

Routing dynamiczny

Routing statyczny vs. dynamiczny



- tablice routingu konfigurowane przez administratora (-ów),
- przewidywalny – trasa po której pakiet jest przesyłany jest dobrze znana i może być kontrolowana,
- łatwy do skonfigurowania w małych sieciach,
- brak skalowalności



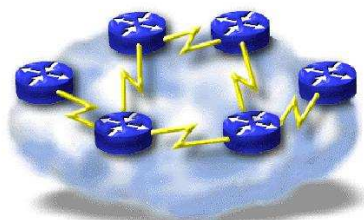
- tablice routingu konfigurowane przez protokoły routingu,
- mniej przewidywalny, gdyż protokoły mają możliwość zmieniania zawartości tablic routingu
- łatwy do skonfigurowania (spora część pracy przerzucona na protokoły),
- łatwiej skalowalny

Routing statyczny vs. dynamiczny

- bardzo ograniczona zdolność dostosowania się do dynamicznych zmian w konfiguracji sieci,
- łącza nie są dodatkowo obciążone wiadomościami służącymi do rutowania,
- brak obsługi redundantnych połączeń
- umiejętność dostosowania się do dynamicznych zmian w konfiguracji sieci,
- konieczność okresowej wymiany danych pomiędzy routerami to z punktu widzenia użytkownika niepotrzebne obciążenie sieci,
- możliwość wykorzystywania wielu ścieżek jednocześnie,
- trudności w implementacji, różnice pomiędzy sprzętem pochodzącym od różnych producentów

Wymagania stawiane protokołom routingu

Routing dynamiczny



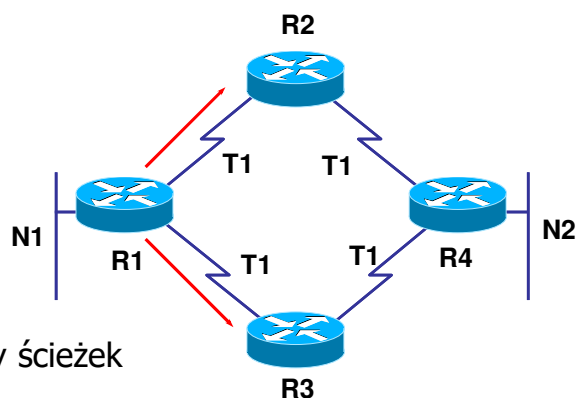
Którędy przesłać pakiet?

- routery (nie administratorzy) wybierają trasę
 - eliminacja pętli
- w wypadku, gdy możliwych jest kilka tras, wybór może być trudny
- kryterium wyboru trasy (tzw. *metryka*) może być np.:
 - liczba routerów po drodze,
 - przepustowość łączy,
 - obciążenie łączy,
 - procent strat na łączach,
 - opóźnienie na łączach,
 - MTU,
 - ...

Zbieżność (konwergencja)

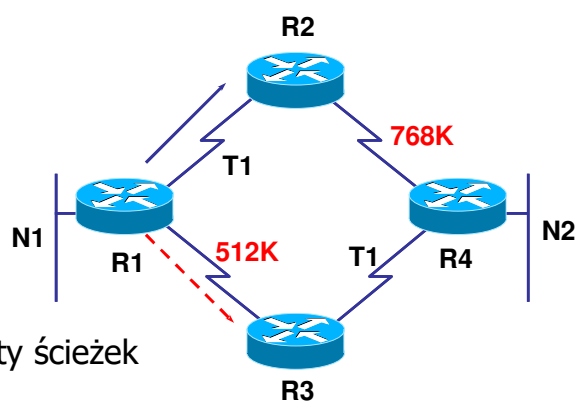
- router potrzebuje czasu na znalezienie alternatywnej ścieżki w wypadku zmiany topologii sieci (np. awaria),
- czas, po którym routery będą miały jednakowy "obraz" sieci (**czas zbieżności**) jest zależny od konfiguracji (np. odstęp między okresami rozsyłanymi pakietami),
- mówimy, że dany protokół routingu dynamicznego (P1) jest **szybciej zbieżny**, niż drugi (P2) wtedy, gdy w sieci o takiej samej topologii cechuje się krótszym czasem zbieżności

Równoważenie obciążenia



- Równe koszty ścieżek

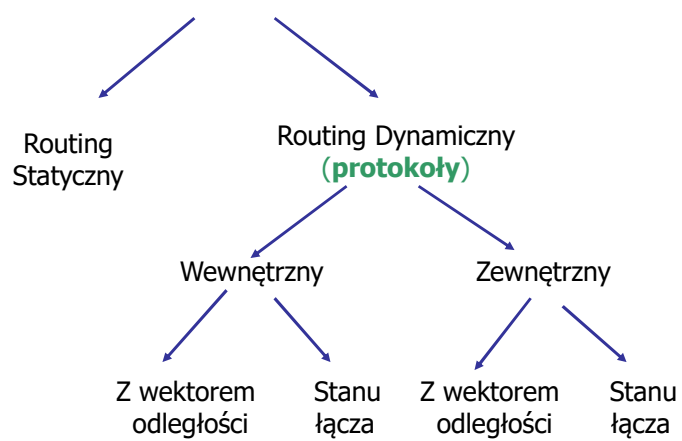
Równoważenie obciążenia



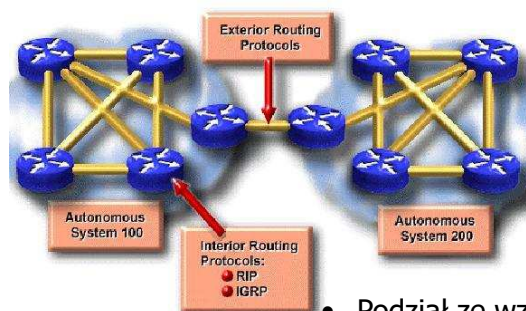
- Zbliżone koszty ścieżek

Protokoły routingu

Podział metod Routingu

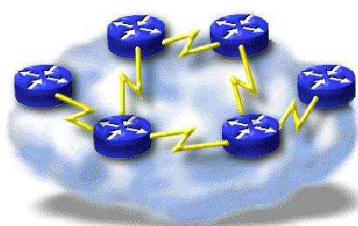


Klasyfikacja protokołów routingu



- Podział ze względu na obszary zastosowań
 - protokoły wewnętrzne
 - protokoły zewnętrzne
- Podział ze względu na charakter wymienianych informacji
 - protokoły dystans-wektor
 - protokoły stanu łącza

System autonomiczny (AS)



AS = autonomous system

- składa się z routerów (i przyłączonych do nich sieci) znajdujących się **pod wspólną administracją**,
- jest identyfikowany przez **16-bitowy numer** nadawany przez NIC (Network Information Center),
- numer systemu autonomicznego bywa wykorzystywany przez protokoły routingu dynamicznego działające zarówno wewnątrz, jak i na zewnątrz **AS**

