Część I
Warstwa sieci — Sieci typu punkt-punkt

Trzecią warstwą w hierarchii protokołów jest warstwa sieci. Jest ona odpowiedzialna za sterowanie działaniem podsieci komunikacyjnej. Podstawowymi problemami projektowymi w tej warstwie są:

1. Interfejs między komputerem głównym i podsiecią.
2. Wybór drogi dla pakietów w podsieci.
3. Zabezpieczenie sieci przed przeciążeniem i zablokowaniem.

Część II
Połączenia wirtualne i datagramy

Aby uprościć oprogramowanie, sieci organizuje się w sposób hierarchiczny. W trakcie realizacji połączenia, komputer „nadawca" komunikuje się z węzłem, który z kolei komunikuje się z innymi węzłami i w końcu z komputerem odbiorczym. Złudzenie, że dowolne komputery mogą się ze sobą porozumiewać bezpośrednio, wywołane jest przez podsieć komunikacyjną.

Część III
Usługi świadczone przez warstwę sieciową

Jednym z przykładów funkcji warstwy sieci jest wykrywanie i korygowanie błędów. Każda podsieć obsługuje błędy łącza danych w sposób „przeźroczysty", mogą jednak wystąpić jeszcze inne błędy, głównie gubienie pakietów lub tworzenie duplikatów pakietów, spowodowane uszkodzeniem sprzętu lub oprogramowania węzła.

Drugim przykładem funkcji, którą ta warstwa sieciowa może wykonywać, jest zapewnienie napływu pakietów do komputera odbiorczego w takiej samej kolejności, w jakiej zostały wysłane. W praktyce stosuje się tu dwa modele: *model połączeń wirtualnych* oraz *model datagramowy*. Pierwszy z nich polega na tym, że warstwa sieciowa dostarcza pakiety do warstwy transportowej przez doskonały kanał (kanał, w którym nie ma błędów). Można tu przytoczyć przykład połączenia telefonicznego, w którym dwaj abonenci rozma-

Tabela 1: Różnice między modelem połączeń wirtualnych a modelem datagramowym

Zagadnienie	**Połączenie wirtualne**	**Datagram**
Adres przeznaczenia	Potrzebny tylko podczas ustalania	Potrzebny w każdym pakiecie
Obsługa błędów	Przeźroczyste dla komputera	Wyraźnie wykonywane przez komputer obliczeniowy
Sterowanie przepływem między punktami	Dostarczone przez podsieć	Nie dostarczone przez podsieć
Kolejność pakietu	Przesyłanie zawsze w kolejności wysyłania	Przesyłanie w kolejności przybycia
Początkowe ustalenia	Wymagane	Niemożliwe

wiają przez telefon. Najpierw musi zostać zestawione połączenie, następnie odbywa się przesyłanie danych (rozmowa) i w końcu połączenie zostaje przerwane poprzez odwieszenie słuchawki). W drugim modelu warstwa sieciowa akceptuje wiadomości nadchodzące z warstwy transportowej i próbuje wysłać każdą z nich jako oddzielną jednostkę. Wiadomości mogą wówczas przybywać nie w kolejności lub w ogóle nie przybywać. Model ten można porównać do systemu pocztowego. Każdy list przesyła się jako oddzielną jednostkę i dlatego musi on zawierać pełny adres przeznaczenia. Zestawienie głównych różnic między obwodem wirtualnym a usługą datagramową przedstawia tabela **??**.

Decyzja, czy warstwa sieciowa ma dostarczać usługi datagramowe, usługi połączenia wirtualnego, czy też obydwa rodzaje, zwykle zależy od typu oferowanego interfejsu warstwy trzeciej, między podsiecią i komputerami.

Z obu opisanych usług, usługa datagramowa jest prymitywniejsza. Jeżeli podsieć dostarcza usługę datagramową, to obsługa błędów i sterowania kolejnością przesyłania steruje warstwa transportowa. Połączenie wirtualne ma tę zaletę, że komputery obliczeniowe nie muszą się troszczyć o uporządkowanie kolejności napływu wiadomości, gdyż nadchodzą one kolejno.

