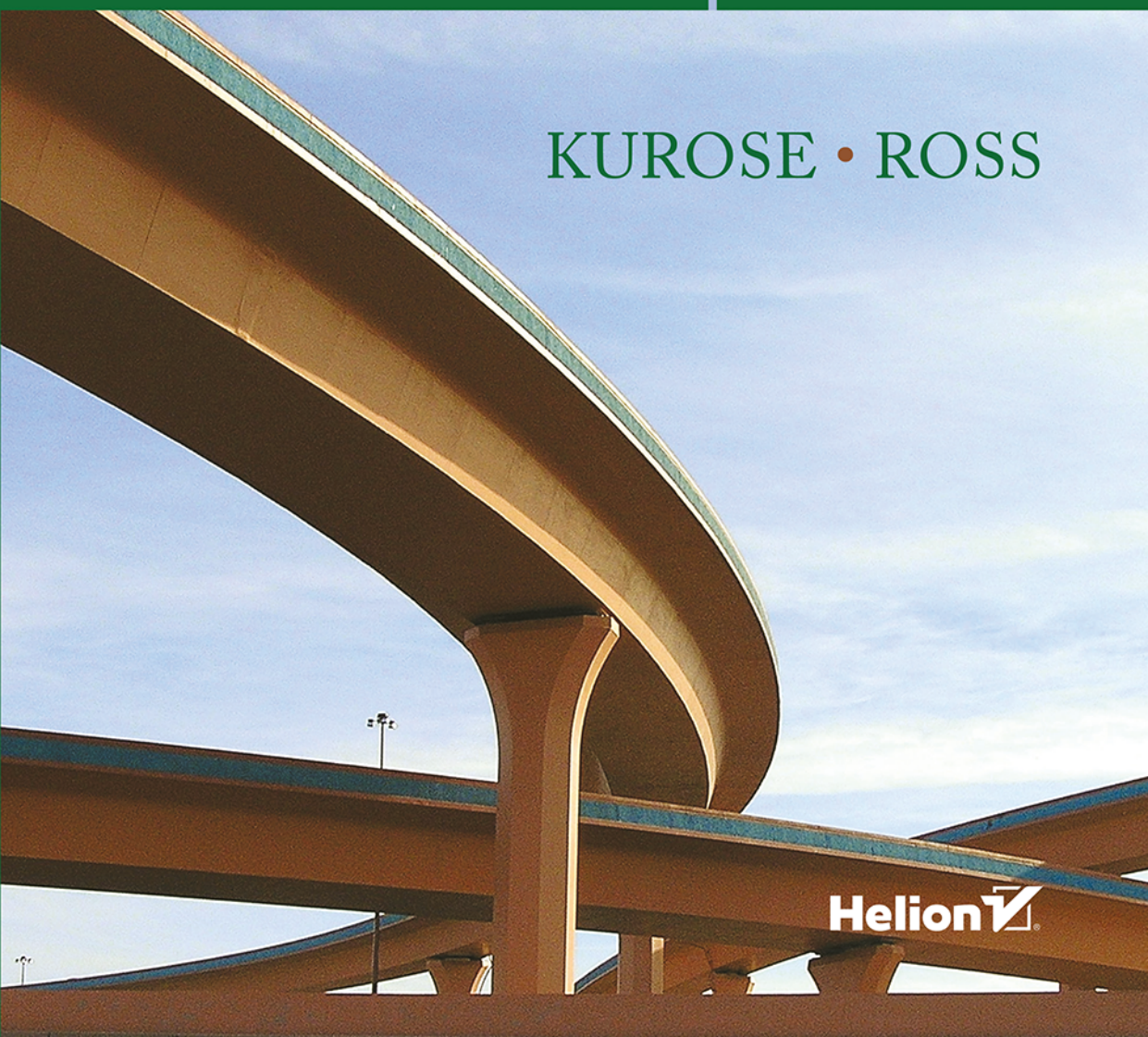
Sieci komputerowe

UJĘCIE CAŁOŚCIOWE | WYDANIE VII



KUROSE • ROSS

Helion

Tytuł oryginału: Computer Networking: A Top-Down Approach (7th Edition)

Tłumaczenie: Tomasz Walczak

ISBN: 978-83-8322-562-3

Authorized translation from the English language edition, entitled: COMPUTER NETWORKING: A TOP-DOWN APPROACH, Seventh Edition; ISBN 0133594149; by James F. Kurose; and by Keith W. Ross; published by Pearson Education, Inc.

Helion S.A.
ul. Kościuszki 1c, 44-100 Gliwice
tel. 32 230 98 63
e-mail: *helion@helion.pl*
WWW: *https://helion.pl* (księgarnia internetowa, katalog książek)

Drogi Czytelniku!
Jeżeli chcesz ocenić tę książkę, zajrzyj pod adres
https://helion.pl/user/opinie/sieu7v
Możesz tam wpisać swoje uwagi, spostrzeżenia, recenzję.

Dodatkowe materiały do książki można znaleźć pod adresem:
https://ftp.helion.pl/przyklady/sieu7v.zip

Printed in Poland.

Spis treści

ROZDZIAŁ 1

Sieci komputerowe i internet

Współczesny internet to prawdopodobnie największy system kiedykolwiek zaprojektowany przez ludzi. Składają się na niego setki miliony powiązanych komputerów, połączeń komunikacyjnych i przełączników. Miliardy osób okresowo łączy się z internetem za pomocą laptopów, tabletów i smartfonów. Ponadto istnieje cały zestaw nowych „rzeczy" podłączanych do internetu: konsol do gier, systemów monitoringu, zegarków, okularów, termostatów, wag czy samochodów. Czy można zrozumieć funkcjonowanie internetu, skoro jest on tak duży, składa się z tak wielu różnorodnych komponentów i ma mnóstwo zastosowań? Czy istnieją zasady i struktury, które pomogą zrozumieć tak niezwykle rozbudowany oraz złożony system? A jeśli tak, to czy możliwe jest, że poznawanie sieci komputerowych może być zarówno interesujące, *jak i* przyjemne? Na szczęście odpowiedź na wszystkie te pytania to stanowcze TAK! Za pomocą tej książki chcemy wprowadzić Czytelników w dynamiczny świat sieci komputerowych oraz zaprezentować zasady i praktyczne wskazówki potrzebne do zrozumienia nie tylko obecnych sieci, ale i tych, które pojawią się w przyszłości.

W niniejszym, pierwszym rozdziale przedstawiamy ogólne omówienie sieci komputerowych i internetu. Naszym celem jest naszkicowanie całościowego obrazu i określenie kontekstu dla następnych części książki. Pozwoli to Czytelnikom ujrzeć las spoza drzew. W rozdziale zamieściliśmy wiele podstawowych informacji i omówiliśmy mnóstwo elementów sieci komputerowej, mając jednocześnie na względzie sieć jako całość.

Przegląd sieci komputerowych dokonany w tym rozdziale ma następujący układ. Po zaprezentowaniu kilku podstawowych terminów i zagadnień w pierwszej kolejności przyjrzymy się elementarnym składnikom sprzętowym i programowym, które tworzą sieć. Rozpoczniemy od obrzeża sieci, po czym omówimy systemy końcowe i aplikacje

uruchamiane w sieci. W dalszej kolejności objaśnimy rdzeń sieci komputerowej, omawiając łącza i przełączniki przesyłające dane, a także sieci dostępowe i fizyczne nośniki łączące systemy końcowe z rdzeniem sieci. Czytelnik dowie się, że internet jest siecią złożoną z wielu sieci, a ponadto omówimy, w jaki sposób są one ze sobą połączone.

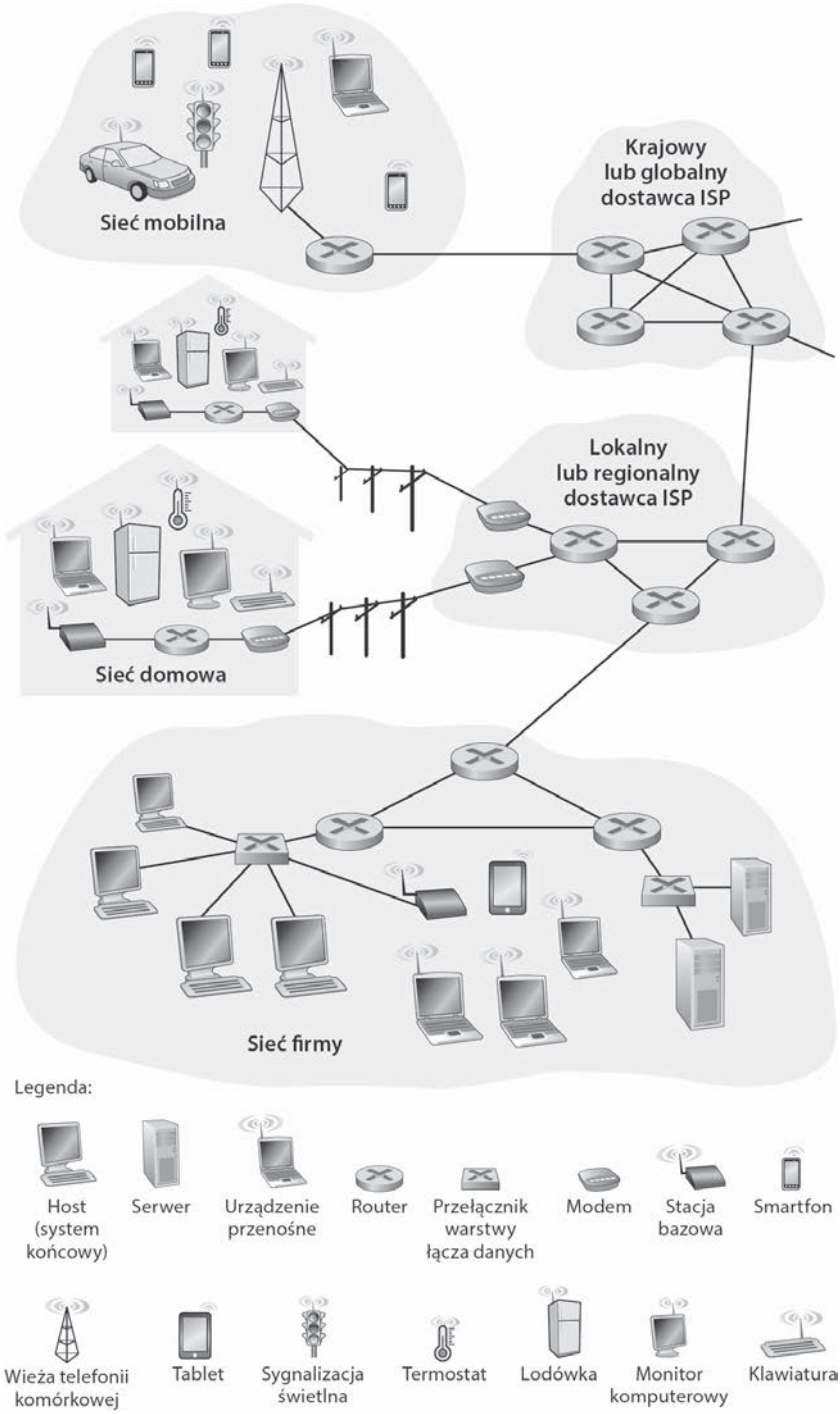
Po zakończeniu przeglądu obrzeża sieci komputerowej i jej rdzenia w drugiej części rozdziału dokonamy szerszego i bardziej abstrakcyjnego omówienia. Przyjrzymy się przyczynom opóźnień w transmisji danych w sieciach, ich utracie i przepustowości. Zaprezentujemy proste modele ilościowe demonstrujące przepustowość i opóźnienia występujące między dwoma węzłami końcowymi. Modele te będą uwzględniały opóźnienia transmisji, propagacji i kolejkowania. Dalej przedstawimy kilka kluczowych zasad dotyczących architektury sieci komputerowych, a dokładniej mówiąc, warstw protokołów i modeli usług. Wyjaśnimy też, że sieci komputerowe są podatne na wiele różnorodnych ataków. Omówimy kilka z nich i pokażemy, jak zabezpieczyć sieć. Na końcu rozdziału zamieszczono krótką historię sieci komputerowych.

1.1. Czym jest internet?

W książce publiczny internet będący specyficzną siecią komputerową traktujemy jako podstawowy fundament, na którym są oparte omówienia sieci i ich protokołów. Jednak czym *jest* internet? Odpowiedzi można udzielić na kilka sposobów. Jeden z nich polega na omówieniu podstawowych sprzętowych i programowych składników tworzących internet. Kolejnym sposobem jest scharakteryzowanie internetu w kategoriach infrastruktury sieciowej, która zapewnia usługi aplikacjom rozproszonym. Zacznijmy od opisu podstawowych składników, korzystając z rysunku 1.1 ilustrującego omówienie.

1.1.1. Opis podstawowych komponentów

Publiczny internet jest siecią komputerową łączącą miliardy urządzeń komputerowych zlokalizowanych w różnych miejscach świata. Nie tak dawno temu tymi urządzeniami były przede wszystkim tradycyjne stacjonarne komputery PC, linuksowe stacje robocze i tak zwane serwery, które przechowują i przesyłają dane, takie jak strony internetowe i wiadomości pocztowe. Jednak wraz z upływem czasu do internetu zaczęto podłączać nietradycyjne rzeczy internetowe, takie jak laptopy, smartfony, tablety, telewizory, konsole do gier, termostaty, systemy zabezpieczeń, sprzęt AGD, zegarki, okulary, samochody, systemy sterowania ruchem itd. Tak naprawdę termin *sieć komputerowa* zaczyna być trochę nieaktualny, gdy pod uwagę weźmie się to, że do internetu podłączanych jest wiele nietradycyjnych urządzeń. W żargonie internetowym wszystkie wymienione urządzenia są nazywane **hostami** lub **systemami końcowymi**. Według szacunków w 2015 r. z internetu korzystało ok. 5 miliardów systemów końcowych, a do 2020 r. ta liczba sięgnie 25 miliardów [Gartner 2014]. Szacuje się, że w 2015 r. na świecie było ponad 3,2 miliarda użytkowników internetu, co stanowi ok. 40% populacji świata [ITU 2015].



Rysunek 1.1. Niektóre składniki internetu

Systemy końcowe są połączone za pomocą **łączy komunikacyjnych** i **przełączników pakietów**. W podrozdziale 1.2 będzie można się dowiedzieć, że istnieje wiele typów łączy komunikacyjnych, które są złożone z różnego rodzaju nośników fizycznych, takich jak kabel koncentryczny, kabel z miedzi, światłowód i widmo radiowe. Różne łącza mogą przesyłać dane z innymi szybkościami. **Szybkość transmisji** łącza jest wyrażana w bitach na sekundę. Kiedy jeden system końcowy wysyła dane do innego takiego systemu, dzieli informacje na segmenty i dołącza do każdego z nich nagłówek. Powstają w ten sposób porcje danych, nazywane w żargonie związanym z sieciami komputerowymi **pakietami**. Są one przesyłane przez sieć do docelowego systemu końcowego, gdzie pakiety są ponownie łączone w pierwotne dane.

Przełącznik pakietów pobiera porcję danych, którą odbierze za pomocą jednego ze swoich wejściowych łączy komunikacyjnych, i przekazuje ją do jednego z łączy wyjściowych. Choć dostępne są różne odmiany przełączników pakietów, aktualnie w internecie najczęściej wykorzystuje się dwa z nich — **routery** i **przełączniki warstwy łącza danych**. Oba typy przełączników przekazują pakiety w kierunku ich ostatecznego miejsca przeznaczenia. Przełączniki warstwy łącza danych zwykle działają w sieciach dostępowych, natomiast routery przeważnie są wykorzystywane w rdzeniu sieci. Droga pokonywana przez pakiet wysłany przez nadawczy system końcowy, a następnie przechodzący przez kolejne łącza komunikacyjne i przełączniki pakietów, aby dotrzeć do odbiorczego systemu końcowego, jest nazywana **trasą** lub **ścieżką** sieciową. Firma Cisco prognozuje, że globalny ruch IP pod koniec 2016 r. przekroczy poziom zetabajta (10^{21} bajtów), a do 2019 r. sięgnie 2 zetabajtów rocznie [Cisco VNI 2015].

Sieci z przełącznikami pakietów (służące do przesyłania pakietów) pod wieloma względami przypominają sieci transportowe składające się z autostrad, mniejszych dróg i skrzyżowań (umożliwiają one ruch pojazdów). Warto wyobrazić sobie fabrykę, która musi przewieźć dużą ilość towaru do magazynu oddalonego o tysiące kilometrów. W fabryce towar jest dzielony i ładowany na poszczególne ciężarówki. Każda z nich niezależnie porusza się siecią autostrad, dróg i skrzyżowań, aby dotrzeć do docelowego magazynu. W nim towar jest rozładowywany i łączony z resztą ładunku z tej samej dostawy. Pakiety to odpowiednik ciężarówek, łącza komunikacyjne to autostrady i drogi, przełączniki pakietów przypominają skrzyżowania, a systemy końcowe są zbliżone do budynków. Podobnie jak ciężarówki poruszają się określoną trasą w ramach sieci transportowej, tak pakiet podąża ścieżką w sieci komputerowej.

Systemy końcowe uzyskują dostęp do internetu za pośrednictwem **dostawcy usług internetowych ISP** (ang. *Internet Service Provider*). Może to być lokalny dostawca, na przykład operator telekomunikacyjny bądź telewizji kablowej. Może to być korporacyjny lub uniwersytecki dostawca ISP bądź taki, który zapewnia dostęp bezprzewodowy na lotniskach, w hotelach, kawiarniach i innych miejscach użyteczności publicznej. Jako dostawcy ISP działają też operatorzy komórkowi, zapewniający dostęp do internetu w smartfonach i innych urządzeniach. Sieć każdego dostawcy ISP jest złożona z przełączników pakietów i łączy komunikacyjnych. Dostawcy ISP oferują systemom końcowym różnego typu dostęp sieciowy, taki jak dostęp szerokopasmowy za pośrednictwem modemu kablowego lub DSL, dostęp za pomocą bardzo szybkiej sieci lokalnej

i dostęp bezprzewodowy. Dostawcy ISP oferują też dostęp firmom zapewniającym informacje, podłączającym witryny WWW oraz serwery z materiałami wideo bezpośrednio do internetu. Aby umożliwić systemom końcowym komunikowanie się ze sobą, wymienieni dostawcy ISP niższego poziomu są połączeni ze sobą przez krajowych lub międzynarodowych dostawców wyższego poziomu, takich jak Level 3 Communications, AT&T, Sprint i NTT. Sieć dostawców wyższego poziomu zawiera bardzo szybkie routery połączone ze sobą przy użyciu szybkich łączy światłowodowych. Każda sieć dostawcy ISP, niezależnie od tego, czy niższego czy wyższego poziomu, jest oddzielnie zarządzana, a ponadto stosuje protokół IP (opisany niżej) i jest zgodna z określonymi konwencjami nazewniczymi i adresowymi. Dostawcom usług internetowych i ich powiązaniom przyjrzymy się bliżej w podrozdziale 1.3.

Systemy końcowe, przełączniki pakietów i inne składniki internetu wykorzystują **protokoły**, które kontrolują proces wysyłania i odbierania informacji transferowanych w internecie. **TCP** (ang. *Transmission Control Protocol*) i **IP** (ang. *Internet Protocol*) to dwa najważniejsze protokoły stosowane w internecie. Protokół IP określa format pakietów wysyłanych i odbieranych przez routery i systemy końcowe. W przypadku tych dwóch głównych protokołów internetowych używa się pojedynczego oznaczenia **TCP/IP**. Choć w tym rozdziale rozpoczniemy omawianie protokołów, jest to zaledwie początek. Spora część książki została poświęcona protokołom sieci komputerowych!

Ze względu na istotną rolę, jaką protokoły odgrywają w przypadku internetu, ważne jest jednoznaczne określenie przeznaczenia każdego protokołu, co pozwala tworzyć współdziałające ze sobą systemy i produkty. Właśnie to umożliwiają standardy. **Standardy internetowe** są opracowywane przez organizację **IETF** (ang. *Internet Engineering Task Force*) [IETF 2016]. Dokumenty standardów tej organizacji są nazywane **dokumentami RFC** (ang. *Requests for Comments*). Pierwotnie dokumenty RFC zawierały ogólne prośby o dołączanie własnych wniosków (stąd wywodzi się termin RFC — prośby o komentarze) umożliwiających rozwiązanie problemów związanych z architekturą sieci i protokołów, z którymi borykali się prekursorzy internetu [Allman 2011]. Dokumenty RFC mają tendencję do bycia dość technicznymi i szczegółowymi. Definiują protokoły, takie jak TCP, IP, HTTP (obsługa witryn WWW) i **SMTP** (ang. *Simple Mail Transfer Protocol*; obsługa poczty elektronicznej). Istnieje ponad 7000 dokumentów RFC. Inne organizacje też definiują standardy dotyczące komponentów sieciowych, a zwłaszcza łączy sieciowych. Przykładowo, organizacja IEEE 802 LAN/MAN Standards Committee [IEEE 802 2016] opracowuje standardy dotyczące Ethernetu i technologii bezprzewodowej Wi-Fi.

1.1.2. Omówienie usług

Dotychczas zidentyfikowano wiele składników tworzących internet. Jednak można go opisać także z zupełnie innej perspektywy — *jako infrastrukturę zapewniającą usługi dla aplikacji*. Obok tradycyjnych aplikacji, takich jak klienty poczty elektronicznej lub przeglądarki internetowe, występują też narzędzia przeznaczone na smartfony i tablety, w tym komunikatory internetowe, mapy z pobieranymi w czasie rzeczywistym

informacjami o ruchu, sieci społecznościowe, systemy strumieniowania danych z chmury, filmów i programów telewizyjnych, gry wieloosobowe, a także systemy rekomendacji oparte na lokalizacji. Są to tak zwane **aplikacje rozproszone**, w przypadku których systemy końcowe mogą przesyłać między sobą dane. Co ważne, aplikacje internetowe działają w systemach końcowych, a nie w przełącznikach pakietów w rdzeniu sieci. Choć przełączniki pakietów ułatwiają wymianę danych między systemami końcowymi, nie uwzględniają tego, jakie aplikacje są nadawcą lub odbiorcą informacji.

Wyjaśnijmy dokładniej, co mamy na myśli, pisząc o infrastrukturze zapewniającej usługi aplikacjom. Załóżmy, że programista ma świetny pomysł na rozproszoną aplikację internetową, która przyniesie korzyści całej ludzkości (lub po prostu zapewni tej osobie bogactwo i sławę). Jak można przekształcić tę ideę na rzeczywisty program internetowy? Aplikacje działają w systemach końcowych, dlatego trzeba napisać odpowiednie oprogramowanie. Programista może to zrobić na przykład za pomocą języka Java, C lub Python. Ponieważ chce utworzyć rozproszoną aplikację internetową, programy działające w różnych systemach końcowych muszą mieć możliwość przesyłania między sobą danych. Dochodzimy w ten sposób do podstawowego zagadnienia związanego z odmiennym opisem internetu — jako platformy dla aplikacji. W jaki sposób oprogramowanie działające w systemie końcowym ma nakazać internetowi przesłanie danych do drugiego programu funkcjonującego w innym systemie?

Systemy końcowe podłączone do internetu udostępniają **interfejs gniazd**. Określa on, w jaki sposób program działający w jednym systemie końcowym ma żądać od internetu przekazania danych do określonej aplikacji docelowej funkcjonującej w innym systemie. Interfejs gniazd internetu obejmuje zestaw reguł, których oprogramowanie wysyłające informacje musi przestrzegać, aby internet mógł przesłać dane do aplikacji docelowej. Interfejs gniazd internetu omawiamy szczegółowo w rozdziale 2. Na razie zaprezentujemy prostą analogię, do której będziemy często wracać w tej książce. Załóżmy, że Alicja chce wysłać do Bartka list pocztą. Alicja nie może oczywiście po prostu napisać wiadomości (dane) i wyrzucić kartki przez okno. Poczta wymaga, aby Alicja umieściła list w kopercie, napisała w jej środkowej części imię i nazwisko Bartka, jego adres oraz kod pocztowy, zakleiła kopertę, nakleiła znaczek w prawym górnym rogu koperty, a następnie wrzuciła ją do oficjalnej skrzynki na listy. Poczta ma więc własny „interfejs usług pocztowych", czyli zestaw reguł, których Alicja musi przestrzegać, aby listonosz dostarczył list do Bartka. Podobnie internet ma interfejs gniazd, którego oprogramowanie wysyłające dane musi przestrzegać, aby internet dostarczył informacje do odbierającej je aplikacji.

Poczta oczywiście świadczy klientom różne usługi — wysyłkę priorytetową, listy polecone, przesyłki ekonomiczne itd. Podobnie internet udostępnia wiele usług aplikacjom. Przy rozwijaniu programów internetowych trzeba wybrać jedną z takich usług. Ich opis przedstawiamy w rozdziale 2.

Właśnie zaprezentowaliśmy dwa opisy internetu. Pierwszy odnosi się do sprzętowych i programowych składników internetu, natomiast drugi do jego usług zapewnianych aplikacjom rozproszonym. Jednak być może w dalszym ciągu Czytelnik nie jest do końca pewien, czym jest internet. Co to jest przełączanie pakietów i protokoły

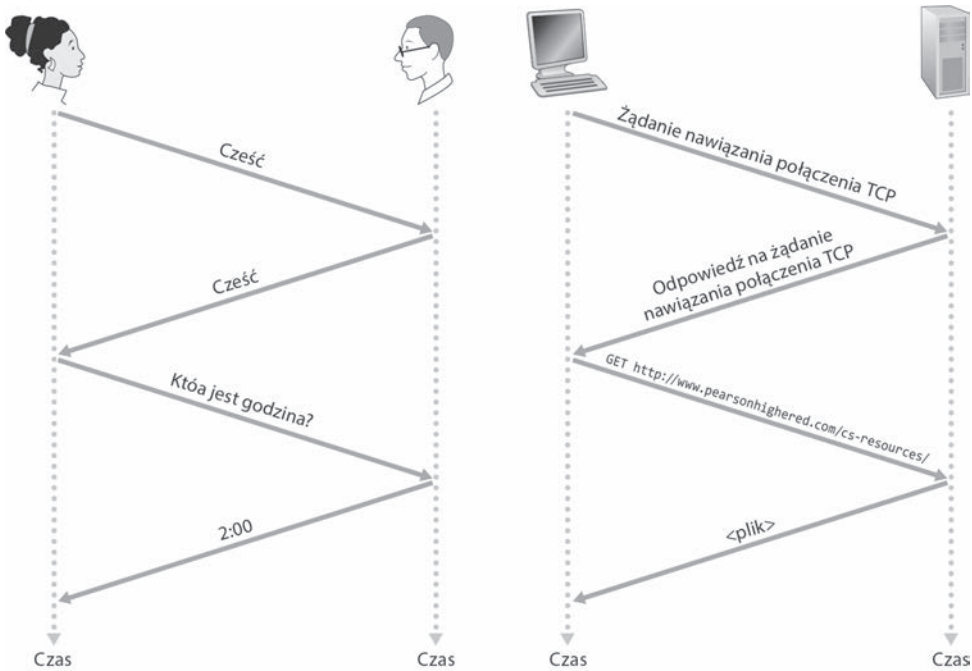
TCP/IP? Czym są routery? Jakiego typu łącza komunikacyjne występują w internecie? Czym jest aplikacja rozproszona? W jaki sposób do internetu można podłączyć termostat lub wagę? Jeśli na chwilę obecną Czytelnik jest trochę przytłoczony tym wszystkim, nie ma powodu do obaw, ponieważ celem książki jest zaprezentowanie zarówno podstawowych składników internetu, jak i zasad, które decydują o jego funkcjonowaniu. W kolejnych podrozdziałach i rozdziałach zostaną omówione wszystkie wymienione zagadnienia i udzielone zostaną odpowiedzi na zadane pytania.

1.1.3. Czym jest protokół?

Gdy już trochę bliżej zapoznaliśmy się z internetem, pod uwagę weźmy kolejny istotny termin związany z sieciami komputerowymi, którym jest *protokół*. Czym jest protokół? Jaka jest jego rola?

Analogia nawiązująca do komunikacji między ludźmi

Prawdopodobniej najprostszym sposobem pozwalającym zrozumieć protokół sieci komputerowej będzie zastosowanie w pierwszej kolejności analogii dotyczącej ludzi, którzy cały czas korzystają z protokołów. Zastanówmy się, co robimy, gdy chcemy zapytać kogoś o godzinę. Na rysunku 1.2 przedstawiono przebieg typowej wymiany. Ludzki protokół (lub przynajmniej dobre maniery) nakazuje, aby najpierw się przywitać (zwrot „Cześć" na rysunku 1.2), zanim rozpocznie się z kimś rozmowę. Typową odpowiedzią na zwrot „Cześć" jest takie samo słowo. Automatycznie przyjmuje się, że po usłyszeniu w odpowiedzi serdecznego zwrotu „Cześć" można kontynuować rozmowę i zapytać o godzinę. Inne różne odpowiedzi na początkowy zwrot „Cześć" (takie jak „Nie przeszkadzaj!", „Nie mówię po polsku" lub inne, których nie można tu przytoczyć) mogą wskazywać na to, że rozmówca nie ma ochoty na konwersację lub nie ma możliwości jej prowadzenia. W tym przypadku ludzki protokół nie umożliwi poznania godziny. Czasami na zadane pytanie nie otrzyma się żadnej odpowiedzi. W tego typu sytuacji zwykle rezygnuje się z pytania o godzinę. Warto zauważyć, że w przypadku ludzkiego protokołu *istnieją określone przekazywane komunikaty i działania podejmowane w odpowiedzi na otrzymane komunikaty lub wystąpienie innych zdarzeń* (takich jak brak odpowiedzi w określonym przedziale czasu). Oczywiście przekazywane i otrzymywane komunikaty, a także czynności wykonywane w momencie transferowania komunikatów lub po wystąpieniu innych zdarzeń w przypadku ludzkiego protokołu odgrywają podstawową rolę. Jeśli ludzie posługują się różnymi protokołami, które nie są ze sobą zgodne (jeśli na przykład jedna osoba posiada maniery, a druga nie lub ktoś rozumie pojęcie czasu, natomiast ktoś inny nie), nie uzyska się żadnego pozytywnego efektu. To samo dotyczy sieci. W celu zrealizowania zadania dwie komunikujące się ze sobą jednostki muszą korzystać z tego samego protokołu.



Rysunek 1.2. Porównanie ludzkiego protokołu i protokołu sieci komputerowej

Rozważmy drugą analogię dotyczącą komunikowania się ludzi. Załóżmy, że jesteśmy na zajęciach (na przykład z sieci komputerowych!). Nauczyciel w nieciekawy sposób mówi o protokołach, co powoduje, że traci się w tym orientację. Prowadzący zajęcia przerywa i pyta się, czy są jakieś pytania (komunikat jest wysyłany i odbierany przez wszystkich studentów, z wyjątkiem tych, którzy śpią). Jeden ze studentów podnosi rękę (przekazując automatycznie komunikat nauczycielowi). Prowadzący reaguje na to uśmiechem i mówi: „Tak, słucham" (przekazany komunikat zachęca studenta do zadania pytania). Nauczyciele *uwielbiają*, gdy zadaje im się pytania. Student zadaje pytanie (przekazuje komunikat nauczycielowi). Po usłyszeniu pytania (otrzymaniu komunikatu) nauczyciel udziela odpowiedzi na nie (przekazuje komunikat studentowi). Również w tym przypadku widać, że przekazywane i otrzymywane komunikaty, a także zestaw typowych działań podejmowanych w chwili wysyłania i odbierania wiadomości to rdzeń protokołu opartego na pytaniach i odpowiedziach.

Protokoły sieciowe

Protokół sieciowy przypomina ludzki protokół, z tą różnicą, że jednostki wymieniające się komunikatami i podejmujące działania są sprzętowymi lub programowymi składnikami określonego urządzenia (na przykład komputera, smartfona, tabletu, routera lub innego urządzenia sieciowego). Wszystkie operacje realizowane w internecie, które angażują dwie lub więcej zdalnych jednostek komunikacyjnych, są zarządzane przez

protokół. Przykładowo, sprzętowe protokoły kart sieciowych dwóch połączonych ze sobą komputerów kontrolują przepływ bitów w kablu znajdującym się między dwoma interfejsami sieciowymi. Protokoły systemów końcowych kontrolujące przeciążenie decydują, z jaką szybkością pakiety będą transmitowane między nadawcą i odbiorcą. Protokoły routerów określają ścieżkę, jaką pakiet będzie podążał od miejsca źródłowego do docelowego. Protokoły są stosowane w każdym miejscu internetu. W związku z tym spora część książki jest poświęcona protokołom sieci komputerowych.

W ramach przykładu protokołu sieciowego, z którym Czytelnik prawdopodobnie jest zaznajomiony, zastanówmy się, co się stanie, gdy do serwera WWW wyśle się żądanie, czyli w oknie przeglądarki internetowej wprowadzi się adres URL witryny WWW. Zostało to zilustrowane w prawej części rysunku 1.2. Najpierw komputer wysyła do serwera WWW żądanie nawiązania połączenia i czeka na odpowiedź. Po otrzymaniu żądania serwer WWW zwraca komunikat odpowiedzi na żądanie. Gdy komputer już „wie", że może żądać pobrania strony internetowej z serwera WWW, w komunikacie `GET` wysyła do niego jej nazwę. Na końcu serwer WWW zwraca komputerowi stronę internetową (plik).

Po przytoczeniu przykładów protokołów ludzkich i sieciowych można powiedzieć, że wymiana komunikatów i działania podejmowane w chwili wysyłania i odbierania komunikatów to kluczowe elementy definiujące protokół.

> ***Protokół** definiuje format i kolejność komunikatów wymienianych między dwoma lub większą liczbą komunikujących się jednostek, a także operacje wykonywane w momencie wysyłania i (lub) odbierania komunikatu bądź innego zdarzenia.*

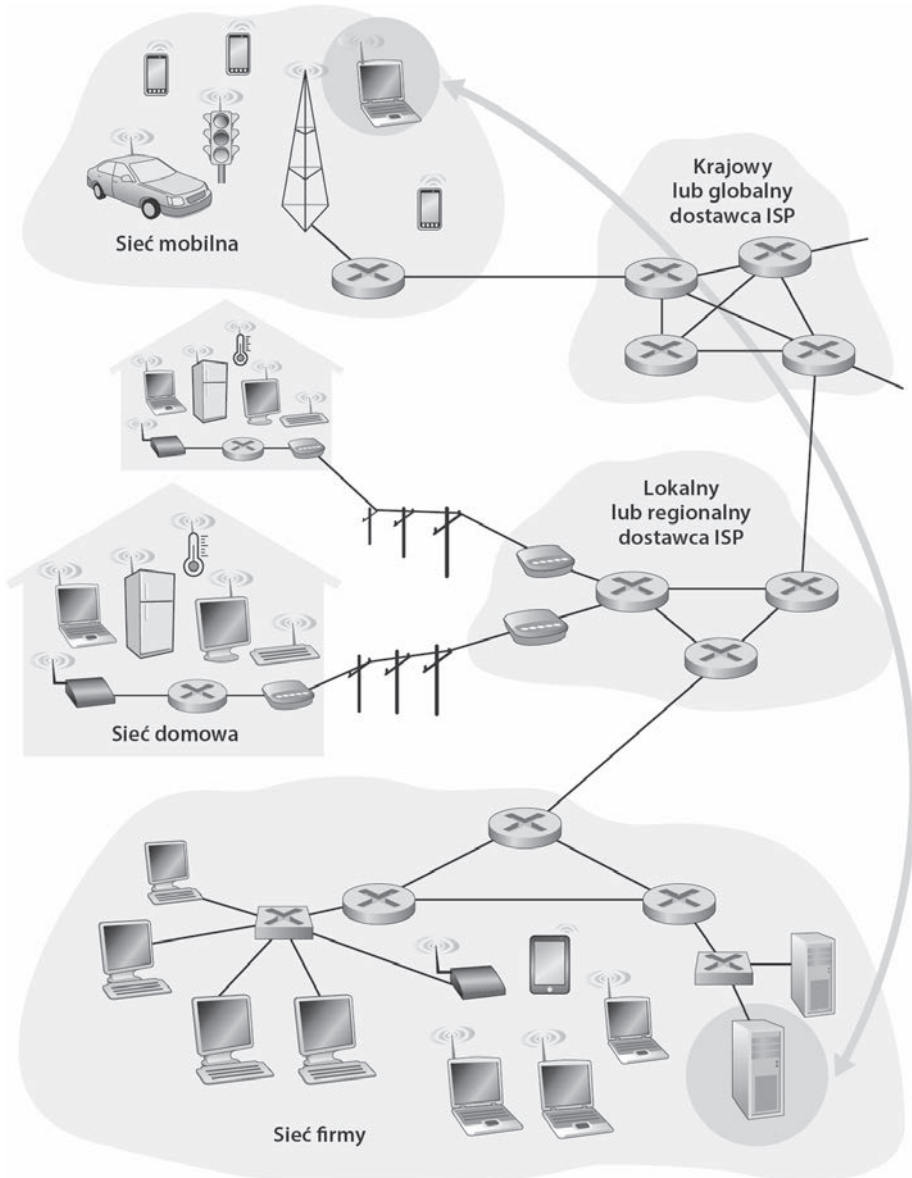
Zasadniczo internet i sieci komputerowe intensywnie korzystają z protokołów. Różne protokoły są używane w celu zrealizowania odmiennych zadań komunikacyjnych. W trakcie lektury książki okaże się, że niektóre protokoły są proste i oczywiste, natomiast inne złożone i zaawansowane intelektualnie. Opanowanie sieci komputerowych jest równoznaczne ze zrozumieniem, jak i dlaczego są stosowane protokoły sieciowe.