Protokoły routingu – wewnętrzne i zewnętrzne

- Wewnętrzne
 - Stosowane wewnątrz jednej domeny administracyjnej
 - Proste, w małym stopniu obciążają routery
 - Mało skalowalne
 - RIP (Routing Information Protocol), IGRP (Interior Gateway Routing Protocol), OSPF (Open Shortest Path First),
- Zewnętrzne
 - Odpowiadają za wymianę informacji pomiędzy dwiema niezależnymi administracyjnie sieciami
 - Dają się skalować, łatwo obsługują duże sieci
 - Są skomplikowane, ilość dodatkowych informacji przesyłanych siecią może szybko zablokować pracę małej lub średniej sieci
 - EGP (exterior gateway protocol), BGP (border gateway protocol)
- Można je zamieniać, ale nie jest to mądre, bo zostały przystosowane do innego trybu pracy

Protokoły z wektorem odległości ang. **Distance Vector**

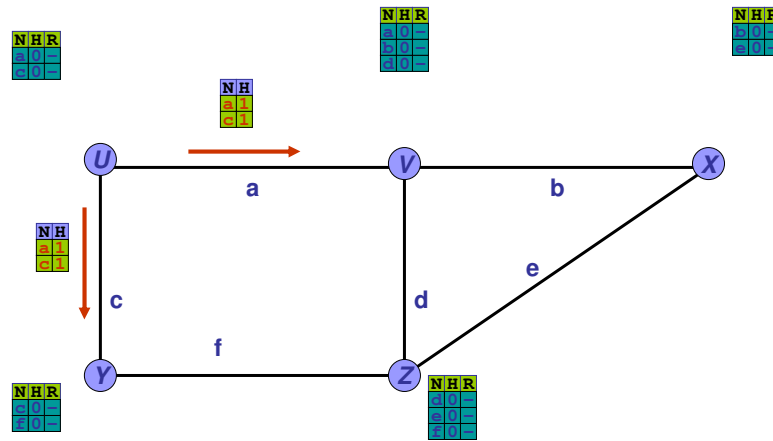
Protokoły dystans-wektor

- Router regularnie wysyła wszystkim swoim sąsiadom informacje na temat każdej dostępnej, znanej sobie sieci:
 - Jak daleko do niej jest (**dystans**)
 - Czas podróży
 - Liczba przeskoków
 - Koszt przesyłu
 - Jak się można do niej dostać (**wektor**)
 - Zwykle — „wyślij do mnie, bo ja wiem, jak to przesłać dalej”.
 - Inny router. Np. gdy router docelowy nie obsługuje danego protokołu routingu.

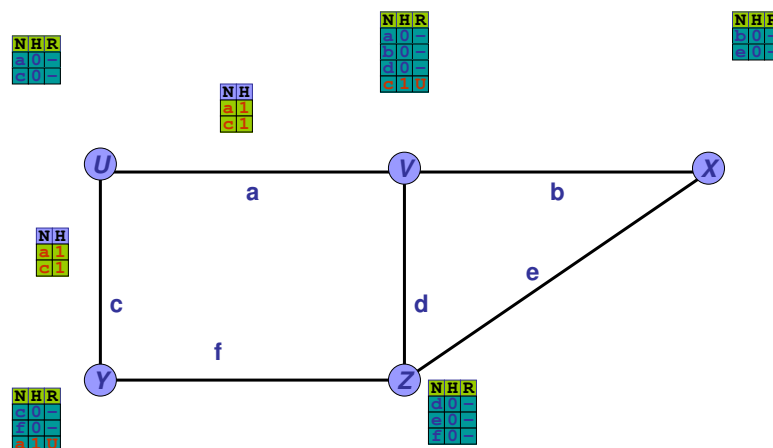
Budowa tablic routingu

- Routery otrzymują tablice od swoich sąsiadów
- Zapamiętują najlepszą znaną odległość do określonego odbiorcy oraz router który przesłał taką informację
- Uaktualniają wpis jeśli odbiorą informację o lepszej odległości
- Obliczają minimalną odległość przyrostowo, nie potrzebują przechowywać wszystkich danych od sąsiadów