Część IV
Wewnętrzna struktura podsieci

Omówione wyżej dwa modele połączeń mogą być użyte do wewnętrznego przesyłania. Przy połączeniu wirtualnym, droga wybierana po zestawieniu

połączenia prowadzi od węzła źródłowego do węzła przeznaczenia. Ta droga służy całemu ruchowi przechodzącemu przez połączenia wirtualne.

Jeżeli pakiety przechodzące przez dane połączenie wirtualne wybierają zawsze tę samą drogę w podsieci, to każdy węzeł musi pamiętać dokąd posłać pakiety dla poszczególnych otwartych połączeń wirtualnych, przechodzących przez dany węzeł. Każdy węzeł musi utrzymywać tablicę z jednym zapisem dla każdego otwartego połączenia wirtualnego. Każdy pakiet przechodzący przez podsieć musi mieć w nagłówku pole z numerem połączenia wirtualnego, oprócz numerów sekwencyjnych, sum kontrolnych itp. Kiedy pakiet przybywa do węzła, wówczas węzeł zna numer linii, z której pakiet przybył i numer połączenia wirtualnego. Na podstawie tej informacji pakiet musi być skierowany do właściwego węzła.

Gdy proces użytkowy powiadomi system operacyjny swojego komputera, że chce nawiązać kontakt z procesem w innej maszynie, wówczas lokalny komputer wybiera dla tej konwersacji nieużywany numer połączenia wirtualnego. Ponieważ każdy komputer wybiera numery połączeń wirtualnych niezależnie, to jakiś pośredni węzeł może użyć tych samych numerów połączeń wirtualnych dla dwóch różnych dróg, co prowadzi do nieporozumień.

Zatem połączenia wirtualne mogą być inicjowane z dowolnego końca, problem powstaje wówczas, kiedy zestawienia wywołań są przenoszone wzdłuż łańcucha węzłów w dwóch kierunkach jednocześnie. W tym samym momencie pakiety przybywają do sąsiednich węzłów. Każdy węzeł musi teraz wybrać numer połączenia wirtualnego, aby zastosować go do połączenia dupleksowego, które próbuje zestawić. Może to spowodować zestawienie dwóch nie związanych ze sobą połączeń wirtualnych na tej samej linii fizycznej, o tym samym numerze. Kiedy później przybędzie pakiet danych, odbiorczy węzeł nie ma możliwości stwierdzenia czy jest on kolejnym pakietem w jednym połączeniu, czy też pakietem zwrotnym w drugim połączeniu. Każdy proces powinien wskazać, kiedy kończy wykorzystywać połączenie wirtualne tak, aby można było usunąć z tablicy to połączenie wirtualne i zwolnić obszar pamięci.

Inną możliwością jest używanie połączeń datagramowych. Przy stosowaniu datagramów, węzły nie mają tablic z jednym zapisem dla każdego otwartego połączenia wirtualnego. Mają natomiast tablice wskazujące linie wyjściowe, których należy użyć dla wszystkich możliwych węzłów przeznaczenia. Te tablice są również potrzebne wtedy, kiedy połączenia wirtualne są używane wewnątrz, aby określić drogę dla pakietu zestawiania połączenia.

Każdy datagram musi zawierać pełny adres odbiorczy (numer maszyny i pewien wskaźnik określający proces, o który chodzi w tej maszynie). Kiedy pakiet przychodzi, wówczas węzeł szuka linii wyjściowej i wysyła go tą drogą. Nic w pakiecie nie jest modyfikowane. Również tworzenie i rozłączanie połączeń wirtualnych warstwy transportowej nie wymaga specjalnej pracy węzłów.