1.2. Obrzeże sieci

W poprzednich punktach dokonaliśmy bardzo ogólnego przeglądu internetu i protokołów sieciowych. Teraz trochę bardziej szczegółowo przyjrzymy się składnikom sieci komputerowej, a w szczególności internetu. W tym podrozdziale zaczniemy omawiać obrzeże sieci i komponenty, z którymi jesteśmy najbardziej zaznajomieni, czyli codziennie używane komputery, smartfony i inne urządzenia. W kolejnym podrozdziale przejdziemy z obrzeża sieci do jej rdzenia i zajmiemy się przełączaniem oraz routingiem wykorzystywanym w sieciach komputerowych.

W żargonie dotyczącym sieci komputerowych komputery podłączone do internetu często są nazywane systemami końcowymi. Wynika to stąd, że takie komputery są

zlokalizowane na obrzeżu internetu (rysunek 1.3). Do internetowych systemów końcowych należy zaliczyć komputery stacjonarne (na przykład biurkowe komputery PC, a także stacje robocze z systemami Macintosh i Linux), serwery (na przykład serwery WWW i pocztowe) i urządzenia przenośne (na przykład laptopy, smartfony i tablety). Ponadto do internetu podłączanych jest coraz więcej innych, nietradycyjnych rzeczy działających jako systemy końcowe (poniższa ramka „Kącik historyczny").



Rysunek 1.3. Interakcja systemów końcowych

Systemy końcowe określa się też terminem *hosty* (ang. *gospodarze*), ponieważ goszczą (uruchamiają) aplikacje, takie jak przeglądarka internetowa, program serwera WWW, program pocztowy lub program serwera poczty. W książce zamiennie będą używane terminy host i system końcowy. Oznacza to, że *host = system końcowy*. Czasami hosty są dzielone na dwie kategorie — **klienty** i **serwery**. Nieoficjalnie za klienty uważa się stacjonarne i przenośne komputery PC, palmtopy itp. Z kolei serwerami są bardziej wydajne komputery, które przechowują i dystrybuują strony internetowe, strumieniowe dane wideo, wiadomości pocztowe itp. Obecnie większość serwerów, z których pobieramy wyniki wyszukiwania, e-maile, strony WWW i filmy, działa w dużych **centrach danych**. Na przykład Google ma 50 – 100 centrów danych, w tym 15 dużych centrów obejmujących ponad 100 000 serwerów każde.

KĄCIK HISTORYCZNY

INTERNET RZECZY

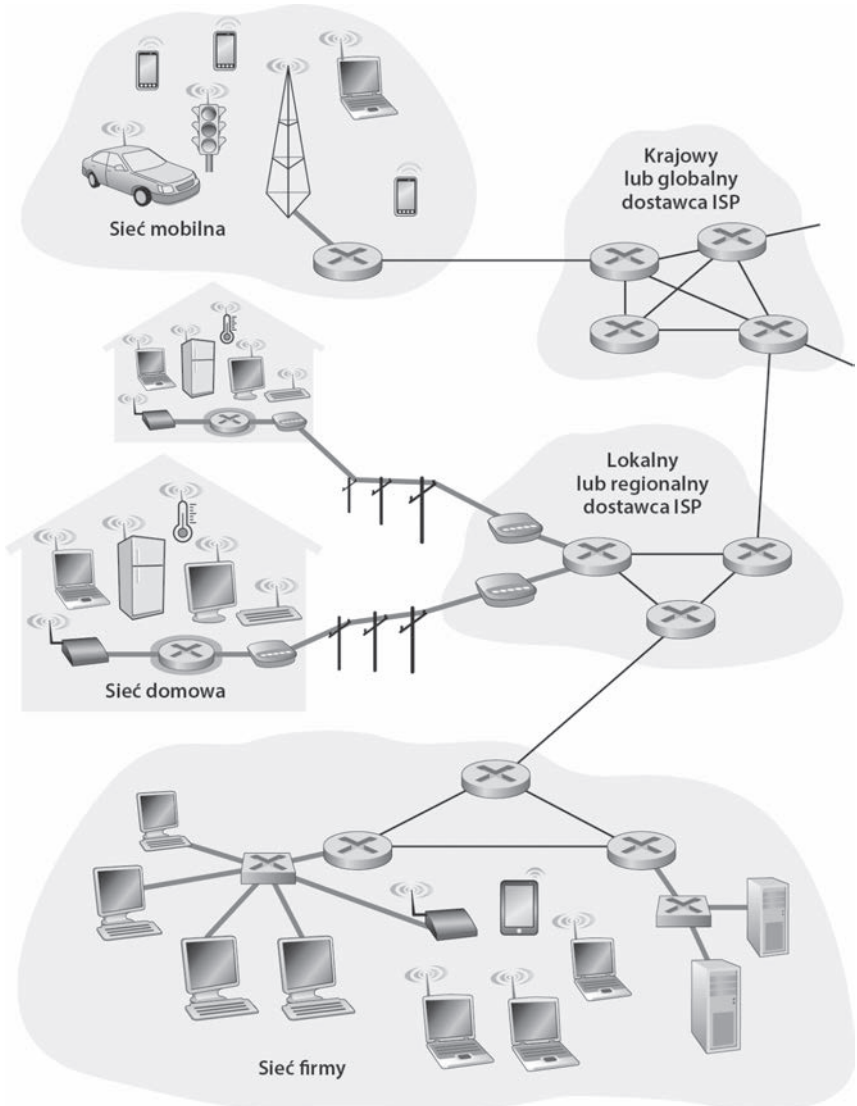
Potrafisz wyobrazić sobie świat, w którym niemal wszystko jest bezprzewodowo podłączone do internetu? Gdzie większość ludzi, samochodów, rowerów, okularów, zegarków, zabawek, sprzętu w szpitalach, czujników w domach, sal szkolnych, systemów monitoringu, czujników pogodowych, produktów na półkach i zwierząt jest w sieci? Możliwe, że świat internetu rzeczy znajduje się tuż za rogiem.

Według niektórych szacunków w 2015 r. do internetu podłączonych było już 5 miliardów rzeczy. Do 2020 r. ta liczba może wzrosnąć do 25 miliardów [Gartner 2014]. Tymi rzeczami są między innymi smartfony, które „śledzą nas" w mieszkaniach, biurach i samochodach, przesyłając naszą lokalizację i dane o użytkowaniu urządzenia do dostawców internetu i aplikacji internetowych. Oprócz smartfonów dostępne są także różnorodne inne nietradycyjne rzeczy z możliwością podłączenia do internetu. Niektóre z nich można nosić na sobie — na przykład zegarki (firmy Apple i wielu innych) czy okulary. Okulary podłączone do internetu mogą na przykład przesyłać wszystko, co widzimy, do chmury, co pozwala dzielić się w czasie rzeczywistym naszymi wrażeniami wizualnymi z ludźmi z całego świata. Dostępne są też podłączone do internetu rzeczy działające w ramach inteligentnych domów. Są to na przykład podłączone do internetu termostaty, które można zdalnie kontrolować za pomocą smartfonów, a także mające dostęp do sieci wagi, pozwalające wyświetlić na smartfonie graficzny podgląd skutków stosowania diety. Istnieją również podłączone do internetu zabawki, w tym lalki, które rozpoznają i interpretują słowa dziecka oraz odpowiednio na nie reagują.

Internet rzeczy może zapewniać użytkownikom rewolucyjne korzyści. Jednak powoduje też poważne ryzyko w obszarze bezpieczeństwa i prywatności. Napastnicy mogą na przykład za pomocą internetu włamywać się do urządzeń działających w internecie rzeczy lub na serwery rejestrujące dane z takich urządzeń. Napastnik może na przykład przejąć kontrolę nad podłączoną do internetu lalką i bezpośrednio rozmawiać z dzieckiem albo włamać się do bazy danych z poufnymi informacjami o stanie zdrowia i aktywności rejestrowanymi za pomocą elektroniki do noszenia. Obawy dotyczące obszaru bezpieczeństwa i prywatności mogą podważyć zaufanie klientów niezbędne do wykorzystania pełnego potencjału technologii. Może to utrudniać upowszechnianie takich rozwiązań [FTC 2015].

1.2.1. Sieci dostępowe

Omówiliśmy już aplikacje i systemy końcowe znajdujące się na „obrzeżach sieci". Przyjrzyjmy się teraz sieciom dostępowym, czyli fizycznym łączom wiążącym dany system końcowy z pierwszym routerem (tak zwanym „routerem brzegowym") na drodze do dowolnego odległego systemu końcowego. Na rysunku 1.4 przedstawiono kilka typów łączy dostępowych poprowadzonych między systemem końcowym i routerem brzegowym. Łącza wyróżniono grubymi cieniowanymi liniami i określono, w jakim kontekście są używane (w sieci domowej, firmowej lub mobilnej o dużym zasięgu).



Rysunek 1.4. Sieci dostępowe

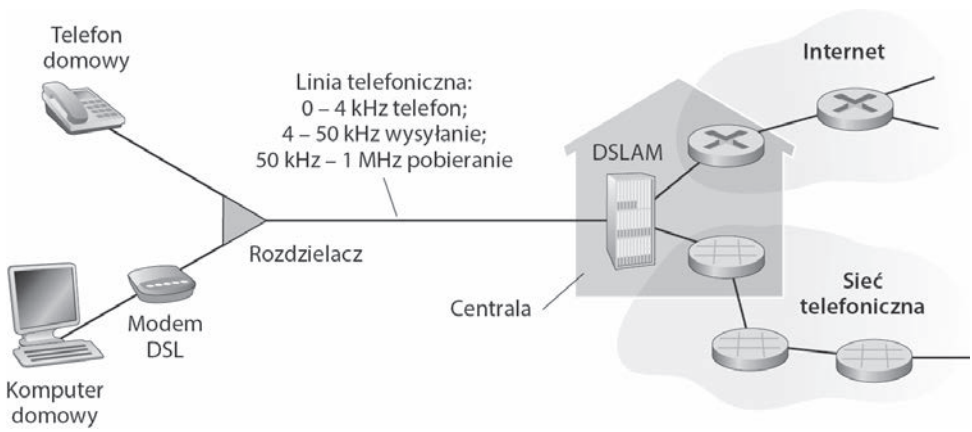
Dostęp do sieci w domu: DSL, sieci kablowe, FTTH, dostęp wdzwaniany i sieci satelitarne

W 2014 r. w krajach rozwiniętych ponad 78% gospodarstw domowych miało dostęp do internetu. Przodują pod tym względem Korea Południowa, Holandia, Finlandia i Szwecja, gdzie internet jest dostępny w ponad 80% gospodarstw domowych, przy czym w zdecydowanej większości są to szybkie połączenia szerokopasmowe [ITU 2015]. Z powodu takiej powszechności sieci dostępowych dla gospodarstw domowych zacznijmy przegląd od omówienia tej właśnie kategorii.

Obecnie dwa najpopularniejsze rozwiązania zapewniające szerokopasmowy dostęp do internetu to **linie DSL** i sieci kablowe. Dostęp za pośrednictwem urządzenia DSL zwykle jest oferowany przez lokalnego operatora telekomunikacyjnego. Dlatego jeśli użytkownik korzysta z linii DSL, operator telekomunikacyjny jest jednocześnie dostawcą ISP. Jak przedstawia to rysunek 1.5, każdy modem DSL za pomocą istniejącej linii telefonicznej (miedzianej skrętki; zob. punkt 1.2.2) wymienia dane z multiplekserem DSLAM, zwykle znajdującym się w centrali operatora. Domowy modem DSL przyjmuje dane cyfrowe i przekształca je na wysokie częstotliwości na potrzeby transferu linią telefoniczną do centrali operatora. Sygnały analogowe z wielu domów są w multiplekserze DSLAM przekształcane ponownie na format cyfrowy.

Linie telefoniczne przenoszą jednocześnie dane i sygnały telefoniczne zakodowane na trzech różnych częstotliwościach:

- pasmo o częstotliwości z przedziału od 50 kHz do 1 MHz — kanał pobierania o dużej szybkości;
- pasmo o częstotliwości z przedziału od 4 do 50 kHz — kanał wysyłania o średniej szybkości;
- pasmo o częstotliwości z przedziału od 0 do 4 kHz — zwykły dwukierunkowy kanał telefoniczny.

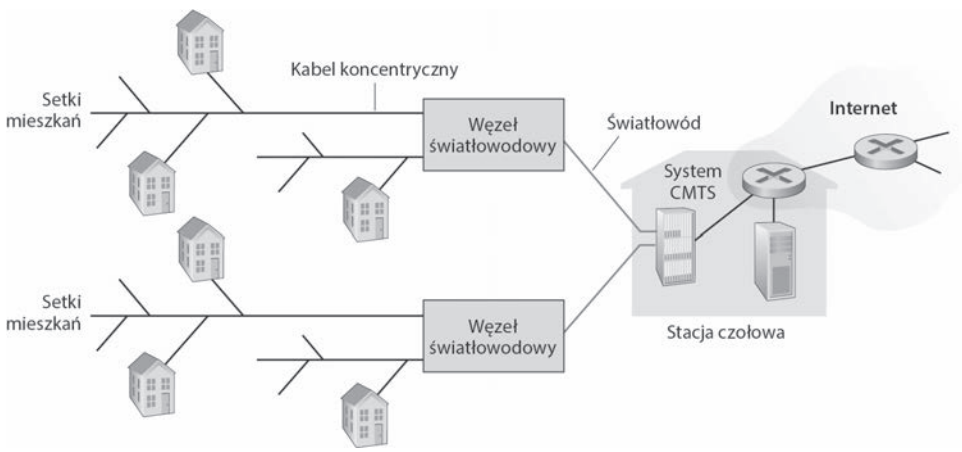


Rysunek 1.5. Dostęp do internetu za pomocą linii DSL

To podejście sprawia, że pojedyncza linia DSL działa jak trzy odrębne łącza, dlatego może jednocześnie obsługiwać połączenia telefoniczne i z internetem (technikę multipleksingu przez podział częstotliwości opisujemy w punkcie 1.3.1). Po stronie klienta rozdzielacz wyodrębnia dane i sygnały telefoniczne oraz kieruje informacje do modemu DSL. Po stronie operatora w centrali multiplekser DSLAM rozdziela dane i sygnały telefoniczne oraz wysyła informacje do internetu. Do jednego multipleksera tego typu podłączone są setki, a nawet tysiące gospodarstw domowych [Dischinger 2007].

W standardach DSL zdefiniowanych jest wiele szybkości transmisji, w tym pobieranie z szybkością 12 Mb/s i przesyłanie z szybkością 1,8 Mb/s [ITU 1999] oraz pobieranie z szybkością 55 Mb/s i przesyłanie z szybkością 15 Mb/s [ITU 2006]. Oznacza to dostęp *asymetryczny*, ponieważ szybkość wysyłania i odbioru są różne. Rzeczywista prędkość pobierania i wysyłania może być niższa od podanych wartości, ponieważ dostawca DSL może ją celowo ograniczać, gdy ma zróżnicowaną ofertę (z różnymi szybkościami i cenami). Szybkość maksymalna jest też ograniczona odległością mieszkania od centrali, średnicą skrętki oraz poziomem zakłóceń elektrycznych. Inżynierowie zaprojektowali linie DSL specjalnie do przesyłu danych na niewielkie odległości między mieszkaniami i centralami. Zwykle jeśli mieszkanie znajduje się w odległości większej niż 10 – 15 kilometrów od centrali, użytkownik musi uzyskać dostęp do internetu w inny sposób.

Przy dostępie do internetu za pomocą modemów DSL i telefonicznych wykorzystywana jest infrastruktura telefoniczna lokalnej firmy telekomunikacyjnej, natomiast w przypadku **sieci kablowych** służy do tego infrastruktura operatora telewizji kablowej. Użytkownik może uzyskać dostęp do internetu od firmy oferującej telewizję kablową. Jak widać na rysunku 1.6, światłowody łączą stację nadawczą z sąsiadującymi z nią węzłami, od których do poszczególnych domów i mieszkań jest poprowadzony zwykły kabel koncentryczny. Każdy węzeł zwykle obsługuje od 500 do 5000 domostw. Ponieważ w tym systemie stosowane są zarówno światłowody, jak i kable koncentryczne, jest on często nazywany siecią **HFC** (ang. *hybrid fiber coax*, czyli hybrydowa sieć światłowodowo-koncentryczna).



Rysunek 1.6. Hybrydowa światłowodowo-koncentryczna sieć dostępowa

Dostęp do internetu za pomocą sieci kablowej wymaga użycia specjalnych modemów nazywanych modemami kablowymi. Takie modemy — podobnie jak modemy DSL — są zwykle zewnętrznym urządzeniem połączonym z domowym komputerem PC za pomocą portu Ethernet (Ethernet bardziej szczegółowo omówiono w rozdziale 6.). W stacji czołowej system **CMTS** (ang. *cable model termination system*) pełni podobną funkcję jak multiplekser DSLAM w sieciach DSL — przekształca sygnał analogowy przesyłany przez modemy kablowe z wielu mieszkań na format cyfrowy. Tego typu modemy dzielą sieć HFC na dwa kanały — pobierania i wysyłania. Tak jak w przypadku technologii DSL, dostęp jest zwykle asymetryczny, a kanał pobierania oferuje większą szybkość transmisji niż kanał wysyłania. W standardzie DOCSIS 2.0 (ang. *Data-Over-Cable Service Interface Specifications*) szybkość pobierania jest ustalona na 42,8 Mb/s, a szybkość wysyłania na 30,7 Mb/s. Jednak podobnie jak jest w sieciach DSL, w praktyce maksymalna szybkość może być nieosiągalna z powodu poziomu wykupionych usług lub niedoskonałości nośników.

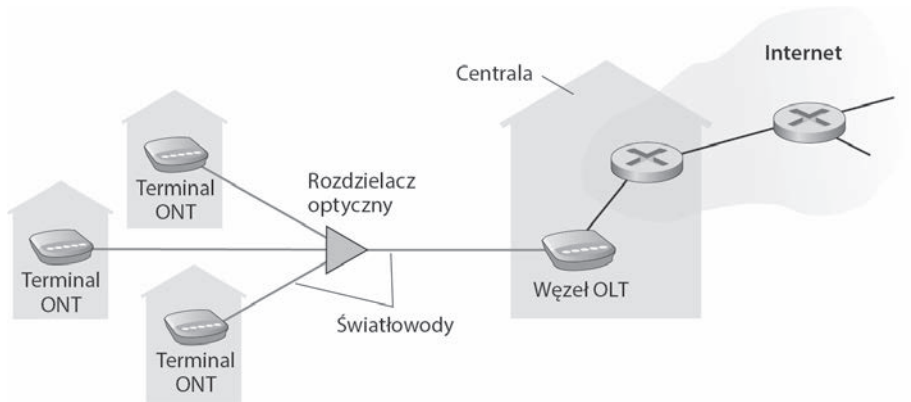
Ważną cechą dostępu kablowego jest to, że używany jest wspólny nośnik transmisyjny. Dokładniej mówiąc, każdy pakiet wysłany przez stację czołową jest przesyłany do każdego domu kanałem pobierania wszystkich łączy. Ponadto każdy pakiet wysłany przez domowy komputer jest transmitowany kanałem wysyłania do stacji czołowej. Z tego powodu, gdy jednocześnie kilku użytkowników będzie pobierało pliki wideo za pośrednictwem kanału pobierania, rzeczywista szybkość przesyłania plików u każdego z nich będzie znacznie mniejsza od oferowanej. Z kolei gdy aktywnych będzie tylko kilku użytkowników, którzy będą przeglądali strony internetowe, każdy z nich może wyświetlać dane z pełną szybkością pobierania. Wynika to stąd, że użytkownicy rzadko w tym samym czasie będą żądali pobrania stron internetowych. Ponieważ kanał wysyłania też jest wspólny, niezbędne jest użycie protokołu rozproszonego wieloźródłowego dostępu, który będzie koordynował transmisje i zapobiegał kolizjom (zagadnieniu kolizji bardziej dokładnie przyjrzymy się w rozdziale 6. podczas omawiania Ethernetu).

Choć w Stanach Zjednoczonych sieci DSL i kablowe zapewniają szerokopasmowy dostęp do sieci w ponad 85% gospodarstw domowych z dostępem do internetu, obecnie wprowadzaną technologią, która zapewnia jeszcze wyższą szybkość, jest **FTTH** (ang. *fiber to the home*, czyli światłowody w domu) [FTTH Council 2016]. Jak wskazuje na to nazwa, sieci FTTH są oparte na prostym podejściu — zapewnieniu bezpośredniego połączenia światłowodowego między mieszkaniem a centralą operatora. Obecnie w wielu krajach, w tym w Zjednoczonych Emiratach Arabskich, Korei Południowej, Hongkongu, Japonii, Singapurze, Tajwanie, Szwecji i na Litwie, stopa rozpowszechnienia sieci FTTH w gospodarstwach domowych przekracza 30% [FTTH Council 2016].

Istnieje kilka konkurencyjnych technologii dostarczania danych światłowodami z centrali do mieszkań. Najprostsze rozwiązania tego typu są oparte na bezpośrednich połączeniach światłowodowych. W tym modelu do każdego mieszkania biegnie jeden światłowód wychodzący z centrali. Częściej jednak światłowody wychodzące z centrali są współużytkowane przez wielu odbiorców. Dopiero w stosunkowo niedużej odległości od mieszkań światłowód jest rozdzielany na kable prowadzące do poszczególnych klientów. Podział ten odbywa się według dwóch konkurencyjnych architektur optycznych

sieci dostępowych — **AON** (ang. *Active Optical Network*, czyli aktywne sieci optyczne) i **PON** (ang. *Passive Optical Network*, czyli pasywne sieci optyczne). Model AON przypomina w działaniu Ethernet z przełączaniem, który omawiamy w rozdziale 6.

Tu pokrótce opisujemy architekturę PON wykorzystywaną w usłudze **FiOS** (ang. *Fiber Optic Service*) firmy Verizon. Rysunek 1.7 przedstawia sieć FTTH opartą na architekturze PON. W każdym mieszkaniu znajduje się terminal **ONT** (ang. *Optical Network Terminator*) podłączony przez dedykowany światłowód do pobliskiego rozdzielacza. Rozdzielacz łączy grupę gospodarstw domowych (zwykle poniżej 100) z jednym, współużytkowanym światłowodem biegnącym do węzła **OLT** (ang. *Optical Line Terminator*) w centrali firmy telekomunikacyjnej. Węzeł OLT przekształca sygnały między postaciami optyczną i elektryczną oraz łączy się z internetem za pośrednictwem routera operatora telekomunikacyjnego. W mieszkaniach użytkownicy podłączają domowy router (zwykle bezprzewodowy) do terminalu ONT i łączą się w ten sposób z internetem. W architekturze PON wszystkie pakiety wysyłane z węzła OLT do rozdzielacza są w nim powielane (podobnie jak w stacji czołowej telewizji kablowej).



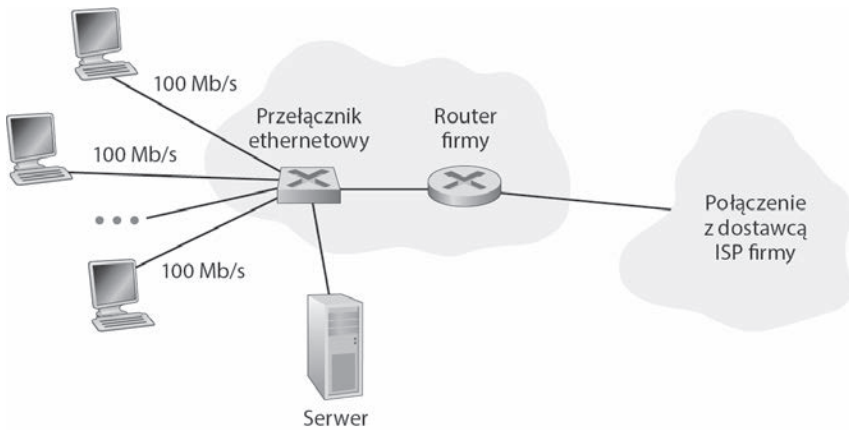
Rysunek 1.7. Dostęp do internetu za pomocą sieci FTTH

Sieci FTTH potencjalnie mogą zapewniać dostęp do internetu z szybkością na poziomie gigabitów na sekundę. Jednak większość dostawców ISP w tym modelu oferuje różną przepustowość, przy czym wyższa szybkość oczywiście więcej kosztuje. W 2011 r. w Stanach Zjednoczonych większość klientów korzystających z sieci FTTH mogła pobierać dane z szybkością ok. 20 Mb/s (w porównaniu z 13 Mb/s w sieciach kablowych i poniżej 5 Mb/s w sieciach DSL) [FTTH Council 2011b].

Stosuje się też dwie inne technologie dostarczania internetu do domów. W miejscach, gdzie sieci DSL, kablowe i FTTH są niedostępne (na przykład w niektórych obszarach wiejskich), można wykorzystać łącza satelitarne działające z szybkością powyżej 1 Mb/s. Taki dostęp oferują na przykład firmy StarBand i HughesNet. Dostęp wdzwaniany z użyciem tradycyjnych linii telefonicznych działa na tej samej zasadzie co dostęp w sieciach DSL — modem domowy łączy się linią telefoniczną z modemem dostawcy ISP. W porównaniu z DSL-em i innymi szerokopasmowymi sieciami dostępowymi sieci wdzwaniane są boleśnie wolne i działają z szybkością 56 kb/s.

Dostęp w przedsiębiorstwach (i domach) — Ethernet oraz Wi-Fi

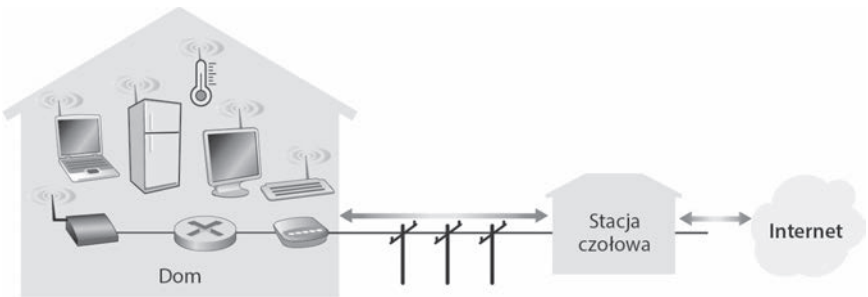
Sieć lokalna **LAN** (ang. *Local Area Network*) znajdująca się w budynkach firm i kampusów uniwersyteckich zwykle służy do łączenia systemów końcowych z routerem brzegowym. Choć istnieje wiele typów sieci lokalnych, w sieciach korporacyjnych i uniwersyteckich najbardziej rozpowszechnionym rozwiązaniem dostępowym jest technologia Ethernet. Jak ilustruje to rysunek 1.8, użytkownicy sieci ethernetowych korzystają z miedzianej skrętki do połączenia komputera z przełącznikiem ethernetowym (technologię tę opisujemy szczegółowo w rozdziale 6.). Przełącznik ethernetowy (lub grupa takich przełączników) jest z kolei podłączony do internetu. Szybkość połączenia z internetem za pomocą Ethernetu wynosi u klientów zwykle 100 Mb/s lub 1 Gb/s, natomiast na serwerach jest to 1, a nawet 10 Gb/s.



Rysunek 1.8. Dostęp do internetu za pomocą sieci Ethernet

Coraz więcej osób korzysta z internetu bezprzewodowego za pomocą laptopów, smartfonów, tabletów i innych rzeczy (zobacz ramkę na temat internetu rzeczy). W bezprzewodowej sieci lokalnej (LAN) korzystający z niej użytkownicy urządzeń bezprzewodowych wysyłają pakiety do punktów dostępowych podłączonych do sieci firmy i odbierają je z nich. Sieć firmy zwykle jest podłączona przewodowym ethernetem do przewodowego internetu. Użytkownicy bezprzewodowych sieci LAN zwykle muszą znajdować się w odległości nie większej niż kilkadziesiąt metrów od punktu dostępowego. Bezprzewodowe sieci lokalne oparte na technologii IEEE 802.11 (stosuje się również nieformalną nazwę Wi-Fi) są obecnie dostępne prawie wszędzie: w uczelniach, biurach firm, kawiarniach, domach i na lotniskach, a nawet w samolotach. W wielu miastach użytkownik stojący na rogu ulicy może znaleźć się w zasięgu 10, a nawet 20 stacji bazowych. W witrynie [wigle.net 2016] znajduje się globalna mapa stacji bazowych 802.11 wykrytych i zarejestrowanych przez osoby, którym taka działalność sprawia wielką satysfakcję. Technologia 802.11, którą dokładnie omówiono w rozdziale 7., oferuje łączną szybkość transmisji przekraczającą 100 Mb/s.

Choć sieci dostępowe oparte na Ethernecie i Wi-Fi pierwotnie opracowano na potrzeby organizacji (korporacji, uczelni), ostatnio stały się stosunkowo powszechnymi komponentami sieci domowych. Obecnie w wielu domach możliwy jest zarówno dostęp szerokopasmowy (modemy kablowe lub DSL), jak i tania technologia bezprzewodowej sieci lokalnej pozwalająca na tworzenie sieci domowych o dużych możliwościach [Edwards 2011]. Na rysunku 1.9 pokazano schemat typowej sieci domowej. Przykładowa sieć składa się z laptopa i komputera PC z okablowaniem, a także ze stacji bazowej (bezprzewodowy punkt dostępowy), która komunikuje się z komputerem i innymi urządzeniami bezprzewodowymi. Ponadto w sieci znajduje się modem kablowy (zapewniający szerokopasmowy dostęp do internetu) i router, który pośredniczy między modemem oraz stacją bazową i stacjonarnym komputerem PC. Sieć umożliwia domownikom dysponowanie szerokopasmowym dostępem do internetu. Jeden z nich może korzystać z laptopa, będąc w kuchni, na podwórku lub w sypialni.



Rysunek 1.9. Schemat typowej sieci domowej

Dostęp bezprzewodowy na większych obszarach: 3G i LTE

Użytkownicy coraz częściej korzystają ze sprzętu takiego jak iPhone lub urządzenia z systemem Android do przesyłania wiadomości, udostępniania zdjęć w sieciach społecznościowych, oglądania filmów i strumieniowego przesyłania muzyki w trakcie przemieszczania się. Te urządzenia wykorzystują infrastrukturę bezprzewodową używaną na potrzeby telefonii komórkowej, aby przesyłać i odbierać pakiety za pośrednictwem stacji bazowej zarządzanej przez dostawcę sieci komórkowej. Różnicą w porównaniu z sieciami Wi-Fi jest to, że odległość użytkownika od stacji bazowej nie może przekraczać kilkudziesięciu kilometrów (a nie kilkudziesięciu metrów).

Firmy telekomunikacyjne zainwestowały bardzo dużo środków w sieci bezprzewodowe trzeciej generacji (3G), które zapewniają bezprzewodowy dostęp do internetu z wykorzystaniem przełączania pakietów na dużym obszarze z szybkością przekraczającą 1 Mb/s. Jednak obecnie wprowadzane są jeszcze szybsze technologie dostępu na dużym obszarze — sieci czwartej generacji (4G). Technologia **LTE** (od ang. *long-term evolution*; jest to kandydat na nagrodę roku za najgorszy akronim) ma źródła w technologii 3G, ale pozwala uzyskać szybkość na poziomie powyżej 10 Mb/s. We wdrożeniach komercyjnych sieci LTE odnotowano szybkość pobierania rzędu kilkudziesięciu Mb/s. Podstawowe zasady działania sieci bezprzewodowych i mobilnych, a także technologii Wi-Fi, 3G i LTE (oraz innych) są przedstawione w rozdziale 7.

1.2.2. Fizyczny nośnik

W poprzednim punkcie dokonaliśmy przeglądu kilku najważniejszych technologii dostępu do sieci stosowanych w internecie. Omawiając je, wspomnieliśmy również o używaniu fizycznego nośnika. Przykładowo, stwierdziliśmy, że technologia HFC korzysta z rozwiązania łączącego światłowód z kablem koncentrycznym. Napisaliśmy, że technologie DSL i Ethernet używają miedzianej skrętki, jak również, że mobilne sieci dostępowe wykorzystują widmo radiowe. W tej części rozdziału dokonamy krótkiego przeglądu tych i innych nośników transmisyjnych powszechnie spotykanych w internecie.

Aby określić, co rozumie się przez fizyczny nośnik, opiszemy krótki żywot bita. Pod uwagę weźmy bit przemieszczający się z jednego systemu końcowego, przez serię łączy i routerów do innego systemu końcowego. Ten nieszczęsny bit jest transmitowany wiele, ale to wiele razy! Najpierw przesyła go źródłowy system końcowy. Chwilę potem bit jest odbierany przez pierwszy router z kilku, który transmituje go do kolejnego routera. Ten odbiera bit i powtarza operację. A zatem bit podróżujący od miejsca źródłowego do docelowego przechodzi przez serię par złożonych z nadajnika i odbiornika. W przypadku każdej takiej pary bit jest przesyłany **fizycznym nośnikiem** w formie fal elektromagnetycznych lub impulsów optycznych. Fizyczny nośnik może przyjmować wiele kształtów i form, a ponadto nie musi być identycznego typu w przypadku poszczególnych par nadajnik-odbiornik, występujących na drodze pokonywanej przez bit. Przykładami fizycznych nośników są: miedziana skrętka, kabel koncentryczny, światłowód wielomodowy, a także naziemne i satelitarne widmo radiowe. Fizyczne nośniki zalicza się do dwóch kategorii — **przewodowe** i **bezprzewodowe**. W przypadku nośników przewodowych fale przemieszczają się wzdłuż ciągłego nośnika, takiego jak światłowód, miedziana skrętka lub kabel koncentryczny. W przypadku nośnika bezprzewodowego fale rozchodzą się w atmosferze i przestrzeni kosmicznej (dotyczy to na przykład bezprzewodowej sieci lokalnej lub cyfrowego kanału satelitarnego).

Zanim zajmiemy się charakterystykami różnego typu nośników, należy w skrócie wspomnieć o ich cenie. Rzeczywisty koszt fizycznego łącza (kabel z miedzi, światłowód itp.) często jest dość nieznaczny w porównaniu z innymi kosztami związanymi z siecią. W rzeczywistości koszty robocizny dotyczące instalacji fizycznego łącza mogą być o rzędy wielkości większe od kosztów materiałów. Z tego powodu wielu wykonawców instaluje skrętkę, światłowód i kabel koncentryczny w każdym pomieszczeniu budynku. Jeśli nawet początkowo będzie używany tylko jeden nośnik, są spore szanse na to, że inny nośnik może okazać się przydatny w najbliższej przyszłości. W takiej sytuacji zaoszczędzi się na braku konieczności ponownego prowadzenia dodatkowego okablowania.

Skrętka miedziana

Najtańszym i najczęściej używanym przewodowym nośnikiem transmisyjnym jest skrętka miedziana. Przez ponad sto lat skrętka była używana w sieciach telefonicznych. W rzeczywistości ponad 99% połączeń kablowych poprowadzonych od aparatu telefonicznego do lokalnego przełącznika jest wykonanych ze skrętki miedzianej.

Większość z nas widziała skrętkę w domach (własnych albo rodziców lub dziadków) i miejscu pracy. Skrętka składa się z dwóch izolowanych przewodów miedzianych (każdy o grubości około 1 mm), które są ułożone w regularny spiralny wzór. Przewody są ze sobą skręcone, aby zredukować zakłócenia elektryczne wywoływane przez bliskość podobnych par. Zwykle kilka par przewodów jest umieszczanych w jednym kablu i otaczanych ochronnym ekranem. Para przewodów tworzy pojedyncze łącze komunikacyjne. **Skrętka nieekranowana** (ang. *Unshielded Twisted Pair* — **UTP**) jest powszechnie stosowana w sieciach komputerowych znajdujących się w budynku, czyli sieciach lokalnych. Obecnie szybkości transferu danych w przypadku sieci lokalnych używających skrętki zawierają się w przedziale od 10 Mb/s do 1 Gb/s. Szybkości, które mogą być uzyskiwane, zależą od grubości przewodu i odległości między nadajnikiem i odbiornikiem.