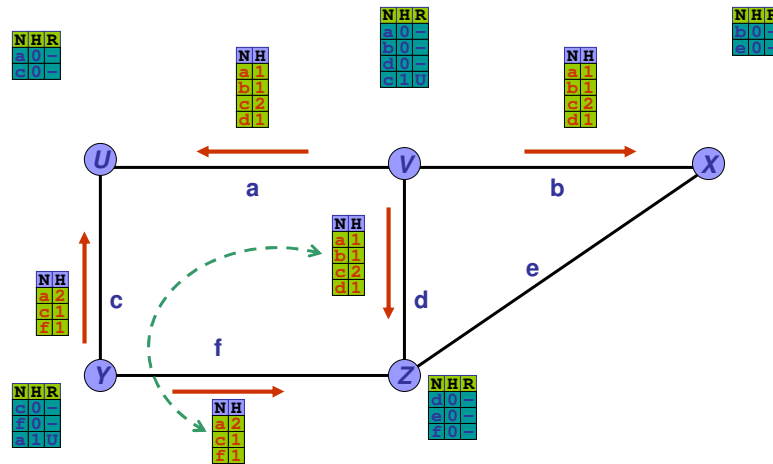
Protokół d-w — stan 0



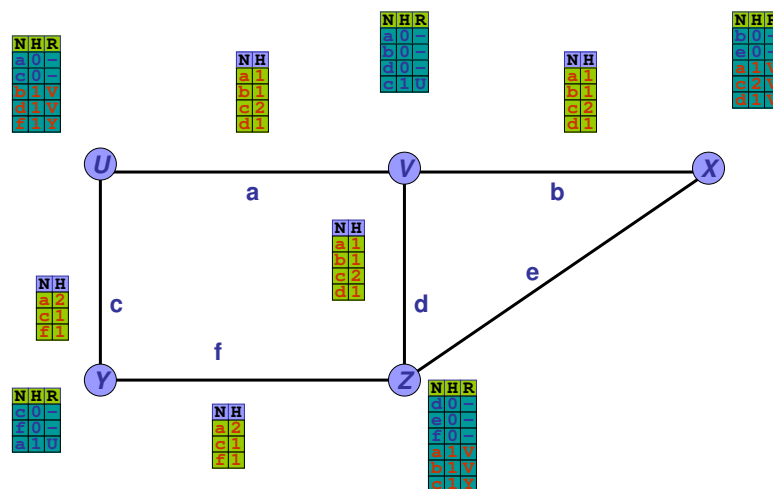
Protokół d-w — stan I



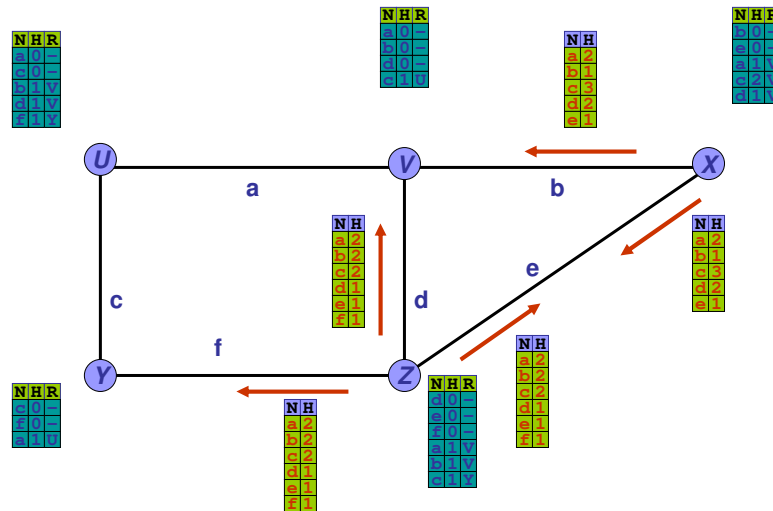
Protokół d-w — stan I



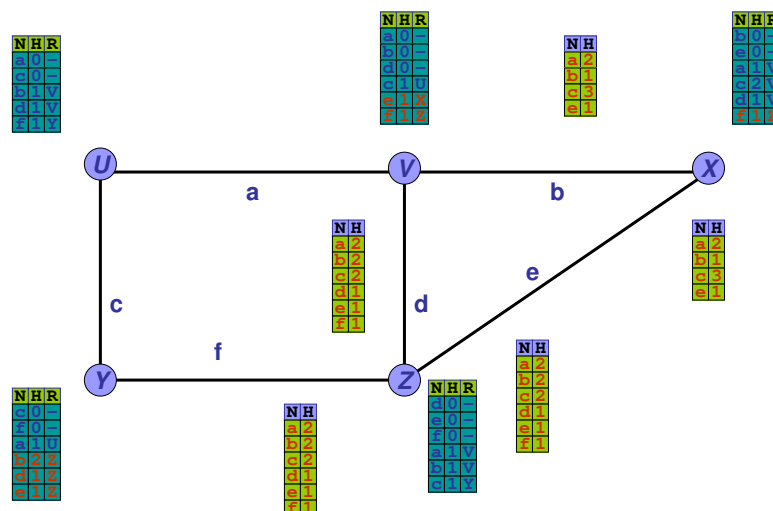
Protokół d-w — stan II



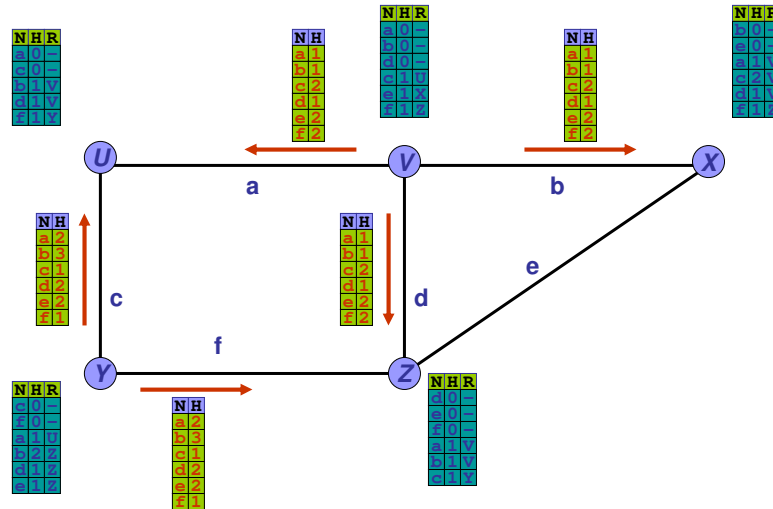
Protokół d-w — stan II



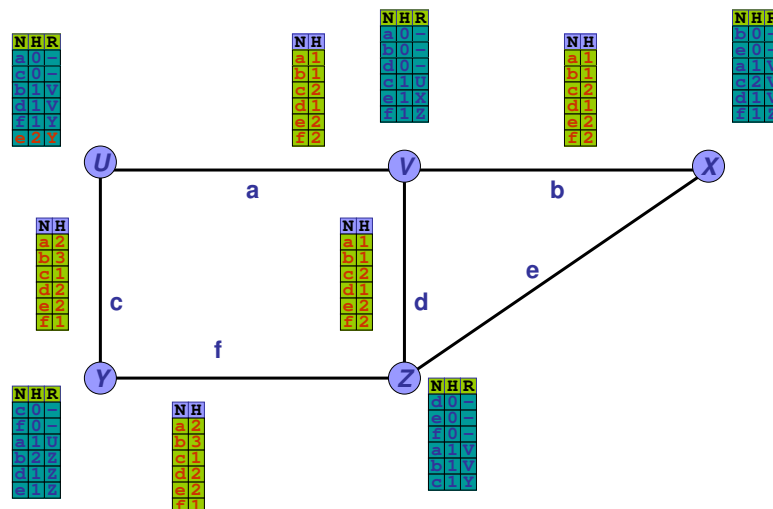
Protokół d-w — stan III



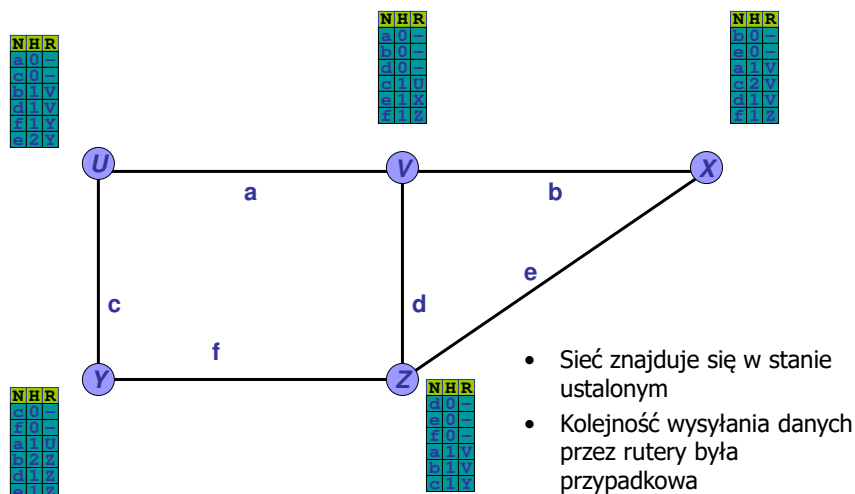
Protokół d-w — stan III



Protokół d-w — stan IV



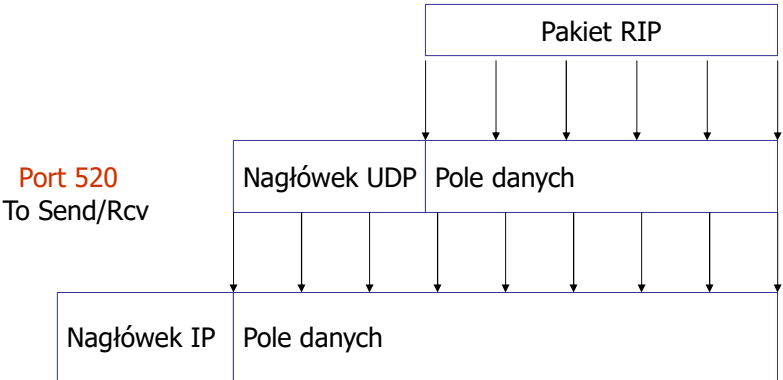
Protokół d-w — stan ustalony



Routing Information Protocol (RIP)

