ZESZYTY NAUKOWE POLITECHNIKI ŚLĄSKIEJ 1993

Seria: INFORMATYKA z. 24 Nr kol. 1222

Grzegorz KOWALCZYK

WYBRANE ZAGADNIENIA ROUTINGU W SIECIACH TCP/IP

Streszczenie. W artykule przedstawiono poglądowy opis zaimplementowanych w sieciach TCP/IP protokołów wyboru drogi pakietu w sieci. Omówione zostały protokoły RIP, SPF i Hello.

CHOSEN PROBLEMS OF ROUTING IN TCP/IP NTWORKS

Summary. The article presents some basic problems of routing in lagre TCP/IP computer networks. It presents most widly used protocols like RIP, SPF and Hello.

PROBLÈMES DE ROUTING DANS RESEAUX TYPE TCP/IP

Resumé. L'article présente une revue de protocoles du choix de chemin pour des paquets utilisès dans reseaux type TCP/IP. On discute les protocoles RIP, SPF et Hello.

1. Wstęp

Rosnąca liczba użytkowników sieci Internet na świecie oraz w Polsce (Naukowe i Akademickie Sieci Komputerowe), a także rozwój lokalnych sieci komputerowych opartych na protokole TCP/IP sprawia, że rośnie liczba zainteresowanych zasadami rządzącymi taką siecią, wśród administratorów systemów, jak i wśród bardziej dociekliwych użytkowników. Wychodząc naprzeciw zainteresowaniu w niniejszej publikacji przedstawiono opis mechanizmów tzw. *routing*'u, rozumianego jako całokształt technik software'owych

zastosowanych przy wyborze drogi datagramu od nadawcy do odbiorcy w rozległej sieci TCP/IP, której powszechnie dostępnym przykładem może być Internet.

Ze względu na niewielką objętość opracowania przyjęto, że czytelnik posiada elementarną wiedzę w zakresie sieci z komutacją pakietów oraz podstawowe wiadomości o protokole IP (struktura datagramu).

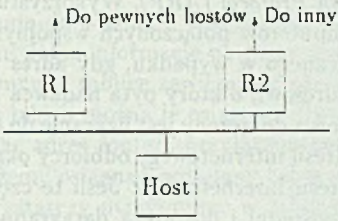
W dalszej części publikacji posłużono się terminem internet dla określenia zarówno sieci Internet, jak i nie będących jego częścią, sieci TCP/IP zawierających routery i różnorakie połączenia fizyczne.

2. Transmisja pakietów w sieci TCP/IP

Sieci TCP/IP bazują na systemie bezpołączeniowej transmisji pakietów, nie zapewniającym pewnego dostarczenia pakietu do odbiorcy (poprawność przesyłanych informacji zapewniają wyższe warstwy), a podstawową jednostką transmisji w takiej sieci jest datagram.

W systemach z komutacją pakietów pojęcie *routing* odnosi się do procesu wyboru drogi jaką będzie przesyłany pakiet, a *router*'ami nazywane są komputery dokonujące tego wyboru, łączące poszczególne fragmenty sieci.

Routing realizowany jest na kilku poziomach. Na przykład sieć rozległa, która posiada wiele fizycznych połączeń (dróg) pomiędzy routerami, odpowiada za wybór drogi pakietu od momentu wejścia pakietu do tej sieci do chwili opuszczenia jej przez pakiet. Na pozostałą część drogi (poza jej obszarem), którą przebywa pakiet, nie ma ona wpływu i droga ta może być wyznaczana z wykorzystaniem innych mechanizmów. Wybór drogi w rozległej sieci o wielu połączeniach fizycznych jest zagadnieniem złożonym. Idealny router powinien brać pod uwagę takie czynniki, jak obciążenie sieci, długość datagramu czy też sposób jego obsługi określony w polu *service type* datagramu (tzn. czy pakiet powinien być kierowany drogą o najmniejszym opóźnieniu, największej przepustowości czy o najmniejszej stopie błędów). Większość pracujących routerów bierze pod uwagę tylko niektóre z tych czynników. Rozległa sieć TCP/IP zbudowana może być z różnych sieci fizycznych połączonych routerami, które stanowią połączenia pomiędzy dwoma bądź większą liczbą takich sieci. W odróżnieniu od routerów, hosty są na ogół przyłączone do jednej sieci fizycznej (spotyka się także tzw. *multihomed hosts* przyłączone jednocześnie do więcej niż jednej sieci). W procesie routingu partycypują zarówno hosty jak i routery. Gdy dany host przystępuje do komunikowania się z innym, musi zadecydować, gdzie wysłać datagram.



Rys. 1. Host musi określić czy datagram ma być przesłany do routera R1 czy R2, ponieważ żaden z routerów nie udostępnia najlepszej drogi do wszystkich hostów w sieci

Fig. 1. Host must choose to send the datagram either to router R1 or to router R2, because no single router provides the best path to all destinations

Jak wynika z powyższego rysunku, host musi powziąć decyzję, gdzie wysłać pakiet, nawet jeśli jest fizycznie przyłączony tylko do jednej sieci. Hosty przyłączone do dwóch lub więcej sieci mogą pracować jako routery i często tak się zdarza, jeśli na danej sieci nie ma maszyny pełniącej jedynie funkcję routera.

