Routing

Słowo „routing” dwa znaczenia:
1. przekazywanie pakietów, ”forwarding”
2. modyfikowanie tablicy routingowej (tym się teraz zajmiemy...)

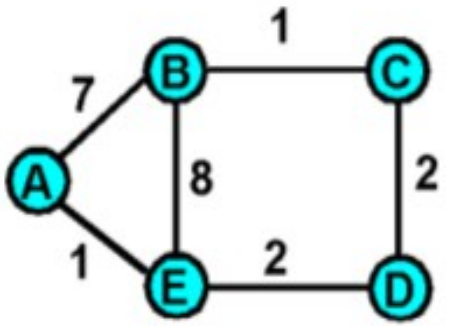
Uproszczony obraz „intersieci”: każda sieć fiz ma 2 routery, istnieje „koszt łącza”

Dwa podejścia: **DV** (Distance Vector) „wektor odległość”, **LS** (Link State) „stanu łącza”
opierają się na „algorytmie najkrótszych ścieżek między wsz parami wierz” (relaksacja)

Tablica routingowa danego węzła zawiera wiersze:
(węzeł_docelowy, sąsiad, odległość/metryka)

DV: każdy wierz wymienia z sąsiadami „tablicę routingową”
oraz aktualizuje swoją tablicę na podstawie otrzymanych informacji
(nowe lub lepsze trasy przez innego sąsiada/gw, itp.)
alg Bellman-Ford, relaksacja, „najkrótsze ścieżki z 1 źródłem do wszy wierz”
LS: każdy wierz sprawdza stan swoich połączeń z sąsiadami,
oraz rozsyła te informacje po całej sieci (czyli każdy wierz ma obraz całej sieci),
następnie każdy wierz lokalnie oblicza przez którego sąsiada prowadzi
najlepsza trasa do danego węzła (alg Dijkstry, relaksacja)

Routing – uproszczenie intersieci...



$D^E()$

destinat.	cost to destination via A	via B	via D
A	(1)	14	5
B	7	8	(5)
C	6	9	(4)
D	4	11	(2)

Uproszczony graf reprezentujący „intersieć”
Wierz grafu to routery (nie ma tu hostów!!)
Łącza mają wagi (mogą być =1), odległość jest ważona
Globalna macierz odległości (ważonych) z pkt widzenia wierz E
Kółkiem oznaczono minimalną trasę
Można ją łatwo przerobić na tabl routingową E...

