REDES DE COMPUTADORAS

Un enfoque descendente

QUINTA EDICIÓN

REDES DE COMPUTADORAS

Un enfoque descendente

QUINTA EDICIÓN

James F. Kurose
University of Massachusetts, Amherst

Keith W. Ross
Politechnic Institute of NYU

REVISIÓN TÉCNICA

Carolina Mañoso Hierro
Profesora Titular de Universidad
Dpto. de Sistemas de Comunicación y Control
Escuela Técnica Superior de Ingeniería Informática
Universidad Nacional de Educación a Distancia

Ángel Pérez de Madrid y Pablo
Profesor Titular de Universidad
Dpto. de Sistemas de Comunicación y Control
Escuela Técnica Superior de Ingeniería Informática
Universidad Nacional de Educación a Distancia

REVISIÓN TÉCNICA PARA LATINOAMÉRICA

Luis Marrone
Ingeniero Electromecánico - Profesor Titular
Fac. de Informática, Universidad Nacional de
La Plata, Buenos Aires (Argentina)

Ingeniero Rubin Ayma Alejo Fedor
Director del Dpto. de Tecnología Informática
Fac. de Ingeniería, Universidad Argentina
de la Empresa, Buenos Aires (Argentina)

Paula Venosa
Licenciada en Informática - Profesora Adjunta
Fac. de Informática, Universidad Nacional de
La Plata, Buenos Aires (Argentina)

Ingeniero Carlos Alberto Binker
Coordinador de la especialidad de Redes de
Ingeniería en Informática
Universidad Nacional de la Matanza,
San Justo, Buenos Aires (Argentina)

Ingeniero Mario Groppo
Coordinador del Laboratorio de Redes del Departamento de Sistemas
Universidad Tecnológica Regional de Córdoba, Córdoba (Argentina)

Addison Wesley
es un sello editorial de

PEARSON

Harlow, England • London • New York • Boston • San Francisco • Toronto • Sydney • Singapore • Hong Kong • Tokyo • Seoul • Taipei • New Delhi • Cape Town Madrid • Mexico City • Amsterdam • Munich • Paris • Milan

REDES DE COMPUTADORAS: UN ENFOQUE DESCENDENTE
James F. Kurose, Keith W. Ross

PEARSON EDUCACIÓN, S. A. 2010
ISBN: 978-84-7829-119-9

Materia:, 004. Computadores

Formato: 195x250 mm Páginas: 844

ISBN: 978-84-7829-119-9

Depósito Legal:

Traducción y Maquetación: Vuelapluma, S.L.U.

Equipo editorial:
Editor: Miguel Martín-Romo
Técnico Editorial: Esther Martín

Equipo de producción:
Director: José A. Clares
Técnico: Isabel Muñoz

Diseño de cubierta: Equipo de diseño de Pearson Educación, S. A.

Impreso por:

IMPRESO EN ESPAÑA - PRINTED IN SPAIN

Este libro ha sido impreso con papel y tintas ecológicos

CAPÍTULO 1

Redes de computadoras e Internet

Hoy día, Internet es casi indiscutiblemente el sistema de ingeniería más grande creado por la mano del hombre, con cientos de millones de computadoras conectadas, enlaces de comunicaciones y switches; cientos de millones de usuarios que se conectan de forma intermitente a través de sus teléfonos móviles y sus PDA; y dispositivos tales como sensores, cámaras web, consolas de juegos, marcos de fotografías e incluso lavadoras que se conectan a Internet. Dado que Internet es una red tan enorme e incluye tantos componentes distintos y tiene tantos usos, ¿es posible tener la esperanza de comprender cómo funciona y, más concretamente, cómo funcionan las redes de computadoras? ¿Existen unos principios y una estructura básicos que puedan proporcionar una base para comprender un sistema tan asombrosamente complejo y grande? Y, en caso afirmativo, ¿es posible que pueda resultar tan interesante y divertido como para dedicarse a estudiar las redes de computadoras? Afortunadamente, la respuesta a todas estas preguntas es un rotundo SÍ. Ciertamente, el objetivo de este libro es el de iniciar al lector en el dinámico campo de las redes de computadoras, proporcionándole los principios y los conocimientos prácticos que necesitará para entender no sólo las redes actuales, sino también las del futuro.

En el primer capítulo se hace una amplia introducción al mundo de las redes de computadoras y de Internet. Nuestro objetivo es proporcionar una visión general y establecer el contexto para el resto del libro, con el fin de poder ver el bosque a través de los árboles. En este capítulo de introducción se abordan muchos de los fundamentos, así como muchos de los componentes que forman una red de computadoras, siempre sin perder de vista la panorámica general.

Vamos a estructurar esta introducción a las redes de computadoras de la siguiente forma: después de exponer algunos conceptos y términos fundamentales, examinaremos los componentes esenciales que forman una red de computadoras. Comenzaremos por la frontera de la red y echaremos un vistazo a los sistemas terminales y aplicaciones que se ejecu-

tan en la red. A continuación, exploraremos el núcleo de una red de computadoras, examinando los enlaces y los switches que transportan los datos, así como las redes de acceso y los medios físicos que conectan los sistemas terminales con el núcleo de la red. Aprenderemos que Internet es una red de redes y cómo estas redes se conectan entre sí.

Una vez completada la introducción sobre la frontera y el núcleo de una red de computadoras, en la segunda mitad del capítulo adoptaremos un punto de vista más amplio y abstracto. Examinaremos los retardos, las pérdidas y la tasa de transferencia de datos en una red de computadoras y proporcionaremos modelos cuantitativos simples para los retardos y tasas de transferencia de terminal a terminal: modelos que tienen en cuenta los retardos de transmisión, de propagación y de cola. A continuación, presentaremos algunos de los principios básicos sobre las arquitecturas de red, en concreto, las capas de protocolos y los modelos de servicios. También veremos que las redes son vulnerables a muchos tipos distintos de ataques; revisaremos algunos de estos ataques y veremos cómo es posible conseguir que las redes sean más seguras. Por último, concluiremos el capítulo con una breve historia de las redes de comunicaciones.

1.1 ¿Qué es Internet?

En este libro, vamos a emplear la red pública Internet, una red de computadoras específica, como nuestro principal vehículo para explicar las redes de computadoras y sus protocolos. Pero, ¿qué es Internet? Hay dos formas de responder a esta pregunta. La primera de ellas es describiendo las tuercas y tornillos que forman la red; es decir, los componentes hardware y software básicos que forman Internet. La segunda es describiéndola en términos de la infraestructura de red que proporciona servicios a aplicaciones distribuidas. Comenzaremos por la descripción de los componentes esenciales, utilizando la Figura 1.1 para ilustrar la exposición.

1.1.1 Descripción de los componentes esenciales

Internet es una red de computadoras que interconecta cientos de millones de dispositivos informáticos a lo largo de todo el mundo. No hace demasiado tiempo, estos dispositivos eran fundamentalmente computadoras PC de escritorio tradicionales, estaciones de trabajo Linux y los llamados servidores que almacenaban y transmitían información tal como páginas web y mensajes de correo electrónico. Sin embargo, cada vez más sistemas terminales no tradicionales como televisiones, computadoras portátiles, consolas de juegos, teléfonos móviles, cámaras web, sistemas de detección medioambientales y de automóviles y dispositivos de seguridad y electrodomésticos están conectados a Internet. Por tanto, el término *red de computadoras* está comenzando a quedar algo desactualizado a causa de la gran cantidad de dispositivos no tradicionales que están conectados a Internet. En la jerga de Internet, todos estos dispositivos se denominan **hosts** o **sistemas terminales**. En julio de 2008, había casi 600 millones de sistemas terminales conectados a Internet [ISC 2009], sin contar los teléfonos móviles, las computadoras portátiles y otros dispositivos que se conectan de forma intermitente a Internet.

Los sistemas terminales se conectan entre sí mediante una red de **enlaces de comunicaciones** y dispositivos de **conmutación de paquetes**. En la Sección 1.2 veremos que existen muchos tipos de enlaces de comunicaciones, los cuales están compuestos por diferentes

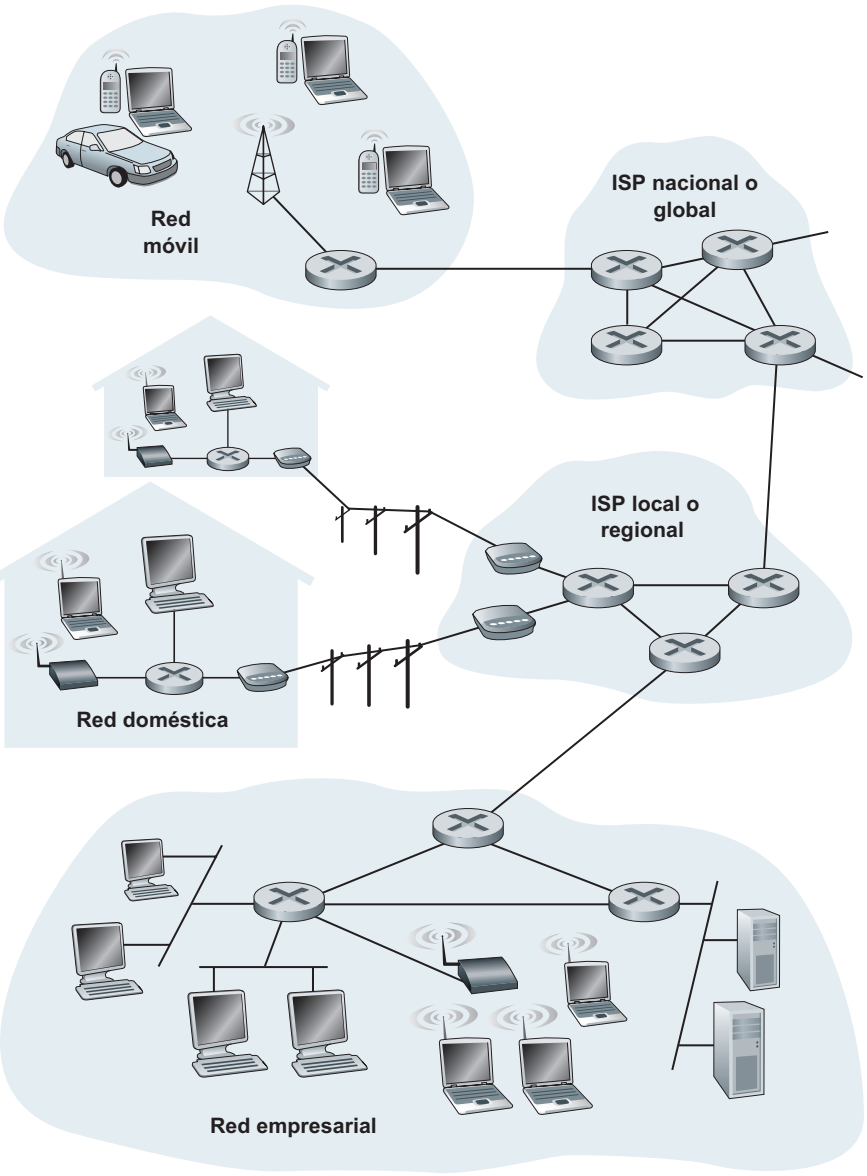


Figura 1.1 • Algunos de los componentes esenciales de Internet.

tipos de medios físicos, entre los que se incluyen el cable coaxial, el hilo de cobre, la fibra óptica y el espectro de radio. Los distintos enlaces pueden transmitir los datos a distintas velocidades y la **velocidad de transmisión** de un enlace se mide en bits/segundo. Cuando un sistema terminal tiene que enviar datos a otro sistema terminal, el emisor segmenta los datos y añade bits de cabecera a cada segmento. Los paquetes de información resultantes, conocidos como **paquetes** en la jerga informática, se envían entonces a través de la red hasta el sistema terminal receptor, donde vuelven a ser ensamblados para obtener los datos originales.

Un conmutador de paquetes toma el paquete que llega de uno de sus enlaces de comunicaciones de entrada y lo reenvía a uno de sus enlaces de comunicaciones de salida. Los dispositivos de conmutación de paquetes se suministran en muchas formas y modelos, pero los dos tipos más utilizados actualmente en Internet son los **routers** y los **switches de la capa de enlace**. Ambos tipos reenvían los paquetes hacia sus destinos finales. Los switches de la capa de enlace normalmente se emplean en las redes de acceso, mientras que los routers suelen utilizarse en el núcleo de la red. La secuencia de enlaces de comunicaciones y conmutadores de paquetes que atraviesa un paquete desde el sistema terminal emisor hasta el sistema terminal receptor se conoce como **ruta** a través de la red. Es difícil estimar la cantidad exacta de tráfico que se transporta a través de Internet [Odylsko 2003]. Pri-Metrica [PriMetrica 2009] estima que, en 2008, los proveedores de Internet emplearon 10 terabits por segundo de capacidad internacional y que dicha capacidad se duplica aproximadamente cada dos años.

Las redes de conmutación de paquetes (que transportan paquetes) son similares en muchos aspectos a las redes de transporte formadas por autopistas, carreteras e intersecciones (que transportan vehículos). Por ejemplo, imagine que una fábrica necesita trasladar un enorme cargamento a un cierto almacén de destino que se encuentra a miles de kilómetros. En la fábrica, el cargamento se reparte y se carga en una flota de camiones. Cada camión hace el viaje hasta el almacén de destino de forma independiente a través de la red de autopistas, carreteras e intersecciones. En el almacén de destino, la carga de cada camión se descarga y se agrupa con el resto del cargamento a medida que va llegando. Luego, en cierto sentido, los paquetes son como los camiones, los enlaces de comunicaciones como las autopistas y carreteras, los dispositivos de conmutación de paquetes como las intersecciones y los sistemas terminales son como los edificios (la fábrica y el almacén). Al igual que un camión sigue una ruta a través de la red de transporte por carretera, un paquete sigue una ruta a través de una red de computadoras.

Los sistemas terminales acceden a Internet a través de los **ISP (*Internet Service Provider*, Proveedor de servicios de Internet)**, incluyendo los ISP residenciales como son las compañías telefónicas o de cable locales; los ISP corporativos; los ISP universitarios y los ISP que proporcionan acceso inalámbrico (WiFi) en aeropuertos, hoteles, cafés y otros lugares públicos. Cada ISP es en sí mismo una red de conmutadores de paquetes y enlaces de comunicaciones. Los ISP proporcionan una amplia variedad de tipos de acceso a red a los sistemas terminales, entre los que se incluyen el acceso a través de módem de acceso telefónico a 56 kbps, el acceso de banda ancha residencial, mediante módem por cable o DSL, el acceso LAN (*Local Area Network*, Red de área local) de alta velocidad y el acceso inalámbrico. Los ISP también proporcionan acceso a Internet a los proveedores de contenido, conectando sitios web directamente a Internet. Internet es todo lo que conecta a los sistemas terminales entre sí, por lo que los ISP que proporcionan el acceso a los sistemas terminales también tienen que estar interconectados entre ellos. Estos ISP de nivel inferior se interco-

nectan a través de los ISP de nivel superior nacionales e internacionales, como AT&T y Sprint. Un ISP de nivel superior consiste en routers de alta velocidad interconectados a través de enlaces de fibra óptica de alta velocidad. La red de cada ISP, sea de nivel inferior o superior, se administra de forma independiente, ejecuta el protocolo IP (véase más adelante) y se ajusta a determinados convenios de denominación y de asignación de direcciones. En la Sección 1.3 examinaremos más detalladamente los ISP y sus interconexiones.

Los sistemas terminales, los conmutadores de paquetes y otros dispositivos de Internet ejecutan **protocolos** que controlan el envío y la recepción de la información dentro de Internet. El protocolo **TCP (*Transmission Control Protocol*, Protocolo de control de transmisión)** y el protocolo **IP (*Internet Protocol*, Protocolo de Internet)** son dos de los protocolos más importantes de Internet. El protocolo IP especifica el formato de los paquetes que se envían y reciben entre los routers y los sistemas terminales. Los principales protocolos de Internet se conocen colectivamente como protocolos **TCP/IP**. En este capítulo de introducción comenzaremos a estudiar los protocolos, pero esto sólo es el principio, ya que gran parte del libro se dedica a los protocolos empleados por las redes de computadoras.

Debido a la importancia de los protocolos en Internet, es importante que todo el mundo esté de acuerdo en qué hacen todos y cada uno de ellos, siendo aquí donde entran en juego los estándares. Los **estándares de Internet** son desarrollados por el IETF (*Internet Engineering Task Force*) [IETF 2009]. Los documentos asociados a estos estándares IETF se conocen como documentos **RFC (*Request For Comments*, Solicitud de comentarios)**. Los RFC nacieron como solicitudes de comentarios de carácter general (de ahí su nombre) para solucionar los problemas de diseño de la red y de los protocolos a los que se enfrentó el precursor de Internet. El contenido de estos documentos suele ser bastante técnico y detallado. Definen protocolos tales como TCP, IP, HTTP (para la Web) y SMTP (para el correo electrónico). Actualmente, existen más de 5.000 documentos RFC. Existen también otros organismos dedicados a especificar estándares para componentes de red, más específicamente para los enlaces de red. El comité de estándares IEEE 802 LAN/MAN [IEEE 802 2009], por ejemplo, especifica los estándares para redes Ethernet y WiFi.

1.1.2 Descripción de los servicios

Hasta el momento hemos identificado muchos de los componentes que forman Internet, pero también podemos describir Internet desde un punto de vista completamente diferente, en concreto como *una infraestructura que proporciona servicios a las aplicaciones.* Entre estas aplicaciones se incluyen el correo electrónico, la navegación web, la mensajería instantánea, Voz sobre IP (VoIP), la radio por Internet, los flujos de vídeo, los juegos distribuidos, la compartición de archivos en redes entre iguales o entre pares (P2P, *Peer-to-peer*), la televisión a través de Internet, las sesiones remotas y otras muchas. Se dice que las aplicaciones son **aplicaciones distribuidas**, porque implican a varios sistemas terminales que intercambian datos entre sí. Es importante saber que las aplicaciones de Internet se ejecutan en los sistemas terminales, no en los conmutadores de paquetes disponibles en el núcleo de la red. Aunque los dispositivos de conmutación de paquetes facilitan el intercambio de datos entre sistemas terminales, no se preocupan de la aplicación que esté actuando como origen o destino de los datos.

Vamos a ahondar un poco más en lo que queremos decir con una infraestructura que proporciona servicios a las aplicaciones. Para ello, supongamos que tenemos una excitante

nueva idea para una aplicación distribuida de Internet, que puede beneficiar enormemente a la humanidad o que simplemente puede hacernos ricos y famosos. ¿Cómo podríamos transformar esa idea en una aplicación real de Internet? Puesto que las aplicaciones se ejecutan en los sistemas terminales, tendremos que escribir programas software que se ejecuten en dichos sistemas. Por ejemplo, podríamos escribir programas en Java, C o Python. Ahora bien, dado que estamos desarrollando una aplicación Internet distribuida, los programas que se ejecuten en los distintos sistemas terminales tendrán que enviarse datos entre sí. Y aquí es cuando llegamos al meollo de la cuestión, a la que nos lleva a la forma alternativa de describir Internet como una plataforma para aplicaciones. ¿Cómo una aplicación que se ejecuta en un sistema terminal instruye a Internet para entregar datos a otro programa que se ejecuta en otro sistema terminal?

Los sistemas terminales conectados a Internet proporcionan una **API (*Application Programming Interface*, Interfaz de programación de aplicaciones)**, que especifica cómo un programa de software que se ejecuta en un sistema terminal pide a la infraestructura de Internet que suministre datos a un programa de software de destino específico que se ejecuta en otro sistema terminal. La API de Internet consta de un conjunto de reglas que el programa que transmite los datos debe cumplir para que Internet pueda entregar esos datos al programa de destino. En el Capítulo 2 se aborda en detalle la API de Internet. Por el momento, veamos una sencilla analogía, una que emplearemos con frecuencia a lo largo de este libro. Supongamos que Alicia desea enviar una carta a Benito utilizando el servicio postal. Por supuesto, Alicia no puede escribir la carta (los datos) y lanzar la carta por la ventana. En lugar de ello, será necesario que Alicia introduzca la carta en un sobre, escriba el nombre completo de Benito, su dirección y código postal en el sobre, lo cierre y pegue un sello en la esquina superior derecha del sobre. Por último, tendrá que introducir el sobre en un buzón del servicio postal. Por tanto, el servicio postal de correos tiene su propia "API de servicio postal", es decir, su propio conjunto de reglas, que Alicia debe seguir para que el servicio de correos entregue su carta a Benito. De forma similar, Internet tiene una API que el programa que envía los datos debe seguir para que Internet entregue los datos al software que los recibirá.

Por supuesto, el servicio de correos proporciona más de un servicio a sus clientes, como correo urgente, acuse de recibo, correo ordinario y otros muchos. Del mismo modo, Internet proporciona múltiples servicios a sus aplicaciones. Cuando desarrolle una aplicación de Internet, también tendrá que seleccionar uno de los servicios de Internet para su aplicación. En el Capítulo 2 describiremos los servicios de Internet.

Esta segunda descripción de Internet, una infraestructura que permite proporcionar servicios a aplicaciones distribuidas, es muy importante. Cada vez más, las necesidades de las nuevas aplicaciones están dirigiendo los avances de los componentes esenciales de Internet. Por tanto, es importante tener presente que Internet es una infraestructura en la que se están inventando e implementando constantemente nuevas aplicaciones.

Aquí sólo hemos dado dos descripciones de Internet; una en términos de sus componentes esenciales y otra como infraestructura que permite proporcionar servicios a aplicaciones distribuidas. Pero es posible que todavía no tenga claro qué es Internet. ¿Qué es la conmutación de paquetes, TCP/IP y una API? ¿Qué son los routers? ¿Qué tipos de enlaces de comunicaciones existen en Internet? ¿Qué es una aplicación distribuida? ¿Cómo puede una tostadora o un sensor de temperatura conectarse a Internet? Si se siente un poco abrumado por todas estas preguntas, no se preocupe, el propósito de este libro es presentarle tanto los componentes hardware como software de Internet, así como los principios que regulan cómo y por qué funciona. En las siguientes secciones y capítulos explicaremos todos estos términos y daremos respuesta a estas cuestiones.

1.1.3 ¿Qué es un protocolo?

Ahora que ya hemos visto por encima para qué sirve Internet, vamos a ocuparnos de otro término importante en el mundo de las redes de computadoras: *protocolo*. ¿Qué es un protocolo? ¿Qué hace un protocolo*?*

Analogía humana

Probablemente, sea más sencillo comprender el concepto de protocolo de red considerando en primer lugar algunas analogías humanas, ya que las personas llevamos a cabo protocolos casi constantemente. Piense en lo que hace cuando necesita preguntar a alguien qué hora es. En la Figura 1.2 se muestra cómo se lleva a cabo un intercambio de este tipo. El protocolo entre personas (o las buenas maneras, al menos) dicta que para iniciar un proceso de comunicación con alguien lo primero es saludar (el primer "Hola" mostrado en la Figura 1.2). La respuesta típica a este saludo será también "Hola". Implícitamente, el saludo de respuesta se toma como una indicación de que se puede continuar con el proceso de comunicación y preguntar la hora. Una respuesta diferente al "hola" inicial (como por ejemplo, ¡No me moleste" o "No hablo su idioma", o cualquier otra respuesta impublicable no debemos escribir) indicará una indisposición o incapacidad para comunicarse. En este caso, el protocolo de las relaciones entre personas establece que no debe preguntarse la hora. En ocasiones, no se obtiene ninguna respuesta, en cuyo caso habrá que renunciar a preguntar a esa persona la hora que es. Tenga en cuenta que, en el protocolo entre personas, *existen mensajes específicos que*

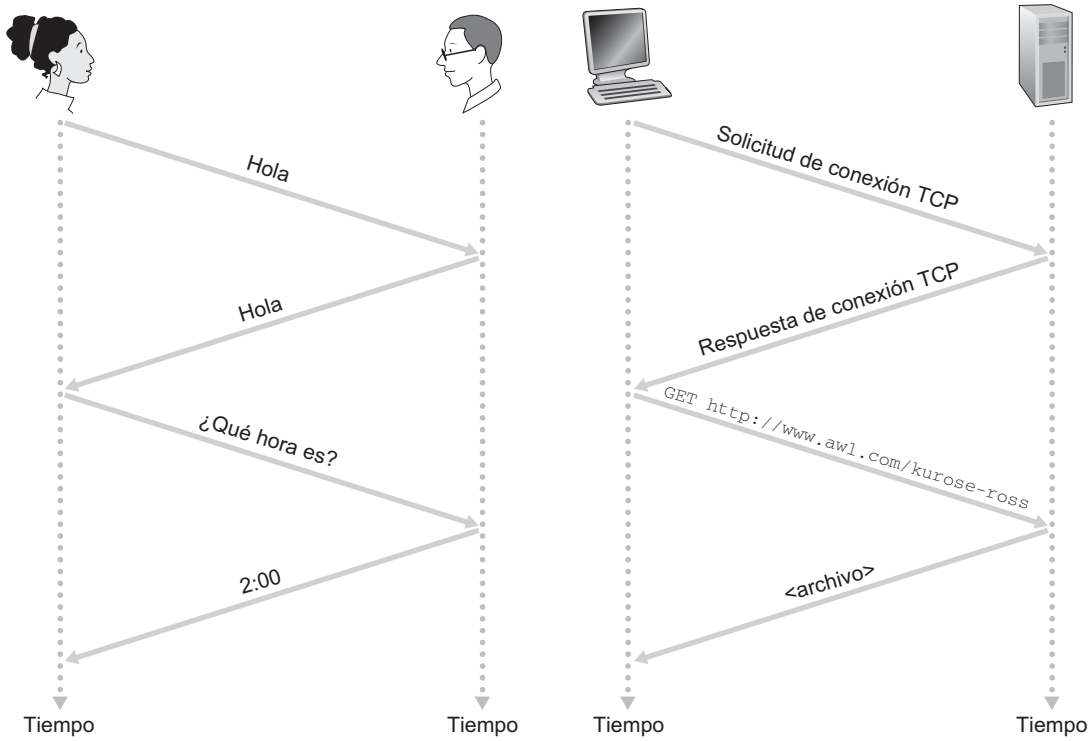


Figura 1.2 • Un protocolo humano y un protocolo de red.

enviamos y acciones específicas que tomamos como respuesta a los mensajes de contestación recibidos o a otros sucesos (como por ejemplo no recibir una respuesta en un periodo de tiempo determinado). Evidentemente, los mensajes transmitidos y recibidos y las acciones tomadas al enviar o recibir estos mensajes u otros sucesos desempeñan un papel principal en el protocolo humano. Si las personas adoptan protocolos diferentes (por ejemplo, si una persona guarda las formas pero la otra no lo hace, o si uno comprende el concepto de tiempo y el otro no), los protocolos no interoperarán y la comunicación no podrá tener lugar. Esta misma idea también es aplicable a las redes: es necesario que las entidades (dos o más) que deseen comunicarse ejecuten el mismo protocolo para poder llevar a cabo la tarea.

Consideremos ahora una segunda analogía humana. Suponga que está asistiendo a una clase (por ejemplo, sobre redes). El profesor está hablando acerca de los protocolos y usted no le comprende. El profesor detiene su explicación y dice: "¿Alguna pregunta?" (un mensaje dirigido a todos los estudiantes, que no están dormidos, y que todos ellos reciben). Usted levanta la mano (transmitiendo un mensaje implícito al profesor). El profesor le dirige una sonrisa y le dice "¿Si . . .?" (mensaje que le anima a plantear su pregunta, ya que los profesores *adoran* que les planteen cuestiones) y, a continuación, usted hace la pregunta (es decir, transmite su mensaje al profesor). El profesor escucha la pregunta (recibe su mensaje) y le responde (le transmite una respuesta). De nuevo, vemos que la transmisión y la recepción de mensajes y el conjunto de acciones convencionales tomadas cuando se envían y reciben estos mensajes, constituyen el núcleo de este protocolo de pregunta-respuesta.

Protocolos de red

Un protocolo de red es similar a un protocolo humano, excepto en que las entidades que intercambian mensajes y llevan a cabo las acciones son los componentes hardware o software de cierto dispositivo (por ejemplo, una computadora, una PDA, un teléfono móvil, un router u otro dispositivo de red). Cualquier actividad de Internet que implique dos o más entidades remotas que se comunican está gobernada por un protocolo. Por ejemplo, los protocolos implementados por hardware en las tarjetas de interfaz de red de dos computadoras conectadas físicamente controlan el flujo de bits a través del "cable" conectado entre las dos tarjetas de interfaz de red; los protocolos de control de congestión de los sistemas terminales controlan la velocidad a la que se transmiten los paquetes entre el emisor y el receptor; los protocolos de los routers determinan la ruta que seguirá un paquete desde el origen al destino. Los protocolos se ejecutan por todas partes en Internet y, en consecuencia, gran parte de este libro está dedicada a los protocolos de redes de computadoras.

Basándonos en un protocolo de red con el que probablemente estará familiarizado, vamos a ver lo que ocurre cuando se hace una solicitud a un servidor web, es decir, cuando usted escribe el URL de una página web en un navegador. Este escenario se ilustra en la mitad derecha de la Figura 1.2. En primer lugar, su computadora enviará un mensaje de solicitud de conexión al servidor web y esperará una respuesta. El servidor web recibirá su mensaje de solicitud y le devolverá un mensaje de respuesta de conexión. Sabiendo ahora que es posible solicitar el documento web, su computadora envía el nombre de la página web que desea extraer del servidor web mediante un mensaje GET. Por último, el servidor web envía la página web (archivo) a su computadora.

Basándonos en los ejemplos anteriores de protocolos humanos y de red, el intercambio de mensajes y las acciones tomadas cuando se envían y reciben estos mensajes constituyen los elementos claves para la definición de un protocolo:

*Un **protocolo** define el formato y el orden de los mensajes intercambiados entre dos o más entidades que se comunican, así como las acciones tomadas en la transmisión y/o la recepción de un mensaje u otro suceso.*

Internet, y las redes de computadoras en general, hacen un uso extensivo de los protocolos. Los distintos protocolos se utilizan para llevar a cabo las distintas tareas de comunicación. Como podrá leer a lo largo del libro, verá que algunos protocolos son simples y directos, mientras que otros son complejos e intelectualmente profundos. Dominar el campo de las redes de computadoras es equivalente a entender el qué, el por qué y el cómo de los protocolos de red.

1.2 La frontera de la red

En la sección anterior hemos presentado una introducción de carácter general sobre Internet y los protocolos de red. Ahora vamos a profundizar un poco más en los componentes de una red de computadoras (y de Internet, en concreto). Comenzaremos la sección en la frontera de una red y nos fijaremos en los componentes con los que estamos más familiarizados, es decir, las computadoras, las PDA, los teléfonos móviles y otros dispositivos que utilizamos a diario. En la siguiente sección nos desplazaremos desde la frontera de la red hasta el núcleo de la misma y examinaremos los procesos de conmutación y enrutamiento que tienen lugar en las redes.

Recuerde de la sección anterior que en la jerga de las redes informáticas, las computadoras y el resto de los dispositivos conectados a Internet a menudo se designan como sistemas terminales, porque se sitúan en la frontera de Internet, como se muestra en la Figura 1.3. Entre los sistemas terminales de Internet se incluyen las computadoras de escritorio (por ejemplo, PC de escritorio, computadoras Mac y equipos Linux), servidores (por ejemplo, servidores web y de correo electrónico) y equipos móviles (por ejemplo, computadoras portátiles, dispositivos PDA y teléfonos con conexiones a Internet inalámbricas). Además, una cantidad creciente de dispositivos alternativos están actualmente conectándose a Internet como sistemas terminales (véase el recuadro de la página siguiente).

Los sistemas terminales también se conocen como *hosts,* ya que albergan (es decir, ejecutan) programas de aplicación tales como navegadores web, servidores web, programas de lectura de mensajes de correo electrónico o servidores de correo electrónico. A lo largo de este libro utilizaremos indistintamente los términos host y sistema terminal; es decir, *host = sistema terminal*. En ocasiones, los hosts se clasifican en dos categorías: **clientes** y **servidores**. En general, podríamos decir que los clientes suelen ser las computadoras de escritorio y portátiles, las PDA, etc., mientras que los servidores suelen ser equipos más potentes que almacenan y distribuyen páginas web, flujos de vídeo, correo electrónico, etc.

1.2.1 Programas cliente y servidor

En el contexto del software de red, existe otra definición para los términos cliente y servidor, definición a la que haremos referencia a lo largo del libro. Un **programa cliente** es un programa que se ejecuta en un sistema terminal que solicita y recibe un servicio de un **programa servidor** que se ejecuta en otro sistema terminal. La Web, el correo electrónico, la

HISTORIA

UNA ASOMBROSA COLECCIÓN DE SISTEMAS TERMINALES DE INTERNET

No hace demasiado tiempo, los sistemas terminales conectados a Internet eran fundamentalmente computadoras tradicionales como los equipos de escritorio y los potentes servidores. Desde finales de la década de 1990 y hasta el momento actual, un amplio rango de interesantes dispositivos cada vez más diversos están conectándose a Internet. Todos estos dispositivos comparten la característica común de necesitar enviar y recibir datos digitales hacia y desde otros dispositivos. Dada la omnipresencia de Internet, sus protocolos bien definidos (estandarizados) y la disponibilidad de productos hardware preparados para Internet, lo lógico es utilizar la tecnología de Internet para conectar estos dispositivos entre sí.