Część V

Porównanie połączeń wirtualnych i datagramów wewnątrz podsieci

Główna różnica między połączeniami wirtualnymi i datagramami wewnątrz podsieci polega na wielkości obszaru pamięci węzła i szerokości pasma. Przy połączeniach wirtualnych pakiety mogą zawierać numery połączeń zamiast pełnego adresu odbiorczego. Jeśli pakiety mają być dosyć krótkie, to pełny adres odbiorcy w każdym pakiecie wymaga znacznej wielkości nagłówka. Zastosowanie wewnętrznych połączeń wirtualnych dla podsieci staje się szczególnie atrakcyjne wówczas, kiedy wiele komputerów jest terminalami konwersacyjnymi, przesyłającymi tylko po kilka znaków w jednym pakiecie. Użycie połączeń wirtualnych pociąga za sobą jednak konieczność użycia tablic w węzłach. Ponadto, jeżeli węzeł ulega zniszczeniu i traci zawartość swojej pamięci, nawet jeśli zostanie odtworzony w sekundę później, to i tak wszystkie połączenia wirtualne przechodzące przez ten węzeł będą musiały być odtworzone. Wykorzystanie datagramów znacznie ogranicza powyższe niedogodności, gdyż w przypadku uszkodzenia węzła tylko ci użytkownicy, których pakiety znajdowały się w danej chwili, poniosą stratę. Dodatkowo przy datagramach jest możliwe wyrównanie przez węzły ruchu w podsieci, ponieważ istnieje możliwość zmiany drogi w połowie konwersacji.

Część VI

Algorytmy wyboru drogi

Algorytm wyboru drogi (routing algorithm) jest częścią trzeciej warstwy oprogramowania. Dobry algorytm wyboru drogi powinien cechować się:

- zapewnianiem poprawności transmisji,
- prostotą,
- odpornością na uszkodzenia,
- stabilnością,
- niezależnością,
- optymalnością.

Ponadto powinien on radzić sobie ze zmianami topologii sieci i natężenia ruchu.

Algorytmy wyboru drogi można podzielić na dwie duże klasy: *adaptacyjne* i *nieadaptacyjne*. Te pierwsze dokonują wyboru drogi na podstawie

pomierzonych lub szacunkowych wielkości bieżącego natężenia ruchu w sieci i jej topologii. Natomiast algorytmy nieadaptacyjne nie korzystają z takich informacji.

Część VII
Algorytm „zalewania" i jego odmiany

Jednym z najprostszych algorytmów wyboru drogi w sieciach komputerowych jest *algorytm „zalewania"* (flooding). Polega on na wysyłaniu pakietu do wszystkich linii poza tą, którą pakiet wszedł do węzła. Jest przy tym generowana ogromna liczba duplikatów pakietów, co wymusza stosowanie różnych metod likwidowania ich (np. użycie licznika kroków). Algorytm ten, mimo że nie jest często stosowany w praktyce, posiada pewne zalety, takie jak duża odporność na uszkodzenia. Poza tym, algorytm „zalewania" zawsze wybierze drogę najkrótszą.

Odmianą algorytmu „zalewania" jest *algorytm selektywnego zalewania* (selective flooding), który jest bardziej praktyczny. W tym algorytmie węzły nie wysyłają wszystkich wejściowych pakietów do wszystkich linii, lecz tylko do tych, które są zbliżone do prawidłowego kierunku.

Innym algorytmem nieadaptacyjnym jest *algorytm losowego wyboru drogi*. Węzeł wybiera losowo linię wyjściową i wysyła po niej pakiet. Jeżeli podsieć ma dobrze rozwiniętą strukturę powiązań, to algorytm ten może z powodzeniem służyć do wyboru dróg alternatywnych.

Część VIII
Statyczny wybór drogi

Statyczny lub bezpośredni wybór drogi (static or directory routing) jest jednym z najszerzej stosowanych algorytmów wyboru drogi. W tym algorytmie każdy węzeł wyposaża się w tablicę opisującą połączenia z innymi węzłami. Przed wysłaniem pakietu, węzeł generuje liczbę losową i według niej wybiera linię, używając wag jako prawdopodobieństw. Tabele dla węzłów muszą być wyznaczone ręcznie przez operatora sieci i raz załadowane do węzłów.

W algorytmie tym, nawet nie znając szczegółów topologii i ruchu w podsieci, można dojść do drogi optymalnej. *Zasada optymalności* mówi, że jeśli węzeł J znajduje się na drodze z węzła I do węzła K, to optymalna droga z J do K również leży na tej samej drodze. Z zasady optymalności wynika, że zbiór optymalnych dróg ze wszystkich źródeł do danego odbiorcy tworzy drzewo grafu, którego korzeniem jest węzeł u odbiorcy. Takie drzewo nazy-

wane jest *drzewem ujścia*.

Główną zaletą algorytmów nieadaptacyjnych jest prostota ich realizacji.

