**Przełączniki**

Przełącznik jest urządzeniem aktywnym, które łączy dwa segmenty sieci w ramach jednej warstwy lub większej liczby warstw modelu OSI. Termin przełącznik (ang. switch) jest stosowany w odniesieniu do różnorodnych urządzeń i w przeciwieństwie do mostu (ang. bridge), opisanego w standardzie IEEE 802.ID, nie jest zdefiniowany w żadnym zaleceniu. W rzeczywistości przełączniki funkcjonujące w warstwie 2. modelu OSI są zgodnie ze standardem IEEE 802.ID mostami, mimo że większość producentów sprzedaje je jako przełączniki. Przełączniki mają zdolność do wyznaczania wirtualnych obwodów dla przenoszonych przez nie ramek, ale często nie są wyposażane w funkcje dynamicznej rekonfiguracji tych obwodów (wymagają interwencji z zewnątrz). Zdolność do dynamicznej rekonfiguracji obwodów polega na możliwości przekazania ruchu z jednego portu na inny w zależności od bieżącej konfiguracji sieci lub w wyniku działania algorytmu optymalizacji.

Przełączniki mogą być urządzeniami zarządzalnymi lub niezarządzalnymi. Komponentów niezarządzalnych nie można konfigurować za pośrednictwem sieci. Natomiast urządzenia zarządzalne taką konfigurację umożliwiają. W przełącznikach zarządzalnych jest zazwyczaj wbudowany moduł agenta SNMP, a także interfejs wiersza poleceń (ang. CLI— Command Line Interface) lub interfejs WWW. Wśród przełączników wyróżnia się te, które udostępniają jedynie niewielki zbiór parametrów konfiguracyjnych, oraz urządzenia klasy enterprise (w pełni zarządzalne), które umożliwiają tworzenie i przechowywanie wielu konfiguracji. Przełączniki korporacyjne (klasy enterprise) zwykle mają większą liczbę portów i pozwalają na budowanie stosów zarządzanych przez administratora jak jedno urządzenie.

Analizując przełącznik, warto zwrócić uwagę na następujące cechy:

* **Porty** — liczba portów, możliwość zmiany priorytetów portów oraz funkcja monitorowania (kopiowania na wskazany port) ruchu (ang. port mirroring).
* **Szybkość** — szybkość portów oraz tryb dupleksu wpływają na przepustowość przełącznika.
* **Agregacja połączeń** — zdolność do tworzenia wirtualnych portów (połączeń) składających się z kilku fizycznych portów (połączeń) w celu zwiększenia przepływności i niezawodności.
* **SNMP** — możliwość zarządzania urządzeniami za pośrednictwem sieci.
* **Filtrowanie** — możliwość klasyfikowania ruchu na podstawie cech urządzeń (na przykład filtrowanie na podstawie adresów MAC). Translacja adresów sieciowych (NAT) jest funkcją fire walla lub routera i zazwyczaj nie jest implementowana w przełącznikach, choć są odstępstwa od tej reguły.
* **Kontrola dostępu do sieci** — zdolność przełącznika do przyznawania urządzeniom dostępu do sieci w określonym zakresie na podstawie zdefiniowanych reguł.
* **VLAN** — zdolność do wydzielania wielu sieci logicznych w ramach większej sieci fizycznej, dzięki czemu uzyskujemy separację ruchu pomiędzy sieciami logicznymi.

Na rynku są dostępne przełączniki o funkcjach właściwych dla wszystkich warstw modelu osi, od warstwy łącza danych do warstwy aplikacji. Jedynie urządzenia warstwy fizycznej — regeneratory i koncentratory — nie są nazywane przełącznikami.

W sieciach Ethernet wszystkie porty koncentratora odbierają te same dane, które w tym przypadku mają charakter rozgłoszeniowy — koncentratory nie wprowadzają segmentacji, co oznacza, że wszystkie połączenia należą do tej samej domeny kolizyjnej. Aby ograniczyć liczbę kolizji, koncentratory pracują w trybie półdupleksowym, współdzieląc połączenie. Przełączniki segmentują ruch w taki sposób, że każdy segment sieciowy dysponuje własną szerokością pasma, działa w ramach własnej domeny kolizyjnej (w której nie wy-stępują kolizje) i w trybie pełnego dupleksu.

Prawdopodobnie najwłaściwszym opisem przełącznika jest przedstawienie jego funkcji w odniesieniu do poszczególnych warstw modelu OSI. W warstwie 2. przełącznik w zasadzie w pełni odpowiada zapisom standardu IEEE 802.ID definiującego most. Funkcje przełącznika realizowane na poziomie warstwy łącza danych zostały szczegółowo opisane w dalszej części rozdziału. Gdy przełącznik wykonuje zadania właściwe dla warstwy sieciowej modelu OSI, w rzeczywistości działa jak router — funkcja routingu również została opisana w dalszej części rozdziału. Wieloportowe urządzenia przełączające o dużym zagęszczeniu portów, nazywane przełącznikami typu director, są urządzeniami warstwy 3., wykorzystywanymi często w sieciach PSTN lub sieciach pamięci masowej bazujących na połączeniach Fibre Channel do łączenia setek urządzeń. Często ich działanie nie podlega prostym regułom kategoryzacji i należy je zaliczyć do usług dwóch lub nawet większej liczby warstw. Przełączniki tego typu są powszechnie nazywane przełącznikami wielowarstwowymi.

Istnieją również przełączniki warstwy 4. i warstwy 7. Przełącznik warstwy 4. odpowiada za translację adresów (NAT — Network Address Translation) oraz za równoważenie obciążenia między portami. Przełączniki warstwy 4. mogą również obejmować firewalle, bramy IPSec i koncentratory VPN. Najczęściej są one sprzedawane właśnie pod nazwą „firewall”, aby zwiększyć ich znaczenie na rynku. Przełączniki warstwy 7. udostępniają usługi warstwy aplikacji i pełnią zazwyczaj funkcję serwerów dostarczania treści lub urządzeń buforowania danych internetowych. Bardzo rzadko spotyka się przełączniki warstwy 7., nazywane rzeczywiście przełącznikami. Przeważnie określa się je mianem serwerów, ponieważ termin ten jest korzystniejszy z marketingowego punktu widzenia.

**Mosty**

Most jest urządzeniem sieciowym, które łączy dwa segmenty sieci (jedną podsieć) na poziomie warstwy łącza danych. Mosty analizują docelowe adresy MAC przekazywanych ramek, ale nie przetwarzają w żaden sposób protokołów warstwy sieciowej, takich jak IP, IPX, NetBEUI lub inne. Mostów można używać także do łączenia sieci o różnych rodzajach medium fizycznego, np. 100Base-T i Wi-Fi lub 100Base-T i 100Base-TX.

W sieci Ethernet most jest przezroczystym, adaptacyjnym urządzeniem sieciowym, które odbiera dane z podłączonych do niego segmentów sieci. Na tej podstawie buduje tablicę mostowania (ang. Forwarding Table), w której przechowuje adresy MAC urządzeń i odpowiadający im segment sieci (port mostu). Porównując adresy MAC przychodzących ramek z tablicą mostowania, podejmuje decyzje o przełączeniu ramki do innego segmentu. Gdy w tabeli mostowania nie ma odpowiedniego wpisu, ramka zostaje wysłana na wszystkie porty z wyjątkiem portu źródłowego. Dopiero przetwarzając odpowiedź na tę ramkę, most rejestruje adres MAC zawarty w odpowiedzi i kojarzy go z określonym segmentem sieci. Adaptacyjne przełączanie jest operacją, która może być realizowana w jednym z trzech trybów:

* **Store and forward**. Technika ta polega na buforowaniu nadchodzących ramek, sprawdzaniu wartości kontrolnych, a następnie przekazywaniu ich do urządzeń docelowych.
* **Cut through.** Działanie tego mechanizmu polega na odczytaniu początkowej części ramki do pola adresu MAC, odszukaniu portu wyjściowego w tablicy przełączania i rozpoczęciu wysyłania przez port wyjściowy przed odebraniem całej ramki na porcie wejściowym. W rozwiązaniu tym nie występuje etap sprawdzania błędów.
* **Fragment free.** Przed rozpoczęciem przekazywania danych odczytywane są 64 bajty ramki. Zakłada się, że sprawdzenie początkowej części ramki wraz z adresem niemal zawsze pozwala wywnioskować, czy dane są nienaruszone lub czy wystąpiła kolizja, która sprawiła, że są one bezużyteczne.

Mosty w sieciach Token Ring wykorzystują inną metodę określania sposobu przekazywania ruchu, nazywaną mostkowaniem na podstawie trasy źródłowej (SRB — source route bridging). Przykładowy most przełącza ramki w odpowiednim kierunku na podstawie informacji, które zostały zawarte w tych ramkach przez urządzenie nadające. Urządzenie nadające przed wysłaniem właściwych pakietów wysyła ramkę, której zadaniem jest zebranie informacji o sieci. Każdy most, który otrzymuje taką ramkę, rozsyła je na wszystkie swoje porty (oprócz źródłowego), jednocześnie dopisując informacje trasowania. Urządzenie docelowe odpowiada na każdą ramkę, używając do tego pełnej informacji trasowania. Urządzenie nadające wybiera trasę na podstawie różnych parametrów — najczęściej jest to trasa z ramki, która najszybciej wróciła.

Mosty znajdują zastosowanie przede wszystkim w konfiguracjach, w których wydzielone są dwie grupy komputerów komunikujących się intensywnie w ramach własnych grup i sporadycznie między grupami. Przykładem może tutaj być sieć obejmująca dwa piętra budynku, na których pracują zespoły finansowy i inżynierski. Innym przykładem wykorzystania może być wydzielenie klastrów systemów Linux i Macintosh. W każdej z opisanych sytuacji mosty zwiększają wydajność sieci przez partycjonowanie większości ruchu — ograniczenie ruchu do obszaru połowy sieci — i zmniejszenie liczby kolizji.

Jedyne urządzenia, które można kupić jako mosty, to punkty dostępu bezprzewodowego skonfigurowane do łączenia dwóch sieci lub dwóch segmentów sieci. W rozwiązaniach przewodowych działanie w charakterze mostu jest trybem pracy przełącznika. Dlatego czasami termin „przełącznik sieciowy” jest używany jako synonim mostu sieciowego. W większości przypadków pojęcie mostu lub mostu sieciowego odnosi się do sieci Ethernet, gdyż opisuje ono każde urządzenie zgodne ze standardem IEEE 802.ID. Opisany również w tym zaleceniu protokół drzewa rozpinającego (STP — Spanning Tree Protocol) został omówiony w dalszej części rozdziału (jest to protokół warstwy 2. zapobiegający powstawaniu pętli).

Oto kilka charakterystycznych cech mostu sieciowego:

* Most nie przetwarza żadnych protokołów sieciowych na poziomie wyższym niż wyznaczony przez protokół odwzorowania adresów (ARP).
* Most wydziela dwie domeny kolizyjne, przetwarzając i regenerując pakiety.
* Niezależnie od liczby dostępnych portów most dysponuje jednym portem, za którego pomocą są przekazywane informacje, oraz drugim, służącym do dystrybuowania informacji. Zatem od strony sieci most ma tylko jeden interfejs sieciowy.
* Most nie wykonuje operacji routingu, ale może filtrować ramki na podstawie ich adresów docelowych.
* W sieci może pracować dowolna liczba mostów, a ograniczenia nakładane na segment sieci nie odnoszą się do segmentu znajdującego się po drugiej stronie mostu.
* Dany port przynależy logicznie tylko do jednego mostu.
* Dodany do mostu port staje się portem niezarządzalnym, ponieważ mosty sieciowe konfigurują się we własnym zakresie.

Mostem sieciowym lub niezarządzalnym przełącznikiem jest urządzenie, które nie dysponuje własnym adresem IP i które nie odpowiada na polecenia sieciowe takie jak ping. Funkcja przenoszenia datagramów między segmentami sieci przyłączonymi do mostu nie wymaga od samego mostu, by był on zarządzalny. Niemniej wiele urządzeń, na przykład przełączniki realizujące zadania logicznych mostów, umożliwia zdalne zarządzanie, dysponuje adresami IP, bierze udział w komunikacji SNMP i jest dostępnych za pośrednictwem usług SSH, TELNET lub RLOGIN. Za pomocą wymienionych mechanizmów można połączyć się z zarządzalnym przełącznikiem, aby przypisać adres IP do wirtualnego interfejsu, który z kolei pozwala na komunikację z innymi interfejsami sieciowymi. Ruch pochodzący z innych punktów końcowych sieci jest przenoszony przez zarządzalny most bez jakichkolwiek modyfikacji.

Osoby, które posługują się systemem Windows XP lub Vista, mogą skonfigurować programowy most sieciowy w ramach systemu operacyjnego. Most sieciowy w systemie Windows jest interfejsem wirtualnym, który obejmuje dwie lub większą liczbę sieci. Jeśli komputer jest wyposażony na przykład w kartę sieci stałej oraz kartę bezprzewodową, można wykorzystać funkcję mostkowania do udostępnienia komputerom działającym w każdej z sieci zasobów wydzielonych na komputerze-moście. Ponadto most sieciowy zapewnia systemom z jednej sieci dostęp do zasobów systemów znajdujących się po jego drugiej stronie.

Nie należy tworzyć mostów między połączeniem internetowym a siecią wewnętrzną, gdyż umożliwia to internautom uzyskanie dostępu do zasobów sieci wewnętrznej.

Aby utworzyć most sieciowy w systemie Windows, należy wykonać przedstawioną poniżej procedurę:

1. Otwórz folder Połączenia sieciowe (Panel sterowania/ Wyświetl stan sieci

i zadania/Zarządzaj połączeniami sieciowymi).

2. Trzymając naciśnięty klawisz Ctrl, kliknij połączenia sieciowe (interfejsy), z których chcesz utworzyć most.

3. Kliknij prawym przyciskiem myszy zaznaczony interfejs i wybierz opcję Połączenia mostkowe. Jeśli to konieczne, wpisz hasło administratora. Na rysunku 9.2 zostało pokazane okno Połączenia sieciowe w systemie Vista po utworzeniu mostu sieciowego.

Obraz zawierający tekst

Opis wygenerowany automatycznie

Most sieciowy jest w tym przypadku wirtualnym interfejsem sieciowym, którym można operować tak samo jak każdym innym interfejsem systemu. Można na przykład wyświetlić odpowiadające mu okno Właściwości, dodać lub usunąć komponenty sieciowe, a nawet dodać kolejne interfejsy. Aby usunąć interfejs z mostu sieciowego lub cały most sieciowy, wystarczy wybrać opcję usunięcia z menu kontekstowego, wyświetlanego po kliknięciu ikony połączenia prawym przyciskiem myszy.

Mostkowanie i routing są dwiema metodami wyboru tras dla danych przekazywanych przez sieć. Jednak routing odnosi się do zadań realizowanych na poziomie warstwy sieciowej. Router kieruje ruch sieciowy do kolejnych węzłów na podstawie adresu logicznego, takiego jak adres IP, podczas gdy most posługuje się jedynie identyfikatorem sprzętowym (adresem MAC). Z tego względu routery dysponują informacją o tym, kiedy komunikacja obejmuje różne sieci, a mosty takich danych nie mają. Jako ogólną zasadę można przyjąć, że mosty są stosowane do łączenia segmentów sieci, natomiast routery pozwalają na łączenie różnych sieci. Mosty są tanimi urządzeniami — droższymi od koncentratorów i regeneratorów, lecz mniej kosztownymi niż przełączniki (zazwyczaj) i routery (zawsze). Ponieważ mosty buforują ramki w celu wyznaczenia ich dalszej trasy, cechują się mniejszą przepustowością niż regeneratory, których działanie ogranicza się do odtworzenia sygnału.

**Routery**

Router sieciowy jest urządzeniem, które łączy dwie różne sieci. Routery wydzielają domeny' kolizyjne, filtrują i blokują ruch rozgłoszeniowy oraz wyznaczają optymalne trasy dla pakietów. Wysokowydajne routery są po prostu efektywnymi komputerami, zdolnymi do przetwarzania dużej ilości danych w krótkim czasie.

Routery jako urządzenia logiczne pierwotnie były uruchamiane w formie serwerów o wielu interfejsach sieciowych. Jednym z pierwszych urządzeń tego typu był produkt firmy BBN Technologies (wcześniej Bolt, Beranek and Newman), zastąpiony ostatecznie przez systemy DEC PDP-11, przeznaczone do trasowania ruchu IP. W latach 80., gdy zaczęto komercjalizować internet, firma Sun Microsystems spopularyzowała tanie serwery SPARC, które wielu

dostawców usług internetowych kupowało z przeznaczeniem do wykonywania zadań routingu. W tym czasie rozpoczęła także swoją działalność firma Cisco, która przekształciła routery w niezależne urządzenia i zdominowała rynek ich producentów.

Funkcje routingu są implementowane w wielu sieciowych systemach operacyjnych, włączając w to systemy UNIX, Linux i Windows Server. Niewielki koszt systemu Linux sprawia, że bardzo często występuje on w roli routera. Również urządzenia firmy Cisco korzystają z systemu operacyjnego (IOS), lecz jest to oprogramowanie specjalnie przygotowane do wykonywania zadań związanych z routingiem i przełączaniem, udostępniające administratorom interfejs wiersza poleceń. Wiele rozwiązań opisanych w tym rozdziale jest wynikiem prac firmy Cisco. Inni dostawcy urządzeń stosują własne systemy operacyjne — w przypadku Juniper Networks jest to system (JUNOS), a firma Extreme Networks korzysta z systemu XOS.

*Jeśli podstawowym przeznaczeniem routera miałaby być segmentacja ruchu (bez wykorzystania funkcji routingu lub translacji protokołów), warto się zastanowić nad wykorzystaniem mostu zamiast niego.*

Wyznaczanie tras dla pakietów w małej sieci nie jest zadaniem szczególnie obciążającym procesor, więc wiele osób przekształca przestarzałe komputery w routery sieciowe. Oprogramowanie, które zapewnia realizację zadań routingu z użyciem komputera PC, to:

* **Quagga** (www.quagga.net) — usługa routingu przeznaczona dla systemów UNIX, Linux i Solaris, która bazując na projekcie Zebra, umożliwia korzystanie z protokołów OSPF, RIP i BGP (projekt typu open source).
* **SmoothWall** (www.smoothwall.org) — dystrybucja systemu Linux udostępniająca łatwy w użyciu graficzny interfejs użytkownika (projekt typu open source).
* **Untangle** (www.untangle.com) — aplikacja bramy, która tworzy router brzegowy wyposażony w funkcje antywirusowe, mechanizmy zapobiegające działaniu programów szpiegowskich, filtry, blokady oraz firewall (projekt typu open source).
* **XORP** (www.xorp.org) — otwarta rozszerzalna platforma routingu (extensible Open Router Platform) obejmująca obsługę protokołów RIP, OSPF, IGMP, BGP i innych. Aplikacja XORP jest dostępna w wersjach dla systemów Linux, Mac OS X (9.2 i późniejszych) oraz Windows Server 2003 (projekt typu open source).

