades de información que provienen de localidades remotas diferentes. Esta disciplina se

basa en el Identificador Uniforme de Recursos (URI), para señalar el recurso al que se debe aplicar un método.

El caso más simple de una comunicación HTTP es la conexión directa entre un usuario y un servidor origen (donde se encuentra físicamente el recurso). Una situación más compleja ocurre cuando uno o más intermediarios se presentan en la cadena solicitud-respuesta. Existen tres tipos de intermediarios: proxy, gateway y túnel. El proxy recibe la solicitud de un URI en su forma absoluta y reescribe todo el mensaje (o parte) para transmitirlo al servidor identificado en el URI. El gateway es un agente receptor, que actúa como una capa por encima de otros servidores y, si es necesario, traduce las solicitudes al protocolo del servidor. Un túnel actúa como punto de transmisión entre dos conexiones, sin cambiar los mensajes.

Usualmente un mensaje de solicitud o respuesta que viaja del usuario al servidor origen, debe pasar a través de varias conexiones separadas. Algunas opciones de la comunicación HTTP aplican sólo al vecino más cercano que no sea túnel, a los puntos finales de la cadena o a todas las conexiones a través de la cadena.

Cualquier parte de la comunicación que no actúe como túnel puede emplear un caché interno para manejar las solicitudes. El caché reduce la cadena de transmisión si uno de los intermediarios tiene una respuesta en su caché que se aplique a la solicitud. No todas las respuestas pueden guardarse en un caché, pero las reglas que consideran esto no son estándares.

El puerto predefinido para HTTP es el TCP 80, pero pueden utilizarse otros. HTTP puede implantarse en cualquier otro protocolo o tipo de red, siempre que provea un transporte confiable.

Sockets

Los sockets se han convertido en un modelo popular que utilizan las aplicaciones para tener acceso a los recursos de una red. Un socket es un extremo de una conexión de dos caminos entre dos programas que corren en una red (son conexiones a bajo nivel). Los sockets se usan para implementar la conexión entre un programa servidor y un cliente; el cliente y el servidor deben acordar un protocolo. Una vez que una aplicación ha establecido un socket, puede usarlo para transmitir y recibir datos.

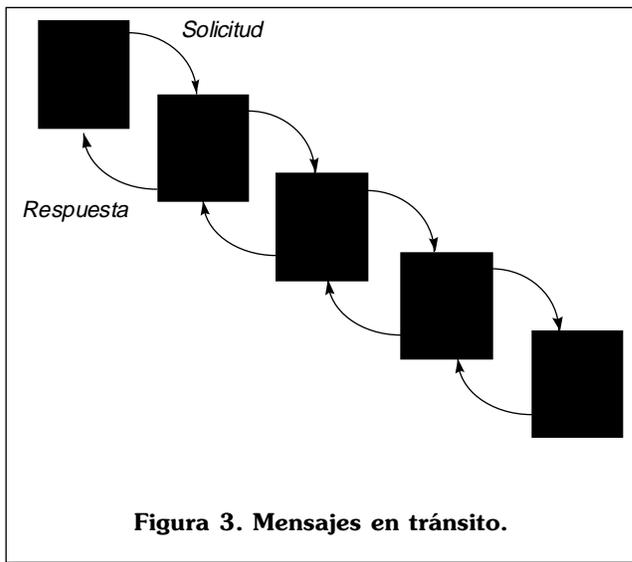


Figura 3. Mensajes en tránsito.

Para las aplicaciones Windows (por ende, para el servidor WEB), el acceso a TCP/IP es posible gracias a la utilización de la interface socket de Windows: el API Winsock. La versión más actual de Winsock es la 2.0 y vale señalar que además de TCP/IP soporta protocolos como IPX/SPX.