- Ruter wysyła informacje co 30s do wszystkich swoich sąsiadów — pakiety typu broadcast — o znanych sobie sieciach i odległości do nich
- Miarą odległości jest liczba routerów jaką należy przejść, żeby dostać się do danej sieci
- Po 180s nie odświeżona droga jest usuwana z tablicy routingu

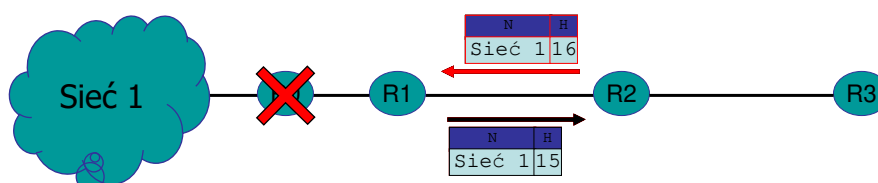
Pakiet RIP – enkapsulacja



Pakiet RIP

0	4	8	12	16	20	24	28	31
polecenie (zapytanie/odpowiedź)			wersja	musi być wypełnione zerami				
identyfikator rodziny adresów				musi być wypełnione zerami				
adres IP								
musi być wypełnione zerami								
musi być wypełnione zerami								
metryka								

Wady RIP — liczenie do nieskończoności



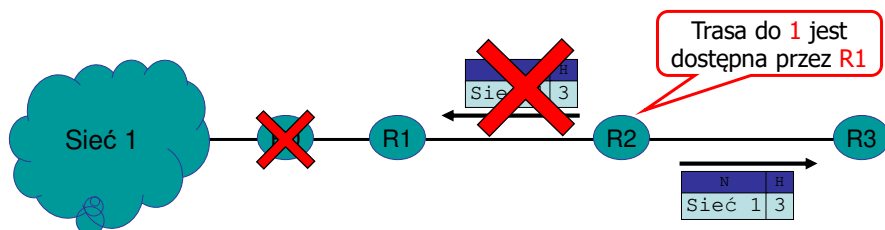
- Ruter R2 wysyła informacje o dostępności sieci 1 co 30 sekund
- Po 180 sekundach R1 wpisuje do swojej tablicy nową drogę do sieci 1

RIP — zmniejszanie prawdopodobieństwa wystąpienia liczenia do nieskończoności

- Uaktualnianie z podzielonym horyzontem
 - router nie propaguje informacji o dostępności sieci na interfejs, przez który prowadzi najlepsza trasa
- Wstrzymanie
 - router wstrzymuje się z akceptacją komunikatów o dostępności sieci, o której awarii otrzymał informację (zazwyczaj na 60 sek.)
- Odświeżanie wymuszone
 - w wypadku zmiany w tablicy router nie czeka 30s tylko rozgłasza nową informację natychmiast

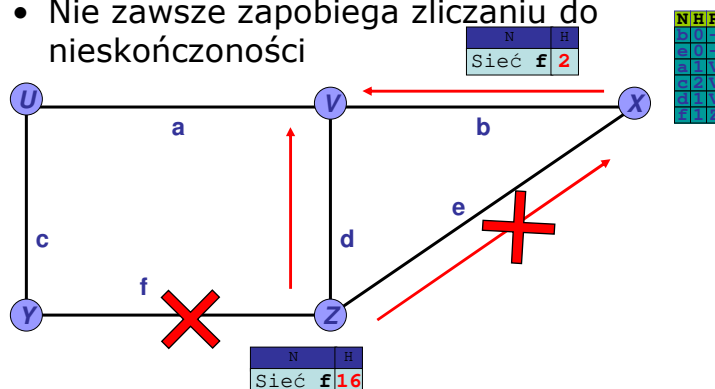
RIP – dzielony horyzont

- Zmniejsza prawdopodobieństwo wystąpienia zjawiska zliczania do nieskończoności



RIP – dzielony horyzont

- Nie zawsze zapobiega zliczaniu do nieskończoności



RIP – wstrzymanie

- Po otrzymaniu komunikatu od routera, że poprzednio dostępna sieć jest niedostępna włącza licznik (hold-down timer)
- Jeśli otrzyma komunikat od tego samego routera, że sieć jest dostępna wyłącza licznik
- Jeśli otrzyma komunikat od innego routera ogłaszający lepszą trasę wyłącza licznik
- Jeśli otrzyma gorsze trasy **ignoruje je**
- Po upływie licznika kasuje wpis

RIP – odświeżanie wymuszone

- Pakiet RIP wysyłany jest natychmiast po zaobserwowaniu zmiany
 - czas zwykle jest opóźniony o kilka sekund, żeby nie spowodować zalewania sieci
- Nie wysyłana jest cała tablica routingu a tylko informacja o zmianach

Wady RIP — synchronizacja

- Co 30 sekund w sieci opartej na protokole RIP następuje znaczny spadek wydajności (synchronizacja komunikatów o tablicach routingu)
 - mniejsza przepustowość lub większy procent zagubionych pakietów
- Rozwiązania:
 - inicjowanie routerów w różnych momentach
 - modyfikacja interwału (15s – 45s; średnio 30s) pomiędzy wysyłaniem kolejnych informacji o zawartości tablicy routingu (losowo)

Wady RIP — rozgłaszanie

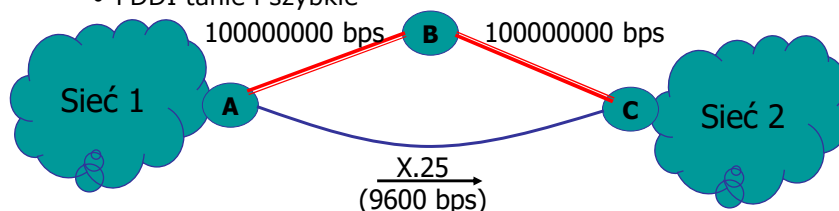
- W przypadku Ethernetu lub FDDI — naturalna metoda przesyłania informacji dotyczących routingu
- W ISDN lub X.25 „milczenie jest złotem”
 - Transmisja wymaga zestawienia kanału transmisyjnego
 - ISDN „B” — 64 kbps
 - kanał wirtualny X.25 — 9,6 kbps
 - przesłanie 2 pakietów RIP zajmuje około 1 sekundy (!)