Osoby interesujące się sieciami z pewnością słyszały określenie brouter. Brouter jest urządzeniem stanowiącym połączenie mostu (ang. bridge) i routera. Gdy do broutera zostaną dostarczone pakiety protokołu routowalnego (takiego jak TCP/IP), będą przetworzone zgodnie z zasadami routingu (zostaną przeniesione z sieci źródłowej do docelowej). Natomiast wszystkie pakiety, które są przenoszone w ramach nieroutowalnego protokołu (takiego jak NetBEUI), są przekazywane zgodnie z zasadami działania mostu.

Sposób działania routerów jest charakteryzowany przez dwa oddzielne systemy funkcjonalne — warstwę sterującą i warstwę przełączania, które wybierają porty i przekazują dane do odpowiedniego interfejsu wyjściowego. Metodologia postępowania jest opisana przez wyrafinowane algorytmy odpowiedzialne za optymalizację wydajności sieci. W zależności od obsługiwanych przez router protokołów administrator może utworzyć różne topologie trasowania. Wiele aspektów funkcjonowania routerów zostało opisanych w punktach tego rozdziału.

**Warstwa sterująca**

Routery składają się z dwóch płaszczyzn operacyjnych — warstwy sterującej, która wskazuje port wyjściowy uczestniczący w przekazywaniu pakietów do sieci docelowej, oraz warstwy przełączania, która przenosi odebrane pakiety z interfejsu wejściowego do wyjściowego. Warstwa sterująca uczestniczy w wymianie danych z innymi urządzeniami sieciowymi, mającej na celu zbudowanie tablicy routingu. Jest także odpowiedzialna za filtrowanie i blokowanie ruchu oraz za ewentualne zaimplementowane przez producenta mechanizmy QoS.

Warstwa sterująca przechowuje tablicę routingu, której podstawowe zadanie polega na gromadzeniu adresów wykorzystywanych w bezpośredniej komunikacji między dwoma punktami końcowymi. Administrator ma możliwość osobistego zdefiniowania tras statycznych lub określenia zasad wykorzystania takich tras. Niektóre wpisy w tablicy routingu mogą się odnosić do logicznych grup systemów, które są wykorzystywane w operacjach multiemisji. Większość routerów, wykonując zadania routingu, polega na tablicach routingu lub bazach danych informacji o trasach (RIB —Routing Information Base). Niemniej niektóre z routerów utrzymują również bazy danych przełączania (FIB -— Forwarding Information Base), które są zapisywane przez warstwę sterującą w pamięci w celu ich wykorzystania z poziomu warstwy przełączania.

Administratorzy sieci zwykle decydują się na włączanie routingu dynamicznego, który umożliwia danemu routerowi wymianę informacji na temat sieci z innymi routerami, a w konsekwencji na samodzielne wyznaczanie najlepszych tras przez sieć. Większość protokołów routingu dynamicznego pozwala na przypisanie routerowi priorytetu, który jest najważniejszym czynnikiem definiującym rolę routera oraz określającym to, w których trasach zostanie on uwzględniony.

Podczas wyznaczania tras przez sieć wykorzystywane są fizyczne połączenia routerów. Jednak nic nie stoi na przeszkodzie, aby utworzyć w routerze również logiczny interfejs sieciowy. Routery umożliwiają przypisywanie dowolnej liczby interfejsów logicznych do interfejsu fizycznego. Dzięki temu mogą obsługiwać wirtualne sieci LAN (VLAN). Obsługa sieci VLAN jest regulowana standardem IEEE 802.1Q. Część routerów współdziała także z protokołami tunelowania ruchu, takimi jak GRE i MPLS. Tunelowanie ruchu jest szczegółowo opisane w rozdziale 29.

**Warstwa przełączania**

Warstwa przełączania (lub warstwa danych) jest tym komponentem routera, który weryfikuje pakiety w interfejsie wejściowym i transportuje je do odpowiedniego interfejsu wyjściowego. Routery zazwyczaj zawierają wiele warstw przełączania połączonych krzyżowo, dzięki czemu mogą przekazywać wiele strumieni równolegle. Poszczególne warstwy przełączania mogą być dodawane w formie kart rozszerzeń z układami ASIC przeznaczonymi do przetwarzania danych. Routery są bowiem wyposażane w płyty montażowe (ang. backplane) (mają budowę kasetową), w które wkłada się odpowiednie karty. Fizyczna budowa wielu routerów przypomina konstrukcję serwerów kasetowych. Aby zmierzyć wydajność routera, można zastosować technikę opracowaną przez grupę IETF Benchmarking Working Group (BMWG, RFC 2544); w technice tej się zakłada, że jedna połowa portów routera jest wykorzystywana do przesyłania pakietów adresowanych do drugiej połowy portów.

Działanie opisywanego podsystemu polega na wyszukiwaniu w tabeli wpisów, które odwzorowują identyfikatory sieciowe (adresy MAC) na trasy. Zgodnie z informacjami przedstawionymi wcześniej system przełączania czasami wykorzystuje bazę FIB zamiast RIB, co przyspiesza wykonanie operacji. Przeszukiwanie baz danych bazuje na algorytmach zoptymalizowanych do przetwarzania adresów IP, w tym drzewach binarnych, strukturach typu radix tree i Patricia tree oraz innych rozwiązaniach opracowywanych przez producentów urządzeń.

Routery przechowują reguły określające, które pakiety należy przekazać, a które odfiltrować. Odfiltrowane pakiety zostają odrzucone. Nie towarzyszy temu odesłanie do nadawcy stosownego komunikatu ICMP. Zadania tego typu są bowiem realizowane po to, by router był niezauważalny dla hakerów. Jeśli w pamięci podręcznej routera lub w tablicy routingu nie występuje informacja, na której podstawie można podjąć decyzje o przesłaniu pakietu, a pakiet nie został odfiltrowany, router odsyła do nadawcy komunikat ICMP informujący o tym, że stacja docelowa jest nieosiągalna.

Ponieważ routery stanowią most między różnymi sieciami na poziomie warstwy sieciowej (warstwy 3.), pakiety o jednakowych protokołach sieciowych mogą być przekazywane bezpośrednio, bez szczególnego przetwarzania. Operacja ta często jest nazywana korzystaniem z szybkiej ścieżki przetwarzania. Pakiety wymagające dodatkowego przetwarzania są przekazywane wzdłuż wolniejszej ścieżki przetwarzania.

Routery realizują również inne zadania. Mogą służyć jako urządzenia zabezpieczające transmisję — szyfrujące transmisję w sposób właściwy dla zaimplementowanej technologii. Moduł odpowiedzialny za wykonywanie tych zadań jest często nazywany warstwą usługową. Sama realizacja wspomnianej funkcji wymaga od routera zdekodowania nagłówka pakietu w warstwie sieciowej, wyodrębnienia i przetworzenia danych zawartych w pakiecie i w razie potrzeby odczytania innych pól pakietu.

Routery mogą również wymuszać zachowanie odpowiedniego poziomu QoS, segregując pakiety, gdy jest to konieczne. Przepełnienie bufora wiąże się z tym, że router nie może przetwarzać nadchodzących danych i musi je odrzucać. Sposób wyboru pakietów do odrzucenia zależy od zastosowanego algorytmu. Najczęściej stosowane techniki to:

* **Algorytm odrzucania ostatnich pakietów** (ang. tail drop). Mechanizm zarządzania kolejką sprawdza zawartość bufora i gdy wykryje przekroczenie maksymalnego poziomu zapełnienia, odrzuca wszystkie nadchodzące pakiety, aż do zwolnienia miejsca. W rozwiązaniu tail drop nie stosuje się rozróżnienia na rodzaje pakietów, źródła danych lub inne właściwości, które miałyby wpływ na podejmowane decyzje.

Gdy system nadawczy przez pewien czas nie będzie odbierał segmentów ACK, uzna, że generowane przez niego pakiety są odrzucane. Wówczas zacznie zmniejszać częstotliwość wysyłania danych aż do uzyskania równomiernego strumienia potwierdzeń. Wadą mechanizmu tail drop jest to, że systemy nadawcze równocześnie rozpoczynają retransmisję pakietów, co prowadzi do zalewu danymi.

* **Wczesne losowe wykrywanie** (RED — Random Early Detection). Zasada działania tego algorytmu polega na monitorowaniu średniej długości kolejki i odrzucaniu pakietów na podstawie funkcji probabilistycznych. Uwzględnienie w działaniu algorytmu RED danych statystycznych daje gwarancję, że pakiety pochodzące ze źródła generującego duże ilości danych będą odrzucane z większym
* prawdopodobieństwem niż pakiety wysyłane przez źródło dostarczające niewiele danych. W przypadku tego rozwiązania nie występuje problem zalewu pakietami (globalnej synchronizacji), charakterystyczny dla techniki tail drop.
* **Ważony algorytm RED i adaptacyjny (aktywny) algorytm RED**. Pierwszy z wymienionych algorytmów wykorzystuje technikę RED, ale dodatkowo przypisuje pakietom pewne wartości priorytetów. Z kolei adaptacyjny lub aktywny algorytm RED wprowadza zmiany w funkcji statystycznej w zależności od stanu kolejki.

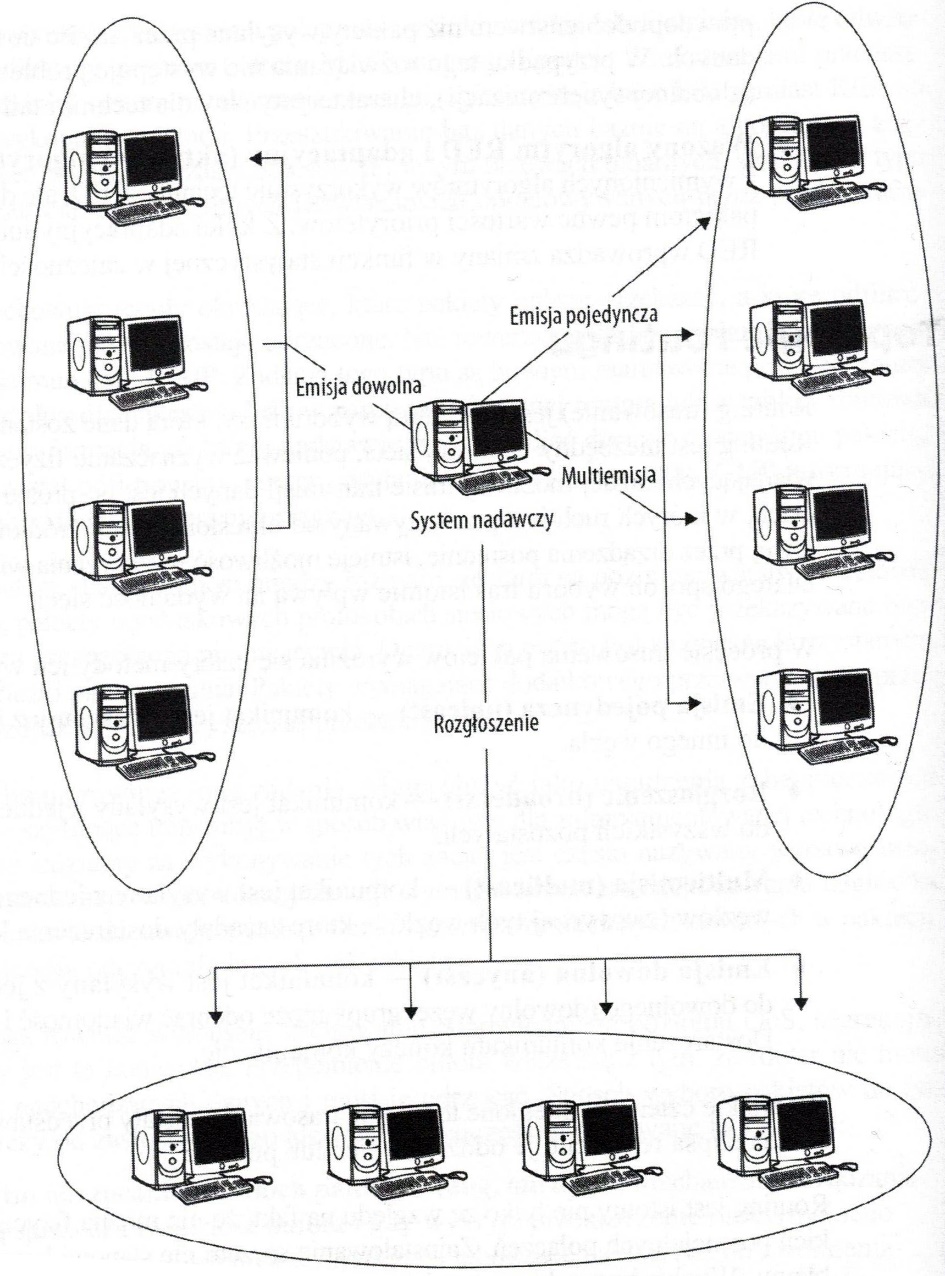
**Topologie routingu**

Routing (trasowanie) jest procedurą wyboru trasy, którą dane zostaną przesłane przez sieć. Routing jest niezbędny w każdej sieci, ponieważ wyznaczanie fizycznych obwodów odpowiadających każdej możliwej trasie transmisji danych jest po prostu niepraktyczne. W sieciach, w których ruch jest przekazywany od określonej stacji źródłowej do jednostki docelowej przez urządzenia pośrednie, istnieje możliwość wyznaczenia więcej niż jednej ścieżki. Dlatego sposób wyboru tras istotnie wpływa na wydajność sieci.

W procesie trasowania pakietów wyróżnia się cztery metody ich wysyłania:

* **Emisja pojedyncza (unicast)** — komunikat jest przesyłany z jednego węzła do innego węzła.
* **Rozgłoszenie (broadcast)** — komunikat jest wysyłany z jednego węzła do wszystkich pozostałych.
* **Multiemisja (multicast)** — komunikat jest wysyłany z jednego węzła do grupy węzłów (zazwyczaj tych węzłów, które zażądały dostarczenia komunikatu).
* **Emisja dowolna (anycast)** — komunikat jest wysyłany z jednego węzła do dowolnego (dowolny węzeł grupy może odebrać wiadomość i ją wykorzystać). Dostarczenie komunikatu kończy komunikację.

Wszystkie cztery wymienione topologie trasowania zostały przedstawione na rysunku.



Każda elipsa reprezentuje oddzielną sieć lub podsieć.

Routing jest istotny nie tylko ze względu na fakt, że nie można fizycznie utworzyć wszystkich potencjalnych połączeń. Zainstalowanie sprzętu nie stanowi bowiem rozwiązania problemu. Wyobraźmy sobie sytuację, w której po zaobserwowaniu dużego natężenia ruchu między dwoma punktami końcowymi administrator tworzy łącze szkieletowe o podobnej pojemności, ale o mniejszej długości i większej szybkości działania. Routery wykrywają nowe połączenie i ustalają, że jest ono trasą o najniższym koszcie. Cały ruch zostaje wówczas skierowany do nowego łącza szkieletowego, wysycając je i ograniczając ogólną wydajność sieci. Opisana sytuacja jest znana jako paradoks Braessa. Gdy wybór tras ogranicza się do wyszukania ścieżki o najniższym koszcie, dodatkowa pojemność sieci zostanie całkowicie wyczerpana, co w niektórych przypadkach może prowadzić do obniżenia wydajności systemu. Analogiczną zależność zaobserwowano w Bostonie, gdzie zamknięcie bardzo obciążonej drogi spowodowało rozłożenie ruchu na inne ulice, prowadząc w konsekwencji do zwiększenia efektywności transportu.

Paradoks Braessa wynika z teorii gier opracowanej przez Johna Forbesa Nasha, fizyka z uniwersytetu w Princeton, który za swoją pracę otrzymał Nagrodę Nobla. Zgodnie z zasadą równowagi Nasha w dowolnym systemie z wieloma graczami, w którym każdy gracz działa we własnym interesie, biorąc pod uwagę postępowanie innych uczestników gry, żadnemu z graczy nie opłaca się jednostronnie zmienić strategię, gdyż nie przyniesie to dodatkowych korzyści. Z twierdzenia tego wynika, że systemy pozostające w równowadze Nasha nie zawsze uzyskująnajlepsze wyniki jako całość. Aby osiągać najlepsze indywidualne rezultaty, od równowagi Nasha muszą odstępować całe grupy systemów.

To jest przestrzeń działania protokołów routingu. Aby routing był efektywny, musi mieć dynamiczny charakter. W dynamicznych systemach sieć reaguje na zdarzenia w celu kontynuowania pracy i wybiera trasy do grup systemów, których poszczególne systemy samodzielnie

nie byłby w stanie wybrać. Na przykład jeśli koparka zniszczy ułożony w ziemi kabel telefoniczny, adaptacyjny protokół routingu doprowadzi do przekazania mchu do innego łącza. Analogicznie, jeśli krótkie, wysoko wydajne łącze stanie się dostępne, mechanizm routingu dynamicznego rozłoży ruch w taki sposób, aby obciążenie całego systemu było równomierne. W konkretnym analizowanym przypadku wybrana trasa może się okazać dłuższa, ale całkowita wydajność systemu będzie optymalna.

**Metody optymalizacji**

W niewielkich sieciach administrator może osobiście zdefiniować preferowane trasy między punktami końcowymi, dodając wpisy do tablicy routingu. Rozwiązanie to okazuje się niepraktyczne w sieciach o dużym rozmiarze. Z tego względu wykorzystuje się w nich trasy, które zostały wstępnie obliczone lub są obliczane na bieżąco w przypadku wystąpienia takiej konieczności. W sieciach PSTN stosuje się rozwiązania, w których tablice są wstępnie wypełniane preferowanymi (wyliczonymi) trasami oraz trasami zapasowymi. Wraz z roz-wojem sieci telefonicznej zaczęto wprowadzać technologie routingu adaptacyjnego, których tablice routingu były generowane przez protokoły routingu i mogły uwzględniać automatyczne zmiany tras. W przypadku intemetu mechanizm routingu jest całkowicie inny —jest w pełni dynamiczny.