Algunos de estos dispositivos parecen haber sido creados exclusivamente para el entretenimiento. Un marco de fotografías IP de escritorio [Ceiva 2009] descarga fotografías digitales de un servidor remoto y las muestra en un dispositivo que parece un marco para fotografías tradicional; una tostadora Internet descarga información metereológica de un servidor y graba una imagen de la previsión del día (por ejemplo, nubes y claros) en su tostada matutina [BBC 2001]. Otros dispositivos proporcionan información útil; por ejemplo, las cámaras web muestran el estado del tráfico y las condiciones meteorológicas o vigilan un lugar de interés, los electrodomésticos conectados a Internet, entre los que se incluyen lavadoras, frigoríficos y hornos, incorporan interfaces de tipo navegador web que permiten su monitorización y control remotos. Los telefónos móviles IP con capacidades GPS (como el nuevo iPhone de Apple) ponen al alcance de la mano la navegación por la Web, el uso del correo electrónico y de servicios dependientes de la ubicación. Una nueva clase de sistemas de sensores de red promete revolucionar la forma en que observaremos e interactuaremos con nuestro entorno. Los sensores en red integrados en nuestro entorno físico permiten la vigilancia de edificios, puentes, de la actividad sísmica, de hábitats de la fauna y la flora, de estuarios y de las capas inferiores de la atmósfera [CENS 2009, CASA 2009]. Los dispositivos biomédicos pueden estar integrados y conectados en red, dando lugar a numerosos problemas de seguridad e intimidad [Halperin 2008]. Un transpondedor RFID (identificación por radiofrecuencia) o un pequeño sensor integrado en cualquier objeto puede hacer que la información acerca del objeto esté disponible en Internet, lo que nos permitirá disfrutar de una "Internet de objetos" [ITU 2005].

transferencia de archivos, las sesiones remotas, los grupos de noticias y muchas otras aplicaciones populares adoptan el modelo cliente-servidor. Puesto que un programa cliente normalmente se ejecuta en una computadora y el programa servidor en otra, las aplicaciones Internet cliente-servidor son, por definición, **aplicaciones distribuidas**. El programa cliente y el programa servidor interactúan enviándose entre sí mensajes a través de Internet. En este nivel de abstracción, los routers, los enlaces y los restantes componentes de Internet sirven de forma colectiva como una caja negra que transfiere mensajes entre los componentes distribuidos entre los que se establece la comunicación de una aplicación de Internet. Este nivel de abstracción se ilustra en la Figura 1.3.

No todas las aplicaciones de Internet actuales están constituidas por programas cliente puros que interactúan con programas servidor puros. Cada vez más aplicaciones son aplica-

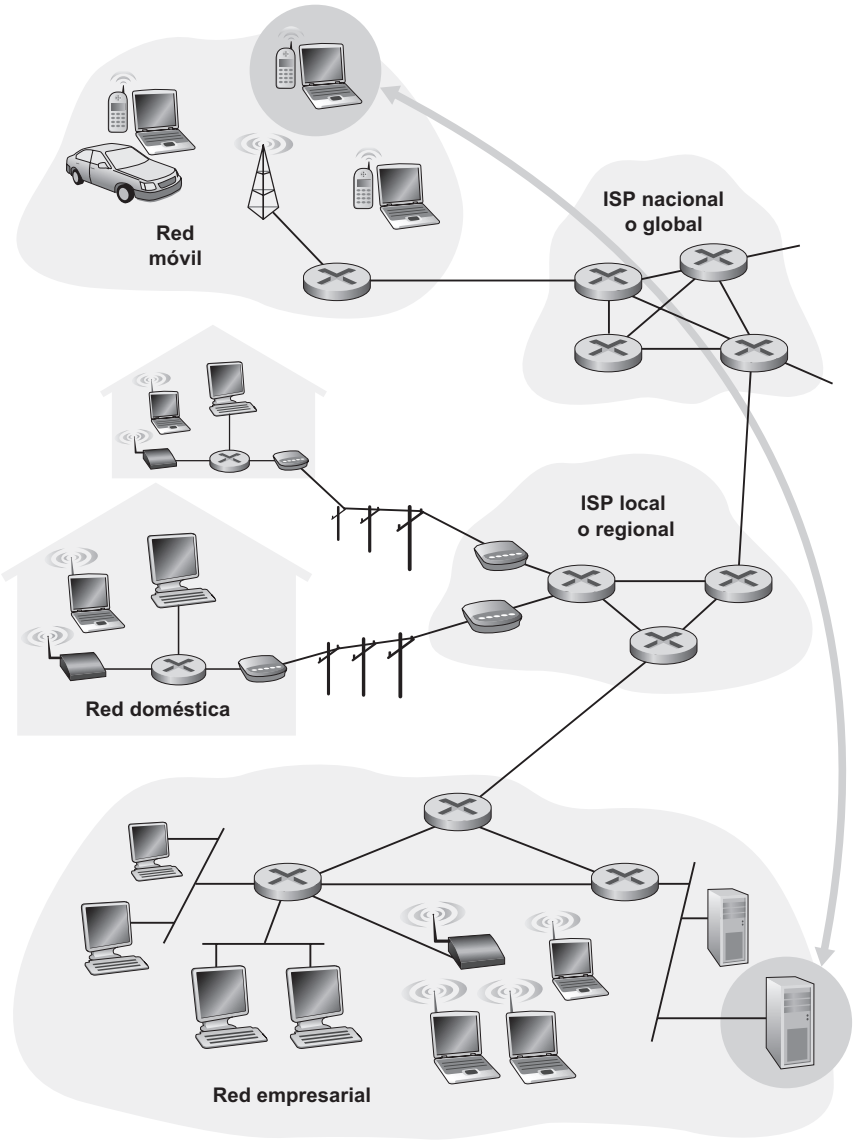


Figura 1.3 • Interacción de los sistemas terminales.

ciones entre iguales o entre pares (P2P, *Peer-to-Peer*), en las que los sistemas terminales interactúan y ejecutan programas que realizan tanto funciones de cliente como de servidor. Por ejemplo, en las aplicaciones de compartición de archivos P2P (como BitTorrent y eMule), el programa disponible en el sistema terminal del usuario actúa como cliente cuando solicita un archivo a un par y como servidor cuando envía un archivo a otro par. En la telefonía por Internet, las dos partes que intervienen en la comunicación interactúan como iguales (la sesión es simétrica, enviando y recibiendo ambas partes datos). En el Capítulo 2, compararemos y contrastaremos en detalle las arquitecturas cliente-servidor y P2P.

1.2.2 Redes de acceso

Una vez vistas las aplicaciones y los sistemas terminales existentes en la "frontera de la red", podemos pasar a ver las redes de acceso, los enlaces físicos que conectan un sistema terminal con el primer router (conocido también como "router de frontera") de una ruta entre el sistema terminal y cualquier otro sistema terminal distante. La Figura 1.4 muestra varios tipos de enlaces de acceso entre un sistema terminal y el router de frontera (los enlaces de acceso están resaltados mediante líneas más gruesas). En esta sección se repasan muchas de

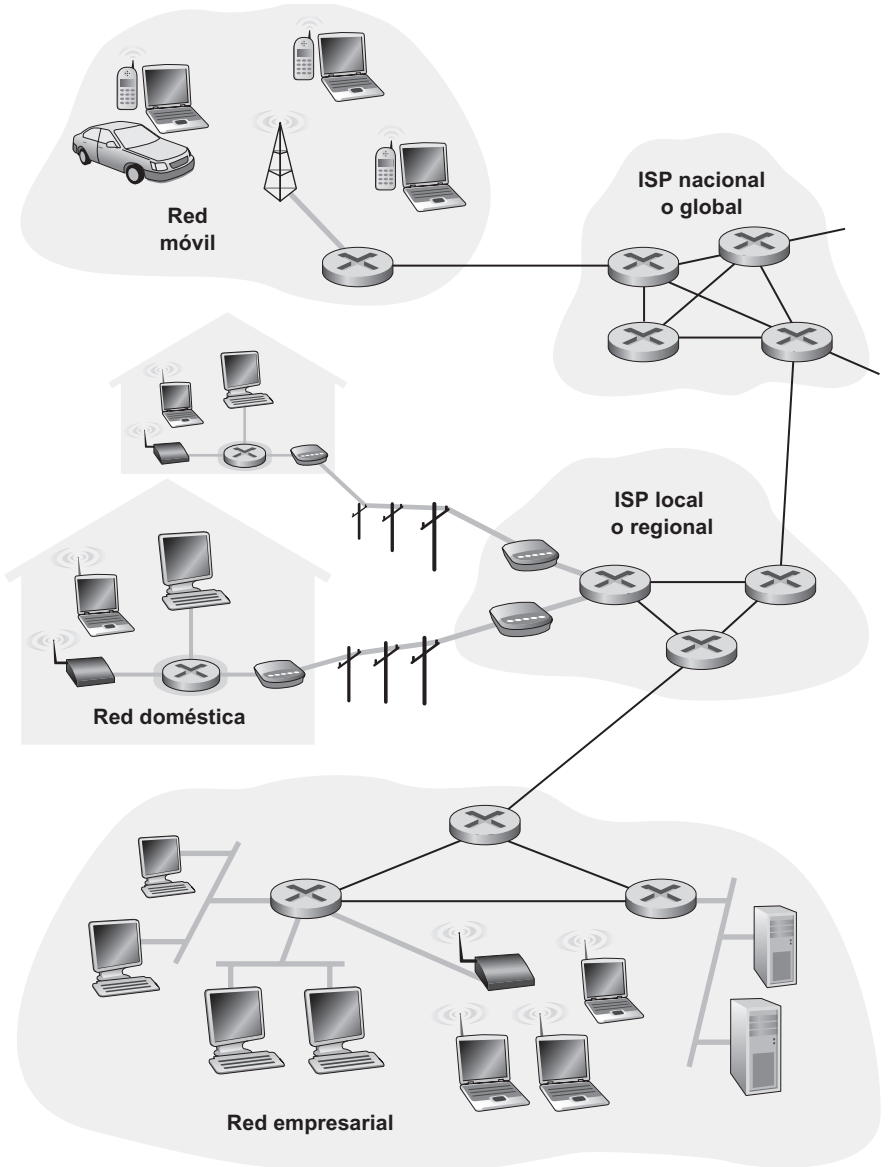


Figura 1.4 • Redes de acceso.

las tecnologías más comunes de las redes de acceso, desde las de baja velocidad hasta las de alta velocidad.

Enseguida veremos que muchas de estas tecnologías de acceso emplean, en distintos grados, partes de la infraestructura de la telefonía cableada tradicional local, la cual es proporcionada por la compañía de telefonía local, a la que haremos referencia simplemente como **telco** local. Algunos ejemplos de estas compañías serían Verizon en Estados Unidos y France Telecom en Francia. Cada residencia (chalet o piso) dispone de un enlace directo de cobre de par trenzado a un switch de la compañía telefónica, el cual se encuentra en un edificio denominado **central telefónica** en la jerga del campo de la telefonía. (Más adelante en esta sección explicaremos lo que es un cable de cobre de par trenzado.) Normalmente, una compañía telefónica local posee cientos de centrales telefónicas y enlaza a cada uno de sus clientes con la central más próxima.

Acceso telefónico

En la década de 1990, casi todos los usuarios residenciales accedían a Internet a través de las líneas telefónicas analógicas normales utilizando un módem de acceso telefónico. Actualmente, muchos usuarios de países subdesarrollados y de áreas rurales en países desarrollados (donde el acceso de banda ancha no está disponible) todavía tienen que acceder a Internet mediante una conexión de acceso telefónico. De hecho, se estima que el 10% de los usuarios residenciales de Estados Unidos utilizaban en 2008 conexiones de acceso telefónico [Pew 2008].

Se utiliza el término "acceso telefónico" (*dial-up*) porque el software del usuario realmente llama al número de teléfono de un ISP y establece una conexión telefónica tradicional con el mismo (por ejemplo, con AOL). Como se muestra en la Figura 1.5, el PC está conectado a un módem de acceso telefónico, que a su vez está conectado a la línea telefónica analógica del domicilio. Esta línea telefónica analógica está hecha de un hilo de cobre de paz trenzado y es la misma línea de teléfono que se emplea para las llamadas telefónicas ordinarias. El módem convierte la salida digital del PC en una señal analógica apropiada para ser transmitida a través de la línea telefónica analógica. En el otro extremo de la conexión, un módem del ISP convierte la señal analógica que recibe en una señal digital que será la señal de entrada para el router del ISP.

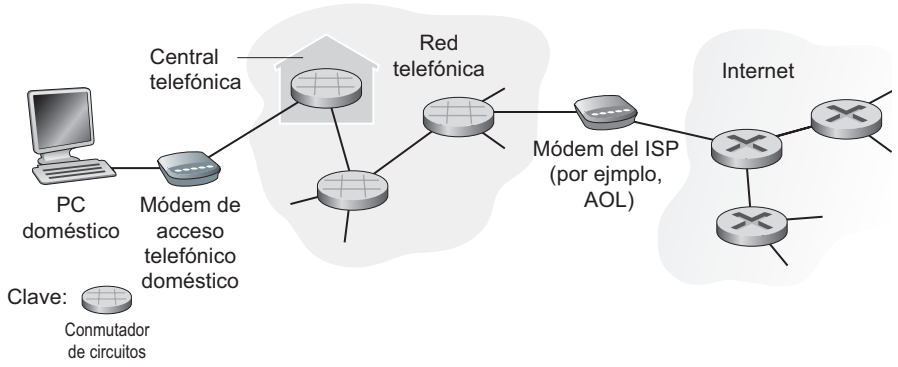


Figura 1.5 • Acceso telefónico a Internet.

El acceso telefónico a Internet presenta dos inconvenientes importantes. El primero y más destacable es que es extremadamente lento, proporcionando una velocidad máxima de 56 kbps. A esta velocidad, se tardan aproximadamente ocho minutos en descargar un archivo MP3 de una canción de tres minutos y se necesitarían varios días para descargar una película de 1 Gbyte. El segundo, un módem de acceso telefónico ocupa la línea telefónica del usuario, así que mientras que un miembro de la familia utiliza el módem de acceso telefónico para navegar por la Web, el resto de la familia no puede recibir ni hacer llamadas telefónicas normales a través de esa línea de teléfono.

DSL

Hoy en día, los dos tipos de acceso residencial de banda ancha predominantes son las líneas DSL (*Digital Subscriber Line*, Línea de abonado digital) y el cable. En la mayoría de los países desarrollados de hoy en día, más del 50% de los domicilios particulares disponen de acceso de banda ancha, con Corea del Sur, Islandia, Holanda, Dinamarca y Suiza a la cabeza con una penetración de más del 74% de los hogares en 2008 [ITIF 2008]. En Estados Unidos, las líneas DSL y cable tienen aproximadamente la misma cuota de mercado para el acceso de banda ancha [Pew 2008]. Fuera de Estados Unidos y Canadá domina la tecnología DSL, especialmente en Europa, donde en muchos países más del 90% de las conexiones de banda ancha se hacen mediante DSL.

Por regla general, los domicilios particulares contratan el servicio DSL de acceso a Internet con la misma empresa que le proporciona el acceso telefónico local (es decir, la compañía telefónica). Por tanto, cuando se utiliza el acceso mediante DSL, la compañía telefónica del cliente también actúa como ISP. Como se muestra en la Figura 1.6, cada módem DSL de un cliente utiliza la línea telefónica existente (hilo de cobre de par trenzado) para intercambiar datos con un multiplexor de acceso DSL (DSLAM), que normalmente se encuentra en la central de la compañía telefónica. La línea telefónica transporta simultáneamente los datos y las señales telefónicas, las cuales se codifican a frecuencias distintas:

- Un canal de descarga de alta velocidad opera en la banda de 50 kHz a 1 MHz.
- Un canal de carga de velocidad media opera en la banda de 4 kHz a 50 kHz.
- Un canal telefónico ordinario bidireccional opera en la banda de 0 a 4 kHz.

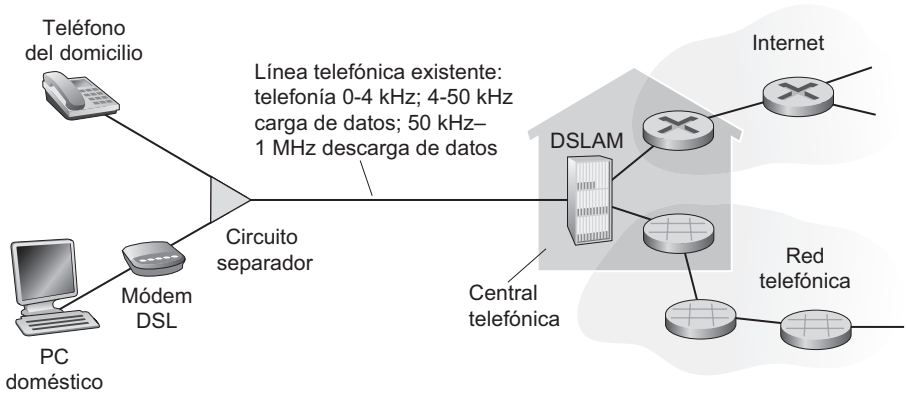


Figura 1.6 • Acceso mediante DSL a Internet.

Este método hace que un único enlace DSL se comporte como tres enlaces separados, de manera que una llamada de teléfono y una conexión a Internet pueden compartir el enlace DSL a un mismo tiempo. En la Sección 1.3.1 describiremos esta técnica de multiplexación por división en frecuencia. En el lado del cliente, las señales que llegan al domicilio son separadas como señales de datos y telefónicas mediante un circuito separador (*splitter*) que reenvía la señal de datos al módem DSL. En el lado de la compañía telefónica, en la central, el multiplexor DSLAM separa las señales de datos y de telefonía y envía los datos a Internet. Cientos o incluso miles de viviendas se conectan a un mismo DSLAM [Cha 2009, Dischinger 2007].

DSL presenta dos ventajas principales en comparación con el método de acceso telefónico a Internet. En primer lugar, puede transmitir y recibir datos a velocidades mucho más altas. Típicamente, un cliente DSL presentará una velocidad de transmisión en el rango comprendido entre 1 y 2 Mbps para las descargas (comunicaciones desde la central al domicilio del usuario) y de entre 128 kbps a 1 Mbps para las cargas (comunicaciones desde el domicilio a la central). Puesto que las velocidades de descarga y carga son diferentes, se dice que el acceso es *asimétrico*. La segunda ventaja importante es que los usuarios pueden hablar por teléfono y acceder a Internet simultáneamente. A diferencia del método de acceso telefónico, el usuario no tiene que llamar al número de teléfono del ISP para tener acceso a Internet; en su lugar, dispone de una conexión permanente "siempre activa" con el DSLAM del ISP (y por tanto con Internet).

Las velocidades de transmisión reales de descarga y de carga disponibles en el domicilio del usuario son función de la distancia entre la casa y la central telefónica, el calibre de la línea de par trenzado y el grado de interferencia eléctrica. Los ingenieros han diseñado expresamente sistemas DSL para distancias cortas entre el domicilio y la central, lo que ha permitido conseguir velocidades de transmisión sustancialmente mayores. Para incrementar la velocidad de transmisión de los datos, el sistema DSL se basa en algoritmos avanzados de procesamiento de señales y de corrección de errores, que pueden conducir a importantes retardos de los paquetes. Sin embargo, si el domicilio no se encuentra en un radio de entre 8 y 16 kilómetros de la central, la tecnología DSL de procesamiento de las señales ya no será tan efectiva y el usuario deberá recurrir a una forma alternativa de acceso a Internet.

Actualmente, existe también una amplia variedad de tecnologías DSL de alta velocidad que gozan de aceptación en muchos países. Por ejemplo, la tecnología VDSL (*Very-high speed DSL*), con la máxima penetración hoy día en Corea del Sur y Japón, proporciona velocidades de entre 12 y 55 Mbps para las descargas y velocidades de carga comprendidas entre 1,6 y 20 Mbps [DSL 2009].

Cable

Muchos domicilios de América del Norte y de muchos otros lugares reciben cientos de canales de televisión a través de redes de cable coaxial (veremos más adelante en esta sección el cable coaxial). En un sistema de televisión por cable tradicional, el **terminal de cabecera de cable** difunde los canales de televisión a través de una red de distribución de cable coaxial y amplificadores hasta los domicilios de los usuarios.

Mientras que la DSL y el acceso telefónico emplean la infraestructura de la telefonía local existente, el acceso por cable a Internet utiliza la infraestructura de la televisión por cable existente. Las casas obtienen el acceso por cable a Internet de la misma compañía que proporciona la televisión por cable. Como se ilustra en la Figura 1.7, la fibra óptica conecta el terminal de cabecera del cable a una serie de nodos de área situados en el vecindario, a

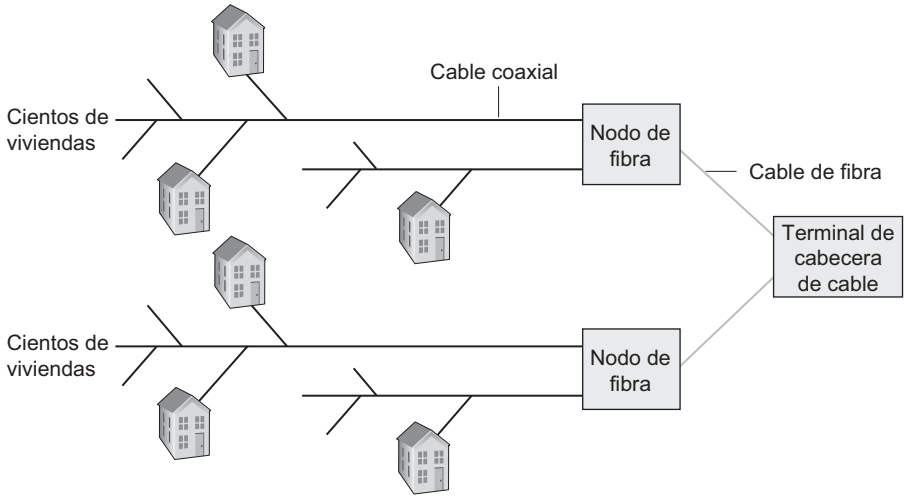


Figura 1.7 • Red de acceso híbrida de fibra óptica y cable coaxial.

partir de los cuales se utiliza el cable coaxial tradicional para llegar a todos los domicilios. Cada nodo de área suele dar soporte a entre 500 y 5.000 viviendas. Puesto que en este sistema se emplea tanto cable coaxial como fibra, a menudo se denomina sistema **HFC (*Hybrid Fiber Coax*, Híbrido de fibra y coaxial)**.

El acceso por cable a Internet requiere el uso de modems especiales, que se conocen como **modems por cable**. Al igual que un módem DSL, normalmente el módem por cable es un dispositivo externo que se conecta a un PC a través de un puerto Ethernet (en el Capítulo 5 veremos más detalles acerca de Ethernet). Los modems por cable dividen la red HFC en dos canales: un canal de descarga y un canal de carga. Al igual que en el caso de la DSL, el acceso suele ser asimétrico, teniendo normalmente el canal de descarga asignada una velocidad de transmisión mayor que el canal de carga.

Una característica importante del acceso a Internet por cable es que se trata de un medio de difusión compartido. Es decir, cada uno de los paquetes enviados por el terminal de cabecera se descargan a través de cada enlace hasta cada vivienda y los paquetes enviados desde las viviendas viajan a través del canal de carga hasta el terminal de cabecera. Así, si varios usuarios descargan simultáneamente un archivo de vídeo a través del canal de descarga, la velocidad real a la que cada usuario recibe su archivo de vídeo será significativamente menor que la velocidad acumulada de descarga por cable. Por el contrario, si sólo hay unos pocos usuarios activos que están navegando por la Web, cada uno de ellos recibirá las páginas web a la velocidad de descarga máxima del cable, ya que los usuarios rara vez solicitarán una página web al mismo tiempo. Puesto que el canal de carga también está compartido, se necesita un protocolo distribuido de acceso múltiple para coordinar las transmisiones y evitar las colisiones (veremos el problema de las colisiones en detalle en el Capítulo 5 al abordar la tecnología Ethernet).

En favor de la tecnología DSL debemos apuntar que se trata de una conexión punto a punto entre la vivienda y el ISP y que, por tanto, toda la capacidad de transmisión del enlace DSL entre el domicilio y el ISP está dedicada en lugar de ser compartida. Sin embargo, podemos decir en favor de la transmisión por cable que una red HFC correctamente dimensionada proporciona velocidades de transmisión más altas que la DSL. Existe una batalla

feroz entre las tecnologías DSL y HFC para el acceso residencial de alta velocidad, especialmente en América del Norte. En las áreas rurales, donde no está disponible ninguna de estas tecnologías, se puede utilizar un enlace vía satélite para conectar una vivienda con Internet a velocidades superiores a 1 Mbps; StarBand y HughesNet son dos proveedores de acceso vía satélite.

Tecnología FTTH (Fiber-To-The-Home, Fibra hasta el hogar)

La fibra óptica (que veremos en la Sección 1.2.3) puede ofrecer velocidades de transmisión significativamente más altas que el cable de cobre de par trenzado o el cable coaxial. En muchos países, algunas compañías telefónicas han tendido recientemente conexiones de fibra óptica desde sus centrales hasta las viviendas, proporcionando acceso a Internet de alta velocidad, así como servicios de telefonía y televisión por fibra óptica. En Estados Unidos, Verizon ha sido especialmente agresiva en el mercado de la tecnología FTTH, a través de su servicio FIOS [Verizon FIOS 2009].

Existen varias tecnologías que compiten por la distribución a través de fibra óptica desde las centrales a los hogares. La red de distribución óptica más simple se denomina **fibra directa**, en la que existe una fibra que sale de la central hasta cada domicilio. Este tipo de distribución puede proporcionar un ancho de banda grande, dado que cada cliente dispone de su propia fibra dedicada todo el camino hasta la central. Sin embargo, lo más habitual es que cada fibra saliente de la central sea compartida por muchas viviendas y ésta no se divida en fibras individuales específicas del cliente hasta llegar a un punto muy próximo a las viviendas. Hay disponibles dos arquitecturas de distribución de fibra óptica que llevan a cabo esta separación: las **redes ópticas activas (AON, *Active Optical Network*)** y las **redes ópticas pasivas (PON, *Passive Optical Network*)**. Las redes AON son fundamentalmente redes Ethernet conmutadas, las cuales abordaremos en el Capítulo 5. Aquí vamos a ver brevemente las redes ópticas pasivas, que se utilizan en el servicio FIOS de Verizon. La Figura 1.8 muestra el uso de la tecnología FTTH utilizando la arquitectura de distribución PON. Cada vivienda dispone de una terminación de red óptica (ONT, *Optical Network Terminator*), que se conecta a un distribuidor del vecindario mediante un cable de fibra óptica dedicado. El distribuidor combina una cierta cantidad de viviendas (normalmente menos de 100) en un único cable de fibra óptica compartido, que se conecta a una terminación de línea óptica (OLT,

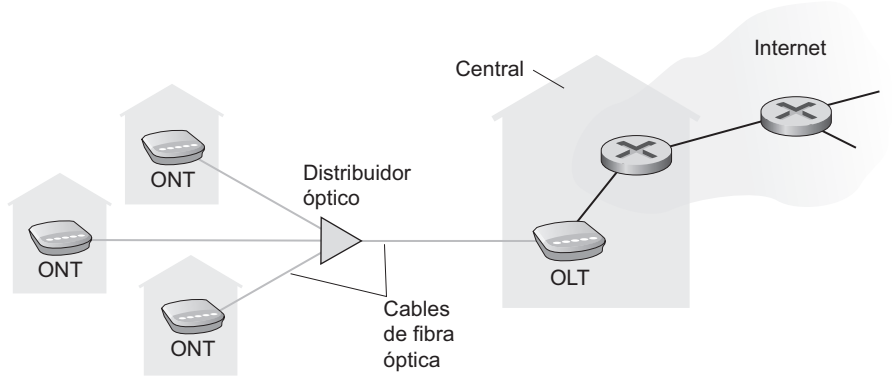


Figura 1.8 • Acceso a Internet mediante FTTH.

Optical Line Terminator) de la central de la compañía telefónica. La OLT, que realiza la conversión de señales ópticas en eléctricas, se conecta a través de Internet mediante un router de la compañía telefónica. En los domicilios, los usuarios conectan su router doméstico (normalmente un router inalámbrico) con la ONT y acceden a Internet a través de este router. En la arquitectura PON, todos los paquetes enviados desde la OLT al distribuidor se replican en este distribuidor (de forma similar a un terminal de cabecera de cable).

En teoría, la tecnología FTTH puede proporcionar velocidades de acceso a Internet del orden de los gigabits por segundo. Sin embargo, la mayoría de los ISP de FTTH ofrecen diferentes velocidades, siendo lógicamente más caras cuanto más altas son. La mayoría de los clientes actuales de la tecnología FTTH disfrutan de velocidades de descarga comprendidas entre 10 y 20 Mbps, y de velocidades de carga de entre 2 y 10 Mbps. Además del acceso a Internet, la fibra óptica permite proporcionar servicios de televisión y el servicio de telefonía tradicional.

Ethernet

En los campus universitarios y corporativos, normalmente se utiliza una red de área local (LAN, *Local Area Network*) para conectar un sistema terminal al router de frontera. Aunque existen muchos tipos de tecnologías LAN, Ethernet es con mucho la tecnología de acceso predominante en las redes corporativas y universitarias. Como se ilustra en la Figura 1.9, los usuarios de Ethernet utilizan cable de cobre de par trenzado para conectarse a un switch Ethernet (tecnología que se verá en detalle en el Capítulo 5). Con acceso Ethernet, normalmente los usuarios disponen de velocidades de acceso de 100 Mbps, y los servidores pueden alcanzar velocidades de 1 Gbps o incluso 10 Gbps.

WiFi

Cada vez es más habitual que los usuarios accedan a Internet a través de conexiones inalámbricas, bien a través de una computadora portátil o mediante un dispositivo móvil, como un iPhone, una Blackberry o un teléfono Google (véase el recuadro anterior "Una asombrosa colección de sistemas terminales de Internet"). Actualmente, existen dos tipos de acceso inalámbrico a Internet. En una **LAN inalámbrica**, los usuarios inalámbricos

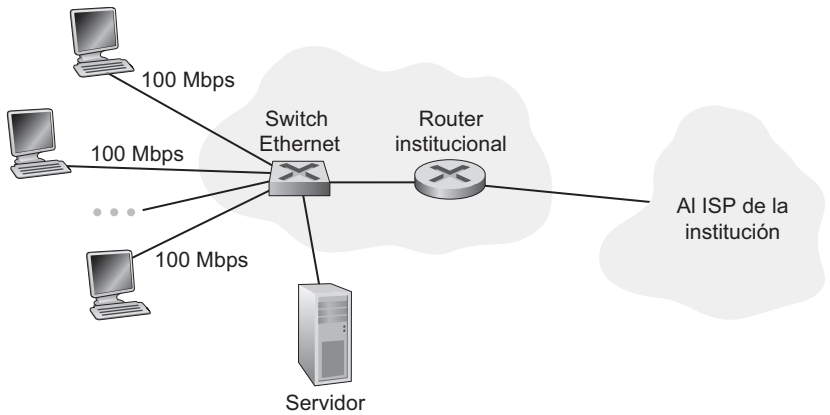


Figura 1.9 • Acceso a Internet utilizando tecnología Ethernet.

transmiten paquetes a (y reciben paquetes de) un **punto de acceso**, el cual a su vez está conectado a la red Internet cableada. Habitualmente, los usuarios de una LAN inalámbrica deben encontrarse a unas pocas decenas de metros del punto de acceso. En las **redes inalámbricas de área extensa**, los paquetes se transmiten a una **estación base** a través de la misma infraestructura inalámbrica utilizada por la telefonía móvil. En este caso, el proveedor de la red móvil gestiona la estación base y, normalmente, el usuario puede estar a unas pocas decenas de kilómetros de la estación base.

Actualmente, el acceso mediante LAN inalámbrica basada en la tecnología IEEE 802.11, es decir WiFi, podemos encontrarlo por todas partes: universidades, oficinas, cafés, aeropuertos, domicilios e incluso en los aviones. La mayor parte de las universidades han instalado estaciones base IEEE 802.11 por sus campus, lo que permite a los estudiantes enviar y recibir mensajes de correo electrónico o navegar por la Web estando en cualquier lugar del campus. En muchas ciudades, alguien puede estar parado en la esquina de una calle y encontrarse dentro del alcance de diez o veinte estaciones base (para ver un mapa global navegable de estaciones base 802.11 descubiertas y registradas en un sitio web por personas que disfrutan haciendo este tipo de cosas, consulte [wigle.net 2009]). Como se explica en el Capítulo 6, actualmente, la tecnología 802.11 proporciona una velocidad de transmisión compartida de hasta 54 Mbps.

Muchas viviendas combinan acceso residencial de banda ancha (es decir, modems por cable o DSL) con tecnología LAN inalámbrica barata para crear redes domésticas potentes. La Figura 1.10 muestra un esquema de una red doméstica típica. Esta red doméstica está formada por un portátil con función de itinerancia (*roaming*) y un PC de sobremesa; una estación base (el punto de acceso inalámbrico), que se comunica con el portátil inalámbrico; un módem por cable, que proporciona el acceso de banda ancha a Internet y un router, que interconecta la estación base y el PC de sobremesa con el módem por cable. Esta red permite a los usuarios de esta red doméstica tener acceso de banda ancha a Internet mediante un dispositivo móvil con el que se puede ir de la cocina a los dormitorios y al jardín.

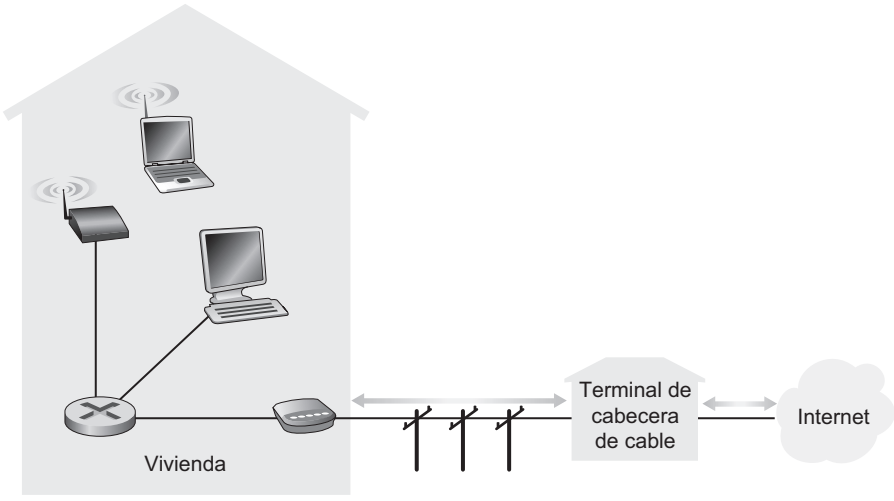


Figura 1.10 • Esquema de una red doméstica típica.