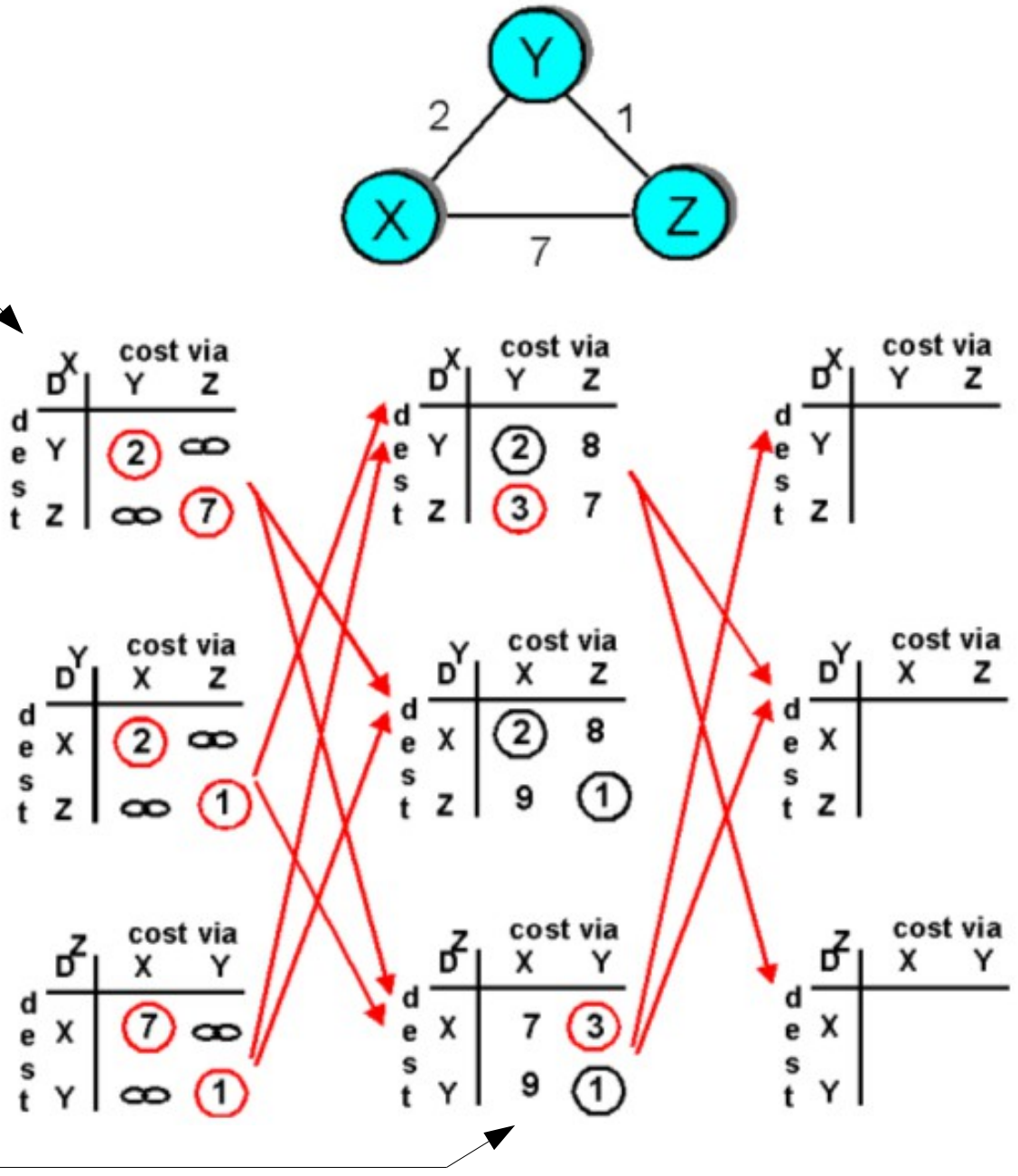
Routing DV

jak uaktualniamy tablice routingowe?

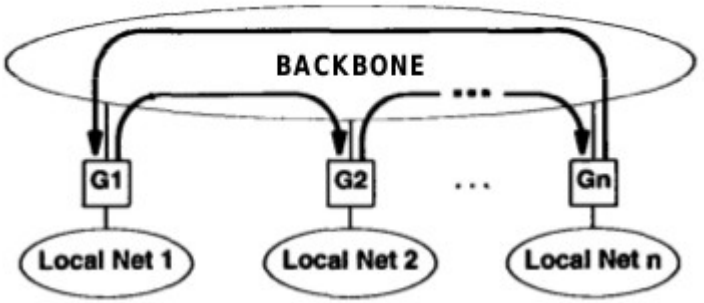
obraz pokazuje zmiany tabl routing
w czasie działania alg...
(tabl routing wyznaczają kółka!)

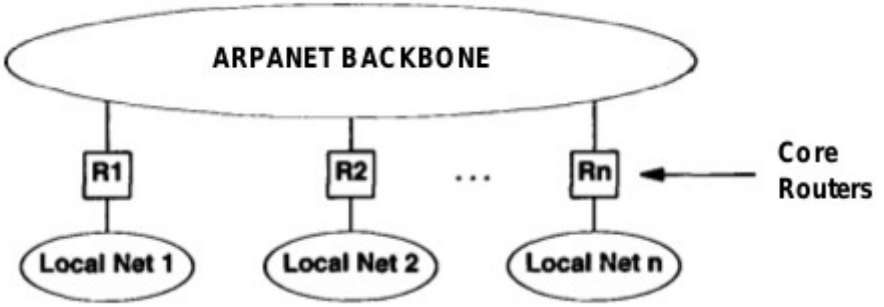
oznaczenie: D^Z(X,Y) to odległość
od Z do X via sąsiad Y
(Y jest sąsiadem Z)

D^Z(X,Y)=nieskoń;
dostalismy info że dist(Y,X)= 2
to wtedy uaktualniamy D^Z(X,Y)
D^Z(X,Y) := c(Z,Y) + dist(Y,X) = 3,
bo 3< nieskoń...