Winsock se basa en el estándar UNIX de Berkeley, por lo que incluye tanto sus funciones puras como extensiones de estas funciones para Windows. Por ejemplo, la función socket() es del estándar UNIX y crea un punto final para la comunicación retornando un descriptor del socket creado. Su extensión para Windows es WSASocket(), que realiza una función idéntica, pero incluye la creación de sockets traslapados o por grupos. Otras de las funciones que más utiliza el servidor son: accept(), que reconoce una conexión de entrada y la asocia al socket; connect(), que inicia una conexión en el socket especificado; listen(), que revisa las conexiones de entrada en un socket, y recv() y send(), que recibe o envía información por un socket.

Interconexión

El diseño consiste de una serie de módulos, contiendo partes públicas y privadas, con sus respectivas restricciones de acceso. La conexión de los módulos del servidor se basa en el modelo de la **figura 4**.

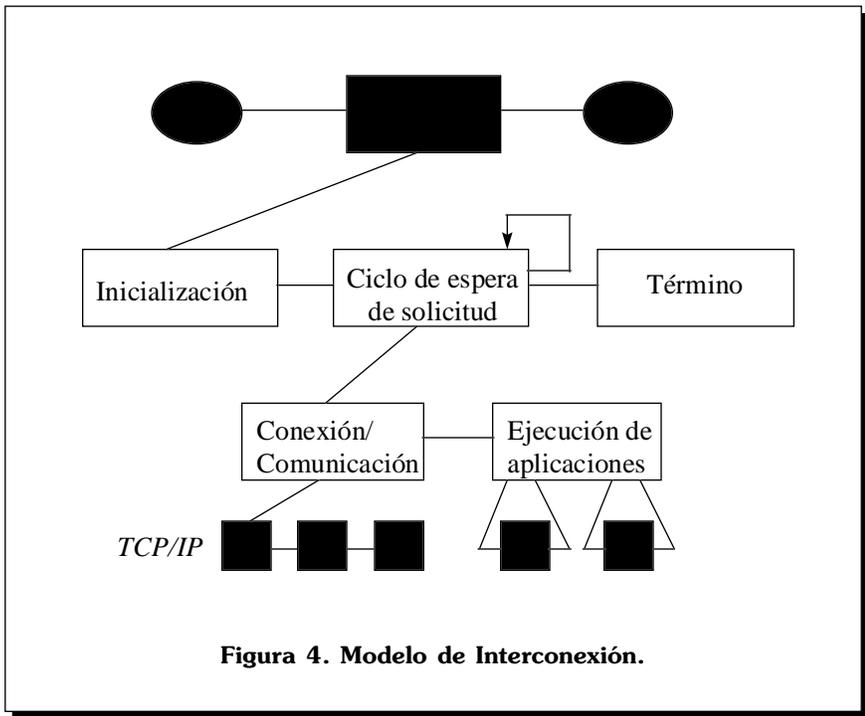


Figura 4. Modelo de Interconexión.

Conclusiones

Actualmente todas las aplicaciones Internet están basadas en los protocolos TCP/IP. Aunque esto puede variar, los nuevos protocolos deberán tener las características de internacionalización que permiten establecer comunicación entre cualquier tipo de máquina y sistema operativo. Es por esto que el conocimiento y manejo de protocolos de este estilo es muy importante para el desarrollo de aplicaciones de comunicaciones externas e internas. Desde el punto de vista académico, nuestra investigación ha permitido preparar personal especializado que actualmente produce aplicaciones de este estilo.

Referencias

- [1] Berners-Lee, T.; Fielding, R.; Frystyk, H. "Request for Comments: 1945, Hypertext Transfer Protocol— HTTP/ 1.0". Network Working Group, May 1996.
- [2] Comer, Douglas E. "Networking with TCP/IP. Principles, protocols and architecture", Volume I, II, III, Second Edition, Prentice Hall, 1991.
- [3] December, John. "Web Development" [www.december.com/web/develop.html] .
- [4] "Hypertext Transfer Protocol - Next Generation (NG)" [http://www.w3.org/pub/WWW/Protocols/HTTP-NG/Overview.html].
- [5] Peterson, David M. "TCP/IP Networking, A guide to the IBM environment", McGraw-Hill, 1995.
- [6] Schildt, Herbert. "Windows 95 Programming in C and C++", McGraw-Hill, 1995.