2.1. Routing pośredni i bezpośredni

Można dokonać ogólnego podziału routingu na routing bezpośredni i pośredni. Na routingu bezpośrednim bazują wszystkie transmisje w sieciach TCP/IP. Występuje on wówczas, gdy dwie wymieniające pomiędzy sobą pakiety maszyny połączone są tym samym systemem transmisyjnym (np. wspólny Ethernet lub łącze szeregowe). Routing pośredni ma miejsce w transmisjach pomiędzy maszynami zlokalizowanych w oddzielnych fragmentach sieci i muszącymi komunikować się poprzez routery.

2.2. Transmisja datagramów pomiędzy komputerami połączonymi wspólną siecią fizyczną

Komputer wysłający datagram zamyka go w polu danych ramki fizycznej, odwzorowuje adres internetowy odbiorcy na jego adres fizyczny i wysyła ramkę fizyczną poprzez interfejs sieciowy. Odwzorowywanie adresu internetowego na adres fizyczny odbywa się w sposób zależny od rodzaju sieci fizycznej, do jakiej dany host jest przyłączony. Przykładowo w sieci token ring proNET-10, jako adresy fizyczne używane są małe liczby całkowite. Przypisanie liczby całkowitej danemu interfejsowi sieciowemu odbywa się podczas jego instalacji w komputerze, a wybór liczby należy do użytkownika bądź administratora sieci. Adresy internetowe w takiej sieci w części określającej hosta będą posiadały wpisany jego adres fizyczny. I tak host o adresie fizycznym np. 9, włączony do sieci IP o adresie 123.76.14, będzie miał adres internetowy 123.76.14.9. Bardziej złożona jest sytuacja w sieciach typu Ethernet. Każdy interfejs sieciowy ma 48-bitowy adres fizyczny przypisany mu przez producenta, nie ma więc możliwości odwzorowania go w 32-bitowy adres internetowy. By dokonać odwzorowania adresu internetowego na fizyczny, używana jest technika

znana jako *Address Resolution Protocol (ARP)*. Wykorzystuje ona fizyczne rozgłaszanie wysyłając do wszystkich komputerów połączonych wspólnym Ethernetem żądanie odesłania sieciowego adresu fizycznego w wypadku, gdy adres internetowy odbiorcy ramki rozgłoszeniowej odpowiada adresowi, o który pyta nadawca ramki.

Host rozpoznaje, czy odbiorca dołączony jest fizycznie do tego samego fragmentu sieci, poprzez porównanie części adresu internetowego odbiorcy określającej adres sieci, z odpowiednią częścią własnego adresu internetowego. Jeśli te części adresów są sobie równe, wówczas stosuje routing bezpośredni i dostarcza datagramy bezpośrednio do odbiorcy, adresując ramkę fizyczną adresem fizycznym odbiorcy datagramu.

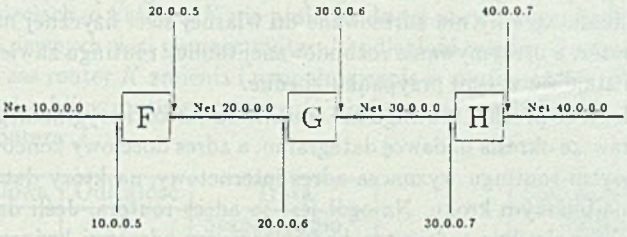
2.3. Transmisja datagramów w przypadku gdy komputery komunikujące się ze sobą nie są zlokalizowane fizycznie w tym samym fragmencie sieci

W przypadku gdy nadawca i odbiorca nie przynależą do tej samej sieci fizycznej (np. wspólny Ethernet lub TokenRing), transmitowane datagramy muszą przejść przez routery. Jeśli na sieci, do której dołączony jest host nadawcy, pracuje więcej niż jeden router, komputer nadawcy sam musi zidentyfikować router, przez który może po optymalnej drodze osiągnąć host odbiorcy. Gdy dokona wyboru routera, wysyła do niego datagram IP zapakowany w ramkę fizyczną, z fizycznym adresem routera otrzymanym sposobem opisanym w poprzednim podrozdziale (ARP). Router odpakowywuje datagram z ramki fizycznej, sprawdza docelowy adres internetowy, sprawdza, czy jest to adres sieci, do której jest on bezpośrednio przyłączony, jeśli nie – to określa, do którego z routerów pracujących na tej samej, co on, sieci fizycznej ma przesłać datagram, pakuje go w ramkę fizyczną z adresem routera i przesyła ją do niego. Jeśli host docelowy dołączony jest do tej samej sieci fizycznej, co jeden z interfejsów sieciowych routera, to router przesyła do niego datagram stosując wcześniej opisany sposób komunikowania się hostów dołączonych do tej samej sieci fizycznej. Opisana idea komunikowania się poprzez routery nosi miano routingu pośredniego. Kluczowym problemem staje się wybór optymalnej drogi pomiędzy hostami w rozległej sieci z wieloma drogami łączącymi poszczególne jej fragmenty, przy uwzględnieniu faktu, że poszczególne łącza fizyczne mogą być wykonane w różnych technologiach i posiadać drastycznie różniące się patrametry transmisyjne.