Acceso inalámbrico de área extensa

Cuando se accede a Internet a través de una red LAN inalámbrica, normalmente es necesario estar a unas pocas decenas de metros del punto de acceso. Esto es viable en viviendas, cafés y, de forma más general, en el interior y los alrededores de un edificio. Pero, ¿qué ocurre cuando se necesita tener acceso a Internet y se está en la playa, en un autobús o en el coche? Para este tipo de acceso de área extensa, los usuarios itinerantes de Internet utilizan la infraestructura de la telefonía móvil para acceder a estaciones base que están separadas entre sí unas decenas de kilómetros.

Las empresas de telecomunicaciones han hecho grandes inversiones en lo que se conoce como redes inalámbricas de tercera generación (3G), que proporcionan acceso inalámbrico a Internet mediante una red de área extensa de conmutación de paquetes a velocidades por encima de 1 Mbps. Actualmente, millones de usuarios emplean estas redes para leer y enviar mensajes de correo electrónico, navegar por la Web y descargar música mientras se desplazan de un lugar a otro.

WiMAX

Como siempre, existe una posible tecnología "definitiva" que espera destronar a estos estándares. WiMAX [Intel WiMAX 2009, WiMAX Forum 2009], también conocido como IEEE 802.16, es un primo lejano del protocolo WiFi 802.11 citado anteriormente. WiMAX opera independientemente de la red de telefonía móvil y promete velocidades comprendidas entre 5 y 10 Mbps o superiores para distancias de decenas de kilómetros. Sprint-Nextel ha invertido miles de millones de dólares en la implantación de WiMAX a partir del año 2007. En el Capítulo 6 se abordan en detalle las tecnologías WiFi, WiMAX y 3G.

1.2.3 Medios físicos

En la subsección anterior hemos proporcionado una panorámica de algunas de las tecnologías de acceso a red más importantes disponibles para Internet. Según hemos ido describiendo estas tecnologías, hemos indicado los medios físicos utilizados. Por ejemplo, hemos dicho que la tecnología HFC emplea una combinación de cable de fibra óptica y de cable coaxial. También hemos señalado que los modems de acceso telefónico a 56 kbps y las DSL utilizan cable de cobre de par trenzado. Asimismo, también hemos comentado que las redes para acceso móvil usan el espectro de radio. En esta subsección vamos a hacer una breve introducción a éstos y otros medios de transmisión que se emplean habitualmente en Internet.

Para definir lo que se entiende por medio físico, reflexionemos sobre la breve vida de un bit. Imagine un bit que viaja desde un sistema terminal atravesando una serie de enlaces y routers hasta otro sistema terminal. Este pobre bit se desplaza de un lado a otro sin descanso. En primer lugar, el sistema terminal de origen transmite el bit y poco tiempo después el primer router de la serie recibe dicho bit; el primer router transmite entonces el bit y poco después lo recibe el segundo router, y así sucesivamente. Por tanto, nuestro bit, al viajar desde el origen hasta el destino, atraviesa una serie de parejas de transmisores y receptores. En cada par transmisor-receptor, el bit se envía mediante ondas electromagnéticas o pulsos ópticos a lo largo de un **medio físico**. Este medio físico puede tener muchas formas y no tiene que ser del mismo tipo para cada par transmisor-receptor existente a lo largo de la ruta. Entre los ejemplos de medios físicos se incluyen el cable de cobre de par

trenzado, el cable coaxial, el cable de fibra óptica multimodo, el espectro de radio terrestre y el espectro de radio por satélite. Los medios físicos se pueden clasificar dentro de dos categorías: **medios guiados** y **medios no guiados**. En los medios guiados, las ondas se transportan a través de un medio sólido, como por ejemplo un cable de fibra óptica, un cable de cobre de par trenzado o un cable coaxial. En los medios no guiados, las ondas se propagan por la atmósfera y el espacio exterior, tal como ocurre en las redes LAN inalámbricas o en un canal de satélite digital.

Pero antes de abordar las características de los distintos tipos de medios, veamos algunos detalles acerca de los costes. El coste real de un enlace físico (cable de cobre, de fibra óptica, o coaxial, etc.) suele ser relativamente pequeño cuando se compara con los restantes costes de la red. En particular, el coste de mano de obra asociado con la instalación del enlace físico puede ser de varios órdenes de magnitud mayor que el coste del material. Por ello, muchos constructores instalan cables de par trenzado, de fibra óptica y coaxial en todas las habitaciones de los edificios. Incluso aunque inicialmente sólo se utilice uno de los medios, existen muchas posibilidades de que se emplee algún otro medio físico en un futuro próximo y, por tanto, se ahorre dinero al no tener que tirar cables adicionales.

Cable de cobre de par trenzado

El medio de transmisión guiado más barato y más comúnmente utilizado es el cable de cobre de par trenzado. Se ha utilizado durante un siglo en las redes telefónicas. De hecho, más del 99 por ciento de las conexiones cableadas utilizan cable de cobre de par trenzado entre el propio teléfono y el conmutador telefónico local. La mayoría de nosotros disponemos de cable de par trenzado en nuestros hogares y entornos de trabajo. Este cable consta de dos hilos de cobre aislados, de un milímetro de espesor cada uno de ellos, que siguen un patrón regular en espiral. Los hilos se trenzan para reducir las interferencias eléctricas procedentes de pares similares próximos. Normalmente, una serie de pares se meten dentro de un cable envolviendo los pares en una pantalla protectora. Un par de hilos constituyen un único enlace de comunicaciones. El **par trenzado no apantallado (UTP, *Unshielded Twisted Pair*)** se utiliza habitualmente en las redes de computadoras ubicadas dentro de un edificio, es decir, para las redes LAN. La velocidad de transmisión de datos de las LAN actuales que emplean cables de par trenzado varían entre 10 Mbps y 1 Gbps. Las velocidades de transmisión de datos que se pueden alcanzar dependen del espesor del cable y de la distancia existente entre el transmisor y el receptor.

Cuando en la década de 1980 surgió la tecnología de la fibra óptica, muchas personas despreciaron el cable de par trenzado a causa de sus relativamente bajas velocidades de transmisión. Algunos pensaron incluso que la fibra óptica desplazaría por completo al cable de par trenzado. Pero el cable de par trenzado no se daría por vencido tan fácilmente. La tecnología moderna del par trenzado, como por ejemplo los cables UTP de categoría 5, pueden alcanzar velocidades de datos de 1 Gbps para distancias de hasta 100 metros. Al final, los cables de par trenzado se han establecido como la solución dominante para las redes LAN de alta velocidad.

Como hemos mencionado anteriormente, los cables de par trenzado también suelen utilizarse para el acceso a Internet de tipo residencial. Hemos dicho que los modems de acceso telefónico permiten establecer conexiones a velocidades de hasta 56 kbps utilizando cables de par trenzado. También hemos comentado que la tecnología DSL (*Digital Subscriber Line*) ha permitido a los usuarios residenciales acceder a Internet a velocidades superiores a

6 Mbps empleando cables de par trenzado (siempre y cuando los usuarios vivan en las proximidades del módem del ISP).

Cable coaxial

Al igual que el par trenzado, el cable coaxial consta de dos conductores de cobre, pero dispuestos de forma concéntrica en lugar de en paralelo. Con esta construcción y un aislamiento y apantallamiento especiales, el cable coaxial puede proporcionar velocidades de transmisión de bit bastante altas. El cable coaxial es bastante común en los sistemas de televisión por cable. Como hemos mencionado anteriormente, recientemente los sistemas de televisión por cable han comenzado a incorporar modems por cable con el fin de proporcionar a los usuarios residenciales acceso a Internet a velocidades de 1 Mbps o superiores. En la televisión por cable y en el acceso a Internet por cable, el transmisor desplaza la señal digital a una banda de frecuencia específica y la señal analógica resultante se envía desde el transmisor a uno a o más receptores. El cable coaxial puede utilizarse como un **medio compartido** guiado; es decir, una serie de sistemas terminales pueden estar conectados directamente al cable, recibiendo todos ellos lo que envíen los otros sistemas terminales.

Fibra óptica

La fibra óptica es un medio flexible y de poco espesor que conduce pulsos de luz, representando cada pulso un bit. Un único cable de fibra óptica puede soportar velocidades de bit tremendamente altas, por encima de decenas o incluso centenas de gigabits por segundo. La fibra óptica es inmune a las interferencias electromagnéticas, presenta una atenuación de la señal muy baja hasta una distancia de 100 kilómetros y es muy difícil que alguien pueda llevar a cabo un "pinchazo" en una de estas líneas. Estas características hacen de la fibra óptica el medio de transmisión guiado a larga distancia preferido, especialmente para los enlaces transoceánicos. Muchas de las redes telefónicas para larga distancia de Estados Unidos y otros países utilizan hoy día exclusivamente fibra óptica. La fibra óptica también es el medio predominante en las redes troncales de Internet. Sin embargo, el alto coste de los dispositivos ópticos, como son los transmisores, receptores y conmutadores, están entorpeciendo su implantación para el transporte a corta distancia, como por ejemplo en el caso de una LAN o en el domicilio de una red de acceso residencial. Las velocidades del enlace estándar de portadora óptica (OC, *Optical Carrier*) están comprendidas en el rango de 51,8 Mbps a 39,8 Gbps; suele hacerse referencia a estas especificaciones como OC-*n*, donde la velocidad del enlace es igual a $n \times 51{,}8$ Mbps. Entre los estándares en uso actuales se encuentran: OC-1, OC-3, OC-12, OC-24, OC-48, OC-96, OC-192, OC-768. [IEC Optical 2009; Goralski 2001; Ramaswami 1998 y Mukherjee 1997] proporcionan información acerca de diversos aspectos de las redes ópticas.

Canales de radio terrestres

Los canales de radio transportan señales en el espectro electromagnético. Constituyen un medio atractivo porque no requieren la instalación de cables físicos, pueden atravesar las paredes, proporcionan conectividad a los usuarios móviles y, potencialmete, pueden transportar una señal a grandes distancias. Las características de un canal de radio dependen de forma significativa del entorno de propagación y de la distancia a la que la señal tenga que ser transportada. Las consideraciones ambientales determinan la pérdida del camino, la atenuación de sombra (lo que disminuye la intensidad de la señal a medida que recorre una dis-

tancia y rodea/atraviesa los objetos que obstruyen su camino), la atenuación multicamino (debida a la reflexión de la señal en los objetos que interfieren) y las interferencias (debidas a otras transmisiones y a las señales electromagnéticas).

Las canales de radio terrestre pueden clasificarse en dos amplios grupos: aquéllos que operan en las áreas locales, normalmente con un alcance de entre diez y unos cientos de metros y los que operan en un área extensa, con alcances de decenas de kilómetros. Las redes LAN inalámbricas descritas en la Sección 1.2.2 emplean canales de radio de área local y las tecnologías celulares utilizan canales de radio de área extensa. En el Capítulo 6 se estudian en detalle los canales de radio.

Canales de radio vía satélite

Las comunicaciones por satélite enlazan dos o más transmisores/receptores de microondas con base en la Tierra, que se conocen como estaciones terrestres. El satélite recibe las transmisiones en una banda de frecuencia, regenera la señal utilizando un repetidor (véase más adelante) y transmite la señal a otra frecuencia. En este tipo de comunicaciones se emplean dos tipos de satélites: los **satélites geoestacionarios** y los **satélites de la órbita baja terrestre (LEO, *Low-Earth Orbiting*)**.

Los satélites geoestacionarios están permanentemente situados en el mismo punto por encima de la Tierra. Esta presencia estacionaria se consigue poniendo el satélite en órbita a una distancia de 36.000 kilómetros por encima de la superficie de la Tierra. La distancia entre la estación terrestre y el satélite más la distancia de vuelta desde el satélite a la estación terrestre introduce un retardo de propagación de la señal de 280 milisegundos. No obstante, los enlaces vía satélite, que pueden operar a velocidades de cientos de Mbps, a menudo se emplean en áreas en las que no hay disponible acceso a Internet mediante DSL o cable.

Los satélites LEO se colocan mucho más cerca de la Tierra y no se encuentran permanentemente en la misma posición, sino que giran alrededor de la Tierra (al igual que la Luna) y pueden comunicarse entre sí, así como con las estaciones terrestres. Para poder proporcionar una cobertura continua a un área, es preciso poner en órbita muchos satélites. Actualmente se están desarrollando muchos sistemas de comunicaciones de baja altitud. La página web Lloyd's satellite constellations [Wood 2009] proporciona y recopila información acerca de los sistemas de constelaciones de satélites para comunicaciones. La tecnología de los satélites de la órbita baja terrestre (LEO) podrá utilizarse, en algún momento en el futuro, para acceder a Internet.

1.3 El núcleo de la red

Una vez que hemos examinado la frontera de Internet, vamos a adentrarnos en el núcleo de la red, la malla de conmutadores de paquetes y enlaces que interconectan los sistemas terminales de Internet. En la Figura 1.11 se ha resaltado el núcleo de la red con líneas más gruesas.

1.3.1 Conmutación de circuitos y conmutación de paquetes

Existen dos métodos fundamentales que permiten transportar los datos a través de una red de enlaces y conmutadores: la **conmutación de circuitos** y la **conmutación de paquetes**. En las redes de conmutación de circuitos, los recursos necesarios a lo largo de una ruta

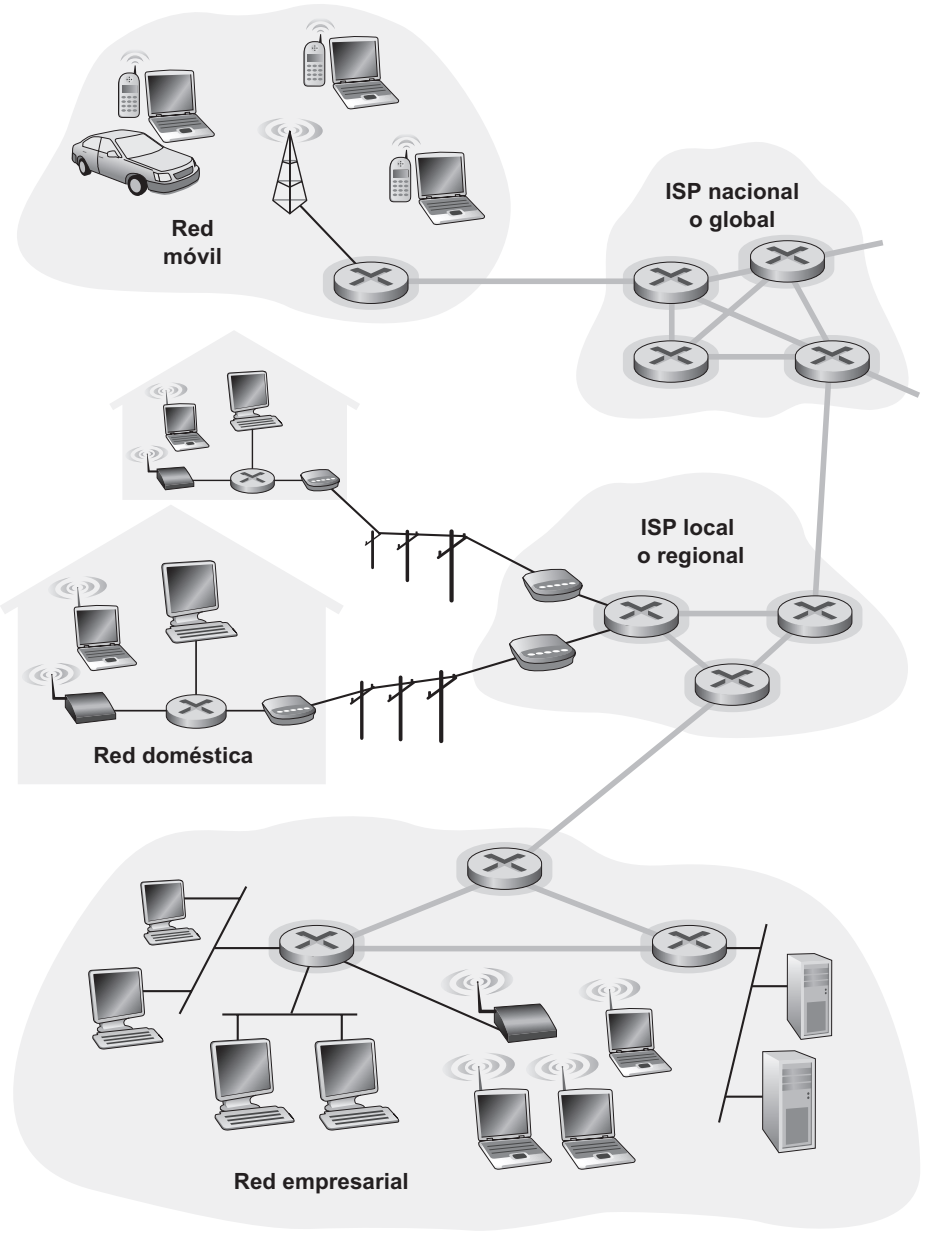


Figura 1.11 • El núcleo de la red.

(buffers, velocidad de transmisión del enlace) que permiten establecer la comunicación entre los sistemas terminales están *reservados* durante el tiempo que dura la sesión entre dichos sistemas terminales. En las redes de conmutación de paquetes, estos recursos *no* están reservados; los mensajes de una sesión utilizan los recursos bajo petición y, en con-

secuencia, pueden tener que esperar (es decir, ponerse en cola) para poder acceder a un enlace de comunicaciones. Veamos una sencilla analogía. Piense en dos restaurantes, en uno de ellos es necesario hacer reserva y en el otro no se requiere hacer reserva ni tampoco las admiten. Para comer en el restaurante que precisa reserva, tenemos que molestarnos en llamar por teléfono antes de salir de casa, pero al llegar allí, en principio, podremos sentarnos y pedir nuestro menú al camarero de manera inmediata. En el restaurante que no admite reservas, no tenemos que molestarnos en reservar mesa, pero al llegar allí, es posible que tengamos que esperar para tener una mesa antes de poder hablar con el camarero.

Las omnipresentes redes telefónicas son ejemplos de redes de conmutación de circuitos. Considere lo que ocurre cuando una persona desea enviar información (de voz o fax-símil) a otra a través de una red telefónica. Antes de que el emisor pueda transmitir la información, la red debe establecer una conexión entre el emisor y el receptor. Se trata de una conexión de *buena fe* en la que los conmutadores existentes en la ruta entre el emisor y el receptor mantienen el estado de la conexión para dicha comunicación. En la jerga del campo de la telefonía, esta conexión se denomina **circuito**. Cuando la red establece el circuito, también reserva una velocidad de transmisión constante en los enlaces de la red para el tiempo que dure la conexión. Dado que el ancho de banda para esta conexión emisor-receptor ha sido reservado, el emisor puede transferir los datos al receptor a la velocidad constante *garantizada*.

La red Internet de hoy día es la quinta esencia de las redes de conmutación de paquetes. Veamos qué ocurre cuando un host desea enviar un paquete a otro host a través de Internet. Al igual que con la conmutación de circuitos, el paquete se transmite a través de una serie de enlaces de comunicaciones. Pero con la técnica de conmutación de paquetes, el paquete se envía a la red sin haber reservado ancho de banda. Si uno de los enlaces está congestionado porque otros paquetes tienen que ser transmitidos a través de él al mismo tiempo, entonces nuestro paquete tendrá que esperar en un buffer en el lado del emisor del enlace de transmisión y, por tanto, sufrirá un retardo. Internet realiza el *máximo esfuerzo* para suministrar los paquetes a tiempo, pero no existe ninguna garantía.

No todas las redes de telecomunicaciones pueden clasificarse como redes puras de conmutación de circuitos o redes puras de conmutación de paquetes. No obstante, esta clasificación es un excelente punto de partida para comprender la tecnología de las redes de telecomunicaciones.

Conmutación de circuitos

Este libro está dedicado a las redes de computadoras, Internet y la conmutación de paquetes, no a las redes telefónicas y la conmutación de circuitos. No obstante, es importante comprender por qué Internet y otras redes de computadoras utilizan la tecnología de conmutación de paquetes en lugar de la tecnología más tradicional de conmutación de circuitos de las redes telefónicas. Por esta razón, a continuación vamos a ofrecer una breve introducción a la conmutación de circuitos.

La Figura 1.12 ilustra una red de conmutación de circuitos. En esta red, los cuatro conmutadores de circuitos están interconectados mediante cuatro enlaces. Cada uno de los enlaces tiene n circuitos, por lo que cada enlace puede dar soporte a n conexiones simultáneas. Cada uno de los hosts (por ejemplo, los PC y estaciones de trabajo) está conectado directamente a uno de los conmutadores. Cuando dos hosts desean comunicarse, la red establece una **conexión terminal a terminal** dedicada entre ellos (por supuesto, las llamadas entre

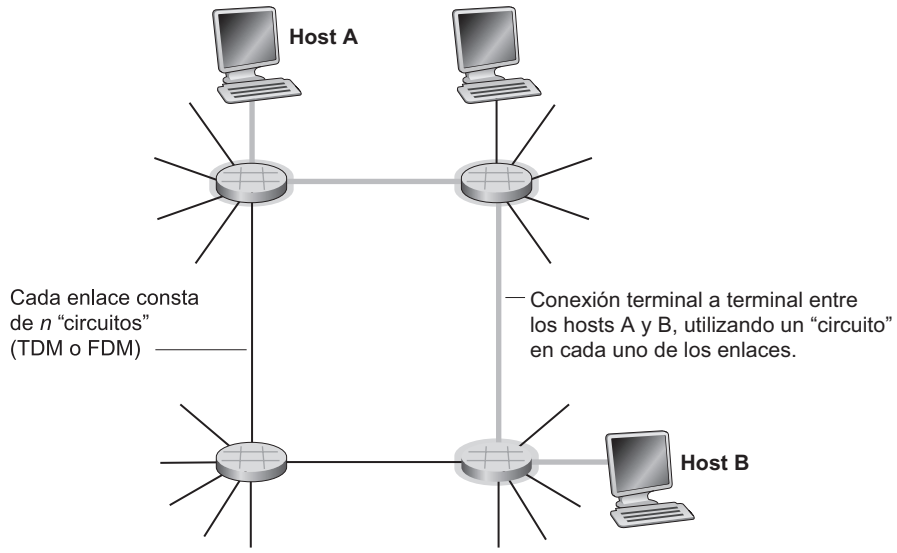


Figura 1.12 • Red de conmutación de circuitos simple formada por cuatro dispositivos de conmutación y cuatro enlaces.

más de dos dispositivos también son posibles, pero con el fin de que el lector comprenda el concepto, vamos a suponer por el momento que sólo intervienen dos hosts en cada conexión). Por tanto, para que el host A envíe mensajes al host B, la red tiene que reservar en primer lugar un circuito en cada uno de los dos enlaces. Dado que cada enlace tiene *n* circuitos, para cada enlace utilizado por la conexión terminal a terminal, la conexión obtiene una fracción 1*/n* del ancho de banda del enlace para el tiempo de duración de la conexión.

Multiplexación en redes de conmutación de circuitos

Un circuito en un enlace se implementa bien mediante **multiplexación por división de frecuencia (FDM, *Frequency-Division Multiplexing*)** o mediante **multiplexación por división en el tiempo (TDM, *Time-Division Multiplexing*)**. Con FDM, el espectro de frecuencia de un enlace se reparte entre las conexiones establecidas a lo largo del enlace. Específicamente, el enlace dedica una banda de frecuencias a cada conexión durante el tiempo que ésta dure. En las redes telefónicas, esta banda de frecuencias normalmente tiene un ancho de 4 kHz (es decir, 4.000 hercios o 4.000 ciclos por segundo). El ancho de esta banda se denomina lógicamente **ancho de banda**. Las estaciones de radio FM también emplean la multiplexación FDM para compartir el espectro de frecuencias entre 88 MHz y 108 MHz, teniendo cada estación asignada una banda de frecuencias específica.

En un enlace TDM, el tiempo se divide en marcos de duración fija y cada marco se divide en un número fijo de particiones. Cuando la red establece una conexión a través de un enlace, la red dedica una partición de cada marco a dicha conexión. Estas particiones están dedicadas para uso exclusivo de dicha conexión con una partición disponible para utilizar (en cada marco) para transmitir los datos de la conexión.

La Figura 1.13 ilustra las multiplexaciones FDM y TDM para un enlace de red específico que da soporte a cuatro circuitos. En el caso de la multiplexación por división de

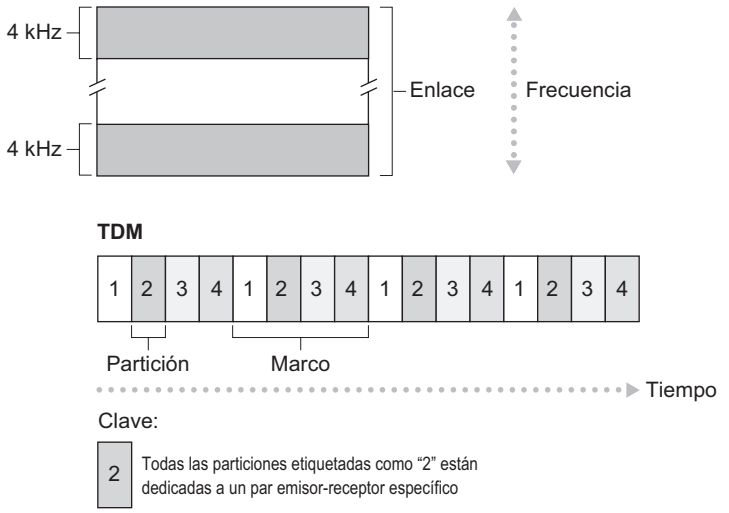


Figura 1.13 • Con FDM cada circuito obtiene de forma continua una fracción del ancho de banda. Con TDM, cada circuito dispone de todo el ancho de banda periódicamente durante breves intervalos de tiempo (es decir, durante las particiones).

frecuencia, el dominio de frecuencia se segmenta en cuatro bandas, siendo el ancho de banda de cada una de ellas de 4 kHz. En el caso de la multiplexación TDM, el dominio del tiempo se divide en cuatro marcos, conteniendo cada uno de ellos cuatro particiones. A cada circuito se le asigna la misma partición dedicada dentro de los marcos, de forma cíclica. En la multiplexación TDM, la velocidad de transmisión de un circuito es igual a la velocidad de marco multiplicada por el número de bits existentes en una partición. Por ejemplo, si el enlace transmite 8.000 marcos por segundo y cada partición consta de 8 bits, entonces la velocidad de transmisión de un circuito es igual a 64 kbps.

Los partidarios de la tecnología de conmutación de paquetes siempre han argumentado que la conmutación de circuitos es derrochadora, porque los circuitos dedicados quedan inutilizados durantes los **periodos de inactividad**. Por ejemplo, cuando una persona deja de hablar durante una llamada telefónica, los recursos de red inactivos (bandas de frecuencia o particiones temporales del enlace a lo largo de la ruta de la conexión) no pueden ser empleados por otras conexiones en curso. Otro ejemplo de cómo estos recursos pueden ser infrautilizados sería un radiólogo que empleara una red de conmutación de circuitos para acceder remotamente a una serie de radiografías de rayos X. El radiólogo establece una conexión, solicita una imagen, la contempla y luego solicita otra imagen. Los recursos de la red están asignados a la conexión pero no se utilizan (es decir, se desperdician) durante el tiempo que el radiólogo contempla las imágenes. Los partidarios de la conmutación de paquetes también disfrutan apuntando que el establecimiento de circuitos terminal a terminal y la reserva de ancho de banda terminal a terminal son procesos complicados que requieren el uso de software de señalización complejo para coordinar el funcionamiento de los switches a lo largo de la ruta terminal a terminal.

Antes de terminar con esta exposición acerca de la conmutación de circuitos, vamos a ver un ejemplo numérico que debería arrojar más luz sobre este tema. Consideremos el

tiempo que se tarda en enviar un archivo de 640.000 bits desde el host A al host B a través de una red de conmutación de circuitos. Supongamos que todos los enlaces de la red utilizan multiplexación TDM con 24 particiones y tienen una velocidad de bit de 1,536 Mbps. Supongamos también que se tardan 500 milisegundos en establecer el circuito terminal a terminal antes de que el host A pueda comenzar a transmitir el archivo. ¿Cuánto tiempo se tarda en transmitir el archivo? La velocidad de transmisión de cada circuito es (1,536 Mbps)/24 = 64 kbps, por lo que se precisan (640.000 bits)/(64 kbps) = 10 segundos en transmitir el archivo. A estos 10 segundos tenemos que sumarles el tiempo de establecimiento del circuito, lo que da como resultado 10,5 segundos de tiempo total de transmisión del archivo. Observe que el tiempo de transmisión es independiente del número de enlaces: el tiempo de transmisión será 10 segundos independientemente de que el circuito terminal a terminal pase a través de un enlace o de cien enlaces. (El retardo real terminal a terminal también incluye un retardo de propagación; véase la Sección 1.4.)

Conmutación de paquetes

Las aplicaciones distribuidas intercambian **mensajes** para llevar a cabo sus tareas. Los mensajes pueden contener cualquier cosa que el diseñador del protocolo desee. Los mensajes pueden realizar una función de control (por ejemplo, los mensajes de saludo "Hola" del ejemplo anterior sobre establecimiento de la comunicación) o pueden contener datos, como por ejemplo un mensaje de correo electrónico, una imagen JPEG o un archivo de audio MP3. En las redes de computadoras modernas, el origen divide los mensajes largos en fragmentos de datos más pequeños que se conocen como **paquetes**. Entre el origen y el destino, cada uno de estos paquetes viaja a través de los enlaces de comunicaciones y de los **conmutadores de paquetes** (de los que existen dos tipos predominantes: los routers y los switches de la capa de enlace). Los paquetes se transmiten a través de cada enlace de comunicaciones a una velocidad igual a la velocidad de transmisión *máxima* del enlace.

La mayoría de los conmutadores de paquetes emplean el método de **transmisión de almacenamiento y reenvío** en las entradas de los enlaces. Transmisión de almacenamiento y reenvío significa que el conmutador tiene que recibir el paquete completo antes de poder comenzar a transmitir el primer bit del paquete al enlace de salida. Por tanto, los conmutadores de paquetes de almacenamiento y reenvío añaden un retardo de almacenamiento y reenvío en la entrada de cada enlace existente a lo largo de la ruta que debe seguir el paquete. Veamos el tiempo que se tarda en enviar un paquete de L bits desde un host a otro host en una red de conmutación de paquetes. Supongamos que existen Q enlaces entre los dos hosts, y que la velocidad en cada uno de ellos es igual a R bps. Suponemos que éste es el único paquete presente en la red. En primer lugar, el paquete tiene que enviarse a través del primer enlace que sale del host A, lo que consume un tiempo de L/R segundos. A continuación, tiene que ser transmitido por cada uno de los $Q - 1$ enlaces restantes; es decir, se tiene que almacenar y reenviar $Q - 1$ veces, añadiéndose cada vez un retardo de almacenamiento y reenvío de L/R. Por tanto, el retardo total es igual a QL/R.

Cada conmutador de paquetes tiene varios enlaces conectados a él y para cada enlace conectado, el conmutador de paquetes dispone de un **buffer de salida** (también denominado **cola de salida**), que almacena los paquetes que el router enviará a través de dicho enlace. El buffer de salida desempeña un papel clave en la conmutación de paquetes. Si un paquete entrante tiene que ser transmitido a través de un enlace, pero se encuentra con que el enlace está ocupado transmitiendo otro paquete, el paquete entrante tendrá que esperar en el buffer de salida. Por tanto, además de los retardos de almacenamiento y reenvío, los paquetes se

ven afectados por los **retardos de cola** del buffer de salida. Estos retardos son variables y dependen del nivel de congestión de la red. Puesto que la cantidad de espacio en el buffer es finita, un paquete entrante puede encontrarse con que el buffer está completamente lleno con otros paquetes que esperan a ser transmitidos. En este caso, se producirá una **pérdida de paquetes**, bien el paquete que acaba de llegar o uno que ya se encuentra en la cola será descartado. Si volvemos a la analogía de los restaurantes vista anteriormente en esta sección, el retardo de cola es análogo a la cantidad de tiempo que usted pierde en el bar del restaurante esperando a que una mesa se quede libre. La pérdida de paquetes es análoga a que el camarero le comunique que es mejor que vaya a otro restaurante porque ya hay demasiadas personas esperando en el bar para conseguir una mesa.

La Figura 1.14 ilustra una red de conmutación de paquetes simple. En esta figura y en las siguientes, los paquetes se han representado mediante bloques tridimensionales. El ancho de un bloque representa el número de bits que contiene el paquete. En esta figura, todos los paquetes tienen el mismo ancho y, por tanto, la misma longitud. Suponga ahora que los hosts A y B están enviando paquetes al host E. En primer lugar, los hosts A y B envían sus paquetes a través de los enlaces Ethernet a 10 Mbps hasta el primer conmutador de paquetes. A continuación, éste dirige los paquetes al enlace de 1,5 Mbps. Si la velocidad de llegada de los paquetes al conmutador excede la velocidad a la que el conmutador puede reenviar los paquetes a través del enlace de salida de 1,5 Mbps, se producirá congestión a medida que los paquetes se pongan en cola en el buffer de salida del enlace antes de poder ser transmitidos. En la Sección 1.4 examinaremos más detalladamente el retardo de cola.