Wady RIP — rozgłaszanie

- Remedium: tablice routingu rozgłaszane są tylko wtedy, gdy zachodzi taka konieczność
 - transmisja z potwierdzeniem
 - problem:
 - jak stwierdzić, że połączenie z sąsiednim routerem działa?
 - założenie osiągalności
 - jeśli próba przesłania pakietu zawiedzie to przerwane
 - „zapominanie” informacji o gorszych drogach
 - po 30 sekundach się nie pojawiają (!!!)
 - rozwiązanie: przechowywanie listy tras wewnątrz routera
 - » stosowane nie tylko w przypadku łącz typu X.25

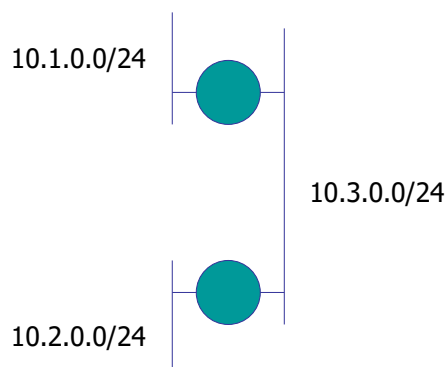
Wady RIP

- Zbyt prosta metryka
 - łączy X.25 i FDDI „tyle samo warte”
 - X.25 drogie i wolne
 - FDDI tanie i szybkie



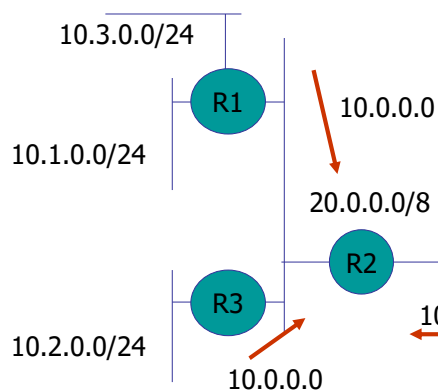
- Dobry do sieci jednorodnych

RIP i podsieci



- Potrafi obsłużyć sytuację, w której istnieją **jednakowo** długie podsieci — na podstawie adresów swoich interfejsów.
 - Wszystkie routery zakładają, że w sieci istnieje jedna długość maski

RIP i podsieci

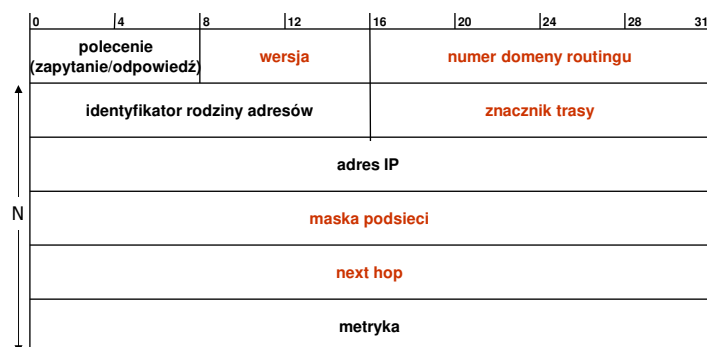


- R1 nie może przekazać informacji o podsieci, bo na tym interfejsie nie ma podsieci — inne routery zinterpretowały by jego komunikat jako drogę do hosta 10.1.0.0
- R2 może nie znać poprawnej drogi do hosta 10.1.1.1

RIP-2

- W ramce zawarta jest również maska podsieci
- Propaguje numer domeny routingu
- Propaguje adres następnego routera (next-hop)

Pakiet RIP-2



RIP-2 — bezpieczeństwo

- Wada RIP-1
 - w RIP-1 każdy komputer nadający z portu 520 jest uznawany za router
 - konieczna ręczna konfiguracja routingu: lista autoryzowanych sąsiadów
- Mechanizm autentykacji RIP-2
 - w ramce RIP-2 można umieścić pole zawierające „hasło”
 - nie powoduje to utraty kompatybilności z RIP-1

RIP-2 — rozgłaszanie

- Wada RIP-1
 - używa broadcastowego adresu MAC do rozsyłania informacji o dostępności do sieci
- RIP-2 używa do tego celu multicastowego adresu IP klasy D (224.0.0.9)
 - nie są konieczne mechanizmy routowania pakietów klasy D, bo informacje dotyczą tylko lokalnej sieci

RIP-2 — podsumowanie

- Oferuje znaczące rozszerzenia względem RIP-1
 - routing bazujący na CIDR i podsieciach
 - autoryzacja pakietów
 - kompatybilność z RIP-1
- Nie eliminuje wszystkich wad RIP-1
 - np. konieczność liczenia do nieskończoności w razie awarii sieci
 - bardzo prosta metryka

Interior Gateway Routing Protocol (IGRP)

- przedstawiciel rodziny „dystans-wektor”
- periodyczne komunikaty o zawartości tablic routingu (co 90 sekund)
- złożone metryki
- właścicielem praw autorskich jest CISCO

Interior Gateway Routing Protocol (IGRP)

- Metryki elementarne

- statyczne

- opóźnienie (ang. delay, ozn. D)
 - mierzone w dziesiątkach mikrosekund
 - suma opóźnień pomiędzy routerem a adresem docelowym
- przepustowość (ang. bandwidth, ozn. B)
 - najmniejsza z przepustowości pomiędzy routerem a adresem docelowym
 - jednostka: 10 000 000 / przepustowość w kbps
 - liczba sekund potrzebna na przesłanie 10 mld bitów
 - 24 bity => zakres przepustowości od 1200 bps do 10 Gbps

Interior Gateway Routing Protocol (IGRP)

- Metryki elementarne - c.d.

- dynamiczne

- niezawodność łącza (ang. reliability, ozn. R)
 - stopień pewności, że pakiet dotrze do celu
 - 8 bitów: 255 odpowiada 100%
- obciążenie łącza (ang. load, ozn. L)
 - dotyczy najbardziej obciążonego łącza na ścieżce
 - 8 bitów: 255 odpowiada 100%
- liczba routerów na ścieżce (ang. hops, ozn. H)
- path MTU

Interior Gateway Routing Protocol (IGRP)

- Rzeczywista (pełna) metryka

$$M = \left(K1 * B + K2 * \frac{B}{256 - L} + K3 * D \right) * \frac{K5}{R + K4}$$

dla $K5=0$

- Współczynniki K1 - K5 są ustalane przez administratora
 - np. ustalenie $K5=0$ powoduje, że współczynnik niezawodności nie jest brany pod uwagę
 - ustawienie domyślne:
 - $K1 = K3 = 1$
 - $K2 = K4 = K5 = 0$
- > $M = B + D$

IGRP — rozważania

- Pakiety IGRP nie przekazują obliczonych metryk tylko składowe wymagane do obliczenia metryki — każdy router stosuje swoje współczynniki
- Dobrze jest ustalić takie same współczynniki we wszystkich routerach w danej domenie routingu.
- Nieskończoność jest zakodowana jako $D=MAX$

Interior Gateway Routing Protocol (IGRP)

- Sposoby zapobiegania pętlom
 - „podzielony horyzont”
 - „wstrzymanie (ang. path holddown)”
 - „odświeżanie wymuszone”
- Routing wielościeżkowy
 - dzielenie obciążenia między kilka możliwych ścieżek
 - pozytywny „efekt uboczny”: zapasowa ścieżka w razie awarii na najlepszym łączy

Enhanced IGRP (EIGRP)

- Wady IGRP
 - synchronizacja
 - nieefektywne algorytmy zapobiegania pętlom
- Algorytm DUAL (*J.J. Garcia-Luna-Aceves*)
(diffusing update algorithm)
 - usuwanie pętli
 - zastosowanie zarówno w protokołach dystans-
wektor, jak i w prot. stanu łącza