3. Organizacja routingu za pomocą tablic sterujących *(Table-Driven Routing)*

Zazwyczaj algorytmy wyboru optymalnej drogi wykorzystują tablice składowane na wszystkich maszynach (hostach i routerach), zawierające informacje o istniejących w sieci celach i drogach, po których cele te można osiągnąć. Tablice takie znajdują się w pamięci hostów i routerów, ponieważ oba rodzaje maszyn biorą udział w decydowaniu o drodze transmitowanego datagramu. Kiedykolwiek oprogramowanie realizujące routing na

hoście czy routerze przystępuje do wysłania datagramu, sprawdza w tablicy routingu, gdzie ma skierować datagram. Jakie informacje powinny znajdować się w takiej tablicy? Bardzo trudno byłoby utrzymywać tablice zawierające adresy wszystkich istniejących w sieci hostów i opisy dróg, po jakich można je osiągnąć. By wysłać datagram w poprawnym kierunku, wystarczy znać adres routera przyłączonego do tej samej sieci fizycznej, co nadawca, przez który możemy osiągnąć pożądany cel w sieci. By zidentyfikować sieć, na której cel się znajduje, wystarczy utrzymywać w tablicy tylko część adresu internetowego, określającą adres sieci, a nie cale adresy wszystkich hostów na niej. Tak więc tablica taka powinna zawierać pary (N, R), gdzie N jest częścią adresu identyfikującą sieć, a R adresem routera, poprzez który należy wysłać datagram kierowany do sieci N. Istotne jest podkreślenie faktu, że tablica zawsze wskazuje na router, który nadawca może osiągnąć poprzez jedną tylko sieć fizyczną. Jeśli na sieci fizycznej, do której dołączona jest maszyna M, pracuje tylko jeden router, to datagramy kierowane do jakiejkolwiek sieci poza własną siecią nadawcy kierowane będą do tego właśnie routera, a w tablicy routingu R we wszystkich parach (N, R) będzie jego adresem internetowym.



(a)

BY OSIĄGNĄĆ HOSTY NA SIECI NR.:	PRZEŚLIJ:
20.0.0.0	bezpośrednio
30.0.0.0	bezpośrednio
10.0.0.0	na adres 20.0.0.5
40.0.0.0	na adres 30.0.0.7

(b)

Rys. 2. (a) Przykładowa sieć składająca się z 4 sieci fizycznych i 3 routerów i (b) tablica routingu dla routera G

Fig. 2. (a) An example network with 4 physical networks and 3 routers, and (b) the routing table for router G

Powyższy rysunek przedstawia cztery sieci połączone za pomocą trzech routerów oraz tablicę routingu dla routera G. Ponieważ G połączony jest bezpośrednio do sieci 20.0.0.0 i 30.0.0.0, może osiągnąć dowolny host na tych sieciach bezpośrednio (np. w przypadku Ethernetu używając ARP, by uzyskać ich adresy fizyczne). Jeśli datagram przeznaczony jest dla hosta na sieci 40.0.0.0, to G przesyła go na adres 30.0.0.7 będący adresem routera H, który może dostarczyć datagram bezpośrednio do adresata. Router G może przesłać

datagram do 30.0.0.7 bezpośrednio, ponieważ połączony jest z routerem H wspólną siecią fizyczną.

Wybór drogi datagramu z wykorzystaniem tablic routingu pociąga za sobą kilka konsekwencji. W większości rozwiązań wszystkie pakiety kierowane do danej sieci będą przebywać taką samą drogę, niezależnie od rodzaju pakietu, nawet jeśli istnieją konkurencyjne drogi o parametrach bardziej odpowiednich dla pewnych rodzajów transmisji. Drugim efektem zastosowania takich tablic routingu jest fakt, że jedynie ostatni router na danej drodze może stwierdzić, czy docelowy host pracuje i czy jest w stanie odebrać datagram. Kolejną konsekwencją takiego sterowania drogą datagramu jest możliwość kierowania datagramu z hosta A do hosta B po innej drodze niż z B do A, co może spowodować, że np. transmisja po szybkich sieciach z A do B może zostać znacznie spowolniona powolnym nadchodzeniem potwierdzeń na wolnej drodze z B do A.

Dodatkową techniką umożliwiającą zmniejszenie zawartości tablic routingu jest wyznaczenie tzw. standardowych tras *(default routes)*. Polega to na wskazaniu jednego z bezpośrednio przyłączonych routerów jako tego, do którego kierowane będą datagramy w przypadku gdy adres sieci, do której są adresowane, nie występuje w tablicy routingu. Gdy sieć fizyczna połączona jest z pozostałą częścią sieci rozległej tylko przez jeden router, wówczas wszystkie datagramy nie adresowane do własnej sieci fizycznej należy przesłać właśnie na ten router, a utrzymywanie rozbudowanej tablicy routingu zawierającej adresy dużej liczby sieci staje się w tym przypadku zbędne.

Zauważyć należy, że protkół routingu nie zmienia zawartości oryginalnego datagramu. Adres źródłowy zawsze określa nadawcę datagramu, a adres docelowy końcowego odbiorcę datagramu. Algorytm routingu wyznacza adres internetowy, na który datagram ma zostać wysłany w najbliższym kroku. Na ogół jest to adres routera. Jeśli datagram może być dostarczony bezpośrednio, adres ten będzie tożsamy adresowi końcowego odbiorcy datagramu i nie jest umieszczany w żadnym z pól oryginalnego datagramu, jedynie jest użyty do powiązania go z adresem fizycznym, jak opisano w punkcie 2.2.