Conmutación de paquetes frente a conmutación de circuitos: multiplexación estadística

Ahora que ya hemos descrito las tecnologías de conmutación de circuitos y de paquetes, vamos a pasar a compararlas. Los detractores de la tecnología de conmutación de paquetes a menudo han argumentado que esta tecnología no es adecuada para los servicios en tiempo real, como por ejemplo las llamadas telefónicas y las videoconferencias, porque sus retardos

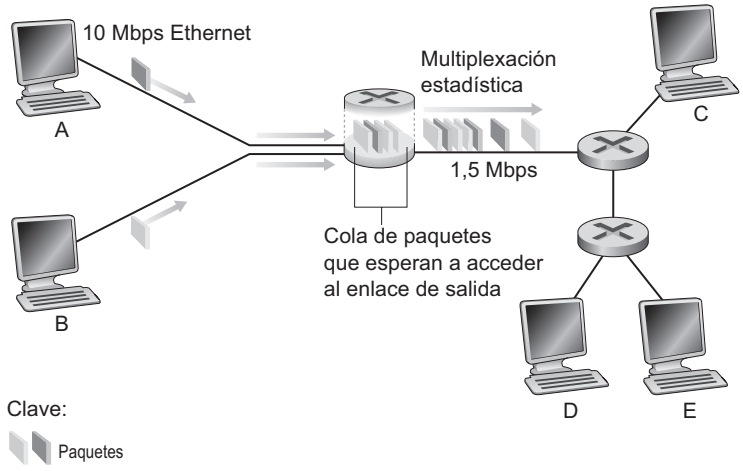


Figura 1.14 • Conmutación de paquetes.

terminal a terminal son variables e impredecibles (a causa principalmente de que los retardos de cola de los paquetes son variables e impredecibles). Por otro lado, los partidarios de la conmutación de paquetes argumentan que (1) ofrece una mejor compartición del ancho de banda que la tecnología de conmutación de circuitos y (2) es más sencilla, más eficiente y menos cara de implementar que la conmutación de circuitos. Puede encontrar una exposición interesante acerca de la conmutación de paquetes frente a la conmutación de circuitos en [Molinero-Fernández 2002]. Generalmente, las personas que no se molestan en reservar mesa en un restaurante prefieren la conmutación de paquetes a la conmutación de circuitos.

¿Por qué es más eficiente la conmutación de paquetes? Veamos un ejemplo sencillo. Suponga que varios usuarios comparten un enlace de 1 Mbps. Suponga también que cada usuario alterna entre periodos de actividad (cuando genera datos a una velocidad constante de 100 kbps) y periodos de inactividad (cuando no genera datos). Además, suponga que un usuario sólo está activo un 10 por ciento del tiempo (y está inactivo tomando café durante el 90 por ciento del tiempo restante). Con la tecnología de conmutación de circuitos, tienen que *reservarse* 100 kbps para *cada* usuario todas las veces. Por ejemplo, en una red de conmutación de circuitos con multiplexación TDM, si un marco de un segundo se divide en 10 particiones de 100 ms, entonces cada usuario tendría asignada una partición por marco.

Por tanto, el enlace de conmutación de circuitos sólo podrá dar soporte a 10 (= 1 Mbps/100 kbps) usuarios simultáneamente. En el caso de utilizar la conmutación de paquetes, la probabilidad de que un determinado usuario esté activo es 0,1 (es decir, del 10 por ciento). Si hay 35 usuarios, la probabilidad de que 11 o más usuarios estén activos simultáneamente es aproximadamente igual a 0,0004. (El Problema P7 indica cómo se obtiene esta probabilidad.) Cuando hay 10 o menos usuarios activos a la vez (lo que ocurre con una probabilidad del 0,9996), la velocidad acumulada de llegada de los datos es menor o igual a 1 Mbps, la velocidad de salida del enlace. Por tanto, cuando el número de usuarios activos es 10 o menor, los paquetes fluyen a través del enlace prácticamente sin retardo, como en el caso de la tecnología de conmutación de circuitos. Cuando hay más de 10 usuarios activos simultáneamente, entonces la velocidad acumulada de llegada de los paquetes excede la capacidad de salida del enlace y la cola de salida comenzará a crecer. (Continúa creciendo hasta que la velocidad acumulada de entrada cae por debajo de 1 Mbps, punto en el que la longitud de la cola comenzará a disminuir.) Puesto que la probabilidad de que haya más de 10 usuarios conectados a la vez es muy baja en este ejemplo, la conmutación de paquetes proporciona prácticamente el mismo rendimiento que la conmutación de circuitos, *pero lo hace permitiendo que haya un número de usuarios más de tres veces superior.*

Consideremos ahora otro ejemplo sencillo. Suponga que hay 10 usuarios y que de repente un usuario genera 1.000 paquetes de 1.000 bits, mientras que los usuarios restantes permanecen inactivos y no generan paquetes. Con la tecnología de conmutación de circuitos con multiplexación TDM con 10 particiones por marco y con cada partición formada por 1.000 bits, el usuario activo sólo puede emplear su partición por marco para transmitir los datos, mientras que las restantes nueve particiones de cada marco permanecen inactivas. Transcurrirán 10 segundos antes de que el millón de bits de datos del usuario activo hayan sido transmitidos. Sin embargo, con la conmutación de paquetes, el usuario activo puede enviar de forma continuada sus paquetes a la velocidad máxima del enlace de 1 Mbps, ya que no hay ningún otro usuario generando paquetes que tengan que ser multiplexados con los paquetes del usuario activo. En este caso, todos los datos del usuario activo se transmitirán en un segundo.

Los ejemplos anteriores han servido para ilustrar dos casos en los que el rendimiento de la tecnología de conmutación de paquetes puede resultar superior a la de la conmutación de circuitos. También ha quedado patente la diferencia crucial entre las dos formas de compartir la velocidad de transmisión del enlace entre varios flujos de datos. La conmutación de circuitos preasigna el uso del enlace de transmisión independientemente de la demanda, con lo que el tiempo de enlace asignado, pero innecesario, se desperdicia. Por el contrario, la conmutación de paquetes asigna el uso del enlace *bajo demanda.* La capacidad de transmisión del enlace se compartirá paquete a paquete sólo entre aquellos usuarios que tienen paquetes que transmitir a través del enlace. La compartición de recursos bajo petición (en lugar de por preasignación) a veces se denomina multiplexación **estadística** de recursos.

Aunque hoy en día las redes de telecomunicaciones predominantes son las de conmutación de circuitos y de paquetes, realmente se está tendiendo hacia las redes de conmutación de paquetes. Incluso muchas de las redes de telefonía de conmutación de circuitos actuales se están migrando lentamente a redes de conmutación de paquetes. En particular, las redes telefónicas suelen emplear la conmutación de paquetes en la parte internacional, que es la más cara de una llamada telefónica.

1.3.2 ¿Cómo atraviesan los paquetes las redes de conmutación de paquetes?

Anteriormente hemos dicho que un router toma un paquete entrante en uno de sus enlaces de comunicaciones, pero ¿cómo el router determina el enlace por el que deberá reenviar el paquete? En realidad, los diferentes tipos de redes pueden hacer esto de diversas formas. En este capítulo de introducción vamos a describir un método popular, el método empleado por Internet.

En Internet, cada paquete que atraviesa la red contiene en su cabecera la dirección del destino del paquete. Al igual que las direcciones postales, esta dirección tiene una estructura jerárquica. Cuando llega un paquete a un router de la red, el router examina una parte de la dirección de destino del paquete y lo reenvía a un router adyacente. Más específicamente, cada router dispone de una **tabla de reenvío** que asigna las direcciones de destino (o una parte de las mismas) a los enlaces salientes. Cuando llega un paquete a un router, éste examina la dirección y busca en su tabla esa dirección de destino para localizar el enlace de salida apropiado. A continuación, el router dirige el paquete a ese enlace de salida.

Acabamos de ver que un router utiliza la dirección de destino de un paquete para indexar una tabla de reenvío y determinar el enlace de salida apropiado. Pero esta afirmación nos lleva a la siguiente pregunta: ¿cómo se definen las tablas de reenvío? ¿Se configuran manualmente en cada router o Internet utiliza un procedimiento más automatizado? Estas cuestiones se abordan en detalle en el Capítulo 4, pero para ir abriendo boca, diremos que Internet dispone de una serie de protocolos de enrutamiento especiales que se utilizan para definir automáticamente las tablas de reenvío. Por ejemplo, un protocolo de enrutamiento determina la ruta más corta desde cada router hasta cada destino y el uso de la ruta más corta da como resultado la configuración de las tablas de reenvío en los routers.

El proceso de enrutamiento de terminal a terminal es análogo al que sigue el conductor de un automóvil que no utiliza un mapa, sino que prefiere preguntar cómo llegar hasta una determinada dirección. Por ejemplo, suponga que Juan sale de Filadelfia y tiene que llegar al 156 de Lakeside Drive en Orlando, Florida. Lo primero que hace Juan es dirigirse a la estación de servicio más próxima y preguntar cómo llegar a su destino. El empleado se

queda con el nombre del estado Florida, y le dice que debe tomar la autopista interestatal I-95 Sur y que existe una entrada a la misma nada más salir de la estación de servicio. También le dice a Juan que una vez que haya entrado en Florida, pregunte a alguien cómo llegar a su destino. Así, Juan toma la I-95 Sur hasta Jacksonville, Florida, lugar donde vuelve a preguntar también en una estación de servicio. El dependiente extrae de la dirección la información que hace referencia a Orlando y le dice que debe continuar por la I-95 hasta Daytona Beach y que luego pregunte. En otra estación de servicio de Daytona Beach, el empleado de nuevo extrae la información referente a Orlando y le dice que tomando la I-4 llegará directamente a Orlando. Juan toma la I-4 y la abandona en la salida que indica a Orlando. De nuevo se detiene en otra gasolinera y esta vez el dependiente extrae la parte de la información de la dirección referente a Lakeside Drive y le indica la carretera que debe seguir para llegar allí. Una vez que Juan se encuentra en Lakeside Drive, pregunta a un niño que va en bicicleta cómo llegar a su destino. El niño extrae el dato 156 de la dirección y le señala una casa. Por fin, Juan ha llegado a su destino.

En esta analogía, los dependientes de las estaciones de servicio y el niño de la bicicleta son los routers, y sus tablas de reenvío, que son sus cerebros, se han ido configurando a lo largo de años de experiencia.

¿Cómo podríamos ver la ruta terminal a terminal que siguen los paquetes en Internet? Le invitamos a que utilice el programa Traceroute, visitando el sitio http://www.traceroute.org. (Para obtener más información acerca de Traceroute, consulte la Sección 1.4.)

1.3.3 Redes troncales de Internet y proveedores ISP

Anteriormente hemos visto que los sistemas terminales (los PC de usuario, las PDA, los servidores web, los servidores de correo electrónico, etc.) se conectan a Internet a través de un ISP local. El ISP puede proporcionar conectividad cableada o inalámbrica, mediante una amplia variedad de tecnologías de acceso, entre las que se incluyen DSL, cable, FTTH, Wi-Fi, celular y WiMAX. Observe que el ISP local no tiene que ser una compañía telefónica ni una compañía de cable: puede ser, por ejemplo, una universidad (que proporciona acceso a Internet a los estudiantes, al personal y a las facultades) o una empresa (que proporciona acceso a sus empleados). Pero la conexión de los usuarios finales y de los proveedores de contenido a los ISP locales es sólo una pequeña parte del problema de conectar los cientos de millones de sistemas terminales y los cientos de miles de redes que conforman Internet. Internet es una *red de redes* y entender esta afirmación es fundamental para resolver este puzzle.

En la red pública Internet, los ISP de acceso situados en la frontera de Internet están conectados al resto de Internet a través de una jerarquía de niveles de proveedores ISP, como se muestra en la Figura 1.15. Los ISP de acceso se sitúan en el nivel inferior de esta jerarquía. En el nivel superior de la jerarquía se encuentran en un número relativamente más pequeño los **ISP de nivel 1**. Por un lado, un ISP de nivel 1 es igual que cualquier red (está formado por enlaces y routers, y está conectado a otras redes). Sin embargo, estos ISP presentan otras características que los hace especiales. La velocidad de enlace suele ser de 622 Mbps o superior, por lo que los ISP de nivel 1 de mayor tamaño disponen de enlaces en el rango comprendido entre 2,5 y 10 Gbps; en consecuencia, sus routers deben poder reenviar los paquetes a velocidades extremadamente altas. Los ISP de nivel 1 también se caracterizan por lo siguiente:

- Están conectados directamente a *cada uno* de los restantes ISP de nivel 1.

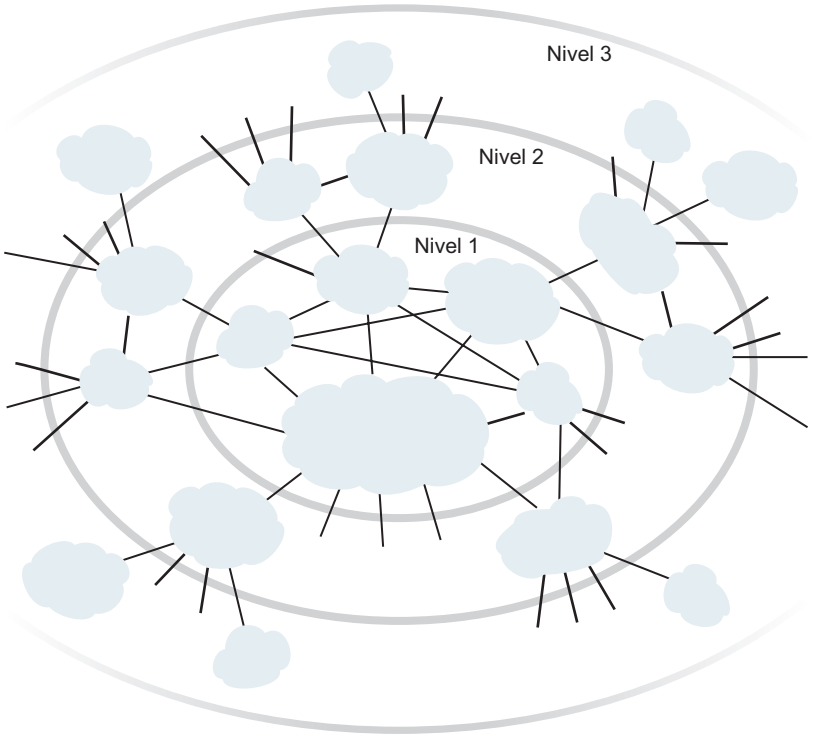


Figura 1.15 • Interconexión de los ISP.

- Están conectados a un gran número de ISP de nivel 2 y a otras redes cliente.
- Proporcionan cobertura internacional.

Los ISP de nivel 1 también se conocen como redes **troncales de Internet**. Entre estos ISP se incluyen Sprint, Verizon, MCI (anteriormente UUNet/WorldCom), AT&T, NTT, Level3, Qwest y Cable & Wireless. Resulta curioso que, oficialmente, no existe ningún grupo que conceda el estatus de nivel 1; como reza el dicho: si tienes que preguntar si eres miembro de un grupo, es que probablemente no lo eres.

Por regla general, un ISP de nivel 2 tiene cobertura regional o nacional y lo que es más importante, sólo está conectado a unos pocos ISP de nivel 1 (véase la Figura 1.15). Por tanto, para llegar a gran parte de la red Internet global, un ISP de nivel 2 tiene que enrutar el tráfico a través de uno de los ISP de nivel 1 a los que está conectado. Se dice que un ISP de nivel 2 es un **cliente** del ISP de nivel 1 al que está conectado y que el ISP de nivel 1 es un **proveedor** de dicho cliente. Muchas instituciones y empresas de gran tamaño conectan sus redes empresariales directamente a un ISP de nivel 1 o de nivel 2, convirtiéndose así en un cliente de dicho ISP. Un ISP proveedor cobra unas determinadas tasas al ISP cliente, que normalmente dependen de la velocidad de transmisión del enlace que los conecta. Una red de nivel 2 también se puede conectar directamente a otras redes de nivel 2, en cuyo caso el tráfico puede fluir entre las dos redes de nivel 2 sin tener que pasar a través de una red de nivel 1. Por debajo de los ISP de nivel 2 se encuentran los ISP de nivel inferior, que se conectan a Internet a través de uno o más ISP de nivel 2. En el nivel más bajo de la jerarquía

se encuentran los ISP de acceso. Para complicar aún más las cosas, algunos proveedores de nivel 1 también son proveedores de nivel 2 (es decir, están integrados verticalmente), que venden directamente acceso a Internet a los usuarios finales y proveedores de contenido, así como a los ISP del nivel inferior. Cuando dos ISP están conectados directamente entre sí en el mismo nivel, se dice que son **igualitarios**. Existe un interesante estudio [Subramanian 2002] que intenta definir la estructura en niveles de Internet de forma más precisa estudiando la topología de Internet en función de las relaciones cliente-proveedor y las relaciones entre iguales. Consulte [Van der Berg 2008] para ver una explicación bastante comprensible acerca de las relaciones entre iguales y cliente-proveedor.

Dentro de la red de un ISP, los puntos en los que el ISP se conecta a otros ISP (sean de nivel inferior, superior o del mismo nivel dentro de la jerarquía) se conocen como **Puntos de presencia (POP, *Point of Presence*)**. Un POP es simplemente un grupo de uno o más routers de la red del ISP en los que los routers de otros ISP o de las redes que pertenecen a los clientes del ISP pueden conectarse. Un proveedor de nivel 1 normalmente tiene muchos POP dispersos por distintas localizaciones geográficas dentro de la red, con múltiples redes y otros ISP conectados a cada POP. Normalmente, cuando una red cliente tiene que conectarse al POP de un proveedor, alquila un enlace de alta velocidad de un proveedor de telecomunicaciones de una tercera empresa y conecta directamente uno de sus routers a un router ubicado en el POP del proveedor. Además, dos ISP pueden disponer de varios puntos de conexión entre iguales, conectándose entre sí en múltiples pares de POP.

En resumen, la topología de Internet es compleja y está formada por docenas de ISP de nivel 1 y de nivel 2 y por miles de ISP de nivel inferior. La cobertura de los ISP puede ser muy variada, pudiéndose extender a varios continentes y océanos hasta estar limitada a pequeñas regiones del mundo. Los ISP del nivel inferior se conectan a los ISP de los niveles superiores y éstos a su vez se interconectan entre sí. Los usuarios y los proveedores de contenido son clientes de los ISP de nivel inferior y éstos son clientes de los ISP de nivel superior.

1.4 Retardos, pérdidas y tasa de transferencia en las redes de conmutación de paquetes

En la Sección 1.1 hemos dicho que Internet puede verse como una infraestructura que proporciona servicios a aplicaciones distribuidas que se ejecutan en sistemas terminales. Idealmente, desearíamos que los servicios de Internet pudieran transportar tantos datos como quisiéramos entre cualesquiera dos sistemas terminales de forma instantánea y sin que tuviera lugar ninguna pérdida de datos. Evidentemente, en la realidad, este objetivo es inalcanzable, ya que necesariamente las redes de computadoras tienen que restringir su tasa de transferencia (la cantidad de datos por segundo que pueden transmitir) entre sistemas terminales, introducir retardos entre los sistemas terminales y perder paquetes. Por una parte, es lamentable que las leyes físicas introduzcan retardos y pérdidas, así como que restrinjan las tasas de transferencia, pero, por otra parte, puesto que las redes de computadoras presentan estos problemas, existen muchas cuestiones interesantes relacionadas con cómo abordarlos, ¡más que suficientes como para llenar un curso sobre redes de computadoras y para motivar cientos de tesis doctorales! En esta sección comenzaremos examinando y cuantificando los retardos, las pérdidas y la tasa de transferencia en las redes de computadoras.

1.4.1 Retardo en las redes de conmutación de paquetes

Recordemos que los paquetes se inician en un host (el origen), atraviesan una serie de routers y terminan su viaje en otro host (el destino). Cuando un paquete viaja de un nodo (host o router) al siguiente nodo (host o router) a lo largo de una ruta, el paquete sufre varios tipos de retardo en *cada* uno de los nodos de dicha ruta. Los más importantes de estos retardos son: el **retardo de procesamiento nodal**, el **retardo de cola**, el **retardo de transmisión** y el **retardo de propagación**; todos estos retardos se suman para proporcionar el **retardo nodal total**. Para adquirir un conocimiento profundo de la tecnología de conmutación de paquetes y de las redes de computadoras, es preciso comprender la naturaleza e importancia de estos retardos.

Tipos de retardos

Utilizaremos la Figura 1.16 para explorar estos retardos. Como parte de la ruta terminal a terminal entre el origen y el destino, un paquete se envía desde el nodo anterior a través del router A hasta el router B. Nuestro objetivo es caracterizar el retardo nodal en el router A. Observe que el router A dispone de un enlace de salida que lleva hasta el router B. Este enlace está precedido por una cola (o buffer). Cuando el paquete llega al router A procedente del nodo anterior, el router A examina la cabecera del paquete para determinar cuál es el enlace de salida apropiado para el paquete y luego dirige dicho paquete a ese enlace. En este ejemplo, el enlace de salida para el paquete es el único que lleva hasta el router B. Un paquete puede transmitirse a través de un enlace sólo si actualmente no se está transmitiendo ningún otro paquete a través de él y si no hay otros paquetes que le precedan en la cola; si el enlace está ocupado actualmente o si existen otros paquetes en la cola esperando para ese enlace, entonces el paquete recién llegado tendrá que ponerse a la cola.

Retardo de procesamiento

El tiempo requerido para examinar la cabecera del paquete y determinar dónde hay que enviarlo es parte del **retardo de procesamiento**. El retardo de procesamiento también incluye otros factores como el tiempo necesario para comprobar los errores de nivel de bit del paquete que se producen al transmitir los bits del paquete desde el nodo anterior al router A. Los retardos de procesamiento en los routers de alta velocidad suelen ser del orden de los microsegundos o menores. Una vez efectuado el procesamiento nodal, el router dirige el paquete a la

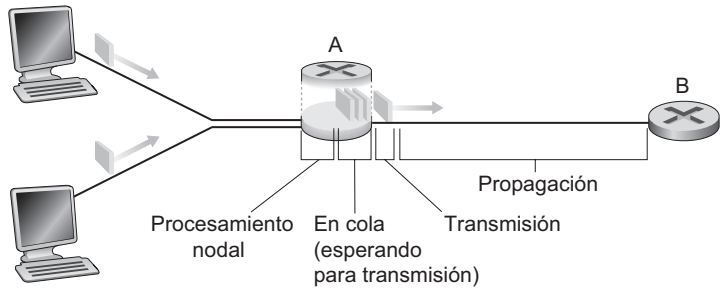


Figura 1.16 • Retardo nodal en el router A.

cola que precede al enlace que lleva al router B. (En el Capítulo 4 estudiaremos los detalles acerca de cómo funciona un router.)

Retardo de cola

En la cola, el paquete experimenta un **retardo de cola** al tener que esperar para ser transmitido a través del enlace. La duración del retardo de cola de un determinado paquete dependerá del número de paquetes que hayan llegado antes a la cola y que están esperando para ser transmitidos por el enlace. Si la cola está vacía y no se está transmitiendo ningún paquete actualmente, entonces el retardo de cola de nuestro paquete será cero. Por el contrario, si hay mucho tráfico y muchos paquetes también están esperando para ser transmitidos, el retardo de cola puede ser grande. Vamos a ver brevemente que el número de paquetes que un paquete entrante puede esperar encontrar es una función de la intensidad y de la naturaleza del tráfico que llega a la cola. En la práctica, los retardos de cola pueden ser del orden de microsegundos a milisegundos.

Retardo de transmisión

Suponiendo que los paquetes se transmiten de manera que el primero que llega es el primero que sale, lo que es una práctica común en las redes de conmutación de paquetes, nuestro paquete sólo puede ser transmitido después de que todos los paquetes que hayan llegado antes que él hayan sido transmitidos. Sea la longitud del paquete igual a L bits y la velocidad de transmisión del enlace del router A hasta el router B igual a R bits/segundo. Entonces, por ejemplo, para un enlace Ethernet a 10 Mbps, la velocidad es $R = 10$ Mbps; para un enlace Ethernet a 100 Mbps, la velocidad será $R = 100$ Mbps. El **retardo de transmisión** (también denominado retardo de almacenamiento y reenvío, como hemos visto en la Sección 1.3) será igual a L/R. Este es el tiempo necesario para introducir (es decir, transmitir) todos los bits del paquete por el enlace. Normalmente, en la práctica, los retardos de transmisión son del orden de los microsegundos a los milisegundos.

Retardo de propagación

Una vez que un bit ha entrado en el enlace, tiene que propagarse hasta el router B. El tiempo necesario para propagarse desde el principio del enlace hasta el router B es el **retardo de propagación**. El bit se propaga a la velocidad de propagación del enlace. Esta velocidad depende del medio físico del enlace (es decir, que el medio sea cable de fibra óptica, cable de cobre de par trenzado, etc.) y está comprendido en el rango entre

$$2 \cdot 10^8 \text{ metros/segundo y } 3 \cdot 10^8 \text{ metros/segundo}$$

que es igual o menor que la velocidad de la luz. El retardo de propagación es igual a la distancia entre dos routers dividida entre la velocidad de propagación. Es decir, el retardo de propagación es igual a d/s, donde d es la distancia entre el router A y el router B y s es la velocidad de propagación del enlace. Una vez que el último bit del paquete se ha propagado hasta el nodo B, éste y todos los bits anteriores del paquete se almacenan en el router B. A continuación, el router B lleva a cabo el reenvío. En las redes de área extensa, los retardos de propagación son del orden de los milisegundos.

Comparación de los retardos de transmisión y de propagación

Los recién llegados al campo de las redes de computadoras en ocasiones tienen dificultades para comprender la diferencia entre el retardo de transmisión y el de propagación. Esta diferencia es sutil pero importante. El retardo de transmisión es la cantidad de tiempo necesario para que el router saque fuera el paquete; es una función de la longitud del paquete y de la velocidad de transmisión del enlace, pero no tiene nada que ver con la distancia existente entre los dos routers. Por el contrario, el retardo de propagación es el tiempo que tarda un bit en propagarse de un router al siguiente; es una función de la distancia entre los dos routers, pero no tiene nada que ver con la longitud del paquete ni con la velocidad de transmisión del enlace.

Veamos una analogía que nos va a permitir clarificar los conceptos de retardo de transmisión y de retardo de propagación. Imagine una autopista en la que hay un puesto de peaje cada 100 kilómetros, como se muestra en la Figura 1.17. Podemos imaginar que los segmentos de autopista entre peajes son los enlaces y las casetas de peaje son los routers. Suponga que los automóviles viajan (es decir, se propagan) por la autopista a una velocidad de 100 km/hora (es decir, cuando un coche sale de un peaje, instantáneamente acelera hasta adquirir la velocidad de 100 km/hora y la mantiene entre puestos de peaje). Supongamos ahora que hay 10 coches que viajan en caravana unos detrás de otros en un orden fijo. Podemos pensar que cada coche es un bit y que la caravana es un paquete. Supongamos también que cada puesto de peaje da servicio (es decir, transmite) a los coches a una velocidad de un coche cada 12 segundos y que es tarde por la noche, por lo que en la autopista sólo se encuentra nuestra caravana de coches. Por último, supongamos que cuando el primer coche de la caravana llega a un peaje, espera en la entrada hasta que los otros nueve coches han llegado y se han detenido detrás de él (así, la caravana completa tiene que almacenarse en el peaje antes de poder ser reenviada). El tiempo necesario para que el peaje ponga a la caravana completa en la autopista es igual a (10 coches)/(5 coches/minuto) = 2 minutos. Este tiempo es análogo al retardo de transmisión de un router. El tiempo necesario para que un coche se desplace desde la salida del peaje hasta el siguiente puesto de peaje es 100 km/(100 km/hora) = 1 hora. Este tiempo es análogo al retardo de propagación. Por tanto, el tiempo que transcurre desde que la caravana queda colocada delante de un peaje hasta que vuelve a quedar colocada delante del siguiente peaje es la suma del tiempo de transmisión y el tiempo de propagación (en este caso, dicho tiempo será igual a 62 minutos).

Profundicemos un poco más en esta analogía. ¿Qué ocurriría si el tiempo de servicio del puesto de peaje invertido en una caravana fuera mayor que el tiempo que tarda un coche en viajar de un peaje al siguiente? Por ejemplo, supongamos que los coches viajan a una velocidad de 1.000 km/hora y que los peajes operan a una velocidad de un coche por minuto. Luego el retardo de desplazarse entre dos puestos de peaje será de 6 minutos y el tiempo que tarda el puesto de peaje en dar servicio a una caravana es de 10 minutos. En este caso, los

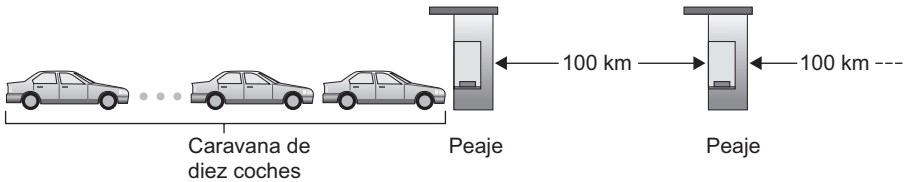


Figura 1.17 • Analogía de la caravana.

primeros coches de la caravana llegarán al segundo puesto de peaje antes de que los últimos coches de la caravana hayan salido del primer peaje. Esta situación también se produce en las redes de conmutación de paquetes: los primeros bits de un paquete pueden llegar a un router mientras que gran parte de los bits restantes del paquete todavía están esperando a ser transmitidos por el router anterior.

Si una imagen vale más que mil palabras, entonces una animación vale más que un millón de palabras. En el sitio web de este libro de texto se proporciona un applet Java interactivo que ilustra y compara los retardos de transmisión y de propagación. Animamos a los lectores a visitar este applet.

Sean d_{proc}, d_{cola}, d_{trans} y d_{prop} los retardos de procesamiento, de cola, de transmisión y de propagación, respectivamente. Entonces el retardo total nodal estará dado por:

$$d_{nodal} = d_{proc} + d_{cola} + d_{trans} + d_{prop}$$

Las contribuciones de estos componentes de retardo pueden variar significativamente. Por ejemplo, d_{prop} puede ser despreciable (digamos un par de microsegundos) para un enlace que conecte dos routers del mismo campus universitario; sin embargo, d_{prop} será igual a cientos de milisegundos para dos routers interconectados mediante un enlace vía satélite geoestacionario y puede ser el término dominante en la expresión que proporciona el retardo d_{nodal}. Del mismo modo, d_{trans} puede ser despreciable o significativo. Su contribución normalmente es despreciable para velocidades de transmisión de 10 Mbps y superiores (por ejemplo, para las redes LAN); sin embargo, puede ser igual a cientos de milisegundos para paquetes grandes de Internet enviados a través del modems de acceso telefónico de baja velocidad. El retardo de procesamiento, d_{proc}, suele ser despreciable; sin embargo, tiene una gran influencia sobre la tasa de transferencia máxima del router, que es la velocidad máxima a la que un router puede reenviar los paquetes.

1.4.2 Retardo de cola y pérdida de paquetes

El componente más complejo e interesante del retardo nodal es el retardo de cola, d_{cola}. De hecho, el retardo de cola es tan importante e interesante en las redes de computadoras que se han escrito miles de artículos y muchos libros sobre él [Bertsekas 1991; Daigle 1991; Kleinrock 1975, 1976; Ross 1995]. Aquí sólo vamos a abordarlo de forma intuitiva y panorámica; los lectores más curiosos pueden echar un vistazo a algunos de los libros que se ocupan de este tema (¡o incluso pueden escribir una tesis doctoral sobre el asunto!). A diferencia de los otros tres retardos (d_{proc}, d_{trans} y d_{prop}), el retardo de cola puede variar de un paquete a otro. Por ejemplo, si llegan 10 paquetes a una cola vacía al mismo tiempo, el primer paquete transmitido no sufrirá retardo de cola, mientras que el último paquete transmitido sufrirá un retardo de cola relativamente largo (mientras espera a que los restantes nueve paquetes sean transmitidos). Por tanto, al caracterizar el retardo de cola, suelen emplearse medidas estadísticas, como el retardo medio de cola, la varianza del retardo de cola y la probabilidad de que el retardo de cola exceda un cierto valor especificado.

¿En qué casos el retardo de cola es grande y en qué casos es insignificante? La respuesta a esta pregunta depende de la velocidad a la que llega el tráfico a la cola, de la velocidad de transmisión del enlace y de la naturaleza del tráfico entrante, es decir, si el tráfico llega periódicamente o a ráfagas. Vamos a profundizar en este punto. Sea a la velocidad media a la que llegan los paquetes a la cola (a se expresa en paquetes/segundo). Recuerde que R es la veloci-

dad de transmisión; es decir, es la velocidad (en bits/segundo) a la que los bits salen de la cola. Con el fin de simplificar, supongamos también que todos los paquetes constan de L bits. Luego la velocidad media a la que llegan los bits a la cola es igual a La bits/segundo. Supongamos por último que la cola es muy grande, por lo que podemos decir que puede almacenar un número infinito de bits. La relación La/R, denominada **intensidad de tráfico**, suele desempeñar un papel importante a la hora de estimar la magnitud del retardo de cola. Si $La/R > 1$, entonces la velocidad media a la que los bits llegan a la cola excede la velocidad a la que los bits pueden ser transmitidos desde la cola. En esta desafortunada situación, la cola tenderá a aumentar sin límite y el retardo de cola se aproximará a ¡infinito! Por tanto, una de las reglas de oro en la ingeniería de tráfico es: *diseñe su sistema de modo que la intensidad de tráfico no sea mayor que 1.*

Veamos ahora el caso en que $La/R \leqslant 1$. Aquí la naturaleza del tráfico entrante influye sobre el retardo de cola. Por ejemplo, si los paquetes llegan periódicamente, es decir, llega un paquete cada L/R segundos, entonces todos los paquetes llegarán a una cola vacía y no habrá retardo de cola. Por el contrario, si los paquetes llegan a ráfagas pero de forma periódica, puede aparecer un retardo medio de cola significativo. Por ejemplo, supongamos que llegan simultáneamente N paquetes cada $(L/R)N$ segundos. En este caso, el primer paquete transmitido no tiene asociado un retardo de cola, el segundo paquete transmitido presentará un retardo de cola de L/R segundos y, de forma más general, el n-ésimo paquete transmitido presentará un retardo de cola de $(n - 1)L/R$ segundos. Dejamos como ejercicio para el lector el cálculo del retardo medio de cola de este ejemplo.