Mechanizmy routingu są wdrażane w ramach systemów autonomicznych lub między systemami autonomicznymi. Systemem autonomicznym (AS —Autonomous System) nazywa się zbiór jednostek działających w ramach wspólnej struktury administracyjnej. Systemem autonomicznym może być sieć, grupa sieci, zbiór sieci dostawcy usług internetowych lub cały internet. Protokoły routingu działające w systemach autonomicznych są nazywane protokołami bram. Protokoły bram wewnętrznych (IGP — Interior Gateway Protocol) są stosowane do przenoszenia pakietów w ramach sieci systemu autonomicznego. Typowymi przykładami protokołów IGP są protokół informowania o trasach (RIP — Routing Information Protocol), opracowany przez firmę Cisco protokół routingu bramy wewnętrznej (IGRP — Interior Gateway Routing Protocol), otwarty protokół wyszukiwania najkrótszych tras (OSPF — Open Shortest Fast First) i protokół łączenia systemów pośrednich (IS-IS — Intermediate System to Intermediate System). Protokoły bram zewnętrznych (EGP — Exterior Gateway Protocol) obejmują pierwotny standard EGP (obecnie uznany za przestarzały) oraz protokół bramy brzegowej (BGP — Border Gateway Protocol). Poszczególne protokoły routingu są omówione dalej.

**Algorytm wektora odległości**

Algorytm wyznaczania wektora odległości (DV — Distance Vector) wyraża koszt trasy liczbą przeskoków pakietu (liczbą węzłów na trasie). Dane są przekazywane wzdłuż trasy, której sumaryczna liczba przeskoków jest najniższa. Każdy węzeł w sieci tworzy własną tablicę odległości i przekazuje ją do węzłów sąsiednich. Routing DV jest bardzo popularnym rozwiązaniem w sieciach pakietowych i jest podstawą działania protokołów RIP (w wersji 1. i 2.) oraz IGRP. Podobna metodologia działania została uwzględniona w protokołach BGP, czyli w podstawowym protokole internetowym, oraz EGP, czyli w starszej i obecnie niestosowanej wersji protokołu routingu zewnętrznego.

Sam algorytm wektora odległości nie wykrywa pętli w sieci i nie eliminuje ich, dlatego protokoły routingu wykorzystujące ten algorytm są wyposażone w szereg dodatkowych mechanizmów pozwalających na wyeliminowanie tego problemu. Tablice routingu są budowane na podstawie informacji o możliwych trasach dostarczania pakietów, zoptymalizowanych w odniesieniu do poszczególnych odcinków połączeniowych. Optymalizacja wyznaczonych tras polega na użyciu algorytmu Bellmana-Forda w odniesieniu do danych zawartych w tablicy routingu. W wyniku jego działania informacje o preferowanych trasach są przekazywane do routerów sąsiednich, które aktualizują własne tablice routingu.

**Algorytm Bellmana-Forda**

Algorytm Bellmana-Forda prowadzi do wyznaczenia najkrótszej trasy w grafie o ważonych krawędziach. Mechanizm został opracowany niezależnie przez Richarda Bellmana w 1958 roku oraz Lestera Forda Jr. w 1956 roku. Większość protokołów bazujących na algorytmie Bellmana-Forda wykorzystuje jego wersję rozproszoną. Rozproszony algorytm Bellmana- -Forda zakłada użycie trzech mechanizmów w celu zapełnienia tablicy routingu w każdym z węzłów:

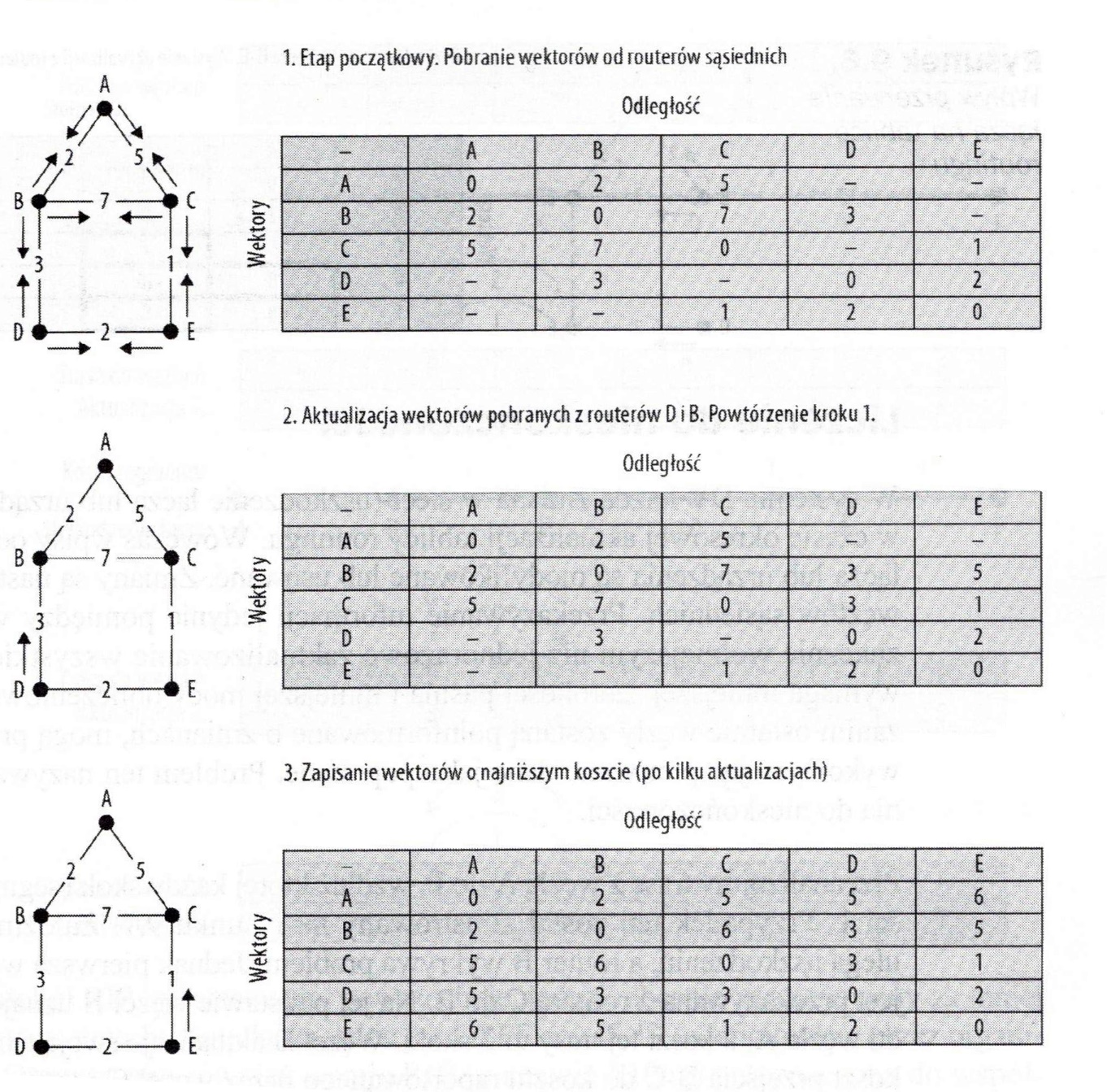
1. **Stan początkowy**. Każdy router zawiera tablicę routingu z trasami (wektorami) do bezpośrednio przyłączonych sieci. Poszczególne wpisy składają się z trzech wartości [cel, dystans, następca]. Następcą jest router lub węzeł o jeden krok bliżej celu (jest to kolejny węzeł na trasie) i jest jednocześnie najbliższym sąsiadem analizowanego routera. Dystans może być wyrażony liczbą przeskoków pakietu lub wartością kosztu wyznaczoną na podstawie przepustowości połączenia lub innego czynnika.

**2.** **Wysyłanie.** Każdy węzeł wysyła znane sobie wektory tras [cel, dystans]

do bezpośrednich sąsiadów. Operacja jest ponawiana okresowo (zależnie od ustawień z interwałem od sekundy do minuty) oraz każdorazowo po zmianie wpisu w tablicy routingu.

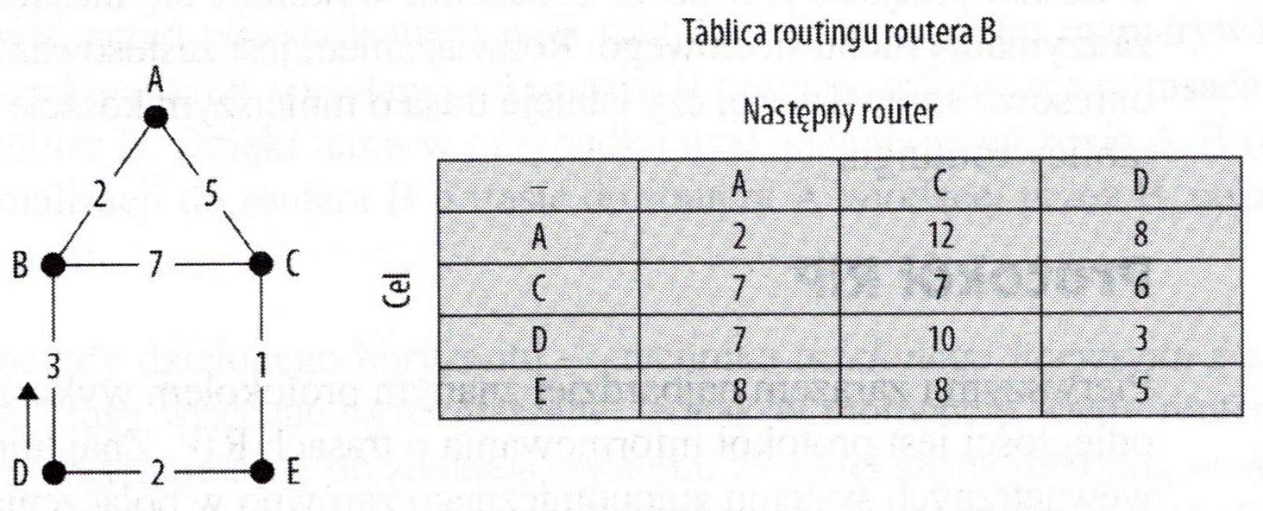
**3. Odbieranie.** Każdy router oblicza koszt tras do sieci odległych na podstawie informacji odebranych od najbliższych sąsiadów. Po zaktualizowaniu tablicy routingu router przystępuje do realizacji kroku 2. i wysyła nowe informacje do najbliższych sąsiadów.

Opisana operacja została zilustrowana na rysunku



Tablica routingu przedstawiona w górnej części rysunku jest wypełniona informacjami o najbliższych sąsiadach. Ponieważ router A nie zna najkrótszej trasy do routera E, pole wektora jest puste. Środkowa tablica routingu reprezentuje stan po przesłaniu pierwszej aktualizacji z węzła D do B. W wyniku tej operacji router B może uzupełnić pole wektora opisującego trasę z B do E wartością 5. Choć nadal nie ma pewności, czy jest to najniższy koszt przejścia z B do E. Do czasu uzupełnienia wektora E-C router B nie ma informacji o tym, że trasa B-D-E rzeczywiście ma najniższy koszt (wartość 5) — trasa B-C-E ma koszt 8. Dolna tablica routingu odzwierciedla stan po zaktualizowaniu tablic routingu najbliższych sąsiadów, przesłaniu aktualizacji z węzła E do C i wykonaniu dostatecznej liczby aktualizacji, aby tabela wypełniła się wektorami o najniższych kosztach.

Tablica routingu przechowywana w poszczególnych routerach ma format podobny do przedstawionego na rysunku



różniący się od zestawienia najniższych kosztów zaprezentowanego na poprzednim rysunku. Jako przykład pokazana została tablica routera B, której wektory [cel, dystans, następca] zaprezentowano w formie tabeli. Każdy wiersz odpowiada określonemu celowi, kolumna wyznacza następcę, a dystans jest wpisany w polu tabeli.

Przeanalizujmy, co się stanie, gdy połączenie między węzłami B i D zostanie uszkodzone.

Obraz zawierający stół

Opis wygenerowany automatycznie

Przerwa zostanie wykryta przez obydwa routery. Każdy z nich wyśle więc natychmiast aktualizację do najbliższych sąsiadów. W wyniku tej operacji w tablicy routingu zmienia się wiele wektorów. Wartość każdego z nich została pogrubiona.

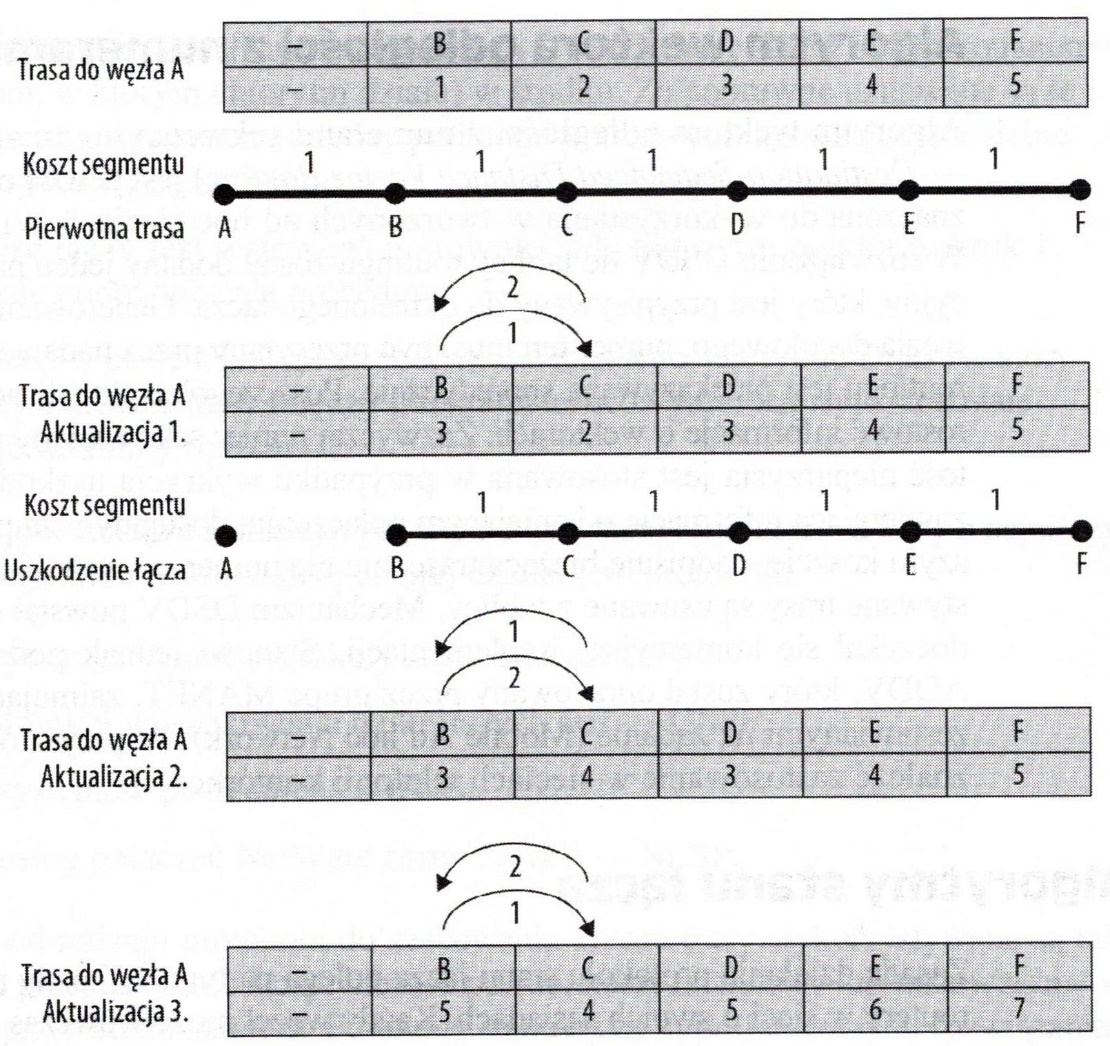
**Liczenie do nieskończoności**

W systemie DV każda zmiana w sieci (uszkodzenie łącza lub urządzenia) jest wykrywana w czasie okresowej aktualizacji tablicy routingu. Wówczas wpisy odnoszące się do danego łącza lub urządzenia są modyfikowane łub usuwane. Zmiany są następnie propagowane do węzłów sąsiednich. Przekazywanie informacji jedynie pomiędzy węzłami jest procesem znacznie wolniejszym niż jednorazowe zaktualizowanie wszystkich węzłów sieci, choć wymaga mniejszej szerokości pasma i mniejszej mocy obliczeniowej. Oznacza jednak, że zanim ostatnie węzły zostaną poinformowane o zmianach, mogą prowadzić komunikację, wykorzystując pierwotne dane jako poprawne. Problem ten nazywa się problemem liczenia do nieskończoności.

Przeanalizujmy trasę z węzła A do F, wzdłuż której każdy skok (segment) ma koszt o wartości 1. Przypadek ten został zilustrowany na rysunku 9.7. Załóżmy, że połączenie A-B ulega uszkodzeniu, a router B wykrywa problem. Jednak pierwsza wymieniana aktualizacja jest przekazywana z routera C do B. Na jej podstawie węzeł B uznaje, że router C zna trasę do węzła A, a koszt tej trasy to 2 skoki. Węzeł B aktualizuje swoją tablicę routingu — dodaje koszt przejścia B-C do kosztu raportowanego przez węzeł C i umieszcza wartość 3 w swojej tablicy routingu. Jednocześnie przyjmuje, że trasa do węzła A wiedzie przez węzeł C. Router C nadal uznaje, że węzeł B jest najbliższym pośrednikiem na trasie do A, i gdy odbiera aktualizację z routera B (o wartości 3), zmienia wartość kosztu na 4. Zmiana jest wprowadzana również we wszystkich routerach za węzłem C. Trzecia aktualizacja powiela błąd pierwszej aktualizacji — węzeł B zmienia koszt przez dodanie do wartości zapisanej w węźle C kosztu przejścia z B do C. Procedura wykonuje się nieskończenie długo, prowadząc do zatrzymania ruchu sieciowego. Rozwiązaniem jest zastosowanie techniki relaksacji, w której okresowo sprawdza się, czy istnieje trasa o mniejszym koszcie niż ten, który jest wpisany do tablicy routingu.

**Protokół RIP**

Pierwszym i zarazem najbardziej znanym protokołem wykorzystującym algorytm wektora odległości jest protokół informowania o trasach RIP. Znajduje on zastosowanie w sieciach wewnętrznych systemu autonomicznego zarówno w połączeniach LAN, jak i WAN. Pierwsza wersja protokołu została opisana w zaleceniu IEEE RFC 1058 w roku 1988. Wersja druga jest zdefiniowana w dokumencie RFC 2453. RIP był pierwszym protokołem routingu stosowanym w Internecie. W rozwiązaniu tym metryką kosztu jest liczba przeskoków. Maksymalna liczba węzłów na trasie wynosi 15, a czas przetrzymywania jednej trasy (bez od-świeżenia wpisu) wynosi 180 sekund. Interwał ponawiania okresowych aktualizacji oscyluje nieznacznie wokół standardowej wartości (30 sekund), dzięki czemu routery nie są obciążane zbyt wieloma aktualizacjami w jednym czasie.