3.1. Inicjalizacja i utrzymanie tablic routingu

W początkach rozwoju rozległych sieci komputerowych bazujących na protokole TCP/IP stosowano ręczną metodę inicjalizacji i uaktualniania zawartości tablic routingu bazując na wiedzy administrujących siecią na temat topologii połączeń, lecz sposób ten z oczywistych powodów stał się niemożliwy do dalszego wykorzystywania. Obecnie zadania te wykonywane są w sposób automatyczny, przy użyciu rozproszonych algorytmów wyznaczających drogi i propagujących informację o nich w sieci.

3.1.1. Algorytm Bellmana–Forda (Vector Distance Routing)

Terminem Vector Distance określa się pewną klasę algorytmów używanych do propagowania informacji o drogach w sieci.

Każdy z routerów rozpoczyna swoją pracę z minimalnym zasobem informacji o sieci. Posiada jedynie zestaw tras dla tych sieci, do których jest bezpośrednio przyłączony. Utrzymuje listę dróg w tablicy, uszeregowanej według adresów sieci, zawierającą poza tym miary odległości do tych sieci wyrażone liczbą pośredniczących routerów oraz spo-

sób dokonania pierwszego kroku na drodze do sieci docelowej, tzn. zawiera bądź adres najbliższego routera, do którego należy przesłać datagram, bądź informację, że dana sieć osiągana jest bezpośrednio.

Zawartość przykładowej tablicy routingu w momencie inicjalizacji systemu routera może mieć postać:

Sieć docelowa	Odległość	Trasa
Sieć 1	0	bezpośrednio
Sieć 2	0	bezpośrednio

Rys. 3. Stan tablicy routingu przy starcie systemu routera z informacją o bezpośrednio dołączonych sieciach

Fig. 3. Contents of an initial routing table with an information entry for each directly connected network

Okresowo, każdy z routerów wysyła kopię swojej tablicy routingu do wszystkich routerów, z którymi połączony jest wspólnymi sieciami fizycznymi. Gdy router K odbierze kopię tablicy nadesłaną przez router J, na jej podstawie uzupełnia zawartość własnej tablicy routingu. Jeśli J zna krótszą drogę do pewnych sieci bądź tablica z J zawiera informacje o sieciach, o których K nie posiadał do tej pory zapisu, bądź jeśli K kierował datagramy do pewnych sieci poprzez router J i odległość między J a tymi sieciami uległa zmianie, wówczas router K zmienia i uzupełnia zapis w swojej tablicy. Na rys.4 przedstawiono przykład tablicy routingu routera K i informację niezbędną do jej uaktualnienia nadesłaną z routera J.

Sieć docelowa	Odległość	Trasa
Sieć 1	0	bezpośrednio
Sieć 2	0	bezpośrednio
Sieć 4	8	przez router R
Sieć 17	5	przez router M
Sieć 24	6	przez router J
Sieć 30	2	przez router Q
Sieć 42	2	przez router J

(a)

Sieć docelowa	Odległość
Sieć 1	2
→Sieć 4	3
Sieć 17	6
→Sieć 21	4
Sieć 24	5
Sieć 30	9
→Sieć 42	3

(b)

Rys. 4. Stan tablicy routera K (a) oraz nadesłana z routera J informacja (b) mająca posłużyć K do uaktualnienia tras. Pozycje oznaczone strzałką zostaną wykorzystane do zmiany istniejących zapisów bądź zostaną dodane jako nowe w tablicy routera K

Fig. 4. An existing route table for router K (a), and an incoming routing update message from router J. The marked entries will be used to update existing entries or add new entries to K's table

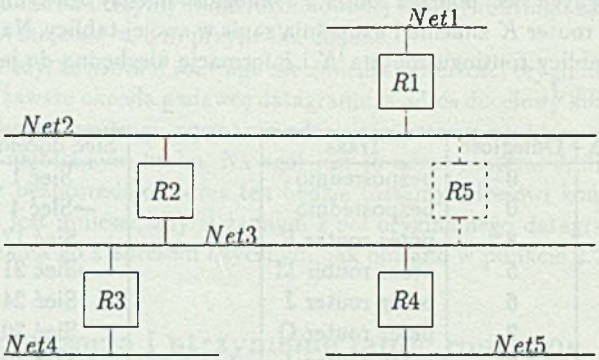
Jeśli w tablicy routera J figuruje odległość N, to K aktualizuje swoją tablicę zapisując odległość $N + 1$. Termin *vector-distance* wywodzi się z faktu, że okresowo propagowane informacje o drogach składają się z par (V, D), gdzie V identyfikuje cel, a D odległość do niego.

Mimo, iż algorytm ten jest łatwy do implementacji, ma pewne słabe strony. W środowisku stabilnym informacje o drogach propagowane są między wszystkimi routerami w

sieci. Gdy stan sieci gwałtownie się zmieni (np. awaria pewnych połączeń), to informacja o takim zdarzeniu rozprzestrzeni się wolno po sieci i przez pewien czas istnieje możliwość błędnego kierowania pakietów wskutek tego, że routery będą posiadały niepoprawne informacje o drogach w sieci.