Los dos ejemplos de llegada periódica de los paquetes que acabamos de describir se corresponden con casos teóricos. Normalmente, el proceso de llegada a una cola es *aleatorio;* es decir, las llegadas no siguen ningún patrón y los paquetes quedan separados por periodos de tiempo aleatorios. En este caso más realista, la cantidad La/R normalmente no es suficiente para caracterizar completamente las estadísticas del retardo de cola. Aun así, resulta útil tener una idea intuitiva de la magnitud del retardo de cola. En particular, si la intensidad de tráfico es próxima a cero, entonces las llegadas de paquetes serán pocas y estarán bastante espaciadas entre sí, por lo que será improbable que un paquete que llegue a la cola se encuentre con que hay otro paquete en la cola. Por tanto, el retardo medio de cola será próximo a cero. Por el contrario, cuando la intensidad de tráfico es próxima a 1, habrá intervalos de tiempo en los que la velocidad de llegada exceda a la capacidad de transmisión (a causa de las variaciones en la velocidad de llegada de los paquetes), por lo que se formará una cola durante estos periodos de tiempo; si la velocidad de llegada es menor que la capacidad de transmisión, la longitud de la cola disminuirá. Sin embargo, cuando la intensidad de tráfico se aproxime a 1, la longitud media de la cola será cada vez mayor. La dependencia cualitativa del retardo medio de cola con relación a la intensidad de tráfico se muestra en la Figura 1.18.

Un aspecto importante de la Figura 1.18 es el hecho de que cuando la intensidad de tráfico se aproxima a 1, el retardo medio de cola aumenta rápidamente. Un pequeño porcentaje de aumento en la intensidad dará lugar a un incremento en porcentaje muy grande del retardo. Es posible que haya experimentado este fenómeno en una autopista. Si viaja regularmente por una autopista que habitualmente está congestionada, quiere decir que la intensidad de tráfico en esa autopista es próxima a 1. En el caso de que se produzca un suceso que dé lugar a una cantidad de tráfico ligeramente mayor que la usual, los retardos que se experimenten pueden llegar a ser enormes.

Con el fin de que entienda bien lo que son los retardos de cola animamos de nuevo al lector a visitar el sitio web dedicado a este libro, donde se proporciona un applet de Java

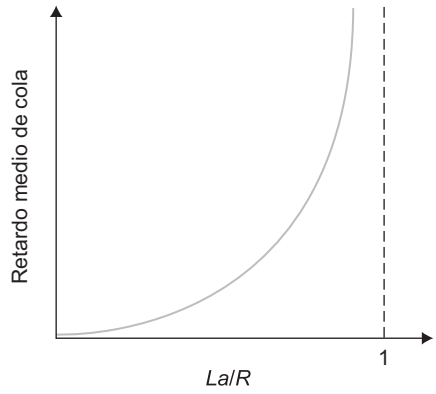


Figura 1.18 • Dependencia del retardo medio de cola con relación a la intensidad de tráfico.

interactivo para una cola. Si establece una velocidad de llegada de los paquetes lo suficientemente alta como para que la intensidad de tráfico sea mayor que 1, comprobará que la cola aumenta lentamente con el tiempo.

Pérdida de paquetes

En la sección anterior hemos supuesto que la cola es capaz de almacenar un número infinito de paquetes. En la práctica, una cola para acceder a un enlace tiene una capacidad finita, aunque la capacidad de la cola depende fundamentalmente del diseño y del coste del router. Puesto que la capacidad de cola es finita, los retardos de los paquetes realmente no se aproximan a infinito cuando la intensidad de tráfico se aproxima a 1. En su lugar, un paquete puede llegar y encontrarse con que la cola está llena. Si no hay sitio para almacenar un paquete, el router lo **elimina**; es decir, el paquete se **pierde**. Este desbordamiento de una cola puede verse también en el applet de Java, cuando la intensidad de tráfico es mayor que 1.

Desde el punto de vista de un sistema terminal, un paquete perdido es un paquete que ha sido transmitido al núcleo de la red pero que nunca sale de la red en su destino. El número de paquetes perdidos aumenta cuando la intensidad de tráfico aumenta. Por tanto, el rendimiento de un nodo suele medirse no sólo en función de los retardos, sino también en función de la probabilidad de pérdida de paquetes. Como veremos en los siguientes capítulos, un paquete perdido puede retransmitirse de terminal a terminal para garantizar que todos los datos sean transferidos desde el origen hasta su destino.

1.4.3 Retardo terminal a terminal

Hasta el momento nos hemos centrado en el retardo nodal, es decir, el retardo en un único router. Ahora vamos a ocuparnos del retardo total entre el origen y el destino. Para entender este concepto, suponga que hay $N - 1$ routers entre el host de origen y el host de destino. Suponga también, por el momento, que la red no está congestionada (por lo que los retardos de cola son despreciables), el retardo de procesamiento en cada router y en el host de origen es d_{proc}, la velocidad de transmisión de salida de cada router y del host de origen es de

R bits/segundo y el retardo de propagación en cada enlace es igual a d_{prop}. Los retardos nodales se suman para proporcionar el retardo terminal a terminal, luego

$$d_{terminal\text{-}terminal} = N\,(d_{proc} + d_{trans} + d_{prop})$$

donde, de nuevo, $d_{trans} = L/R$, siendo L el tamaño del paquete. Dejamos para el lector el ejercicio de generalizar esta fórmula para el caso en que los retardos en los nodos sean diferentes y exista un retardo medio de cola en cada nodo.

Traceroute

Para ver el orden de magnitud del retardo terminal a terminal de una red de computadoras, podemos utilizar el programa Traceroute. Se trata de un programa simple que se puede ejecutar en cualquier host de Internet. Cuando el usuario especifica un nombre de host de destino, el programa del host de origen envía varios paquetes especiales al destino. A medida que estos paquetes se dirigen a su destino, pasan a través de una serie de routers. Cuando un router recibe uno de estos paquetes especiales, devuelve al origen un mensaje corto que contiene el nombre y la dirección del router.

Más concretamente, suponga que hay $N - 1$ routers entre el origen y el destino. Luego el origen enviará N paquetes especiales a la red dirigidos al destino final. Estos N paquetes especiales se marcan de *1* a *N,* quedando el primer paquete marcado como *1* y el último como *N*. Cuando el router n-ésimo recibe el paquete n-ésimo marcado como *N,* el router no reenvía el paquete hacia su destino, sino que devuelve un mensaje al origen. Cuando el host de destino recibe el paquete n-ésimo, también devuelve un mensaje al origen. El origen registra el tiempo transcurrido entre el momento en que envió un paquete y el momento en que recibe el correspondiente mensaje de vuelta; también registra el nombre y la dirección del router (o del host de destino) que devuelve el mensaje. De esta forma, el origen puede reconstruir la ruta seguida por los paquetes que fluyen desde el origen hasta el destino, y el origen puede determinar los retardos de ida y vuelta de todos los routers que intervienen en el proceso. Traceroute repite el proceso que acabamos de describir tres veces, de modo que el origen realmente envía *3* • N paquetes al destino. El documento RFC 1393 describe en detalle el programa Traceroute.

He aquí un ejemplo de la salida proporcionada por el programa Traceroute, en el que se ha trazado la ruta desde el host de origen gaia.cs.umass.edu (en la Universidad de Massachusetts) al host cis.poly.edu (en la Universidad Politécnica de Brooklyn). La salida consta de seis columnas: la primera de ellas contiene el valor n descrito anteriormente, es decir, el número del router a lo largo de la ruta; la segunda columna especifica el nombre del router; la tercera indica la dirección del router (con el formato xxx.xxx.xxx.xxx); las tres últimas columnas especifican los retardos de ida y vuelta correspondientes a los tres experimentos. Si el origen recibe menos de tres mensajes de cualquier router (debido a la pérdida de paquetes en la red), Traceroute incluye un asterisco justo después del número de router y proporciona menos de tres tiempos de ida y vuelta para dicho router.

```
1 cs-gw (128.119.240.254) 1.009 ms 0.899 ms 0.993 ms
2 128.119.3.154 (128.119.3.154) 0.931 ms 0.441 ms 0.651 ms
3 border4-rt-gi-1-3.gw.umass.edu (128.119.2.194) 1.032 ms 0.484 ms 0.451 ms
4 acr1-ge-2-1-0.Boston.cw.net (208.172.51.129) 10.006 ms 8.150 ms 8.460 ms
5 agr4-loopback.NewYork.cw.net (206.24.194.104) 12.272 ms 14.344 ms 13.267 ms
```

```
6 acr2-loopback.NewYork.cw.net (206.24.194.62) 13.225 ms 12.292 ms 12.148 ms
7 pos10-2.core2.NewYork1.Level3.net (209.244.160.133) 12.218 ms 11.823 ms 11.793 ms
8 gige9-1-52.hsipaccess1.NewYork1.Level3.net (64.159.17.39) 13.081 ms 11.556 ms 13.297 ms
9 p0-0.polyu.bbnplanet.net (4.25.109.122) 12.716 ms 13.052 ms 12.786 ms
10 cis.poly.edu (128.238.32.126) 14.080 ms 13.035 ms 12.802 ms
```

Podemos ver en esta traza que existen nueve routers entre el origen y el destino. La mayor parte de estos routers tiene un nombre y todos ellos tienen direcciones. Por ejemplo, el nombre del router 3 es `border4-rt-gi-1-3.gw.umass.edu` y su dirección es `128.119.2.194`. Si observamos los datos proporcionados para este mismo router, vemos que en la primera de las tres pruebas el retardo de ida y vuelta entre el origen y el router ha sido de 1,03 milisegundos. Los retardos de ida y vuelta para las dos pruebas siguientes han sido 0,48 y 0,45 milisegundos, respectivamente. Estos retardos de ida y vuelta contienen todos los retardos que acabamos de estudiar, incluyendo los retardos de transmisión, de propagación, de procesamiento del router y de cola. Puesto que el retardo de cola varía con el tiempo, el retardo de ida y vuelta del paquete n envíado al router n puede, en ocasiones, ser mayor que el retardo de ida y vuelta del paquete $n+1$ enviado al router $n+1$. Efectivamente, puede observar este fenómeno en el ejemplo anterior: los retardos correspondientes al router 6 son mayores que los correspondientes al router 7.

¿Desea probar el programa Traceroute? Le recomendamos *vivamente* que visite el sitio http:// www.traceroute.org, donde se proporciona una interfaz web a una extensa lista de orígenes para el trazado de rutas. Seleccione un origen y especifique el nombre de host de cualquier destino. El programa Traceroute hará entonces todo el trabajo. Hay disponibles diversos programas software gratuitos que proporcionan una interfaz gráfica para Traceroute; uno de nuestros programa favoritos es PingPlotter [PingPlotter 2009].

Retardos de los sistemas terminales, de las aplicaciones y otros

Además de los retardos de procesamiento, de transmisión y de propagación, en los sistemas terminales pueden existir retardos adicionales significativos. Por ejemplo, los modems de acceso telefónico introducen un retardo de modulación/codificación, que puede ser del orden de decenas de milisegundos (los retardos de modulación/codificación para otras tecnologías de acceso, como Ethernet, modems por cable y DSL, son menos significativos y suelen ser despreciables). Un sistema terminal que desea transmitir un paquete a través de un medio compartido (por ejemplo, en un escenario WiFi o Ethernet) puede retardar su transmisión *a propósito* como parte de su protocolo, para compartir el medio con otros sistemas terminales. Veremos estos protocolos en detalle en el Capítulo 5. Otro retardo importante es el retardo de empaquetamiento del medio, que aparece en las aplicaciones de Voz sobre IP (VoIP, *Voice-over-IP*). En VoIP, el lado emisor debe, en primer lugar, rellenar un paquete con voz digitalizada codificada antes de pasar el paquete a Internet. El tiempo que se tarda en rellenar un paquete (lo que se denomina retardo de empaquetamiento) puede ser significativo y puede repercutir en la calidad percibida por el usuario de una llamada VoIP. Este problema se abordará más detalladamente en uno de los problemas del final del capítulo.

1.4.4 Tasa de transferencia en las redes de computadoras

Además de los retardos y la pérdida de paquetes, otra medida crítica de rendimiento de las redes de computadoras es la tasa de transferencia de terminal a terminal. Para definir la tasa

de transferencia, consideremos la transferencia de un archivo de gran tamaño desde el host A al host B a través de una red. Por ejemplo, esta transferencia podría consistir en transferir un clip de vídeo de gran tamaño desde un par (*peer*) a otro en un sistema de compartición de archivos P2P. La **tasa de transferencia instantánea** en cualquier instante de tiempo es la velocidad (en bits/segundo) a la que el host B recibe el archivo. (Muchas aplicaciones, incluyendo muchos sistemas de compartición de archivos P2P, muestran la tasa de transferencia instantánea durante las descargas en la interfaz del usuario; ¡es posible que ya se haya fijado anteriormente en este detalle!) Si el archivo consta de F bits y la transferencia dura T segundos hasta que el host B recibe los F bits, entonces la **tasa media de transferencia** del archivo es igual a F/T bits/segundo. En algunas aplicaciones, tales como la telefonía por Internet, es deseable tener un retardo pequeño y una tasa de transferencia instantánea por encima de un cierto umbral (por ejemplo, por encima de 24 kbps para ciertas aplicaciones de telefonía por Internet y por encima de 256 kbps para las aplicaciones de vídeo en tiempo real). Para otras aplicaciones, entre las que se incluyen aquéllas que implican la transferencia de archivos, el retardo no es crítico, pero es deseable que la tasa de transferencia sea lo más alta posible.

Con el fin de comprender mejor el importante concepto de tasa de transferencia, vamos a ver algunos ejemplos. La Figura 1.19(a) muestra dos sistemas terminales, un servidor y un cliente, conectados mediante dos enlaces de comunicaciones y un router. Consideremos la tasa de transferencia para transmitir un archivo desde el servidor al cliente. Sea R_s la velocidad del enlace entre el servidor y el router, y sea R_c la velocidad del enlace entre el router y el cliente. Supongamos que los únicos bits que están siendo enviados a través de la red son los que van desde el servidor al cliente. En este escenario ideal, ¿cuál es la tasa de transferencia del servidor al cliente? Para responder a esta pregunta, podemos pensar en los bits como en un *fluido* y en los enlaces de comunicaciones como en las *tuberías*. Evidentemente, el servidor no puede bombear los bits a través de su enlace a una velocidad mayor que R_s bps; y el router no puede reenviar los bits a una velocidad mayor que R_c bps. Si $R_s < R_c$, entonces los bits bombeados por el servidor "fluirán" a través del router y llegarán al cliente a una velocidad de R_s bps, obteniéndose una tasa de transferencia de R_s bps. Si, por el contrario, $R_c < R_s$, entonces el router no podrá reenviar los bits tan rápidamente como los recibe. En

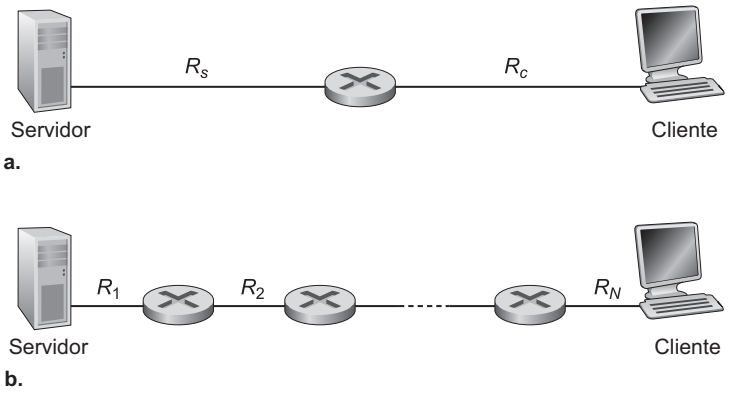


Figura 1.19 • Tasa de transferencia para la transmisión de un archivo desde un servidor a un cliente.

este caso, los bits abandonarán el router a una velocidad de sólo R_c, dando lugar a una tasa de transferencia de terminal a terminal igual a R_c. (Observe también que si continúan llegando bits al router a una velocidad R_s, y siguen abandonando el router a una velocidad igual a R_c, la cantidad de bits en el router que están esperando a ser transmitidos hacia el cliente aumentará constantemente, lo que es una situación nada deseable.) Por tanto, en esta sencilla red de dos enlaces, la tasa de transferencia es mín$\{R_c, R_s\}$, es decir, la velocidad de transmisión del **enlace cuello de botella**. Una vez determinada la tasa de transferencia, podemos aproximar el tiempo que se tardará en transferir un archivo de gran tamaño de F bits desde el servidor al cliente como F/mín$\{R_s, R_c\}$. Veamos un ejemplo concreto. Suponga que está descargando un archivo MP3 de $F = 32$ millones de bits, el servidor tiene una velocidad de transmisión de $R_s = 2$ Mbps y la velocidad del enlace es $R_c = 1$ Mbps. El tiempo necesario para transferir el archivo será igual a 32 segundos. Por supuesto, estas expresiones para la tasa de transferencia y el tiempo de transferencia son únicamente aproximaciones, ya que no se han tenido en cuenta las cuestiones relativas al protocolo y a nivel de paquete.

La Figura 1.19(b) muestra una red con N enlaces entre el servidor y el cliente, siendo las velocidades de transmisión de los N enlaces iguales a R_1, R_2,..., R_N. Aplicando el mismo análisis que para la red de dos enlaces, podemos determinar que la tasa de transferencia para transferir un archivo desde el servidor al cliente es mín$\{R_1, R_2,...,R_N\}$, que es de nuevo la velocidad de transmisión del enlace cuello de botella existente en la ruta entre el servidor y el cliente.

Veamos ahora otro ejemplo inspirado en la red Internet de hoy día. La Figura 1.20(a) muestra dos sistemas terminales, un servidor y un cliente, conectados a una red de computadoras. Considere la tasa de transferencia para transmitir un archivo desde el servidor al cliente. El servidor está conectado a la red mediante un enlace de acceso cuya velocidad es R_s y el cliente está conectado a la red mediante un enlace de acceso de velocidad R_c. Supongamos ahora que todos los enlaces existentes en el núcleo de la red de comunicaciones tienen velocidades de transmisión muy altas, muy por encima de R_s y R_c. Ciertamente, hoy en día, el núcleo de Internet está sobredimensionado, con enlaces de alta velocidad que experimentan una congestión muy baja [Akella 2003]. Supongamos también que únicamente se están enviando a la red los bits que se transfieren desde el servidor al cliente. Dado que el núcleo de la red es como una tubería ancha en este ejemplo, la velocidad a la que los bits pueden fluir desde el origen hasta el destino de nuevo es el mínimo de R_s y R_c, es decir, la tasa de transferencia es igual a mín$\{R_s, R_c\}$. Por tanto, normalmente, el factor de restricción de la tasa de transferencia en Internet actualmente es la red de acceso.

Veamos un último ejemplo. Vamos a utilizar la Figura 1.20(b); en ella vemos que hay 10 servidores y 10 clientes conectados al núcleo de la red de computadoras. En este ejemplo, tienen lugar 10 descargas simultáneas, lo que implica a 10 pares cliente-servidor. Supongamos que estas 10 descargas son el único tráfico existente en la red a un mismo tiempo. Como se muestra en la figura, hay un enlace en el núcleo que es atravesado por las 10 descargas. Sea R la velocidad de transmisión de este enlace R. Supongamos que los enlaces de acceso de todos los servidores tienen la misma velocidad R_s, los enlaces de acceso de todos los clientes tienen la misma velocidad R_c y las velocidades de transmisión de todos los enlaces del núcleo (excepto el enlace común de velocidad R) tienen velocidades mucho mayores que R_s, R_c y R. Ahora deseamos saber cuáles son las tasas de transferencia de las descargas. Evidentemente, si la velocidad del enlace común, R, es por ejemplo cien veces mayor que R_s y R_c, entonces la tasa de transferencia de cada descarga será de nuevo mín$\{R_s, R_c\}$. Pero, ¿qué ocurre si la velocidad del enlace común es del mismo orden que R_s

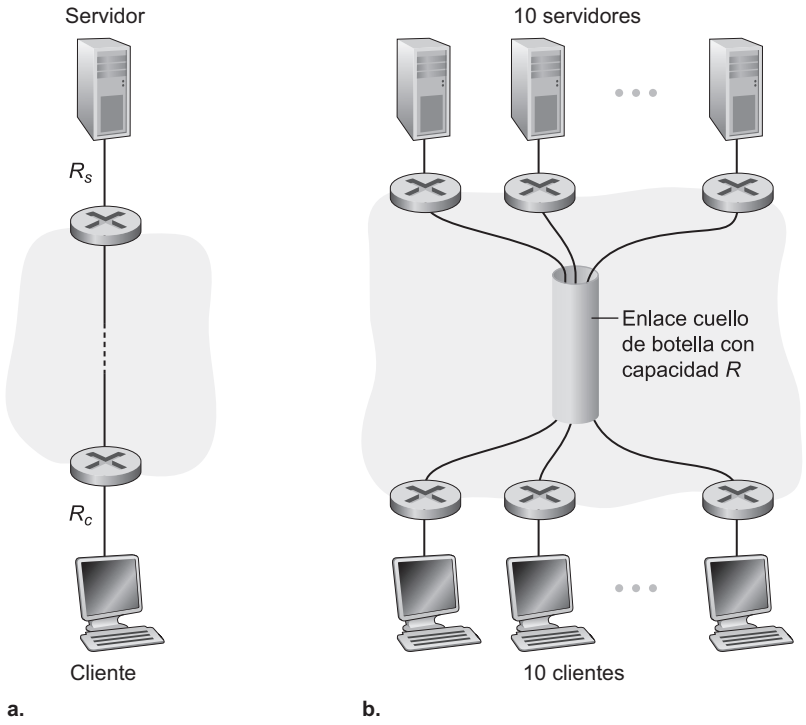


Figura 1.20 • Tasa de transferencia terminal a terminal: (a) un cliente descarga un archivo de un servidor; (b) 10 clientes descargando con 10 servidores.

y R_c? ¿Cuál será la tasa de transferencia en este caso? Veamos un ejemplo concreto. Supongamos que $R_s = 2$ Mbps, $R_c = 1$ Mbps, $R = 5$ Mbps y que el enlace común divide su velocidad de transmisión en partes iguales entre las 10 descargas. Luego, ahora, el cuello de botella para cada descarga ya no se encuentra en la red de acceso, sino en el enlace compartido del núcleo que sólo proporciona una tasa de transferencia de 500 kbps a cada descarga. Luego la tasa de transferencia terminal a terminal para cada descarga ahora se ha reducido a 500 kbps.

Los ejemplos de las Figuras 1.19 y 1.20(a) demuestran que la tasa de transferencia depende de las velocidades de transmisión de los enlaces a través de los que fluyen los datos. Hemos visto que cuando no existe ningún otro tráfico, la tasa de transferencia puede simplemente aproximarse a la velocidad mínima de transmisión a lo largo de la ruta entre el origen y el destino. El ejemplo de la Figura 1.20(b) muestra que, generalmente, la tasa de transferencia depende no sólo de las velocidades de transmisión de los enlaces a lo largo de la ruta, sino también del tráfico existente. En particular, un enlace con una velocidad de transmisión alta puede ser el enlace cuello de botella para la transferencia de un archivo si hay muchos otros flujos de datos atravesando también ese enlace. Examinaremos más detalladamente la tasa de transferencia en las redes de computadoras en los problemas que el lector realizará en su casa y en los capítulos siguientes.

1.5 Capas de protocolos y sus modelos de servicio

Después de lo que hemos visto hasta aquí, parece que Internet es un sistema *extremadamente* complicado. Hemos visto que son muchas las piezas que conforman Internet: numerosas aplicaciones y protocolos, distintos tipos de sistemas terminales, dispositivos de conmutación de paquetes y diversos tipos de medios para los enlaces. Dada esta enorme complejidad, ¿tenemos alguna esperanza de poder organizar una arquitectura de red o al menos nuestra exposición sobre la misma? Afortunadamente, la respuesta a ambas preguntas es sí.

1.5.1 Arquitectura en capas

Antes de intentar organizar nuestras ideas sobre la arquitectura de Internet, vamos a ver una analogía humana. Realmente, de forma continua estamos tratando con sistemas complejos en nuestra vida cotidiana. Imagine que alguien le pide que describa, por ejemplo, cómo funciona el sistema de líneas aéreas. ¿Qué estructura utilizaría para describir este complejo sistema que emplea agencias de viajes para la venta de billetes, personal para el control de equipajes, puertas de embarque, pilotos, aviones, control de tráfico aéreo y un sistema de ámbito mundial para dirigir los aviones? Una forma de describir este sistema podría consistir en describir la serie de acciones que usted realiza (o que otros realizan para usted) cuando se vuela en un avión. En primer lugar, usted compra un billete, luego factura el equipaje, se dirige a la puerta de embarque y por último sube al avión. El avión despega y se dirige a su destino. Una vez que el avión ha tomado tierra, usted desembarca y recoge su equipaje. Si el viaje ha sido malo, se quejará de ello a la agencia de viajes (lo que, por supuesto, no le servirá de nada). Este escenario se muestra en la Figura 1.21.

Podemos ver algunas analogías con las redes de computadoras: la compañía área le traslada desde un origen hasta un destino, al igual que Internet trasporta un paquete desde un origen hasta un destino. Pero ésta no es la analogía que estábamos buscando. Lo que queremos es encontrar una cierta *estructura* en la Figura 1.21. Si nos fijamos en esta figura, observaremos que hay una función Billete en cada extremo; también existe una función Equipaje

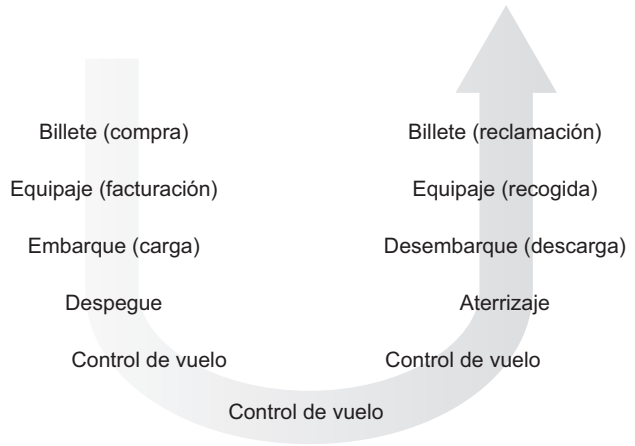


Figura 1.21 • Acciones para realizar un viaje en avión.

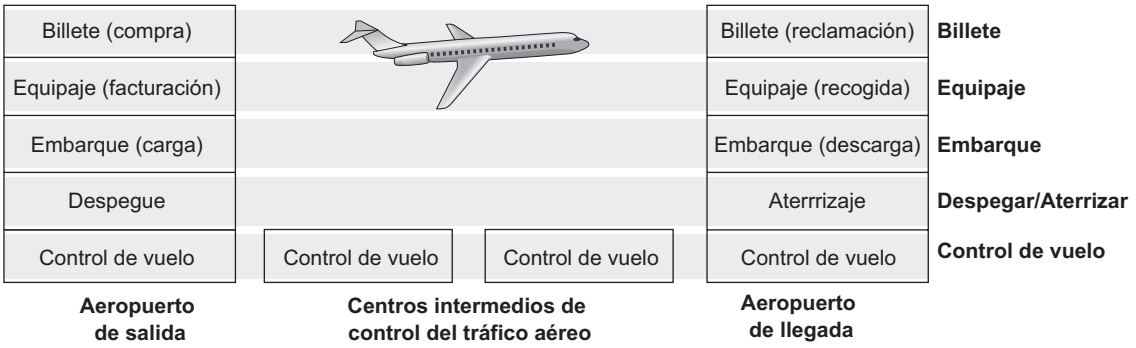


Figura 1.22 • Disposición de capas en horizontal de las funcionalidades de una compañía área.

para los pasajeros que tienen un billete y una función Embarque para los pasajeros que tienen billete y han facturado su equipaje. Para los pasajeros que han embarcado (es decir, que han adquirido un billete, han facturado las maletas y han embarcado), existe una función de despegue y aterrizaje y durante el vuelo, hay una función de control del avión. Esto sugiere que podemos fijarnos en las funcionalidades señaladas en la Figura 1.21 de forma *horizontal*, como se muestra en la Figura 1.22.

En la Figura 1.22 se han separado las distintas funciones de la compañía aérea en capas, proporcionando un marco de trabajo en el que podemos explicar cómo se realiza un viaje en avión. Observe que cada capa, combinada con las capas que tiene por debajo, implementa una cierta funcionalidad, un cierto *servicio.* En las capas Billete e inferiores se lleva a cabo la transferencia de una persona de un mostrador de línea aérea a otro. En las capas Equipaje e inferiores se realiza la transferencia de una persona y su equipaje dese el punto de facturación hasta la recogida de equipaje. Observe que la capa Equipaje sólo proporciona este servicio a las personas que ya han adquirido un billete. En la capa Embarque, se realiza la transferencia embarque/desembarque de una persona y su equipaje. En la capa Despegue/Aterrizaje, se realiza la transferencia pista a pista de personas y equipajes. Cada capa proporciona su servicio (1) llevando a cabo determinadas acciones dentro de dicha capa (por ejemplo, en la capa Embarque, se hace subir y bajar al pasaje del avión) y (2) utilizando los servicios de la capa que tiene directamente debajo de ella (por ejemplo, en la capa Embarque, utilizando el servicio de transferencia de pasajeros pista a pista de la capa Despegue/Aterrizaje).

Una arquitectura de capas nos permite estudiar una parte específica y bien definida de un sistema más grande y complejo. Esta simplificación por sí misma tiene un valor considerable al proporcionar modularidad, haciendo mucho más fácil modificar la implementación del servicio suministrado por la capa. Dado que la capa proporciona el mismo servicio a la capa que tiene por encima de ella y emplea los mismos servicios de la capa que tiene por debajo, el resto del sistema permanece invariable cuando se modifica la implementación de una capa. (Tenga en cuenta que cambiar la implementación de un servicio es muy diferente a cambiar el propio servicio.) Por ejemplo, si se modificara la función Embarque para que las personas embarcaran y desembarcaran por alturas, el resto del sistema de la compañía aérea no se vería afectado, ya que la capa Embarque continuaría llevando a cabo la misma función (cargar y descargar personas); simplemente implementa dicha función de una forma

diferente después de realizar el cambio. En sistemas complejos de gran tamaño que se actualizan constantemente, la capacidad de modificar la implementación de un servicio sin afectar al resto de los componentes del sistema es otra importante ventaja de la disposición en capas.

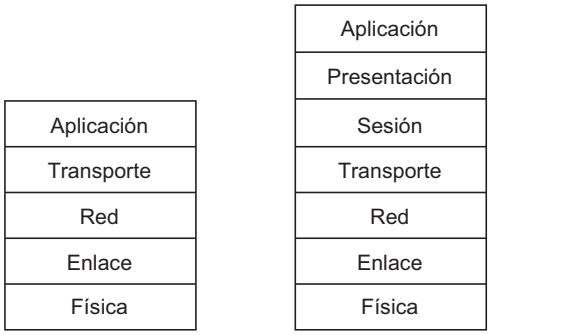
Capas de protocolos

Pero ya hemos hablado suficiente de compañías aéreas. Dirijamos ahora nuestra atención a los protocolos de red. Para proporcionar una estructura al diseño de protocolos de red, los diseñadores de redes organizan los protocolos (y el hardware y el software de red que implementan los protocolos) en **capas**. Cada protocolo pertenece a una de las capas, del mismo modo que cada función en la arquitectura de la compañía aérea de la Figura 1.22 pertenecía a una capa. De nuevo, estamos interesados en los **servicios** que ofrece una capa a la capa que tiene por encima, lo que se denomina **modelo de servicio** de capa. Como en el caso del ejemplo de la compañía aérea, cada capa proporciona su servicio (1) llevando a cabo ciertas acciones en dicha capa y (2) utilizando los servicios de la capa que tiene directamente debajo de ella. Por ejemplo, los servicios proporcionados por la capa n pueden incluir la entrega fiable de mensajes de un extremo de la red al otro. Esto podría implementarse mediante un servicio no fiable de entrega de mensajes terminal a terminal de la capa $n - 1$, y añadiendo la funcionalidad de la capa n para detectar y retransmitir los mensajes perdidos.

Una capa de protocolo puede implementarse por software, por hardware o mediante una combinación de ambos. Los protocolos de la capa de aplicación, como HTTP y SMTP, casi siempre se implementan por software en los sistemas terminales, al igual que los protocolos de la capa de transporte. Puesto que la capa física y las capas de enlace de datos son responsables de manejar la comunicación a través de un enlace específico, normalmente se implementan en las tarjetas de interfaz de red (por ejemplo, tarjetas Ethernet o WiFi) asociadas con un determinado enlace. La capa de red a menudo es una implementación mixta de hardware y software. Observe también que al igual que las funciones de la arquitectura de la compañía aérea estaban distribuidas entre los distintos aeropuertos y el centro de control de vuelo que formaban el sistema, también un protocolo de capa n está *distribuido* entre los sistemas terminales, los dispositivos de conmutación de paquetes y los restantes componentes que conforman la red. Es decir, suele haber una parte del protocolo de capa n en cada uno de estos componentes de red.

Las capas de protocolos presentan ventajas conceptuales y estructurales. Como hemos visto, las capas proporcionan una forma estructurada de estudiar los componentes del sistema. Además, la modularidad facilita la actualización de los componentes del sistema. Sin embargo, tenemos que comentar que algunos investigadores e ingenieros de redes se oponen vehemente a la estructura de capas [Wakeman 1992]. Un potencial inconveniente de la estructura de capas es que una capa puede duplicar la funcionalidad de la capa inferior. Por ejemplo, muchas pilas de protocolos proporcionan una función de recuperación de errores tanto por enlace como extremo a extremo. Un segundo potencial inconveniente es que la funcionalidad de una capa puede precisar información (por ejemplo, un valor de una marca temporal) que sólo existe en otra capa, y esto viola el objetivo de la separación en capas.