Gdy w latach 80. pojawiła się technologia światłowodowa, wiele osób zlekceważyło skrętkę, ponieważ oferuje stosunkowo niewielkie szybkości transmisji. Niektórzy uznali nawet, że światłowody powinny całkowicie zastąpić skrętkę. Jednak skrętka nie tak łatwo ustępowała pola. Nowsze technologie skrętki, takie jak kategoria UTP 6a, pozwalają osiągnąć szybkość transmisji 1 Gb/s przy maksymalnej odległości wynoszącej kilkaset metrów. Ostatecznie skrętka zaczęła pełnić rolę dominującego rozwiązania dla bardzo szybkich sieci lokalnych.

Jak wspomniano w punkcie dotyczącym sieci dostępowych, skrętka jest też powszechnie wykorzystywana na potrzeby oferowania dostępu do internetu prywatnym użytkownikom. Stwierdziliśmy, że modem telefoniczny za pośrednictwem skrętki pozwala uzyskać maksymalną szybkość 56 kb/s. Ponadto technologia DSL wykorzystująca skrętkę umożliwiła prywatnym użytkownikom łączenie się z internetem z szybkościami na poziomie dziesiątków MB/s (gdy domostwa znajdują się blisko modemu dostawcy ISP).

Kabel koncentryczny

Kabel koncentryczny podobnie jak skrętka jest złożony z dwóch miedzianych przewodów. Jednak są one prowadzone koncentrycznie, a nie równolegle. Dzięki takiemu rozwiązaniu, a także specjalnemu izolowaniu i ekranowaniu kabel koncentryczny może oferować duże szybkości transmisji. Kabel koncentryczny jest dość rozpowszechniony w systemach telewizji kablowych. Jak wcześniej wspomniano, w ostatnim czasie w tego typu systemach zastosowano modemy kablowe, które prywatnym użytkownikom zapewniają dostęp do internetu z szybkościami rzędu dziesiątków Mb/s. W przypadku telewizji kablowej i kablowego dostępu internetowego nadajnik przesuwa sygnał cyfrowy na określone pasmo częstotliwości, a następnie uzyskany sygnał analogowy jest wysyłany do jednego lub kilku odbiorników. Kabel koncentryczny może pełnić rolę przewodowego **wspólnego nośnika**. Dokładniej mówiąc, kilka systemów końcowych może zostać bezpośrednio podłączonych do kabla. Każdy taki system odbierze dowolne dane przesyłane przez pozostałe systemy.

Światłowód

Światłowód jest cienkim i elastycznym nośnikiem, który przenosi impulsy świetlne. Każdy impuls reprezentuje bit. Pojedynczy światłowód jest w stanie oferować niebywałe szybkości transmisji, które maksymalnie mogą wynosić dziesiątki, a nawet setki gigabitów na sekundę. Światłowody są odporne na zakłócenia elektromagnetyczne. Na odległości do 100 kilometrów cechują się bardzo niewielkim tłumieniem sygnału, a ponadto bardzo trudno się do nich „podpiąć". Wymienione cechy sprawiły, że światłowody stały się preferowanym przewodowym nośnikiem transmisyjnym, zwłaszcza w przypadku łączy międzykontynentalnych. Aktualnie wiele długodystansowych sieci telefonicznych istniejących w Stanach Zjednoczonych i innych państwach wykorzystuje wyłącznie światłowody. Są one też rozpowszechnione w sieci szkieletowej internetu. Jednak wysoki koszt urządzeń optycznych, takich jak transmitery, odbiorniki i przełączniki, zmniejsza tempo wdrażania technologii światłowodowej w przypadku mniejszych odległości (dotyczy to sieci lokalnych lub dostępu sieciowego uzyskiwanego z domostw prywatnych użytkowników). Szybkości łączy w standardzie **OC** (ang. *Optical Carrier*) wynoszą od 51,8 Mb/s do 39,8 Gb/s. Ich specyfikacje często są podawane w systemie OC-*n*, gdzie szybkość łącza wynosi *n*·51,8 Mb/s. Obecnie stosowane są standardy OC-1, OC-3, OC-12, OC-24, OC-48, OC-96, OC-192 i OC-768. W źródłach [Mukherjee 2006, Ramaswami 2010] można znaleźć omówienie różnych aspektów dotyczących sieci optycznych.

Naziemne kanały radiowe

Kanały radiowe przenoszą sygnały za pośrednictwem widma elektromagnetycznego. Jest to atrakcyjny nośnik, ponieważ nie wymaga prowadzenia żadnego okablowania, radzi sobie ze ścianami, zapewnia łączność użytkownikom mobilnym i ewentualnie może być użyty do przesyłania sygnału na duże odległości. Charakterystyki kanału radiowego w dużym stopniu zależą od środowiska propagacji i odległości, na jaką sygnał będzie przenoszony. Ze środowiskiem są związane takie kwestie jak utrata ścieżki i „zacienianie" (osłabianie sygnału w trakcie jego przesyłania spowodowane przez przeszkody znajdujące się na drodze sygnału lub w jego pobliżu), zanikanie wielodrożne (spowodowane przez odbicia sygnału od przeszkód) i zakłócenia (wywołane przez inne kanały radiowe lub sygnały elektromagnetyczne).

Naziemne kanały radiowe można ogólnie zaklasyfikować do dwóch grup. Pierwsze działają na bardzo krótkie odległości (na przykład jednego lub dwóch metrów). Są to kanały wykorzystywane lokalnie, które zwykle swoim zasięgiem obejmują od 10 do kilkuset metrów. Z kolei do drugiej grupy zaliczają się kanały używane na większych odległościach liczących dziesiątki kilometrów. Urządzenia osobiste, na przykład słuchawki bezprzewodowe, klawiatury i urządzenia medyczne, działają na krótkie dystanse. Z lokalnych kanałów radiowych korzystają technologie LAN omówione w punkcie 1.2.2. Technologie oparte na dostępie do internetu za pośrednictwem sieci komórkowych stosują kanały radiowe zaliczane do drugiej grupy. Kanały radiowe zostaną szczegółowo opisane w rozdziale 7.

Satelitarne kanały radiowe

Satelita komunikacyjny łączy ze sobą dwa lub więcej znajdujących się na ziemi mikrofalowych transmiterów lub odbiorników nazywanych też stacjami naziemnymi. Satelita odbiera transmisje na jednym paśmie częstotliwości, a następnie regeneruje sygnał za pomocą wzmacniacza (omówiony niżej) i przesyła sygnał, korzystając z innej częstotliwości. W komunikacji są stosowane dwa typy satelitów — **geostacjonarne** i **niskoorbitalne** [Wiki Satellite 2016].

Satelity geostacjonarne cały czas pozostają nad tym samym miejscem Ziemi. Taką stacjonarność osiąga się przez umieszczenie satelity na orbicie położonej na wysokości 36 000 kilometrów nad powierzchnią Ziemi. Tak duża odległość, która musi być pokonana przez sygnał (od stacji naziemnej do satelity i z powrotem), powoduje znaczne opóźnienie jego propagacji wynoszące 280 ms. Niemniej jednak łącza satelitarne, które mogą działać z szybkościami równymi setkom megabitów na sekundę, są często wykorzystywane w obszarach, gdzie nie jest możliwy dostęp do internetu za pomocą modemów DSL i kablowych.

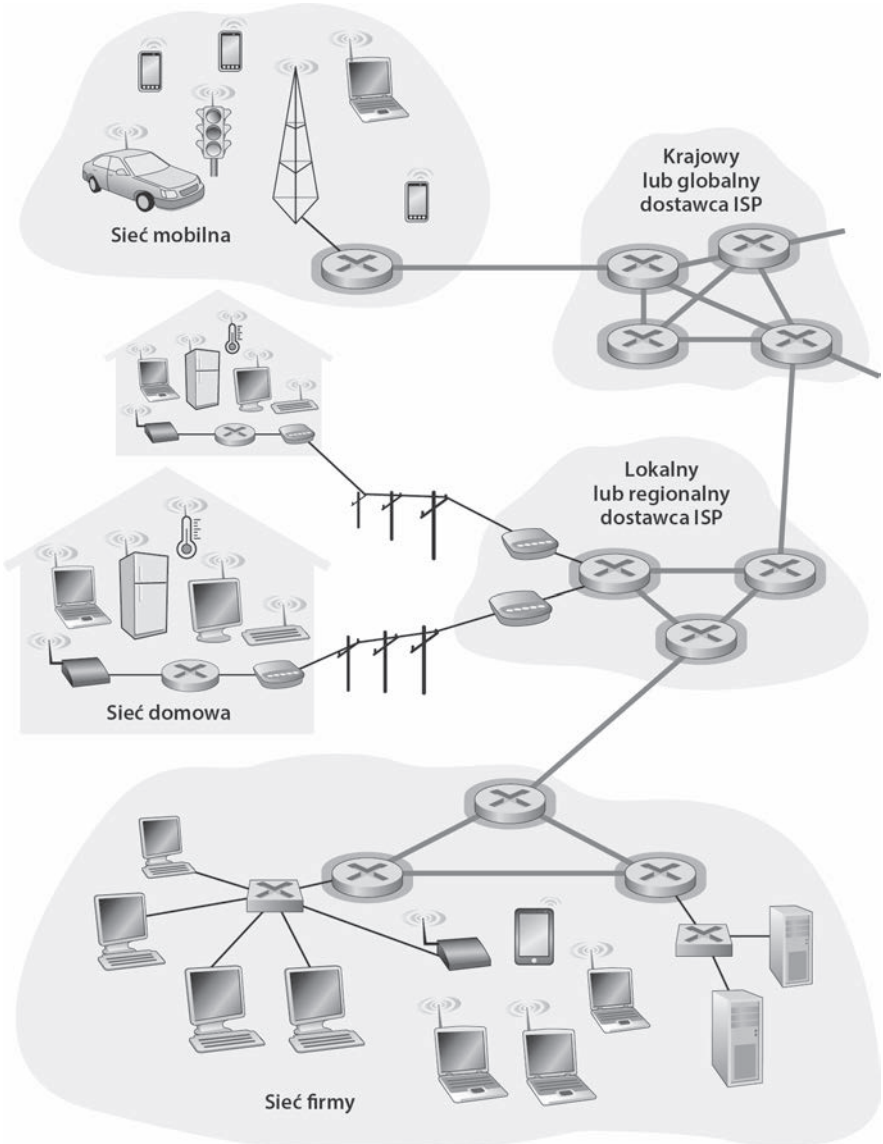
Satelity niskoorbitalne są zlokalizowane znacznie bliżej Ziemi i nie znajdują się cały czas nad jednym jej punktem. Tego typu satelity obracają się wokół Ziemi, podobnie jak Księżyc, i mogą komunikować się ze sobą, jak również ze stacjami naziemnymi. Aby zapewnić niezmienne obejmowanie przez satelitę określonego obszaru, na orbicie musi być umieszczonych wiele satelitów. Aktualnie jest wdrażanych wiele systemów komunikacyjnych używających satelitów niskoorbitalnych. Być może w przyszłości satelity niskoorbitalne zostaną zastosowane na potrzeby oferowania dostępu internetowego.

1.3. Rdzeń sieci

Po omówieniu obrzeży internetu zajmijmy się rdzeniem sieci, czyli systemem przełączników pakietów i odwodów, które wiążą ze sobą systemy końcowe podłączone do internetu. Na rysunku 1.10 grubymi cieniowanymi liniami wyróżniono rdzeń sieci.

1.3.1. Przełączanie pakietów

W celu zrealizowania zadania aplikacje rozproszone wymieniają się **komunikatami**. Komunikaty mogą zawierać wszystko, czego zażąda twórca protokołu. Komunikaty mogą pełnić funkcję kontrolną (komunikaty „Cześć" zastosowane w przykładzie prezentującym komunikowanie się dwóch osób; zob. rysunek 1.2) lub przechowywać dane, takie jak wiadomość pocztowa, obraz JPEG lub plik audio MP3. Aby przesłać komunikat ze źródłowego systemu końcowego do docelowego, nadawca dzieli długie wiadomości na mniejsze porcje danych nazywane **pakietami**. Między nadawcą i odbiorcą każdy pakiet jest przesyłany przy użyciu łączy komunikacyjnych i **przełączników pakietów** (wśród nich dominują dwa typy — **routery** i **przełączniki warstwy łącza danych**). Pakiety są przesyłane przez każde łącze komunikacyjne z szybkością równą jego *maksymalnej* szybkości. Jeśli więc źródłowy system końcowy lub przełącznik pakietów wysyła pakiet o długości D bitów połączeniem o szybkości transmisji S bitów na sekundę, czas transferu pakietu wynosi D/S sekund.



Rysunek 1.10. Rdzeń sieci

Transmisja buforowana

Większość przełączników pakietów na wejściu łącza stosuje **transmisję buforowaną** (ang. *store-and-forward transmission*). Tego typu transmisja powoduje, że przełącznik musi odebrać cały pakiet, zanim będzie mógł rozpocząć wysyłanie jego pierwszego bitu do łącza wyjściowego. Aby lepiej zrozumieć transmisję buforowaną, rozważ prostą sieć obejmującą dwa systemy końcowe połączone jednym routerem (zob.

rysunek 1.11). Router zwykle ma wiele łączy wejściowych, ponieważ jego zadaniem jest przekazywanie przychodzących pakietów do łącza wyjściowego. W tym prostym przykładzie router ma łatwe zadanie, polegające na przekazywaniu przesyłanych pakietów z jednego łącza (wejściowego) do jedynego innego łącza. Tu w źródle znajdują się trzy pakiety po *D* bitów, które należy przesłać do docelowej lokalizacji. W momencie przedstawionym na rysunku 1.11 źródło przesłało fragment pakietu 1., a początek tego pakietu dotarł już do routera. Ponieważ router stosuje transmisję buforowaną, nie może przesłać otrzymanych bitów. Zamiast tego musi najpierw zapisać bity pakietu. Dopiero po tym, jak router odbierze *wszystkie* bity pakietu, może zacząć przesyłać (transmitować) pakiet łączem wyjściowym. Aby uzyskać lepszy wgląd w transmisję buforowaną, obliczmy, ile czasu upłynie od rozpoczęcia wysyłania pakietu przez punkt źródłowy do momentu odebrania całego pakietu przez punkt docelowy. Pomijamy tu opóźnienie propagacji, czyli czas potrzebny na transfer bitów przewodem z prędkością bliską prędkości światła; to zagadnienie jest opisane w punkcie 1.4. Punkt źródłowy rozpoczyna transmisję w czasie 0. Po czasie *D/S* sekund źródło prześle cały pakiet, a ten zostanie odebrany i zapisany w routerze (ponieważ pomijamy opóźnienie propagacji). Tak więc po czasie *D/S* sekund router otrzymał cały pakiet i może zacząć przesyłać go łączem wyjściowym do punktu docelowego. Po 2*D/S* sekund router przesłał cały pakiet, a ten został odebrany w punkcie docelowym. Całkowite opóźnienie wynosi więc 2*D/S*. Gdyby przełącznik zaczął przekazywać bity bezpośrednio po ich otrzymaniu (bez oczekiwania na cały pakiet), całkowite opóźnienie wyniosłoby *D/S*, ponieważ bity nie są wtedy przechowywane w routerze. Jednak, co opisano w podrozdziale 1.4, routery przed przekazaniem pakietu muszą go w całości odebrać, zapisać i *przetworzyć*.

Rysunek 1.11. Przełączanie pakietów z transmisją buforowaną

Obliczmy teraz czas, jaki upływa od momentu rozpoczęcia przesyłania pierwszego pakietu przez punkt źródłowy do chwili otrzymania wszystkich trzech pakietów przez punkt docelowy. Tak jak wcześniej w czasie *D/S* router zaczyna przekazywać pierwszy pakiet. W tym samym momencie punkt źródłowy rozpoczyna przesyłanie drugiego pakietu, ponieważ właśnie skończył transfer pierwszego. Tak więc w czasie 2*D/S* punkt docelowy otrzymał pierwszy pakiet, a router — drugi pakiet. Podobnie w czasie 3*D/S* punkt docelowy pobrał dwa pierwsze pakiety, a router — trzeci pakiet. W czasie 4*D/S* punkt docelowy otrzymał wszystkie trzy pakiety.

Rozważmy teraz ogólny przypadek wysyłania jednego pakietu z punktu źródłowego do docelowego ścieżką obejmującą N łączy, każde o szybkości S (między punktami źródłowym a docelowym jest więc N–1 routerów). Po zastosowaniu tej samej logiki co wcześniej widać, że opóźnienie między punktami końcowymi wynosi:

$$o_{\text{węz–węz}} = N\frac{D}{S} \tag{1.1}$$

Możesz teraz zechcieć obliczyć opóźnienie dla P pakietów przesyłanych przez N łączy.

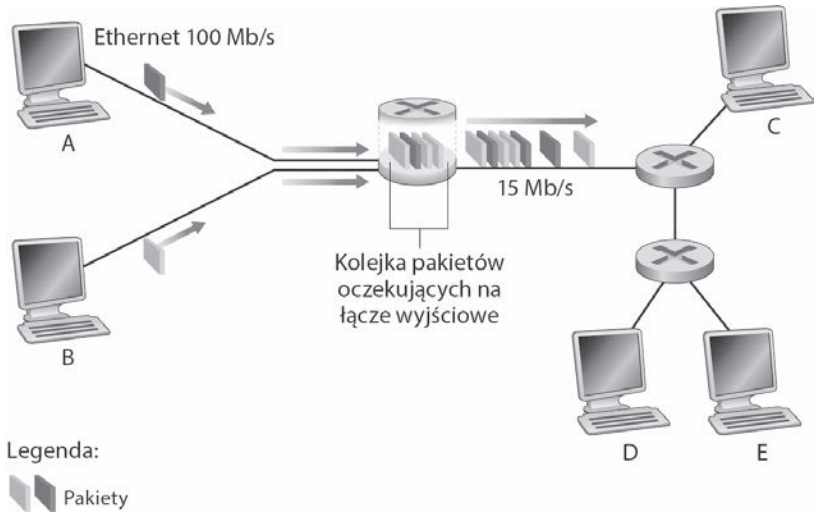
Opóźnienia kolejkowania i utrata pakietów

Do każdego przełącznika pakietów jest podłączonych wiele łączy. Każdemu takiemu łączu przełącznik przydziela **bufor wyjściowy** (nazywany też **kolejką wyjściową**) przechowujący pakiety, które router prześle do łącza. W przypadku przełączania pakietów bufory wyjściowe odgrywają kluczową rolę. Jeśli odebrany pakiet musi zostać przesłany łączem, które akurat jest zajęte transmisją innego pakietu, otrzymany pakiet będzie musiał poczekać w buforze wyjściowym. W związku z tym poza opóźnieniami związanymi z transmisją buforowaną na pakiet będą miały też wpływ **opóźnienia kolejkowania** w buforze wyjściowym. Tego typu opóźnienia są zmienne i zależą od poziomu przeciążenia sieci. Ponieważ pojemność bufora jest ograniczona, przy odbiorze pakietu może się okazać, że bufor jest całkowicie zapełniony innymi pakietami oczekującymi na transmisję. W tym przypadku dojdzie do **utraty pakietu**. Zostanie odrzucony otrzymany pakiet lub jeden z tych, które są już kolejkowane.

Na rysunku 1.12 przedstawiono prostą sieć z przełączaniem pakietów. Podobnie jak na rysunku 1.11 pakiety są reprezentowane przez trójwymiarowe płytki. Szerokość płytki identyfikuje liczbę bitów pakietu. Na rysunku wszystkie pakiety są identycznej szerokości, a zatem również długości. Załóżmy, że hosty A i B wysyłają pakiety do hosta E. Hosty A i B najpierw swoje pakiety przesyłają łączami Ethernet o szybkości 100 Mb/s do pierwszego przełącznika pakietów. Przełącznik kieruje pakiety do łącza o szybkości 15 Mb/s. Jeśli szybkość, z jaką pakiety są transmitowane do przełącznika, przekracza jego szybkość przekazywania pakietów za pomocą łącza wyjściowego 15 Mb/s, dojdzie do przeciążenia, ponieważ przed przesłaniem pakietów łączem będą one kolejkowane w jego buforze wyjściowym. Na przykład: jeśli hosty A i B wysyłają serie po pięć pakietów jeden po drugim, większość tych pakietów przez pewien czas będzie oczekiwać w kolejce. Jest to sytuacja analogiczna do codziennych scenariuszy — jak wtedy, gdy czekamy w kolejce lub w punkcie poboru opłat. Opóźnienia związane z kolejkowaniem opisujemy szczegółowo w podrozdziale 1.4.

Tablice przekazywania i protokoły routingu

Wcześniej wyjaśniliśmy, że router przyjmuje pakiet przesłany jednym z łączy komunikacyjnych i przekazuje go do następnego węzła za pomocą innego łącza. Jak jednak router określa, którym łączem powinien wysłać pakiet? W różnych sieciach odbywa się to w odmienny sposób. W niniejszym wprowadzającym rozdziale opiszemy jedno z popularnych rozwiązań — model stosowany w internecie.



Rysunek 1.12. Przełączanie pakietów

W internecie każdy pakiet przesyłany w sieci ma w nagłówku swój docelowy adres — adres IP. Gdy źródłowy system końcowy chce przesłać pakiet do docelowego systemu końcowego, system źródłowy zapisuje w nagłówku pakietu docelowy adres IP. Ten adres — podobnie jak adres pocztowy — ma budowę hierarchiczną. Kiedy pakiet trafia do routera w sieci, ten sprawdza adres docelowy pakietu i przekazuje go do przyległego routera. Wyjaśnijmy to bardziej szczegółowo. Każdy router ma **tablicę przekazywania**, w której adresy docelowe (lub ich części) są odwzorowane na łącza wyjściowe. Kiedy pakiet dotrze do routera, ten sprawdzi adres i wyszuka go w tablicy przekazywania, aby znaleźć odpowiednie łącze wyjściowe. Następnie router skieruje pakiet do tego łącza.

Proces routingu między dwoma punktami można przyrównać do kierowcy, który zamiast korzystać z mapy, woli zapytać kogoś o drogę. Dla przykładu załóżmy, że Joe jedzie z miasta Philadelphia do Orlando w stanie Floryda, z zamiarem zlokalizowania domu o numerze 156 znajdującego się przy ulicy Lakeside Drive. Joe najpierw udaje się na pobliską stację benzynową, aby dowiedzieć się, jak dojechać do wspomnianej ulicy. Pracownik stacji zwraca uwagę na część adresu dotyczącą stanu Floryda i mówi Joemu, że musi wjechać na międzystanową autostradę I-95 South, od której prowadzi zjazd do kolejnej stacji. Ponadto radzi Joemu, aby po wjechaniu do stanu Floryda zapytał kogoś o drogę. Joe jedzie autostradą I-95 South aż do chwili dotarcia do Jacksonville w stanie Floryda. Tam pyta o drogę pracownika kolejnej stacji benzynowej. Ten przygląda się części adresu dotyczącego Orlando i mówi Joemu, że powinien dalej jechać autostradą I-95 aż do Daytona Beach i tam zapytać kogoś o drogę. W Daytona Beach kolejny pracownik stacji patrzy na część adresu związaną z Orlando i informuje Joego, że autostradą I-4 dojedzie bezpośrednio do tego miasta. Joe rusza autostradą I-4 i zjeżdża z niej do Orlando. Udaje się na następną stację benzynową. Tym razem jej

pracownik zwraca uwagę na część adresu z ulicą Lakeside Drive i pokazuje Joemu drogę, którą musi pojechać, aby się tam dostać. Po dotarciu na ulicę Lakeside Drive Joe pyta napotkane dziecko na rowerze, gdzie znajduje się szukany dom. Dziecko przygląda się części adresu zawierającej numer 156 i wskazuje dom. Joe wreszcie lokalizuje cel swojej podróży. W tej analogii pracownicy stacji benzynowych i dziecko na rowerze odpowiadają routerom. Ich tablice przekazywania (mózgi) zostały skonfigurowane w wyniku wieloletnich doświadczeń.

Wyjaśniliśmy właśnie, że router stosuje adres docelowy pakietu do przeszukiwania tablicy przekazywania i ustalania odpowiedniego łącza wyjściowego. Jednak to stwierdzenie prowadzi do następnego pytania — jak powstają takie tablice? Czy są konfigurowane ręcznie w każdym routerze, czy w internecie stosowane jest bardziej automatyczne rozwiązanie? To zagadnienie omawiamy szczegółowo w rozdziale 5. Jednak aby zaostrzyć apetyt Czytelników, nadmieńmy, że w internecie funkcjonuje wiele specjalnych **protokołów routingu** służących do automatycznego konfigurowania tablic przekazywania. Protokół routingu może na przykład określić najkrótszą ścieżkę od każdego routera do lokalizacji docelowej i wykorzystać efekty tej operacji do utworzenia tablic przekazywania w routerach.

Jak można sprawdzić, jaką trasę pokonują w internecie pakiety między dwoma punktami? Proponujemy poświęcić trochę czasu na zapoznanie się z narzędziem *Traceroute*. Wystarczy przejść pod adres *http://www.traceroute.org*, wybrać system źródłowy w konkretnym kraju i prześledzić trasę od tego systemu do Twojego komputera (więcej na jego temat można znaleźć w podrozdziale 1.4).

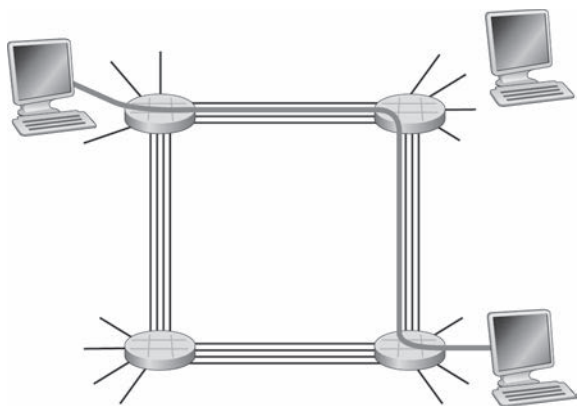
1.3.2. Przełączanie obwodów

Są dwie podstawowe metody przesyłania danych w sieci łączy i przełączników — **przełączanie obwodów** i **przełączanie pakietów**. Po omówieniu w poprzednim punkcie sieci z przełączaniem pakietów pora skupić się na sieciach z przełączaniem obwodów.

W sieciach z przełączaniem obwodów zasoby wymagane przez ścieżkę (bufory, szybkość transmisji łącza) zapewniającą komunikację między systemami końcowymi są *rezerwowane* na czas trwania sesji. W sieciach z przełączaniem pakietów zasoby *nie są rezerwowane*. Komunikaty przesyłane w ramach sesji korzystają z zasobów na żądanie. W efekcie mogą być zmuszone do czekania w kolejce na uzyskanie dostępu do łącza komunikacyjnego. W celu przytoczenia prostej analogii rozważmy dwie restauracje, z których jedna wymaga rezerwacji, natomiast druga nie żąda ich, ani też ich nie przyjmuje. W przypadku pierwszej restauracji przed wyjściem z domu trzeba będzie zadzwonić. Jednak dzięki temu po pojawieniu się w restauracji w zasadzie będzie można od razu wezwać kelnera i zamówić danie. W przypadku drugiej restauracji nie musimy pamiętać o rezerwowaniu stolika. Jednak po przybyciu do niej możemy być zmuszeni do czekania na stolik, zanim będzie można przywołać kelnera.

Tradycyjne sieci telefoniczne są przykładem sieci z przełączaniem obwodów. Zastanówmy się nad tym, co się stanie, gdy jedna osoba będzie chciała wysłać dane (głos lub faks) do drugiej za pośrednictwem sieci telefonicznej. Zanim nadawca będzie w stanie wysłać informacje, w sieci musi zostać nawiązane *prawdziwe* połączenie między nim i odbiorcą, dla którego przełączniki znajdujące się na ścieżce ustanowionej między nadawcą i odbiorcą przechowują dane dotyczące stanu połączenia. W żargonie związanym z sieciami telefonicznymi tego typu połączenie jest nazywane **obwodem**. Gdy w sieci zostanie utworzony obwód, na czas połączenia w łączach sieciowych jest rezerwowana stała szybkość transmisji (odpowiadająca ułamkowi możliwości transmisyjnych każdego łącza). Ponieważ dla takiego połączenia między nadawcą i odbiorcą jest rezerwowana przepustowość, nadawca może przesyłać dane do odbiorcy z *gwarantowaną* stałą szybkością.

Na rysunku 1.13 zilustrowano sieć z przełączaniem obwodów. W przypadku tej sieci cztery przełączniki obwodów są ze sobą połączone za pomocą takiej samej liczby łączy. Każde łącze posiada *n* obwodów. W związku z tym każde łącze może obsługiwać *n* jednoczesnych połączeń. Hosty (na przykład komputery PC i stacje robocze) są bezpośrednio podłączone do jednego z przełączników. Gdy dwa hosty zamierzają się ze sobą komunikować, sieć ustanawia między nimi dedykowane **połączenie punkt-punkt**. Aby host A mógł wysłać komunikaty do hosta B, sieć musi najpierw dla każdego z dwóch łączy zarezerwować jeden obwód. W tym przykładzie dedykowane połączenie punkt-punkt używa drugiego obwodu w pierwszym łączu i czwartego obwodu w drugim łączu. Ponieważ każde łącze ma cztery obwody, dla każdego łącza używanego przez omawiane połączenie punkt-punkt przypada 1/4 przepustowości łącza na czas trwania połączenia. Na przykład: jeśli szybkość transmisji każdego łącza między przyległymi przełącznikami wynosi 1 Mb/s, to każde połączenie międzywęzłowe w sieci z przełączaniem obwodów otrzyma na własny użytek 250 kb/s.



Rysunek 1.13. Prosta sieć z przełączaniem obwodów złożona z czterech przełączników i czterech łączy

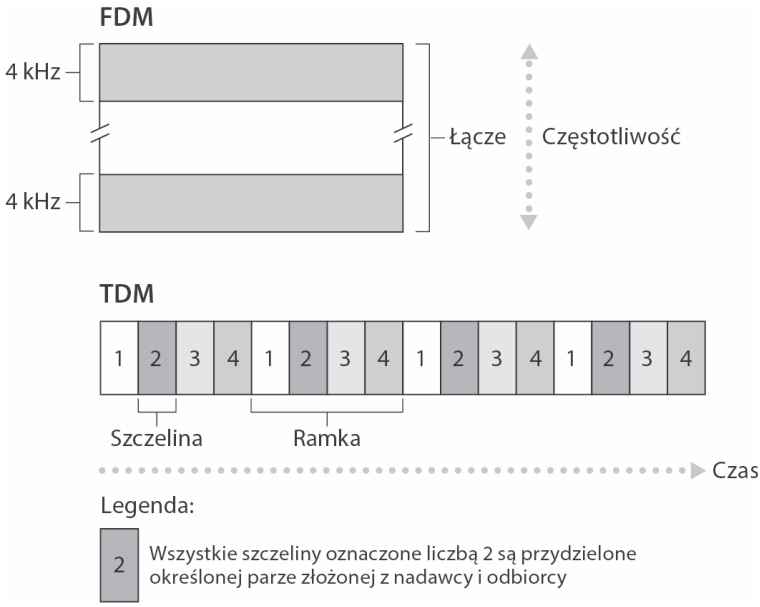
Rozważmy, co się stanie, gdy jeden host będzie chciał przesłać pakiet do innego hosta za pośrednictwem sieci z przełączaniem pakietów, na przykład internetu. Podobnie jak w przypadku przełączania obwodów pakiet jest przesyłany przy użyciu kilku łączy komunikacyjnych. Jednak w przypadku przełączania pakietów pakiet jest transferowany w sieci bez rezerwowania żadnej przepustowości. Jeśli jedno z łączy będzie przeciążone, ponieważ w tym samym czasie konieczne jest przesłanie za jego pośrednictwem innych pakietów, transmitowany pakiet będzie musiał poczekać w buforze znajdującym się po stronie nadawczej łącza transmisyjnego. W efekcie wystąpi opóźnienie. Choć w internecie są czynione wszelkie starania, aby pakiety dostarczać w odpowiednim czasie, nie ma na to żadnej gwarancji.

Multipleksowanie w sieciach z przełączaniem obwodów

Obwód w łączu jest implementowany za pomocą **multipleksowania z podziałem częstotliwości** (ang. *Frequency-Division Multiplexing* — **FDM**) lub **multipleksowania z podziałem czasu** (ang. *Time-Division Multiplexing* — **TDM**). W przypadku multipleksowania FDM widmo częstotliwości łącza jest dzielone między połączenia nawiązane za pośrednictwem tego łącza. Dokładniej mówiąc, łącze przydziela pasmo częstotliwości każdemu połączeniu na czas jego trwania. W sieciach telefonicznych pasmo to zwykle ma szerokość wynoszącą 4 kHz (inaczej 4000 herców lub 4000 cykli na sekundę). Szerokość pasma jest też nazywana **przepustowością**. Stacje radiowe FM również korzystają z multipleksowania FDM w celu współużytkowania widma częstotliwości z przedziału od 88 do 108 MHz (każda stacja otrzymuje określone pasmo częstotliwości).

W przypadku łącza z multipleksowaniem TDM czas jest dzielony na ramki o stałej długości. Każda ramka składa się z niezmiennej liczby szczelin czasowych. Gdy w sieci jest ustanawiane połączenie za pośrednictwem łącza, połączeniu jest przydzielana jedna szczelina czasowa każdej ramki. Szczeliny te są przeznaczone wyłącznie na potrzeby tego konkretnego połączenia. Udostępniona szczelina czasowa każdej ramki służy do transmisji danych połączenia.

Na rysunku 1.14 zilustrowano użycie multipleksowania FDM i TDM dla określonego łącza sieciowego obsługującego maksymalnie cztery obwody. W przypadku multipleksowania FDM domena częstotliwości jest segmentowana na cztery pasma, z których każde ma szerokość 4 kHz. W przypadku multipleksowania TDM domena czasowa jest dzielona na ramki, z których każda posiada cztery szczeliny czasowe. Każdemu obwodowi jest przydzielana ta sama dedykowana szczelina w powtarzających się ramkach TDM. Gdy stosuje się multipleksowanie TDM, szybkość transmisji obwodu jest równa liczbie ramek przypadających na sekundę pomnożonej przez liczbę bitów w szczelinie. Jeśli na przykład łącze przesyła 8000 ramek w ciągu sekundy i każda szczelina zawiera 8 bitów, szybkość transmisji obwodu wyniesie 64 kb/s.



Rysunek 1.14. W przypadku multipleksowania FDM każdemu obwodowi nieprzerwanie jest przydzielana część przepustowości. W przypadku multipleksowania TDM każdy obwód okresowo podczas krótkich odcinków czasu (identyfikowanych przez szczeliny) dysponuje całą dostępną przepustowością

Zwolennicy przełączania pakietów zawsze uważali, że przełączanie obwodów jest marnotrawstwem, ponieważ podczas **okresów ciszy** dedykowane obwody znajdują się w stanie bezczynności. Przykładowo, gdy ktoś zakończy rozmawiać przez telefon, bezczynne zasoby sieciowe (pasma częstotliwości lub szczeliny w łączach tworzących ścieżkę połączenia) nie mogą być wykorzystane przez inne utworzone połączenia. Kolejnym przykładem tego, jak zasoby mogą być niedostatecznie użytkowane, jest radiolog, który w celu uzyskania zdalnego dostępu do serii zdjęć rentgenowskich korzysta z sieci z przełączaniem obwodów. Radiolog nawiązuje połączenie, żąda pobrania zdjęcia, a następnie ogląda je i żąda kolejnego. Zasoby sieciowe są marnotrawione w czasie, gdy specjalista przegląda zdjęcia. Zwolennicy przełączania pakietów z przyjemnością zaznaczą, że definiowanie obwodów punkt-punkt i rezerwowanie dla nich przepustowości jest skomplikowane i wymaga zastosowania złożonego oprogramowania przetwarzającego sygnały, które koordynuje pracę przełączników znajdujących się na ścieżce biegnącej między dwoma węzłami.

Zanim zakończymy omawianie przełączania obwodów, przytoczymy liczbowy przykład, który powinien w większym stopniu przybliżyć to zagadnienie. Zastanówmy się, ile czasu zajmie wysłanie za pośrednictwem sieci z przełączaniem obwodów z hosta A do hosta B pliku liczącego 640 000 bitów. Załóżmy, że wszystkie łącza sieciowe korzystają z multipleksowania TDM oferującego 24 szczeliny i posiadają szybkość transmisji wynoszącą 1,536 Mb/s. Dodatkowo przyjmijmy, że 500 milisekund zajmuje

utworzenie obwodu punkt-punkt, jeszcze zanim host A będzie mógł rozpocząć wysyłanie pliku. Ile czasu zajmie przesłanie pliku? Ponieważ szybkość transmisji każdego obwodu wynosi (1,536 Mb/s)/24 = 64 kb/s, transmisja pliku zajmie 10 sekund (640 000 bitów)/(64 kb/s)). Do 10 sekund należy dodać czas tworzenia obwodu, co powoduje, że czas wysyłania pliku zwiększa się do 10,5 sekundy. Warto zauważyć, że czas transmisji jest niezależny od liczby łączy. Czas ten wyniesie 10 sekund, zarówno gdy obwód punkt-punkt będzie korzystał z jednego, jak i stu łączy (rzeczywiste opóźnienie występujące między dwoma punktami uwzględnia też opóźnienie propagacji; więcej o tym można znaleźć w podrozdziale 1.4).