Dawniej protokół RIP był stosowany powszechnie, jednak obecnie jest uznawany za mniej efektywny niż protokoły stanu łącza, takie jak OSPF czy IS-IS (zdefiniowany przez organizację ISO). Opracowano również wersję RIP (o nazwie RIPng) przeznaczoną do współdziałania z protokołem IPv6, w której dodatkowo zaimplementowano rozwiązania gwarantujące, że przedawnione lub niekompletne informacje nie są przekazywane w sieci. Jeden z mechanizmów — nazywany dzielonym horyzontem — uniemożliwia routerowi przesyłanie informacji o trasie do węzła, od którego pozyskał informację o tej trasie.

Mechanizm dzielonego horyzontu skutecznie eliminuje problem liczenia do nieskończoności i zabezpiecza sieć przed powstawaniem pętli routingu. W przypadku rozpatrywanym powyżej router C w aktualizacji wysyłanej do routera B pominie informacje o trasach prowadzących przez router B. Dzięki temu w przypadku uszkodzenia połączenia A-B router C nie przekaże aktualizacji do routera B o trasie do routera A wiodącej przez B, co zapobiegnie powstaniu pętli.

Istnieje odmiana metody dzielonego horyzontu — technika dzielonego horyzontu z zatruwaniem wstecznym. Trasy wsteczne są rozgłaszane, jednak w momencie, kiedy router wykryje, że regularnie otrzymuje trasę do zdalnego systemu ze zwiększoną metryką, w swojej aktualizacji zwrotnej zatruwa tę trasę poprzez zwiększenie metryki do 16 (dla protokołu RIP — sieć nieosiągalna). Technika dzielonego horyzontu jest wykorzystywana również w protokołach IGRP, EIGRP oraz w usłudze wirtualnych sieci prywatnych (VPLS — Virtual Private LAN Service).

**Algorytm wektora odległości z numerami sekwencyjnymi**

Algorytm wektora odległości z numerami sekwencyjnymi sieci docelowych (DSDV — Destination-Sequenced Distance Vector Routing} jest pewną odmianą systemu DV przeznaczoną do wykorzystania w tworzonych ad hoc sieciach Wi-Fi (sieciach mobilnych). W rozwiązaniu DSDV do tablicy routingu został dodany jeden parametr — numer sekwencyjny, który jest przypisywany do określonego łącza. Generowanie numeru należy do zadań węzła docelowego; numer ten musi być przesyłany przez nadawcę aktualizacji. Cała tablica routingu jest przekazywana sporadycznie. Poszczególne aktualizacje zawierają jedynie przyrostowe informacje o wektorach. Zazwyczaj numer sekwencyjny ma wartość parzystą. Wartość nieparzysta jest stosowana w przypadku wykrycia uszkodzenia łącza. Aktualizacja zawierająca informacje o istniejącym połączeniu, dostępnym za pośrednictwem trasy o niższym koszcie, nadpisuje bieżącą trasę, ale nie numer sekwencyjny. Okresowo niewykorzystywane trasy są usuwane z tablicy. Mechanizm DSDV powstał dość dawno, ale nigdy nie doczekał się komercyjnej implementacji. Stanowi jednak podstawę działania protokołu AODV, który został opracowany przez grupę MANET, zajmującą się mobilnymi sieciami zestawianymi na żądanie (Mobile Ad hoc Network). W przyszłości protokół AODV może znaleźć zastosowanie w sieciach telefonii komórkowej.

**Algorytmy stanu łącza**

Zasada działania protokołu stanu łącza polega na tym, że każdy router informuje pozostałe routery w sieci o swoich sąsiadach. Każdy węzeł może wówczas utworzyć mapę (graf) sieciowych powiązań poszczególnych routerów ze sobą w centrum tej topologii. Utworzone mapy służą do obliczenia najkrótszych tras. Zazwyczaj za ustalenie najkrótszej trasy odpowiada algorytm Dijkstry. Protokoły wektora odległości wymuszają na routerach wymianę tablic routingu, natomiast protokoły stanu łącza przekazują jedynie informacje, w których zawiera się identyfikator routera wysyłającego informację, dane o stanie łączy, jakie posiada, i informację o sąsiadach, którzy są podłączeni przez te łącza. Gdy stan łącza ulegnie zmianie (z włączonego na wyłączony lub z wyłączonego na włączony), generowana jest aktualizacja i informacja o bieżącym stanie łącza jest dostarczana do wszystkich routerów w sieci w obszarze działania protokołu.

Działanie protokołu stanu łącza przebiega zgodnie z poniższą procedurą:

1. Router wysyła do wszystkich węzłów informacje o stanie swoich łączy, a odebrane informacje zapisuje w bazie topologii.

2. Algorytm routingu przypisuje każdej trasie numer sekwencyjny.

3. Ogłoszenie o stanie łącza (LSA —Link State Advertisement) jest wysyłane okresowo do wszystkich węzłów w obszarze działania protokołu.

4. Jeśli nie został wcześniej zarejestrowany numer sekwencyjny aktualizacji pochodzącej z określonego węzła, nowa informacja jest zapisywana w bazie topologii. Jeśli odebrana informacja ma wyższy numer sekwencyjny niż zapisany w bazie topologii, nadpisuje wcześniejszą informację.

Kroki 3. i 4. są powtarzane przez wszystkie routery w domenie routingu. Aktualizacje są przesyłane do wszystkich węzłów za pomocą multiemisji. Router z włączonym protokołem stanu łącza okresowo wysyła pakiet o nazwie HELLO w celu sprawdzenia stanu łączy.

5. Algorytm stanu łącza analizuje bazę topologii i tworzy mapę sieci z danym routerem (z routerem, w którym algorytm działa) w środku. Za poprawne uznawane są te łącza, które są ustanowione między punktami końcowymi zgłaszającymi siebie nawzajem jako węzły sąsiednie.

6. Dostępność łączy jest testowana ponownie, gdy algorytm powtórzy krok 1.

i rozpocznie ponownie całą procedurę.

7. Dane zebrane w bazie topologii są poddawane działaniu algorytmu Dijkstry, który wyznacza najkrótsze trasy między punktami końcowymi i zapisuje stosowne informacje w tablicy routingu.

Tablica routingu w routerze z działającym protokołem stanu łącza zawiera trasy o najniższym koszcie do wszystkich węzłów sieci objętej działaniem protokołu.

Najczęściej wykorzystywane protokoły stanu łącza to:

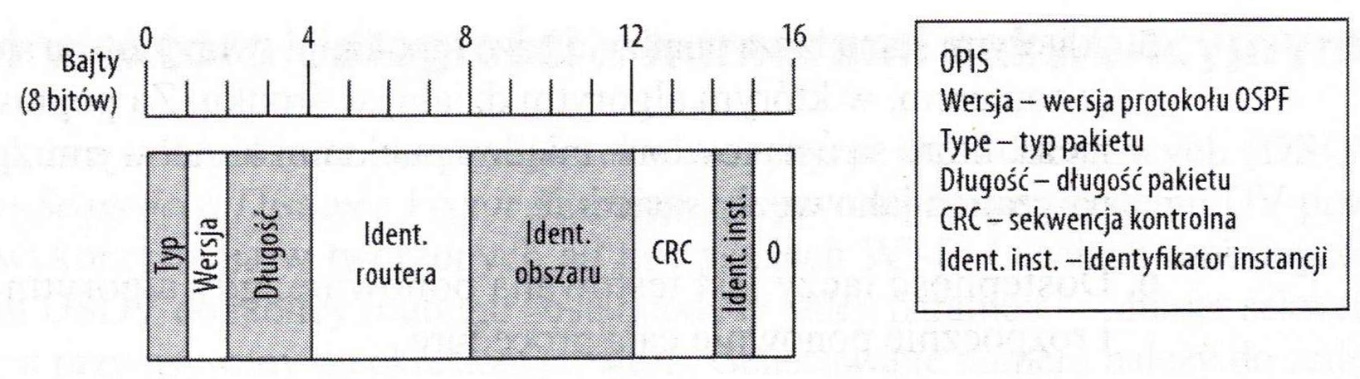
* otwarty protokół wyszukiwania najkrótszych tras — OSPF,
* protokół systemów pośrednich — IS-IS,
* protokół usług połączeń NetWare firmy Novell — NLSP.

W zależności od rodzaju protokołu do szacowania kosztu trasy wykorzystywane są takie parametry, jak przepustowość łączy, bieżące pasmo, koszt finansowy lub inne właściwości zdefiniowane przez administratora. Protokoły stanu łącza są preferowanymi rozwiązaniami w dużych sieciach komputerowych, ponieważ gwarantują szybszą reakcję na zmiany w sieci niż protokoły wektora odległości. Są dominującymi protokołami stosowanymi wewnątrz systemów autonomicznych.

**Protokół OSPF**

Protokół OSPF jest najczęściej stosowanym protokołem routingu z algorytmem stanu łącza. Jest powszechnie wykorzystywany jako protokół wewnętrzny w ramach AS-ów oraz w wielu innych sieciach. Ostatnia opracowana — trzecia — wersja standardu OSPF jest opisana w dokumencie RFC 5340, wydanym w roku 2008. Wersja ta obsługuje protokół IPv6.

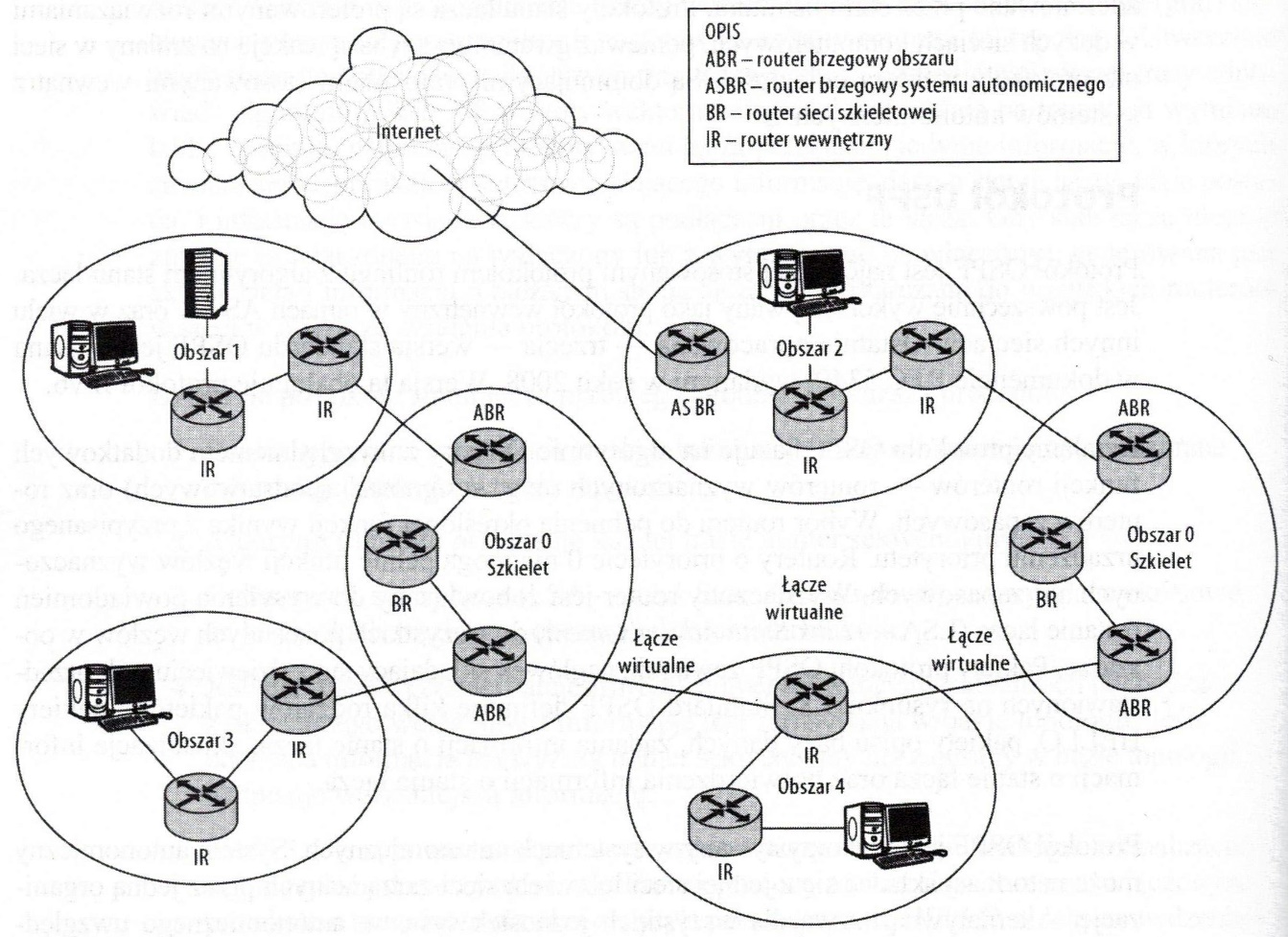
Działanie protokołu OSPF bazuje na algorytmie Dijkstry z uwzględnieniem dodatkowych funkcji routerów — routerów wyznaczonych (ang. designated) (podstawowych) oraz routerów zapasowych. Wybór routera do pełnienia określonej funkcji wynika z przypisanego urządzeniu priorytetu. Routery o priorytecie 0 nie mogą pełnić funkcji węzłów wyznaczonych ani zapasowych. Wyznaczony router jest zobowiązany do wysyłania powiadomień o stanie łącza (LSA — Link State Advertisement) do wszystkich pozostałych węzłów w obszarze. Pakiety protokołu OSPF zawierają nagłówek składający się z dziewięciu pól, przedstawionych na rysunku.



Standard OSPF definiuje kilka rodzajów pakietów: pakiety HELLO, pakiety opisu bazy danych, żądania informacji o stanie łącza, aktualizacje informacji o stanie łącza oraz potwierdzenia informacji o stanie łącza.

Protokół OSPF jest wykorzystywany w systemach autonomicznych. System autonomiczny może natomiast składać się z jednej sieci lub wielu sieci zarządzanych przez jedną organizację. Alternatywną nazwą dla wszystkich jednostek systemu autonomicznego uwzględnionych w tej samej mapie topologicznej jest „domena routingu”. Protokół OSPF wydziela obszary jako osobne jednostki topologiczne, dzięki czemu dany obszar nie ma informacji o routingu w innym obszarze. Rozwiązanie ma na celu obniżenie natężenia mchu sieciowego oraz przyspieszenie procesu wyszukiwania najlepszych tras do poszczególnych obszarów.

Grupy obszarów są łączone ze sobą przez routery brzegowe obszarów w sieć szkieletową OSPF. Sieć szkieletowa sama jest zorganizowana na zasadzie obszaru OSPF, a informacje o routingu w ramach tego obszaru są niezależne od informacji wymienianych w obszarach dołączonych. Można również zorganizować sieć szkieletową OSPF w taki sposób, aby stanowiła ona dwie lub większą liczbę niepołączonych grup. Aby przekształcić sieć szkieletową na ciągłą, należy wyznaczyć łącza wirtualne wiodące przez routery w obszarach innych niż szkieletowy, które funkcjonują jako połączenia między routerami grup szkieletowych. W obszarze szkieletowym OSPF są wyznaczane routery brzegowe, które zapewniają połączenie z routerami obsługującymi protokoły routingu zewnętrznego, takie jak BGP lub EGP. Na rysunku



została przedstawiona sieć z kilkoma obszarami, obszarem szkieletowym oraz łączem wirtualnym.

**Protokół IS-IS**

Protokół IS-IS jest drugim pod względem popularności protokołem stanu łącza wykorzystywanym w sieciach pakietowych. Najczęściej znajduje zastosowanie w sieciach dostawców usług internetowych oraz dużych sieciach korporacyjnych, gdzie funkcjonuje jako protokół wewnętrzny w ramach systemu autonomicznego. Wymianę danych z innymi systemami autonomicznymi zapewniają protokoły routingu zewnętrznego.

Protokół IS-IS został opracowany przez Digital Equipment Corporation jako element protokołu DECnet w latach 80. i został przyjęty jako standard ISO — ISO/IEC 10589.2002. Ponieważ IS-IS nie jest publicznym standardem, nie jest stosowany w Internecie — choć w 1990 roku organizacja IETF opublikowała zalecenie 10589.2002 jako dokument RFC 1142. Pierwotna wersja protokołu IS-IS została rozszerzona o obsługę protokołu IP w sieciach TCP/IP i w starszej literaturze była wyróżniana jako zintegrowany protokół IS-IS (ang. Integrated IS-IS).

Protokół IS-IS stanowi konkurencję dla protokołu OSPF —jego działanie również opiera się na algorytmie Dijkstry. Choć w obydwu rozwiązaniach uwzględniono wiele identycznych funkcji, mechanizm IS-IS jest uznawany za nieco stabilniejszy niż OSPF, ale o gorszych parametrach wydajnościowych. Dodatkowe funkcje protokołu OSPF wprowadzają większe obciążenie związane z przetwarzaniem komunikatów i prawdopodobnie dlatego protokół IS-IS lepiej się skaluje.

W standardzie IS-IS wyróżniono trzy rodzaje obszarów routingu: poziom 1. (wewnętrzny), poziom 2. (zewnętrzny) oraz poziom 1.-2. (wewnętrzny-zewnętrzny). Routery poziomów 1. i 2. mogą wymieniać informacje jedynie z routerami tego samego poziomu. Obydwie grupy mogą również komunikować się z routerami poziomu 1.-2. W przeciwieństwie do protokołu OSPF, w którym do wymiany informacji między obszarami wykorzystuje się obszar szkieletowy (obszar 0) i w którym router brzegowy obszaru jest punktem wspólnym dla obydwu obszarów, w rozwiązaniu IS-IS nie wydziela się obszarów szkieletowych, a obszary sieci nigdy na siebie nie nachodzą.