3.1.2. Routing Information Protocol (RIP)

Jedną z najbardziej znanych, zmodyfikowanych implementacji algorytmu *vector–distance* jest protokół znany jako RIP (w niektórych implementacjach unixowych występuje pod nazwą *routed*). Protokół ten angażuje do pracy nad wyznaczaniem dróg w sieci i utrzymaniem ich poprawności wszystkie pracujące w sieci maszyny (hosty i routery). Routery pracują w trybie aktywnym – co oznacza, że wysyłają i odbierają rozgłaszane informacje o drogach, a hosty w trybie pasywnym – jedynie odbierają informacje wysyłane przez routery. Routery rozgłaszają kopie swoich tablic routingu standardowo co 30 s. Przyjęto, że odległość między routerami przyłączonymi do tej samej sieci fizycznej wynosi 1, jeśli oddziela je jeden router pośredniczący 2, itd. Wszystkie komputery oczekują na rozgłaszane informacje o drogach i na ich podstawie aktualizują zawartość swych tablic routingu. Routery do rozprowadzenia tablic routingu wykorzystują fizyczny broadcast.



Rys. 5. Wprowadzenie routera $R5$ tworzy nową, alternatywną drogę pomiędzy sieciami 2 i 3. Oprogramowanie realizujące routing powinno w przypadku uszkodzenia routera $R2$ szybko zaadoptować się do nowej sytuacji i kierować pakiety na alternatywną drogę

Fig. 5. The addition of router $R5$ introduces an alternate path between networks 2 and 3. Routing software can quickly adapt to failure of one router and automatically switch routes to the alternate path

Rozgłaszana przez routery wiadomość zawiera pary: adres sieci i odległość do tej sieci. Wykorzystywanie jako miary odległości liczby opartej na ilości routerów na danej drodze nie prowadzi zawsze do wyznaczenia optymalnej drogi. Przykładowo droga poprzez trzy routery połączone szybką siecią typu Ethernet będzie efektywniejsza od drogi poprzez dwa routery połączone linią telefoniczną, lecz RIP skieruje pakiety na tę drugą trasę.

Wszystkie maszyny współpracujące w ramach protokołu RIP odbierają rozgłaszane wiadomości o drogach i wg reguły opisanej w p. 3.1.1 aktualizują swoje tablice routingu. Na powyższym rysunku router $R1$ będzie rozgłaszał na sieć 2 wiadomość zawierającą parę $(1, 1)$, co oznacza, że osiąga sieć 1 po koszcie 1. Routery $R2$ i $R5$ odbiorą tę informację i w

swoich tablicach routingu wprowadzą zapis o drodze do sieci 1 przez router $R1$ z kosztem 2. Następnie routery te rozgłoszą w sieci posiadane informacje o drogach, włączając parę (1,2).

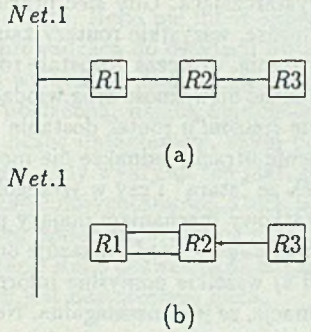
Protokół RIP wprowadza wiele mechanizmów poprawiających funkcjonowanie algorytmu *vector-distance*. Router musi utrzymywać zapis o drodze do danej sieci, aż nie nadejdzie informacja o krótszej drodze do tej sieci. Zapobiega to oscylacjom pomiędzy dwoma bądź większą liczbą dróg o tej samej długości.

Kolejnym zagrożeniem są tzw. zapętlenia drogi pakietu *(routing loops)* polegające na cyklicznym obiegu pakietu na drodze pomiędzy dwoma routerami.

By zapobiec niestabilności algorytmu, należy ograniczyć maksymalną wartość licznika odległości – w implementacji protokołu ograniczono do 16, co jednocześnie powoduje jego nieprzydatność w dużych sieciach o wielu routerach.

Jeszcze jednym problemem wynikającym z algorytmu *vector-distance* jest problem powolnej zbieżności. W wypadku zmiany konfiguracji dróg w sieci informacja o tym fakcie rozprzestrzenia się po sieci wolno i przez dłuższy czas od momentu zaistnienia zmiany routery i hosty posiadać będą niepoprawną informację o trasach. Wybór niewielkiej wartosci maksymalnej licznika odległości pomaga ograniczyć problem, lecz nie usuwa go całkowicie. Należy dodać, że odległość równa 16 oznacza, że do celu jest "nieskończenie daleko", tzn. że liczba pośredniczących routerów jest większa niż 15, bądź cel jest nieosiągalny z innych przyczyn.

Chwilowa niepoprawność zawartości tablic routingu jest nie do uniknięcia. By zrozumieć ten problem, rozpatrzmy przypadek przedstawiony na rys.6 .