Część IX
Algorytmy scentralizowanego wyboru drogi

Najważniejszą wadą algorytmów nieadaptacyjnych jest fakt, że nie dostosowują się one do zmieniających się warunków w sieci. Tablice statycznego wyboru drogi projektowane są na podstawie średniego natężenia ruchu, jednakże dla nagłych skoków natężenia ruchu wartości średnie nie są użyteczne. W algorytmie scentralizowanego wyboru drogi najistotniejszą rolę odgrywa sposób tworzenia tablic. W pewnym miejscu w sieci znajduje się *centrum sterowania wyborem drogi* (RCC). Każdy węzeł przesyła co jakiś czas do RCC informację o swoim natężeniu ruchu. W RCC wszystkie informacje są zbierane, a następnie opracowywana jest nowa tablica połączeń optymalnych (podobnie jak w algorytmie statycznego wyboru drogi).

Algorytm ten umożliwia podejmowanie bardzo trafnych decyzji, ale posiada również bardzo istotne wady. Jeżeli podsieć ma adaptować się do zmian natężenia ruchu, to obliczenia należy wykonywać często. Powoduje to znaczne obciążenie sieci, wymuszone koniecznością stałego odświerzania tablic. Inną wadą jest wrażliwość na nieprawidłowe działanie RCC lub na odzielenie RCC od części sieci (np. z powodu awarii).

Część X
Izolowany wybór drogi

Istotą *wyizolowanych adaptacyjnych algorytmów wyboru drogi* (isolated adaptive) jest fakt, że węzły próbują dostosować się do zmian zachodzących w topologii i obciążeniu sieci, nie wymieniając przy tym informacji z innymi węzłami.

Przykładem wyizolowanego algorytmu może być np. algorytm Barana, zwany również *algorytmem gorącego ziemniaka*. Polega on na tym, że gdy do węzła przybędzie pakiet, to węzeł umieszcza go w najkrótszej kolejce wyjściowej, aby pozbyć się go w najkrótszym czasie.

Innym izolowanym algorytmem wyboru drogi jest tzw. *uczenie zwrotne* (backward learning). Polega ono na dołączaniu do każdego pakietu identyfikatora węzła źródłowego i licznika etapów przebytej drogi. Po pewnym czasie, każdy węzeł będzie znał najkrótszą drogę do każdego innego węzła. Wadą te-

go podejścia jest niemożność wykrycia uszkodzenia lub przeładowania, gdyż węzły rejestrują tylko zmiany na lepsze. Aby rozwiązać ten problem, co jakiś czas usuwa się dane przechowywane w węzłach.

Algorytm wyznaczania drogi *metodą delta* stanowi połączenie algorytmu scentralizowanego i izolowanego. Każdy węzeł mierzy „koszt" każdej linii i okresowo przesyła pakiet zawierający te wartości do RCC. RCC oblicza następnie k najlepszych dróg dla wszystkich węzłów i rozsyła listy dróg do wszystkich węzłów, przy czym potrzebne są tylko początkowe linie, a nie całe drogi.

Część XI
Algorytm rozproszonego wyboru dróg

Algorytm rozproszonego wyboru dróg polega najogólniej na tym, że każdy węzeł wymienia okresowo dokładną informację o wyborze drogi ze wszystkimi swoimi sąsiadami. W węźle przetrzymuje się tablicę wyboru drogi zawierającą jeden zapis dla każdego innego węzła w podsieci, składający się z zalecanej linii wyjściowej i pewnej szacunkowej wartości czasu lub odległości do tego odbiorcy.

Część XII
Aktualizacja topologii sieci

Jedną z metod przyśpieszających reakcję sieci na informacje o uszkodzeniu węzła lub linii jest metoda *utrzymaj się* (hold down). Polega ona na tym, że w momencie, gdy najlepsza dotychczas droga nagle ulegnie pogorszeniu, to zapis w tablicy wyboru drogi jest ustawiany na bieżącą wartość odebraną z linii, która do tej chwili była najlepszą drogą, nawet gdy istnieje inna droga o krótszym opóźnieniu. Jeżeli węzeł lub linia ulegną całkowitemu uszkodzeniu, to węzeł wykrywający błędne działanie powinien wygenerować specjalny pakiet wyboru drogi, zawierający czas nieskończenie długi dla wszystkich zapisów w tablicach dotyczących tej linii.