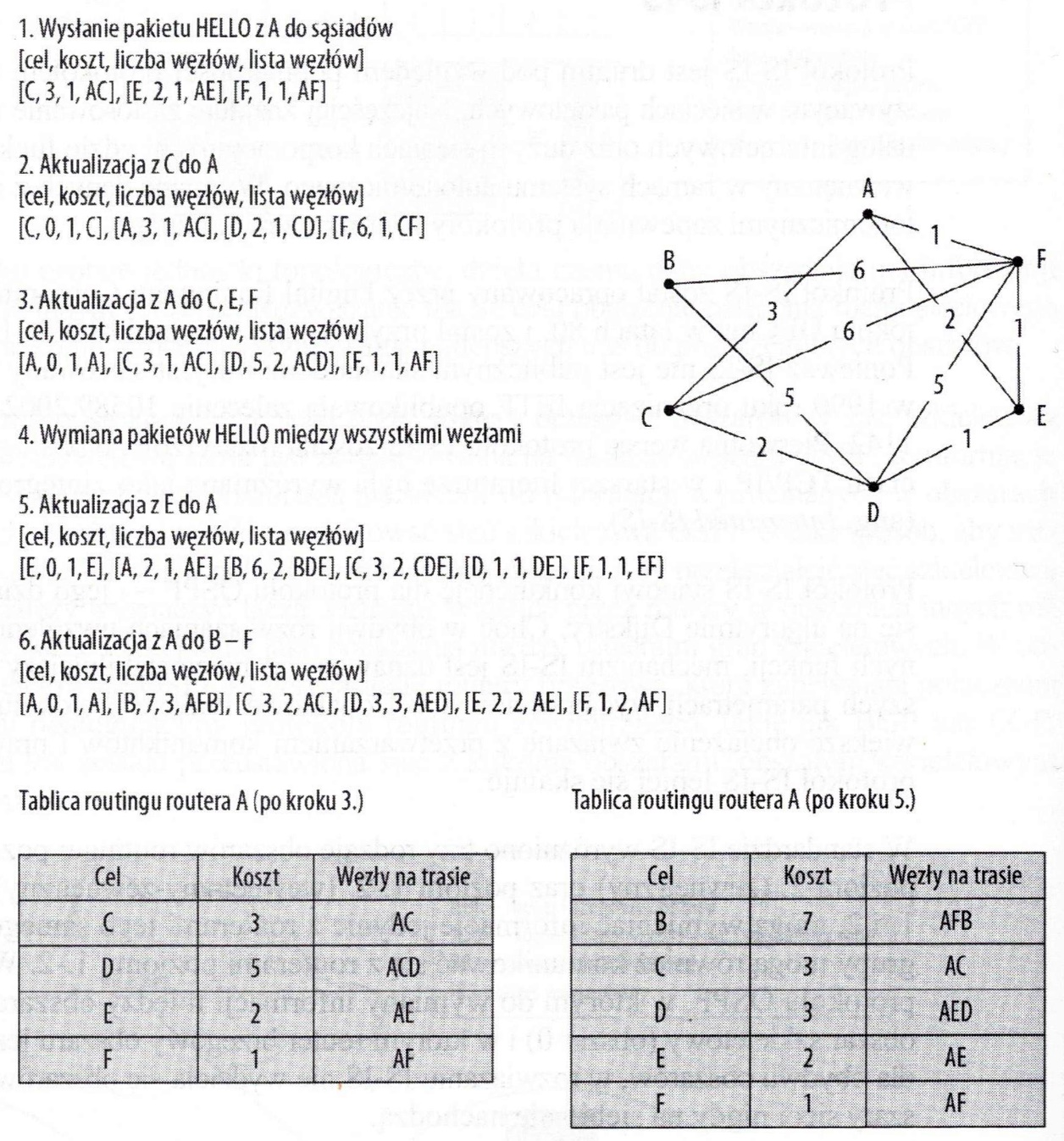
**Algorytm wektora ścieżki**

Algorytm routingu typu wektor ścieżki (ang. path vector) jest ostatnim z trzech omówionych w tej książce najważniejszych mechanizmów budowania tablic routingu. Poprzednie to algorytmy wektora odległości i stanu łącza. Algorytm wektora ścieżki jest pochodną mechanizmu DV. W rozwiązaniu tym wektory odległości od celu węzeł otrzymuje od węzłów sąsiednich wraz z pełną informacją o ścieżce, która musi zostać pokonana, aby dojść do elementu docelowego. Znając pełną ścieżkę, algorytm może znacznie łatwiej wykryć pętlę i dostosować swoją pracę do tej sytuacji niż w przypadku mechanizmów wektora odległości. W systemach obsługujących ten typ routingu przechowywane są dwie tablice — tablica bieżących ścieżek do wszystkich węzłów oraz tablica routingu z identyfikatorem najbliższego routera w ramach każdej trasy.

Przykład wykorzystania algorytmu

Rozważmy nieskomplikowany przykład wykorzystania algorytmu wektora ścieżki, przedstawiony na rysunku 9.10. Wektory przyjmują postać:

[Cel, Koszt, Liczba węzłów ścieżki, Lista węzłów ścieżki]



W kroku 1. węzeł A wysyła pakiet HELLO i gromadzi informacje na temat węzłów sąsiednich. Analogiczną operację wykonuje węzeł C. Następnie router C przesyła własne wektory do węzła A. W kroku 3. węzeł A przebudowuje swoją tablicę routingu zgodnie z pozyskanymi informacjami. Wektory routera C umożliwiają węzłowi A zdefiniowanie trasy do węzła D, ale nie wpływają na żadne inne trasy, co zostało pokazane w lewej dolnej tabeli na rysunku 9.10. W kroku 4. został przedstawiony stan po tym, jak wszystkie węzły zgromadziły informacje o sąsiadach dzięki wykorzystaniu pakietów HELLO. W kroku 5. router E wysyła własne wektory do routera A. Ten z kolei może utworzyć tablicę routingu obejmującą wszystkie węzły sieci. Informacje dostarczone z węzłów E i F powodują dodanie trasy do routera B (A-F-B) — informacja z węzła F — oraz zmianę trasy D (A-E-D) — informacja z węzła E. W ostatnim kroku węzeł A może wysłać tablicę routingu, widoczną w prawym dolnym narożniku rysunku, do wszystkich pozostałych węzłów.

W systemie routingu działającym na podstawie algorytmu wektora ścieżki jeden lub większa liczba węzłów — nazywanych węzłami rozgłaszającymi (ang. speaker node) — przechowuje tablice routingu przeznaczone dla innych przyłączonych węzłów. Odległości zawarte w tych tablicach są wyznaczane właśnie przez węzły rozgłaszające. Węzły rozgłaszające wysyłają powiadomienia o dostępnych ścieżkach do innych węzłów rozgłaszających. Algorytm wektora ścieżki ma zadanie zminimalizowania liczby domen, przez które są transportowane komunikaty, co sprawia, że jest on odpowiedni, aby stosować go jako mechanizm routingu między systemami autonomicznymi. Opisana tutaj metodologia działania jest podstawą funkcjonowania protokołu BGP.

Spośród trzech zaprezentowanych algorytmów routingu — wektor odległości, wektor ścieżki i wektor stanu łącza —jedynie protokoły wektora ścieżki znajdują praktyczne zastosowanie w routingu międzydomenowym. W przypadku algorytmu wektora odległości każdy dodatkowy węzeł pokonywany przez komunikat zwiększa prawdopodobieństwo wyboru trasy przedawnionej lub uszkodzonej. Działanie protokołów stanu łącza wymaga przenoszenia przez sieć intensywnego ruchu w przypadku zaistnienia zmiany, której prawdopodobieństwo wzrasta w zależności od liczby obsługiwanych systemów oraz dostępności w każdym węźle odpowiedniej ilości zasobów do sporządzenia mapy sieci.

**Protokół BGP**

Protokół BGP jest niezwykle łatwym w skalowaniu protokołem routingu zewnętrznego, wykorzystywanym do wymiany informacji między systemami autonomicznymi. Jego działanie bazuje na algorytmie wektora ścieżki, który został opisany we wcześniejszym punkcie rozdziału. BGP jest podstawowym protokołem routingu w Internecie — następcą protokołu EGP. Mechanizm EGP był pierwszym protokołem routingu internetowego, opracowanym przez firmę BBN Technologies w 1980 roku. Z kolei aktualna wersja protokołu BGP (wersja 4.) została opublikowana w dokumencie RFC 4271 w 2006 roku.

Rozwiązanie to jest stosowane zazwyczaj przez dostawców usług internetowych oraz przez korporacje, w których funkcjonują sieci o bardzo dużym zasięgu. Jednak z uwagi na fakt, że BGP reguluje ruch w Internecie, warto poświęcić trochę czasu na zapoznanie się z zasadami jego działania. BGP jest jedynym protokołem, który wykorzystuje protokół TCP jako mechanizm transportowy. Wymienia pakiety na porcie 179. Protokół BGP jest uruchamiany w dwóch odmianach — eBGP (zewnętrzny BGP) i iBGP (wewnętrzny BGP). Router BGP pracujący wewnątrz systemu autonomicznego jest routerem iBGP. Natomiast routery działające pomiędzy systemami autonomicznymi są routerami eBGP. Każdy router pracujący wewnątrz systemu autonomicznego, komunikujący się z jednostkami innych systemów autonomicznych, jest nazywany routerem brzegowym. Z kolei routery działające poza systemami autonomicznymi są nazywane routerami szkieletu internetowego.

Routery największych dostawców usług internetowych przechowują obecnie tablice routingu BGP złożone z około 150 000 tras. Jeśli więc ktoś uruchamia router BGP i dysponuje szybkim połączeniem (np. El) z dużą firmą telekomunikacyjną, może pobrać 150 000 tras z każdego serwisu, do którego jest przyłączony. Mechanizm BGP dzieli trasy na podstawie atrybutów (parametrów trasy), dzięki czemu można nimi wydajniej zarządzać. Wspomniane atrybuty to:

* Adres następnego routera — adres pierwszego routera na trasie do wskazanej sieci.
* Źródło danych — informacja o tym, czy dane pochodzą z routingu eBGP, czy iBGP.
* Ścieżka systemów autonomicznych — identyfikatory systemów autonomicznych zarejestrowanych na trasie powiadomienia.
* Preferowane wyjście lokalne — preferowany router wyjściowy w danym systemie autonomicznym.
* Dyskryminator wielu wyjść — atrybut niestandardowy, obsługiwany w urządzeniach firmy Cisco.
* Wyróżnik społeczności — dopuszczalne wartości to no-export (bez eksportu), no-advert1se (bez powiadomień) lub Internet (powiadomienia do wszystkich).
* Koszt lub waga trasy — atrybut niestandardowy firmy Cisco, która stosuje określenie „waga”.

Ponieważ zgodnie z hierarchią sieci internetowych dalsze uszczegółowienie tablic routingu zapewniają protokoły routingu bezklasowego (CIDR), między routerami BGP można wymieniać całe bloki adresów. System CIDR, który zastępuje starszą notację klas sieci, został szczegółowo omówiony w rozdziale 18.

**Protokół drzewa rozpinającego**

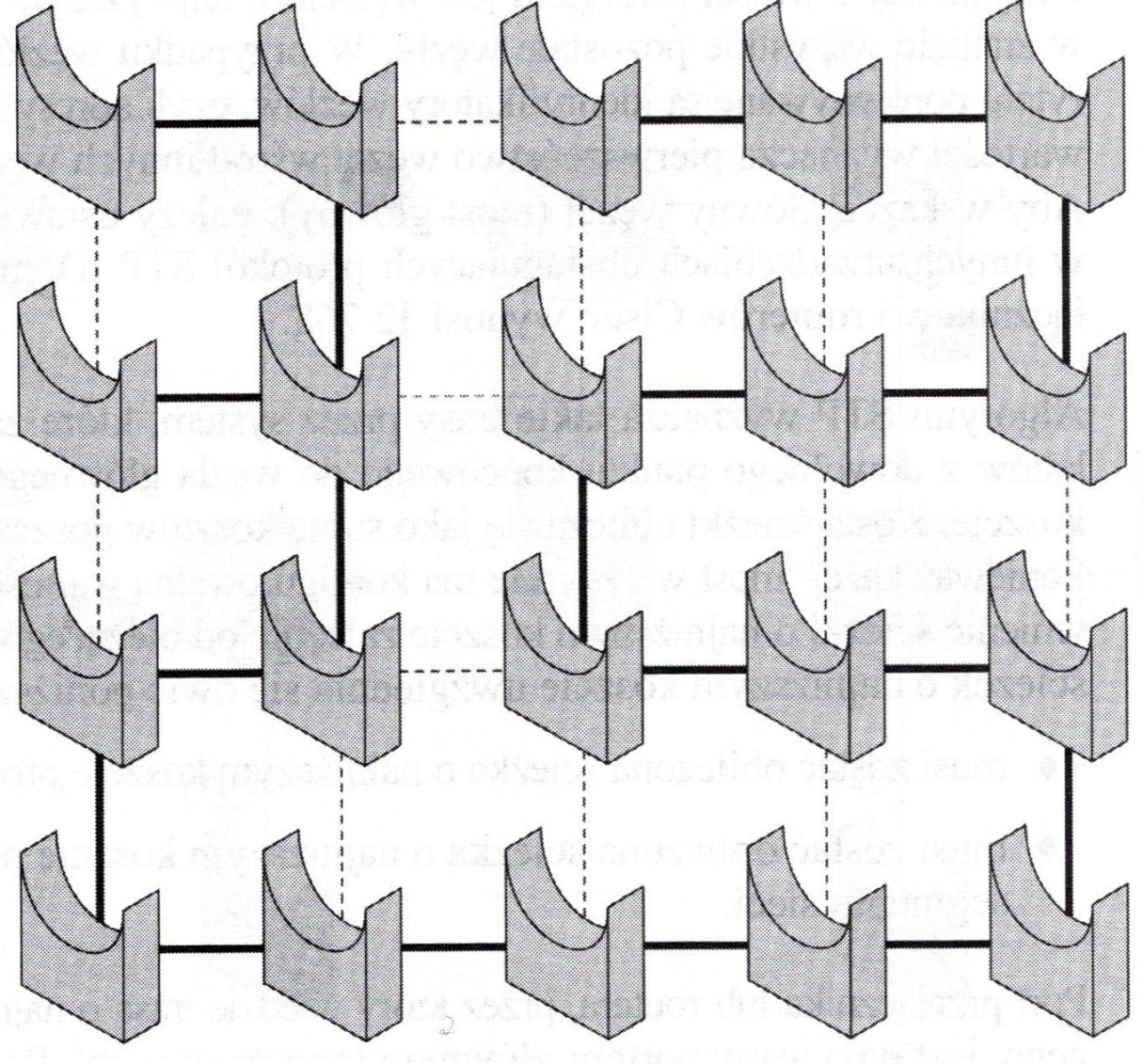
Protokół drzewa rozpinającego (STP — Spanning Tree Protocol) został opisany w standardzie IEEE 802.ID i jest technologią adaptacyjnego wydzielania ścieżek komunikacyjnych, która rozwiązuje problem pętli w sieci przez zastosowanie adaptacyjnego i dynamicznego mechanizmu wytyczania połączeń. STP jest podstawową technologią wykorzystywaną w sieciach przełączanych, która zapewnia tworzenie wirtualnych obwodów z pominięciem ewentualnych wykrytych pętli. Połączenia te są tworzone przez węzły pełniące funkcję mostów. W praktyce zazwyczaj oznacza to wykorzystanie przełączników, które wykonują zadania mostów sieciowych. Rolę tę mogą również odgrywać routery uruchomione w trybie mostów. Dodatkowa funkcja jest przez nie wykorzystywana również w różnych mechanizmach routingu.

Algorytm STP (DEC STP) został opracowany przez Radię Perlman w 1985 roku w firmie Digital Equipment Corporation przed rozwojem sieci WWW. Protokół STP działa na poziomie warstwy łącza danych modelu OSI, ponad warstwą fizyczną, i jest implementowany w urządzeniach takich jak przełączniki i routery.

W sieci hierarchicznej główny węzeł jest połączony z pewną liczbą węzłów pierwszego poziomu, a te z kolei z następnymi węzłami. Opisana hierarchia przypomina drzewo odwrócone korzeniem do góry, w którym wyłączenie jednego węzła powoduje niedostępność węzłów umiejscowionych na niższym poziomie hierarchii. W konfiguracji w pełni hierarchicznej komunikacja między danym węzłem a węzłem innej gałęzi wymaga przekazania informacji w górę drzewa, do poziomu węzła głównego, a następnie w dół, do węzła docelowego. Z tego względu jedynie sieci o niewielkich rozmiarach są w pełni zgodne ze strukturą drzewa.

Rozwiązanie stosowane w dużych sieciach polega na utworzeniu połączeń krzyżowych między poszczególnymi gałęziami drzewa. Połączenia tego typu skracają trasy i zapewniają większą wydajność komunikacji. Dodatkowo zapewniają nadmiarowość połączeń, ponieważ większość potencjalnych połączeń może być zrealizowana za pomocą kilku ścieżek komunikacyjnych. Jednak połączenia krzyżowe mogą doprowadzić do powstania pętli w sieci.

W teorii grafów drzewa rozpinające są tworzone za pomocą algorytmu, który wyznacza zbiór tras przez system połączonych ze sobą węzłów, ale w taki sposób, że każdy węzeł należy przynajmniej do jednej gałęzi, a wśród połączeń nie występują pętle. Przykład drzewa rozpinającego został pokazany na rysunku.



Linie ciągłe reprezentują połączenia wchodzące w skład drzewa rozpinającego, natomiast linie przerywane symbolizują odcinki wykluczone.

Do wyznaczenia drzewa rozpinającego można zastosować wiele algorytmów. W jednym z rozwiązań każdej krawędzi przypisuje się pewną wagę (otrzymujemy graf ważony), a algorytm drzewa rozpinającego wyznacza ścieżki przez system, które cechują się najniższą sumą wag, generując w rezultacie minimalne drzewo rozpinające. W systemach wielodo- menowych połączenie kilku minimalnych drzew rozpinających jest nazywane minimalnym lasem rozpinającym.

Dopuszczalne jest wprowadzanie także innych mechanizmów optymalizacyjnych, takich jak wyznaczanie drzewa rozpinającego o największej liczbie krawędzi, minimalnej średnicy, najmniejszej liczbie liści lub minimalnej dylatacji. Krawędź jest trasą między dwoma węzłami, obliczoną przez algorytm drzewa. Liśćmi są węzły końcowe gałęzi. Średnica określa liczbę przełączników, które występują na trasie łączącej dwa przełączniki sieci. Dylatacja reprezentuje różnicę między najkrótszą ścieżką między dwoma węzłami drzewa a ścieżką wyliczoną przez algorytm drzewa rozpinającego.

**Hierarchia węzeł-most**

Celem administratora sieci jest utworzenie połączeń, w których nie będą występowały pętle, ale przy jednoczesnym zapewnieniu połączeń nadmiarowych, które umożliwią wymianę danych nawet w przypadku awarii jednego z węzłów lub połączeń. W praktycznych rozwiązaniach zamiast ważonych krawędzi stosuje się pojęcie ścieżek o najniższym koszcie. Aby wyznaczyć ścieżki o najniższym koszcie, definiowane są dwa parametry:

* priorytet węzła,
* identyfikator węzła.