Rys. 6. Problem powolnej zbieżności. Droga do sieci 1 wiedzie poprzez trzy routery (a). Połączenie z siecią 1 ulega przerwaniu (b) i router $R2$ powoduje zapętlenie drogi pakietu, rozgłaszając informację o ciągłym istnieniu połączenia

Fig. 6. The slow convergence problem. In (a) three routers each have a route to network 1. In (b) the connection to network 1 has vanished, but $R2$ causes a loop by advertising it

Router $R1$ jest bezpośrednio połączony z siecią 1 i rozgłasza parę (1,1), router $R2$ odbiera tę informację, umieszcza w swojej tablicy zapis (1,2), a następnie rozgłasza ją . Router $R3$ podobnie, z tym że zapisuje i rozgłasza parę (1,3).

W przypadku przerwania połączenia pomiędzy siecią 1 a $R1$, router $R1$ zmieni w

swej tablicy zapis o odległości do tej sieci, umieszczając wartość 16 (nieskończoność) i w najbliższym cyklu rozgłaszania wyśle taką informację w kierunku *R*2.

Jednakże w razie, gdyby protokół nie posiadał dodatkowych mechanizmów zapobiegających negatywnym skutkom takiego wypadku, mogłoby się zdarzyć, że inny z routerów rozgłosiłby informację o drodze do sieci 1, wyprzedzając w tej czynności router *R*1. Załóżmy, że router *R*2 rozgłosił informacje o drogach krótko po przerwaniu łącza między *R*1 a siecią 1, zanim dokonał tego router *R*1. Wówczas *R*1 po odebraniu informacji z *R*2 umieści w swojej tablicy routingu zapis o drodze do sieci 1 wiodącej przez router *R*2 po koszcie 3 (2 między *R*2 a siecią 1 i dodatkowo 1 między *R*2 a *R*1). Jeśli *R*2 odbierze datagram skierowany do sieci 1, prześle go do *R*1 a ten ponownie wyśle go do *R*2. Datagram będzie w ten sposób oscylował, aż wartość w polu *time-to-live* osiągnie zero (zawartość tego pola jest dekrementowana przy przejściu przez router) i datagram zostanie usunięty. Stan taki będzie utrzymywał się stosunkowo długo. Po następnym cyklu rozgłaszania *R*2 odnotuje drogę do sieci 1 przez *R*1 po koszcie 4, o której następnie poinformuje *R*1, ten zpisze sobie drogę do sieci 1 przez *R*2 po koszcie 5 itd., aż miara odległości osiągnie wartość 16, wówczas routery stwierdzą, że sieć 1 jest nieosiągalna.

By zapobiec takim przypadkom, RIP dysponuje techniką noszącą nazwę *split horizon update*. Router zapamiętuje, którym z interfejsów sieciowych otrzymał informację o danej drodze i gdy podejmie rozgłaszanie informacji z własnej tablicy routingu, nie wyśle informacji o drodze poprzez ten z interfejsów, przez który ją otrzymał. Wracając do poprzedniego przykładu *R*2 nie prześle do *R*1 pary (1,2), więc pochodząca od *R*1 informacja (1,16) zostanie w krótkim czasie dostarczona wszystkim routerom w sieci. Jednakże technika ta nie jest wystarczająca. Gdy sieć pracuje prawidłowo, jeśli router rozgłasza informację o krótkiej drodze, wszystkie routery zauważą ten fakt szybko. Jeśli jednak router zaprzestanie rozgłaszania, wówczas pozostałe routery muszą odczekać ustalony odcinek czasu *(timeout)*, by uznać nieważność dróg wiodących przez ten router. Może jednak zdarzyć się, że po upływie *timeout*'u router dostanie informację o alternatywnej drodze do sieci, do której połączenie utracił, jednakże nie może on wiedzieć, czy "nowa" droga nie ma odcinków wspólnych ze "starą" i czy w szczególności wspólnym odcinkiem nie jest odcinek uszkodzony. Dotatkowy mechanizm mający pomóc w rozwiązaniu problemów tego rodzaju nosi nazwę *hold down period*. Nakazuje on routerom ignorować przez ustalony okres czasu (typowo 60 s) wszelkie pomyślne informacje o drogach do sieci, o której otrzymały wcześniej informacje, że jest nieosiągalna. Negatywnym efektem takiego rozwiązania jest fakt utrzymywania zapętlenia dróg (jeśli wystąpi) przez cały ten okres. Kolejny mechanizm polega na tym, że router, który otrzymał informację o utracie połączenia z pewną siecią, ma natychmiast przystąpić do rozgłaszania, nie czekając, aż nadejdzie chwila wynikająca z normalnego cyklu rozgłaszania.

Wszystkie te techniki pomagają rozwiązać pewne problemy, wprowadzają jednak inne. Na przykład jeśli większa liczba routerów połączona jest tą samą siecią fizyczną i jeśli jeden z nich rozgłosi informację o utracie połączenia z pewną siecią, wymusi to na pozostałych odpowiednią zmianę w tablicach routingu i natychmiastowe rozgłoszenie poprzez wspólną sieć. Jeśli druga runda rozgłaszania spowoduje zmiany w tablicach kolejnych routerów, może wywołać to kolejną falę rozgłaszania.

Wykorzystywanie fizycznego rozgłaszania, możliwość powstawania zapętleń, mechanizm *hold down* mogą spowodować, że w sieciach rozległych RIP stanie się nieefektywny.