Część XIII
Hierarchiczny wybór drogi

W miarę rozrastania się sieci, rosną proporcjonalnie tablice wyboru drogi. Powoduje to zwiększanie się pamięci zajmowanej przez tablicę w węźle oraz wzrost czasu potrzebny na ich przeglądanie. Aby uniknąć nadmiernego rozrostu węzłów sieć dzieli się na regiony. Każdy węzeł zawiera informację o wyborze dróg wewnątrz własnego regionu. W pewnych przypadkach konieczne jest wprowadzenie hierarchii wielopoziomowej. Pakiety grupuje się w regiony, regiony w pęki, pęki w strefy itd. Korzyści wynikające z ograniczenia obszaru „pamiętanego" przez węzły pociągają niestety za sobą pewne wady, np. wzrost długości połączenia pomiędzy regionami. Gdy pojedyncza sieć jest za duża, należy podjąć decyzję o ilości poziomów hierarchii, na które podzieli się sieć. Optymalną liczbą poziomów dla podsieci z N węzłami jest $\ln N$. W węźle potrzebnych jest zatem $e \ln N$ zapisów. Wzrost efektywnej średniej długości drogi powodowany przez hierarchiczny wybór drogi jest pomijalnie mały.

Część XIV
Rozsiewczy wybór drogi

W przypadku, gdy zachodzi potrzeba rozesłania wiadomości do wszystkich, np. aby rozesłać informacje potrzebne do aktualizacji tablic wyboru drogi, stosuje się tzw. algorytmy *rozsiewania* lub *rozgłaszania* (broadcasting).

Jedną z metod rozsiewania, jest metoda polegająca na wysłaniu z węzła źródłowego oddzielnego pakietu do każdego odbiorcy. Nie jest to metoda oszczędna, a do jej realizacji wymaga się posiadania pełnej listy wszystkich odbiorców.

Zalewanie jest jedną z metod rozsiewania. Polega ono na wysłaniu pakietów do wszystkich linii, oprócz tej, którą pakiet wszedł do węzła.

Innym algorytmem rozsiewania są tzw. *wieloadresowe wybory drogi* (multidestination routing). W metodzie tej każdy pakiet zawiera listę odbiorców lub mapę bitową wskazującą żądanych odbiorców. Kiedy pakiet przybędzie do węzła, wówczas dokonywane jest sprawdzenie wszystkich odbiorców, w celu ustalenia zbioru linii wyjściowych. Następnie węzeł generuje nową kopię pakietu dla tylko tych odbiorców, którzy będą nadal używali tej linii. Po odpowiedniej liczbie kroków każdy pakiet będzie przenosił informacje tylko dla jednego odbiorcy i może być traktowany jako normalny pakiet.

Czwartym algorytmem rozsiewczym jest algorytm, który do zapoczątkowania rozsiewania używa drzewa ujścia lub innego dogodnego drzewa roz-

pinającego. Jeżeli każdy węzeł ma informacje, które z jego linii należą do drzewa rozpinającego, to może kopiować wchodzący pakiet rozsiewany do wszystkich linii drzewa rozpinającego oprócz linii, z której ten pakiet przybył. Aby zatem stosować tę metodę, każdy węzeł musi mieć informacje o pewnym drzewie rozpinającym, z którego będzie korzystał. Istnieje jednak sposób ominięcia tej niedogodności, umożliwiający pracę algorytmu w sytuacji, gdy węzły nie mają żadnych informacji o drzewach rozpinających. Gdy pakiet rozsiewczy przybędzie do węzła, wtedy węzeł sprawdza, czy pakiet przybył z tej linii, która normalnie służy wysyłaniu pakietów do źródła rozsiewania. Jeżeli tak, to węzeł przesyła kopie pakietu do wszystkich linii oprócz tej, z której pakiet przybył. W przeciwnym przypadku pakiet jest traktowany jak duplikat i jest odrzucany.