Porównanie przełączania pakietów i obwodów oraz multipleksowanie statystyczne

Po omówieniu przełączania obwodów i pakietów dokonamy porównania obu rozwiązań. Krytycy przełączania pakietów często twierdzą, że ze względu na zmienne i nieprzewidywalne opóźnienia (chodzi głównie o opóźnienia kolejkowania) między dwoma punktami nie nadaje się ono do zastosowania w przypadku usług czasu rzeczywistego (na przykład rozmowy telefoniczne i wideokonferencje). Zwolennicy przełączania pakietów argumentują, że po pierwsze, w porównaniu z przełączaniem obwodów oferuje lepsze współużytkowanie przepustowości, a po drugie, przełączanie pakietów można wdrożyć prościej, efektywniej i taniej. Interesujące porównanie przełączania pakietów i przełączania obwodów znajduje się w tekście [Molinero-Fernandez 2002]. Ogólnie mówiąc, osoby, które nie lubią dokonywać rezerwacji w restauracjach, nad przełączanie obwodów przedkładają przełączanie pakietów.

Dlaczego przełączanie pakietów jest efektywniejsze? Przytoczmy prosty przykład. Załóżmy, że użytkownicy korzystają ze wspólnego łącza 1 Mb/s. Ponadto przyjmijmy, że każdy użytkownik na zmianę ma do czynienia z okresem aktywności (generowane są dane ze stałą szybkością 100 kb/s) i nieaktywności (nie są generowane żadne dane). Załóżmy dodatkowo, że użytkownik jest aktywny tylko przez 10% czasu (przez pozostałe 90% tylko pije kawę). W przypadku przełączania obwodów przepustowość 100 kb/s musi być cały czas *zarezerwowana* dla *każdego* użytkownika. Przykładowo, gdy korzysta się z przełączania obwodów z multipleksowaniem TDM, po podzieleniu jednosekundowej ramki na 10 szczelin czasowych 100 ms, każdemu użytkownikowi zostanie przydzielona jedna szczelina ramki.

A zatem łącze może jednocześnie obsłużyć tylko 10 użytkowników (= 1 Mb/s/ 100 kb/s). W przypadku przełączania pakietów prawdopodobieństwo, że określony użytkownik jest aktywny, wynosi 0,1, czyli 10%. Jeśli istnieje 35 użytkowników, prawdopodobieństwo, że w tym samym czasie aktywnych jest 11 lub więcej użytkowników w przybliżeniu wynosi 0,0004 (w problemie 8. zamieszczonym na końcu rozdziału wyjaśniono, w jaki sposób uzyskuje się wartość prawdopodobieństwa). Gdy jednocześnie aktywnych jest 10 lub mniej użytkowników (prawdopodobieństwo takiej sytuacji wynosi 0,9996), całkowita szybkość, z jaką są wysyłane dane będzie mniejsza lub równa wyjściowej

szybkości łącza wynoszącej 1 Mb/s. A zatem, jeśli aktywnych jest 10 lub mniej osób, wysyłane przez nie pakiety w zasadzie będą transmitowane przez łącze bez opóźnień, tak jak w przypadku przełączania obwodów. Gdy w danej chwili aktywnych będzie ponad 10 użytkowników, łączna szybkość, z jaką pakiety są wysyłane, przekroczy szybkość łącza. W efekcie zacznie się tworzyć kolejka wyjściowa (będzie to trwało do momentu, gdy całkowita szybkość wejściowa spadnie poniżej 1 Mb/s; wtedy kolejka zacznie się zmniejszać). Ponieważ w przytoczonym przykładzie prawdopodobieństwo tego, że jednocześnie aktywnych będzie ponad 10 użytkowników, jest znikome, przełączanie pakietów w zasadzie zapewnia taką samą wydajność, co przełączanie obwodów. *Jednak przełączanie pakietów dodatkowo umożliwia obsłużenie ponadtrzykrotnie większej liczby użytkowników.*

Omówmy teraz kolejny prosty przykład. Załóżmy, że istnieje 10 użytkowników i jeden z nich generuje nagle tysiąc pakietów 1000-bitowych, natomiast pozostałe osoby nie wykonują żadnej operacji powodującej tworzenie pakietów. W przypadku przełączania obwodów z multipleksowaniem TDM i ramki złożonej z 10 szczelin czasowych, z których każda zawiera 1000 bitów, w celu przesłania danych aktywny użytkownik będzie mógł zastosować tylko jedną szczelinę każdej ramki. Pozostałe dziewięć szczelin czasowych znajdujących się w każdej ramce będzie bezczynnych. Upłynie 10 sekund, zanim zostaną przesłane wszystkie z miliona bitów wygenerowanych przez aktywnego użytkownika. Gdy użyje się przełączania pakietów, aktywny użytkownik może cały czas przesyłać swoje pakiety z pełną szybkością łącza wynoszącą 1 Mb/s. Wynika to stąd, że żaden inny użytkownik nie generuje pakietów, które muszą być multipleksowane z pakietami aktywnego użytkownika. W tym przypadku wszystkie jego dane zostaną przesłane w ciągu sekundy.

Powyższe przykłady demonstrują dwa przypadki, w których wydajność przełączania pakietów może być wyższa od oferowanej przez przełączanie obwodów. Ponadto przykłady uwidaczniają zasadniczą różnicę między dwoma odmianami współużytkowania przepustowości łącza przez wiele strumieni danych. Przełączanie obwodów odgórnie decyduje o wykorzystaniu łącza transmisyjnego niezależnie od zapotrzebowania. W efekcie marnowana jest część przydzielonego łącza, która w danej chwili nie jest używana. Z kolei przełączanie pakietów przydziela przepustowość łącza *na żądanie*. Pojemność łącza transmisyjnego w przypadku kolejnych pakietów będzie dzielona tylko przez użytkowników tworzących pakiety, które muszą zostać przesłane przy użyciu łącza.

Choć w obecnie istniejących sieciach telekomunikacyjnych rozpowszechnione jest zarówno przełączanie pakietów, jak i przełączanie obwodów, z pewnością obowiązuje tendencja zmierzania w stronę przełączania pakietów. Nawet wiele aktualnie stosowanych sieci telefonicznych z przełączaniem obwodów powoli jest migrowanych do technologii przełączania pakietów. Z przełączania pakietów sieci telefoniczne często korzystają zwłaszcza w celu obsługi kosztownej części połączenia międzykontynentalnego.

1.3.3. Sieć sieci

Wcześniej wspomnieliśmy, że systemy końcowe (komputery PC, smartfony, serwery WWW, serwery poczty itp.) są łączone z internetem za pośrednictwem lokalnego dostawcy ISP. Może on udostępniać połączenie przewodowe lub bezprzewodowe za pomocą jednej z wielu technologii, takich jak sieć DSL, kablowa, FTTH, Wi-Fi lub komórkowa. Warto zauważyć, że lokalnym dostawcą ISP nie musi być operator telekomunikacyjny ani sieci kablowej. Może to być na przykład uczelnia (zapewniająca dostęp do internetu studentom, pracownikom i wykładowcom) lub firma (gwarantująca dostęp pracownikom). Jednak podłączanie użytkowników końcowych i dostawców treści do sieci dostępowych stanowi tylko niewielki element układanki łączącej ze sobą miliardy systemów końcowych tworzących internet. Dokończenie układanki wymaga połączeń także między sieciami dostępowymi dostawców ISP. Internet jest *siecią sieci*. Zrozumienie tego pojęcia jest kluczem do poradzenia sobie z tą układanką.

Przez lata sieć sieci tworząca internet wyewoluowała w bardzo złożoną strukturę. Duża część tej ewolucji jest napędzana ekonomią i polityką państw, a nie uwzględnianiem wydajności. Aby opisać obecną strukturę sieci internetowej, przedstawiamy po kolei zestaw struktur, z których każda następna jest lepszym przybliżeniem dzisiejszego złożonego internetu. Pamiętaj, że ogólnym celem jest połączenie sieci dostępowych dostawców ISP w taki sposób, aby wszystkie systemy końcowe mogły przesyłać między sobą pakiety. Naiwne rozwiązanie polega na *bezpośrednim* połączeniu każdego dostawcy ISP z wszystkimi innymi dostawcami. Taki projekt w postaci kraty jest oczywiście zbyt kosztowny dla dostawców ISP, ponieważ wymagałby, aby każdy dostawca utrzymywał odrębne łącze komunikacyjne z każdym z setek tysięcy innych dostawców z całego świata.

Pierwsza omawiana struktura sieciowa, *struktura sieci nr 1*, łączy wszystkie sieci dostępowe dostawców ISP z *jednym globalnym tranzytowym dostawcą ISP*. Ten (wyimaginowany) dostawca to sieć routerów i łączy komunikacyjnych, która nie tylko pokrywa cały świat, ale obejmuje przynajmniej jeden router w pobliżu każdej z setek tysięcy sieci dostępowych dostawców ISP. Oczywiście zbudowanie tak dużej sieci byłoby bardzo kosztowne. Aby osiągać zyski, globalny dostawca musiałby naliczać wszystkim dostawcom sieci dostępowych opłaty, przy czym ceny powinny odzwierciedlać (choć niekoniecznie w pełni proporcjonalnie) ruch generowany przez dostawców sieci dostępowych w sieci dostawcy globalnego. Ponieważ w tym ujęciu dostawca sieci dostępowych płaci globalnemu dostawcy sieci tranzytowej, ten pierwszy jest **klientem**, a ten drugi — **dostawcą**.

Jeśli jakaś firma zbuduje i będzie eksploatować zyskowną globalną sieć tranzytową, oczywiście inne organizacje też zaczną tworzyć podobne sieci i konkurować z pierwszym dostawcą. To prowadzi do *struktury sieci nr 2*, składającej się z setek tysięcy dostawców sieci dostępowych i *wielu* dostawców globalnych sieci tranzytowych. Dostawcy sieci dostępowych naturalnie preferują strukturę sieci nr 2, ponieważ mogą wybierać spośród konkurujących ze sobą dostawców globalnych sieci dostępowych, uwzględniając ceny i oferowane usługi. Zauważ, że same globalne sieci tranzytowe

muszą być ze sobą połączone. W przeciwnym razie dostawcy sieci dostępowych połączeni z jednym dostawcą sieci tranzytowej nie będą mogli komunikować się z sieciami dostępowymi powiązanymi z innymi sieciami tranzytowymi.

Opisana tu struktura sieci nr 2 to dwupoziomowa hierarchia z globalnymi sieciami tranzytowymi w górnej warstwie i sieciami dostępowymi w dolnej. Zakładamy tu, że dostawcy globalnych sieci tranzytowych nie tylko potrafią dotrzeć blisko każdej sieci dostępowej, ale też że jest to dla nich opłacalne. W praktyce jest tak, że choć niektórzy dostawcy mają imponujący zasięg globalny i bezpośrednie połączenia z wieloma sieciami dostępowymi, żaden dostawca nie jest obecny we wszystkich miastach świata. Na danym obszarze może działać **regionalny dostawca ISP**, z którym łączą się sieci dostępowe z tego regionu. Każdy regionalny dostawca ISP łączy się z **dostawcami ISP pierwszej warstwy**. Dostawcy ISP pierwszej warstwy działają podobnie jak (wyimaginowany) dostawca globalnej sieci tranzytowej, jednak istnieją naprawdę i nie są obecnie we wszystkich miastach świata. Działa kilkunastu takich dostawców, w tym firmy Level 3 Communications, AT&T, Sprint i NTT. Co ciekawe, nikt nie sankcjonuje statusu dostawców ISP pierwszej warstwy (zgodnie z powiedzeniem: jeśli musisz pytać, czy jesteś członkiem grupy, prawdopodobnie nim nie jesteś).

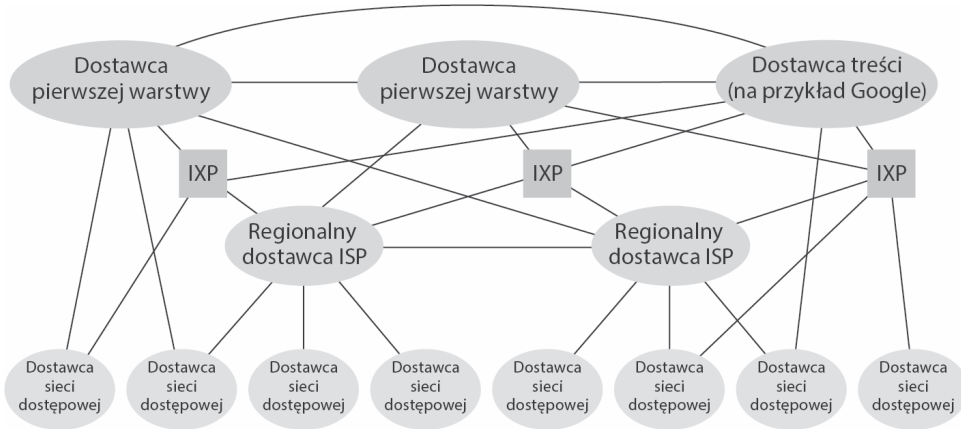
Wróćmy do sieci sieci. Nie tylko istnieje wielu konkurujących ze sobą dostawców ISP pierwszej warstwy, ale na danym obszarze może działać też wielu rywalizujących między sobą dostawców regionalnych. W takiej hierarchii każdy dostawca sieci dostępowej płaci regionalnemu dostawcy ISP, z którym się łączy, a każdy dostawca regionalny płaci dostawcy pierwszej warstwy. Dostawca sieci dostępowej może też łączyć się bezpośrednio z dostawcą pierwszej warstwy; wtedy płaci bezpośrednio jemu. Na każdym poziomie hierarchii występuje więc relacja klient – dostawca. Zauważ, że dostawcy pierwszej warstwy nie płacą nikomu, ponieważ znajdują się na szczycie hierarchii. Sytuację dodatkowo komplikuje to, że w niektórych obszarach działają duzi regionalni dostawcy ISP (czasem pokrywający cały kraj), z którymi łączą się mniejsi regionalni dostawcy. Duży regionalny dostawca łączy się z dostawcą pierwszej warstwy. Na przykład w Chinach w każdym mieście znajduje się dostawca sieci dostępowej, który łączy się z dostawcą z poziomu prowincji, łączącemu się z dostawcą krajowym połączonym z dostawcą pierwszej warstwy [Tian 2012]. Jest to hierarchia wielowarstwowa, *struktura sieci nr 3*, która nadal jest tylko zgrubnym przybliżeniem współczesnego internetu.

Aby uzyskać sieć bardziej przypominającą obecny internet, do hierarchicznej struktury sieci nr 3 trzeba dodać punkty **PoP** (ang. *points of presence*), multi-homing, peering i punkty **IXP** (ang. *internet exchange points*). Punkty PoP znajdują się na wszystkich poziomach hierarchii oprócz dolnej (sieci dostępowych). Punkt **PoP** to router lub grupa routerów w tej samej lokalizacji w sieci dostawcy, gdzie klient może połączyć się z siecią dostawcy. Klient, aby powiązać swoją sieć z punktem PoP dostawcy, może wydzierżawić szybkie łącze od niezależnego dostawcy telekomunikacyjnego i bezpośrednio połączyć jeden ze swoich routerów z routerem z punktu PoP. Każdy dostawca ISP (z wyjątkiem dostawców pierwszej warstwy) może stosować **multi-homing**, czyli utrzymywać połączenie z więcej niż jednym dostawcą z wyższej warstwy.

Na przykład dostawca sieci dostępowej może być połączony z dwoma dostawcami regionalnymi, a dodatkowo także z dostawcą pierwszej warstwy. Podobnie dostawca regionalny może utrzymywać połączenie z kilkoma dostawcami pierwszej warstwy. W tym modelu dostawca ISP może wciąż wysyłać i otrzymywać pakiety nawet po awarii jednego z jego dostawców.

Wiesz już, że dostawcy z niższej warstwy płacą dostawcom z warstwy wyżej za zapewnianie globalnych połączeń internetowych. Kwota, jaką klient płaci swojemu dostawcy, jest zależna od generowanego ruchu. Aby zmniejszyć koszty, para pobliskich dostawców z tego samego poziomu może zastosować **peering**, czyli bezpośrednio powiązać swoje sieci, tak aby cały ruch między nimi był przekazywany przez bezpośrednie połączenie, a nie za pomocą pośredników z wyższych warstw. W takim modelu równorzędni dostawcy zwykle nie płacą sobie nawzajem. Jak wcześniej wspomniano, także dostawcy pierwszej warstwy stosują peering między sobą bez naliczania sobie opłat. Przystępne omówienie peeringu oraz relacji klient – dostawca znajdziesz w [Van der Berg 2008]. Na podobnej zasadzie niezależna firma może utworzyć **punkt IXP**, czyli punkt zbiorczy, w którym wielu dostawców ISP może się ze sobą łączyć. Punkt IXP zwykle znajduje się w niezależnym budynku z własnymi przełącznikami [Ager 2012]. W internecie działa obecnie ponad 400 punktów IXP [IXP List 2016]. Ten ekosystem, obejmujący dostawców sieci dostępowych, dostawców regionalnych, dostawców pierwszej warstwy, punkty PoP, multi-homing, peering i punkty IXP, nazywamy *strukturą sieci nr 4*.

Na zakończenie dochodzimy do *struktury sieci nr 5*, opisującej obecny internet. Ta struktura, przedstawiona na rysunku 1.15, to struktura sieci nr 4 wzbogacona o **sieci dostawców treści**. Obecnie jednym z najlepszych przykładów twórców takich sieci jest Google. W czasie, gdy powstaje ta książka, Google ma 50 – 100 centrów danych rozproszonych po Ameryce Północnej, Europie, Azji, Ameryce Południowej i Australii. Niektóre z tych centrów danych obejmują ponad sto tysięcy serwerów, natomiast inne są mniejsze i mieszczą tylko setki serwerów. Wszystkie centra danych Google'a są połączone prywatną siecią TCP/IP tej firmy, która pokrywa cały świat, ale jest odrębna od publicznego internetu. Ważne jest to, że prywatna sieć Google'a obsługuje tylko ruch kierowany do serwerów tej firmy i wychodzący z nich. Na rysunku 1.15 pokazano, jak prywatna sieć Google'a „omija" górne warstwy internetu dzięki peeringowi (bezkosztowemu) z dostawcami z niższych warstw — albo dzięki bezpośrednim połączeniom z tymi dostawcami, albo z wykorzystaniem punktów IXP [Labovitz 2010]. Jednak ponieważ do wielu sieci dostępowych można dotrzeć tylko przez sieci pierwszej warstwy, sieć Google'a łączy się też z dostawcami pierwszej warstwy i płaci za ruch, jaki do nich kieruje. Dzięki utworzeniu własnej sieci dostawca treści nie tylko zmniejsza kwoty płacone dostawcom z wyższych warstw, ale też zachowuje większą kontrolę nad tym, jak jego usługi są dostarczane do użytkowników końcowych. Infrastruktura sieciowa Google'a jest szczegółowo opisana w podrozdziale 2.6.



Rysunek 1.15. Połączenia między dostawcami ISP

Podsumowując, topologia internetu — sieci złożonej z sieci — jest złożona. Składa się z dziesiątków dostawców ISP pierwszej warstwy, a także tysięcy dostawców niższych warstw. Dostawcy ISP różnią się skalą działania. Niektórzy z nich swoim zakresem obejmują wiele kontynentów i oceanów, natomiast inni ograniczają się do niewielkich regionów świata. Dostawcy ISP niższych warstw łączą się z dostawcami wyższych warstw, natomiast te są połączone ze sobą. Użytkownicy i dostawcy treści są klientami dostawców ISP niższych warstw, a te klientami dostawców wyższych warstw. W ostatnich latach duzi dostawcy treści zaczęli tworzyć własne sieci i tam, gdzie to możliwe, łączyć je z dostawcami ISP niższych warstw.

1.4. Opóźnienie, utrata pakietów i przepustowość w sieciach z przełączaniem pakietów

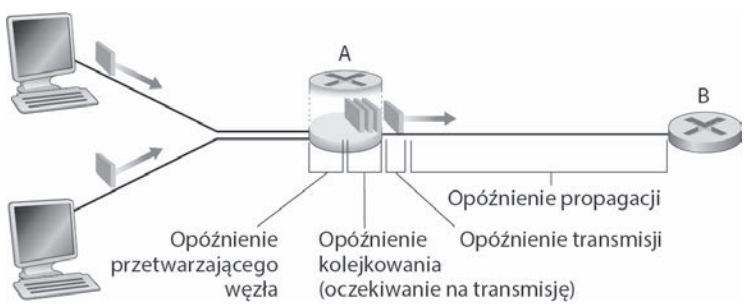
W podrozdziale 1.1 stwierdziliśmy, że internet można traktować jako infrastrukturę świadczącą usługi aplikacjom rozproszonym działającym w systemach końcowych. Chcielibyśmy, aby usługi internetowe przekazywały dowolną ilość danych między dwoma systemami natychmiast i bez utraty pakietów. Jednak są to wygórowane oczekiwania, niemożliwe do zrealizowania w praktyce. Sieci komputerowe zawsze mają ograniczoną przepustowość (ilość danych, jaką można przesłać w ciągu sekundy) między dwoma systemami, działają z opóźnieniem i tracą pakiety. Z jednej strony niefortunne jest, że prawa fizyki powodują opóźnienia i utratę danych, a także ograniczają przepustowość. Z drugiej strony z uwagi na występowanie tych problemów w sieciach komputerowych powstała fascynująca dziedzina związana z ich rozwiązywaniem. Obejmuje ona tyle zagadnień, że można poświęcić im cały kurs na temat sieci komputerowych i napisać na ich temat setki doktoratów! W tym podrozdziale rozpoczynamy badanie i omawianie opóźnień, utraty pakietów oraz przepustowości w sieciach komputerowych.

1.4.1. Omówienie opóźnień w sieciach z przełączaniem pakietów

Jak wcześniej wspomniano, pakiet jest najpierw wysyłany przez host źródłowy, a następnie przechodzi przez serię routerów i swoją podróż kończy, docierając do hosta docelowego. Po drodze pakiet jest transmitowany od jednego węzła (host lub router) do kolejnego. W tym czasie po trafieniu do *każdego* kolejnego węzła na pakiet ma wpływ kilka typów opóźnień. Do najważniejszych z nich należy zaliczyć **opóźnienie przetwarzającego węzła**, **kolejkowania**, **transmisji** i **propagacji**. Razem wymienione opóźnienia tworzą **całkowite opóźnienie węzła**. Wydajność wielu aplikacji internetowych (na przykład wyszukiwarek, przeglądarek, klientów poczty elektronicznej, komunikatorów, telefonii VoIP) jest w wysokim stopniu zależna od opóźnień w sieci. Aby dogłębnie zrozumieć przełączanie pakietów i sieci komputerowe, trzeba poznać naturę tych opóźnień i ich istotną rolę.

Typy opóźnień

Opóźnienia objaśnijmy, korzystając z rysunku 1.16. W trakcie pokonywania trasy od jednego punktu (źródłowego) do drugiego (docelowego) pakiet jest transmitowany kolejno przez routery A i B. Naszym celem jest scharakteryzowanie opóźnienia węzła występującego w routerze A. Warto zauważyć, że router A posiada łącze wyjściowe prowadzące do routera B. Przed łączem znajduje się kolejka (inaczej bufor). Gdy pakiet wysłany przez węzeł dotrze do routera A, ten sprawdzi nagłówek pakietu, aby identyfikować dla niego odpowiednie łącze wyjściowe i skierować go do niego. W omawianym przykładzie łączem wyjściowym pakietu jest łącze prowadzące do routera B. Pakiet może być transmitowany łączem, tylko gdy w danej chwili nie przesyła ono żadnego innego pakietu, a ponadto w kolejce nie ma pakietów. Jeśli łącze akurat jest zajęte lub są już kolejkowane inne pakiety przeznaczone dla tego łącza, nowy pakiet, który odebrano, zostanie umieszczony w kolejce.



Rysunek 1.16. Opóźnienie węzła występujące w routerze A

Opóźnienie przetwarzającego węzła

Czas wymagany do sprawdzenia nagłówka pakietu i stwierdzenia, gdzie należy go skierować, jest elementem **opóźnienia przetwarzającego węzła**. Opóźnienie to może też uwzględniać inne czynniki, takie jak czas niezbędny do sprawdzenia błędów na

poziomie bitów, które wystąpiły podczas przesyłania bitów pakietu od węzła nadawczego do routera A. Opóźnienie przetwarzającego węzła w przypadku bardzo szybkich routerów jest określane w mikrosekundach lub mniejszych jednostkach czasu. Po przetworzeniu pakietu router umieszcza go w kolejce znajdującej się przed łączem prowadzącym do routera B (w rozdziale 4. szczegółowo wyjaśniono zasady działania routera).

Opóźnienie kolejkowania

Opóźnienie kolejkowania dotyczy pakietu oczekującego w kolejce na przesłanie łączem. Wielkość opóźnienia w przypadku określonego pakietu będzie zależała od liczby pakietów wcześniej umieszczonych w kolejce i oczekujących na transmisję przy użyciu łącza. Jeśli kolejka jest pusta i w danej chwili nie jest przesyłany żaden inny pakiet, opóźnienie kolejkowania dla rozważanego pakietu będzie zerowe. Z kolei gdy obciążenie sieci jest duże i wiele innych pakietów oczekuje na przesłanie, opóźnienie będzie znaczne. Wkrótce okaże się, że liczba pakietów, które mogą się znaleźć w kolejce przed odebranym pakietem, jest funkcją natężenia i charakteru danych umieszczanych w kolejce. W praktyce wartość opóźnienia kolejkowania może być wyrażona w mikrosekundach lub milisekundach.

Opóźnienie transmisji

Zakładając, że pakiety są transmitowane zgodnie z zasadą „pierwszy na wejściu, pierwszy na wyjściu", powszechnie wykorzystywaną w sieciach z przełączaniem pakietów, rozważany pakiet będzie mógł być przesłany dopiero, gdy zostaną wysłane wszystkie pakiety, które wcześniej odebrano. Określmy długość pakietu jako D bitów, natomiast szybkość łącza między routerem A i B jako S bitów/s. Przykładowo, w przypadku łącza Ethernet 10 Mb/s S wyniesie 10 Mb/s. W przypadku łącza Ethernet 100 Mb/s S = 100 Mb/s. **Opóźnienie transmisji** wynosi D/S. Opóźnienie to określa czas niezbędny do umieszczenia (wysłania) w łączu wszystkich bitów pakietu. W praktyce wartość opóźnienia transmisji zwykle jest wyrażana w mikrosekundach lub milisekundach.

Opóźnienie propagacji

Po umieszczeniu bitu w łączu trzeba go przesłać do routera B. Czas potrzebny na propagację bitu z początku łącza do routera B wyznacza **opóźnienie propagacji**. Bit jest przemieszczany z szybkością propagacji łącza. Zależy ona od fizycznego nośnika (światłowód, skrętka miedziana itp.) łącza i zawiera się w następującym przedziale:

$$2 \cdot 10^8 m/s - 3 \cdot 10^8 m/s$$

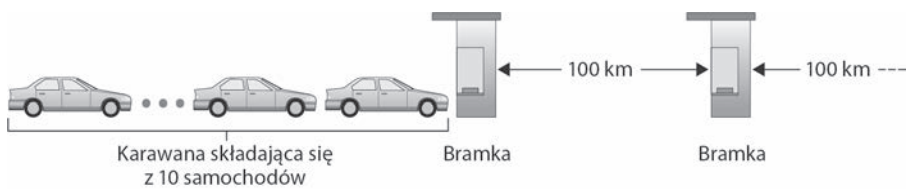
Szybkość jest równa lub trochę mniejsza od prędkości światła. Opóźnienie propagacji jest odległością między dwoma routerami podzieloną przez szybkość propagacji. Oznacza to, że opóźnienie to określa zależność d/s, w której d jest dystansem między routerem A i B, natomiast s szybkością propagacji łącza. Gdy zakończy się propagacja do routera B ostatniego bitu pakietu, wraz z wszystkimi wcześniejszymi bitami zostanie

zapisany w routerze B. Gdy to nastąpi, router B będzie kontynuował proces, wykonując operację przekazywania pakietu. W sieciach rozległych opóźnienia propagacji są wyrażane w milisekundach.

Porównanie opóźnień transmisji i propagacji

Nowicjusze w dziedzinie sieci komputerowych czasami mają problemy ze zrozumieniem różnicy występującej między opóźnieniami transmisji i propagacji. Różnica jest subtelna, ale istotna. Opóźnienie transmisji określa czas, jakiego router potrzebuje na wysłanie pakietu. Opóźnienie to jest funkcją długości pakietu i szybkości transmisji łącza. Jednak nie ma nic wspólnego z odległością między dwoma routerami. Z kolei opóźnienie propagacji jest czasem, jaki jest wymagany na propagację bitu z jednego routera do drugiego. Opóźnienie jest funkcją odległości między dwoma routerami. Jednak nie ma związku z długością pakietu lub szybkością transmisji łącza.

Analogia powinna rozwiać wątpliwości dotyczące różnic między opóźnieniami transmisji i propagacji. Weźmy pod uwagę autostradę, na której co 100 kilometrów znajduje się bramka (rysunek 1.17). Segmenty autostrady między kolejnymi bramkami można potraktować jak łącza, natomiast bramki jak routery. Załóżmy, że samochody przemieszczają się (propagują) po autostradzie z prędkością 100 km/h (oznacza to, że po minięciu bramki pojazd natychmiast przyspieszy do prędkości 100 km/h i utrzyma ją do kolejnej bramki). Przyjmijmy, że autostradą przemieszcza się karawana złożona z 10 samochodów jadących w określonej niezmiennej kolejności. Każdy pojazd można potraktować jak bit, natomiast karawanę jak pakiet. Ponadto załóżmy, że każda bramka w ciągu 12 sekund obsługuje (transmituje) jeden samochód. Ponieważ jest późna pora, samochody tworzące karawanę są jedynymi znajdującymi się na autostradzie. Przyjmijmy jeszcze, że gdy pierwszy pojazd grupy dotrze do bramki, poczeka przy niej do momentu, aż pozostałe 9 samochodów ustawi się za nim (w związku z tym przed ruszeniem w dalszą drogę cała karawana musi być pomieszczona przy bramce). Czas potrzebny na przepuszczenie przez bramkę całej karawany wynosi (10 samochodów)/(5 pojazdów na minutę) = 2 minuty. Czas ten można przyrównać do opóźnienia transmisji występującego w routerze. Czas potrzebny na przemieszczenie się samochodu od wylotu jednej bramki do kolejnej bramki wynosi 100 km/(100 km/h) = 1 godzina. Czas ten odpowiada opóźnieniu propagacji. A zatem czas, jaki upływa od momentu zgromadzenia całej karawany przy wlocie bramki do chwili znalezienia się pojazdów przy wlocie kolejnej bramki, jest sumą opóźnienia transmisji i propagacji. W omawianym przykładzie czas ten wyniesie 62 minuty.



Rysunek 1.17. Analogia do karawany samochodów

Rozwińmy bardziej przytoczoną analogię. Co by się stało, gdyby czas obsługi karawany przez bramkę był dłuższy niż czas pokonania przez samochód dystansu dzielącego dwie bramki? Dla przykładu załóżmy, że teraz pojazdy jadą z prędkością 1000 km/h, natomiast bramka obsługuje samochody z szybkością jednego na minutę. W tym przypadku opóźnienie związane z przejechaniem odcinka między dwoma bramkami wyniesie 6 minut, natomiast czas obsługi karawany 10 minut. Pierwsze samochody karawany pojawią się przy drugiej bramce, zanim ostatnie jej pojazdy miną pierwszą bramkę. Taka sytuacja występuje również w sieciach z przełączaniem pakietów. Pierwsze bity pakietu mogą zostać odebrane przez router, gdy wiele z pozostałych bitów w dalszym ciągu oczekuje na wysłane przez poprzedni router.

Jeśli obraz jest wart tysiąca słów, animacja musi być warta miliona. W poświęconej książce witrynie znajduje się interaktywna animacja dobrze ilustrująca opóźnienia transmisji i propagacji. Ta animacja pozwala też porównać oba typy opóźnienia. Gorąco zachęcamy Czytelników do uruchomienia tej animacji. Ponadto w [Smith 2009] w bardzo przystępny sposób opisano opóźnienie propagacji, kolejkowania i transmisji.

Jeśli opóźnienia przetwarzającego węzła, kolejkowania, transmisji i propagacji oznaczy się odpowiednio symbolami o_{przet}, o_{kolejk}, o_{trans} i o_{prop}, całkowite opóźnienie węzła określi następujące równanie:

$$o_{\text{węz}} = o_{\text{przet}} + o_{\text{kolejk}} + o_{\text{trans}} + o_{\text{prop}}$$

Udział poszczególnych opóźnień może się bardzo znacząco różnić. Przykładowo, opóźnienie o_{prop} może być nieznaczne (kilka mikrosekund) w przypadku łącza poprowadzonego między dwoma routerami znajdującymi się na tym samym kampusie uniwersyteckim. Jednak to samo opóźnienie będzie miało wartość setek milisekund, gdy dwa routery będą ze sobą połączone przy użyciu łącza obsługiwanego przez satelitę geostacjonarnego. W efekcie opóźnienie to może odgrywać dominującą rolę w równaniu określającym wartość opóźnienia $o_{\text{węz}}$. Podobnie opóźnienie o_{trans} może być nieznaczne, ale też dość istotne. Wkład wnoszony przez to opóźnienie zwykle jest pomijalny w przypadku szybkości transmisji rzędu 10 Mb/s i wyższych (dotyczy to na przykład sieci lokalnych). Jednak wartość opóźnienia o_{trans} może wynieść setki milisekund, gdy za pośrednictwem wolnego łącza obsługiwanego przez modem telefoniczny przesyła się duże pakiety internetowe. Choć opóźnienie o_{przet} jest często nieznaczne, ma duży wpływ na maksymalną przepustowość routera, która jest największą szybkością, z jaką urządzenie to może przekazywać pakiety.

1.4.2. Opóźnienie kolejkowania i utrata pakietów

Najbardziej skomplikowanym i interesującym składnikiem opóźnienia węzła jest opóźnienie kolejkowania o_{kolejk}. W rzeczywistości opóźnienie kolejkowania występujące w sieci komputerowej jest tak ważne i ciekawe, że poświęcono mu tysiące artykułów i wiele książek [Bertsekas 1991; Daigle 1991; Kleinrock 1975, 1976; Ross 1995]. W tym miejscu dokonamy jedynie bardzo ogólnego omówienia opóźnienia kolejkowania. Bardziej zainteresowani tym zagadnieniem mogą sprawdzić kilka książek, a ostatecznie

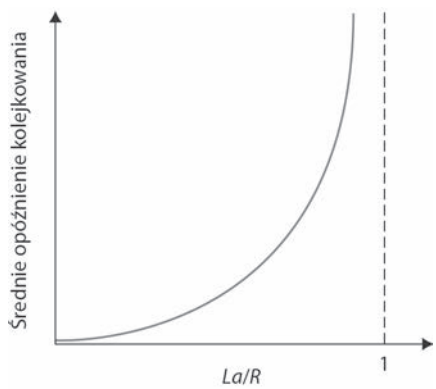
nawet napisać pracę doktorską na ten temat! W przeciwieństwie do pozostałych trzech opóźnień (o_{przet}, o_{trans} i o_{prop}) opóźnienie kolejkowania może być inne dla każdego pakietu. Jeśli na przykład jednocześnie w pustej kolejce pojawi się 10 pakietów, te, które będą transmitowane w pierwszej kolejności, nie doświadczą żadnego opóźnienia. Z kolei w przypadku pakietu wysłanego jako ostatni opóźnienie kolejkowania będzie stosunkowo duże (wynika to z faktu czekania pakietu na przesłanie pozostałych dziewięciu). A zatem opisując opóźnienie kolejkowania, zwykle korzysta się ze wskaźników statystycznych, takich jak średnie opóźnienie, wariancja opóźnienia i prawdopodobieństwo tego, że opóźnienie przekroczy określoną wartość.