Obciążenie sieci rozgłaszanymi przez routery informacjami będzie rosło wraz ze wzrostem liczby routerów. Również zagrożenie zapętleniami w wypadku kanałów transmisyjnych o małej przepustowości rodzi możliwość nasycenia łącza krążącymi pakietami, uniemożliwiając routerom wymianę informacji koniecznych do przerwania tych pętli.

3.1.3. Protokół Hello

Protokół Hello bazuje na algorytmie *vector-distance*, jednakże wykorzystuje jako miary odległości wartości opóźnień wnoszonych przez sieć na poszczególnych drogach. Protokół ten wykonuje dwie funkcje: synchronizuje zegary pomiędzy współpracującymi maszynami i umożliwia każdej z maszyn wyznaczenie najkrótszych dróg. Każda z maszyn utrzymuje tablicę z przybliżonymi wskazaniami zegarów systemowych sąsiadujących z nią maszyn. W każdym z transmitowanych pakietów komputer umieszcza aktualną wartość swojego zegara. Na podstawie tej wartości oraz zawartości tablicy odbiorca wylicza bieżące opóźnienie na łączu. Okresowo maszyny uaktualniają w swych tablicach zapisy wskazań zegarów sąsiadów. Routery co pewien ustalony czas wysyłają do swych sąsiadów tablice zawierające przybliżone wartości opóźnień na drogach do wszystkich pozostałych routerów. Przypuśćmy, że maszyna A wysyła do maszyny B tablicę zawierającą adresy sieci i wartości opóźnień na drogach do nich. Router B porównuje zawartość swojej tablicy z wartościami nadesłanymi przez A. Jeśli dotychczas zapisana wartość opóźnienia na drodze z B do D jest większa niż wartość opóźnienia z A do D powiększonego o opóźnienie między B a A, to B zmienia zapis w tablicy routingu i będzie przesyłać datagramy kierowane do D poprzez A.

Jak każdy algorytm routingu, Hello nie może zmieniać tras zbyt szybko, może spowodować to niestabilność prowadzącą do oscylacji dwustanowych. Oscylacje te polegają na cyklicznym włączaniu i wyłączaniu obciążenia datagramami pewnego odcinka drogi. Router znajduje najmniej obciążoną, dającą najmniejsze opóźnienia drogę do pewnego celu i kieruje na nią dużą liczbę datagramów, czym powoduje znaczny wzrost obciążenia drogi i wzrost opóźnienia. Gdy wykrywa ten fakt, przełącza całe obciążenie na inną, alternatywną drogę, odciążając dotychczasową, która ponownie staje się drogą wnoszącą najmniejsze opóźnienia, więc router znów kieruje na nią datagramy itd. By zapobiec takiemu zjawisku, implementacje Hello pozwalają na zmianę dotychczasowej drogi na nową, jeśli różnica w opóźnieniach jest dostatecznie duża.

3.2. Protokół SPF

Alternatywną grupą protokołów dla protokołów opartych na algorytmie *vector-distance* są protokoły wykorzystujące algorytm znany jako *Shortest Path First (SPF)*. Każda z współpracujących w ramach takiego protokołu maszyn musi posiadać kompletną informację o topologii sieci, w postaci mapy ukazującej wszystkie routery i połączenia między nimi. W formie abstrakcyjnej można przedstawić sieć jako graf, w którym wierzchołki odpowiadają routerom, a krawędzie połączeniom między routerami. Wierzchołki połączone są krawędzią tylko wtedy, gdy między odpowiadającymi im routerami istnieje fizyczne połączenie.

Zamiast przesyłać między routerami informacje o wszystkich możliwych drogach, ro-

utery realizujące SPF wykonują jedynie dwa zadania. Pierwsze, to aktywna kontrola statusu sąsiadujących routerów, a drugie - okresowe rozesłanie raportu o statusie sąsiednich routerów do wszystkich pozostałych routerów w sieci.

By skontrolować status sąsiedniego routera, router okresowo wysyła do niego datagram z zapytaniem, czy pracuje i jeśli otrzyma odpowiedź pozytywną, oznacza połączenie jako poprawne, a w wypadku odpowiedzi negatywnej - jako przerwane. W praktyce, by zapobiec oscylacjom stanu połączenia między stanami "poprawne" a "przerwane", połączenie przyjmuje się jako poprawne w przypadku gdy m na n datagramów z zapytaniem o status sąsiada da odpowiedź pozytywną. Każdy router okresowo wysyła raport o statusie połączeń do pozostałych routerów. Jeśli sieć nie umożliwia rozgłaszania, raport wysyłany jest z wykorzystaniem połączeń dwupunktowych. Gdy router odbierze raport od innego routera, aktualizuje swoją mapę połączeń i wyznacza optymalne drogi do wszystkich sieci docelowych, wykorzystując *algorytm wyboru najkrótszej drogi Dijkstra'y*.

Główną zaletą protokołów wykorzystujących algorytm SPF jest fakt, że każdy z uczestniczących komputerów wyznacza optymalną drogę sam, opierając się na tych samych bazowych informacjach, co reszta maszyn w sieci, oraz niewielka objętość wymienianych między routerami danych, nie powodująca znacznego wzrostu obciążenia sieci.