Koszt (waga) węzła jest wyznaczany na podstawie obydwu wymienionych parametrów. W protokole STP jako pierwszy (ważniejszy) jest uwzględniany priorytet węzła. Węzeł o najniższej wartości priorytetu jest węzłem o najwyższym priorytecie — poprzedzającym w analizie wszystkie pozostałe węzły. W przypadku węzłów o jednakowej wartości priorytetu porównywane są identyfikatory węzłów, czyli adresy MAC. Adres MAC o najniższej wartości wyznacza pierwszeństwo węzła wśród innych węzłów o tym samym priorytecie. Aby wskazać główny węzeł (most główny), należy ustawić niższą wartość priorytetu niż w innych urządzeniach obsługujących protokół STP. Domyślnie wartość priorytetu przełączników i routerów Cisco wynosi 32 768.

Algorytm STP wyznacza takie trasy przez system, które zapewniają przesyłanie komunikatów z dowolnego punktu końcowego do węzła głównego wzdłuż ścieżek o najniższym koszcie. Koszt ścieżki oblicza się jako sumę kosztów poszczególnych segmentów na trasie. Ponieważ każdy most w systemie ma konfigurowalną wartość priorytetu, protokół STP może zmienić ścieżki o najniższym koszcie zależnie od bieżącego stanu sieci. Podczas obliczania ścieżek o najniższym koszcie uwzględnia się dwie poniższe zasady:

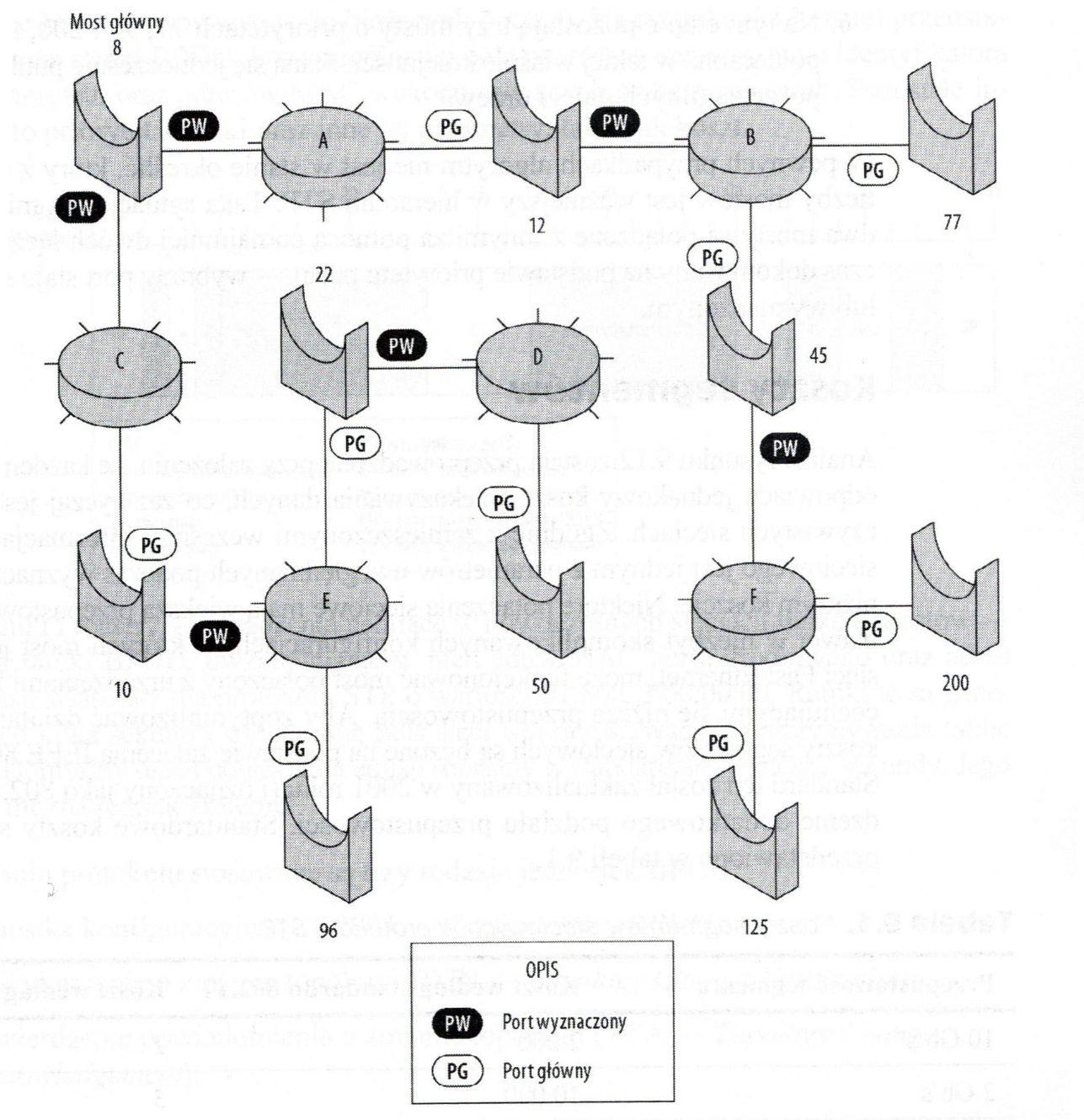
* musi zostać obliczona ścieżka o najniższym koszcie prowadząca z każdego węzła,
* musi zostać obliczona ścieżka o najniższym koszcie prowadząca do każdego segmentu sieci.

Port przełącznika lub routera, przez który wiedzie trasa o najniższym koszcie do węzła głównego, jest nazywany portem głównym (ang. root port). Port występujący na trasie o najniższym koszcie, ale prowadzący do segmentu sieciowego, jest nazywany portem wyznaczonym (ang. designated port). Na potrzeby tej analizy można przyjąć, że segment sieciowy składa się z węzłów przyłączonych do wspólnej warstwy fizycznej i wykorzystujących jednakowy model zabezpieczeń. W tym rozumieniu dwie podsieci jednej sieci LAN będą stanowiły dwa segmenty, podobnie jak dwie grupy robocze (lub domeny).

Gdy algorytm STP określi porty główne i wyznaczone, wszelkie pozostałe porty zostają zablokowane. Często zdarza się sytuacja, w której dwie ścieżki (lub większa liczba ścieżek) z mostu do węzła głównego mają jednakowy (najniższy) koszt. Wówczas wybierana jest trasa wiodąca przez most o niższym identyfikatorze. Port tego mostu staje się portem głównym. Jeśli do danego segmentu są przyłączone dwa mosty (lub większa liczba mostów), które dysponują ścieżkami o najniższym koszcie prowadzącymi do mostu głównego, portem wyznaczonym staje się port tego mostu, który jest połączony z mostem o najniższej wartości identyfikatora.

Na rysunku 9.12 została przedstawiona sieć z uruchomionym protokołem drzewa rozpinającego. Aby ułatwić analizę rysunku, przyjęto założenie, że każdemu segmentowi sieci przypisano ten sam koszt jednostkowy. Uzyskanie konfiguracji pokazanej na rysunku 9.12 zostało poprzedzone poniższą analizą.

1. Most o wartości priorytetu 8 ma najwyższy priorytet (najniższą wartość), więc zostaje wybrany mostem głównym. Most główny niekoniecznie musi być urządzeniem o największej wydajności. Zazwyczaj jest to centralny komponent sieci. Można



przyjąć zasadę, że węzłem głównym powinno być urządzenie najrzadziej przenoszone lub modyfikowane. Z tego względu węzłami głównymi są z reguły przełączniki rdzeniowe. Warto zauważyć, że most główny jest jedynym mostem w sieci, który nie ma portu głównego.

2. Od mostu głównego odchodzą dwie ścieżki do węzłów o priorytetach 10 i 12. Ponieważ wartość 10 jest najwyższym priorytetem, most o tym priorytecie łączy się z następnym mostem o najwyższym priorytecie (o wartości 22).

3. Spośród dwóch niepodłączonych węzłów o wartościach 12 i 22 węzeł o priorytecie 12 ma pierwszeństwo. Most ten może utworzyć połączenie pozbawione pętli z mostem o priorytecie 45.

4. Na tym etapie przyłączenia nie są poprowadzone do mostów o priorytetach 22

i 45. Zatem most o priorytecie 22 zostanie połączony z węzłem o kolejnym najwyższym priorytecie (50).

5. Z dwóch pozostałych priorytetów, tj. 45 i 50, pierwszeństwo ma 45. Most o tym priorytecie jest więc łączony z węzłem o priorytecie 125.

6. Na tym etapie pozostajątrzy mosty o priorytetach 77, 96 i 200, które zostaną podłączone w takiej właśnie kolejności. Staną się jednocześnie punktami końcowymi poszczególnych gałęzi drzewa.

W pewnych przypadkach algorytm nie jest w stanie określić, który z dwóch lub większej liczby mostów jest ważniejszy w hierarchii STP. Taka sytuacja ma miejsce wówczas, gdy dwa mosty są połączone z innymi za pomocą co najmniej dwóch łączy. Wybór jest wówczas dokonywany na podstawie priorytetu portu — wybrany port staje się portem głównym lub wyznaczonym.

**Koszty segmentów**

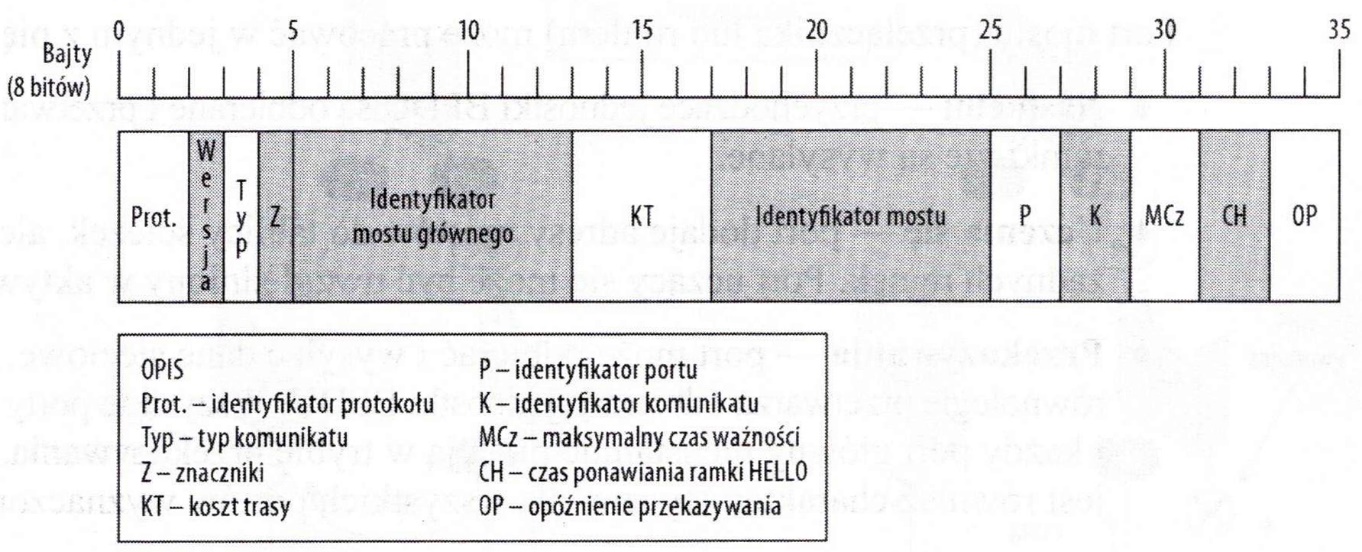
Analiza rysunku poprzednim została przeprowadzona przy założeniu, że każdemu segmentowi sieci odpowiada jednakowy koszt przekazywania danych, co zazwyczaj jest nieprawdą w rzeczywistych sieciach. Zgodnie z zamieszczonymi wcześniej informacjami koszt segmentu sieciowego jest jednym z parametrów uwzględnianych podczas wyznaczania ścieżki o najniższym koszcie. Niektóre połączenia sieciowe mają większą przepustowość, inne mniejszą. Nawet w niezbyt skomplikowanych konfiguracjach, w których most jest przyłączony do sieci Fast Ethernet, może funkcjonować most połączony z urządzeniami bezprzewodowymi, cechującymi się niższą przepustowością. Aby zoptymalizować działanie protokołu STP, koszty segmentów sieciowych są liczone na podstawie zalecenia IEEE 802.ID z 1998 roku. Standard ten został zaktualizowany w 2001 roku (i oznaczony jako 802.1T) przez wprowadzenie dodatkowego podziału przepustowości. Standardowe koszty segmentów zostały przedstawione w tabeli.

Obraz zawierający stół

Opis wygenerowany automatycznie

**Dynamiczna optymalizacja**

W kolejnym z prezentowanych przykładów został przedstawiony stan sieci po uwzględnieniu przez protokół STP wszystkich priorytetów mostów. Ma on na celu zilustrowanie funkcji, która istotnie podnosi wartość protokołu STP — zdolności do adaptacji do zmieniającej się topologii sieci. Jednym ze sposobów wykrywania zmian w sieci jest użycie protokołu mostów, który rozsyła specjalne ramki, nazywane jednostkami danych protokołu mostów (BPDU — Bridge Protocol Data Units), zawierające informacje o koszcie segmentu oraz identyfikatorach dostępnych węzłów. Dzięki aktualizacji informacji trasa do mostu głównego może być dostosowywana do bieżących potrzeb. Na rysunku



został przedstawiony format ramki BPDU. Poszczególne jej pola przenoszą dane na temat identyfikatora mostu, priorytetu oraz adresów MAC wykorzystywanych przez przełącznik. Pozostałe informacje to priorytet ścieżki oraz inne jej parametry, takie jak koszt.

Przełączniki i routery — urządzenia pełniące w nowoczesnych sieciach funkcję mostów — rozsyłają ramki BPDU, uwzględniając w nich adres MAC portu źródłowego oraz adres multiemisji właściwy dla protokołu STP o wartości 01:80:C2:00:00:00. Ramki te są generowane co kilka sekund i wyznaczają puls sieci wykorzystywany do utrzymywania tablic ścieżek. Domyślny okres ponawiania emisji (opisany w standardzie) wynosi 2 sekundy. Jego wartość można jednak zmienić.

W działaniu protokołu stosowane są trzy rodzaje jednostek BPDU:

* jednostka konfiguracyjna (CBPDU — Configuration BPDU),
* powiadomienie o zmianie topologii (TCN — Topology Change Notification),
* potwierdzenie powiadomienia o zmianie topologii (TCA — Topology Change Acknowledgement).

Za każdym razem, gdy nowe urządzenie jest dodawane do sieci, jego port jest wprowadzany w stan nasłuchu, w trakcie którego urządzenie odbiera jednostki BPDU i poznaje konfigurację sieci. Urządzeniami tymi mogą być dowolne punkty końcowe wyposażone w interfejs sieciowy. Domyślny czas nasłuchiwania wynosi 15 sekund. Po jego upływie urządzenie przechodzi w stan uczenia się, który również trwa 15 sekund. Całkowity czas nasłuchiwania i uczenia się jest parametrem konfigurowalnym, znanym jako opóźnienie przekazywania (ang. forward delay). Dzięki opóźnieniu włączane do sieci urządzenia mają czas, aby odebrać informacje z węzła głównego. Gdy urządzeniem tym jest komputer, serwer lub drukarka (jednostka niepełniąca funkcji mostu), wraz z zakończeniem procesu uczenia się port przechodzi do trybu przekazywania i rozpoczyna nadawanie jednostek BPDU.

Gdy do portu działającego mostu dodawany jest nowy most, procedura uruchamiania portu jest nieco inna. Każdy nowy przełącznik lub router może wprowadzić pętle do topologii sieci, więc jego port pozostaje w trybie blokowania aż do zakończenia cyklu nasłuchu i uczenia się. Port ten wysyła również ramki TCN do głównego mostu. Po odebraniu powiadomienia TCN most główny odsyła potwierdzenie za pomocą komunikatu TCA, w którym określa status nowego portu. Od tego momentu nowy port wysyła jednostki BPDU w standardowych odstępach czasu, aby wszystkie pozostałe węzły mogły zaktualizować odpowiednio swoje tablice tras. Most główny modyfikuje standardową jednostkę BPDU w taki sposób, aby było wiadomo, że sieć jest w trakcie zmian, a następnie przesyła informację o zmianie do wszystkich węzłów. Każdy z mostów modyfikuje tablicę tras i potwierdza wykonanie polecenia.

Port mostu (przełącznika lub routera) może pracować w jednym z pięciu trybów:

* **Nasłuchu** — przychodzące jednostki BPDU są odbierane i przetwarzane, ale żadne ramki nie są wysyłane.
* **Uczenia się** — port dodaje adresy mostów do tablicy ścieżek, ale nie przekazuje żadnych ramek. Port uczący się może być uwzględniony w aktywnej topologii.
* **Przekazywania** — port może odbierać i wysyłać dane sieciowe. Proces STP równolegle przetwarza odbierane jednostki BPDU. Wszystkie porty mostu głównego i każdy port główny nieustannie pracują w trybie przekazywania. Ten tryb pracy jest również charakterystyczny dla wszystkich portów wyznaczonych.
* **Blokowania** — ustawienie portu nie pozwala na wysyłanie ani odbieranie danych. Niemniej odbierane są jednostki BPDU, a port może zmienić stan w razie konieczności. Każdy port mostu, który obsługuje połączenie z innym mostem i nie jest portem głównym lub wyznaczonym, musi być zablokowany.
* **Wyłączenia** — port może być wyłączony programowo (na przykład w wyniku wykonania polecenia SMTP), ale nie w ramach mechanizmu STP.