Protokoły stanu łącza
ang. **Link State**

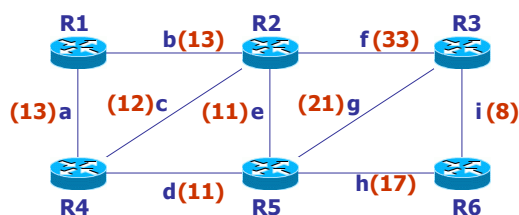
Protokoły stanu łącza

- Router wysyła informację do wszystkich routerów o stanie swoich łączy
- Router przechowuje w pamięci „mapę sieci”
 - mapa jest regularnie uaktualniana
 - możliwe centralne wyznaczanie ścieżki do celu
- Szybka reakcja na zmiany topologii
 - relatywnie (w stosunku do protokołów dystans-wektor) małe pakiety informacyjne
- Brak pętli
 - brak konieczności „liczenia do nieskończoności”
- Możliwość stosowania wielu metryk

Tworzenie tablic routingu

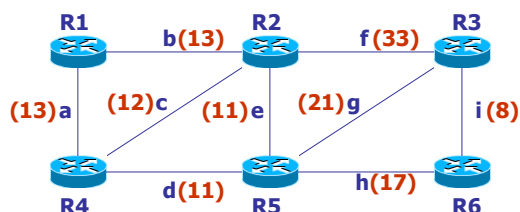
- Wykrywanie sąsiadów
- Współdzielenie informacji o sąsiadach ze wszystkimi routerami
- Tworzenie tablic zawierających pełną informację o sieci (mapa sieci) oraz na ich podstawie tablicy routingu

Wykrywanie sąsiadów



- R3 wysyła przez wszystkie swoje interfejsy pakiet HELLO ('Hello, to ja R3')
- R3 otrzymuje od swoich sąsiadów pakiety HELLO
- Pakiety HELLO muszą być potwierdzone

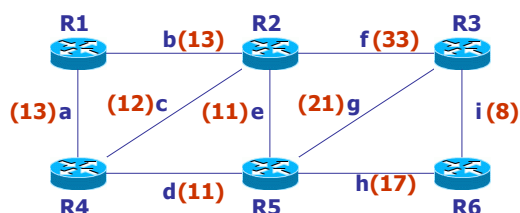
Pakiety LSA



- Budowa pakietów LSA (Link State Advertisement)
 - Przykład dla routera R3
- Zalewanie sieci pakietami
 - Nowy pakiet
 - Zapamiętany
 - Przesłany na wszystkie łącza oprócz tego, które otrzymało pakiet
 - Stary pakiet
 - Odrzucony

Dokąd	Sieć	Metryka
R2	f	33
R5	g	21
R6	i	8

Tablica LSD



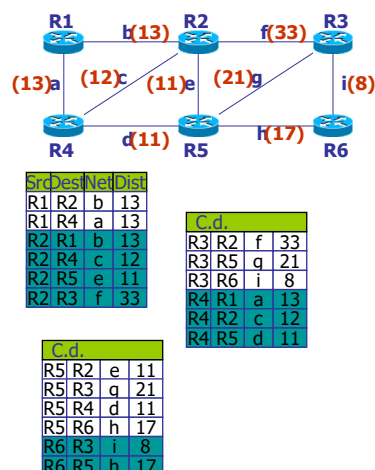
- Budowa tablic LSD (Link State Database)

Src	Dest	Net	Dist
R1	R2	b	13
R1	R4	a	13
R2	R1	b	13
R2	R4	c	12
R2	R5	e	11
R2	R3	f	33

C.d.			
R3	R2	f	33
R3	R5	g	21
R3	R6	i	8
R4	R1	a	13
R4	R2	c	12
R4	R5	d	11

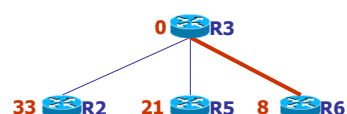
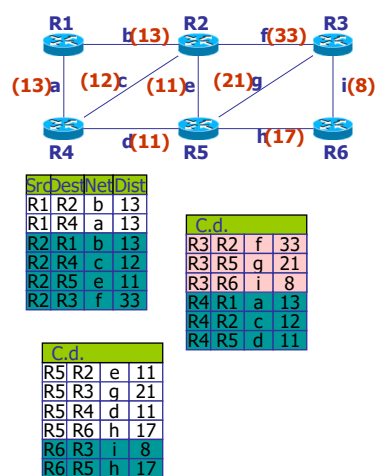
C.d.			
R5	R2	e	11
R5	R3	g	21
R5	R4	d	11
R5	R6	h	17
R6	R3	i	8
R6	R5	h	17

Tworzenie tablicy routingu



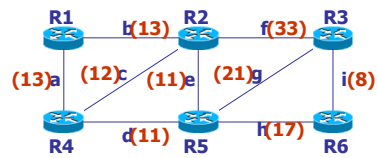
- Algorytm Dijkstry
 - Tworzenie drzewa
 - Korzeń — system tworzący drzewo
 - Gałęzie — połączenia do innych systemów
- Analizujemy budowę drzewa z punktu widzenia routera R3

Algorytm Dijkstry



- Tworzy drzewo tymczasowe i ostateczne
 - Tymczasowe — elementy, do których jeszcze nie zna lepszej (koszt) drogi
 - Ostateczne — elementy do których zna najkrótszą drogę
- Koszt — suma kosztów od elementu do korzenia

Algorytm Dijkstry

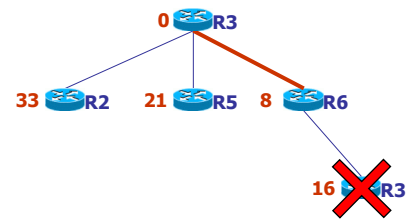


Src	Des	Net	Dist
R1	R2	b	13
R1	R4	a	13
R2	R1	b	13
R2	R4	c	12
R2	R5	e	11
R2	R3	f	33

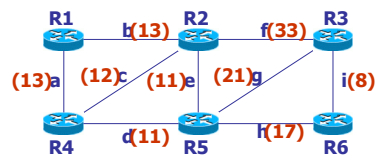
C.d.

R4	R1	a	13
R4	R2	c	12
R4	R5	d	11

Src	Des	Net	Dist
R5	R2	e	11
R5	R3	g	21
R5	R4	d	11
R5	R6	h	17
R6	R3	i	8
R6	R5	h	17



Algorytm Dijkstry

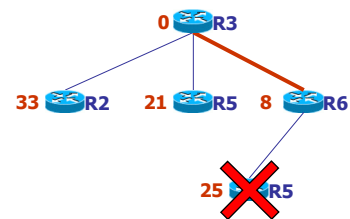


Src	Des	Net	Dist
R1	R2	b	13
R1	R4	a	13
R2	R1	b	13
R2	R4	c	12
R2	R5	e	11
R2	R3	f	33

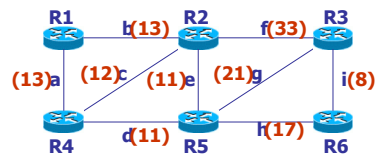
C.d.