Mimo zalet tego protokołu jest on stosunkowo rzadko wykorzystywany w praktyce. Jedna z jego pierwszych implementacji została wykorzystana w sieci ARPANET, obecnie można spotkać implementację algorytmu SPF w protokole o nazwie OSPF *(Open SPF)*.

Protkól OSPF umożliwia lepsze wykorzystanie możliwości, jakie daje sieć IP. Pozwala on bowiem na wyróżnienie odrębnych dróg dla datagramów wymagających różnych typów obsługi (np. transmisja drogą o najmniejszym opóźnieniu lub transmisja drogą o największej przepustowości). OSPF przy wyborze drogi opiera się nie tylko na adresie docelowym datagramu, lecz także na zawartym w strukturze każdego datagramu polu *type of service*. W przypadku istnienia wielu dróg wiodących do tego samego celu o tym samym koszcie, OSPF potrafi rozłożyć równomiernie obciążenie tych połączeń. Umożliwia on podział całej sieci na obszary, których topologia pozostaje ukryta dla pozostałej części sieci, ułatwiając w ten sposób zarządzanie i zmniejszając ilość wymienianej między routerami informacji. Możliwa jest także kontrola wiarygodności wymienianej informacji o drogach. Eliminuje to z współpracy routery administrowane przez nieupoważnionych nadzorców, mogące zdezorganizować pracę sieci. Wprowadzone zostały także mechanizmy umożliwiające współpracę w ramach sieci wielodostępnych (np. Ethernet). Jak wcześniej opisano, SPF opierał się na grafie, w którym dwa sąsiednie wierzchołki odpowiadały routerom połączonym siecią fizyczną, a każdy z routerów okresowo przepytywał sąsiadujące routery o ich status. Jeśli K routerów przyłączonych jest do sieci Ethernet, wówczas zostanie rozgłoszonych K^2 zapytań o status. Protokół OSPF minimalizuje rozgłaszanie, pozwalając na wyznaczenie routera, który w imieniu pozostałych rozgłasza status połączenia.

By zapewnić większą elastyczność, OSPF pozwala administratorom na opisanie topologii wirtualnej, umożliwiając wprowadzenie wirtualnego połączenia pomiędzy routerami, między którymi nie istnieje bezpośrednie fizyczne połączenie.

4. Podsumowanie

Powyższe omówienie nie wyczerpuje, oczywiście, problematyki routingu w sieciach TCP/IP. Przedstawione jednak zagadnienia pozwalają na pewne zorientowanie się w tematyce. Jak można było zauważyć, dla poprawnego funkcjonowania sieci oraz pełnego wykorzystania jej możliwości technicznych duże znaczenie ma sposób, szybkość i poprawność wyznaczenia drogi datagramów w sieci. Wenątrz małych, rzadko rekonfigurowanych sieci możliwe jest ustawienie stałych tras datagramów, dokonywane manualnie przez administaratora sieci podczas jej zestawiania. Wraz ze wzrostem liczby komputerów współpracujących poprzez sieć, a szczególnie przy rosnącej liczbie połączeń pomiędzy routerami, zagadnienie komplikuje się. Przedstawione w artykule protokoły umożliwiają automatyczną inicjalizację i utrzymywanie tablic routingu. Najszerzej przedstawiony został protokół RIP, jako najczęściej spotykany. Obarczony jest on wieloma niedoskonałościami i mimo iż w zamierzeniach jego autorów (University of California) nie był przeznaczony do wykorzystywania w rozległych sieciach o wielu routerach, stał sie bardzo popularny dzięki rozprzestrzenieniu się różnych modyfikacji systemu 4BSD UNIX, w którego skład wchodził również RIP. Pozostałe omówione w niniejszym opracowaniu protokoły są może doskonalsze, lecz RIP góruje nad nimi popularnością.

LITERATURA

[1] Douglas E. Comer. *Internetworking with TCP/IP.* Prentice–Hall International 1991.

[2] Sun Microsystems. *System and Network Administration.*

[3] Marshall T. Rose *The Simple Book - an Introduction to Management of TCP/IP Based Internets.* Prentice–Hall International 1991.

Recenzent: Dr inż. Andrzej Kwiecień

Wpłynęło do Redakcji 27 września 1993 r.

Abstract

The TCP/IP network connected by routers, and consists of many types of physical links, is the basis to exchange IP datagrams. In the packet switching systems (as TCP/IP networs), *routing* refers to process of choosing a path over which to send packets, and *router* refers to computer making such a chois. Router's software, ideally, would establish

and maintain set of routes to all destination in network, and examine network load, datagram lenght or the type of service, when selecting the best path. The article presents basic method for establish routes *Table-Driven Routing*, and some protocols for automatic initiation and update of routing tables. The most widely used protocol RIP, is based on vector-distance algorithm. The RIP is using a *hop count* - number of routers which datagram must pass by to reach the destination, to meter the distance between source and destination networks. It causes the RIP not scale well when paths differ in technologies. Second protocol named Hello, is also based on vector-distance algorithm, however uses network delay instead of *hop count*. The last protocol described, is SPF (and OSPF). The SPF scales better than protocols based on vector-distance algorithms. It actively tests the status of all neighbor routers connected to common network, and propagates the link status information to all others routers. Routers use this information to compute the best path to all destinations.