Kiedy opóźnienie kolejkowania jest duże, a kiedy nieznaczne? Odpowiedź na to pytanie zależy od szybkości, z jaką dane są umieszczane w kolejce, szybkości transmisji łącza i sposobu odbierania danych (chodzi o to, czy pojawiają się w kolejce cyklicznie czy impulsowo). Aby trochę bardziej to przybliżyć, literą *a* oznaczmy średnią szybkość (wyrażoną w pakietach na sekundę), z jaką pakiety pojawiają się w kolejce. Jak pamiętamy, *S* jest szybkością transmisji (w bitach/s), z jaką bity są pobierane z kolejki. Dla uproszczenia załóżmy, że wszystkie pakiety składają się z *D* bitów. W związku z tym średnia szybkość, z jaką bity będą umieszczane w kolejce, wyniesie *Da* bitów/s. Przyjmijmy jeszcze, że kolejka jest bardzo duża, na tyle, że w zasadzie może pomieścić nieskończoną liczbę bitów. Współczynnik *Da/S* nazywany **natężeniem ruchu** często odgrywa istotną rolę w szacowaniu zakresu opóźnienia kolejkowania. Jeśli $Da/S > 1$, średnia szybkość, z jaką bity pojawiają się w kolejce, przekracza szybkość pobierania z niej bitów. W takiej niekorzystnej sytuacji kolejka będzie miała tendencję do powiększania się bez ograniczeń i opóźnienie kolejkowania będzie zbliżało się do nieskończoności! W związku z tym jedna z podstawowych zasad obowiązujących w przypadku inżynierii ruchu sieciowego brzmi: „*System należy tak projektować, aby natężenie ruchu nie przekroczyło wartości 1*”.

Rozważmy teraz sytuację, gdy $Da/S \leq 1$. W tym przypadku charakterystyka ruchu wejściowego ma wpływ na opóźnienie kolejkowania. Jeśli na przykład pakiety są odbierane cyklicznie, czyli co *D/S* sekund pojawia się jeden pakiet, każdy z nich będzie umieszczany w pustej kolejce, dzięki czemu nie wystąpi opóźnienie kolejkowania. Z kolei gdy pakiety są odbierane w sposób okresowy, ale też impulsowy, może pojawić się średnie opóźnienie o znacznej wartości. Dla przykładu załóżmy, że co *(D/S)N* sekund jednocześnie odbieranych jest *N* pakietów. Pierwszy transmitowany pakiet nie posiada żadnego opóźnienia kolejkowania. Drugi przesyłany pakiet ma opóźnienie wynoszące *D/S* sekund. Mówiąc bardziej ogólniej, z *n*-tym transmitowanym pakietem jest związane opóźnienie kolejkowania równe *(n–1)D/S* sekund. W ramach ćwiczenia Czytelnikowi pozostawiamy obliczenie dla przytoczonego przykładu średniego opóźnienia kolejkowania.

Powyższe dwa przykłady dotyczące okresowego odbierania danych są trochę akademickie. Zwykle proces umieszczania bitów w kolejce ma przebieg *losowy*. Oznacza to, że bity nie są odbierane według żadnego schematu i pakiety pojawiają się w przypadkowych odstępach czasu. W tym bardziej realistycznym wariancie zależność *Da/S* zwykle nie jest wystarczająca do pełnego scharakteryzowania statystyk dotyczących opóźnienia.

Niemniej jednak zależność jest pomocna w zrozumieniu zakresu oddziaływania opóźnienia kolejkowania. W szczególności wtedy, gdy natężenie ruchu jest bliskie zeru, w kolejce pojawia się niewiele pakietów i jest między nimi duży odstęp. Ponadto mało prawdopodobne jest, że odebrany pakiet napotka w kolejce na inny pakiet. A zatem średnie opóźnienie kolejkowania będzie niemal zerowe. Z kolei jeśli natężenie ruchu jest zbliżone do wartości 1, wystąpią okresy, w których szybkość pojawiania się bitów przekroczy możliwości transmisyjne (na skutek impulsowego charakteru odbierania danych), co w efekcie spowoduje utworzenie kolejki. W miarę zbliżania się wartości natężenia ruchu do 1 średnia długość kolejki będzie coraz większa. Na rysunku 1.18 przedstawiono zależność średniego opóźnienia kolejkowania od natężenia ruchu.



Rysunek 1.18. Zależność średniego opóźnienia kolejkowania od natężenia ruchu

Ważną kwestą uwidocznioną na rysunku 1.18 jest to, że wraz ze zbliżaniem się wartości natężenia ruchu do 1 szybko zwiększa się średnie opóźnienie kolejkowania. Niewielki procentowy wzrost natężenia powoduje znacznie większy procentowy wzrost wartości opóźnienia. Być może z takim zjawiskiem miało się do czynienia na autostradzie. Jeśli regularnie jeździ się drogą, która zazwyczaj jest zatłoczona, oznacza to, że w jej przypadku wartość natężenia ruchu jest bliska 1. Jeśli jakieś zdarzenie spowoduje nawet jeszcze większy ruch na drodze niż zwykle, opóźnienia mogą być ogromne.

Aby dobrze zrozumieć, czego dotyczą opóźnienia kolejkowania, warto odwiedzić poświęconą książce witrynę. Znajduje się w niej interaktywny aplet Java ilustrujący kolejki. Po ustawieniu częstotliwości przesyłania pakietów na wysokim poziomie (tak aby natężenie ruchu przekraczało 1) można zauważyć, że z czasem kolejka zaczyna się wydłużać.

Utrata pakietów

W powyższej analizie założyliśmy, że kolejka jest w stanie przechowywać nieskończoną liczbę pakietów. W rzeczywistości kolejka zlokalizowana przed łączem posiada ograniczoną pojemność, która w znacznym stopniu zależy od konstrukcji przełącznika i jego ceny. Ponieważ pojemność kolejki jest ograniczona, tak naprawdę przy zbliżaniu się

wartości natężenia ruchu do 1 opóźnienia pakietów nie zmierzają do nieskończoności. Może się okazać, że po odebraniu pakietu kolejka będzie już pełna. Gdy nie będzie możliwe przechowanie pakietu, router **odrzuci** go, co jest równoznaczne z **utratą** pakietu. Aplet Java ilustrujący kolejki pozwala zaobserwować także ich przepełnienie. Zjawisko to wystąpi, jeśli natężenie ruchu przekroczy 1.

Z punktu widzenia systemu końcowego wygląda to tak, jakby pakiet został przesłany do rdzenia sieci, lecz nigdy nie dotarł do miejsca przeznaczenia. Liczba utraconych pakietów zwiększa się wraz ze wzrostem natężenia ruchu. A zatem wydajność węzła jest często określana nie tylko w kategoriach opóźnienia, ale też pod względem prawdopodobieństwa utraty pakietu. W kolejnych rozdziałach Czytelnik dowie się, że utracony pakiet może być ponownie przesłany między poszczególnymi węzłami, co gwarantuje przesłanie wszystkich danych od punktu źródłowego do docelowego.

1.4.3. Opóźnienie międzywęzłowe

Do tego momentu skoncentrowaliśmy się na opóźnieniu węzła występującego w pojedynczym routerze. W ramach podsumowania pobieżnie przyjrzymy się opóźnieniu między węzłem źródłowym i docelowym. Aby przybliżyć to zagadnienie, załóżmy, że między źródłowym i docelowym hostem znajduje się N–1 routerów. Przyjmijmy również, że sieć nie jest przeciążona (opóźnienia kolejkowania są pomijalne), opóźnienie każdego przetwarzającego routera i źródłowego hosta wynosi o_{przet}, szybkość wysyłania danych przez routery i źródłowy host jest równa S bitów/s, natomiast opóźnienie propagacji dla każdego łącza ma wartość o_{prop}. Po zsumowaniu opóźnień węzła uzyska się opóźnienie międzywęzłowe wyrażone następującą zależnością:

$$o_{\text{węz-węz}} = N(o_{\text{przet}}+o_{\text{trans}}+o_{\text{prop}}) \tag{1.2}$$

W równaniu $o_{\text{trans}} = D/S$, gdzie D jest długością pakietu. Zauważ, że wzór 1.2 to uogólnienie wzoru 1.1, nieuwzględniającego opóźnienia przetwarzania i opóźnienia propagacji. Czytelnikowi pozostawiamy uogólnienie powyższego równania pod kątem różnorodnych opóźnień występujących w węzłach i istnienia w każdym z nich średniego opóźnienia kolejkowania.

Traceroute

Aby od strony praktycznej przedstawić opóźnienie występujące w sieci komputerowej, skorzystamy z programu diagnostycznego *Traceroute*. Jest to proste narzędzie, które można uruchomić na dowolnym komputerze podłączonym do internetu. Gdy użytkownik zidentyfikuje docelowy host, program uaktywniony na źródłowym komputerze wyśle w jego kierunku wiele specjalnych pakietów. W trakcie zmierzania do docelowego hosta pakiety przechodzą przez serię routerów. Gdy router odbierze jeden z takich pakietów, do źródłowego hosta odeśle krótki komunikat zawierający nazwę i adres routera.

Załóżmy, że między źródłowym i docelowym hostem znajduje się *N*–1 routerów. Źródłowy host wyśle do sieci *N* specjalnych pakietów, z których każdy posiada adres hosta docelowego. Pakiety te są ponumerowane od 1 do *N* (pierwszy pakiet ma numer 1, natomiast ostatni *N*). Gdy *n*-ty router odbierze *n*-ty pakiet o numerze *n*, nie przekaże go w kierunku jego miejsca przeznaczenia, lecz wyśle komunikat do źródłowego hosta. Po odebraniu *n*-tego pakietu docelowy host również prześle komunikat źródłowemu hostowi. Źródłowy host rejestruje czas, jaki upłynie od chwili wysłania pakietu do momentu otrzymania powiązanego z nim komunikatu zwrotnego. Ponadto host zapisuje nazwę i adres routera lub docelowego hosta, który zwrócił komunikat. W ten sposób źródłowy host może odtworzyć trasę pokonaną przez pakiety transmitowane z miejsca źródłowego do docelowego. Poza tym źródłowy host jest w stanie określić opóźnienia całkowitej trasy generowane przez pośredniczące routery. Program *Traceroute* trzykrotnie wykonuje wyżej opisany proces. Oznacza to, że źródłowy host wysyła do docelowego $3 * N$ pakietów. Narzędzie zostało szczegółowo omówiono w dokumencie RFC 1393.

Przedstawimy teraz przykład wyników zwróconych przez program *Traceroute*. W tym przypadku trasa wiodła od źródłowego hosta *gaia.cs.umass.edu* (zlokalizowany na uniwersytecie w Massachusetts) do hosta *cis.poly.edu* (znajduje się na politechnice w Brooklynie). Wynik ma postać sześciu kolumn. W pierwszej jest podana wyżej opisana wartość *n* określająca liczbę routerów znajdujących się na trasie pakietów. W drugiej kolumnie jest widoczna nazwa routera. W trzeciej kolumnie jest wstawiony adres routera (w formacie *xxx.xxx.xxx.xxx*), natomiast w trzech ostatnich opóźnienia trasy dla trzech prób. Jeśli źródłowy host od dowolnego routera otrzyma mniej niż trzy komunikaty (na skutek utraty pakietu w sieci), program *Traceroute* umieści znak gwiazdki tuż za numerem routera i nie poda dla niego trzech wartości opóźnień całkowitej trasy.

```
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  12.082 ms
```

W powyższym wyniku znajduje się dziewięć routerów zlokalizowanych między hostem źródłowym i docelowym. Wszystkie routery posiadają adres, natomiast część z nich nazwę. Przykładowo, nazwą routera 3. jest `border4-rt-gi-1-3.gw.umass.edu`, a jego adresem `128.119.2.194`. Po przyjrzeniu się wynikom podanym dla tego routera stwierdzi się, że pierwsze z trzech opóźnień całkowitej trasy związanych z przesłaniem pakietu między źródłowym hostem i routerem wynosi 1,03 ms. Pozostałe dwa opóźnienia mają wartości 0,48 i 0,45 ms. Wszystkie trzy opóźnienia całkowitej trasy uwzględniają wszystkie wcześniej omówione opóźnienia, takie jak opóźnienia transmisji, propagacji,

przetwarzania przez router i kolejkowania. Ponieważ opóźnienie kolejkowania zmienia się w czasie, opóźnienie całkowitej trasy pakietu *n* wysłanego do routera *n* w rzeczywistości może być większe od opóźnienia całkowitej trasy pakietu *n*+1 przesłanego do routera *n*+1. W powyższym wyniku należy zauważyć, że w przypadku routera 6. opóźnienia okazują się być większe od opóźnień routera 7.

Czy Czytelnik chciałby sam wypróbować program *Traceroute*? *Gorąco* namawiamy do odwiedzenia strony internetowej (*http://www.traceroute.org/*), na której zamieszczono interfejs udostępniający rozbudowaną listę źródeł umożliwiających śledzenie trasy pokonywanej przez pakiety. Po wybraniu źródła podaje się nazwę dowolnego hosta docelowego. Program *Traceroute* zajmie się całą resztą. Dostępnych jest wiele bezpłatnych aplikacji pełniących funkcję graficznego interfejsu dla tego programu. Jednym z naszych ulubionych jest PingPlotter [PingPlotter 2016].

Opóźnienia związane z systemami końcowymi, aplikacjami i inne

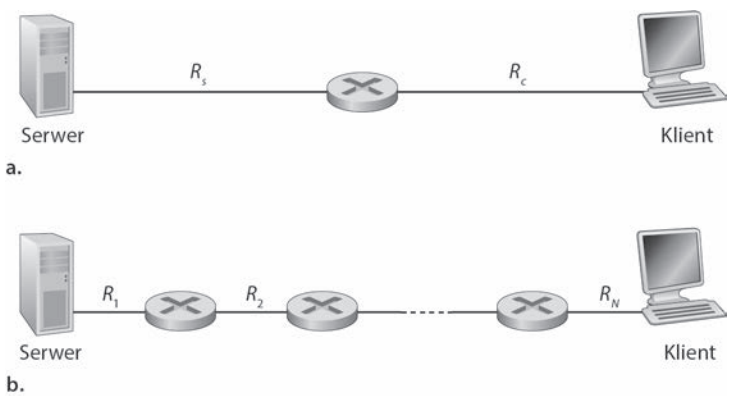
Oprócz opóźnień związanych z przetwarzaniem, transmisją i propagacją występują też dodatkowe opóźnienia w systemach końcowych. System końcowy przesyłający pakiet do wspólnego nośnika (na przykład w sieciach Wi-Fi lub z użyciem modemu kablowego) może *celowo* opóźniać transmisję z uwagi na protokół współużytkowania nośnika z innymi systemami (takie protokoły omawiamy szczegółowo w rozdziale 6.). Inne ważne opóźnienie związane jest z tworzeniem pakietów na potrzeby danego nośnika i występuje w aplikacjach VoIP. W systemach VoIP po stronie nadawcy trzeba zapełnić pakiet zakodowanym, cyfrowym zapisem mowy. Dopiero potem można wysłać dane do internetu. Czas potrzebny na zapełnienie pakietu — opóźnienie związane z tworzeniem pakietów — może być długi i wpływać na postrzeganą przez użytkownika jakość połączenia w systemie VoIP. To zagadnienie rozwijamy w jednym z ćwiczeń w końcowej części rozdziału.

1.4.4. Przepustowość w sieciach komputerowych

Obok opóźnienia i utraty pakietów następną kluczową miarą wydajności sieci komputerowych jest przepustowość między węzłami końcowymi. Aby zdefiniować przepustowość, załóżmy, że w sieci komputerowej z hosta A do hosta B przesyłany jest duży plik. Może to być na przykład długi film wideo przekazywany między węzłami w systemie wymiany plików P2P. **Przepustowość chwilowa** w każdym momencie to szybkość (w bitach na sekundę) pobierania danych przez host B. Wiele aplikacji, w tym liczne systemy wymiany plików P2P, wyświetlają w interfejsie użytkownika przepustowość chwilową w czasie pobierania danych — Czytelnik prawdopodobnie już to zaobserwował! Jeśli plik składa się z *F* bitów, a ich pobranie zajmuje hostowi B *T* sekund, **przepustowość średnia** transferu wyniesie *F*/*T* b/s. W niektórych aplikacjach, na przykład w obszarze telefonii internetowej, pożądane jest niskie opóźnienie i utrzymanie przepustowości chwilowej powyżej określonego progu (na przykład ponad 24 kb/s przy internetowych połączeniach telefonicznych i ponad 256 kb/s przy przekazywaniu

nagrań wideo w czasie rzeczywistym). W innych programach, na przykład służących do transferu plików, opóźnienie nie jest krytyczne, natomiast korzystne jest osiągnięcie jak najwyższej przepustowości.

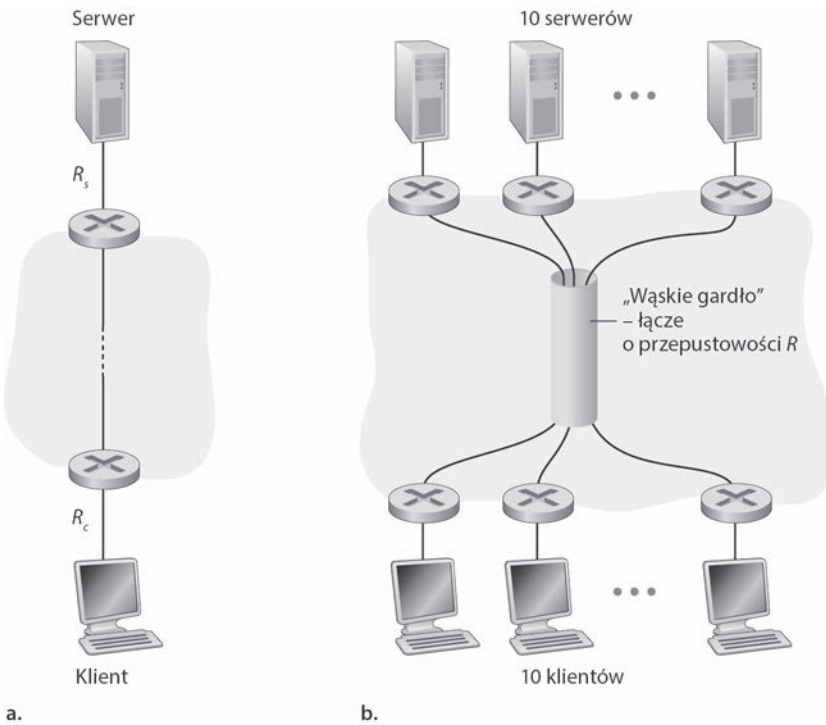
Aby lepiej wyjaśnić to ważne zagadnienie, jakim jest przepustowość, posłużymy się kilkoma przykładami. Rysunek 1.19(a) przedstawia dwa systemy końcowe (serwer i klient) połączone dwoma łączami komunikacyjnymi i routerem. Zastanówmy się nad przepustowością przy transferze pliku z serwera do klienta. Przyjmijmy, że szybkość połączenia między routerem i serwerem wynosi R_s, a szybkość połączenia między routerem i klientem — R_c. Załóżmy, że jedyne bity przesyłane w całej sieci to te przekazywane z serwera do klienta. Jaka jest przepustowość na trasie serwer – klient w tych idealnych warunkach? Aby odpowiedzieć na to pytanie, wyobraźmy sobie, że bity to *ciecz*, a łącza komunikacyjne to *rury*. Serwer oczywiście nie może „pompować" bitów przez połączenie z szybkością większą niż R_s b/s, a router nie może przekazywać danych szybciej niż z prędkością R_c b/s. Jeśli $R_s < R_c$, bity „pompowane" przez serwer „przepłyną" przez router i dotrą do klienta z szybkością R_s b/s, co oznacza przepustowość na tym samym poziomie. Natomiast jeżeli $R_s > R_c$, router nie będzie mógł przekazywać bitów tak szybko, jak je odbiera. Wtedy bity opuszczą router z szybkością wynoszącą tylko R_c, co oznacza, że przepustowość między węzłami wyniesie także R_c. Warto zauważyć, że jeśli bity będą docierać do routera z szybkością R_s, a opuszczać go w tempie R_c, liczba bajtów w routerze oczekujących na przesłanie do klienta będzie ciągle rosła, co jest wysoce niepożądane! Dlatego w tej prostej, dwupołączeniowej sieci przepustowość wynosi $\min\{R_s, R_c\}$, czyli jest równa szybkości transmisji **łącza stanowiącego wąskie gardło**. Po określeniu przepustowości można oszacować czas potrzebny na przesłanie dużego pliku o F bitach z serwera do klienta — $F/\min\{R_s, R_c\}$. Przedstawmy teraz konkretny przykład. Załóżmy, że użytkownik pobiera plik MP3 o wielkości $F = 32$ milionów bitów, szybkość transmisji dla serwera wynosi $R_s = 2$ Mb/s, a połączenie dostępowe działa z prędkością $R_c = 1$ Mb/s. Czas potrzebny na przesłanie tego pliku wynosi 32 sekundy. Oczywiście takie obliczenia przepustowości i czasu transferu to tylko przybliżone szacunki, ponieważ nie uwzględniliśmy w nich wielkości pakietów ani kwestii związanych z protokołem.



Rysunek 1.19. Przepustowość przy transferze pliku z serwera do klienta

Rysunek 1.19(b) przedstawia sieć o N łączach między serwerem i klientem. Szybkość transmisji dla tych łączy wynosi $R_1, R_2, \ldots, R_N$. Po zastosowaniu tego samego toku rozumowania, co w przypadku sieci dwupołączeniowej, można ustalić, że przepustowość przy transferze pliku z serwera do klienta wyniesie $\min\{R_1, R_2, \ldots, R_N\}$. Także tu jest to szybkość transmisji dla łącza będącego wąskim gardłem na drodze między serwerem i klientem.

Zastanówmy się teraz nad następnym przykładem związanym ze współczesnym internetem. Rysunek 1.20(a) przedstawia dwa systemy końcowe (serwer i klient) podłączone do sieci komputerowej. Jaka będzie przepustowość przy transferze pliku z serwera do klienta? Serwer jest połączony z siecią łączem dostępowym o szybkości R_s, a łącze dostępowe klienta działa z prędkością R_c. Załóżmy, że wszystkie łącza w rdzeniu sieci komunikacyjnej zapewniają bardzo wysoką szybkość transferu, znacznie powyżej poziomu R_s i R_c. Rzeczywiście, obecnie w rdzeniu internetu dostępny jest zapas szybkich łączy, dlatego przeciążenia zdarzają się rzadko. Przyjmijmy też, że jedyne bity przesyłane w całej sieci to te przekazywane z serwera do klienta. Ponieważ rdzeń sieci komputerowej w tym przykładzie przypomina szeroką rurę, szybkość przepływu danych z lokalizacji źródłowej do docelowej to ponownie minimum z wartości R_s i R_c, a więc przepustowość = $\min\{R_s, R_c\}$. Dlatego czynnikiem ograniczającym przepustowość we współczesnym internecie jest zwykle sieć dostępowa.



Rysunek 1.20. Przepustowość między węzłami — (a) klient pobiera plik z serwera; (b) 10 klientów pobiera dane z 10 serwerów

Oto ostatni przykład. Przyjrzyjmy się rysunkowi 1.20(b). Widnieje na nim 10 serwerów i 10 klientów połączonych z rdzeniem sieci komputerowej. W tym przykładzie jednocześnie trwa 10 operacji pobierania danych w 10 parach klient-serwer. Załóżmy, że w danym momencie ruch generują tylko te elementy. Na rysunku widać, ze w rdzeniu znajduje się łącze używane przy wszystkich 10 operacjach pobierania. Szybkość tego łącza to S. Przyjmijmy, że łącza dostępowe wszystkich serwerów działają z szybkością S_s, łącza dostępowe wszystkich klientów mają prędkość S_c, a szybkość transmisji wszystkich pozostałych łączy w rdzeniu (oprócz wspólnego łącza o prędkości S) znacznie przekracza wartości S_s, S_c i S. Jaka będzie przepustowość operacji pobierania? Oczywiste jest, że jeśli szybkość S wspólnego łącza jest wysoka, na przykład stukrotnie przekracza prędkości S_s i S_c, to przepustowość ponownie wyniesie $\min\{S_s, S_c\}$. Co się jednak stanie, jeśli wspólne łącze działa z szybkością porównywalną do S_s i S_c? Jaka wtedy będzie przepustowość? Przyjrzyjmy się konkretnemu przykładowi. Załóżmy, że $S_s = 2$ Mb/s, $S_c = 1$ Mb/s, $S = 5$ Mb/s, a szybkość transmisji wspólnego łącza jest równo rozdzielona między 10 operacji pobierania. Wtedy wąskim gardłem dla każdej z tych operacji nie będzie sieć dostępowa, ale wspólne łącze w rdzeniu, które zapewnia poszczególnym parom przepustowość na poziomie tylko 500 kb/s. Dlatego przepustowość między węzłami dla każdej operacji pobierania jest obniżona do 500 kb/s.

Przykłady przedstawione na rysunkach 1.19 i 1.20(a) pokazują, że przepustowość zależy od szybkości transmisji łączy, którymi przesyłane są dane. Zobaczyliśmy, że kiedy nie ma innego ruchu, można w przybliżeniu przyjąć, iż przepustowość to minimalna szybkość transmisji na ścieżce między lokalizacją źródłową i docelową. Przykład zilustrowany na rysunku 1.20(b) pokazuje, że ogólnie przepustowość zależy nie tylko od szybkości transmisji łączy na ścieżce, ale też od konkurencyjnego ruchu. Łącze o wysokiej szybkości może być wąskim gardłem w procesie transferu pliku, jeśli przekazuje także duże ilości innych danych. Przepustowości w sieciach komputerowych przyjrzymy się bliżej w ćwiczeniach i w następnych rozdziałach.

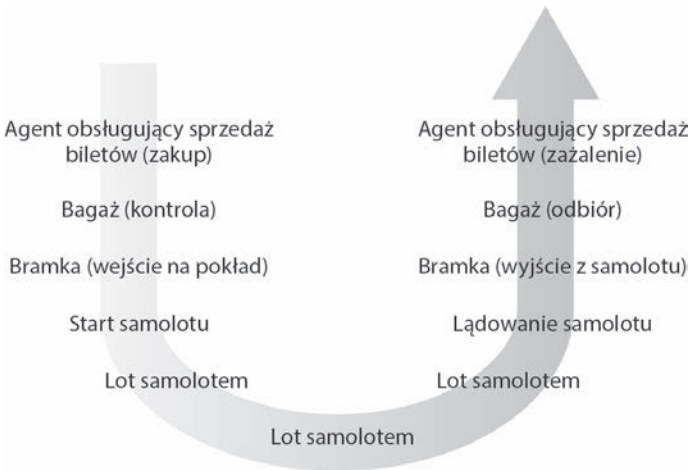
1.5. Warstwy protokołów i modele ich usług

Z naszej dotychczasowej analizy wynika, że internet jest *wyjątkowo* złożonym systemem. Przekonaliśmy się, że internet składa się z licznych elementów, takich jak wiele aplikacji i protokołów, różnego typu systemy końcowe i połączenia między nimi, przełączniki pakietów i różnorodne odmiany nośników łączy. Biorąc pod uwagę tę złożoność, czy istnieje jakakolwiek szansa na zorganizowanie architektury sieci lub przynajmniej na jej omówienie? Na szczęście odpowiedź na obie części pytania jest twierdząca.

1.5.1. Architektura warstwowa

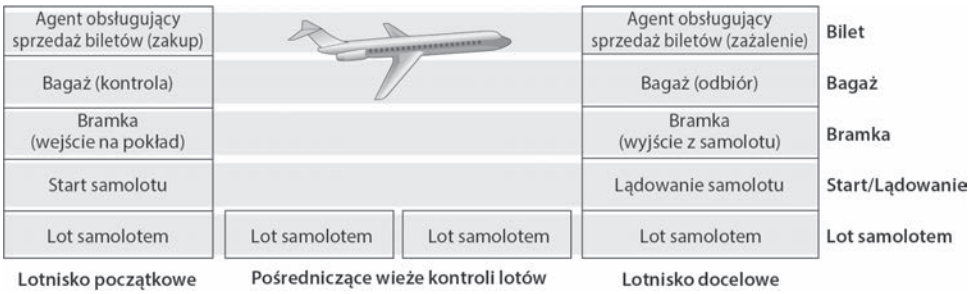
Zanim podejmiemy się próby uporządkowania przemyśleń dotyczących architektury internetu, przeanalizujmy analogię ze świata ludzi. W rzeczywistości nieustannie w codziennym życiu mamy do czynienia ze złożonymi systemami. Wyobraźmy sobie,

że na przykład ktoś poprosi nas o opisanie systemu linii lotniczych. W jaki sposób określiłoby się strukturę charakteryzującą tak skomplikowany twór złożony z agentów obsługujących bilety, kontrolerów bagażu, personelu bramek, pilotów, samolotów, wieży kontroli ruchu i systemu obsługującego w skali globalnej trasy lotu samolotów? Jeden ze sposobów może polegać na opisaniu serii działań podejmowanych samemu lub przez innych, gdy zamierza się polecieć samolotem. W tym celu kupuje się bilet, oddaje bagaż do kontroli, udaje do bramki i ostatecznie wchodzi na pokład samolotu. Samolot startuje i jest kierowany do docelowego lotniska. Po wylądowaniu pasażer przemieszcza się do bramki i odbiera bagaż. Jeśli podróż nie była dla nas zadowalająca, agentowi obsługującemu bilety zgłasza się zażalenie (nic nie załatwiając). Zostało to zilustrowane na rysunku 1.21.



Rysunek 1.21. Działania podejmowane, gdy planuje się polecieć samolotem

Już można zauważyć, występują pewne analogie z siecią komputerową. Pasażer jest przewożony samolotem z miejsca początkowego do docelowego. Pakiet jest przesyłany z hosta źródłowego do docelowego za pośrednictwem internetu. Jednak nie do końca chodzi tu o taką analogię. Na rysunku 1.21 szukamy pewnej *struktury*. Przyglądając się mu, zauważymy, że na obu końcach strzałki znajduje się obsługa biletów. Z obu stron widoczny jest też punkt obsługi bagażu pasażerów posiadających bilet i bramki, do których podchodzą osoby mające bilet i będące już po kontroli bagażu. Pasażerowie, którzy znajdą się za bramką (posiadający bilet i sprawdzony bagaż oraz mający za sobą kontrolę przy bramce), będą mogli znaleźć się w samolocie. Samolot wystartuje, a następnie wyląduje. Dodatkowo osoby te będą uczestniczyły w locie samolotem. Można z tego wywnioskować, że funkcje przedstawione na rysunku 1.21 mogą być postrzegane *horyzontalnie* (rysunek 1.22).



Rysunek 1.22. Poziome warstwy odpowiadające funkcjom systemu linii lotniczych

Na rysunku 1.22 funkcje systemu linii lotniczych podzielono na warstwy, tworząc strukturę pozwalającą omówić podróż samolotem. Warto zauważyć, że każda warstwa połączona ze znajdującymi się niżej oferuje określoną funkcję lub *usługę*. W warstwie agentów obsługujących bilety i niżej jest realizowany transfer pasażera na poziomie kas sprzedaży należących do linii lotniczej. W warstwie bagażowej i niżej ma miejsce transfer pasażera i bagażu na poziomie punktu kontroli i odbioru bagażu. Warto zauważyć, że warstwa bagażowa świadczy usługę tylko tym pasażerom, którzy posiadają już bilet. W warstwie bramki jest wykonywana operacja transferu pasażera i bagażu na poziomie bramek zlokalizowanych w sali odlotów i przylotów. W przypadku warstwy startu i lądowania samolotu ma miejsce przemieszczanie ludzi i bagażu na poziomie pasa startowego. Każda warstwa świadczy usługę przez, po pierwsze, wykonanie w swoim obrębie określonych operacji (na przykład w warstwie bramki jest to wchodzenie pasażerów do samolotu i wychodzenie z niego) i, po drugie, skorzystanie z usług warstwy znajdującej się bezpośrednio poniżej (na przykład w przypadku warstwy bramki zostanie użyta usługa transferu pasażerów po pasie startowym oferowana przez warstwę startu i lądowania samolotu).

Architektura warstwowa umożliwia nam omówienie dobrze zdefiniowanego wybranego elementu dużego i złożonego systemu. Ponadto takie rozwiązanie przedstawia znaczną wartość, ponieważ zapewnia modularność. Dzięki temu znacznie łatwiej zmodyfikować sposób realizowania usługi oferowanej przez warstwę. Dopóki warstwa świadczy taką samą usługę warstwie znajdującej się nad nią i korzysta z identycznych usług oferowanych przez warstwę położoną niżej, reszta systemu nie ulegnie zmianie, gdy wprowadzi się w warstwie modyfikacje dotyczące sposobu realizowania usługi (warto zauważyć, że jest to coś zupełnie innego niż zmiana samej usługi!). Jeśli na przykład zmodyfikowano funkcje bramki (pasażerowie wchodzą do samolotu i wychodzą z niego w kolejności określanej przez ich wzrost), pozostała część systemu linii lotniczych będzie taka sama, ponieważ warstwa bramki nadal pełni identyczną rolę, czyli wpuszczanie ludzi do samolotu i wypuszczanie ich z niego. Po prostu po dokonaniu zmiany warstwa realizuje funkcję w inny sposób. W przypadku dużych i skomplikowanych systemów, które są nieustannie uaktualniane, możliwość zmiany sposobu świadczenia usługi bez wpływu na inne składniki systemu jest kolejną istotną zaletą architektury warstwowej.

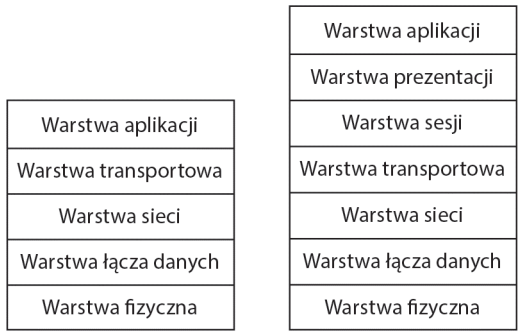
Warstwy protokołów

Na razie wystarczy o liniach lotniczych. Skoncentrujmy się teraz na protokołach sieciowych. Aby zapewnić strukturę wykorzystywaną przy tworzeniu protokołów sieciowych, projektanci umieścili je w **warstwach** wraz z urządzeniami i oprogramowaniem sieciowym obsługującym protokoły. Każdy protokół należy do jednej z warstw, podobnie jak funkcje systemu linii lotniczej widoczne na rysunku 1.22. Ponownie interesują nas **usługi**, które warstwa świadczy warstwie znajdującej się ponad nią. Jest to tak zwany **model usług** warstwy. Tak jak w przypadku systemu linii lotniczej każda warstwa sieci oferuje swoje usługi przez, po pierwsze, wykonanie w swoim obrębie określonych operacji i, po drugie, skorzystanie z usług warstwy znajdującej się bezpośrednio poniżej. Przykładowo, usługi świadczone przez warstwę *n* mogą uwzględniać niezawodne dostarczanie komunikatów z obrzeża jednej sieci do drugiej. Usługa może być realizowana przez zawodną usługę warstwy *n*–1 polegającą na dostarczaniu komunikatów między sieciami, a także przez funkcję warstwy *n*, która wykrywa utracone komunikaty i ponownie je przesyła.

Warstwa protokołów może być obsługiwana na poziomie programowym, sprzętowym lub tak i tak. Protokoły warstwy aplikacji, takie jak HTTP i SMTP, prawie zawsze są umieszczane w oprogramowaniu systemów końcowych. Tak samo jest w przypadku protokołów warstwy transportowej. Ponieważ warstwy fizyczna i łącza danych są odpowiedzialne za obsługę komunikacji realizowanej przy użyciu określonego łącza, zwykle są oferowane przez kartę sieciową (na przykład kartę Ethernet lub Wi-Fi) powiązaną z łączem. Warstwa sieci często jest obsługiwana przez kombinację sprzętu i oprogramowania. Warto zauważyć, że podobnie do funkcji systemu linii lotniczej o architekturze warstwowej, które były rozmieszczone na różnych lotniskach i centrach kontroli lotów tworzących system, protokół warstwy *n* jest *rozproszony* między systemami końcowymi, przełącznikami pakietów i innymi składnikami sieci. Oznacza to, że często w każdym z wymienionych komponentów sieciowych występuje element protokołu warstwy *n*.

Warstwy protokołów oferują korzyści pod względem pojęciowym i strukturalnym [RFC 3439]. Jak pokazaliśmy, warstwy pozwalają w strukturalny sposób omawiać składniki systemu. Modularność ułatwia uaktualnianie składników. Jednak trzeba wspomnieć, że niektórzy naukowcy i inżynierowie sieciowi są zdecydowanie przeciwni stosowaniu warstw [Wakeman 1992]. Potencjalną wadą warstw jest to, że jedna warstwa może powielać funkcje innej położonej niżej. Przykładowo, wiele stosów protokołów zapewnia funkcję usuwania błędów zarówno na poziomie łącza, jak i węzłów końcowych. Kolejną potencjalną wadą jest to, że funkcja jednej warstwy może wymagać danych (na przykład wartości znacznika czasu) dostępnych tylko w innej warstwie. Coś takiego przeczy idei oddzielania warstw.

Kombinacja protokołów różnych warstw jest nazywana **stosem protokołów**. Stos protokołów internetowych jest złożony z protokołów pięciu warstw — fizycznej, łącza danych, sieci, transportowej i aplikacji (rysunek 1.23(a)). Po zajrzeniu do spisu treści można zauważyć, że ogólną strukturę książki oparliśmy na warstwach stosu protokołów internetowych. Zastosowaliśmy przy tym podejście „**od góry do dołu**”, ponieważ zaczynamy od omówienia warstwy aplikacji, a następnie przechodzimy do warstw położonych niżej.