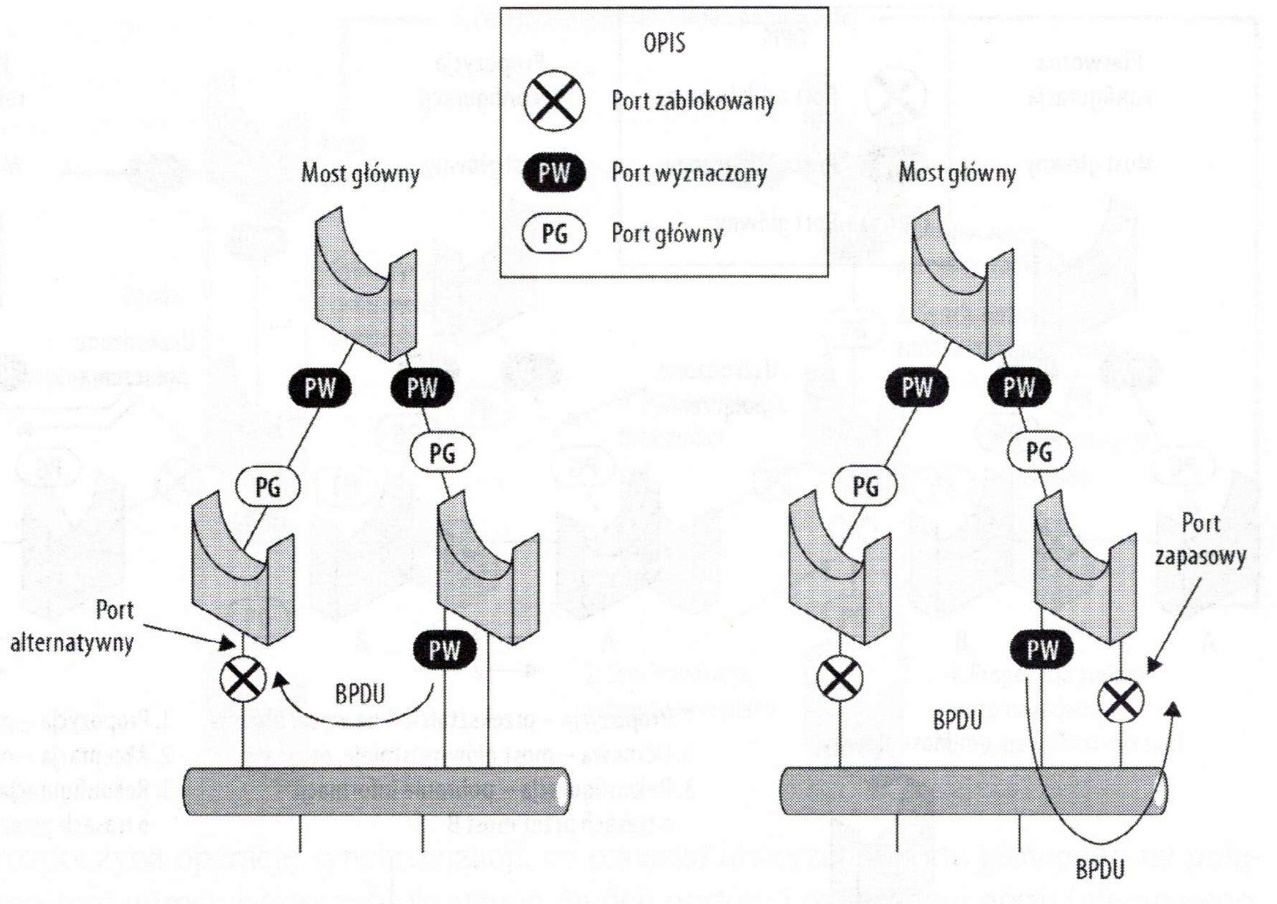
**Szybki protokół drzewa rozpinającego**

W przypadku zastosowania klasycznego protokołu STP reakcja na zmianę topologii może zająć mostom nawet jedną minutę. W roku 1995 takie rozwiązanie wydawało się wystarczające. Organizacja IEEE zdefiniowała szybszą odmianę protokołu STP — RSTP (Rapid Spanning Tree Protocol). Standard otrzymał oznaczenie 802.Iw i został opublikowany w 1998 roku. W roku 2004 organizacja IEEE połączyła zalecenia 802.ID, 802.11-2001 oraz 802.Iw w jeden standard 802.1D-2004. Wiele zmian uwzględnionych w protokole RSTP stało się częścią implementacji STP w przełącznikach ethemetowych firmy Cisco.

Działanie RSTP bazuje na mechanizmach STP, ale z istotnymi zmianami w porównaniu z pierwotnym protokołem. Modyfikacje umożliwiły skrócenie czasu rekonfiguracji do okresu poniżej dwóch sekund w przypadku awarii węzła głównego.

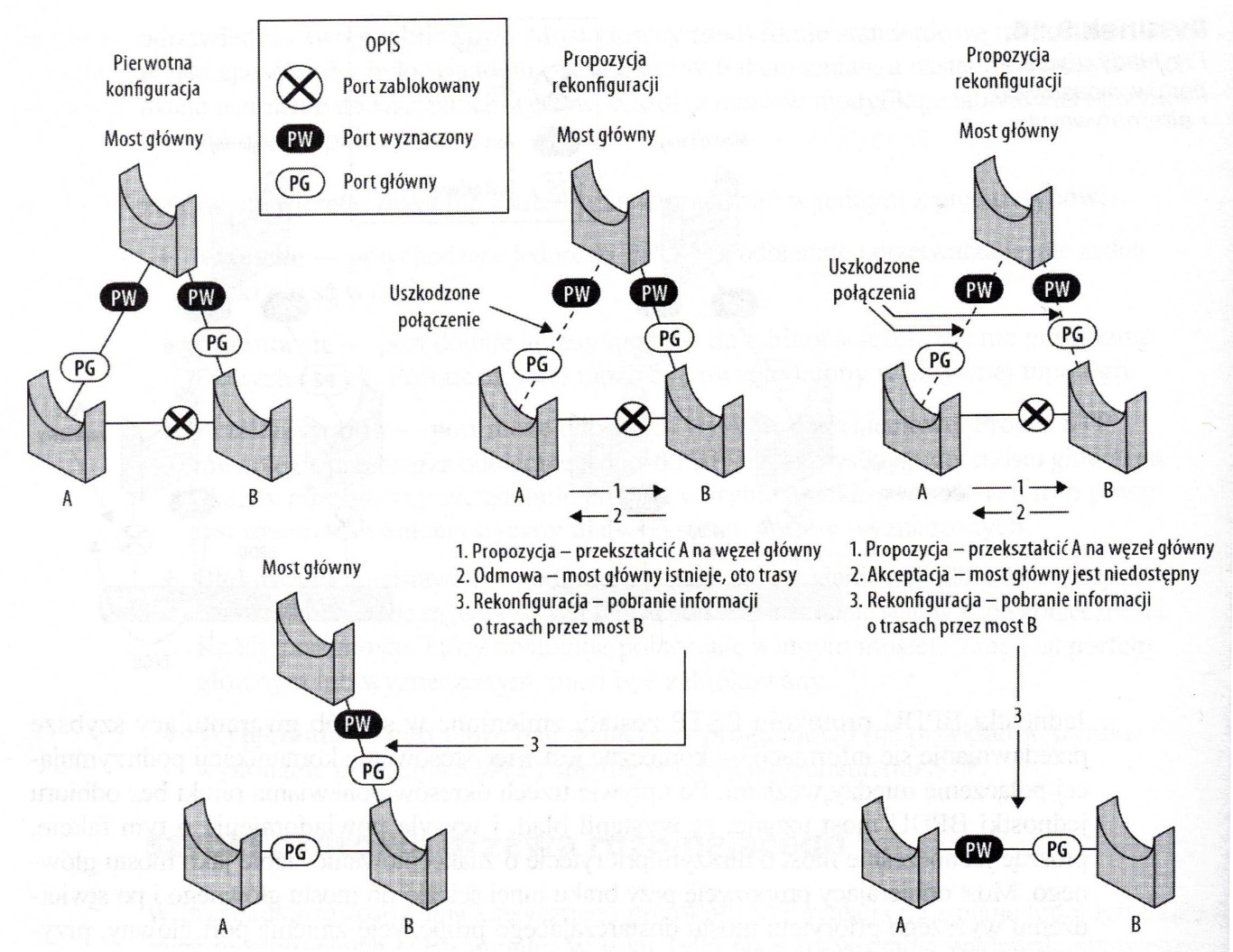
*Włączenie więcej niż jednego wariantu protokołu STP może doprowadzić do nieprzewidywalnego zachowania sieci.*

W RSTP blokowane porty są podzielone na dwie kategorie — porty alternatywne i porty zapasowe. Port alternatywny odbiera jednostki BPDU od innego mostu o wyższym priorytecie. Port zapasowy otrzymuje jednostki BPDU od tego samego mostu co inny port danego mostu. Rozróżnienie to pozwala na znacznie szybsze użycie alternatywnych ścieżek do mostu głównego po uszkodzeniu portu głównego. Port zapasowy zapewnia nadmiarowe połączenie z tym samym segmentem sieci, ale nie gwarantuje połączenia z mostem głównym. Pozostałe rozwiązania RSTP pokrywają się z funkcjami STP. Przykład zastosowania portów alternatywnych i zapasowych został przedstawiony na rysunku.



Jednostki BPDU protokołu RSTP zostały zmienione w sposób gwarantujący szybsze przedawnianie się informacji -— konieczne jest więc stosowanie komunikacji podtrzymującej połączenie między węzłami. Po upływie trzech okresów ponawiania ramki bez odbioru jednostki BPDU most uznaje, że wystąpił błąd, i wysyła powiadomienie o tym fakcie, prosząc jednocześnie most o niższym priorytecie o zaakceptowanie siebie jako mostu głównego. Most odbierający propozycję przy braku innej ścieżki do mostu głównego i po stwierdzeniu wyższego priorytetu mostu dostarczającego propozycję zmienia port główny, przypisując tę funkcję do portu połączonego z nadawcą propozycji. Jeśli jednak most odbierający propozycję dysponuje poprawną ścieżką do pierwotnego mostu głównego, wysyła jednostkę BPDU do mostu, który wygenerował propozycję, informując ten most o statusie oryginalnego mostu głównego i powodując aktualizację tras oraz rekonfigurację procesu STP w tym węźle.

Na rysunku



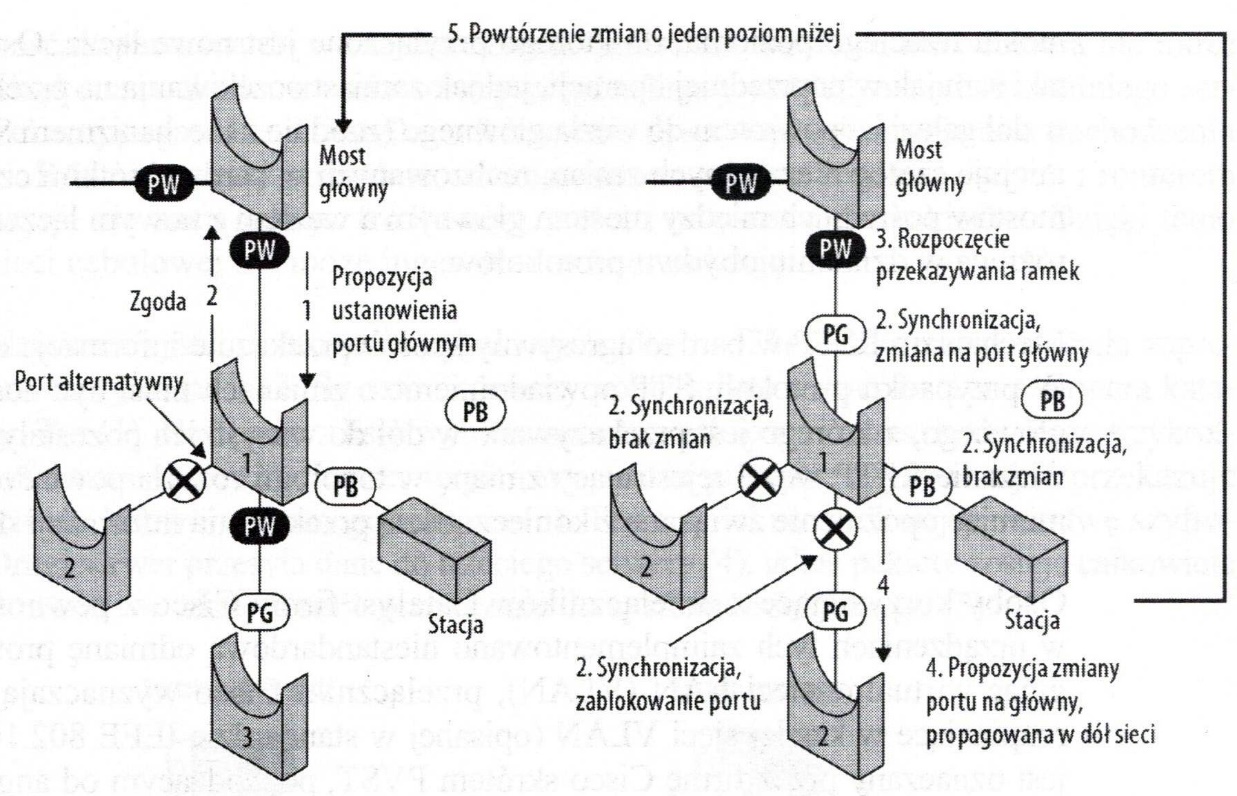
zaprezentowana została idea proponowania zmiany mostu głównego oraz operacja wyboru łącza zapasowego. W lewym górnym narożniku widoczna jest pierwotna konfiguracja. Gdy łącze zostaje uszkodzone (jak na środkowym górnym przykładzie), węzeł A przekazuje do węzła B propozycję przekształcenia siebie na węzeł główny. Ponieważ węzeł B wciąż dysponuje połączeniem z mostem głównym, propozycja zostaje odrzucona, a do węzła A są przesyłane informacje o ścieżkach węzła B. Przełączenie tras kończy się w chwili zmiany konfiguracji w węźle A (pokazanej w lewej dolnej części rysunku), która wyznacza trasę z węzła A przez węzeł B do mostu głównego.

Gdy następuje uszkodzenie dwóch połączeń, odpowiedź na propozycję jest inna. Po dostarczeniu przez węzeł A propozycji do węzła B most B akceptuje zmianę węzła głównego, ponieważ nie ma już kontaktu z dotychczasowym węzłem głównym. Wynik operacji został przedstawiony w prawym dolnym narożniku rysunku. Węzeł A ustawia jeden z portów jako port wyznaczony, a węzeł B przekształca port łącza w port główny. W wyniku przełączenia istnieje bezpośrednie łącze między mostami A i B.

W sieci pracującej pod kontrolą protokołu RSTP żaden port przyłączony do stacji końcowej nie może utworzyć pętli, ponieważ jednostki końcowe są z założenia jednoportowe. Wszystkie porty tego typu są oznaczane jako porty brzegowe i przełączane w tryb przekazywania bez konieczności uruchamiania trybu nasłuchiwania i uczenia się. Porty brzegowe pozostają takimi nawet wówczas, gdy protokół RSTP przelicza i modyfikuje topologię drzewa rozpinającego. Niemniej gdy tylko port odbierze ramkę BPDU, zostanie natychmiast przekształcony na port drzewa.

Natychmiastowe przejście do stanu przekazywania może być wykonane również w przypadku portów, które obsługują połączenia punkt-punkt. Porty działające w trybie pełnego dupleksu są uznawane za elementy połączeń punkt-punkt. Z kolei połączenia półdupleksowe są realizowane przez porty określane jako współdzielone. Ponieważ niemal wszystkie nowoczesne przełączniki wykorzystują pełnodupleksowy tryb pracy portów, protokół RSTP może w bardzo krótkim czasie przełączyć je w tryb przekazywania. Szybka zmiana stanu jest możliwa dzięki mechanizmowi propozycji-akceptacji, który propaguje przez sieć, zmieniając kolejno stan portów na każdym łączu.

Technika szybkiej zmiany stanu portu w protokole RSTP została zilustrowana na rysunku.



Propozycja wyznaczenia portu głównego (oznaczona numerem 1) zostaje zaakceptowana (w odpowiedzi o numerze 2), tak jak to zostało pokazane na lewym przykładzie. Przełącznik numer 2 rozpoczyna operację synchronizacji, co oznacza utworzenie portu głównego na połączeniu z mostem numer l oraz zablokowanie dwóch portów i oznaczenie portu brzegowego. Propozycja wyznaczenia portu głównego jest wysyłana dalej w dół sieci, bez konieczności przekazywania informacji o tej zmianie do mostu głównego. Dzięki temu zadanie jest wykonywane w bardzo krótkim czasie.

Protokół RSTP w nieco inny sposób niż STP obsługuje także wprowadzanie nowych łączy. W klasycznej konfiguracji sieci dodanie połączenia między mostem głównym a węzłem trzeciego poziomu (lub niższego poziomu hierarchii) prowadzi do powstania pętli. Mechanizm STP wykrywa tę pętlę i blokuje porty mostu głównego oraz węzła drugiego poziomu, przełączając te węzły w tryb nasłuchiwania. Dodatkowo wyłącza zainstalowane łącze do czasu rozpoznania nowej topologii. Węzeł drugiego poziomu odbiera informacje bezpo-średnio z mostu głównego i wymienia z nim dane. Następnie przesyła jednostki BPDU w dół sieci do węzła trzeciego poziomu (do którego zostało przyłączone nowe łącze), co powoduje zablokowanie portu prowadzącego do węzła drugiego poziomu. Węzeł trzeciego poziomu przekazuje jednostkę BPDU z powrotem do węzła głównego. Po odebraniu komunikatu BPDU węzeł główny akceptuje nową topologię, uruchamia łącze i podtrzymuje blokadę na porcie prowadzącym do mostu, który wcześniej był węzłem drugiego poziomu. Węzeł drugiego poziomu po dodaniu nowego łącza staje się mostem trzeciego poziomu. Wadą opisanego rozwiązania jest to, że wprowadzenie zmian trwa dwukrotnie dłużej, niż wynosi opóźnienie przekazywania.

Ten sam problem jest rozwiązywany przez protokół RSTP w nieco inny sposób. Mechanizm RSTP wykrywa nowe łącze i blokuje porty między mostem głównym a węzłem drugiego poziomu (podobnie jak w poprzednim przypadku). Węzeł główny wysyła propozycję rekonfiguracji do węzła drugiego poziomu, co powoduje zablokowanie wszystkich portów wyznaczonych w moście drugiego poziomu — działanie to nazywa się operacją synchronizacji. Węzeł drugiego poziomu przekazuje do węzła głównego żądanie odblokowania portu i ustawienia go w trybie przekazywania. Zmiana stanu jest realizowana natychmiast. Na tym etapie procedury blokady są ustanowione o jeden poziom niżej od połączenia mostu głównego z mostem drugiego poziomu. Procedura się powtarza aż do zablokowania portów

mostu trzeciego poziomu, do którego przyłączone jest nowe łącze. Ostateczny wynik jest taki sam jak w poprzedniej operacji, jednak zamiast oczekiwania na przekazanie komunikatu w dół gałęzi i z powrotem do węzła głównego (zgodnie z mechanizmem STP) protokół RSTP inicjuje szereg niezależnych zmian, realizowanych w bardzo krótkim czasie. Im więcej jest mostów pośrednich między mostem głównym a węzłem z nowym łączem, tym większa jest różnica w działaniu obydwu protokołów.

Mechanizm RSTP w bardzo agresywny sposób przekazuje informacje o zmianie topologii. W przypadku protokołu STP powiadomienie o zmianach musi być dostarczone do mostu głównego, z którego jest przekazywane w dół do wszystkich pozostałych węzłów. W rozwiązaniu RSTP węzeł rejestrujący zmianę w topologii rozsyła powiadomienia w sieci, eliminując opóźnienie związane z koniecznością przekazania informacji do węzła głównego.