R4	R1	a	13
R4	R2	c	12
R4	R5	d	11

Src	Des	Net	Dist
R5	R2	e	11
R5	R3	g	21
R5	R4	d	11
R5	R6	h	17
R6	R3	i	8
R6	R5	h	17



Algorytm Dijkstry

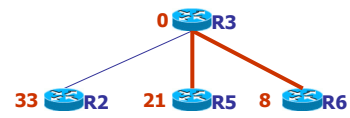


Src	Des	Net	Dist
R1	R2	b	13
R1	R4	a	13
R2	R1	b	13
R2	R4	c	12
R2	R5	e	11
R2	R3	f	33

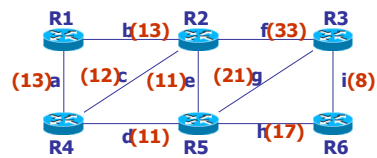
C.d.

R4	R1	a	13
R4	R2	c	12
R4	R5	d	11

Src	Des	Net	Dist
R5	R2	e	11
R5	R3	g	21
R5	R4	d	11
R5	R6	h	17
R6	R3	i	8
R6	R5	h	17



Algorytm Dijkstry

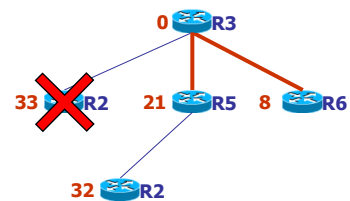


Src	Des	Net	Dist
R1	R2	b	13
R1	R4	a	13
R2	R1	b	13
R2	R4	c	12
R2	R5	e	11
R2	R3	f	33

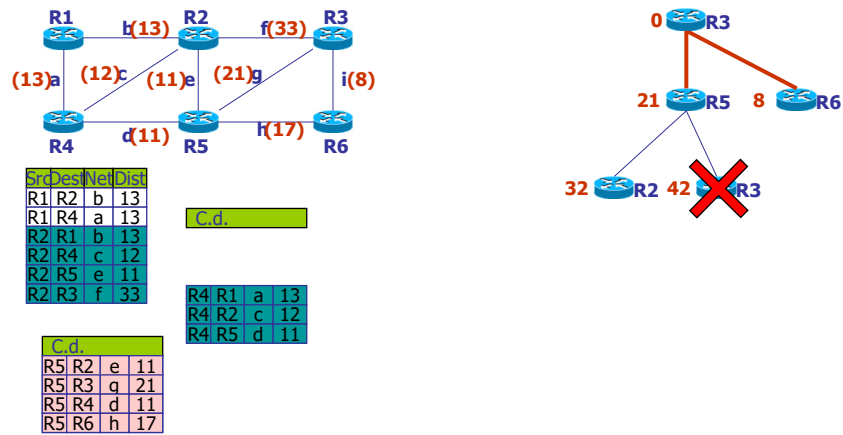
C.d.

R4	R1	a	13
R4	R2	c	12
R4	R5	d	11

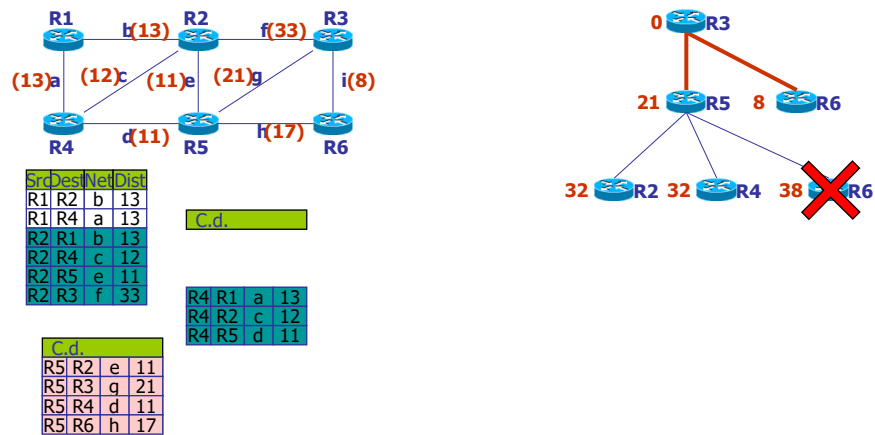
Src	Des	Net	Dist
R5	R2	e	11
R5	R3	g	21
R5	R4	d	11
R5	R6	h	17