Rysunek 1.23. Stos protokołów internetowych (a) i model referencyjny OSI (b)

Warstwa aplikacji

W warstwie aplikacji znajdują się aplikacje sieciowe i ich protokoły. Internetowa warstwa aplikacji uwzględnia wiele protokołów, takich jak HTTP (obsługuje żądania pobrania stron internetowych), SMTP (umożliwia przesyłanie wiadomości pocztowych) i FTP (pozwala na transfer plików między dwoma systemami końcowymi). Okaże się, że niektóre funkcje sieciowe, takie jak translacja przyjaznych dla użytkownika nazw internetowych systemów końcowych (na przykład *www.ietf.org*) na 32-bitowe adresy sieciowe, są realizowane przy użyciu protokołu warstwy aplikacji oferującego usługę **DNS** (ang. *Domain Name System*). W rozdziale 2. Czytelnik przekona się, że w bardzo prosty sposób można tworzyć własne nowe protokoły warstwy aplikacji.

Protokoły warstwy aplikacji działają w wielu systemach końcowych. Aplikacje z jednego takiego systemu korzystają z protokołu do wymiany pakietów z aplikacjami z innych systemów końcowych. Pakiety informacji w warstwie aplikacji nazywamy **komunikatami**.

Warstwa transportowa

Internetowa warstwa transportowa przesyła komunikaty warstwy aplikacji między klientem i serwerem tworzącym aplikację. Można wyróżnić dwa protokoły internetowej warstwy transportowej — TCP i UDP. Oba są w stanie transportować komunikaty warstwy aplikacji. Protokół TCP zapewnia aplikacjom usługę zorientowaną na połączenie. Tego typu usługa gwarantuje dostarczenie komunikatów warstwy aplikacji do miejsca ich przeznaczenia, a także oferuje kontrolę przepływu (zapewnia zgodność szybkości transmisji nadawcy i odbiorcy). Ponadto protokół TCP dzieli długie komunikaty na krótsze segmenty i zapewnia mechanizm kontroli przeciążenia sieci, dzięki czemu w momencie jego wystąpienia źródłowy węzeł zmniejszy szybkość transmisji. Protokół UDP świadczy swoim aplikacjom usługę bezpołączeniową. Jest ona bardzo uproszczona — nie zapewnia niezawodności ani kontroli przepływu i przeciążenia. W książce pakiet warstwy transportowej określa się mianem **segmentu**.

Warstwa sieci

Internetowa warstwa sieci jest odpowiedzialna za przesyłanie pakietów nazywanych **datagramami** od jednego hosta do drugiego. Protokół internetowej warstwy transportowej (TCP lub UDP) źródłowego hosta przekazuje segment i adres docelowy warstwie sieci, tak jak nadawca zostawiający na poczcie list z adresem odbiorcy. Warstwa sieci oferuje następnie segmentowi usługę polegającą na dostarczeniu go do warstwy transportowej docelowego hosta.

Internetowa warstwa sieci obejmuje znany protokół IP. Definiuje on pola datagramu, a także to, jak mogą być przetwarzane przez systemy końcowe i routery. Istnieje tylko jeden taki protokół i musi być on używany przez wszystkie składniki internetu posiadające warstwę sieci. Internetowa warstwa sieci zawiera również protokoły routingu określające trasy, którymi podążają datagramy między źródłowymi i docelowymi węzłami. W internecie jest stosowanych wiele protokołów routingu. Jak wspomniano w podrozdziale 1.3, internet jest siecią sieci. W obrębie sieci jej administrator może uaktywnić dowolny żądany protokół routingu. Choć warstwa sieci zawiera zarówno protokół IP, jak i kilka protokołów routingu, często jest po prostu nazywana warstwą IP, co odzwierciedla fakt, że protokół IP pełni rolę spoiwa wiążącego internet.

Warstwa łącza danych

Internetowa warstwa sieci przesyła datagram za pośrednictwem serii routerów znajdujących się między źródłowym i docelowym węzłem. Aby przemieścić pakiet z jednego węzła (hosta lub routera) do kolejnego zlokalizowanego na trasie, warstwa sieci korzysta z usług warstwy łącza danych. W szczególności warstwa sieci każdego węzła przekazuje datagram niżej położonej warstwie łącza danych, która dostarcza go do kolejnego węzła znajdującego się na trasie. Warstwa łącza danych tego węzła przemieszcza datagram do warstwy sieci.

To, jakie usługi są zapewniane przez warstwę łącza danych, zależy od jej określonego protokołu, który obsługuje łącze. Przykładowo, niektóre protokoły oferują niezawodne dostarczanie danych na poziomie łącza, które łączy ze sobą węzeł nadawczy i odbiorczy. Warto zauważyć, że taka usługa różni się od usługi niezawodnego dostarczania świadczonej przez protokół TCP, która dotyczy przesyłania danych między dwoma systemami końcowymi. Ethernet, Wi-Fi i protokół DOCSIS z kablowych sieci dostępowych są przykładami protokołów warstwy łącza danych. Ponieważ zwykle datagramy w celu przemieszczenia się od źródłowego do docelowego węzła muszą pokonać kilka łączy, w różnych łączach wchodzących w skład trasy mogą być obsługiwane przez inne protokoły warstwy łącza. Przykładowo, w jednym łączu datagram może być obsługiwany przez protokół Ethernet, natomiast w kolejnym przez protokół PPP. Przez każdy odmienny protokół warstwy łącza danych warstwie sieci zostanie zaoferowana inna usługa. W książce pakiety warstwy łącza danych są nazywane **ramkami**.

Warstwa fizyczna

Zadaniem warstwy łącza danych jest przemieszczanie całych ramek od jednego elementu sieci do kolejnego z nim sąsiadującego, natomiast rolą warstwy fizycznej jest przesyłanie *poszczególnych bitów* ramki pomiędzy dwoma węzłami. Protokoły warstwy fizycznej są zależne od łącza, a także od rzeczywistej szybkości transmisji oferowanej przez nośnik łącza (na przykład skrętka lub światłowód jednomodowy). Przykładowo, standard Ethernet korzysta z wielu protokołów warstwy fizycznej. Poszczególne protokoły są powiązane ze skrętką, kablem koncentrycznym, światłowodem i innymi nośnikami. W każdym przypadku bit jest w inny sposób przesyłany łączem.

Model OSI

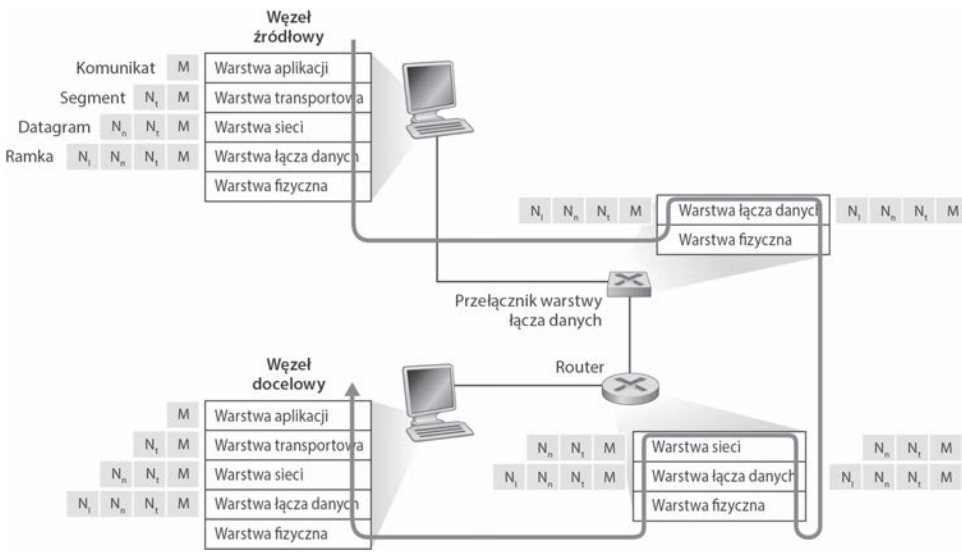
Po szczegółowym omówieniu stosu protokołów internetowych warto wspomnieć, że nie jest to jedyny taki stos. W latach 70. organizacja ISO zaproponowała architekturę sieci opartą na siedmiu warstwach, nazwaną modelem **OSI** (ang. *Open Systems Interconnection*) [ISO 2016]. Model OSI ukształtowano w czasie, kiedy późniejsze protokoły internetowe dopiero powstawały. Pomysłodawcy pierwotnej wersji tego modelu prawdopodobnie nie planowali jego zastosowania w internecie. Mimo to, począwszy od drugiej połowy lat 70., wiele jednostek szkoleniowych oraz uniwersytetów zaczęło uwzględniać zalecenia organizacji ISO i budować kursy na podstawie modelu siedmiowarstwowego. Z uwagi na jego znaczenie w edukacji w początkowym okresie rozwoju sieci komputerowych model ten nadal pojawia się w niektórych podręcznikach i programach kursów.

Siedem warstw referencyjnego modelu OSI (rysunek 1.23(b)) to: warstwa aplikacji, warstwa prezentacji, warstwa sesji, warstwa transportowa, warstwa sieci, warstwa łącza danych i warstwa fizyczna. Funkcje pięciu z tych warstw są podobne jak w ich internetowych odpowiednikach o identycznych nazwach. Dlatego omówimy tylko dwie dodatkowe warstwy referencyjnego modelu OSI — prezentacji i sesji. Warstwa prezentacji ma świadczyć usługi, które umożliwiają komunikującym się aplikacjom zrozumienie wymienianych danych. Te usługi to między innymi kompresja i szyfrowanie informacji (te funkcje są oczywiste), a także opis danych (zwalnia on aplikacje z konieczności obsługi wewnętrznego formatu reprezentacji lub przechowywania danych, który na poszczególnych komputerach może być inny). Warstwa sesji służy do wyznaczania ram czasowych wymiany danych i synchronizowania tego procesu. Między innymi zapewnia narzędzia do tworzenia punktów kontrolnych i systemów przywracania.

Ponieważ internet nie obejmuje dwóch warstw referencyjnego modelu OSI, można zadać kilka ciekawych pytań. Czy usługi świadczone przez te warstwy są nieistotne? Co się stanie, jeśli aplikacja *potrzebuje* jednej z takich usług? W kontekście internetu odpowiedź na oba te pytania jest taka sama — wszystko zależy od autora aplikacji. To on musi zdecydować, czy usługa ma znaczenie. Jeśli *tak*, programista musi wbudować w aplikację odpowiednie mechanizmy.

1.5.2. Kapsułkowanie

Na rysunku 1.24 zilustrowano fizyczną ścieżkę, jaką dane pokonują w dół stosu protokołów nadawczego systemu końcowego, w stosach protokołów pośrednich przełączników warstwy łącza danych i routerów, a następnie w górę stosu protokołów odbiorczego systemu końcowego. Dalej w książce wyjaśnione jest, że w routerach i przełącznikach warstwy łącza danych sprzęt i oprogramowanie sieciowe są uporządkowane w warstwach. Jednak takie urządzenia zwykle nie implementują *wszystkich* warstw stosu protokołów; przeważnie obsługują tylko dolne warstwy. Na rysunku 1.24 pokazano, że przełączniki warstwy łącza danych implementują warstwy pierwszą i drugą, a routery — warstwy od pierwszej do trzeciej. Oznacza to na przykład, że routery internetowe mogą implementować protokół IP (protokół warstwy trzeciej), natomiast przełączniki warstwy łącza danych tego nie potrafią. Dalej zobaczysz, że choć przełączniki warstwy łącza danych nie rozpoznają adresów IP, potrafią obsługiwać adresy warstwy drugiej, takie jak adresy ethernetowe. Zauważ, że hosty implementują wszystkie pięć warstw. Jest to spójne z tym, że w architekturze internetu duża część złożoności występuje na obrzeżach sieci.



Rysunek 1.24. Hosty, routery i przełączniki warstwy łącza danych. Każdy z tych węzłów posiada inny zestaw warstw, które są odzwierciedleniem ich różnego przeznaczenia

Na rysunku 1.24 zilustrowano też ważne zagadnienie, jakim jest **kapsułkowanie**. **Komunikat warstwy aplikacji** (litera K na rysunku 1.24) nadawczego hosta jest przekazywany warstwie transportowej. W najprostszym wariancie warstwa transportowa pobiera komunikat i dodaje do niego dodatkowe informacje (jest to zawartość tak zwanego nagłówka warstwy transportowej, oznaczonego na rysunku 1.24 symbolem N_t), które zostaną wykorzystane przez warstwę transportową węzła odbiorczego.

Komunikat warstwy aplikacji i nagłówek warstwy transportowej razem tworzą **segment warstwy transportowej**. Segment kapsułkuje komunikat warstwy aplikacji. Dodatkowe informacje zawarte w nagłówku mogą uwzględniać dane umożliwiające warstwie transportowej strony odbiorczej dostarczenie komunikatu do odpowiedniej aplikacji, a także bity wykrywania błędów, które pozwalają węzłowi odbiorczemu stwierdzić, czy bity komunikatu zostały zmodyfikowane podczas przesyłania. Warstwa transportowa przekazuje następnie segment warstwie sieci, która dodaje własny nagłówek (na rysunku 1.24 symbol N_s) zawierający takie informacje jak adresy źródłowego i docelowego systemu końcowego. W efekcie jest tworzony **datagram warstwy sieci**. Datagram jest następnie przekazywany warstwie łącza danych, która (oczywiście!) dołączy swój nagłówek i utworzy **ramkę warstwy łącza danych**. Widać więc, że w każdej warstwie pakiet ma pola dwóch rodzajów — nagłówkowe i **z treścią**. Treścią jest zwykle pakiet z wyższej warstwy.

W tym miejscu przydatną analogią będzie wysłanie służbowej notatki z jednego oddziału od drugiego za pośrednictwem ogólnie dostępnej usługi pocztowej. Załóżmy, że Alicja (pracownica jednej filii) chce wysłać wiadomość do Bartka (pracownika innego oddziału). *Notatka* to odpowiednik *komunikatu warstwy aplikacji*. Alicja umieszcza notatkę w biurowej kopercie. Na jej przedzie podaje nazwisko Bartka i dział, w którym ten pracuje. *Biurowa koperta* to odpowiednik *segmentu warstwy transportowej* — obejmuje informacje nagłówkowe (nazwisko Bartka i numer jego działu) i kapsułkuje komunikat warstwy aplikacji (notatkę). Dział oddziału firmy zajmujący się korespondencją otrzymuje notatkę służbową, umieszcza ją w kopercie przeznaczonej do wysyłania za pośrednictwem publicznej poczty, a następnie podaje na kopercie adresy oddziału nadawcy i odbiorcy oraz przykleja znaczek. Tę *pocztową kopertę* można porównać z *datagramem*, który kapsułkuje segment warstwy transportowej (biurowa koperta wraz z jej zawartością). Z kolei segment kapsułkuje oryginalny komunikat (notatka). Poczta dostarcza zwykłą kopertę działowi oddziału firmy odbiorcy zajmującego się korespondencją. Na tym etapie rozpoczyna się proces dekapsułkowania. Pracownicy działu otwierają kopertę i wyjmują biurową kopertę z notatkę służbową. Koperta jest przekazywana właściwemu pracownikowi, Bartkowi, który ją otwiera i wyciąga notatkę.

Proces kapsułkowania może być bardziej złożony od właśnie omówionego. Przykładowo: duży komunikat może zostać podzielony na wiele segmentów warstwy transportowej (one same mogą być podzielone na wiele datagramów warstwy sieci). Po stronie odbiorczej taki segment musi być odtworzony z tworzących go datagramów.

1.6. Sieci pod atakiem

Dostęp do internetu jest obecnie niezbędny wielu instytucjom, w tym dużym i małym firmom, uniwersytetom oraz agencjom rządowym. Także wiele osób prywatnych potrzebuje internetu do celów zawodowych, społecznych i osobistych. Do internetu podłączone są miliardy rzeczy, w tym elektronika do noszenia i urządzenia używane w domach. Jednak obok przydatnych i budzących ekscytację aspektów tej sieci istnieje

też jej ciemna strona. Związana jest ona z „czarnymi charakterami", które próbują utrudnić codzienne życie użytkowników przez uszkodzenie podłączonych do internetu komputerów, naruszenie prywatności i unieruchomienie potrzebnych usług internetowych.

Dziedzina dotycząca bezpieczeństwa sieci obejmuje wiedzę o tym, w jaki sposób napastnicy mogą atakować sieci, a także jak Czytelnicy, przyszli eksperci od sieci komputerowych, mogą chronić je przed napaścią lub — co jeszcze lepsze — projektować nowe, odporne na ataki architektury. Z uwagi na częstotliwość i różnorodność znanych ataków, a także na zagrożenie nowymi, bardziej niebezpiecznymi napaściami, bezpieczeństwo sieci stało się w ostatnich latach głównym zagadnieniem w obszarze sieci komputerowych. Jedną z cech tego podręcznika jest zwracanie uwagi na kwestie związane z bezpieczeństwem.

Ponieważ na tym etapie Czytelnicy nie mają jeszcze bogatej wiedzy z dziedziny sieci komputerowych i protokołów internetowych, zaczniemy od przedstawienia wybranych spośród najczęściej występujących problemów z bezpieczeństwem. Powinno to zaostrzyć apetyt na bardziej dogłębne analizy z dalszych rozdziałów. Zacznijmy od prostego pytania — gdzie mogą wystąpić problemy? W których miejscach sieci komputerowe są podatne na atak? Jakie metody najczęściej stosują napastnicy?

„Czarne charaktery" mogą za pośrednictwem internetu umieścić na hoście szkodliwe oprogramowanie

Użytkownicy podłączają urządzenia do internetu, ponieważ chcą pobierać i wysyłać informacje. Mogą to być różne wartościowe dane — strony internetowe, listy elektroniczne, pliki MP3, rozmowy telefoniczne, przesyłane na żywo nagrania wideo, wyniki wyszukiwania itd. Jednak, niestety, wraz z takimi danymi przekazywane są też niebezpieczne elementy, nazywane ogólnie **szkodliwym oprogramowaniem** (ang. *malware*). Także one mogą dotrzeć do urządzeń i zainfekować je. Kiedy szkodliwe oprogramowanie znajdzie się w urządzeniu, może przeprowadzić dowolne złośliwe operacje, na przykład usunąć pliki lub zainstalować oprogramowanie szpiegowskie rejestrujące prywatne informacje (takie jak numer PESEL, hasła lub listy wciśniętych klawiszy) i wysyłające je — oczywiście za pośrednictwem internetu! — do napastnika. Zainfekowany host może zostać włączony do sieci składającej się z tysięcy innych podobnych urządzeń (nazwa takiej sieci to **botnet**). Kontrolę nad nią sprawują napastnicy, którzy mogą wykorzystać ją do rozsyłania spamu lub przeprowadzania rozproszonych ataków DoS (omawiamy je dalej) na wybrane hosty.

Obecnie duża część szkodliwego oprogramowania ma zdolność do **samoreplikacji**. Po zainfekowaniu hosta takie programy starają się uzyskać za pośrednictwem internetu dostęp do innych węzłów i wykorzystać je do ataku na kolejne hosty. W ten sposób samoreplikujące się szkodliwe oprogramowanie może rozprzestrzeniać się w tempie wykładniczym. Szkodliwe oprogramowanie może rozprzestrzeniać się w formie wirusa lub robaka. **Wirusy** to szkodliwe oprogramowanie, które może zainfekować urządzenie tylko przy współudziale użytkownika. Klasycznym przykładem jest załącznik do listu elektronicznego zawierający szkodliwy kod wykonywalny. Jeśli użytkownik otrzyma i otworzy taki załącznik, nieświadomie uruchomi szkodliwe oprogramowanie.

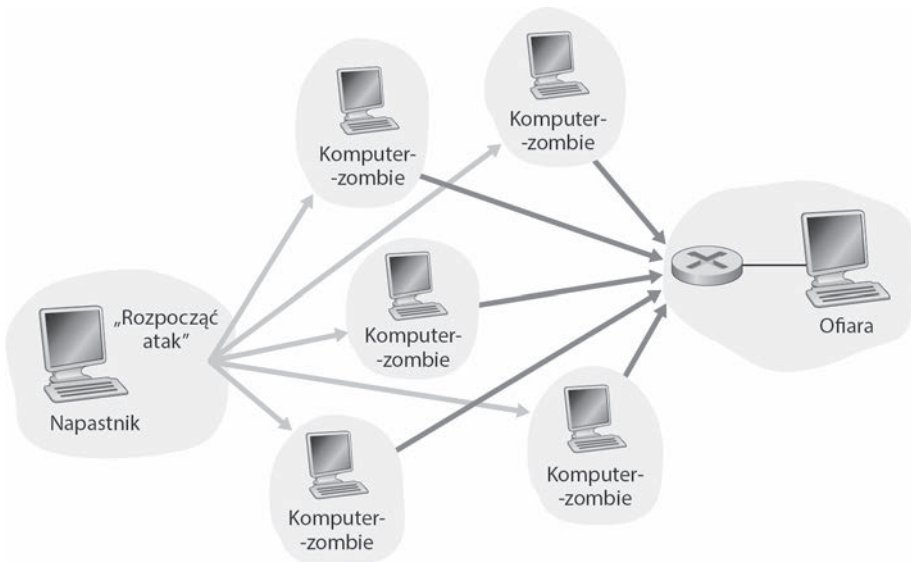
Zwykle wirusy przesyłane w e-mailach mają zdolność do samoreplikacji. Taki wirus po uruchomieniu rozsyła identyczną wiadomość z jednakowymi szkodliwymi załącznikami na przykład do wszystkich odbiorców z książki adresowej danej osoby. **Robaki** to szkodliwe oprogramowanie potrafiące zainfekować urządzenie bez pomocy użytkownika. Załóżmy, że dana osoba korzysta z podatnej na atak aplikacji sieciowej, do której napastnik przesłał szkodliwe oprogramowanie. Aplikacja ta może bez interwencji użytkownika zaakceptować szkodliwy kod wysłany z internetu i uruchomić go, przez co tworzy robaka. Następnie robak za pomocą świeżo zainfekowanego systemu skanuje internet w poszukiwaniu innych hostów z uruchomioną tą samą podatną na atak aplikacją sieciową. Kiedy znajdzie niezabezpieczony węzeł, wysyła do niego swoją kopię. Obecnie szkodliwe oprogramowanie jest wszechobecne, a obrona przed nim dużo kosztuje. Zachęcamy Czytelników, aby w czasie lektury tego podręcznika zastanowili się nad następującym pytaniem — co projektanci sieci komputerowych mogą zrobić, aby zabezpieczyć urządzenia z dostępem do internetu przed atakami ze strony szkodliwego oprogramowania?

„Czarne charaktery" mogą zaatakować serwery i infrastrukturę sieciową

Szeroką kategorię zagrożeń z obszaru bezpieczeństwa stanowią **ataki DoS** (ang. *Denial of Service*, czyli ataki przez odmowę usługi). Jak wskazuje na to nazwa, ataki DoS powodują, że uprawnieni użytkownicy nie mogą korzystać z sieci, hosta lub innego elementu infrastruktury. Na takie ataki narażone są serwery WWW, serwery pocztowe, serwery nazw DNS (omawiamy je w rozdziale 2.) i sieci organizacji. Ataki DoS są w internecie niezwykle częste. Każdego roku przeprowadzane są ich tysiące [Moore 2001]. W witrynie Digital Attack Map możesz wyświetlić wizualizację najpoważniejszych ataków DoS z danego dnia [DAM 2016]. Większość internetowych ataków DoS należy do jednej z trzech kategorii:

- *Wykorzystanie luki*. Te ataki polegają na wysłaniu kilku specjalnie zaprojektowanych wiadomości do podatnej aplikacji lub niezabezpieczonego systemu operacyjnego w docelowym hoście. Jeśli do takiej jednostki napastnik wyśle odpowiednią sekwencję pakietów, może to doprowadzić do zatrzymania usługi lub — co gorsza — awarii hosta.
- *Zalewanie przepustowości* (ang. *bandwidth flooding*). Napastnik wysyła do docelowego hosta dużą liczbę pakietów. Jest ich tak wiele, że łącza dostępowe zaatakowanego celu zostają zablokowane, co uniemożliwia właściwym pakietom dotarcie do serwera.
- *Zalewanie połączenia* (ang. *connection flooding*). Napastnik nawiązuje dużą liczbę połowicznie lub w pełni otwartych połączeń TCP (takie połączenia omawiamy w rozdziale 3.) z docelowym hostem. Taki system może zostać tak spowolniony przez fałszywe połączenia, że zaprzestanie przyjmować właściwe połączenia.

Przyjrzyjmy się teraz dokładniej atakowi przez zalewanie przepustowości. Z analizy opóźnienia i utraty pakietów z punktu 1.4.2 wynika, że jeśli szybkość dostępowa serwera to *S* b/s, napastnik — aby wyrządzić szkody — musi wysyłać dane w podobnym tempie. Jeśli wartość *S* jest bardzo wysoka, pojedyncze źródło ataku może nie wystarczyć do wygenerowania odpowiedniego ruchu. Ponadto jeśli wszystkie dane pochodzą z jednego źródła, router zlokalizowany na początku ścieżki może wykryć atak i zablokować cały ruch z danej lokalizacji, zanim dotrze on w pobliże serwera. W **rozproszonym ataku DoS** (ang. *Distributed DoS* — **DDoS**), zilustrowanym na rysunku 1.25, napastnik kontroluje wiele źródeł i za pomocą każdego z nich wysyła dane do docelowej maszyny. W tej metodzie łączny ruch z wszystkich kontrolowanych źródeł musi mieć natężenie bliskie *S*, aby unieruchomić usługę. Obecnie często przeprowadzane są ataki DDoS z wykorzystaniem botnetów obejmujących tysiące zainfekowanych hostów [DAM 2016]. Ataki DDoS są dużo trudniejsze do wykrycia i powstrzymania niż ataki DoS przeprowadzane za pomocą pojedynczego hosta.



Rysunek 1.25. Atak DDoS

Zachęcamy do zastanowienia się w czasie lektury książki nad następującym pytaniem — co projektanci sieci komputerowych mogą zrobić, aby zabezpieczyć je przed atakami DoS? W dalszej części książki pokazujemy, że trzy różne typy takich ataków wymagają odmiennych środków obrony.

Napastnicy mogą podglądać pakiety

Wielu użytkowników korzysta z internetu za pomocą urządzeń bezprzewodowych, takich jak laptopy z połączeniami Wi-Fi lub małe urządzenia przenośne łączące się z internetem za pośrednictwem sieci komórkowych (omawiamy je w rozdziale 7.).

Choć powszechny dostęp do internetu jest niezwykle wygodny i umożliwia działanie fantastycznych nowych aplikacji dla użytkowników urządzeń mobilnych, powoduje też powstanie poważnej luki w zabezpieczeniach. Dzięki umieszczeniu pasywnego odbiornika w pobliżu bezprzewodowego transmitera napastnik może otrzymać kopię każdego przesyłanego pakietu! Te pakiety mogą zawierać różnorodne poufne informacje, w tym hasła, numery PESEL, tajemnice handlowe i wiadomości osobiste. Pasywny odbiornik rejestrujący kopię każdego przesyłanego pakietu to **program analizujący pakiety** (ang. *packet sniffer*).

Programy analizujące pakiety można zainstalować także w środowisku kablowym. W takich systemach, na przykład w ethernetowej sieci lokalnej, omawiane aplikacje rejestrują kopie wszystkich przesyłanych pakietów. W podrozdziale 1.2 wyjaśniliśmy, że w kablowych technologiach dostępowych także są przesyłane pakiety. Dlatego również takie sieci są podatne na przechwytywanie danych. Ponadto napastnik, który uzyska dostęp do firmowego routera lub łącza dostępowego powiązanego z internetem, może zdołać zainstalować program analizujący pakiety w celu utworzenia kopii wszystkich danych pobieranych przez organizację i wysyłanych z niej. Następnie można poza siecią poddać przechwycone pakiety analizie w poszukiwaniu poufnych informacji.

Oprogramowanie do przechwytywania pakietów jest dostępne bezpłatnie w wielu witrynach internetowych. Istnieją też komercyjne aplikacje tego typu. Wykładowcy prowadzący kursy poświęcone sieciom komputerowym czasem zadają ćwiczenia polegające na napisaniu programu do przechwytywania pakietów lub rekonstrukcji danych z warstwy aplikacji. Zadania związane z programem *Wireshark* [Wireshark 2016] z tej książki (zobacz wprowadzające ćwiczenia dotyczące tej aplikacji zamieszczone na końcu tego rozdziału) są oparte na właśnie takim programie analizującym pakiety!

Ponieważ programy analizujące pakiety są pasywne (nie umieszczają pakietów w kanale), trudno je wykryć. Dlatego przy przesyłaniu pakietów drogą bezprzewodową trzeba pogodzić się z myślą, że napastnik może rejestrować ich kopie. Łatwo się domyślić, że najlepsza obrona przed przechwytywaniem pakietów polega na zastosowaniu szyfrowania. Wykorzystanie kryptografii do zabezpieczania sieci omawiamy w rozdziale 8.

„Czarne charaktery" mogą podać się za zaufaną jednostkę

Zaskakująco łatwo można utworzyć pakiet o dowolnym adresie źródłowym, zawartości i adresie docelowym, a następnie przesłać taki ręcznie przygotowany pakiet do internetu, który sumiennie przekaże go pod wskazany adres. Czytelnicy wkrótce *sami* będą potrafili to zrobić, kiedy przeczytają dalszą część książki! Można wyobrazić sobie niepodejrzewającego niczego odbiorcę (na przykład router w internecie), który otrzyma taki pakiet, uzna (fałszywy) adres źródłowy za prawdziwy, a następnie wykona polecenie zapisane w treści pakietu (na przykład zmodyfikuje tablicę przekazywania). Przesyłanie w internecie pakietów o fałszywym adresie źródłowym to tak zwane **fałszowanie adresu IP**. Jest to tylko jedna z wielu technik, które użytkownik może zastosować do podania się za kogoś innego.

Aby rozwiązać ten problem, należy wykorzystać *uwierzytelnianie punktów końcowych*. Mechanizm ten pozwala się upewnić, że wiadomość podchodzi z danego źródła. Ponownie zachęcamy do zastanowienia się w czasie lektury książki nad zabezpieczaniem aplikacji i protokołów sieciowych. Mechanizmy uwierzytelniania punktów końcowych przedstawiamy w rozdziale 8.

Ten podrozdział kończymy przemyśleniami na temat tego, w jaki sposób internet stał się tak niebezpiecznym miejscem. Odpowiedź sprowadza się do tego, że zaprojektowano go w ten sposób. Oparto go na modelu „grupy wzajemnie ufających sobie użytkowników podłączonych do transparentnej sieci" [Blumenthal 2001]. W tym modelu zabezpieczenia z definicji nie są potrzebne. Wiele aspektów pierwotnej architektury internetu odzwierciedla politykę „wzajemnego zaufania". Na przykład użytkownicy mogą domyślnie wysyłać pakiety do innych osób. Nie trzeba w tym celu prosić odbiorcy o zgodę. Ponadto domyślnie zadeklarowana tożsamość nadawcy jest przyjmowana bez uwierzytelniania.

Jednak współczesny internet z pewnością nie obejmuje „wzajemnie ufających sobie użytkowników". Mimo to nadal muszą oni komunikować się ze sobą, choć nie zawsze sobie wierzą. Użytkownicy mogą chcieć kontaktować się anonimowo, komunikować się za pośrednictwem niezależnych jednostek (na przykład omówionych w rozdziale 2. buforów sieciowych lub opisanych w rozdziale 7. agentów wspomagających działanie sieci mobilnych), a także nie ufać sprzętowi, oprogramowaniu, a nawet powietrzu. W czasie lektury Czytelnicy przekonają się, że stoi przed nimi wiele zadań związanych z bezpieczeństwem — powinni szukać zabezpieczeń przed przechwytywaniem pakietów, podawaniem się za inne punkty końcowe, atakami „człowiek pośrodku", atakami DDoS, szkodliwym oprogramowaniem itd. Warto pamiętać, że komunikacja między wzajemnie ufającymi sobie użytkownikami to raczej wyjątek niż reguła. Witamy w świecie współczesnych sieci komputerowych!

1.7. Historia sieci komputerowych i internetu

W podrozdziałach od 1.1 do 1.6 dokonano przeglądu technologii sieci komputerowych i internetu. Czytelnik powinien teraz dysponować wiedzą wystarczającą do tego, aby zrobić wrażenie na rodzinie i znajomych! Jeśli jednak chciałoby się naprawdę zabłysnąć na kolejnym przyjęciu, powinno się wzbogacić planowane przemówienie o ciekawostki dotyczące fascynującej historii internetu [Segaller 1998].

1.7.1. Rozwój technologii przełączania pakietów: 1961 – 1972

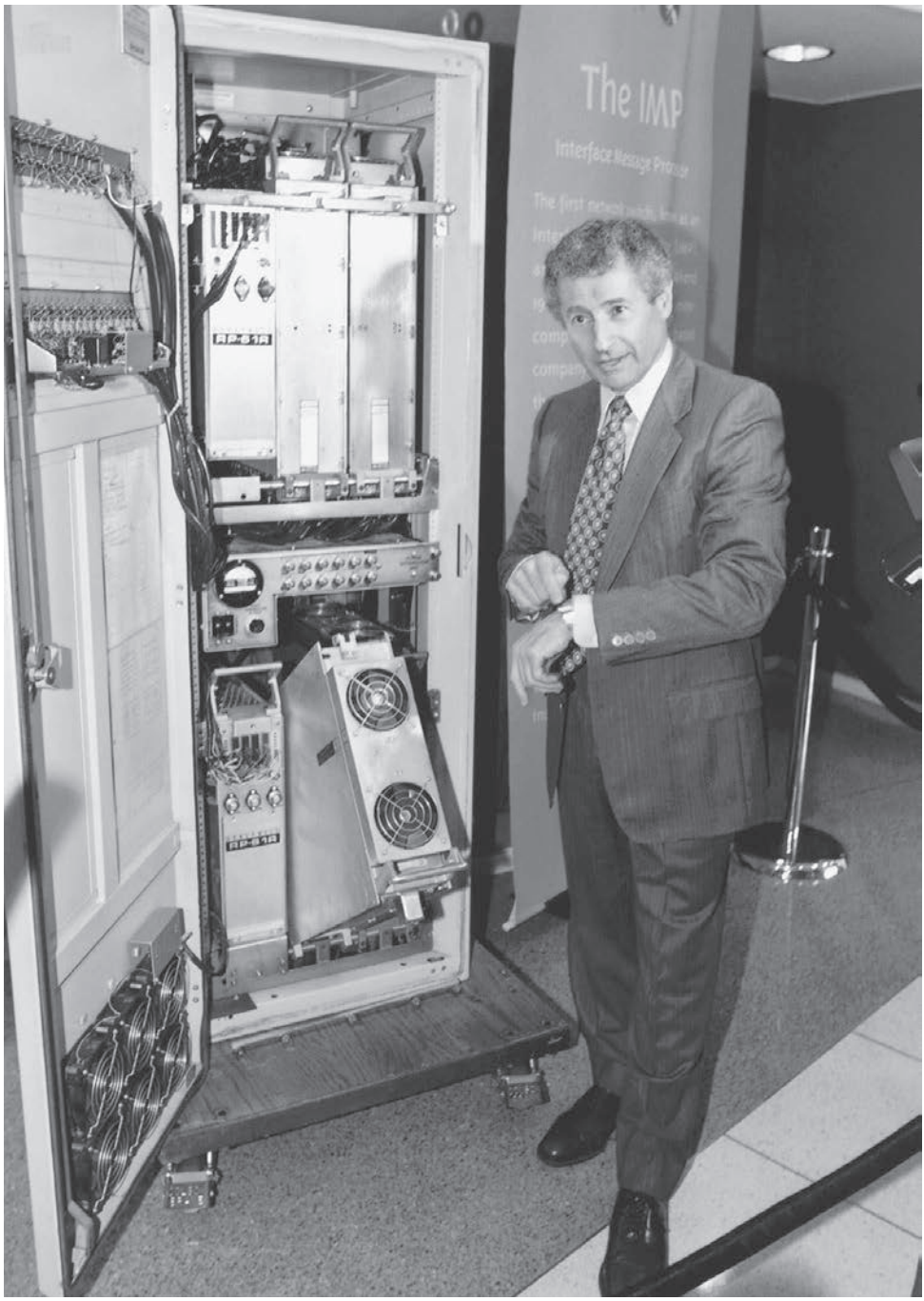
Branżę związaną z sieciami komputerowymi i fundamenty obecnej postaci internetu zaczęto tworzyć z początkiem lat 60., gdy sieć telefoniczna była dominującą w skali światowej siecią komunikacyjną. W podrozdziale 1.3 wspomniano, że sieć telefoniczna korzysta z przełączania obwodów w celu przesłania informacji od nadawcy do odbiorcy (jest to właściwa opcja, biorąc pod uwagę to, że głos między dwoma węzłami

jest transmitowany ze stałą szybkością). Uwzględniając coraz większe znaczenie (i znaczny koszt) komputerów na początku lat 60. i pojawienie się komputerów z podziałem czasu, być może oczywistym (lepiej późno niż wcale!) było zastanowienie się nad tym, w jaki sposób połączyć ze sobą komputery, aby mogły być wykorzystywane przez użytkowników znajdujących się w różnych lokalizacjach geograficznych. Ruch generowany przez takich użytkowników miał raczej charakter *impulsowy*. Polegało to na tym, że po okresach aktywności, takich jak przesłanie polecenia do zdalnego komputera, następowały okresy bezczynności, podczas których oczekiwano na odpowiedź lub ją sprawdzano.