Cuando los protocolos de las distintas capas se toman en conjunto se habla de **pila de protocolos**. La pila de protocolos de Internet consta de cinco capas: capa física, capa de enlace, capa de red, capa de transporte y capa de aplicación, como se muestra en la



a. Pila de protocolos de Internet de cinco capas

b. Modelo de referencia OSI de ISO de siete capas

Figura 1.23 • Pila de protocolos de Internet (a) y modelo de referencia OSI (b).

Figura1.23(a). Si examina el Contenido, comprobará que hemos organizado el libro utilizando las capas de la pila de protocolos de Internet. Vamos a aplicar el **enfoque descendente**, abordando en primer lugar la capa de aplicación y continuando hacia abajo por la pila.

Capa de aplicación

La capa de aplicación es donde residen las aplicaciones de red y sus protocolos. La capa de aplicación de Internet incluye muchos protocolos, tales como el protocolo HTTP (que permite la solicitud y transferencia de documentos web), SMTP (que permite la transferencia de mensajes de correo electrónico) y FTP (que permite la transferencia de archivos entre dos sistemas terminales). Veremos que determinadas funciones de red, como la traducción de los nombres legibles que utilizamos las personas para los sistemas terminales de Internet (por ejemplo, www.ietf.org) en direcciones de red de 32 bits se realiza también con la ayuda de un protocolo específico de la capa de aplicación, en concreto, el Sistema de nombres de dominio (DNS, *Domain Name System*). En el Capítulo 2 veremos que es muy fácil crear e implantar nuestros propios protocolos de la capa de aplicación.

Un protocolo de la capa de aplicación está distribuido a lo largo de varios sistemas terminales, estando la aplicación en un sistema terminal que utiliza el protocolo para intercambiar paquetes de información con la aplicación de otro sistema terminal. A este paquete de información de la capa de aplicación lo denominaremos **mensaje**.

Capa de transporte

La capa de transporte de Internet transporta los mensajes de la capa de aplicación entre los puntos terminales de la aplicación. En Internet, existen dos protocolos de transporte, TCP y UDP, pudiendo cada uno de ellos transportar los mensajes de la capa de aplicación. TCP ofrece a sus aplicaciones un servicio orientado a la conexión. Este servicio proporciona un suministro garantizado de los mensajes de la capa de aplicación al destino y un mecanismo de control del flujo (es decir, adaptación de las velocidades del emisor y el receptor). TCP también divide los mensajes largos en segmentos más cortos y proporciona un mecanismo de control de congestión, de manera que un emisor regula su velocidad de

transmisión cuando la red está congestionada. El protocolo UDP proporciona a sus aplicaciones un servicio sin conexión. Es un servicio básico que no ofrece ninguna fiabilidad, ni control de flujo, ni control de congestión. En este libro, denominaremos a los paquetes de la capa de transporte **segmentos**.

Capa de red

La capa de red de Internet es responsable de trasladar los paquetes de la capa de red, conocidos como **datagramas**, de un host a otro. El protocolo de la capa de transporte (TCP o UDP) de Internet de un host de origen pasa un segmento de la capa de transporte y una dirección de destino a la capa de red, al igual que damos al servicio de correo postal una carta con una dirección de destino. Luego, la capa de red proporciona el servicio de suministrar el segmento a la capa de transporte del host de destino.

La capa de red de Internet incluye el conocido protocolo IP, que define los campos del datagrama, así como la forma en que actúan los sistemas terminales y los routers sobre estos campos. Existe un único protocolo IP y todos los componentes de Internet que tienen una capa de red deben ejecutar el protocolo IP. La capa de red de Internet también contiene los protocolos de enrutamiento que determinan las rutas que los datagramas siguen entre los orígenes y los destinos. Internet dispone de muchos protocolos de enrutamiento. Como hemos visto en la Sección 1.3, Internet es una red de redes y, dentro de una red, el administrador de la red puede ejecutar cualquier protocolo de enrutamiento que desee. Aunque la capa de red contiene tanto el protocolo IP como numerosos protocolos de enrutamiento, suele hacerse referencia a ella simplemente como la capa IP, lo que refleja el hecho de que IP es el pegamento que une todo Internet.

Capa de enlace

La capa de red de Internet encamina un datagrama a través de una serie de routers entre el origen y el destino. Para trasladar un paquete de un nodo (host o router) al siguiente nodo de la ruta, la capa de red confía en los servicios de la capa de enlace. En concreto, en cada nodo, la capa de red pasa el datagrama a la capa de enlace, que entrega el datagrama al siguiente nodo existente a lo largo de la ruta. En el siguiente nodo, la capa de enlace pasa el datagrama a la capa de red.

Los servicios proporcionados por la capa de enlace dependen del protocolo de la capa de enlace concreto que se emplee en el enlace. Por ejemplo, algunos protocolos de la capa de enlace proporcionan una entrega fiable desde el nodo transmisor, a través del enlace y hasta el nodo receptor. Observe que este servicio de entrega fiable es diferente del servicio de entrega fiable de TCP, que lleva a cabo una entrega fiable desde un sistema terminal a otro. Entre los ejemplos de protocolos de la capa de enlace se incluyen Ethernet, WiFi y el Protocolo punto a punto (PPP, *Point-to-Point Protocol*). Puesto que normalmente los datagramas necesitan atravesar varios enlaces para viajar desde el origen hasta el destino, un datagrama puede ser manipulado por distintos protocolos de la capa de enlace en los distintos enlaces disponibles a lo largo de la ruta. Por ejemplo, un datagrama puede ser manipulado por Ethernet en un enlace y por PPP en el siguiente enlace. La capa de red recibirá un servicio diferente por parte de cada uno de los distintos protocolos de la capa de enlace. En este libro, denominaremos a los paquetes de la capa de enlace **tramas**.

Capa física

Mientras que el trabajo de la capa de enlace es mover las tramas completas de un elemento de la red hasta el elemento de red adyacente, el trabajo de la capa física es el de mover los *bits individuales* dentro de la trama de un nodo al siguiente. Los protocolos de esta capa son de nuevo dependientes del enlace y, por tanto, dependen del medio de transmisión del enlace (por ejemplo, cable de cobre de par trenzado o fibra óptica monomodo). Por ejemplo, Ethernet dispone de muchos protocolos de la capa física: uno para cable de cobre de par trenzado, otro para cable coaxial, otro para fibra, etc. En cada caso, los bits se desplazan a través del enlace de forma diferente.

El modelo OSI

Una vez vista en detalle la pila de protocolos de Internet, deberíamos mencionar que no es la única pila de protocolos existente. En concreto, a finales de la década de 1970, la Organización Internacional de Estandarización (ISO, *International Organization for Standardization*) propuso que las redes de computadoras fueran organizadas utilizando siete capas, en lo que se vino a denominar modelo OSI (*Open Systems Interconnection*, Interconexión de sistemas abiertos) [ISO 2009]. El modelo OSI tomó forma cuando los protocolos que se convertirían en los protocolos de Internet estaban en su infancia y eran simplemente uno de los muchos conjuntos de protocolos diferentes que estaban en desarrollo; de hecho, probablemente los inventores del modelo OSI original no estaban pensando en Internet cuando lo crearon. No obstante, a partir de la década de 1970, muchos cursos universitarios y de formación, siguiendo las recomendaciones de ISO, organizaron cursos sobre el modelo de siete capas. A causa de su temprano impacto sobre la formación en redes, el modelo de siete capas todavía perdura en algunos libros de texto y cursos de formación sobre redes.

Las siete capas del modelo de referencia OSI, mostrado en la Figura 1.23(b), son: capa de aplicación, capa de presentación, capa de sesión, capa de transporte, capa de red, capa de enlace de datos y capa física. La funcionalidad de cinco de estas capas es básicamente la misma que sus contrapartidas del mismo nombre de Internet. Por tanto, vamos a centrarnos en las dos capas adicionales del modelo de referencia OSI: la capa de presentación y la capa de sesión. La función de la capa de presentación es la de proporcionar servicios que permitan a las aplicaciones que se comunican interpretar el significado de los datos intercambiados. Estos servicios incluyen la compresión y el cifrado de los datos (funciones cuyos nombres son autoexplicativos), así como la descripción de los datos (lo que, como veremos en el Capítulo 9, libera a la aplicación de tener que preocuparse por el formato interno en el que los datos se representan y almacenan, formatos que pueden diferir de una computadora a otra). La capa de sesión permite delimitar y sincronizar el intercambio de datos, incluyendo los medios para crear un punto de restauración y un esquema de recuperación.

El hecho de que en Internet falten dos de las capas existentes en el modelo de referencia OSI plantea un par de cuestiones interesantes: ¿acaso los servicios proporcionados por estas dos capas no son importantes? ¿Qué ocurre si una aplicación *necesita* uno de estos servicios? La respuesta de Internet a ambas preguntas es la misma: es problema del desarrollador de la aplicación. El desarrollador de la aplicación tiene que decidir si un servicio es importante y si lo es, será su problema el incorporar dicha funcionalidad a la aplicación.

1.5.2 Mensajes, segmentos, datagramas y tramas

La Figura 1.24 muestra la ruta física que siguen los datos al descender por la pila de protocolos de un sistema terminal emisor, al ascender y descender por las pilas de protocolos de un switch de la capa de enlace y de un router, para finalmente ascender por la pila de protocolos del sistema terminal receptor. Como veremos más adelante en el libro, los routers y los switches de la capa de enlace operan como dispositivos de conmutación de paquetes. De forma similar a los sistemas terminales, los routers y los switches de la capa de enlace organizan su hardware y software de red en capas. Pero estos dispositivos no implementan todas las capas de la pila de protocolos; habitualmente sólo implementan las capas inferiores. Como se muestra en la Figura 1.24, los switches de la capa de enlace implementan las capas 1 y 2; y los routers implementan las capas 1 a 3. Esto significa, por ejemplo, que los routers de Internet son capaces de implementar el protocolo IP (un protocolo de la capa 3) y los switches de la capa de enlace no. Veremos más adelante que aunque los switches de la capa de enlace no reconocen las direcciones IP, pueden reconocer las direcciones de la capa 2, como por ejemplo las direcciones Ethernet. Observe que los hosts implementan las cinco capas, lo que es coherente con la idea de que la arquitectura de Internet es mucho más compleja en las fronteras de la red.

La Figura 1.24 también ilustra el importante concepto de **encapsulación**. En el host emisor, un **mensaje de la capa de aplicación** (M en la Figura 1.24) se pasa a la capa de transporte. En el caso más simple, la capa de transporte recibe el mensaje y añade información

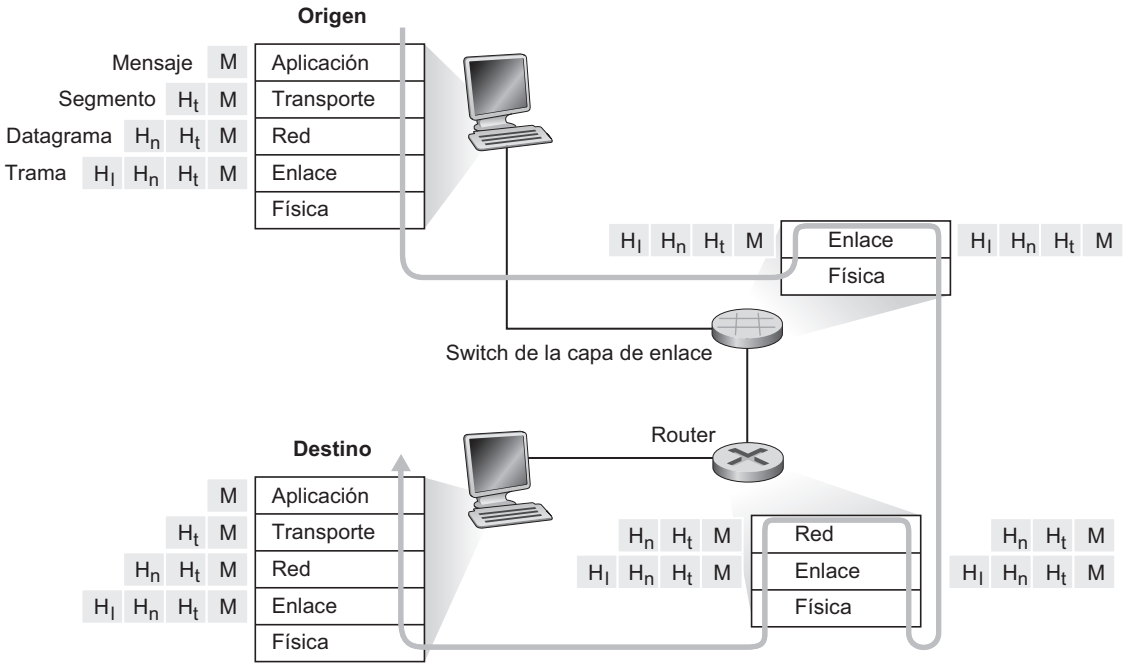


Figura 1.24 • Hosts, routers y switches de la capa de enlace. Cada uno de ellos contiene un conjunto distinto de capas, lo que refleja sus distintas funcionalidades.

adicional (denominada información de cabecera de la capa de transporte, H_t en la Figura 1.24) que será utilizada por la capa de transporte del lado receptor. El mensaje de la capa de aplicación y la información de cabecera de la capa de transporte constituyen el **segmento de la capa de transporte**. El segmento de la capa de transporte encapsula el mensaje de la capa de aplicación. La información añadida debe incluir información que permita a la capa de transporte del lado receptor entregar el mensaje a la aplicación apropiada y los bits de detección de errores que permitan al receptor determinar si los bits del mensaje han cambiado a lo largo de la ruta. A continuación, la capa de transporte pasa el segmento a la capa de red, que añade información de cabecera de la capa de red (H_n en la Figura 1.24) como son las direcciones de los sistemas terminales de origen y de destino, creando un **datagrama de la capa de red**. Este datagrama se pasa entonces a la capa de enlace, que (¡por supuesto!) añadirá su propia información de cabecera dando lugar a una **trama de la capa de enlace**. Así, vemos que en cada capa, un paquete está formado por dos tipos de campos: los campos de cabecera y un **campo de carga útil**. Normalmente, la carga útil es un paquete de la capa superior.

Una buena analogía para ilustrar este tema sería el envío de un informe interno de empresa desde una sucursal a otra a través del servicio postal público. Supongamos que Alicia, que se encuentra en una sucursal, quiere enviar un informe a Benito, que se encuentra en la otra sucursal. El *informe* es análogo al *mensaje de la capa de aplicación*. Alicia introduce el informe en un sobre para correo interno de la empresa y escribe en él el nombre y el número de departamento de Benito. El *sobre para correo interno* es análogo a un *segmento de la capa de transporte* (contiene la información de cabecera, el nombre y el departamento de Benito) y encapsula el mensaje de la capa de aplicación (el informe). Cuando en la sala de correo de la sucursal emisora se recibe este sobre, lo meten en otro sobre adecuado para enviar el informe mediante el servicio público de correos. En la sala de correo de la empresa también se escriben en el segundo sobre las direcciones postales tanto de la sucursal emisora como de la sucursal receptora. Aquí, el *sobre postal* es análogo al *datagrama*, encapsula el segmento de la capa de transporte (el sobre para correo interno), que encapsula el mensaje original (el informe). El servicio postal entrega el sobre postal a la sala de correos de la sucursal receptora, donde comienza el proceso de desencapsulación. En esta sala se extrae el informe y se envía a Benito. Por último, Benito abre el sobre para correo interno y saca el informe.

El proceso de encapsulación puede ser más complejo que el que acabamos de describir. Por ejemplo, un mensaje largo puede dividirse en varios segmentos de la capa de transporte (los cuales a su vez pueden dividirse en varios datagramas de la capa de red). En el extremo receptor, cada segmento tiene entonces que ser reconstruido a partir de sus datagramas constituyentes.

1.6 Ataques a las redes

Internet se ha convertido en una herramienta crítica para muchas instituciones actuales, incluyendo empresas pequeñas y medianas, universidades y organismos gubernamentales. Muchas personas individuales también confían en Internet para llevar a cabo muchas de sus actividades profesionales, sociales y personales. Pero detrás de todas estas utilidades y emociones, hay un lado oscuro, un lado donde los “chicos malos” intentan hacer estragos en nuestras vidas diarias dañando nuestras computadoras conectadas a Internet, violando nues-

tra privacidad y volviendo inoperables los servicios de Internet de los que dependemos [Skoudis 2006].

El campo de la seguridad de red se ocupa de ver cómo los "chicos malos" pueden atacar a las redes de computadoras y de cómo nosotros, que pronto seremos expertos en redes, podemos defendernos frente a estos ataques, o mejor todavía, de cómo diseñar nuevas arquitecturas que sean inmunes a tales ataques. Dada la frecuencia y variedad de ataques existentes, así como la amenaza de nuevos y más destructivos ataques futuros, la seguridad de red se ha convertido en un tema principal en el campo de las redes de comunicaciones en los últimos años. Una de las características de este libro de texto es que lleva las cuestiones sobre la seguridad en las redes a primer plano. En esta sección comenzaremos nuestra incursión en el campo de la seguridad de red, describiendo brevemente algunos de los ataques actuales más dañinos y predominantes en Internet. A continuación, en los siguientes capítulos nos ocuparemos en detalle de las tecnologías y los protocolos de red y consideraremos los diversos problemas relacionados con la seguridad, asociados con dichas tecnologías y protocolos. Por último, en el Capítulo 8, armados con nuestra experiencia recién adquirida en las redes de computadoras y los protocolos de Internet, estudiaremos en profundidad cómo las redes se pueden defender frente a los ataques, o diseñarse y operar para hacer que esos ataques sean imposibles en primera instancia.

Puesto que todavía no tenemos experiencia ni en redes ni en los protocolos de Internet, comenzaremos haciendo un repaso de algunos de los problemas de seguridad que predominan en la actualidad, con el fin de abrir boca para las explicaciones más sustanciosas que proporcionaremos en los capítulos siguientes. Así que podemos preguntarnos, ¿qué es lo que no funciona? ¿En qué sentido son vulnerables las redes de computadoras? ¿Cuáles son los principales tipos de ataques hoy día?

Los "malos" pueden introducir software malicioso en su host a través de Internet

Conectamos nuestros dispositivos a Internet porque deseamos recibir y enviar datos a Internet, lo que incluye todo tipo de cosas, páginas web, mensajes de correo electrónico, archivos MP3, llamadas telefónicas, vídeos en directo, resultados de motores de búsqueda, etc. Pero, lamentablemente, junto con todos estos elementos beneficiosos también existen elementos maliciosos, lo que se conoce de forma colectiva como **software malicioso** o **malware**, que puede acceder a nuestros dispositivos e infectarlos. Una vez que el malware ha infectado un dispositivo puede hacer todo tipo de maldades, como borrar nuestros archivos, instalar software espía que recopile nuestra información personal, como el número de la seguridad social, contraseñas y pulsaciones de teclas y luego enviar estos datos (a través de Internet, por supuesto) a los "chicos malos", a los atacantes. Nuestro host comprometido también puede pertenecer a una red de miles de dispositivos comprometidos de forma similar, lo que se conoce de forma colectiva como **botnet** (red robot), que los atacantes controlan y aprovechan para la distribución de correo electrónico basura (*spam*) o para llevar a cabo ataques distribuidos de denegación de servicio (que pronto explicaremos) contra los hosts objetivo.

Gran parte del malware que existe actualmente es **auto-replicante**: una vez que infecta un host, busca cómo acceder desde dicho host a otros hosts a través de Internet, y de nuevo desde esos hosts que acaba de infectar, busca cómo acceder a otros. De esta forma, el malware auto-replicante puede extenderse rápidamente de forma exponencial. Por ejemplo, el número de dispositivos infectados por el gusano 2003 Saphire/Slammer se replicaba cada

8,5 segundos en los primeros minutos después del brote, infectando más del 90 por ciento de los hosts vulnerables en 10 minutos [Moore 2003]. El malware puede extenderse en forma de virus, de un gusano o de un caballo de Troya [Skoudis 2004]. Un **virus** es un software malicioso que requiere cierta interacción del usuario para infectar el dispositivo. El ejemplo clásico es un adjunto de correo electrónico que contiene código ejecutable malicioso. Si un usuario recibe y abre un adjunto de este tipo, el usuario inadvertidamente ejecutará el malware en el dispositivo. Normalmente, tales virus enviados en los mensajes de correo electrónico se replican a sí mismos: una vez que se ha ejecutado, el virus puede enviar un mensaje idéntico con el mismo adjunto malicioso a, por ejemplo, todos los contactos de su libreta de direcciones. Un **gusano** (como el gusano Slammer) es malware que puede entrar en un dispositivo sin que el usuario interaccione de forma explícita. Por ejemplo, un usuario puede estar ejecutando una aplicación de red vulnerable a la que un atacante puede enviar software malicioso. En algunos casos, sin que el usuario intervenga, la aplicación puede aceptar el malware de Internet y ejecutarlo, creando un gusano. El gusano instalado ahora en el dispositivo recién infectado explora entonces Internet, buscando otros hosts que ejecuten la misma aplicación de red vulnerable. Cuando encuentra otros hosts vulnerables, envía una copia de sí mismo a esos hosts. Por último, un **caballo de Troya** es un malware que está oculto dentro de otro software que es útil. Hoy día, el malware está generalizado y es difícil defenderse de él. A lo largo del libro, le animaremos a que piense en la siguiente cuestión: ¿qué pueden hacer los diseñadores de redes de computadoras para defender a los dispositivos conectados a Internet de los ataques de malware?

Los "malos" pueden atacar a los servidores y a la infraestructura de red

Muchas de las amenazas de seguridad pueden clasificarse como **ataques de denegación de servicio (DoS, *Denial-of-Service*)**. Como su nombre sugiere, un ataque DoS vuelve inutilizable una red, un host o cualquier otro elemento de la infraestructura para los usuarios legítimos. Los servidores web, los servidores de correo electrónico, los servidores DNS (que se estudian en el Capítulo 2) y las redes institucionales pueden ser todos ellos objeto de ataques DoS. Los ataques DoS de Internet son muy comunes, teniendo lugar miles de ataques de este tipo cada año [Moore 2001; Mirkovic 2005]. La mayoría de los ataques DoS de Internet pueden clasificarse dentro de una de las tres categorías siguientes:

- *Ataque de vulnerabilidad.* Este ataque implica el envío de unos pocos mensajes bien construidos a una aplicación o sistema operativo vulnerable que esté ejecutándose en un host objetivo. Si se envía la secuencia de paquetes correcta a una aplicación o un sistema operativo vulnerable, el servicio puede detenerse o, lo que es peor, el host puede sufrir un fallo catastrófico.
- *Inundación del ancho de banda.* El atacante envía una gran cantidad de paquetes al host objetivo, de modo que comienzan a inundar el enlace de acceso del objetivo, impidiendo que los paquetes legítimos puedan alcanzar al servidor.
- *Inundación de conexiones.* El atacante establece un gran número de conexiones TCP completamente abiertas o semi-abiertas (estas conexiones se estudian en el Capítulo 3) en el host objetivo. El host puede comenzar a atascarse con estas conexiones fraudulentas impidiéndose así que acepte las conexiones legítimas.

Vamos a ver a continuación el ataque por inundación del ancho de banda más detalladamente. Recuerde el análisis que hemos realizado en la Sección 1.4.2 sobre los retardos y la

pérdida de paquetes; es evidente que si el servidor tiene una velocidad de acceso de *R* bps, entonces el atacante tendrá que enviar el tráfico a una velocidad de aproximadamente *R* bps para causar daños. Si *R* es muy grande, es posible que un único origen de ataque no sea capaz de generar el tráfico suficiente como para dañar al servidor. Además, si todo el tráfico procede de un mismo origen, un router situado en un punto anterior de la ruta puede detectar el ataque y bloquear todo el tráfico procedente de ese origen antes de que llegue cerca del servidor. En un ataque DoS distribuido **(DDoS, *Distributed DoS*)**, como el mostrado en la Figura 1.25, el atacante controla varios orígenes y hace que cada uno de ellos bombardee el objetivo con tráfico. Con este método, la tasa acumulada de tráfico para todos los orígenes controlados tiene que ser aproximadamente igual a *R* para inutilizar el servicio. Actualmente, se producen de forma común ataques DDoS que utilizan botnets con miles de hosts comprometidos [Mirkovic 2005]. Los ataques DDoS son mucho más difíciles de detectar y es mucho más complicado defenderse de ellos que de los ataques DoS procedentes de un único host.

Le animamos a que piense en la siguiente pregunta según vaya leyendo el libro: ¿qué pueden hacer los diseñadores de redes de computadoras para defenderlas de los ataques DoS? Veremos que son necesarias diferentes defensas para cada uno de los tres tipos de ataques DoS.

Los "malos" pueden examinar y analizar los paquetes

Actualmente, muchos usuarios acceden a Internet a través de dispositivos inalámbricos, tales como computadoras portátiles con conexión WiFi o dispositivos de mano con conexiones Internet móviles (lo que veremos en el Capítulo 6). Aunque el omnipresente acceso a Internet es extremadamente útil y habilita maravillosas aplicaciones para los usuarios móviles, también crea una importante vulnerabilidad de seguridad, al colocar un receptor pasivo en las vecindades del transmisor inalámbrico, que puede recibir una copia de todos los paquetes que se están transmitiendo. Estos paquetes pueden contener todo tipo de información confidencial, incluyendo contraseñas, números de la seguridad social, secretos comerciales

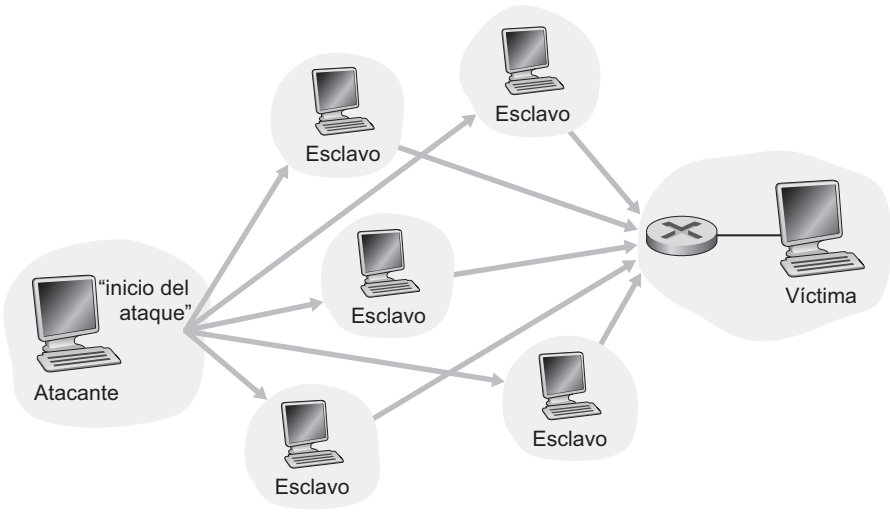


Figura 1.25 • Ataque de denegación de servicio distribuido.

y mensajes personales privados. Un receptor pasivo que registra una copia de todos los paquetes que pasan por él se conoce como **packet sniffer (**husmeador de paquetes).

Los sniffers también pueden implantarse en entornos cableados. En entornos cableados de multidifusión, como en muchas redes LAN Ethernet, un sniffer puede obtener copias de todos los paquetes enviados a través de la LAN. Como se ha descrito en la Sección 1.2, las tecnologías de acceso por cable también difunden paquetes y por tanto son vulnerables a ser monitorizados y analizados. Además, un atacante que consigue acceder al enlace de acceso o al router de acceso de una institución puede colocar un sniffer que haga una copia de todos los paquetes entrantes y salientes de la organización. Los paquetes así monitorizados pueden ser analizados después para obtener la información confidencial.

El software sniffer está disponible de forma gratuita en varios sitios web y como productos comerciales. Los profesores que imparten cursos sobre redes suelen realizar prácticas de laboratorio que implican escribir un programa sniffer y un programa de reconstrucción de datos de la capa de aplicación. Por supuesto, las prácticas de laboratorio con Wireshark [Wireshark 2009] asociadas con este texto (véase la práctica de laboratorio introductoria de Wireshark al final del capítulo) utilizan exactamente un programa sniffer así para monitorizar y analizar paquetes.

Puesto que los programas sniffer son pasivos, es decir, no inyectan paquetes en el canal, son difíciles de detectar. Por tanto, cuando envíamos paquetes a un canal inalámbrico, tenemos que aceptar que existe la posibilidad de que algún atacante pueda registrar copias de nuestros paquetes. Como es posible que haya adivinado, una de las mejores formas de defenderse frente a los programas sniffer son las técnicas criptográficas. En el Capítulo 8 veremos cómo se aplica la criptografía a la seguridad de la red.

Los "malos" pueden suplantar identidades

Es sorprendentemente fácil (y el lector tendrá los conocimientos necesarios para hacerlo muy pronto a medida que vaya leyendo este texto) crear un paquete con una dirección de origen, un contenido de paquete y una dirección de destino arbitrarios y luego transmitir dicho paquete a Internet, que reenviará el paquete a su destino. Imagine que el receptor confiado (por ejemplo, un router de Internet) que recibe tal paquete, toma la dirección de origen (falsa) como buena y luego ejecuta algún comando integrado en el contenido del paquete (por ejemplo, modifica la tabla de reenvío). La capacidad para inyectar paquetes en Internet con una dirección de origen falsa se conoce como **suplantación IP** y es una de las muchas formas en las que un usuario puede hacerse pasar por otro.

Para resolver este problema, necesitaremos aplicar un medio de *autenticación en el punto terminal,* es decir, un mecanismo que nos permita determinar con seguridad si un mensaje tiene su origen donde creemos que lo tiene. De nuevo, animamos a los lectores a que, a medida que avanzan por los capítulos del libro, piensen en cómo pueden hacer esto las aplicaciones y protocolos de red. En el Capítulo 8 exploraremos los mecanismos de autenticación en el punto terminal.

Los "malos" pueden modificar o borrar los mensajes

Terminamos este breve repaso sobre los ataques de red describiendo los **ataques de interposición**. En este tipo de ataques, los atacantes se introducen en la ruta de comunicaciones existente entre dos entidades que han establecido una conexión. Sean por ejemplo estas entidades Alicia y Benito, que pueden ser personas reales o entidades de red, como por ejemplo dos routers o dos servidores de correo. El atacante podría ser por ejemplo un router del que

el atacante haya tomado el control en la ruta de comunicación, o un módulo software residente en uno de los hosts terminales en una capa inferior de la pila de protocolos. En los ataques de interposición, el atacante no sólo tiene la capacidad de examinar y analizar los paquetes que pasan entre Benito y Alicia, sino que también puede inyectar, modificar o borrar paquetes. En la jerga utilizada al hablar de la seguridad de las redes, se dice que un ataque de interposición puede comprometer la *integridad* de los datos enviados entre Alicia y Benito. Como veremos en el Capítulo 8, los mecanismos que proporcionan confidencialidad (protección frente al husmeo de los paquetes) y autenticación en el punto terminal (lo que permite al receptor verificar con certeza al originador del mensaje) no necesariamente proporcionan integridad de los datos. Por tanto, necesitaremos otras técnicas para proporcionar esta funcionalidad.

Para terminar con este sección, vale la pena comentar cuál es la razón de que Internet se haya convertido en un lugar inseguro. Básicamente, la respuesta es que Internet fue diseñada originalmente para ser insegura, ya que se basaba en el modelo de un "grupo de usuarios que confiaban entre sí conectados a una red transparente" [Blumenthal 2001], un modelo en el que (por definición) no había necesidad de pensar en la seguridad. Muchos aspectos de la arquitectura de Internet original reflejan esta idea de confianza mutua. Por ejemplo, la posibilidad de que un usuario envíe un paquete a cualquier otro usuario es la opción predeterminada, en lugar de ser una capacidad solicitada/concedida, al igual que lo normal es creer que la identidad del usuario es la que declara, en lugar de autenticarle por defecto.

Pero, actualmente Internet no implica realmente "usuarios de confianza mutua". Sin embargo, los usuarios de hoy día necesitan comunicarse aunque no necesariamente confíen entre sí, pueden desear comunicarse de forma anónima, pueden comunicarse indirectamente a través de terceros (por ejemplo, cachés web, que estudiaremos en el Capítulo 2, o asistentes de movilidad, que veremos en el Capítulo 6), y deben desconfiar del hardware, el software e incluso del aire a través del que se comunican. A lo largo del libro vamos a encontrarnos con muchos retos relacionados con la seguridad, buscaremos formas de defendernos frente a los sniffer, la suplantación de identidades en el punto terminal, los ataques de interposición, los ataques DDoS, el software malicioso, etc. Tenemos que tener presente que la comunicación entre usuarios de mutua confianza es la excepción más que la regla. ¡Bienvenido al mundo de las redes modernas de comunicaciones!

1.7 Historia de Internet y de las redes de computadoras

En las Secciones 1.1 a 1.6 se ha hecho una presentación de las tecnologías utilizadas en las redes de comunicaciones e Internet. Ahora ya sabe lo suficiente como para impresionar a sus familiares y amigos. Sin embargo, si realmente desea causar una gran impresión en la siguiente fiesta a la que asista, debería salpicar su discurso con algunos detalles interesantes acerca de la fascinante historia de Internet [Segaller 1998].

1.7.1 El desarrollo de la conmutación de paquetes: 1961-1972

Tanto Internet como las redes de computadoras de hoy día tienen sus inicios a principios de la década de 1960, cuando la red telefónica era la red de comunicaciones dominante

en el mundo. Recordemos de la Sección 1.3 que la red telefónica utiliza mecanismos de conmutación de circuitos para transmitir la información entre un emisor y un receptor, una opción apropiada en la que la voz se transmite a una velocidad constante entre el emisor y el receptor. Debido a la creciente importancia (y los enormes costes) de las computadoras en los primeros años de la década de 1960 y a la aparición de las computadoras de tiempo compartido, quizá fue natural (¡al menos retrospectivamente!) considerar la cuestión de cómo enlazar las computadoras con el fin de que pudieran ser compartidas entre usuarios geográficamente distribuidos. El tráfico generado por esos usuarios probablemente era a *ráfagas,* compuesto por periodos de actividad como el envío de un comando a una computadora remota, seguido de periodos de inactividad mientras se espera a obtener una respuesta o mientras se contempla la respuesta recibida.