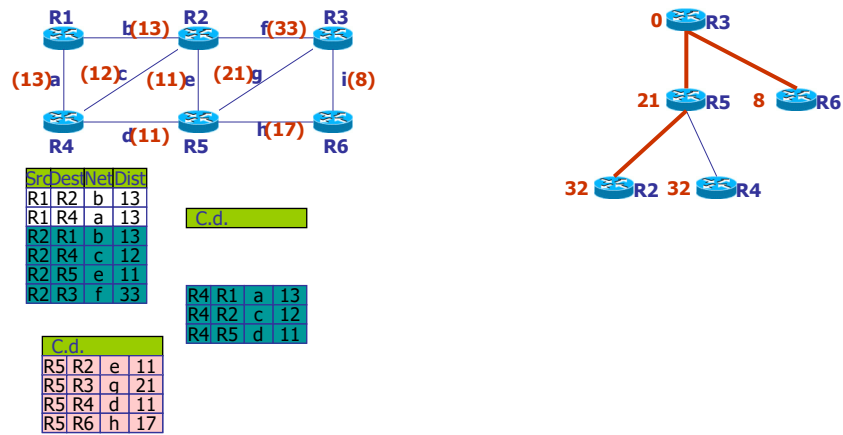
Algorytm Dijkstry



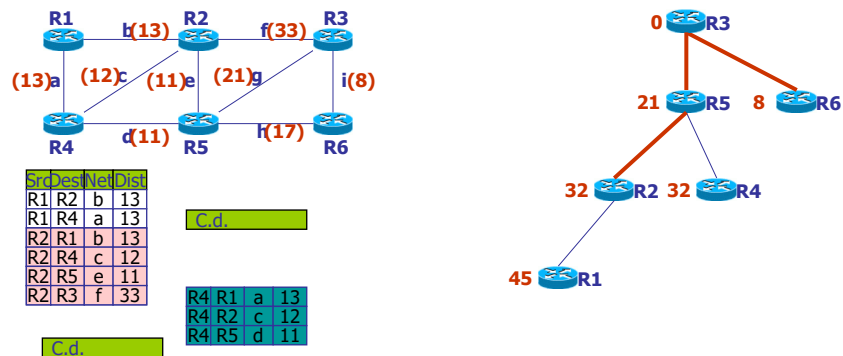
Algorytm Dijkstry



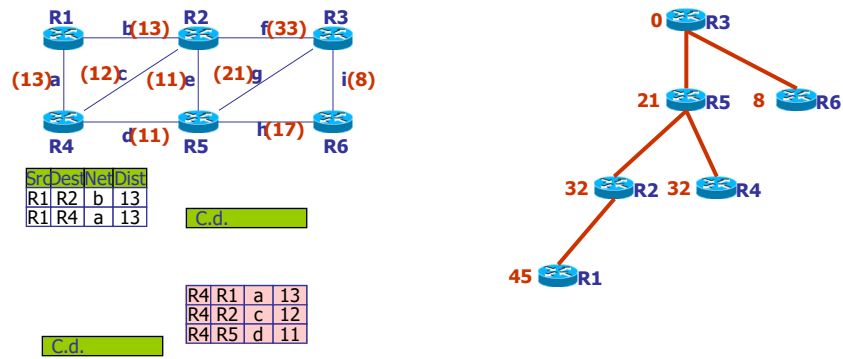
Algorytm Dijkstry



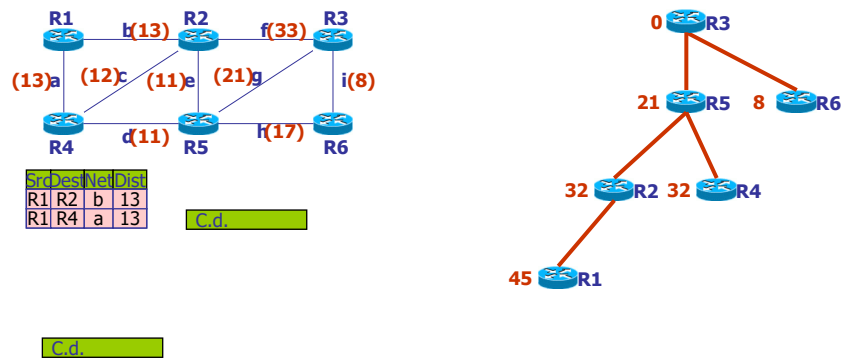
Algorytm Dijkstry



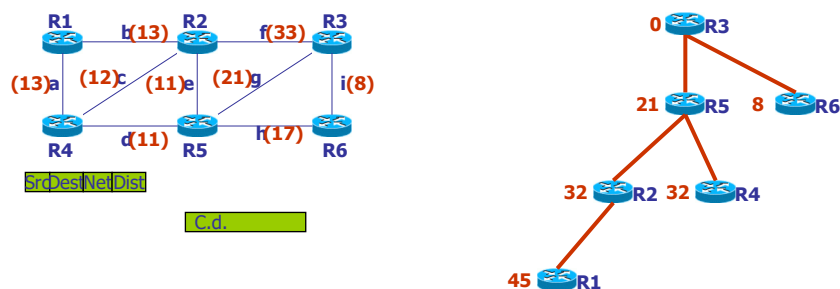
Algorytm Dijkstry



Algorytm Dijkstry



Algorytm Dijkstry



- Zostało stworzone pełne drzewo — na jego podstawie budujemy tablicę routingu
- Zawartość tablicy LSD nie jest usuwana!

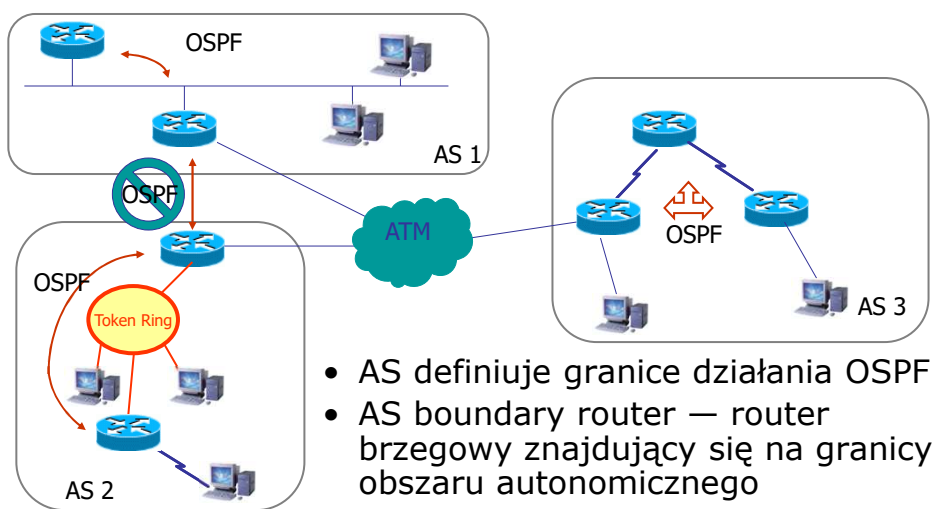
Protokół OSPF (Open Shortest Path First)

- OSPF jest protokołem IGP - *Interior Gateway Protocols* (routing w obrębie jednego systemu autonomicznego).
- Początek prac przez IETF w 1987
- Wersja OSPF opublikowana w 1991 (RFC 1247)
- Wersja OSPFv2 – RFC2328
- Wersja obsługująca IPv6 – OSPFv3 (RFC2740)
- Metryką jest koszt trasy:
 - 10Mbit/s = 10
 - 1Gbit/s (100Mbit/s) = 1
- Rozwiązanie bezklasowe, obsługuje CIDR i VLSM
- Otwarty standard
- Szybka konwergencja i brak pętli
- Hierarchiczny podział na obszary
- Brak zdefiniowanej maksymalnej liczby przeskoków

OSPF i organizacja sieci

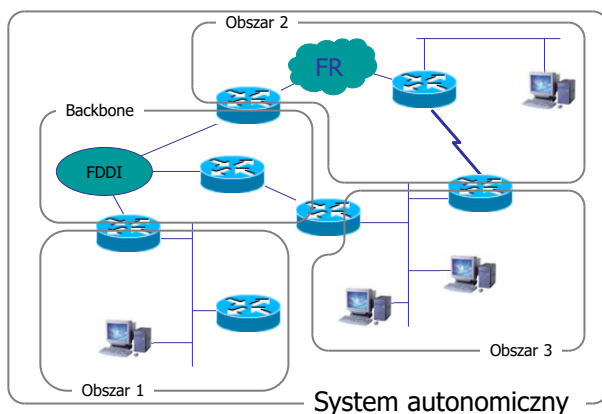
- Sieć Internet można traktować jako jedną całość
 - Każdy router wymienia informacje z każdym innym
 - Setki tysięcy systemów
 - Obciążenie sieci
 - Olbrzymi rozmiar tablic LSD
 - Bardzo długi czas obliczania algorytmu Dijkstry
- Wprowadzenie hierarchii
 - Systemy autonomiczne
 - Obszary

Systemy autonomiczne



Obszary OSPF

- AS może być zbyt duży



- Podział dowolny
 - Elementy obszaru muszą być połączone
- Wewnątrz obszaru OSPF działa normalnie
- Area border router — router brzegowy dla obszaru OSPF
 - Rozgłasza *summary LSA* o innych obszarach

Backbone

- Obszar o specjalnym znaczeniu
- Koncentrator dla wszystkich obszarów
- Wszystkie obszary są połączone z backbone
- W przypadku straty połączenia pomiędzy elementami istnieje możliwość skonfigurowania wirtualnego połączenia pomiędzy *backbone router*

Podsumowanie - protokoły stanu łącza

- Skomplikowane
- Bardziej wydajne
- Skalowalne
- Brak pętli
- Dobra zbieżność

Literatura

- *Routing in the Internet*, Christian Huitema, Prentice Hall PTR
- *IPng and the TCP/IP Protocols* Stephen A. Thomas
- *Zarządzanie sieciami IP za pomocą ruterów Cisco*, Scott M. Ballew, O'Reilly
- *Sieci komputerowe TCP/IP, t.1*, Douglas E. Comer, Wydawnictwa Naukowo-Techniczne
- RFC 1058 (RIP), RFC 2453 (RIP-2), RFC 1131 (OSPF-1), RFC 2178 (OSPF-2)