Trzy grupy badawcze z różnych miejsc świata, które nie wiedziały, czym pozostałe grupy się zajmują [Leiner 1998] zaczęły prace nad przełączaniem pakietowym, które miało stanowić efektywną i pewną alternatywę dla przełączania obwodów. Pierwsza opublikowana praca na temat metod przełączania pakietów została napisana przez Leonarda Kleinrocka [Kleinrock 1961; Kleinrock 1964], który wówczas ukończył studia na uczelni MIT. Wykorzystująca teorię kolejkowania praca Kleinrocka znakomicie demonstrowała efektywność technologii przełączania pakietów w przypadku źródeł danych o charakterze impulsowym. W 1964 r. Paul Baran [Baran 1964] rozpoczął w instytucie Rand Institute prace nad zastosowaniem przełączania pakietów do bezpiecznej transmisji głosu w sieciach wojskowych. Donald Davies i Roger Scantlebury, pracujący w angielskiej organizacji **NPL** (ang. *National Physical Laboratory*), także rozwijali swoje pomysły dotyczące przełączania pakietów.

Wyniki prac zrealizowanych na uczelni MIT, w instytucie Rand i organizacji NPL stanowią fundament obecnego internetu. Z internetem jest też związana długa historia działań typu „zbudujmy to i zademonstrujmy", której początki sięgają wczesnych lat 60. J.C.R. Licklider [DEC 1990] i Lawrence Roberts, będący kolegami Kleinrocka z czasów studiów na uczelni MIT, kierowali projektem informatycznym realizowanym przez amerykańską agencję **ARPA** (ang. *Advanced Research Projects Agency*). Roberts udostępnił ogólny plan sieci ARPAnet [Roberts 1967], która była pierwszą siecią komputerową z przełączaniem pakietów i bezpośrednim protoplastą obecnego internetu. Pierwszego maja 1969 r. na uczelni UCLA pod kierownictwem Kleinrocka zainstalowano pierwszy przełącznik pakietów. Wkrótce kolejne trzy takie urządzenia umieszczono w instytucie **SRI** (ang. *Stanford Research Institute*) uniwersytetu UC położonego w Santa Barbara, a także na Uniwersytecie Utah (rysunek 1.26). Nowo powstały prekursor internetu pod koniec 1969 r. liczył cztery węzły. Kleinrock wspomina, że pierwsze użycie sieci w celu zdalnego zalogowania się na komputerze znajdującym się w instytucie SRI przy użyciu komputera zlokalizowanego na uczelni UCLA zakończyło się zawieszeniem systemu [Kleinrock 2004].

W 1972 r. sieć ARPAnet w przybliżeniu liczyła 15 węzłów. W tym samym roku sieć została po raz pierwszy publicznie zademonstrowana przez Roberta Kahna. Stworzono pierwszy protokół **NCP** (ang. *Network-Control Protocol*) [RFC 001], którego użyto w sieci ARPAnet do obsługi komunikacji między systemami końcowymi. Po udostępnieniu protokołu międzywęzłowego można było pisać aplikacje. W 1972 r. Ray Tomlinson z firmy BBN stworzył pierwszy program pocztowy.



Rysunek 1.26. Wczesny przełącznik pakietów

1.7.2. Sieci zastrzeżone i łączenie sieci: 1972 – 1980

Pierwotnie sieć ARPAnet była pojedynczą zamkniętą siecią. Aby można było się skomunikować z hostem sieci ARPAnet, trzeba było go podłączyć do kolejnego procesora IMP. W pierwszej połowie lat 70. poza siecią ARPAnet pojawiły się następujące niezależne sieci z przełączaniem pakietów: ALOHAnet, sieć mikrofalowa łącząca uniwersytety wysp hawajskich [Abramson 1970], a także satelitarną sieć organizacji **DARPA** (ang. *Defense Advanced Research Projects Agency*) [RFC 829] i sieci radiowe [Kahn 1978]; Telenet, komercyjna sieć firmy BBN z przełączaniem pakietów, oparta na technologii sieci ARPAnet; Cyclades, francuska sieć z przełączaniem pakietów stworzona przez Louisa Pouzina [Think 2012]; sieci z podziałem czasu, takie jak Tymnet i GE Information Services, które istniały od końca lat 60. do początku lat 70. [Schwartz 1977]; Sieć SNA firmy IBM (1969 – 1974), która przypominała sieć ARPAnet [Schwartz 1977].

Liczba sieci zwiększała się. Z perspektywy czasu można stwierdzić, że wówczas był odpowiedni moment na to, aby zacząć projektować ogólną architekturę umożliwiającą połączenie sieci. Pionierskie prace w tym zakresie (sponsorowane przez organizację DARPA), w zasadzie polegające na utworzeniu *sieci sieci*, zostały wykonane przez Vintona Cerfa i Roberta Kahna [Cerf 1974]. W celu opisania rezultatu tych prac wymyślono termin *internetting*.

Te architektoniczne podstawy zostały uwzględnione w protokole TCP. Jednak pierwsze wersje protokołu TCP dość znacznie różniły się od jego obecnej postaci. W pierwszych wersjach protokołu połączono mechanizm niezawodnego sekwencyjnego dostarczania danych przy użyciu retransmisji realizowanej przez system końcowy (nadal stanowi część obecnej wersji protokołu) z funkcjami przekazywania (aktualnie są wykonywane przez protokół IP). Wczesne doświadczenia związane z protokołem TCP, a także stwierdzenie potrzeby użycia zawodnej usługi transportu między węzłami bez kontroli przepływu w przypadku zastosowań takich jak pakietowe przesyłanie głosu, spowodowało wydzielenie protokołu IP z protokołu TCP i stworzenie protokołu UDP. Pod koniec lat 70. istniały już koncepcje dotyczące trzech podstawowych protokołów internetowych, które są stosowane obecnie (TCP, UDP i IP).

Poza badaniami związanymi z internetem prowadzonymi przez organizację DARPA realizowanych było wiele innych istotnych działań związanych z sieciami. Na Hawajach Norman Abramson stworzył pakietową sieć radiową ALOHAnet, która umożliwiała komunikowanie się wielu zdalnym jednostkom znajdującym się na wyspach hawajskich. Protokół ALOHA [Abramson 1970] był pierwszym protokołem wielodostępowym zezwalającym użytkownikom rozproszonym w różnych geograficznych lokalizacjach na współużytkowanie pojedynczego nośnika transmisyjnego (częstotliwość radiowa). Projektując protokół Ethernet [Metcalfe 1976] przeznaczony dla kablowych sieci o wspólnym nośniku transmisyjnym, Metcalfe i Boggs korzystali z wyników prac Abramsona dotyczących jego protokołu wielodostępowego. Interesujące jest to, że protokół Ethernet został stworzony przez Metcalfe'a i Boggsa, ponieważ zależało im na połączeniu ze sobą wielu komputerów PC, drukarek i współużytkowanych dysków [Perkins 1994]. Dwadzieścia pięć lat temu, jeszcze sporo przed rewolucją związaną z komputerami PC i eksplozją sieci, Metcalfe i Boggs stworzyli podwaliny pod obecnie stosowane sieci lokalne.

1.7.3. Popularyzacja sieci: 1980 – 1990

Pod koniec lat 70. do sieci ARPAnet było podłączonych w przybliżeniu dwieście hostów. Z końcem lat 80. liczba hostów podłączonych do publicznego internetu (grupa sieci bardzo przypominająca obecny internet) osiągnęła sto tysięcy. Na lata 80. przypadł okres niebywałego rozwoju sieci.

W znacznej części rozwój wynikał z kilku różnych działań mających na celu stworzenie sieci komputerowych łączących ze sobą uniwersytety. Sieć BITNET kilku wyższym uczelniom znajdującym się w północno-wschodniej części Stanów Zjednoczonych zapewniała usługę poczty elektronicznej i transferu plików. Sieć **CSNET** (ang. *Computer Science Network*) została zbudowana w celu umożliwienia komunikacji naukowcom uniwersyteckim, którzy nie dysponowali dostępem do sieci ARPAnet. W 1986 r. stworzono sieć NSFNET, aby udostępnić centra superkomputerowe sponsorowane przez organizację NSF. Początkowo szkielet sieci NSFNET oferował szybkość 56 kb/s. Jednak pod koniec dekady szybkość sieci NSFNET wynosiła już 1,5 Mb/s, dzięki czemu pełniła ona rolę szkieletu łączącego sieci regionalne.

W sieci ARPAnet stosowano wiele ostatecznych elementów obecnej architektury internetu. Pierwszego stycznia 1983 r. w sieci ARPAnet oficjalnie wdrożono nowy standard protokołów TCP/IP przeznaczonych dla hostów, który zastąpił protokół NCP. Przejście z protokołu NCP na protokoły TCP/IP [RFC 801] miało miejsce w amerykański dzień flagi. Tego dnia wszystkie hosty musiały zacząć korzystać z protokołów TCP/IP. Pod koniec lat 80. w protokole TCP wprowadzono istotne rozszerzenia uwzględniające kontrolę przeciążenia realizowaną na poziomie hosta [Jacobson 1988]. Opracowano również system DNS [RFC 1034] służący do translacji przyjaznych dla użytkownika nazw internetowych (na przykład *gaia.cs.umass.edu*) na odpowiadające im 32-bitowe adresy IP.

Równolegle z rozwojem sieci ARPAnet (w zdecydowanej części znajdującej się w Stanach Zjednoczonych) na początku lat 80. Francja rozpoczęła realizację ambitnego projektu Minitel, który miał umożliwić uzyskanie dostępu do sieci z każdego domu. System tworzony w ramach projektu Minitel sponsorowanego przez francuski rząd składał się z publicznej sieci z przełączaniem pakietów (oparta na zestawie protokołów X.25), serwerów Minitel i tanich terminali z wbudowanymi modemami o niewielkiej szybkości. W 1984 r. system ten okazał się wielkim sukcesem, gdy rząd francuski za darmo przekazywał terminal każdemu obywatelowi, który chciał go użyć. Wśród witryn Minitela istniały witryny darmowe świadczące usługi, takie jak książka telefoniczna, a także prywatne witryny, które od każdego użytkownika pobierały opłatę uzależnioną od stopnia korzystania z usług. W szczytowej fazie rozwoju, który przypadł na połowę lat 90., system Minitel oferował ponad 20 000 usług, począwszy od domowej bankowości, a skończywszy na specjalistycznych bazach wyszukujących. System był używany przez ponad 20% mieszkańców Francji i każdego roku generował przychód przekraczający 1 miliard dolarów. Dziesięć lat przed tym, zanim większość Amerykanów w ogóle usłyszała o internecie, system Minitel był dostępny w sporej części francuskich domostw.

1.7.4. Eksplozja internetu: lata 90.

Lata 90. przyniosły kilka wydarzeń, które symbolizują ciągły rozwój i bliską już komercjalizację internetu. Sieć ARPAnet, będąca przodkiem internetu, przestała istnieć. W 1991 r. sieć NSFNET zniosła ograniczenia dotyczące zastosowania jej do celów komercyjnych. Sieć NSFNET zlikwidowano w 1995 r., gdy szkielet internetu zaczął być obsługiwany przez komercyjnych dostawców usług internetowych.

Głównym wydarzeniem lat 90. było pojawienie się technologii **WWW** (ang. *World Wide Web*), która sprawiła, że internet zagościł w domach i firmach milionów osób z całego świata. Technologia ta pełniła też rolę platformy umożliwiającej realizowanie i wdrażanie setek nowych zastosowań, które są dziś powszechnie dostępne. Są to między innymi wyszukiwarki (na przykład Google i Bing), platformy do handlu w internecie (na przykład Amazon i eBay) i sieci społecznościowe (na przykład Facebook).

Technologia WWW została opracowana przez Tima Berners-Lee (pracownik CERN-u) między 1989 i 1991 r. [Berners-Lee 1989] w oparciu o pomysły zawarte we wcześniejszych pracach dotyczących hipertekstu, pisanych w latach 40. przez Vannevara Busha [Bush 1945] i, począwszy od lat 60., przez Teda Nelsona [Xanadu 2012]. Berners-Lee i jego współpracownicy stworzyli pierwotne wersje języka HTML, protokołu HTTP, serwera WWW i przeglądarki, czyli czterech kluczowych składników technologii WWW. Mniej więcej z końcem 1993 r. czynnych było około 200 serwerów WWW. Taka kolekcja serwerów stanowiła jedynie zwiastun tego, co miało nastąpić. W tym czasie kilku naukowców projektowało przeglądarki internetowe z graficznym interfejsem użytkownika. Wśród nich był Marc Andreesen, który razem z Jimem Clarkiem stworzyli firmę Mosaic Communications (później Netscape Communications Corporation) [Cusumano 1998; Quittner 1998]. W 1995 r. studenci uniwersytetów w codziennej pracy używali przeglądarki Netscape do wyświetlania stron internetowych. Mniej więcej w tym czasie duże i małe firmy zaczęły udostępniać serwery WWW i za ich pośrednictwem prowadzić działalność handlową. W 1996 r. Microsoft zaczął prace nad własnymi przeglądarkami. Zapoczątkowało to rywalizację między przeglądarkami firm Netscape i Microsoft, którą kilka lat później wygrała ta druga [Cusumano 1998].

W drugiej połowie lat 90. nastąpił okres niebywałego rozwoju internetu i wprowadzania związanych z nim innowacji. W tym czasie duże korporacje i tysiące nowych firm tworzyło produkty i usługi internetowe. Pod koniec drugiego tysiąclecia internet obsługiwał setki popularnych funkcji, uwzględniających następujące cztery najważniejsze:

- poczta elektroniczna umożliwiająca wysyłanie załączników i udostępniająca wiadomości z poziomu przeglądarki internetowej;
- technologia WWW pozwalająca na przeglądanie stron internetowych i prowadzenie handlu elektronicznego;
- komunikatory zawierające listy kontaktowe;
- wymiana plików (na przykład MP3) za pomocą sieci równoległych (pionierem było narzędzie Napster).

Interesujące jest to, że pierwsze dwa z wyżej wymienionych zastosowań wywodzą się ze społeczności naukowców, natomiast dwa pozostałe zostały zapoczątkowane przez kilku młodych przedsiębiorców.

Okres od 1995 do 2001 r. stoi pod znakiem bardzo dynamicznego wykorzystania internetu przez rynki finansowe. Zanim jeszcze zaczęły przynosić zyski, setki nowych firm internetowych robiły publiczne oferty i ich akcje pojawiały się na giełdach papierów wartościowych. Wiele spółek, które nie miały żadnych znaczących przychodów, wyceniano na miliardy dolarów. „Bańka internetowa" pękła w latach 2000 – 2001 i w efekcie wiele nowych firm zbankrutowało. Niemniej jednak kilka spółek, takich jak Microsoft, Cisco, Yahoo, e-Bay, Google i Amazon, odniosło sukces w branży internetowej.

1.7.5. Ostatnie dokonania

W dalszym ciągu jest utrzymywane duże tempo wprowadzania innowacji w sieciach komputerowych. Postępy są czynione we wszystkich dziedzinach uwzględniających zwiększanie szybkości transmisji w sieciach dostępnych i szkieletowych oraz szybsze routery. Jednak na szczególną uwagę zasługują następujące obszary:

- Od początku naszego wieku widoczne jest dynamiczne rozpowszechnianie się szerokopasmowego dostępu do internetu w domach — nie tylko za pomocą modemów kablowych i sieci DSL, ale też dzięki technologii FTTH (zob. podrozdział 1.2). Szybki dostęp do internetu umożliwił powstanie wielu aplikacji związanych z wideo, w tym dystrybucję filmów nagranych przez użytkowników (na przykład YouTube), strumieniowanie na żądanie filmów i programów telewizyjnych (na przykład Netflix) oraz telekonferencji z udziałem wielu osób (na przykład Skype, Facetime i Google Hangouts).
- Rosnąca powszechność szybkich (od 54 Mb/s w górę) publicznych sieci Wi-Fi oraz dostępu do internetu ze średnią szybkością (na poziomie dziesiątek Mb/s) za pomocą sieci komórkowych 4G nie tylko pozwala na stałe połączenie z internetem w trakcie poruszania się, ale też na wprowadzanie nowych aplikacji opartych na lokalizacji, takich jak Yelp, Tinder, Yik Yak i Waz. Od 2011 r. do internetu podłączonych jest więcej urządzeń bezprzewodowych niż przewodowych. Szybki dostęp bezprzewodowy otworzył drogę do szybkiego wprowadzenia „komputerów" podręcznych (iPhone'ów, urządzeń z Androidem, iPadów itd.) ze stałym i nieograniczonym dostępem do internetu.
- Internetowe sieci społecznościowe (Facebook, Instagram, Twitter i bardzo popularny w Chinach WeChat) prowadzą do powstawania bazujących na internecie olbrzymich sieci ludzi. Liczne z tych sieci społecznościowych są powszechnie używane do wymiany wiadomości i udostępniania zdjęć. Obecnie wielu użytkowników internetu „żyje" głównie w sieciach społecznościowych. Dzięki interfejsom API internetowe sieci społecznościowe stanowią platformy do tworzenia nowych aplikacji i gier działających w sieci.

- W punkcie 1.3.3 opisano, że dostawcy usług internetowych, na przykład Google i Microsoft, opracowali własne rozbudowane sieci prywatne, które nie tylko łączą globalnie rozproszone centra danych tych firm, ale też służą do „obchodzenia" internetu dzięki peeringowi z dostawcami ISP niższych warstw. Wskutek tego Google udostępnia wyniki wyszukiwania i zapewnia dostęp do poczty elektronicznej niemal natychmiast, tak jakby centra danych tej firmy działały na komputerze użytkownika.
- Wiele firm prowadzących handel w internecie uruchamia obecnie swoje aplikacje w chmurze — na przykład w usługach EC2 Amazonu, Application Engine Google'a czy Azure Microsoftu. Także liczne firmy i uczelnie przeniosły swoje aplikacje internetowe (na przykład serwery poczty elektronicznej i WWW) do chmury. Firmy udostępniające chmurę zapewniają aplikacjom nie tylko skalowalne środowiska przetwarzania i składowania danych, ale też bezpośredni dostęp do wydajnych sieci prywatnych tych firm.

1.8. Podsumowanie

W niniejszym rozdziale zawarliśmy ogromną ilość materiału! Przyjrzeliśmy się różnym sprzętowym i programowym składnikom tworzącym internet w szczególności i ogólnie sieci komputerowe. Zaczęliśmy od omówienia obrzeża sieci, omawiając systemy końcowe i aplikacje, a także usługę transportową oferowaną aplikacjom uruchamianym na systemach końcowych. Przyjrzeliśmy się też technologiom warstwy łącza danych i fizycznym nośnikom spotykanym zwykle w sieciach dostępowych. W dalszej kolejności zagłębiliśmy się bardziej w strukturę sieci, analizując jej rdzeń, identyfikując przełączanie pakietów i obwodów jako dwie podstawowe metody transportowania danych w sieciach telekomunikacyjnych. Omówiliśmy również wady i zalety obu tych podejść. Zaprezentowaliśmy też strukturę globalnego internetu, stwierdzając, że jest to sieć sieci. Czytelnicy dowiedzieli się, że hierarchiczna struktura internetu składająca się z dostawców ISP niższych i wyższych warstw może być skalowana, aby możliwe było uwzględnienie w niej tysięcy sieci.

W drugiej części tego wprowadzającego rozdziału omówiliśmy kilka zagadnień odgrywających ważną rolę w przypadku sieci komputerowych. Najpierw przeanalizowaliśmy przyczyny opóźnień i utraty pakietów w sieci z przełączaniem pakietów oraz przepustowość. Utworzyliśmy proste modele ilościowe dotyczące opóźnienia transmisji, propagacji i kolejkowania, jak również przepustowości. Będziemy z nich intensywnie korzystać, rozwiązując problemy zamieszczone w różnych miejscach książki. W dalszej kolejności przyjrzeliśmy się warstwom protokołów i modelom usług, będącym kluczowymi elementami architektury sieci, do których będziemy się w książce odwoływali. Omówiliśmy też wybrane z najczęściej przeprowadzanych w epoce internetu ataków na sieci. Na zakończenie wprowadzenia przedstawiliśmy w skrócie historię sieci komputerowych. Pierwszy rozdział sam w sobie stanowi minikurs z zakresu sieci komputerowych.

W pierwszym rozdziale zawarliśmy naprawdę niebywałą ilość podstawowych informacji! Jeśli Czytelnik czuje się trochę przytłoczony, nie ma powodu do obaw. W kolejnych rozdziałach ponownie znacznie bardziej szczegółowo zostaną omówione wszystkie zagadnienia (to jest obietnica, a nie groźba!). Mamy nadzieję, że po przeczytaniu tego rozdziału Czytelnik orientuje się w składnikach tworzących sieć i terminologii z nią związanej (należy śmiało powracać do niniejszego rozdziału), a także ma ochotę na dalsze pogłębianie wiedzy na temat sieci. Temu mają służyć pozostałe rozdziały książki.

Struktura książki

Przed rozpoczęciem każdej podróży zawsze powinno się spojrzeć na mapę, aby zapoznać się z głównymi drogami i pobliskimi połączeniami. W przypadku podróży, w którą my się wybierzemy, ostatecznym celem jest pełne zrozumienie sieci komputerowych. Nasza mapa drogowa jest złożona z następujących rozdziałów książki:

1. Sieci komputerowe i internet.
2. Warstwa aplikacji.
3. Warstwa transportowa.
4. Warstwa sieci — aspekt danych.
5. Warstwa sieci — aspekt sterowania.
6. Warstwa łącza danych i sieci lokalne.
7. Sieci bezprzewodowe i mobilne.
8. Bezpieczeństwo sieci komputerowych.
9. Multimedia i sieci.

Rozdziały od 2. do 6. to pięć podstawowych rozdziałów książki. Należy zwrócić uwagę na to, że rozdziały te poświęcono czterem górnym warstwom pięciowarstwowego stosu protokołów internetowych. Dodatkowo należy zauważyć, że nasza podróż rozpocznie się od najwyżej położonej warstwy stosu protokołów, czyli od warstwy aplikacji, a następnie będzie przebiegała przez kolejne niższe warstwy. Powodem obrania takiego kierunku podróży (od góry do dołu) jest to, że po zrozumieniu aplikacji będzie łatwiej zrozumieć usługi sieciowe wymagane przez aplikacje. Przyjrzymy się różnym metodom, które można zastosować do wdrażania takich usług przy użyciu architektury sieciowej. Omówienie aplikacji w pierwszej kolejności będzie stanowiło motywację do zapoznania się z resztą książki.

W drugiej części książki (rozdziały od 7. do 9.) skupiono się na trzech wyjątkowo istotnych (i w pewnym stopniu niezależnych) zagadnieniach dotyczących nowoczesnych sieci komputerowych. W rozdziale 7. omówimy technologie bezprzewodowe i mobilne, takie jak bezprzewodowe sieci lokalne (w tym Wi-Fi i Bluetooth), sieci operatorów komórkowych (GSM, 3G i 4G) i sieci mobilne (IP i GSM). W rozdziale 8., poświęconym zabezpieczeniom w sieciach komputerowych, najpierw przeanalizujemy

podstawy szyfrowania i zabezpieczeń sieciowych, a następnie sprawdzimy, jak teoria jest wykorzystywana w praktyce w wielu różnych zastosowaniach internetowych. W ostatnim rozdziale, dotyczącym multimediów w sieci, przyjrzymy się zastosowaniom z obszaru audio-wideo, takim jak telefonia internetowa, wideokonferencje i transmisja strumieniowa magazynowanych multimediów. Zastanowimy się też, w jaki sposób mogą być projektowane sieci z przełączaniem pakietów, aby były w stanie zapewnić aplikacjom audio-wideo stałą jakość usług.

Problemy do rozwiązania i pytania

Rozdział 1. Pytania kontrolne

PODROZDZIAŁ 1.1

1. Jaka jest różnica między hostem i systemem końcowym? Podaj typy systemów końcowych. Czy serwer WWW jest systemem końcowym?
2. Termin *protokół* jest często używany do opisania stosunków dyplomatycznych. Podaj przykład takiego protokołu.
3. Dlaczego standardy są ważne przy tworzeniu protokołów?

PODROZDZIAŁ 1.2

4. Wymień sześć technologii sieci dostępowych. Każdą z nich zaklasyfikuj do jednej z kategorii: sieci dostępowe prywatnych użytkowników, sieci dostępowe firm i bezprzewodowe sieci dostępowe.
5. Czy szybkość transmisji oferowana przez technologię HFC jest dedykowana czy dzielona między wszystkimi użytkownikami? Czy w kanale pobierania urządzenia HFC mogą wystąpić kolizje? Dlaczego do tego może dojść lub dlaczego nie?
6. Wymień technologie dostępowe dla użytkowników prywatnych oferowane w miejscu Twojego zamieszkania. Dla każdej z nich zapisz oficjalnie podawaną szybkość pobierania i wysyłania danych oraz cenę miesięcznego abonamentu.
7. Jaka jest szybkość transmisji danych w ethernetowych sieciach lokalnych?
8. Jakie fizyczne nośniki mogą być wykorzystane w przypadku technologii Ethernet?
9. Na potrzeby sieci dostępowych prywatnych użytkowników używa się modemów telefonicznych, a także modemów HFC, DSL i FTTH. Dla każdej z tych technologii dostępowej określ zakres szybkości transmisji i stwierdź, czy szybkość jest dedykowana, czy współużytkowana.
10. Opisz najpopularniejsze obecnie bezprzewodowe technologie zapewniające dostęp do internetu. Wymień podobieństwa i różnice między nimi.

PODROZDZIAŁ 1.3

11. Załóżmy, że między hostem nadawczym i odbiorczym znajduje się dokładnie jeden przełącznik pakietów. Szybkości transmisji między hostem nadawczym i przełącznikiem oraz między przełącznikiem i hostem odbiorczym wynoszą odpowiednio R_1 i R_2. Przyjmując, że przełącznik stosuje przełączanie pakietów z buforowaniem, jakie będzie całkowite opóźnienie międzywęzłowe w przypadku wysyłania pakietu o długości D (należy zignorować opóźnienia kolejkowania, propagacji i przetwarzania)?
12. Jakie w porównaniu z siecią z przełączaniem pakietów są zalety sieci z przełączaniem obwodów? Jakie w porównaniu z multipleksowaniem FDM są zalety multipleksowania TDM sieci z przełączaniem obwodów?
13. Załóżmy, że użytkownicy korzystają ze wspólnego łącza o szybkości 2 Mb/s. Przyjmijmy ponadto, że każdy z nich przesyła dane ze stałą prędkością 1 Mb/s, jednak robi to tylko przez 20% czasu. Zapoznaj się z omówieniem multipleksowania statystycznego, zamieszczonym w podrozdziale 1.3.
    a. Ilu użytkowników może być obsługiwanych przez sieć z przełączaniem obwodów?
    b. Na potrzeby następnych pytań załóżmy, że w sieci działa przełączanie pakietów. Dlaczego opóźnienie kolejkowania przed łączem będzie prawie niezauważalne, jeśli dane przesyła jednocześnie nie więcej niż dwóch użytkowników? Dlaczego takie opóźnienie się pojawi, kiedy informacje zaczną przekazywać trzy osoby?
    c. Określ prawdopodobieństwo tego, że określony użytkownik transmituje dane.
    d. Załóżmy, że z sieci korzystają trzy osoby. Określ prawdopodobieństwo tego, że w dowolnej chwili wszystkie trzy jednocześnie przesyłają dane. Ustal, przez jaki procent czasu kolejka będzie się wydłużać.
14. Dlaczego dwóch dostawców ISP z tego samego poziomu hierarchii często stosuje peering między sobą? Z czego czerpią zyski operatorzy punktów IXP?
15. Niektórzy dostawcy treści tworzą własne sieci. Opisz sieć Google'a. Po co dostawcy treści budują własne sieci?

PODROZDZIAŁ 1.4

16. Pod uwagę weźmy przesłanie pakietu z hosta nadawczego do odbiorczego za pośrednictwem ustalonej trasy. Wymień składniki opóźnienia międzywęzłowego. Które ze składowych opóźnień są stałe, a które zmienne?
17. Uruchom animację Transmission Versus Propagation Delay ilustrującą opóźnienia transmisji i propagacji (znajduje się ona w witrynie poświęconej książce). Znajdź kombinację szybkości, opóźnienia propagacji i wielkości pakietów, przy której nadawca zakończy transmisję przed dotarciem pierwszego bitu pakietu do nadawcy. Ustal inną kombinację, przy której pierwszy bit pakietu trafi do nadawcy przed zakończeniem wysyłania danych przez nadawcę.

18. Ile czasu zajmie propagacja pakietu o długości 1000 bajtów na dystansie 2500 kilometrów przez łącze o prędkości propagacji na poziomie $2{,}5 \cdot 10^8$ m/s i szybkości transmisji 2 Mb/s? Ujmijmy to bardziej ogólnie — ile potrwa propagacja pakietu o długości D na dystansie d przez łącze o prędkości propagacji na poziomie p i szybkości transmisji S? Czy opóźnienie zależy od wielkości pakietu? Czy ma na nie wpływ szybkość transmisji?
19. Załóżmy, że użytkownik hosta A chce przesłać duży plik do hosta B. Ścieżka między nimi obejmuje trzy łącza, działające z szybkością S_1 = 500 kb/s, S_2 = 2 Mb/s i S_3 = 1 Mb/s.
    a. Jaka będzie przepustowość przy transferze pliku, jeśli w sieci nie są przesyłane żadne inne dane?
    b. Załóżmy, że plik ma cztery miliony bajtów. Ile w przybliżeniu potrwa transfer tego pliku do hosta B (podziel rozmiar przez przepustowość)?
    c. Odpowiedz na pytania z punktów (a) i (b) dla niższej S_2, równej 100 kb/s.
20. Załóżmy, że użytkownik systemu końcowego A chce przesłać duży plik do systemu końcowego B. Opisz na bardzo ogólnym poziomie, w jaki sposób system A podzieli plik na pakiety. Jakie informacje z pakietu wykorzystuje przełącznik pakietów, aby ustalić, którym łączem ma przesłać dalej dane? Dlaczego przełączanie pakietów w internecie przypomina podróżowanie między miastami i pytanie po drodze o wskazówki?
21. Otwórz animację Queuing and Loss ilustrującą kolejkowanie i utratę pakietów, dostępną w witrynie poświęconej książce. Jaka jest maksymalna szybkość emisji i minimalna szybkość transmisji? Jakie jest natężenie ruchu przy tych wartościach? Uruchom animację z takimi ustawieniami i sprawdź, po jakim czasie nastąpi utrata pakietów. Powtórz ten eksperyment. Czy tym razem czas był inny? Dlaczego tak się stało?

PODROZDZIAŁ 1.5

22. Podaj pięć zadań, które może zrealizować warstwa. Czy możliwe jest, aby jedno lub więcej zadań mogło zostać wykonanych przez dwie lub więcej warstw?
23. Jakich pięć warstw wchodzi w skład internetowego stosu protokołów? Jakie są podstawowe funkcje każdej z warstw?
24. Czym jest komunikat warstwy aplikacji? Co to jest segment warstwy transportowej? Czym jest datagram warstwy sieci, a czym ramka warstwy łącza danych?
25. Które warstwy internetowego stosu protokołów są używane przez router? Które warstwy są wykorzystywane przez przełącznik warstwy łącza danych? Które warstwy są stosowane przez host?

PODROZDZIAŁ 1.6

26. Opisz różnice między wirusem i robakiem.
27. Opisz, jak powstają botnety i jak można je wykorzystać do przeprowadzenia ataku DDoS.
28. Załóżmy, że Alicja i Bartek wysyłają do siebie pakiety przez sieć komputerową. Przyjmijmy, że Teresa podłączyła się do sieci, aby móc przechwytywać wszystkie pakiety nadane przez Alicję i wysyłać do Bartka dowolne dane. Teresa może też przechwytywać wszystkie pakiety od Bartka i przekazywać Alicji dowolne informacje. Opisz, jakie szkody Teresa może wyrządzić po zajęciu takiej pozycji.

Problemy

1. Utwórz i opisz protokół warstwy aplikacji, który będzie pośredniczył między zautomatyzowanym bankomatem i komputerem znajdującym się w centrali banku. Protokół powinien umożliwiać weryfikację karty i hasła użytkownika, sprawdzanie salda konta (przechowywane na komputerze w centrali) i realizowanie operacji na koncie (wypłacanie pieniędzy użytkownikowi). Elementy protokołu powinny obsługiwać częsty przypadek, w którym na koncie nie ma wystarczającej kwoty, aby zrealizować wypłatę. Zdefiniuj protokół przez podanie wymienianych komunikatów i działań podejmowanych przez bankomat lub komputer w centrali banku podczas wysyłania i odbierania komunikatów. Narysuj schemat przedstawiający funkcjonowanie protokołu w przypadku zwykłej bezbłędnej operacji wypłaty (skorzystaj z diagramu pokazanego na rysunku 1.2). Wyraźnie określ założenia protokołu dotyczące używanej przez niego międzywęzłowej usługi transportowej.
2. W równaniu 1.1 określony jest wzór na opóźnienie międzywęzłowe przy przesyłaniu pakietu o długości D przez N łączy z szybkością transmisji równą S. Uogólnij ten wzór, aby uwzględnić przesyłanie P takich pakietów jeden po drugim przez N łączy.
3. Weźmy pod uwagę aplikację, która transmituje dane ze stałą szybkością (przykładowo, nadawca generuje N-bitową jednostkę danych co k jednostek czasu, gdzie k jest wartością niewielką i niezmienną). Ponadto, gdy aplikacja zostanie uruchomiona, będzie aktywna przez stosunkowo długi okres. Udziel odpowiedzi na następujące pytania i krótko je uzasadnij:
   a. Czy w przypadku tej aplikacji bardziej odpowiednia byłaby sieć z przełączaniem pakietów czy z przełączaniem obwodów? Dlaczego jedna z sieci będzie lepsza?
   b. Załóżmy, że używana jest sieć z przełączaniem pakietów, w której jedyne przesyłane dane są generowane przez wyżej opisaną aplikację. Dodatkowo przyjmijmy, że suma szybkości przesyłania danych aplikacji jest mniejsza od oferowanej przez każde łącze. Czy będzie potrzebna jakiegoś typu kontrola przeciążenia? Jeśli tak, to dlaczego?

4. Weźmy pod uwagę sieć z przełączaniem obwodów pokazaną na rysunku 1.13. W jej przypadku na każde łącze przypadało *n* obwodów. Oznacz cztery przełączniki jako A, B, C i D (zgodnie z ruchem wskazówek zegara).
   a. Jaka jest maksymalna liczba jednoczesnych połączeń, które w danej chwili mogą być aktywne w sieci?
   b. Przyjmijmy, że wszystkie połączenia są nawiązywane między przełącznikami A i C. Jaka jest maksymalna liczba jednoczesnych aktywnych połączeń?
   c. Załóżmy, że chcemy nawiązać cztery połączenia między przełącznikami A i C oraz cztery między przełącznikami B i D. Czy można tak pokierować dane, aby za pomocą czterech łączy obsłużyć wszystkich osiem połączeń?
5. Zapoznaj się z analogią karawany samochodów przedstawionej w podrozdziale 1.4. Ponownie przyjmijmy, że prędkość propagacji (przemieszczania) wynosi 100 km/h.
   a. Załóżmy, że karawana pokonuje dystans 150 kilometrów. Trasa podróży zaczyna się od wejścia jednej bramki, przebiega przez drugą bramkę i kończy tuż przed trzecią bramką. Jakie jest opóźnienie międzywęzłowe?
   b. Powtórnie rozpatrz przypadek przedstawiony w punkcie (a), ale przyjmij, że zamiast dziesięciu samochodów karawana liczy osiem?
6. Przedstawiony w tym punkcie elementarny problem rozpoczyna rozpatrywanie opóźnień propagacji i transmisji, będących dwoma głównymi zagadnieniami związanymi z przesyłaniem danych w sieci. Weźmy pod uwagę hosty A i B połączone za pomocą pojedynczego łącza o szybkości *S* b/s. Załóżmy, że dwa hosty są oddalone od siebie o *m* metrów, a ponadto, że szybkość propagacji w przypadku łącza wynosi *s* m/s. Host A zamierza wysłać do hosta B pakiet liczący *D* bitów.
   a. Posługując się wielkościami *m* i *s*, określ opóźnienie propagacji o_{prop}.
   b. Używając wielkości *D* i *S*, wyznacz opóźnienie transmisji pakietu o_{trans}.
   c. Ignorując opóźnienia przetwarzania i kolejkowania, sformułuj wyrażenie określające opóźnienie międzywęzłowe.
   d. Zakładając, że host A rozpocznie transmisję pakietu w chwili czasu $t = 0$, gdzie znajdzie się ostatni bit pakietu w chwili czasu $t = o_{trans}$?
   e. Przyjmijmy, że o_{prop} jest większe od o_{trans}. Gdzie znajdzie się pierwszy bit pakietu w chwili czasu $t = o_{trans}$?
   f. Przyjmijmy, że o_{prop} jest mniejsze od o_{trans}. Gdzie znajdzie się pierwszy bit pakietu w chwili czasu $t = o_{trans}$?
   g. Załóżmy, że $s = 2{,}5 \cdot 10^8$, $D = 120$ bitów i $S = 56$ kb/s. Wyznacz odległość *m*, tak aby o_{prop} było równe o_{trans}.
7. W przypadku tego problemu pod uwagę weźmiemy przesyłanie głosu z hosta A do hosta B za pośrednictwem sieci z przełączaniem pakietów (na przykład używanej przez telefonię internetową VoIP). Host A zamienia „w locie" głos z postaci analogowej na cyfrowy strumień bitów transmitowany z szybkością 64 kb/s. Host A grupuje następnie bity w 56-bajtowe pakiety. Między hostami znajduje się tylko jedno łącze.