Osoby korzystające z przełączników Catalyst firmy Cisco z pewnością zauważyły, że w urządzeniach tych zaimplementowano niestandardową odmianę protokołu STP. Obsługując wirtualne sieci LAN (VLAN), przełączniki Cisco wyznaczają oddzielne drzewo rozpinające w każdej sieci VLAN (opisanej w standardzie IEEE 802.1Q). Mechanizm ten jest oznaczany przez firmę Cisco skrótem PVST, pochodzącym od angielskich słów określających drzewo rozpinające w każdym VLAN-ie (Per-VLAN Spanning Tree). W przełącznikach Catalyst jest również implementowana wersja protokołu PVST+, wykorzystująca dodatkowo mechanizm tunelowania ruchu. Także organizacja IEEE opracowała standard protokołu wielokrotnych drzew rozpinających (MSTP — Multiple Spanning Tree Protocol) — zdefiniowany w dokumentach IEEE 802.1s/Q — który umożliwia tworzenie niezależnych drzew rozpinających w każdej grupie sieci VLAN. Wersja mechanizmu MSTP imple-mentowana w urządzeniach Cisco jest nazywana protokołem drzewa rozpinającego o wielu instancjach (MISTP — Multiple Instances Spanning Tree Protocol). Z kolei inne rozwiązanie firmy Cisco — R-PVST —jest połączeniem protokołów RSTP i PVST zapewniającym utworzenie jednego drzewa w każdym VLAN-ie.

**Routery cebulowe**

Wszyscy widzieliśmy takie sceny w filmach. Źli chłopcy wysyłają komunikat do dobrych chłopców. Dobrzy chłopcy próbują zlokalizować serwer w Nowym Jorku, ale gdy są gotowi do schwytania złych chłopców, nadchodzi kolejna wiadomość. Tym razem z Singapuru, a potem następna z Berlina. Każda przesyłka jest wysyłana z innego serwera, co sprawia, że ustalenie lokalizacji nadawcy jest niemożliwe. Anonimowa komunikacja stanowi właśnie cel działania routerów cebulowych.

W systemie routerów cebulowych komunikaty sieciowe podlegają wielokrotnemu (trzykrotnemu) szyfrowaniu w jednostce źródłowej i są wysyłane w losowy sposób przez sieć routerów IP (serwery cebulowe). Każdy router usuwa jedną warstwę szyfrowania — tak, jakby się obierało cebulę. Serwer wejściowy jest wybierany losowo spośród mniejszego zbioru serwerów nazywanych strażnikami wejść. Każdy z trzech serwerów —jeden losowo wybrany spośród strażników wejść i dwa inne losowo wskazane spośród ogólnoświatowej sieci serwerów cebulowych — wykorzystuje własny klucz prywatny do usunięcia jednej warstwy szyfrowania.

Wiadomość dostarczana do węzła docelowego jest rozszyfrowana, ale odbiorca nie może ustalić, skąd ona pochodzi ani jaką trasę przebyła. Znany jest jedynie adres ostatniego serwera przekazującego dane. Nie tylko odbiorca nie dysponuje informacjami o pochodzeniu przesyłki. Również węzły pośrednie, umiejscowione między źródłem danych i routerem wyjściowym, nie znają adresu źródłowego, treści i przeznaczenia pakietów. Dzięki temu nikt w sieci cebulowej nie może ingerować w komunikację.

Zasada działania systemu routerów cebulowych (Tor — The onion router) została zaprezentowana na rysunku.

Obraz zawierający mapa

Opis wygenerowany automatycznie

System nadawczy pobiera listę serwerów Tor z serwera katalogowego Tor (1). Następnie z listy wybiera serwer wejściowy i przesyła do niego trzykrotnie zaszyfrowane dane (2). Serwer wejściowy usuwa pierwsze szyfrowanie i przekazuje informacje do losowo wybranego serwera (3). Ten z kolei usuwa kolejną warstwę szyfrowania. Drugi serwer przesyła dane do trzeciego serwera (4), gdzie pakiety zostają całkowicie rozszyfrowane i w takiej formie przekazane do systemu odbiorczego (5).

Celem sieci Tor jest uchronienie użytkowników przed atakami wynikającymi z analizy ruchu. Ataki tego typu polegają na analizowaniu grupy komunikatów po obydwu stronach trasy pakietów w celu ustalenia, które routery sieci są wykorzystywane, i w celu wyznaczenia pewnych wzorców ruchu. Im większą liczbę pakietów uda się przechwycić, tym lepiej. Format komunikatów (to, czy są zaszyfrowane, czy nie) nie ma znaczenia. Celem jest przechwycenie określonej wiadomości. Rozszyfrować można ją później. Często celem bywa również przerwanie komunikacji.

Jeden z rodzajów ataków wymagających przejęcia komunikacji służy do utworzenia połączenia SSH z jednostką atakowaną i określenia zależności czasowych między zwracanymi komunikatami. Czas między poszczególnymi znakami jest analizowany pod względem statystycznym z wykorzystaniem modelu Markowa, na którego podstawie można wydedukować hasło. Systemy Tor znacznie utrudniają przeprowadzenie podobnych ataków, ale nie uniemożliwiają ich. Trzeba również pamiętać, że ruch opuszczający system routerów cebulowych nie podlega szyfrowaniu i może zostać spreparowany, podobnie jak inne wiadomości.

**Sieci Tor**

Routing cebulowy jest technologią, która obecnie została wdrożona jedynie w ramach projektu Tor. Jak nietrudno się domyślić, bezpieczna komunikacja jest kluczowa dla działań wojskowych. Z tego względu początkowo rozwojem routingu cebulowego zajmowało się laboratorium badawcze marynarki wojennej Stanów Zjednoczonych. Projekt drugiej generacji o nazwie The Onion Router (Tor) został zapoczątkowany przez organizację Electronic Frontier Foundation (http://www.eff.org) w 2004 roku, a w roku 2006 stał się projektem otwartym o nazwie The Tor Project (http://www.torproject.org), prowadzonym przez organizację non profit.

Choć mechanizm routingu cebulowego jest rozwiązaniem, które może zostać zaimplementowane przez każdego, sieć Tor jest jedyną, w której praktycznie je wdrożono. Obecnie lista ogólnoświatowych serwerów Tor zawiera ponad 1800 pozycji — chociaż liczba jednostek działających w danym czasie jest zmienna.

**Jednostki klienckie Tor**

Ruch Tor jest generowany przez cebulowe serwery proxy, instalowane w systemie nadawczym. Komponent proxy komunikuje się z usługą katalogową Tor i ustala wirtualny obwód w sieci. Oprogramowanie to jest interfejsem mechanizmu SOCKS. Zatem aplikacje, które tworzą gniazda, mogą skorzystać z komponentu proxy do przesyłania ruchu w sieci Tor w ramach wirtualnego obwodu. Wiadomości są multipleksowane, a następnie wysyłane ustaloną trasą. Wśród aplikacji zdolnych do korzystania z mechanizmu SOCKS są przeglądarki, komunikatory internetowe oraz aplikacje klientów IRC.

Aby w pełni skonfigurować komponent proxy sieci Tor, potrzebne są następujące aplikacje:

* **Privoxy** (http://www.privoxy.org). Aplikacja Privoxy jest filtrującym, pozbawionym buforowania komponentem proxy dla sieci Web. Ułatwia zachowanie prywatności, zarządza danymi cookies, modyfikuje treść stron WWW, przechwytując wyskakujące okna, banery itp. Jest to program darmowy, bazujący na kodzie komponentu Internet Junkbuster. Obecna jego wersja to 3.0.17.
* **Tor** (http://www.torproject.org). Klient Tor zapewnia obsługę protokołu Tor oraz innych komponentów, które umożliwiają korzystanie z sieci Tor.
* **Torbutton** (https://www.torproject.org/torbutton). Program instaluje w przeglądarce Firefox przycisk, który umożliwia włączanie i wyłączanie trybu pracy z siecią Tor.
* **Vidalia** (http://www.torproject.org/vidalia). Graficzny interfejs, który umożliwia monitorowanie i zarządzanie konfiguracją Tor.

Programiści skupieni w ramach projektu Tor ułatwili instalację wymienionych powyżej komponentów i udostępnili pojedynczy instalator ze wszystkimi niezbędnymi elementami. Aby pobrać program kliencki Tor, wystarczy skorzystać z jednej z poniższych stron WWW:

Instalator dla systemu Windows

— https ://www. torproject. org/docs/tor-doc-windows. html. en.

Instalator dla systemu Mac OS X

— https: //ww. torproject. org/docs/tor-doc-osx. html. en.

Instalator dla systemów Linux, BSD, UNIX

— https ://www. torproject. org/docs/tor-doc-unix. html. en.

Po zainstalowaniu klienta Tor należy sprawdzić, czy działa poprawnie. Jeden ze sposobów polega na odwołaniu się do ukrytego serwera (opisanego w kolejnym punkcie) w sieci Tor. Wystarczy wpisać w przeglądarce adres http://duskgytldkx1uqc6.onion i po transferze danych trwającym około minuty sieć Tor powinna odwzorować adres.

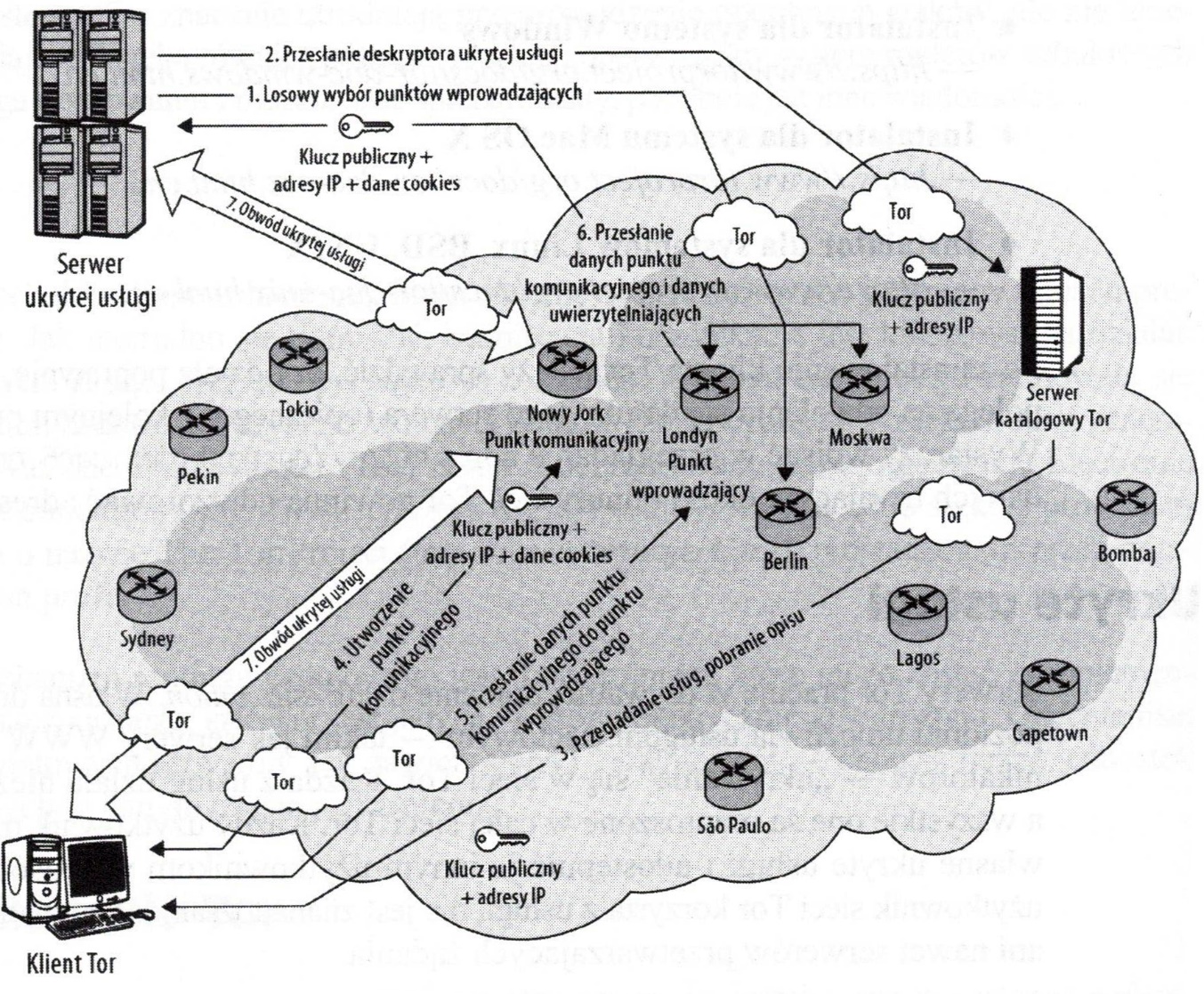
**Ukryte usługi**

Serwery Tor pracują w prywatnej domenie o sufiksie .onion. Własna domena najwyższego poziomu umożliwia usługom sieciowym — takim jak serwery WWW lub serwery komunikatorów — „ukrywanie” się w sieci Tor. Każda z usług działa niezależnie od innych, a wszystkie one są rozproszone w całej sieci Tor. Każdy użytkownik może skonfigurować własne ukryte usługi i udostępnić je innym użytkownikom w sposób anonimowy. Gdy użytkownik sieci Tor korzysta z usługi, nie jest znana tożsamość ani nadawcy, ani odbiorcy, ani nawet serwerów przetwarzających żądania.

Aby utworzyć ukrytą usługę, wystarczy działający komponent kliencki Tor oraz serwer WWW. Programiści Tor polecają pakiety Samant lub Apache dla systemu Windows bądź thttpd dla systemów UNIX i Mac OS X. Serwer WWW musi być skojarzony z portem 5222 jednostki lokalnej. Przypisanie niestandardowego portu powoduje, że system zewnętrzny nie może ustalić, czy usługa jest uruchomiona w systemie lokalnym. Serwer WWW powinien zostać uruchomiony jako niezależny proces — inny niż ewentualne wcześniejsze serwery WWW, szczególnie te, które są udostępniane w Internecie lub intranecie.

Schemat działania ukrytych usług zostałrysunku 9.18. Zainstalowana usługa powiadamia jednostki klienckie o swojej dostępności (1) za pomocą rozgłoszeń, dostarczanych za pośrednictwem specjalnego protokołu (losowo wybranymi trasami) do serwerów katalogowych Tor, które przechowują taką informację wraz z przypisanym usłudze kluczem publicznym. Wspomniane serwery przyjmują rolę punktów wprowadzających i utrzymują klucz publiczny ukrytej usługi (2). Ponieważ trasa między serwerem ukrytej usługi i punktem wprowadzającym obejmuje losowo wybierane obwody wirtualne, klient nie może skojarzyć dwóch systemów, aby poznać adres IP ukrytego serwera.

Klient Tor dowiaduje się, że usługa pochodzi z serwera katalogowego Tor (3) i tworzy punkt komunikacyjny. Następnie ustanawia połączenie z jednym z serwerów wprowadzających (4). Punkt komunikacyjny dysponuje zarówno kluczem publicznym, jak i danymi cookies, które są potrzebne do szyfrowania (lub deszyfrowania) informacji oraz do generowania informacji umożliwiających przekazanie danych z serwera usługi do klienta Tor. Gdy punkt wprowadzający przekaże informacje klienckie do serwera ukrytej usługi, tworzony jest wirtualny obwód (oznaczony cyfrą 7 i grubszą strzałką). System oddziela punkt wprowadzający od punktu komunikacyjnego, gwarantując zachowanie anonimowości klienta.



Ukryta usługa tworzy deskryptor zawierający klucz publiczny i opis usługi, a następnie podpisuje go kluczem prywatnym. Deskryptor ukrytej usługi jest przekazywany w formie zaszyfrowanej do serwera katalogowego Tor, gdzie jest replikowany i rozsyłany w sieci. Operacja ta zapewnia ukrycie usługi. Serwer katalogowy automatycznie generuje nazwę domenową <UkrytaUsługa>.onion, która może być wykorzystywana przez oprogramowanie klienckie Tor. Na tym etapie konfiguracja ukrytej usługi jest zakończona.

Gdy klient chce sprawdzić działanie usługi, pobiera jej deskryptor z serwera katalogowego. Deskryptor zawiera listę punktów wprowadzających oraz klucz publiczny usługi. Klient ustanawia połączenie z losowo wybranym serwerem Tor, żąda od niego pracy w charakterze punktu komunikacyjnego i przesyła do serwera dane cookies z jednorazowym hasłem. Do zaszyfrowania wiadomości wprowadzającej wykorzystywany jest klucz publiczny ukrytego serwera. Sama wiadomość składa się z adresu punktu komunikacyjnego oraz danych cookies z jednorazowym hasłem. Wymiana informacji jest realizowana przez sieć Tor w klasyczny sposób.

Ukryty serwer wykorzystuje informacje zawarte w wiadomości wprowadzającej do utworzenia obwodu z punktem komunikacyjnym, a następnie przesyła do niego hasło, które potwierdza połączenie z klientem Tor. Punkt komunikacyjny przesyła do klienta komunikat informujący o ustanowieniu połączenia. Dwustronna komunikacja między klientem i serwerem ukrytej usługi jest realizowana w ramach wirtualnego obwodu z wykorzystaniem punktu komunikacyjnego jako pośrednika. Obwód składa się z sześciu przekaźników, z których trzy są wybierane przez kliencki komponent obsługi obwodu wirtualnego, a trzy kolejne są wskazywane przez analogiczny komponent po stronie serwera. Wspólnym przekaźnikiem jest punkt wprowadzający.

**Bramy**

Brama sieciowa jest urządzeniem lub oprogramowaniem umożliwiającym komunikację między dwoma różnymi sieciami. Bramy tłumaczą adresy, protokoły sieciowe, a nawet dane. Czasami kupuje się je w formie odrębnych urządzeń. W innych przypadkach są instalowane jako oprogramowanie komputera, który pełni funkcję łącznika sieci. Przykładem bramy może być program, który pobiera dane z modułu składania zamówień serwisu WWW i przekazuje stosowne informacje do serwera obsługi płatności, nazywanego bramą kart kredytowych. Innym przykładem może być firewall lub serwer proxy. Adres bramy sieciowej musi zostać określony w każdym interfejsie sieci TCP/IP. Powszechnie są także stosowane bramy pocztowe i bramy stacji roboczych.

Określenie „brama” jest więc raczej terminem marketingowym i musi być postrzegane w szerszym kontekście. Pewne elementy bramy są zawarte w routerze. Nawet funkcja udostępniania połączenia internetowego jest pewnym rodzajem bramy. Tym, co wyróżnia bramy wśród innych urządzeń sieciowych (np. routerów), jest zdolność do wykorzystywania wyższych warstw modelu OSI. Bramy działają w warstwie transportowej, a często również w warstwie aplikacji, czyli najwyższej w stosie protokołów — routery mogą działać w warstwie 4., ale nigdy w warstwie 7.