Tres grupos de investigación repartidos por el mundo, cada uno de ellos ignorante de la existencia de los otros [Leiner 1998], comenzaron a trabajar en la conmutación de paquetes como en una alternativa eficiente y robusta a la conmutación de circuitos. El primer trabajo publicado sobre las técnicas de conmutación de paquetes fue el de Leonard Kleinrock [Kleinrock 1961; Kleinrock 1964], un estudiante graduado en el MIT. Basándose en la teoría de colas, el trabajo de Kleinrock demostraba de forma elegante la efectividad de la técnica de conmutación de paquetes para las fuentes que generaban tráfico a ráfagas. En 1964, Paul Baran [Baran 1964] en el Rand Institute había comenzado a investigar el uso de la conmutación de paquetes para las comunicaciones de voz seguras en redes militares y, en el National Physical Laboratory (NPL) de Inglaterra, Donald Davies y Roger Scantlebury también estaban desarrollando sus ideas acerca de la conmutación de paquetes.

Los trabajos realizados en el MIT, en el instituto Rand y en los laboratorios NPL establecieron las bases de la red Internet actual. Pero Internet también tiene una larga tradición de "hagámoslo y demostremos que funciona" que también se remonta a la década de 1960. J. C. R. Licklider [DEC 1990] y Lawrence Roberts, ambos colegas de Kleinrock en el MIT, dirigieron el programa de Ciencias de la Computación en la Agencia de Proyectos de Investigación Avanzada (ARPA, *Advanced Research Projects Agency*) de Estados Unidos. Roberts publicó un plan global para la red ARPAnet [Roberts 1967], la primera red de computadoras de conmutación de paquetes y un ancestro directo de la red Internet pública actual. Los primeros conmutadores de paquetes se conocían como procesadores de mensajes de interfaz (**IMP, *Interface Message Processors***), y el contrato para construir estos conmutadores fue concedido a la empresa BBN. El Día del Trabajo de 1969, se instaló el primer IMP en UCLA bajo la supervisión de Kleinrock y poco después se instalaron tres IMP adicionales en el Instituto de Investigación de Stanford (SRI) en UC Santa Barbara y en la Universidad de Utah (Figura 1.26). Hacia finales de 1969 estaba disponible la red precursora de Internet, que estaba formada por cuatro nodos. Kleinrock recuerda la primera vez que utilizó la red para llevar a cabo un inicio de sesión remoto desde UCLA al SRI, consiguiendo que el sistema fallara [Kleinrock 2004].

Hacia 1972, la red ARPAnet había crecido aproximadamente hasta 15 nodos y la primera demostración pública fue realizada por Robert Kahn en 1972 en la *International Conference on Computer Communications*. Se completó el primer protocolo host a host entre sistemas terminales de ARPAnet, conocido como el protocolo de control de red (NCP, *Network Control Protocol*) [RFC 001]. Disponiendo de un protocolo terminal a terminal, ahora podían escribirse aplicaciones. Ray Tomlinson de BBN escribió el primer programa de correo electrónico en 1972.

Figura 1.26 • Uno de los primeros procesadores de mensajes de interfaz (IMP) y L. Kleinrock (Mark J. Terrill, AP/Wide World Photos).

1.7.2 Redes propietarias e interredes: 1972-1980

La red inicial ARPAnet era una única red cerrada. Para establecer una comunicación con un host de la red ARPAnet, había que estar realmente conectado a otro IMP de ARPAnet. A mediados de la década de 1970, comenzaron a surgir otras redes de conmutación de paquetes autónomas además de ARPAnet, entre las que se incluyen las siguientes:

- ALOHANet, una red de microondas que enlazaba universidades de las islas Hawai [Abramson 1970], así como redes de conmutación de paquetes vía satélite de DARPA [RFC 829] y redes de conmutación de paquetes vía radio [Kahn 1978].
- Telenet, una red de conmutación de paquetes comercial de BBN basada en la tecnología ARPAnet.
- Cyclades, una red de conmutación de paquetes francesa dirigida por Louis Pouzin [Think 2009].

- Redes de tiempo compartido tales como Tymnet y la red de Servicios de información GE, entre otras, a finales de la década de 1960 y principios de la década de 1970 [Schwartz 1977].
- La red SNA de IBM (1969-1974), que constituía un desarrollo paralelo al de ARPAnet [Schwartz 1977].

El número de redes fue creciendo. Retrospectivamente, podemos ver que había llegado el momento de desarrollar una arquitectura completa para la interconexión de redes. El trabajo pionero sobre interconexión de redes (realizado bajo el patrocinio de la Agencia DARPA, *Defense Advanced Research Projects Agency*), que en esencia es la creación de una *red de redes,* fue realizado por Vinton Cerf y Robert Kahn [Cerf 1974]; el término *internetting* (interredes o interconexión de redes) fue acuñado para describir este trabajo.

Estos principios arquitectónicos fueron integrados en TCP. Sin embargo, las primeras versiones de TCP eran bastante distintas al TCP de hoy día. Esas primeras versiones combinaban una entrega fiable en secuencia de los datos mediante retransmisiones del sistema terminal (esta función todavía la realiza TCP hoy día) con funciones de reenvío (que actualmente son realizadas por IP). Los primeros experimentos con TCP, combinados con el reconocimiento de la importancia de disponer de un servicio de transporte terminal a terminal no fiable y sin control de flujo para aplicaciones tales como voz empaquetada, llevaron a la separación de IP de TCP y al desarrollo del protocolo UDP. Los tres protocolos clave de Internet que se emplean actualmente (TCP, UDP e IP) fueron concebidos a finales de la década de 1970.

Además de las investigaciones relativas a Internet de la agencia DARPA, también se llevaron a cabo muchas otras importantes actividades de red. En Hawai, Norman Abramson desarrolló ALOHAnet, una red de paquetes vía radio que permitía a varios sitios remotos de las islas Hawai comunicarse entre sí. El protocolo ALOHA [Abramson 1970] fue el primer protocolo de acceso múltiple, que permitió a usuarios distribuidos geográficamente compartir un mismo medio de comunicación de difusión (una frecuencia de radio). Metcalfe y Boggs se basaron en el protocolo de acceso múltiple de Abramson para desarrollar el protocolo Ethernet [Metcalfe 1976] para redes de difusión compartidas basadas en cable; véase la Figura 1.27. Es interesante comentar que el protocolo Ethernet de Metcalfe y Boggs fue motivado por la necesidad de conectar varios PC, impresoras y discos compartidos [Perkins 1994]. Hace veinticinco años, bastante antes de la revolución de los PC y de la explosión de las redes, Metcalfe y Boggs establecieron las bases para las redes LAN de computadoras PC actuales. La tecnología Ethernet también representó un paso importante en la interconexión de redes (interred). Cada red de área local Ethernet era una red por sí misma, y dado que el número de redes LAN proliferaba, la necesidad de interconectar estas redes LAN adquiría cada vez más importancia. En el Capítulo 5 veremos en detalle Ethernet, ALOHA y otras tecnologías LAN.

1.7.3 Proliferación de las redes: 1980-1990

A finales de la década de 1970, había unos doscientos hosts conectados a la red ARPAnet. A finales de la década de 1980, el número de hosts conectados a la red Internet pública, una confederación de redes similar a la Internet actual, llegaría a los cien mil. La década de 1980 fue una época de enorme crecimiento.

Gran parte de este crecimiento fue el resultado de varios y distintos esfuerzos por crear redes de computadoras que enlazaran universidades. BITNET proporcionaba servicios de

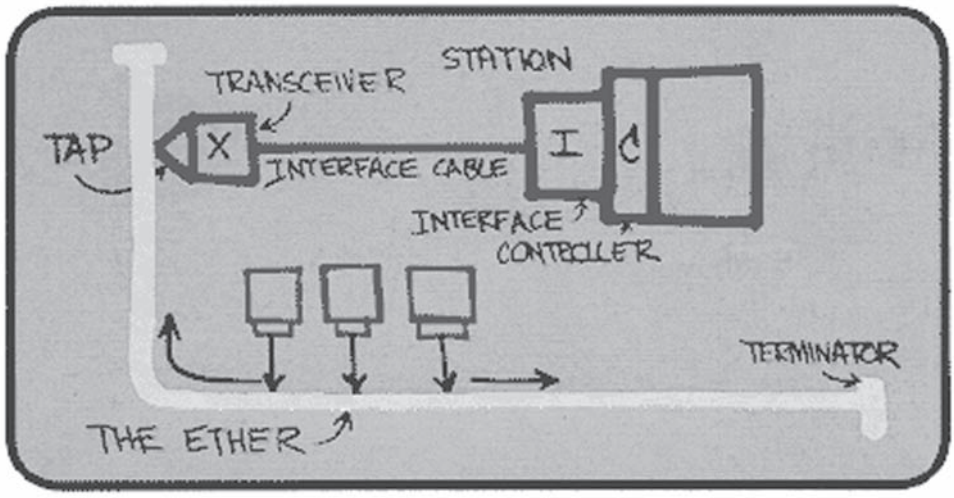


Figura 1.27 • Concepción original de Ethernet de Metcalfe.

correo electrónico y de transferencia de archivos entre varias universidades. CSNET (*Computer Science Network*) se formó para que investigadores universitarios que no tenían acceso a la red ARPAnet pudieran comunicarse. En 1986 se creó NSFNET para proporcionar acceso a los centros de supercomputación patrocinados por NSF. Inicialmente, con una velocidad en la red troncal de 56 kbps, la red troncal de NSFNET llegaría a operar a 1,5 Mbps a finales de la década y serviría como una red troncal primaria para enlazar redes regionales.

En la comunidad ARPAnet, muchas de las piezas finales de la arquitectura de Internet actual fueron encajando. El 1 de enero de 1983 se llevó a cabo el lanzamiento oficial de TCP/IP como el nuevo protocolo de host estándar para ARPAnet (reemplazando al protocolo NCP). La transición [RFC 801] de NCP a TCP/IP fue un suceso señalado: a todos los hosts se les requirió pasar a utilizar TCP/IP ese día. A finales de la década de 1980, se realizaron importantes extensiones en TCP con el fin de implementar el control de congestión basado en host [Jacobson 1988]. También se desarrolló el sistema DNS, que se emplea para establecer la correspondencia entre los nombres de Internet que usamos las personas (por ejemplo, gaia.cs.umass.edu) y sus direcciones IP de 32 bits [RFC 1034].

En paralelo con el desarrollo de ARPAnet (realizado fundamentalmente en Estados Unidos), a principios de la década de 1980, los franceses lanzaron el proyecto Minitel, un ambicioso plan para llevar las redes de datos a todos los hogares. Patrocinado por el gobierno francés, el sistema Minitel consistía en una red pública de conmutación de paquetes (basada en la serie de protocolos X.25), servidores Minitel y terminales económicos que incorporaban modems de baja velocidad. Minitel alcanzó un gran éxito en 1984 cuando el gobierno francés proporcionó un terminal Minitel gratuito a todo aquel que deseara tener uno en su casa. Los sitios Minitel incluían sitios gratuitos, como por ejemplo el directorio telefónico, y sitios privados, que cobraban sus tarifas basándose en la utilización del servicio. A mediados de la década de 1990, se produjo el pico de servicios ofertados ofreciendo más de 20.000 servicios, desde servicios de banca hasta bases de datos de investigación especializadas. Esta red era utilizada por más del 20 por ciento de la población de Francia,

generó más de 1.000 millones de dólares de ingresos anuales y dió lugar a la creación de 10.000 puestos de trabajo. Minitel se utilizó en muchos hogares de Francia 10 años antes de que los americanos ni siquiera hubieran oído hablar de Internet.

1.7.4 La explosión de Internet: década de 1990

La década de 1990 estuvo marcada por una serie de acontecimientos que simbolizaron la continua evolución y la pronta llegada de la comercialización de Internet. ARPAnet, el progenitor de Internet, dejó de existir. MILNET y la Red de Datos de Defensa habían crecido durante la década de 1980 transportando la mayor parte del tráfico relacionado con el Departamento de Defensa de Estados Unidos y NSFNET había empezado a servir como red troncal para conectar las redes regionales de Estados Unidos y las redes nacionales de ultramar. En 1991, NSFNET retiró sus restricciones sobre el uso de NSFNET para propósitos comerciales. La propia NSFNET fue eliminada del servicio activo en 1995, siendo el tráfico troncal de Internet transportado por los Proveedores de servicios de Internet comerciales.

Sin embargo, el principal acontecimiento de la década de 1990 fue la aparición de la World Wide Web, que llevaría Internet a los hogares y negocios de millones de personas de todo el mundo. La Web sirvió como plataforma para posibilitar e implantar cientos de nuevas aplicaciones, que hoy damos por sentadas. Si desea conocer una breve historia sobre la primera época de la Web, consulte [W3C 1995].

La Web fue inventada en el CERN por Tim Berners-Lee entre 1989 y 1991 [Berners-Lee 1989], basándose en las ideas de los primeros trabajos acerca del hipertexto desarrolladas en la década de 1940 por Vannevar Bush [Bush 1945] y luego en la década de 1960 por Ted Nelson [Xanadu 2009]. Berners-Lee y sus socios desarrollaron las versiones iniciales de HTML, HTTP, un servidor web y un navegador (los cuatro componentes clave de la Web). Hacia finales del año 1993 estaban operativos aproximadamente doscientos servidores web. Esta colección de servidores sólo sería un presagio de lo que estaba por venir. Al mismo tiempo, había varios investigadores desarrollando navegadores web con interfaces GUI, entre los que se encontraba Marc Andreessen, que lideró el desarrolló del popular navegador con interfaz GUI Mosaic. En 1994, Marc Andreessen y Jim Clark crearon Mosaic Communications, que más tarde se convertiría en Netscape Communications Corporation [Cusumano 1998; Quittner 1998]. Hacia 1995, los estudiantes universitarios empleaban los navegadores Mosaic y Netscape para navegar por la Web diariamente. También por esa época, las empresas, grandes y pequeñas, comenzaron a trabajar con servidores web y a realizar transacciones comerciales a través de la Web. En 1996, Microsoft empezó a desarrollar navegadores y comenzaría la guerra de los navegadores entre Netscape y Microsoft, que terminaría ganando Microsoft unos pocos años después [Cusumano 1998].

La segunda mitad de la década de 1990 fue un periodo de gran crecimiento e innovación para Internet, con las principales corporaciones y miles de empresas creando productos y servicios Internet. El correo electrónico a través de Internet continuó evolucionando con lectores de correo cada vez con más funcionalidades que proporcionaban libretas de direcciones, adjuntos, hipervínculos activos y transporte multimedia. Al final del milenio, Internet daba soporte a cientos de aplicaciones populares, entre las que se incluyen:

- Correo electrónico, incluyendo los adjuntos y el correo electrónico accesible a través de la Web.
- La Web, incluyendo la navegación web y el comercio por Internet.

- La mensajería instantánea, con listas de contactos, primeramente introducida por ICQ.
- La compartición igualitaria de archivos MP3, de la que Napster fue la pionera.

Es interesante saber que las dos primeras de esta aplicaciones estrella fueron creadas por la comunidad de investigadores, mientras que las dos últimas lo fueron por unos pocos jóvenes emprendedores.

El periodo comprendido entre 1995 y 2001 fue un paseo en montaña rusa para Internet en los mercados financieros. Antes de que fueran incluso rentables, cientos de nuevas empresas de Internet fueron vendidas en el mercado bursátil con oferta pública inicial. Muchas empresas fueron valoradas en miles de millones de dólares sin tener ingresos significativos. Las acciones de Internet se hundieron en 2000-2001 y muchas de esas empresas de nueva creación cerraron. No obstante, algunas de ellas emergieron como los grandes ganadores en el espacio Internet, entre las que se incluyen Microsoft, Cisco, Yahoo, e-Bay, Google y Amazon.

1.7.5 Desarrollos recientes

La innovación en las redes de computadoras continúa a buen paso. Se han hecho avances en todos los frentes, incluyendo la implementación de nuevas aplicaciones, distribución de contenidos, telefonía por Internet, velocidades de transmisión más altas en las redes LAN y routers más rápidos. Pero los tres desarrollos que merecen una atención especial son la proliferación de métodos de acceso de alta velocidad (incluido el acceso inalámbrico), la seguridad y las redes P2P.

Como se ha visto en la Sección 1.2, la creciente penetración del acceso a Internet residencial mediante módem por cable y DSL abre una etapa de abundancia para nuevas aplicaciones multimedia, incluyendo voz y vídeo sobre IP [Skype 2009], compartición de vídeos [YouTube 2009] y televisión sobre IP [PPLive 2009]. La creciente omnipresencia de las redes públicas Wi-Fi de alta velocidad (11 Mbps y superior) y el acceso a Internet a velocidades medias (cientos de kbps) mediante redes de telefonía móvil no hacen sólo posible estar constantemente conectado, sino que también habilitan un nuevo y excitante conjunto de servicios específicos de la ubicación. En el Capítulo 6 nos ocuparemos de las redes inalámbricas y de la movilidad.

Después de una serie de ataques de denegación de servicio en los servidores web más importantes a finales de la década de 1990 y de la proliferación de ataques mediante gusanos (como el gusano Blaster), la seguridad de las redes se convirtió en un tema extremadamente importante. Estos ataques dieron lugar al desarrollo de los sistemas de detección de intrusiones, que son capaces de advertir en una fase temprana de un ataque, y al uso de cortafuegos para filtrar el tráfico no deseado antes de que entre en la red. En el Capítulo 8 veremos unos cuantos temas importantes relacionados con la seguridad.

La última innovación de la que tomaremos nota son las redes P2P. Una aplicación de red P2P explota los recursos disponibles en las computadoras de los usuarios (almacenamiento, contenido, ciclos de CPU y presencia humana) y tiene una autonomía significativa con respecto a los servidores centrales. Normalmente, las computadoras de los usuarios (es decir, los pares) disponen de una conectividad intermitente. Son numerosas las historias de éxito P2P que han tenido lugar en los últimos años, entre las que se incluyen la compartición de archivos P2P (Napster, Kazaa, Gnutella, eDonkey, LimeWire, etc.), la distribución de

archivos (BitTorrent), Voz sobre IP (Skype) e IPTV (PPLive, ppStream). En el Capítulo 2 veremos muchas de estas aplicaciones P2P.

1.8 Resumen

En este capítulo hemos presentado una gran cantidad de material. Hemos hablado de las distintas piezas hardware y software que forman Internet en particular y las redes de computadoras en general. Hemos partido de la frontera de la red, fijándonos en los sistemas terminales y las aplicaciones, y en el servicio de transporte proporcionado a las aplicaciones que se ejecutan en los sistemas terminales. También hemos hablado de las tecnologías de la capa de enlace y de los medios físicos que pueden encontrarse normalmente en la red de acceso. A continuación, nos hemos adentrado en el interior de la red, en su núcleo, y hemos identificado los mecanismos de conmutación de paquetes y de conmutación de circuitos como los dos métodos básicos utilizados para el transporte de datos a través de una red de telecomunicaciones, y hemos examinado las fortalezas y las debilidades de cada método. También hemos examinado la estructura global de Internet, y hemos aprendido que es una red de redes. Hemos visto que la estructura jerárquica de Internet, formada por los ISP de nivel superior e inferior, ha permitido que Internet crezca hasta incluir miles de redes.

En la segunda parte de este capítulo de introducción, hemos abordado varios temas fundamentales del campo de las redes. En primer lugar, hemos estudiado las causas de los retardos, la tasa de transferencia y la pérdida de paquetes en una red de conmutación de paquetes. Hemos desarrollado modelos cuantitativos simples para determinar los retardos de transmisión, de propagación y de cola, así como para la tasa de transferencia; hemos hecho un uso exhaustivo de estos modelos de retardo en los problemas de repaso incluidos al final del capítulo. A continuación, hemos visto las capas de protocolos y los modelos de servicio, las principales claves arquitectónicas de las redes a las que se volverá a hacer referencia a lo largo del libro. Asimismo, hemos repasado los ataques de seguridad más habituales actualmente en Internet. Hemos terminado nuestra introducción con un breve repaso a la historia de las redes de computadoras. Este primer capítulo en sí mismo constituye un minicurso sobre redes de computadoras.

Por tanto, hemos cubierto una enorme cantidad de conceptos básicos en este primer capítulo. Si está un poco abrumado, no se preocupe, en los siguientes capítulos revisaremos todas estas ideas, viéndolas con mucho más detalle (lo que es una promesa, no una amenaza). En este punto, esperamos que termine el capítulo teniendo ya una idea sobre las piezas que forman una red, teniendo un conocimiento (que todavía deberá desarrollar) del vocabulario del campo de las redes de computadoras (no se asuste por tener que volver a consultar este capítulo) y habiendo incrementado su deseo por aprender más acerca de las redes. Ésta es la tarea que tenemos por delante durante el resto del libro.

Mapa de este libro

Antes de iniciar cualquier viaje, siempre debería echarse un vistazo a un mapa de carreteras para familiarizarse con las principales carreteras y desvíos con los que nos encontraremos más adelante. En el viaje en el que nos hemos embarcado nosotros, el destino final es conocer en profundidad el cómo, el qué y el por qué de las redes de computadoras, y nuestro mapa son los capítulos de este libro:

1. Redes de computadoras e Internet
2. La capa de aplicación
3. La capa de transporte
4. La capa de red
5. La capa de enlace y las redes de área local
6. Redes inalámbricas y móviles
7. Redes multimedia
8. Seguridad en las redes de computadoras
9. Gestión de redes

Los Capítulos 2 a 5 son los cuatro capítulos centrales del libro. Observará que están organizados según las cuatro capas superiores de la pila de protocolos de Internet de cinco capas, un capítulo por capa. Observará también que nuestro viaje va a comenzar por la parte superior de la pila de protocolos de Internet, es decir, por la capa de aplicación, y luego continuaremos nuestro trabajo descendiendo por la pila. El razonamiento de hacer este recorrido de arriba-abajo es porque una vez que conozcamos las aplicaciones, estaremos en condiciones de comprender los servicios de red necesarios para dar soporte a esas aplicaciones. Podremos así examinar las distintas formas en que tales servicios pueden ser implementados por una arquitectura de red. Ocuparnos de las aplicaciones en primer lugar nos va a proporcionar la motivación necesaria para abordar el resto del texto.

En la segunda mitad del libro, los Capítulos 6 a 9, se abordan cuatro temas enormemente importantes (y en cierto modo independientes) en las redes de comunicaciones modernas. En el Capítulo 6, examinaremos las redes inalámbricas y celulares, incluyendo las redes LAN inalámbricas (lo que incluye las tecnologías WiFi, WiMAX y Bluetooth), las redes de telefonía móvil (lo que incluye las redes GSM) y la movilidad (tanto en redes IP como GSM). En el Capítulo 7, "Redes multimedia", examinaremos aplicaciones de audio y de vídeo tales como la telefonía Internet, la videoconferencia y los flujos de información multimedia almacenada. También veremos cómo pueden diseñarse las redes de conmutación de paquetes para proporcionar una calidad de servicio coherente a las aplicaciones de audio y de vídeo. En el Capítulo 8, "Seguridad en las redes de computadoras", veremos en primer lugar los fundamentos de los mecanismos de cifrado y de seguridad de red y después examinaremos cómo está siendo aplicada la teoría básica a un amplio rango de contextos de Internet. El último capítulo, "Gestión de redes", examina los problemas fundamentales de la administración de red, así como los principales protocolos de Internet empleados para administrar la red.

Problemas y cuestiones de repaso

Capítulo 1 Cuestiones de repaso

SECCIÓN 1.1

R1. ¿Cuál es la diferencia entre un host y un sistema terminal? Enumere los tipos de sistemas terminales. ¿Es un servidor web un sistema terminal?

R2. El término *protocol*o a menudo se emplea para describir las relaciones diplomáticas. Proporcione un ejemplo de un protocolo diplomático.

SECCIÓN 1.2

R3. ¿Qué es un programa cliente? ¿Qué es un programa servidor? ¿Un programa servidor solicita y recibe servicios de un programa cliente?

R4. Enumere seis tecnologías de acceso. Clasifíquelas como de acceso residencial, acceso empresarial o acceso móvil.

R5. ¿La velocidad de transmisión en un sistema HFC es dedicada o compartida entre los usuarios? ¿Pueden producirse colisiones en un canal de descarga HFC? ¿Por qué?

R6. Enumere las tecnologías de acceso residencial disponibles en su ciudad. Para cada tipo de acceso, detalle la velocidad de descarga, la velocidad de carga y el precio mensual.

R7. ¿Cuál es la velocidad de transmisión en las redes LAN Ethernet? Para una determinada velocidad de transmisión, ¿pueden los usuarios de la LAN transmitir continuamente a dicha velocidad?

R8. Cite algunos de los medios físicos sobre los que se puede emplear la tecnología Ethernet.

R9. Para el acceso residencial se emplean los modems de acceso telefónico, los sistemas HFC, DSL y FTTH. Para cada una de estas tecnologías de acceso, detalle el rango de velocidades de transmisión e indique si la velocidad de transmisión es dedicada o compartida.

R10. Describa las tecnologías de acceso inalámbrico a Internet más populares hoy día.

SECCIÓN 1.3

R11. ¿Qué ventajas presenta una red de conmutación de circuitos frente a una red de conmutación de paquetes? ¿Qué desventajas tiene la multiplexación TDM frente a la multiplexación FDM en una red de conmutación de circuitos?

R12. ¿Por qué se dice que la conmutación de paquetes emplea multiplexación estadística? Compare la multiplexación estadística con la multiplexación por división en el tiempo (TDM).

R13. Suponga que hay un único dispositivo de conmutación de paquetes entre un host emisor y un host receptor. Las velocidades de transmisión entre el host emisor y el dispositivo de conmutación (switch) y entre el switch y el host receptor son R_1 y R_2, respectivamente. Suponiendo que el switch utiliza el mecanismo de conmutación de paquetes de almacenamiento y reenvío, ¿cuál es el retardo total terminal a terminal si se envía un paquete de longitud L? (Ignore los retardos de cola, de propagación y de procesamiento.)

R14. ¿Cuál es la diferencia entre un ISP de nivel 1 y un ISP de nivel 2?

R15. Suponga que los usuarios comparten un enlace de 2 Mbps y que cada usuario transmite a una velocidad de 1 Mbps continuamente, pero sólo durante el 20 por ciento del tiempo. (Véase la explicación sobre la multiplexación estadística de la Sección 1.3.)

a. Si se utiliza la conmutación de circuitos, ¿a cuántos usuarios puede darse soporte?

b. Para el resto del problema, suponga que se utiliza la conmutación de paquetes. ¿Por qué prácticamente no habrá retardo de cola antes del enlace si dos o menos usuarios transmiten a un mismo tiempo? ¿Por qué existirá retardo de cola si tres usuarios transmiten simultáneamente?

c. Calcule la probabilidad de que un usuario dado esté transmitiendo.

d. Suponga ahora que hay tres usuarios. Calcule la probabilidad de que en un instante determinado los tres usuarios estén transmitiendo simultáneamente. Halle la fracción de tiempo durante la que la cola crece.

SECCIÓN 1.4

R16. Considere el envío de un paquete desde un host emisor a un host receptor a través de una ruta fija. Enumere los componentes del retardo terminal a terminal. ¿Cuáles de estos retardos son constantes y cuáles son variables?

R17. Visite el applet *Transmission Versus Propagation Delay* (transmisión frente a retardo de propagación) disponible en el sitio web del libro. Utilizando las velocidades, retardos de propagación y tamaños de paquete disponibles, determine una combinación para la cual el emisor termine la operación de transmisión antes de que el primer bit del paquete haya llegado al receptor. Halle otra combinación para la que el primer bit del paquete haga llegado al receptor antes de que el emisor haya terminado de transmitir.

R18. ¿Cuánto tiempo tarda un paquete cuya longitud es de 1.000 bytes en propagarse a través de un enlace a una distancia de 2.500 km, siendo la velocidad de propagación igual a $2{,}5 \cdot 10^8$ m/s y la velocidad de transmisión a 2 Mbps? De forma más general, ¿cuánto tiempo tarda un paquete de longitud L en propagarse a través de un enlace a una distancia d, con una velocidad de propagación s y una velocidad de transmisión de R bps? ¿Depende este retardo de la longitud del paquete? ¿Depende este retardo de la velocidad de transmisión?

R19. Suponga que el host A desea enviar un archivo de gran tamaño al host B. La ruta desde el host A al host B está formada por tres enlaces, cuyas velocidades son $R_1 = 500$ kbps, $R_2 = 2$ Mbps y $R_3 = 1$ Mbps.

a. Suponiendo que no hay tráfico en la red, ¿cuál es la tasa de transferencia para el archivo?

b. Suponga que el tamaño del archivo es de 4 millones de bytes. Dividiendo el tamaño del archivo entre la tasa de transferencia, ¿cuánto tiempo tardará aproximadamente en transferirse el archivo al host B?

c. Repita los apartados (a) y (b), pero ahora con R_2 igual a 100 kbps.

R20. Suponga que el sistema terminal A desea enviar un archivo de gran tamaño al sistema terminal B. Sin entrar en detalles, describa cómo crea el sistema terminal A los paquetes a partir del archivo. Cuando uno de estos paquetes llega a un conmutador de paquetes, ¿qué información del mismo utiliza el conmutador para determinar el enlace por el que debe ser reenviado el paquete? ¿Por qué la conmutación de paquetes en Internet es análoga a viajar de una ciudad a otra preguntando por la dirección a la que nos dirigimos?

R21. Visite el applet *Queuing and Loss* (colas y pérdida de paquetes) en el sitio web del libro. ¿Cuáles son las velocidades de transmisión máxima y mínima? Para esas velocidades, ¿cuál es la intensidad de tráfico? Ejecute el applet con esas velocidades y determine cuánto tiempo tiene que transcurrir para que tenga lugar una pérdida de paquetes. A continuación, repita el experimento una segunda vez y determine de nuevo cuánto

tiempo pasa hasta producirse una pérdida de paquetes. ¿Son diferentes los valores obtenidos? ¿Por qué?

SECCIÓN 1.5

R22. Enumere cinco tareas que puede realizar una capa. ¿Es posible que una (o más) de estas tareas pudieran ser realizadas por dos (o más) capas?

R23. ¿Cuáles son las cinco capas de la pila de protocolos Internet? ¿Cuáles son las responsabilidades principales de cada una de estas capas?

R24. ¿Qué es un mensaje de la capa de aplicación? ¿Y un segmento de la capa de transporte? ¿Y un datagrama de la capa de red? ¿Y una trama de la capa de enlace?

R25. ¿Qué capas de la pila de protocolos de Internet procesa un router? ¿Qué capas procesa un switch de la capa de enlace? ¿Qué capas procesa un host?

SECCIÓN 1.6

R26. ¿Cuál es la diferencia entre un virus, un gusano y un caballo de Troya?

R27. Describa cómo puede crearse una red robot (botnet) y cómo se puede utilizar en un ataque DDoS.

R28. Suponga que Alicia y Benito están enviándose paquetes entre sí a través de una red. Imagine que Victoria se introduce en la red de modo que puede capturar todos los paquetes enviados por Alicia y que luego envía lo que ella quiere a Benito. Además, también puede capturar todos los paquetes enviados por Benito y luego enviar a Alicia lo que le parezca. Enumere algunos de los daños que Victoria puede ocasionar desde su posición.

Problemas

P1. Diseñe y describa un protocolo de nivel de aplicación que será utilizado entre un cajero automático y la computadora central de un banco. El protocolo deberá permitir verificar la tarjeta y la contraseña del usuario, consultar el saldo de la cuenta (que se almacena en la computadora central) y hacer un apunte en la cuenta por la cantidad retirada por el usuario. Las entidades del protocolo deben poder controlar todos los casos en los que no hay suficiente saldo en la cuenta como para cubrir el reembolso. Especifique el protocolo enumerando los mensajes intercambiados y las acciones realizadas por el cajero automático o la computadora central del banco al transmitir y recibir mensajes. Haga un boceto del funcionamiento del protocolo para el caso de una retirada de efectivo sin errores, utilizando un diagrama similar al mostrado en la Figura 1.2. Establezca explícitamente las suposiciones hechas por el protocolo acerca del servicio de transporte terminal a terminal subyacente.

P2. Considere una aplicación que transmite datos a una velocidad constante (por ejemplo, el emisor genera una unidad de datos de N bits cada k unidades de tiempo, donde k es un valor pequeño y fijo). Además, cuando esta aplicación se inicia, se ejecutará durante un periodo de tiempo relativamente largo. Responda a las siguientes cuestiones de forma breve y justificando su respuesta:

a. ¿Qué sería más apropiado para esta aplicación, una red de conmutación de circuitos o una red de conmutación de paquetes? ¿Por qué?

b. Suponga que se utiliza una red de conmutación de paquetes y el único tráfico que existe en la misma procede de la aplicación que acabamos de describir. Además, suponga que la suma de las velocidades de datos de la aplicación es menor que las capacidades de cada uno de los enlaces. ¿Será necesario algún mecanismo de control de congestión? ¿Por qué?

P3. Considere la red de conmutación de circuitos de la Figura 1.12. Recuerde que hay n circuitos en cada enlace.

a. ¿Cuál es el número máximo de conexiones simultáneas que pueden estar en curso en un determinado instante de tiempo en esta red?

b. Suponga que todas las conexiones se encuentran entre el dispositivo de conmutación de la esquina superior izquierda y el dispositivo de conmutación de la esquina inferior derecha. ¿Cuál será el número máximo de conexiones que puede haber en curso?

P4. Repase la analogía de la caravana de coches de la Sección 1.4. Suponga una velocidad de propagación de 100 km/hora.

a. Suponga que la caravana se mueve a una velocidad de 150 km, partiendo de la caseta de peaje, pasando por una segunda caseta de peaje y deteniéndose justo después de la tercera caseta de peaje. ¿Cuál es el retardo terminal a terminal?

b. Repita el apartado (a) suponiendo ahora que en la caravana hay ocho coches en lugar de diez.