Jego szybkość transmisji wynosi 2 Mb/s, natomiast opóźnienie propagacji 10 ms. Od razu po utworzeniu pakietu host A wysyła go do hosta B. Natychmiast po odebraniu całego pakietu host B dokonuje konwersji jego bitów na sygnał analogowy. Ile czasu upłynie od utworzenia bitu przez host A (przy użyciu oryginalnego sygnału analogowego) do chwili poddania bitu dekodowaniu przez host B (w ramach konwersji na sygnał analogowy)?

8. Załóżmy, że użytkownicy korzystają ze wspólnego łącza o szybkości 3 Mb/s, a ponadto, że każdy z nich do transmisji wymaga szybkości 150 kb/s. Jednak każdy użytkownik przesyła dane tylko przez 10% czasu pracy (zapoznaj się z omówieniem multipleksowania statystycznego, zamieszczonym w podrozdziale 1.3).
   a. Ilu użytkowników może być obsługiwanych, gdy używa się przełączania obwodów?
   b. W przypadku pozostałej części tego problemu przyjmijmy, że jest wykorzystywane przełączanie pakietów. Określ prawdopodobieństwo tego, że określony użytkownik transmituje dane.
   c. Załóżmy, że istnieje 120 użytkowników. Określ prawdopodobieństwo tego, że w dowolnej chwili czasu dokładnie n użytkowników jednocześnie przesyła dane (*wskazówka:* należy skorzystać z rozkładu dwumianowego).
   d. Wyznacz prawdopodobieństwo tego, że 21 lub więcej użytkowników jednocześnie transmituje dane.
9. Pod uwagę weźmy omówienie multipleksowania statystycznego, znajdujące się w podrozdziale 1.3. Zawarto w nim przykład łącza o szybkości 1 Mb/s. Gdy użytkownicy są zajęci, generują dane z szybkością 100 kb/s. Jednak prawdopodobieństwo tego, że tak jest, wynosi zaledwie $p = 0{,}1$. Załóżmy, że łącze 1 Mb/s zostanie zastąpione łączem o szybkości 1 Gb/s.
   a. Jaka jest maksymalna liczba N użytkowników, którzy jednocześnie mogą być obsługiwani przez sieć z przełączaniem obwodów?
   b. Weźmy pod uwagę przełączanie pakietów i populację liczącą M użytkowników. Podaj wzór (używając wielkości p, M i N) określający prawdopodobieństwo tego, że więcej niż N użytkowników przesyła dane.
10. Załóżmy, że pakiet o długości D z systemu końcowego A jest przesyłany przez trzy łącza do docelowego systemu końcowego. Te trzy łącza są połączone dwoma przełącznikami pakietów. Niech d_i, s_i i S_i oznaczają długość, szybkość propagacji i szybkość transmisji łącza i, gdzie $i = 1, 2, 3$. Opóźnienie w każdym przełączniku pakietów wynosi o_{prze}. Przy założeniu, że nie występują opóźnienia kolejkowania, podaj łączne opóźnienie międzywęzłowe dla pakietu. Posłuż się wartościami d_i, s_i i S_i (dla $i = 1, 2, 3$) oraz D. Teraz przyjmijmy, że pakiet ma długość 1500 bajtów, szybkość propagacji łączy wynosi $2{,}5 \cdot 10^8$ m/s, szybkość transmisji jest równa 2 Mb/s, a opóźnienie w przełączniku pakietów to 3 milisekundy. Długości łączy to odpowiednio 5000, 4000 i 1000 kilometrów. Jakie będzie opóźnienie międzywęzłowe dla tych wartości?

11. Przyjmijmy, że w poprzednim problemie $S_1 = S_2 = S_3 = S$, a $o_{prze} = 0$. Załóżmy też, że przełącznik pakietów nie buforuje ich, ale natychmiast przesyła każdy otrzymany bit, nie czekając na pobranie całego pakietu. Ile wtedy wyniesie opóźnienie międzywęzłowe?
12. Przełącznik pakietów odbiera pakiet i określa łącze wyjściowe, do którego należy przekazać dane. W momencie nadejścia pakietu inny pakiet został już w połowie przesłany określonym łączem, a cztery następne oczekują na transmisję. Pakiety są przesyłane w kolejności ich nadejścia. Przyjmijmy, że wszystkie pakiety mają po 1500 bajtów, a szybkość łącza jest równa 2 Mb/s. Ile wyniesie opóźnienie kolejkowania dla nowego pakietu? Ujmijmy to bardziej ogólnie — jakie jest opóźnienie kolejkowania, jeśli wszystkie pakiety mają długość D, szybkość transmisji wynosi S, zostało przesłanych x bitów obecnie przekazywanego pakietu, a w kolejce znajduje się n pakietów?
13. (a) Załóżmy, że N pakietów jednocześnie dotarło do łącza, które obecnie nie przesyła żadnych innych danych. Ponadto żadne pakiety nie oczekują w kolejce na transmisję. Każdy pakiet ma długość D, a szybkość transmisji łącza wynosi S. Jakie jest średnie opóźnienie kolejkowania dla N pakietów?

    (b) Teraz przyjmijmy, że co DN/S sekund do każdego łącza dociera N takich pakietów. Jakie jest średnie opóźnienie kolejkowania pakietu?
14. Uwzględnijmy opóźnienie kolejkowania bufora routera. Niech I określa natężenie ruchu wynoszące Da/S. Przyjmijmy, że dla $I < 1$ opóźnienie kolejkowania przyjmuje postać $ID/S\ (1–I)$.
    a. Podaj wzór dla całkowitego opóźnienia, które jest sumą opóźnienia kolejkowania i transmisji.
    b. Sporządź wykres całkowitego opóźnienia jako funkcję D/S.
15. Niech a oznacza częstotliwość docierania pakietów do łącza (w pakietach na sekundę), a μ — szybkość transmisji łącza (w tej samej jednostce). Na podstawie wzoru na opóźnienie całkowite (czyli opóźnienie kolejkowania plus opóźnienie transmisji) wyprowadzonego w poprzednim zadaniu podaj wzór na opóźnienie całkowite wyrażone w a i μ.
16. Weźmy pod uwagę bufor routera przed łączem wyjściowym. W tym problemie wykorzystasz wzór Little'a (jest to znana formuła z teorii kolejkowania). Załóżmy, że N oznacza średnią liczbę pakietów w buforze plus przesyłany pakiet, natomiast a to szybkość docierania pakietów do łącza. Niech d określa średnie opóźnienie całkowite (czyli opóźnienie kolejkowania plus opóźnienie transmisji) dla pakietu. Wzór Little'a to $N = a{\cdot}d$. Przyjmijmy, że średnio bufor zawiera 10 pakietów, a średnie opóźnienie kolejkowania dla pakietu wynosi 10 milisekund. Szybkość transmisji łącza to 100 pakietów na sekundę. Jaka będzie szybkość docierania pakietów według wzoru Little'a przy założeniu, że nie występuje ich utrata?

17. (a) Uogólnij wzór 1.2 zamieszczony w punkcie 1.4.3 i określający opóźnienie międzywęzłowe, aby uwzględniał różne szybkości przetwarzania, szybkości transmisji i opóźnienia propagacji.

    (b) Powtórnie rozpatrz przypadek przedstawiony w punkcie (a), ale przyjmij dodatkowo, że w każdym węźle występuje średnie opóźnienie kolejkowania o_{kolejk}.

18. Za pomocą narzędzia *Traceroute* w trzech różnych porach dnia prześledź trasę pokonywaną przez pakiety między węzłem źródłowym i docelowym, które znajdują się na tym samym kontynencie.

    a. W przypadku każdej z trzech pór dnia dla opóźnień całkowitej trasy określ średnie i standardowe odchylenie.

    b. W przypadku każdej z trzech pór dnia wyznacz liczbę routerów znajdujących się na drodze pakietów. Czy poszczególne ścieżki różnią się od siebie?

    c. Spróbuj określić liczbę sieci dostawców ISP, przez które narzędzie *Traceroute* przesyła pakiety od węzła źródłowego do docelowego. Routery o podobnych nazwach i (lub) adresach IP powinny być uznane za należące do tych samych dostawców ISP. Czy w przeprowadzonych doświadczeniach największe opóźnienia występują w przypadku interfejsów węzłów znajdujących się między sąsiadującymi ze sobą sieciami dostawców ISP?

    d. Powyższe trzy punkty powtórnie wykonaj dla węzłów źródłowego i docelowego zlokalizowanych na różnych kontynentach. Porównaj wyniki uzyskane dla węzłów położonych na tym samym i na innych kontynentach.

19. (a) Odwiedź witrynę *http://www.traceroute.org* i zbadaj trasy między dwoma miastami we Francji i tym samym docelowym hostem w Stanach Zjednoczonych. Ile łączy na obu trasach jest takich samych? Czy używane jest to samo łącze transatlantyckie?

    (b) Wykonaj to samo ćwiczenie co w punkcie (a), ale tym razem wybierz jedno miasto we Francji i jedno w Niemczech.

    (c) Wybierz miasto w Stanach Zjednoczonych i zbadaj trasy do dwóch hostów z różnych chińskich miast. Ile wspólnych łączy występuje w obu trasach? Czy trasy zaczynają się różnić przed dotarciem do Chin?

20. Zastanów się nad przykładem dotyczącym przepustowości zilustrowanym na rysunku 1.20(b). Załóżmy, że zamiast 10 jest M par klient-serwer. S_s, S_c i S to oznaczenia szybkości łącza serwera, łącza klienta i łącza sieci. Przyjmijmy, że wszystkie łącza mają wystarczającą pojemność, a jedyny ruch w sieci generuje M par klient-serwer. Wyprowadź ogólny wzór na przepustowość wyrażoną miarami S_s, S_c, S i M.

21. Przyjrzyj się rysunkowi 1.19(b). Załóżmy, że jest M ścieżek między serwerem i klientem. Żadna z nich nie wykorzystuje wspólnych łączy. Ścieżka k (k = 1, …, M) obejmuje N łączy o szybkości transmisji S_1^k, S_2^k, …, S_N^k. Jaka będzie maksymalna przepustowość serwera, jeśli może on używać tylko jednej ścieżki do przesyłania

danych klientowi? Jak zmieni się ta przepustowość, jeżeli serwer będzie mógł wykorzystać do transferu danych wszystkie M ścieżek?

22. Zastanów się nad rysunkiem 1.19(b). Załóżmy, że dla każdego łącza między serwerem i pakietem występuje prawdopodobieństwo utraty pakietu równe p, a prawdopodobieństwa dla poszczególnych łączy są od siebie niezależne. Jakie jest prawdopodobieństwo poprawnego odebrania przez nadawcę pakietu wysłanego przez serwer? Jeśli pakiet zostanie utracony na trasie między serwerem i klientem, serwer ponownie wyśle dane. Ile średnio razy serwer musi ponownie przesłać pakiet, aby dotarł on do klienta?
23. Przyjrzyj się rysunkowi 1.19(a). Przyjmijmy, że wąskim gardłem na trasie między serwerem i klientem jest pierwsze łącze. Działa ono z szybkością S_s b/s. Załóżmy, że użytkownik wysyła jeden za drugim dwa pakiety z serwera do klienta, a daną trasą nie są przekazywane żadne inne dane. Każdy z tych pakietów ma wielkość D bitów, a w obu łączach występuje takie samo opóźnienie propagacji równe o_{prop}.
    a. Ile wynosi czas między dotarciem kolejnych pakietów (ile czasu upłynie od pobrania ostatniego bitu pierwszego pakietu do otrzymania ostatniego bitu drugiego pakietu)?
    b. Teraz załóżmy, że wąskim gardłem jest drugie łącze (czyli $S_c < S_s$). Czy możliwe jest, że drugi pakiet znajdzie się w kolejce wejściowej drugiego łącza? Odpowiedź uzasadnij. Następnie przyjmijmy, że serwer wysyła drugi pakiet po T sekundach od wysłania pierwszego. Jak duże musi być T, aby przed drugim łączem nie utworzyła się kolejka? Wyjaśnij odpowiedź.
24. Załóżmy, że musisz pilnie przesłać 40 terabajtów danych z Gdańska do Krakowa. Do transferu możesz wykorzystać dedykowane łącze o szybkości 100 Mb/s. Czy wolisz przesłać dane tym łączem, czy raczej skorzystasz z usług firmy kurierskiej zapewniającej dostawę na następny dzień? Odpowiedź uzasadnij.
25. Załóżmy, że dwa hosty A i B są od siebie oddalone o 20 000 kilometrów i połączone bezpośrednim łączem o szybkości S = 2 Mb/s. Przyjmijmy też, że w przypadku łącza szybkość propagacji wynosi $2{,}5 \cdot 10^8$ m/s.
    a. Oblicz iloczyn przepustowości i opóźnienia propagacji $S \cdot o_{\text{prop}}$.
    b. Weźmy pod uwagę przesłanie pliku liczącego 800 000 bitów z hosta A do hosta B. Załóżmy, że plik zostanie wysłany jako jeden wielki komunikat. Jaka jest maksymalna liczba bitów, które w danej chwili znajdą się w łączu?
    c. Przedstaw interpretację iloczynu przepustowości i opóźnienia propagacji.
    d. Jaka jest długość bita (w metrach) znajdującego się w łączu? Czy jest dłuższy od boiska piłkarskiego?
    e. Uwzględniając szybkość propagacji s, szybkość transmisji S i długość łącza m, utwórz zależność określającą długość bitu.

26. Nawiązując do problemu 25., załóżmy, że można zmodyfikować wartość S. Dla jakiej wartości S długość bitu będzie równa długości łącza?
27. Weźmy pod uwagę problem 25., z tym że niech szybkość łącza wynosi S = 1 Gb/s.
    a. Oblicz iloczyn przepustowości i opóźnienia propagacji $S \cdot o_{prop}$.
    b. Weźmy pod uwagę przesłanie pliku liczącego 800 000 bitów z hosta A do hosta B. Załóżmy, że plik zostanie wysłany jako jeden wielki komunikat. Jaka jest maksymalna liczba bitów, które w danej chwili znajdą się w łączu?
    c. Jaka jest długość bita (w metrach) znajdującego się w łączu?
28. Ponownie odnieśmy się do problemu 25.
    a. Ile czasu zajmie wysłanie pliku przy założeniu, że operacja jest wykonywana bez żadnych przerw?
    b. Przyjmijmy, że plik podzielono na 20 pakietów, z których każdy liczy 40 000 bitów. Ponadto załóżmy, że dla każdego pakietu węzeł odbiorczy odsyła potwierdzenie i czas jego transmisji jest nieznaczny. Przyjmijmy wreszcie, że nadawca nie może wysłać kolejnego pakietu, dopóki dla poprzedniego nie otrzyma potwierdzenia dostarczenia. Ile czasu zajmie przesłanie pliku?
    c. Porównaj ze sobą wyniki uzyskane w punktach (a) i (b).
29. Załóżmy, że między geostacjonarnym satelitą i jego naziemną stacją bazową znajduje się łącze mikrofalowe o szybkości 10 Mb/s. Co minutę satelita wykonuje cyfrowe zdjęcie i przesyła je do stacji. Przyjmijmy, że szybkość propagacji wynosi $2{,}4 \cdot 10^8$ m/s.
    a. Jakie jest dla łącza opóźnienie propagacji?
    b. Jaka jest wartość iloczynu przepustowości i opóźnienia propagacji $R \cdot o_{prop}$?
    c. Niech x oznacza wielkość zdjęcia. Jaka musi być dla łącza mikrofalowego minimalna wartość x, aby transmisja była ciągła?
30. Pod uwagę weźmy analogię dotyczącą podróży samolotem przytoczoną podczas omawiania warstw w podrozdziale 1.5, a także dodawanie nagłówków do zawartości pakietów przemieszczających się przez kolejne warstwy stosu protokołów. Czy istnieją informacje odpowiadające danym znajdującym się w nagłówkach, które są kojarzone z bagażem i pasażerami podczas pokonywania przez nich kolejnych poziomów stosu protokołów systemu linii lotniczej?
31. W nowoczesnych sieciach z przełączaniem pakietów źródłowy host segmentuje długie komunikaty warstwy aplikacji (na przykład obraz lub plik audio) na mniejsze pakiety i umieszcza je w sieci. Host odbiorczy łączy pakiety z powrotem, aby uzyskać oryginalny komunikat. Proces ten określa się mianem *segmentacji komunikatów*. Na rysunku 1.27 zilustrowano transport komunikatu między węzłami bez stosowania segmentacji i z jej wykorzystaniem. Pod uwagę weźmy komunikat liczący $8 \cdot 10^6$ bitów, który zostanie przesłany między źródłowym i docelowym hostem (rysunek 1.27). Załóżmy, że każde łącze widoczne na rysunku ma szybkość 2 Mb/s. Opóźnienia propagacji, kolejkowania i przetwarzania należy zignorować.

a. Rozważmy przesłanie komunikatu między źródłowym i docelowym węzłem *bez* zastosowania segmentacji. Ile czasu zajmie przemieszczenie komunikatu ze źródłowego hosta do pierwszego przełącznika pakietów? Pamiętając o tym, że każdy przełącznik korzysta z przełączania pakietów z buforowaniem, jaki będzie całkowity czas potrzebny na przesłanie komunikatu między hostem źródłowym i docelowym?

b. Przyjmijmy teraz, że komunikat jest segmentowany na 800 pakietów, z których każdy liczy 10 000 bitów. Ile potrwa przemieszczenie pierwszego pakietu ze źródłowego hosta do pierwszego przełącznika? Gdy pierwszy pakiet jest przesyłany między pierwszym i drugim przełącznikiem, drugi pakiet jest transmitowany od źródłowego hosta do pierwszego przełącznika. W jakim czasie drugi pakiet zostanie całkowicie odebrany przez pierwszy przełącznik?

c. Ile czasu zajmie przesłanie pliku między źródłowym i docelowym hostem, gdy zostanie przeprowadzona segmentacja komunikatu? Uzyskany wynik porównaj z odpowiedzią udzieloną w przypadku punktu (a) i skomentuj to.

d. Jakie — oprócz ograniczenia opóźnienia — są powody segmentacji komunikatów?

e. Omów wady segmentacji komunikatów.

Rysunek 1.27. Transport komunikatu między węzłami: (a) — bez zastosowania segmentacji; (b) — z wykorzystaniem segmentacji

32. Poeksperymentuj z interaktywną animacją ilustrującą segmentowanie komunikatów, Message Segmentation, która znajduje się na stronie internetowej poświęconej książce. Czy opóźnienia występujące w animacji odpowiadają tym, które pojawiły się w poprzednim problemie? Jak opóźnienia propagacji powiązane z łączem wpływają na ogólne opóźnienie międzywęzłowe w przypadku przełączania pakietów (z segmentacją komunikatów) i przełączania komunikatów?

33. Pod uwagę weźmy wysłanie między hostami A i B dużego pliku liczącego *F* bitów. Między hostami znajdują się trzy łącza i dwa przełączniki. Łącza nie są przeciążone (oznacza to, że nie występują opóźnienia kolejkowania). Host A dzieli plik na segmenty, z których każdy liczy *Seg* bitów i do każdego dodaje 80-bitowy nagłówek. W efekcie powstają pakiety o długości $S = 80+Seg$ bitów. Każde łącze oferuje szybkość transmisji wynoszącą *S* b/s. Określ wartość *Seg*, która zminimalizuje opóźnienie związane z przesyłaniem pliku między hostami A i B. Opóźnienie propagacji należy zignorować.
34. Firma Skype oferuje usługę umożliwiającą wykonywanie połączeń z komputerów PC na zwykły telefon. To oznacza, że połączenie głosowe musi przejść przez internet i sieć telefoniczną. Wyjaśnij, jak może to być realizowane.

Ćwiczenie realizowane za pomocą narzędzia Wireshark

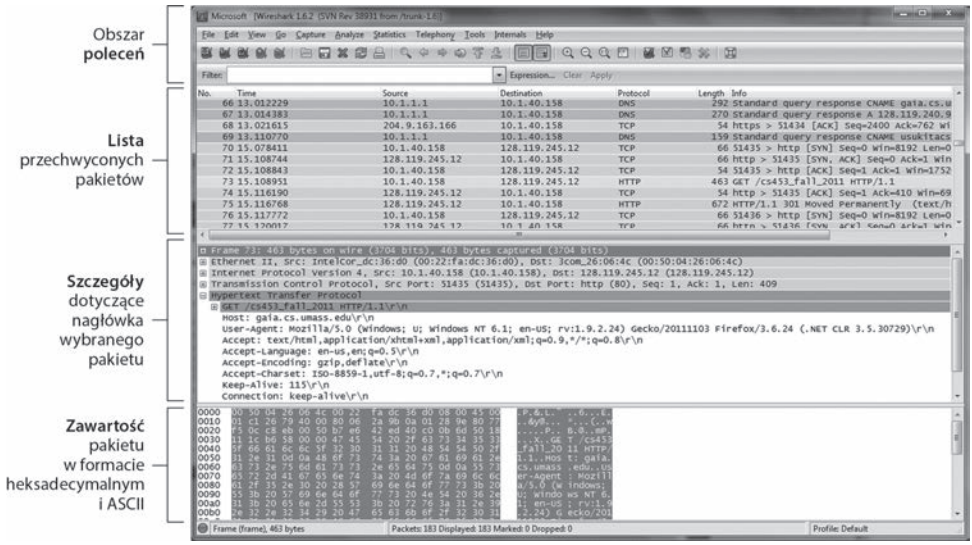
Powiedz mi, a zapomnę. Pokaż mi, a zapamiętam.
Zaangażuj mnie, a zrozumiem

— przysłowie chińskie

Poziom zrozumienia protokołów sieciowych często można znacznie zwiększyć przez zobaczenie ich w akcji i poeksperymentowaniu z nimi (obserwacja sekwencji komunikatów wymienianych między dwoma składnikami protokołu, zagłębianie się w szczegóły dotyczące działania protokołu, powodowanie wykonywania przez protokół określonych operacji, a także monitorowanie ich efektów). Można to zrealizować przy użyciu symulacji lub rzeczywistego środowiska sieciowego, jakim jest internet. W przypadku pierwszego wariantu zostaną użyte interaktywne animacje zamieszczone na stronie internetowej poświęconej książce. Drugi wariant będzie realizowany w ramach ćwiczeń wykorzystujących narzędzie *Wireshark*. Czytelnik będzie obserwował protokoły sieciowe na poziomie komputera współpracujące i wymieniające się komunikatami ze składnikami protokołów funkcjonujących w innym miejscu internetu. Oznacza to, że posiadany komputer będzie integralną częścią rzeczywistych doświadczeń. Przez obserwację będzie się zdobywało wiedzę.

Podstawowym narzędziem umożliwiającym obserwację komunikatów wymienianych między aktywnymi składnikami protokołu jest **program analizujący pakiety**. Tego typu narzędzie w pasywny sposób kopiuje komunikaty, które są wysyłane i odbierane przez komputer. Ponadto program wyświetla zawartość różnych pól protokołu znajdujących się w przechwyconych komunikatach. Na rysunku 1.28 przedstawiono zrzut ekranu okna programu *Wireshark* służącego do analizowania pakietów. Jest to darmowe narzędzie dostępne w wersji dla systemów Windows, Linux/Unix i Mac. W książce zamieszczono ćwiczenia wykorzystujące program *Wireshark*, które umożliwią Czytelnikowi zapoznanie się z kilkoma protokołami omawianymi w danym rozdziale.

Niniejsze pierwsze ćwiczenie polega na pobraniu i zainstalowaniu kopii narzędzia *Wireshark*, wyświetleniu strony internetowej, a następnie przechwyceniu i przeanalizowaniu komunikatów protokołu wymienianych między przeglądarką i serwerem WWW.



Rysunek 1.28. Okno programu Wireshark

Szczegółowe informacje o pierwszym ćwiczeniu (wraz z instrukcjami pobierania i instalowania programu Wireshark) znajdziesz na stronie *http://www.pearsonhighered.com/cs-resources/*.

WYWIAD Z…

Leonard Kleinrock

Leonard Kleinrock jest profesorem informatyki na Uniwersytecie Kalifornijskim znajdującym się w Los Angeles. W 1969 r. jego uczelniany komputer stał się pierwszym węzłem internetu. Stworzone przez niego w 1961 r. podstawy technologii przełączania pakietów stanowiły fundament internetu. Leonard ukończył inżynierię elektryczną na uczelni CCNY (*City College of New York*), a na uczelni MIT zdobył tytuł magistra i doktora z tej samej dziedziny.

Co sprawiło, że postanowił Pan specjalizować się w technologiach związanych z sieciami i internetem?

Po uzyskaniu w 1959 r. na uczelni MIT tytułu doktora zorientowałem się, że większość znajomych ze studiów zajmuje się badaniami z zakresu teorii informacji i kodowania. Na uczelni MIT pracował Claude Shannon, znakomity naukowiec, który był pionierem w obu dziedzinach i poradził sobie już z większością istotnych problemów. Problemy badawcze, które nadal nie były rozwiązane, zaliczały się do trudnych, ale też mniej ważnych. W związku z tym postanowiłem zainicjować badania w nowej dziedzinie, o której nikt jeszcze nie pomyślał. Trzeba pamiętać, że na uczelni MIT wokół mnie znajdowało się mnóstwo komputerów. Jasne było dla mnie, że wkrótce będą musiały się ze sobą komunikować. W tamtym czasie nie istniała skuteczna metoda, która by to umożliwiała, dlatego postanowiłem opracować technologię pozwalającą na tworzenie w efektywny sposób sieci danych.

Jaka była Pana pierwsza praca w branży komputerowej? Na czym polegała?

Od 1951 do 1957 r. uczęszczałem na wieczorowe zajęcia prowadzone na uczelni CCNY, których zwieńczeniem było uzyskanie przeze mnie wykształcenia w zakresie inżynierii elektrycznej. W ciągu dnia pracowałem najpierw jako technik, a później jako inżynier w niewielkiej firmie Photobell, zajmującej się elektroniką przemysłową. W tamtym czasie do linii produkcyjnej firmy wprowadziłem technologię cyfrową. Przede wszystkim korzystaliśmy z fotoelektrycznych urządzeń wykrywających obecność określonych obiektów (skrzynek, ludzi itp.). Układ nazywany *multiwibratorem bistabilnym* był tą technologią, której potrzebowaliśmy, aby w dziedzinie detekcji wykorzystać cyfrowe przetwarzanie. Tego typu układy okazały się elementami tworzącymi komputery. W obecnie używanej terminologii układy te nazywa się *przerzutnikami* lub *przełącznikami*.

Co Panu przyszło na myśl po przesłaniu pierwszego komunikatu między dwoma hostami zlokalizowanymi na uczelni UCLA i w instytucie Stanford Research Institute?

Szczerze mówiąc, nie zdawałem sobie sprawy ze znaczenia tego wydarzenia. Nie przygotowaliśmy specjalnego komunikatu o historycznej wadze, jak zrobiło to wielu wcześniejszych wynalazców (na przykład „What hath God wrought”[1] Samuela Morse, „Watson, come here! I want you”[2] Aleksandra Grahama Bella lub „That’s one small step for a man, one giant leap for mankind”[3] Neala Amstronga). Ci ludzie byli naprawdę *sprytni*! Rozumieli znaczenie mediów i public relations. My chcieliśmy tylko zalogować się do komputera instytutu SRI. Dlatego wpisaliśmy literę „L”, która została poprawnie zarejestrowana, i „o”, także przyjęte. Następnie wpisaliśmy „g”, co spowodowało awarię komputera w instytucie SRI! Okazało się więc, że nasz tekst był najkrótszym, a możliwe, iż także najbardziej proroczym komunikatem w historii, ponieważ brzmiał „Lo”, jak w zwrocie „Lo and behold!”[4].

Wcześniej w tym samym roku na konferencji prasowej odbywającej się na uczelni UCLA powtórzono moje słowa, w których stwierdziłem, że gdy sieć powstanie i zacznie funkcjonować, uzyskanie dostępu z naszych domów i biur do zasobów komputerowych będzie tak proste, jak w przypadku eksploatacji urządzeń elektrycznych i telefonicznych. W tamtym czasie moja wizja była taka, że internet stanie się wszechobecny, niedostrzegalny, zawsze aktywny i dostępny, a ponadto umożliwi podłączenie się do niego każdemu z dowolnego miejsca przy użyciu każdego urządzenia. Jednak nie przewidziałem, że moja 99-letnia mama będzie dzisiaj korzystać z internetu. Naprawdę to robi!

Jaka jest Pana wizja dotycząca przyszłości sieci?

Łatwym do przewidzenia aspektem wizji jest infrastruktura. Przewiduję szybki rozwój w obszarach przetwarzania mobilnego, urządzeń mobilnych i inteligentnych miejsc. Dostępność prostych, tanich, przenośnych i bardzo wydajnych urządzeń obliczeniowych oraz komunikacyjnych, a także wszechobecność internetu pozwoliły nam stać się nomadami. Mobilne technologie komputerowe umożliwiają użytkownikom przemieszczającym się z miejsca w miejsce korzystanie z usług internetowych w bezproblemowy sposób, niezależnie od tego, dokąd się udadzą i z jakich urządzeń korzystają. Trudniej jest przewidzieć rozwój aplikacji i usług — ich pojawianie się (poczty elektronicznej, technologii wyszukiwania, sieci WWW, blogów, sieci społecznościowych, technik generowania treści przez użytkowników, wymiany plików muzycznych, zdjęć, filmów wideo itd.) nieustannie nas zaskakuje. W bardzo niedalekiej przyszłości pojawi się nowa kategoria zaskakujących i innowacyjnych aplikacji mobilnych na urządzenia przenośne.

Kolejny pozwoli nam przenieść się ze świata cyberprzestrzeni do fizycznego świata inteligentnych miejsc. Nasze otoczenie (biurka, ściany, pojazdy, zegarki, pasy itp.) ożyje dzięki technologii reprezentowanej przez urządzenia uruchamiające, czujniki, układy logiczne, procesory, magazyny danych, kamery, mikrofony, głośniki, wyświetlacze i sprzęt komunikacyjny. Taka wbudowana technologia umożliwi otaczającemu nas środowisku zaoferowanie usług opartych na protokole IP, których zażądamy. Gdy wejdę do pokoju, ten będzie o tym „wiedział”. Będę w stanie w naturalny sposób komunikować się z otoczeniem, posługując się naszym językiem.

---

[1] Czyli „Co Bóg uczynił” — *przyp. tłum.*

[2] Czyli „Panie Watson, proszę przyjść! Potrzebuję pana” — *przyp. tłum.*

[3] Czyli „To jeden mały krok dla człowieka, ale wielki skok dla ludzkości” — *przyp. tłum.*

[4] Czyli „No i proszę!” — *przyp. tłum.*

Moje żądania spowodują wygenerowanie odpowiedzi przekazujących zawartość stron internetowych za pośrednictwem ściennych wyświetlaczach, szkieł kontaktowych, mowy, hologramów itp.

Sięgając w trochę dalszą przyszłość, wyobrażam sobie sieć uwzględniającą dodatkowe kluczowe komponenty. Widzę inteligentne agenty programowe zastosowane w sieci, których celem jest analizowanie i przetwarzanie danych, obserwowanie trendów, a także realizowanie zadań w sposób dynamiczny i adaptacyjny. Ponadto zdecydowanie większy ruch sieciowy będzie generowany nie przez ludzi, a przez wbudowane urządzenia i ich inteligentne agenty. Mogę sobie wyobrazić, że duży zbiór systemów, które same sobą zarządzają, będzie sprawować kontrolę nad taką rozległą i szybką siecią. Ogromna ilość informacji natychmiast będzie przesyłana w sieci. Będzie z tym związane intensywne przetwarzanie i filtrowanie. W zasadzie internet będzie rozprzestrzeniającym się globalnym układem nerwowym. Według mnie wszystko to i nie tylko nastanie już w XXI wieku.

Kto był dla Pana inspiracją podczas kariery zawodowej?

Przede wszystkim Claude Shannon pracujący na uczelni MIT. Był to znakomity naukowiec, który w bardzo intuicyjny sposób potrafił połączyć swoje matematyczne pomysły ze światem fizycznym. Claude wchodził w skład komisji związanej z moim doktoratem.

Czy ma Pan jakieś rady dla studentów zaczynających zajmować się sieciami komputerowymi i internetem?

Internet i wszystko, co pozwala z niego korzystać, jest nowym, rozległym obszarem pełnym niesamowitych wyzwań. Jest w nim miejsce na sporo innowacji. Nie należy być ograniczonym przez obecną technologię. Warto sięgać w przyszłość i wyobrażać sobie, co mogłoby w niej być, a następnie dążyć do tego, aby tak było.

Skorowidz

A

B

C

D

E

F

G

H

I

J

K

L

Ł

M

N

O

P

R

S

Ś

T

U

V

W

Y

Z

Ż

Aplikacje sieciowe, protokoły, internet — wszystko, co musisz wiedzieć.

Zagadnienia związane z sieciami komputerowymi są wyjątkowo złożone. Opanowanie tej tematyki wymaga przyswojenia wielu pojęć oraz zrozumienia licznych protokołów i technologii, które dodatkowo są ze sobą powiązane w zawiły sposób. Konieczne jest również uwzględnienie gwałtownego rozwoju technologii sieciowych i rosnącej złożoności nowych aplikacji. Aby poradzić sobie z tymi zagadnieniami, konieczne jest całościowe ujęcie tematyki sieci komputerowych.

Ta książka jest siódmym, zaktualizowanym i ulepszonym wydaniem znakomitego podręcznika. Zrozumienie zagadnień ułatwia wykorzystana przez autorów metoda omawiania zagadnień „od góry do dołu", od ogółu do szczegółu, a więc prezentują pierwszą warstwę aplikacji, nastepnie kolejne, niższe warstwy — aż do warstwy fizycznej. W książce szczególnie dużo miejsca poświęcono wiedzy o działaniu internetu, jego architekturze i protokołach. Zaprezentowano tu także fundamentalne zasady budowy i działania sieci oraz informacje o podstawowych problemach sieciowych i metodach ich rozwiązywania. W efekcie ten podręcznik pozwala na zdobycie gruntownej wiedzy, umożliwiającej zrozumienie niemal każdej technologii sieciowej.

W tej książce między innymi:

- warstwowość architektury sieciowej
- warstwa aplikacji, w tym strumieniowanie i sieci CDN
- działanie routerów i sterowanie logiką warstwy sieciowej
- bezpieczeństwo sieci
- administrowanie siecią

Dr Jim Kurose pracuje w US National Science Foundation. Wcześniej był redaktorem naczelnym „IEEE Transactions on Communications" i „IEEE/ACM Transactions on Networking". Jest członkiem IEEE i ACM. Zajmuje się protokołami, architekturą sieciową, pomiarami sieciowymi i komunikacją multimedialną.
Prof. Keith Ross jest dziekanem Instytutu Inżynierii i Informatyki uczelni NYU Shanghai. Wcześniej pracował na Uniwersytecie Pensylwanii. Zajmuje się zagadnieniami prywatności, sieci społecznościowych, sieci P2P, pomiarów sieciowych, sieci dystrybucji treści i modelowania stochastycznego.

Helion
helion.pl
HELION SA
ul. Kościuszki 1c
44-100 Gliwice
tel.: 32 230 98 63
helion@helion.pl
KOD KORZYŚCI
Sięgnij po więcej!
ISBN 978-83-8322-562-3
9 788383 225623
Cena: 149,00 zł

PEARSON ALWAYS LEARNING