Część XV
Przeciążenie

Zjawisko *przeciążenia* (congestion) występuje przy zbyt dużej liczbie pakietów w podsieci lub w tej części obniża się wydajność przesyłania. Dopóki liczba pakietów w podsieci nie przekracza jej pojemności przesyłowej, to wszystkie pakiety (pozbawione błędów transmisji) trafiają do miejsca przeznaczenia. W przypadku jednak, gdy liczba pakietów zbytnio wzrośnie, wówczas węzły zaczynają „gubić" pakiety. Spowodowane to jest kilkoma czynnikami. Węzły mogą być zbyt wolne albo szybkość ruchu wejściowego przekroczy przepustowość linii, co powoduje wydłużenie kolejek. Rozwiązanie powyższego problemu sprowadza się w zasadzie do zastosowania dostatecznej liczby buforów.

Część XVI
Wstępne rozmieszczenie buforów

Jeżeli w sieci istnieją połączenia wirtualne, to problem przeciążenia można całkowicie rozwiązać w następujący sposób: przy zestawieniu połączenia wirtualnego, pakiet żądania wywołania, jest transmitowany przez podsieć tworząc zapisy w tablicach węzłów, przez które „przechodzi". Po przybyciu pakietu do odbiorcy zostaje określona droga dla całego dalszego ruchu danych i są już zapisy w tablicach wyboru drogi wszystkich węzłów pośrednich.

W przypadku stałego przydziału buforów dla wszystkich połączeń wirtualnych, w każdym węźle zapewnia się miejsce do zapamiętania wchodzącego pakietu do czasu, kiedy będzie go można wysłać.

Część XVII
Odrzucanie pakietów

Drugi mechanizm sterowania przeciążeniem jest właściwie przeciwieństwem pierwszego. Zamiast rezerwowania zawczasu wszystkich buforów, wcześniej nie rezerwuje się nic. Jeżeli nadejdzie pakiet i nie ma dla niego miejsca, to węzeł po prostu go odrzuca. W przypadku połączeń datagramowych unika się w ten sposób przeciążenia przez dowolne odrzucanie pakietów. Dla połączeń wirtualnych kopia każdego pakietu musi być gdzieś trzymana tak, aby w przypadku błędów transmisji mogła być retransmitowana później.

Odrzucanie pakietów, pomimo prostoty realizacji, ma jednak pewne wady. Najważniejszą z nich jest potrzeba posiadania dodatkowej szerokości pasma przenoszenia dla duplikatów. Jeżeli dodatkowo przedział czasu jest zbyt krótki, to duplikaty pakietów będą generowane wówczas, kiedy nie są potrzebne, co spowoduje dodatkowe zwiększanie obciążenia sieci. Przy zbyt długim przedziale czasu opóźnienie przesyłania pakietów będzie się zwiększało.

Część XVIII
Sterowanie przeciążeniem izorytmicznym

Jeżeli przeciążenie powstaje w przypadku, gdy w sieci znajduje się zbyt wiele pakietów, to istnieje możliwość sterowania przeciążeniem przez ograniczanie liczby pakietów. Metoda *izorytmiczna* (isarithmic) polega na utrzymaniu pewnej stałej liczby pakietów. Istnieją zezwolenia na transmisję, które krążą w podsieci. W momencie, gdy węzeł chce wysłać pakiet właśnie przekazany, musi poczekać na zezwolenie i je zniszczyć. Gdy pakiet trafi do odbiorcy, to węzeł odbiorczy usuwa pakiet i generuje zezwolenie na przesyłanie. Przy zastosowaniu tej metody, liczba pakietów w sieci nigdy nie przekroczy liczby zezewoleń początkowo istniejących w sieci.

Algorytm izorytmiczny nie gwarantuje jednak, że znajdujące się w sieci pakiety nie trafią nagle w tym samym czasie do jednego węzła i nie obciążą go nadmiernie. Innym problemem związanym z tą metodą jest problem rozdziału zezwoleń na przesyłanie. Aby uchronić nowo wygenerowane pakiety przed zbytnim opóźnieniem, spowodowanym oczekiwaniem na zezwolenie, należy zezwolenia przydzielać jednolicie. Poza tym przypadkowa wędrówka dodatkowych zezwoleń powoduje obciążenie sieci, a zezwolenia są podatne na uszkodzenia. W warstwie sieci możliwe jest też stosowanie protokołów bezkolizyjnych, z ograniczoną rywalizacją o dostęp do kanału i całego szeregu innych protokółów, dotyczących alokacji kanału.