P5. En este problema se exploran los retardos de propagación y de transmisión, dos conceptos fundamentales en las redes de datos. Considere dos hosts, A y B, conectados por un enlace cuya velocidad es de R bps. Suponga que los dos hosts están separados m metros y que la velocidad de propagación a lo largo del enlace es igual a s metros/segundo. El host A envía un paquete de tamaño L bits al host B.

a. Exprese el retardo de propagación, d_{prop}, en función de m y s.

b. Determine el tiempo de transmisión del paquete, d_{trans}, en función de L y R.

c. Ignorando los retardos de procesamiento y de cola, obtenga una expresión para el retardo terminal a terminal.

d. Suponga que el host A comienza a transmitir el paquete en el instante $t = 0$. En el instante $t = d_{\text{trans}}$, ¿dónde estará el último bit del paquete?

e. Suponga que d_{prop} es mayor que d_{trans}. En el instante $t = d_{\text{trans}}$, ¿dónde estará el primer bit del paquete?

f. Suponga que d_{prop} es menor que d_{trans}. En el instante $t = d_{\text{trans}}$, ¿dónde estará el primer bit del paquete?

g. Suponga que $s = 2{,}5 \cdot 10^8$ metros/segundo, $L = 120$ bits y $R = 56$ kbps. Determine la distancia m de modo que d_{prop} sea igual a d_{trans}.

P6. En este problema vamos a considerar la transmisión de voz en tiempo real desde el host A al host B a través de una red de conmutación de paquetes (VoIP). El host A convierte sobre la marcha la voz analógica en un flujo de bits digital a 64 kbps. A continuación, el host A agrupa los bits en paquetes de 56 bytes. Entre el host A y el host B existe un enlace, cuya velocidad de transmisión es de 2 Mbps y su retardo de propaga-

ción es igual a 10 milisegundos. Tan pronto como el host A forma un paquete, lo envía al host B. Cuando el host B recibe un paquete completo, convierte los bits del paquete en una señal analógica. ¿Cuánto tiempo transcurre desde el momento en que se crea un bit (a partir de la señal analógica en el host A) hasta que se decodifica (como parte de la señal analógica en el host B)?

P7. Suponga que varios usuarios comparten un enlace de 3 Mbps. Suponga también que cada usuario requiere 150 kbps para transmitir y que sólo transmite durante el 10 por ciento del tiempo. (Véase la explicación sobre la multiplexación estadística de la Sección 1.3.)

a. Si se utiliza la conmutación de circuitos, ¿a cuántos usuarios puede darse soporte?

b. Para el resto de este problema, suponga que se utiliza una red de conmutación de paquetes. Halle la probabilidad de que un determinado usuario esté transmitiendo.

c. Suponga que hay 120 usuarios. Determine la probabilidad de que en un instante determinado haya exactamente n usuarios transmitiendo simultáneamente. (*Sugerencia*: utilice la distribución binomial.)

d. Calcule la probabilidad de que haya 21 o más usuarios transmitiendo simultáneamente.

P8. Consulte la explicación acerca de la multiplexación estadística de la Sección 1.3, en la que se proporciona un ejemplo con un enlace a 1 Mbps. Los usuarios están generando datos a una velocidad de 100 kbps cuando están ocupados, pero sólo lo están con una probabilidad de $p = 0{,}1$. Suponga que el enlace a 1 Mbps se sustituye por un enlace a 1 Gbps.

a. ¿Cuál es el valor de N, el número máximo de usuarios a los que se les puede dar soporte simultáneamente, cuando se emplea una red de conmutación de circuitos?

b. Considere ahora que se utiliza una red conmutación de paquetes y que el número de usuarios es M. Proporcione una fórmula (en función de p, M, N) para determinar la probabilidad de que más de N usuarios estén enviando datos.

P9. Considere un paquete de longitud L que tiene su origen en el sistema terminal A y que viaja a través de tres enlaces hasta un sistema terminal de destino. Estos tres enlaces están conectados mediante dos dispositivos de conmutación de paquetes. Sean d_i, s_i y R_i la longitud, la velocidad de propagación y la velocidad de transmisión del enlace i, para $i = 1, 2, 3$. El dispositivo de conmutación de paquetes retarda cada paquete d_{proc}. Suponiendo que no se produce retardo de cola, ¿cuál es el retardo total terminal a terminal del paquete en función de d_i, s_i, R_i, ($i = 1, 2, 3$) y L? Suponga ahora que la longitud del paquete es de 1.500 bytes, la velocidad de propagación en ambos enlaces es igual a $2{,}5 \cdot 10^8$ m/s, la velocidad de transmisión en los tres enlaces es de 2 Mbps, el retardo de procesamiento en el conmutador de paquetes es de 3 milisegundos, la longitud del primer enlace es de 5.000 km, la del segundo de 4.000 km y la del último enlace es de 1.000 km. Para estos valores, ¿cuál es el retardo terminal a terminal?

P10. En el problema anterior, suponga que $R_1 = R_2 = R_3 = R$ y $d_{\text{proc}} = 0$. Suponga también que el conmutador de paquetes no almacena los paquetes y los reenvía, sino que transmite inmediatamente cada bit que recibe sin esperar a que llegue el paquete completo. ¿Cuál será el retardo terminal a terminal?

P11. Un conmutador de paquetes recibe un paquete y determina el enlace saliente por el que deberá ser reenviado. Cuando el paquete llega, hay otro paquete ya transmitido hasta la mitad por el mismo enlace de salida y además hay otros cuatro paquetes esperando para ser transmitidos. Los paquetes se transmiten según el orden de llegada. Suponga que todos los paquetes tienen una longitud de 1.500 bytes y que la velocidad del enlace es de 2 Mbps. ¿Cuál es el retardo de cola para el paquete? En sentido más general, ¿cuál es el retardo de cola cuando todos los paquetes tienen una longitud L, la velocidad de transmisión es R, x bits del paquete que se está transmitiendo actualmente ya han sido transmitidos y hay n paquetes en la cola esperando a ser transmitidos?

P12. Suponga que N paquetes llegan simultáneamente a un enlace en el que actualmente no se está transmitiendo ningún paquete ni tampoco hay ningún paquete en cola. Cada paquete tiene una longitud L y el enlace tiene una velocidad de transmisión R. ¿Cuál es el retardo medio de cola para los N paquetes?

P13. Considere el retardo de cola en el buffer de un router (que precede a un enlace de salida). Suponga que todos los paquetes tienen L bits, que la velocidad de transmisión es R bps y que llegan simultáneamente N paquetes al buffer cada LN/R segundos. Calcule el retardo medio de cola de un paquete. (*Sugerencia*: el retardo de cola para el primer paquete es igual a cero; para el segundo paquete es L/R; para el tercero es $2L/R$. El paquete N ya habrá sido transmitido cuando el segundo lote de paquetes llegue.)

P14. Considere el retardo de cola en el buffer de un router. Sea I la intensidad de tráfico; es decir, $I = La/R$. Suponga que el retardo de cola puede expresarse como $IL/R\ (1 - I)$ para $I < 1$.

a. Determine una fórmula para calcular el retardo total, es decir, el retardo de cola más el retardo de transmisión.

b. Dibuje el retardo total en función de L/R.

P15. Sea a la velocidad de llegada de los paquetes a un enlace en paquetes/segundo y sea μ la velocidad de transmisión del enlace en paquetes/segundo. Basándose en la fórmula del retardo total (es decir, el retardo de cola más el retardo de transmisión) obtenida en el problema anterior, deduzca una fórmula para el retardo total en función de a y μ.

P16. Considere el buffer de un router que precede a un enlace de salida. En este problema utilizaremos la fórmula de Little, una fórmula famosa en la teoría de colas. Sea N el número medio de paquetes que hay en el buffer más el paquete que está siendo transmitido. Sea a la velocidad a la que los paquetes llegan al enlace. Sea d el retardo medio total (es decir, el retardo de cola más el retardo de transmisión) experimentado por un paquete. La fórmula de Little es $N = a \cdot d$. Suponga que como media, el buffer contiene 10 paquetes y que el retardo medio de cola de un paquete es igual a 10 milisegundos. La velocidad de transmisión del enlace es igual a 100 paquetes/segundo. Utilizando la fórmula de Little, ¿cuál es la velocidad media de llegada de los paquetes suponiendo que no se produce pérdida de paquetes?

P17. a. Generalice la fórmula del retardo terminal a terminal dada en la Sección 1.4.3 para velocidades de procesamiento, velocidades de transmisión y retardos de propagación heterogéneos.

b. Repita el apartado (a), pero suponiendo ahora que existe un retardo medio de cola d_{cola} en cada nodo.

P18. Realice un trazado de la ruta (Traceroute) entre un origen y un destino situados en un mismo continente para tres horas del día diferentes.

a. Determine la media y la desviación estándar de los retardos de ida y vuelta para cada una de las horas.

b. Determine el número de routers existente en la ruta para cada una de las horas. ¿Ha variado la ruta para alguna de las horas?

c. Intente identificar el número de redes de ISP que los paquetes de Traceroute atraviesan desde el origen hasta el destino. Los routers con nombres similares y/o direcciones IP similares deben considerarse como parte del mismo ISP. En sus experimentos, ¿los retardos más largos se producen en las interfaces situadas entre proveedores ISP adyacentes?

d. Repita el apartado anterior para un origen y un destino situados en diferentes continentes. Compare los resultados para ubicaciones en un mismo continente y en distintos continentes.

P19. Considere el ejemplo sobre la tasa de transferencia correspondiente a la Figura 1.20(b). Suponga que hay M parejas cliente-servidor en lugar de 10. Sean R_s, R_c y R las velocidades de los enlaces de servidor, de los enlaces de cliente y del enlace de red. Suponga que todos los enlaces tienen la capacidad suficiente y que no existe otro tráfico en la red que el generado por las M parejas cliente-servidor. Deduzca una expresión general para la tasa de transferencia en función de R_s, R_c, R, y M.

P20. Considere la Figura 1.19(b). Suponga que existen M rutas entre el servidor y el cliente. No hay dos rutas que compartan ningún enlace. La ruta k ($k = 1, \ldots, M$) consta de N enlaces con velocidades de transmisión iguales a R_1^k, R_2^k, . . ., R_N^k. Si el servidor sólo puede utilizar una ruta para enviar datos al cliente, ¿cuál será la máxima tasa de transferencia que puede alcanzar dicho servidor? Si el servidor puede emplear las M rutas para enviar datos, ¿cuál será la máxima tasa de transferencia que puede alcanzar el servidor?

P21. Considere la Figura 1.19(b). Suponga que cada enlace entre el servidor y el cliente tiene una probabilidad de pérdida de paquetes p y que las probabilidades de pérdida de paquetes de estos enlaces son independientes. ¿Cuál es la probabilidad de que un paquete (enviado por el servidor) sea recibido correctamente por el receptor? Si un paquete se pierde en el camino que va desde el servidor hasta el cliente, entonces el servidor volverá a transmitir el paquete. Como media, ¿cuántas veces tendrá que retransmitir el paquete el servidor para que el cliente lo reciba correctamente?

P22. Considere la Figura 1.19(a). Suponga que sabemos que el enlace cuello de botella a lo largo de la ruta entre el servidor y el cliente es el primer enlace, cuya velocidad es R_s bits/segundo. Suponga que envíamos un par de paquetes uno tras otro desde el servidor al cliente y que no hay más tráfico que ese en la ruta. Suponga que cada paquete tiene un tamaño de L bits y que ambos enlaces presentan el mismo retardo de propagación d_{prop}.

a. ¿Cuál es el periodo entre llegadas de paquetes al destino? Es decir, ¿cuánto tiempo transcurre desde que el último bit del primer paquete llega hasta que lo hace el último bit del segundo paquete?

b. Suponga ahora que el enlace cuello de botella es el segundo enlace (es decir, $R_c < R_s$). ¿Es posible que el segundo paquete tenga que esperar en la cola de entrada del

segundo enlace? Explique su respuesta. Suponga ahora que el servidor envía el segundo paquete T segundos después de enviar el primero. ¿Qué valor debe tener T para garantizar que el segundo paquete no tenga que esperar en la cola de entrada del segundo enlace? Explique su respuesta.

P23. Suponga que necesita enviar de forma urgente 40 terabytes de datos de Boston a Los Ángeles. Dispone de un enlace dedicado a 100 Mbps para transferencia de datos. ¿Qué preferiría, transmitir los datos a través del enlace o utilizar FedEx para hacer el envío por la noche? Explique su respuesta.

P24. Se tienen dos hosts, A y B, separados 20.000 kilómetros y conectados mediante un enlace directo con $R = 2$ Mbps. Suponga que la velocidad de propagación por el enlace es igual a $2{,}5 \cdot 10^8$ metros/segundo.

a. Calcule el producto ancho de banda-retardo, $R \cdot d_{\text{prop}}$.

b. Se envía un archivo cuyo tamaño es de 800.000 bits desde el host A al host B. Suponga que el archivo se envía de forma continua como un mensaje de gran tamaño. ¿Cuál es el número máximo de bits que habrá en el enlace en un instante de tiempo determinado?

c. Haga una interpretación del producto ancho de banda-retardo.

d. ¿Cuál es el ancho (en metros) de un bit dentro del enlace? ¿Es más grande que un campo de fútbol?

e. Deduzca una expresión general para la anchura de un bit en función de la velocidad de propagación s, la velocidad de transmisión R y la longitud del enlace m.

P25. Continuando con el Problema P24, suponga que podemos modificar R. ¿Para qué valor de R es el ancho de un bit tan grande como la longitud del enlace?

P26. Considere el Problema P24 pero ahora para un enlace con $R = 1$ Gbps.

a. Calcule el producto ancho de banda-retardo, $R \cdot d_{\text{prop}}$.

b. Considere el envío de un archivo de 800.000 bits desde el host A al host B. Suponga que el archivo se envía de forma continua como si fuera un mensaje de gran tamaño. ¿Cuál es el número máximo de bits que puede haber en el enlace en cualquier instante de tiempo dado?

c. ¿Cuál es el ancho (en metros) de un bit dentro del enlace?

P27. Haciendo referencia de nuevo al problema P24.

a. ¿Cuánto tiempo tarda en enviarse el archivo suponiendo que se envía de forma continua?

b. Suponga ahora que el archivo se divide en 20 paquetes conteniendo cada uno de ellos 40.000 bits. Suponga también que el receptor confirma la recepción de cada paquete y que el tiempo de transmisión de un paquete de confirmación es despreciable. Por último, suponga que el emisor no puede transmitir un paquete hasta que el anterior haya sido confirmado. ¿Cuánto tiempo se tardará en enviar el archivo?

c. Compare los resultados obtenidos en los apartados (a) y (b).

P28. Suponga que existe un enlace de microondas a 10 Mbps entre un satélite geoestacionario y su estación base en la Tierra. El satélite toma una fotografía digital por minuto y la envía a la estación base. La velocidad de propagación es $2{,}4 \cdot 10^8$ metros/segundo.

a. ¿Cuál es el retardo de propagación del enlace?

b. ¿Cuál es el producto ancho de banda-retardo, $R \cdot d_{\text{prop}}$?

c. Sea x el tamaño de la fotografía. ¿Cuál es el valor mínimo de x para que el enlace de microondas esté transmitiendo continuamente?

P29. Considere la analogía de la compañía aérea utilizada en la Sección 1.5 dedicada a las capas y la adición de cabeceras a las unidades de datos del protocolo a medida que fluyen en sentido descendente por la pila de protocolos. ¿Existe algún concepto equivalente a la información de cabecera que pueda añadirse a los pasajeros y al equipaje a medida que descienden por la pila de protocolos de la compañía aérea?

P30. En las redes de conmutación de paquetes modernas, el host de origen segmenta los mensajes largos de la capa de aplicación (por ejemplo, una imagen o un archivo de música) en paquetes más pequeños y los envía a la red. Después, el receptor ensambla los paquetes para formar el paquete original. Este proceso se conoce como *segmentación de mensajes*. La Figura 1.28 ilustra el transporte terminal a terminal de un mensaje con y sin segmentación del mensaje. Imagine que se envía un mensaje cuya longitud es de $8 \cdot 10^6$ bits desde el origen hasta el destino mostrados en la Figura 1.28. Suponga que cada enlace de los mostrados en la figura son enlaces a 2 Mbps. Ignore los retardos de propagación, de cola y de procesamiento.

a. Suponga que el mensaje se transmite desde el origen al destino *sin* segmentarlo. ¿Cuánto tiempo tarda el mensaje en desplazarse desde el origen hasta el primer conmutador de paquetes? Teniendo en cuenta que cada conmutador de paquetes utiliza el método de conmutación de almacenamiento y reenvío, ¿cuál el tiempo total que invierte el mensaje para ir desde el host de origen hasta el host de destino?

b. Suponga ahora que el mensaje se segmenta en 4.000 paquetes y que la longitud de cada paquete es de 2.000 bits. ¿Cuánto tiempo tarda el primer paquete en transmitirse desde el origen hasta el primer conmutador de paquetes? Cuando se está enviando el primer paquete del primer conmutador al segundo, el host de origen envía un segundo paquete al primer conmutador de paquetes. ¿En qué instante de tiempo habrá recibido el primer conmutador el segundo paquete completo?

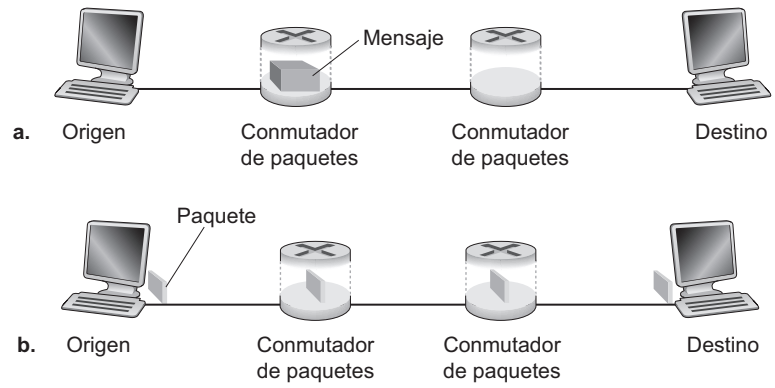


Figura 1.28 • Transporte de mensajes terminal a terminal: (a) sin segmentación de mensajes; (b) con segmentación de mensajes.

c. ¿Cuánto tiempo tarda en transmitirse el archivo desde el host de origen al host de destino cuando se emplea la segmentación de mensajes? Compare este resultado con la respuesta del apartado (a) y coméntelo.

d. Comente los inconvenientes de la segmentación de mensajes.

P31. Experimente con la applet *Message Segmentation* (segmentación de mensajes) disponible en el sitio web del libro. ¿Se corresponden los retardos indicados en el applet con los retardos del problema anterior? ¿Cómo afectan los retardos de propagación del enlace al retardo global terminal a terminal de la conmutación de paquetes (con segmentación de mensajes) y de la conmutación de mensajes?

P32. Se envía un archivo de gran tamaño de F bits desde el host A al host B. Entre los hosts A y B hay tres enlaces (y dos dispositivos de conmutación) y los enlaces no están congestionados (es decir, no existen retardos de cola). El host A divide el archivo en segmentos de S bits y añade 80 bits de cabecera a cada segmento, formando paquetes de $L = 80 + S$ bits. La velocidad de transmisión de cada enlace es de R bps. Calcule el valor de S que minimiza el retardo al transmitir el archivo desde el host A al host B. No tenga en cuenta el retardo de propagación.

Preguntas para la discusión

D1. ¿Qué tipos de servicios inalámbricos móviles hay disponibles en la región donde vive?

D2. Utilizando la tecnología LAN inalámbrica 802.11, diseñe una red doméstica para su casa o la de sus padres. Enumere los modelos específicos de los productos, así como los costes correspondientes de su red doméstica.

D3. Describa los servicios Skype de PC a PC. Pruebe el servicio de vídeo de Skype de PC a PC y redacte un informe narrando la experiencia.

D4. Skype ofrece un servicio que permite realizar una llamada telefónica desde un PC a un teléfono tradicional. Esto significa que la llamada de voz debe pasar por Internet y una red telefónica. Explique cómo puede hacerse esto.

D5. ¿Qué es un Servicio de mensajes cortos (SMS, *Short Message Service*)? ¿En qué países y continentes es popular este servicio? ¿Es posible enviar un mensaje SMS desde un sitio web a un teléfono móvil?

D6. ¿Qué es la transmisión de flujos de vídeo almacenado? ¿Cuáles son algunos de los sitios web más populares que suministran actualmente flujos de vídeo?

D7. ¿Qué son los flujos P2P de vídeo en directo? ¿Cuáles son algunos de los sitios web más populares que proporcionan actualmente este servicio?

D8. Localice cinco empresas que proporcionen servicios de compartición de archivos P2P. Para cada empresa, ¿con qué tipo de archivos trabajan, es decir, qué tipo de contenido gestionan?

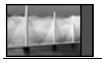
D9. ¿Quién inventó ICQ, es decir, el primer servicio de mensajería instantánea? ¿Cuándo fue inventado y cuál era la edad de sus inventores? ¿Quién inventó Napster, cuándo y qué edades tenían sus inventores?

D10. Compare el acceso a Internet inalámbrico WiFi y el acceso a Internet inalámbrico 3G. ¿Cuáles son las velocidades de bit de estos dos servicios? ¿Cuáles son los costes? Comente brevemente el concepto de itinerancia y ubicuidad de acceso.

D11. ¿Por qué ya no existe el servicio de compartición de archivos P2P de Napster original? ¿Qué es la RIAA y qué medidas se están tomando para limitar la compartición P2P de archivos con derechos de propiedad intelectual? ¿Cuál es la diferencia entre la infracción directa e indirecta de los derechos de propiedad intelectual?

D12. ¿Qué es BitTorrent? ¿Qué es lo que le hace fundamentalmente diferente de un servicio de compartición de archivos P2P como eDonkey, LimeWire o Kazaa?

D13. ¿Cree que dentro de 10 años seguirán compartiéndose de modo habitual archivos con derechos de propiedad intelectual a través de las redes de comunicaciones? ¿Por qué? Razone su respuesta.



Prácticas de laboratorio con Wireshark

"Dímelo y lo olvidaré. Enséñamelo y lo recordaré. Implícame y lo entenderé."

Proverbio chino

Para comprender mejor los protocolos de red se puede profundizar enormemente en ellos viéndolos en acción y observando, por ejemplo, la secuencia de mensajes intercambiados entre dos entidades, examinando los detalles de la operación del protocolo, haciendo que lleven a cabo determinadas acciones y observando dichas acciones y sus consecuencias. Esto puede hacerse en escenarios simulados o en un entorno de red real, como Internet. Los applets Java disponibles en el sitio web del libro aplican el primero de estos métodos. En el laboratorio con Wireshark se aplicará el segundo método. Se ejecutan aplicaciones de red en diversos escenarios utilizando una computadora doméstica, de una oficina o de un laboratorio de prácticas. Podrá observar los protocolos de red en su equipo, interactuando e intercambiando mensajes con entidades que se ejecutan en cualquier punto de Internet. Así, usted y su computadora serán una parte integral de estas prácticas de laboratorio. Podrá observar practicando y, de ese modo, aprender.

La herramienta básica para observar los mensajes intercambiados entre entidades que ejecutan protocolos es un *husmeador de paquetes* (***packet sniffer***). Como su nombre sugiere, un husmeador de paquetes copia de forma pasiva los mensajes que están siendo enviados y recibidos por una computadora; también muestra el contenido de los distintos campos de protocolo de los mensajes capturados. En la Figura 1.29 se muestra una captura de pantalla del software Wireshark. Wireshark es un husmeador de paquetes gratuito que se ejecuta en sistemas Windows, Linux/Unix y Mac. A lo largo del libro, encontrará prácticas de laboratorio con Wireshark que le permitirán explorar los protocolos estudiados en el capítulo. En la primera práctica de laboratorio con Wireshark, tendrá que conseguir e instalar una copia de Wireshark, acceder a un sitio web y capturar y examinar los mensajes de protocolo que estén siendo intercambiados entre su navegador web y el servidor web.

Puede encontrar todos los detalles acerca de esta primera práctica de laboratorio con Wireshark (incluyendo las instrucciones acerca de cómo obtener e instalar Wireshark) en el sitio web http://www.awl.com/kurose-ross.

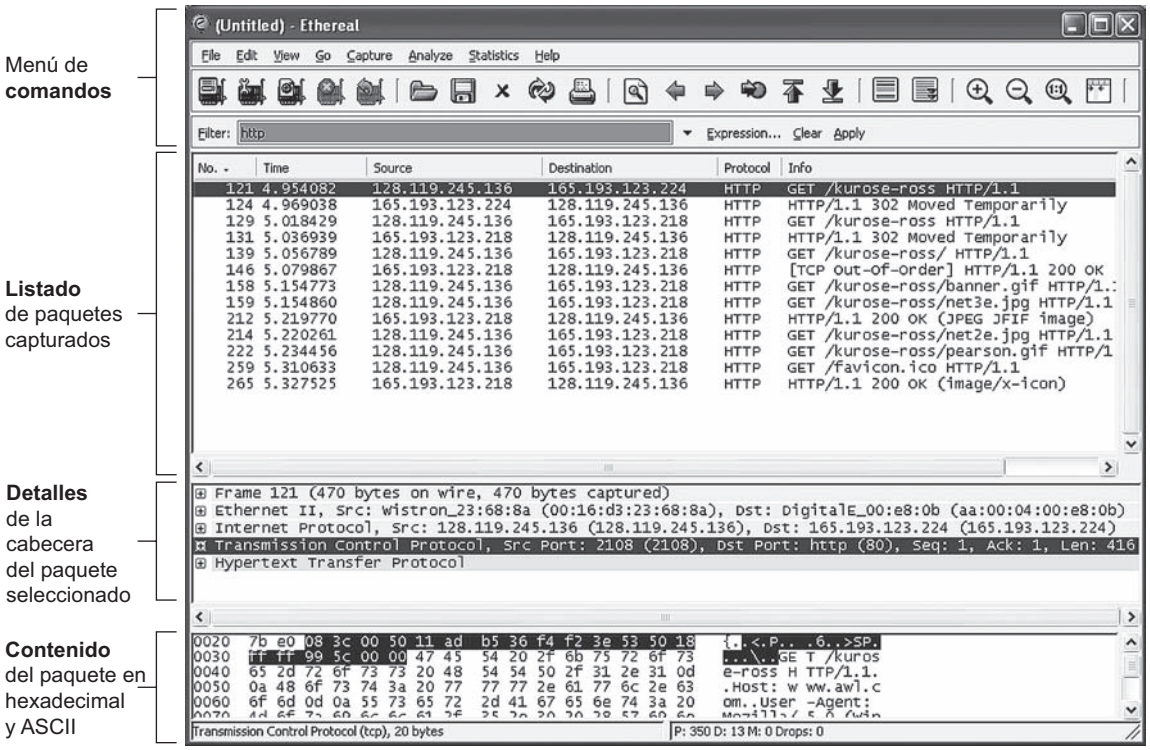


Figura 1.29 • Una captura de pantalla de Wireshark.

UNA ENTREVISTA CON...

Leonard Kleinrock

Leonard Kleinrock es catedrático de Ciencias de la Computación en la Universidad de California, Los Ángeles. En 1969, su computadora en UCLA se convirtió en el primer nodo de Internet. Su definición de los principios de la conmutación de paquetes en 1961 se convirtió en la tecnología en la que hoy se basa Internet. Recibió su licenciatura en el City College of New York (CCNY) y sus títulos de máster y doctor en Ingeniería Eléctrica en el MIT.

¿Qué le hizo decidirse a especializarse en la tecnología de redes e Internet?

Como estudiante de doctorado en el MIT en 1959, me di cuenta de que la mayor parte de mis compañeros de clase estaban realizando sus investigaciones en el área de la teoría de la información y de la teoría de la codificación. En el MIT se encontraba el gran investigador Claude Shannon, quien había abierto estos campos de investigación y que ya había resuelto la mayor parte de los problemas importantes. Los problemas de investigación que quedaban eran muy complicados y de consecuencias mucho menos importantes. Así que decidí iniciarme en una nueva área en la que nadie había pensado todavía. En el MIT, estaba rodeado de montones de computadoras y vi claramente que pronto esas máquinas necesitarían comunicarse entre sí. En aquel momento, no existía una forma efectiva de hacerlo, por lo que decidí desarrollar la tecnología que permitiría crear redes de datos eficientes y fiables.

¿Cuál fue su primer trabajo en la industria informática? ¿Qué significó para usted?

Entre 1951 y 1957 realicé en el CCNY los estudios de grado en Ingeniería Eléctrica en el turno de tarde. Durante el día, trabajé primero como técnico y luego como ingeniero en una pequeña empresa de electrónica industrial llamada Photobell. Mientras estuve allí, introduje la tecnología digital en sus líneas de productos. Básicamente, utilizábamos dispositivos fotoeléctricos para detectar la presencia de ciertos elementos (cajas, personas, etc.). El uso de un circuito que por entonces se denominaba *multivibrador biestable* era la clase de tecnología que necesitábamos para llevar el procesamiento digital al campo de la detección. Estos circuitos resultaron ser la base de las computadoras y ahora se conocen como *flip-flops, biestables* o *conmutadores* en la jerga actual.

¿Qué pasó por su cabeza cuando envió el primer mensaje de un host a otro (desde UCLA al Instituto de Investigación de Stanford)?

Francamente, no teníamos ni idea de la importancia de aquel suceso. No teníamos preparado un mensaje especial que pasara a la historia, como tantos otros inventores del pasado (Samuel Morse con "¡Lo que ha hecho Dios!", Alexander Graham Bell con "Watson, ¡ven aquí! Te necesito" o Neal Amstrong con "Un pequeño paso para el hombre, pero un gran paso para la Humanidad") ¡Aquellos tipos sí eran *inteligentes*! Conocían a los medios de comunicación y sabían lo que eran las relaciones públicas. Todo lo que nosotros queríamos hacer era iniciar una sesión en la computadora del SRI. Por lo que escribimos "L", lo que fue correctamente recibido, escribimos luego la letra "o", que fue recibida, y después la letra "g", que hizo que la computadora host del SRI fallara estrepitosamente. Así, nuestro mensaje fue el más corto de la historia: "Lo!".

Anteriormente, aquel año, yo había sido citado en una revista de UCLA diciendo que una vez que la red estuviera activa y funcionando, sería posible acceder a las utilidades informáticas desde nuestros

hogares y oficinas tan fácilmente como ya era disponer de electricidad y teléfono. Por tanto, mi visión en aquel momento era que Internet estaría en todas partes, siempre en funcionamiento, siempre disponible, que cualquiera con cualquier dispositivo podría conectarse desde cualquier lugar y que sería invisible. Sin embargo, nunca pude prever que mi madre con 99 años utilizaría Internet, y realmente la utilizó.

¿Cómo ve el futuro de las redes?

La parte fácil de la visión es predecir la propia infraestructura. Preveo que veremos una considerable implantación de la computación nómada, los dispositivos móviles y los espacios inteligentes. De hecho, la disponibilidad de dispositivos informáticos ligeros, baratos, de altas prestaciones y portátiles, y la disponibilidad de dispositivos de comunicaciones (combinado con la omnipresencia de Internet) nos ha permitido convertirnos en nómadas. La computación nómada hace referencia a la tecnología que permite a los usuarios finales ir de un lugar a otro obteniendo acceso a los servicios de Internet de forma transparente, independientemente de por dónde viajen e independientemente del dispositivo que utilicen. La parte más complicada de esta visión de futuro es predecir las aplicaciones y servicios, que siempre nos han sorprendido de forma increíble (correo electrónico, tecnologías de búsqueda, la world-wide-web, los blogs, las redes sociales, la generación y compartición por parte de los usuarios de música, fotografías y vídeos, etc.). Nos encontramos en el amanecer de una nueva era de aplicaciones móviles sorprendentes e innovadoras, que podremos utilizar con nuestros dispositivos de mano.

La siguiente ola nos permitirá pasar del mundo virtual del ciberespacio al mundo físico de los espacios inteligentes. Nuestros entornos (mesas, paredes, vehículos, relojes, etc.) cobrarán vida con la tecnología, mediante actuadores, sensores, lógica, procesamiento, almacenamiento, cámaras, micrófonos, altavoces, pantallas y comunicación. Esta tecnología integrada permitirá a nuestro entorno proporcionar los servicios IP que deseemos. Cuando accedamos a una habitación, la habitación sabrá que hemos entrado. Podremos comunicarnos con nuestro entorno de forma natural, hablando en nuestra lengua materna; nuestras solicitudes generarán respuestas que nos presentarán páginas web en pantallas de pared, en los cristales de las gafas, en forma de texto hablado, de hologramas, etc.

Mirando un poco más lejos, preveo un futuro en red que incluya los siguientes componentes adicionales fundamentales. Veo agentes software inteligentes por toda la red, cuya función será escarbar en los datos, actuar de acuerdo con ellos, detectar tendencias y llevar a cabo tareas de forma dinámica y adaptativa. Preveo que habrá una cantidad de tráfico de red considerablemente mayor, generada no tanto por los seres humanos, sino por estos dispositivos integrados y esos agentes software inteligentes. Preveo que habrá grandes conjuntos de sistemas dotados de auto-organización y encargados de controlar esa red tan enorme y tan rápida. Preveo que habrá enormes cantidades de información viajando instantáneamente a través de esa red y viéndose sometida en el camino a enormes cantidades de procesamiento y de filtrado. Internet será, esencialmente, un sistema nervioso global que llegará a todas partes. Preveo que sucederán todas estas cosas y muchas más a medida que nos vayamos adentrando en el siglo XXI.

¿Qué personas le han inspirado profesionalmente?

Con diferencia, el que más me ha inspirado ha sido Claude Shannon del MIT, un brillante investigador que tenía la capacidad de relacionar sus ideas matemáticas con el mundo físico de una forma extremadamente intuitiva. Estuvo en el tribunal de mi tesis doctoral.

¿Tiene algún consejo para los estudiantes que se inician ahora en el campo de las redes y de Internet?

Internet y todo lo que esa red hace posible constituye una nueva frontera de grandes dimensiones, llena de desafíos asombrosos. Hay grandes oportunidades para la innovación y no hay que sentirse restringido por la tecnología actual. Lo que hay que hacer es abrir la mente e imaginar cómo podrían ser las cosas, y luego hacer que eso suceda.