Routing - historia





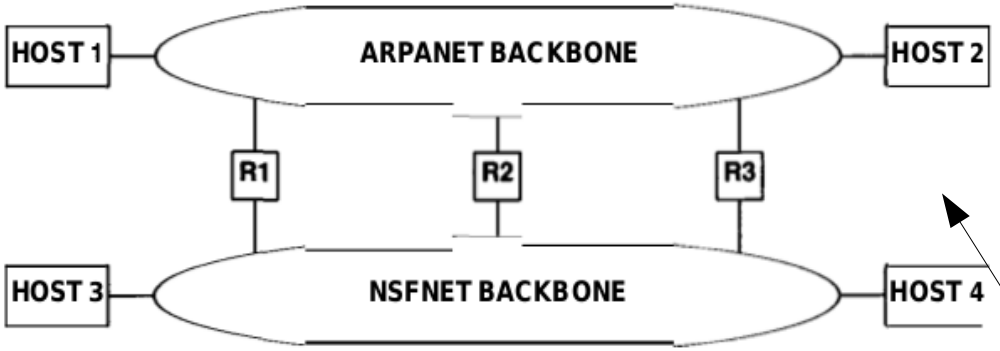
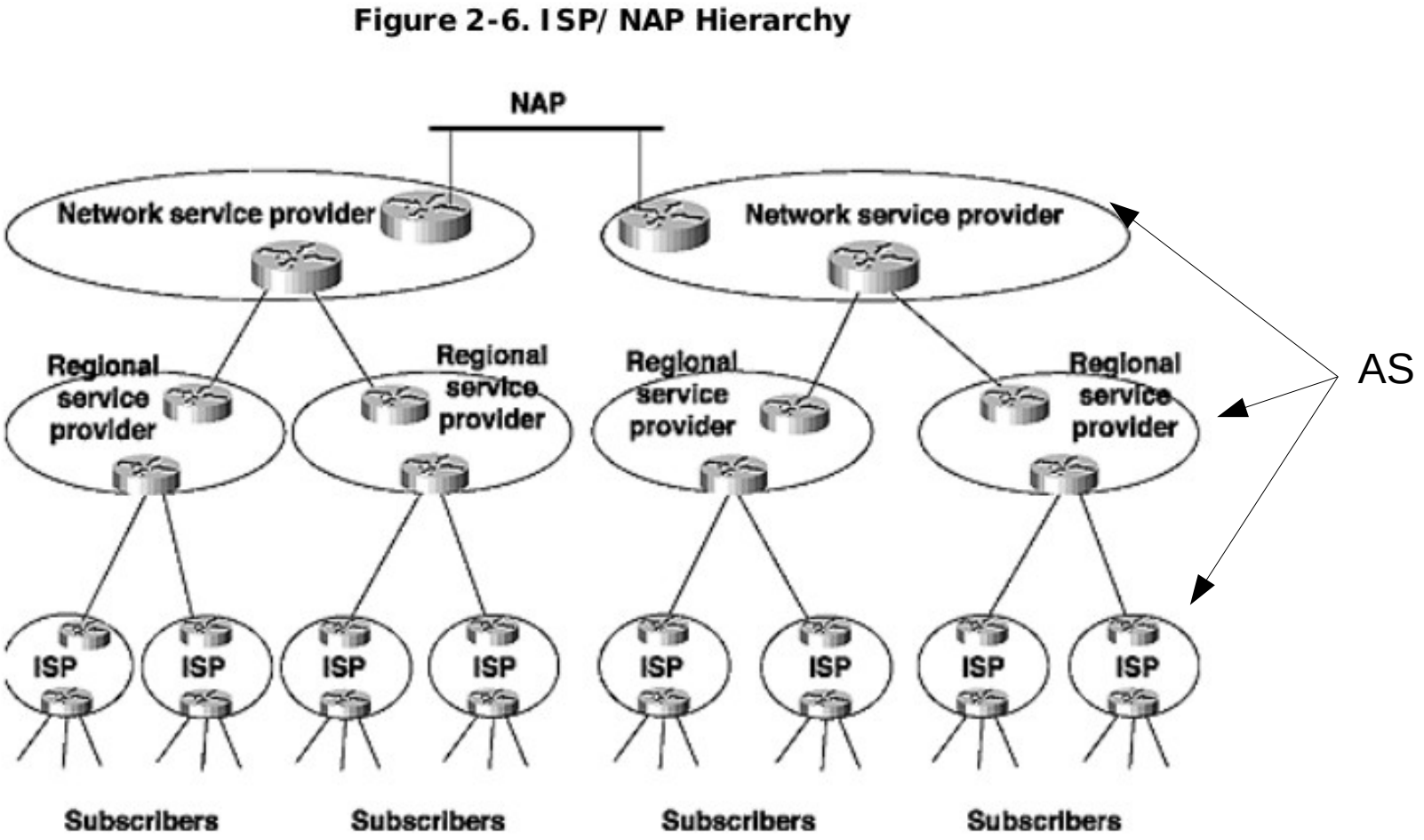


Figure 14.4 An example of peer backbones interconnected through multiple routers. The diagram illustrates the architecture of the Internet in 1989. In later generations, parallel backbones were each owned by an ISP.

System routerów podstawowych
Znają trasy do wszystkich sieci
Są także routery poboczne/wewn
(w sieciach „wewnętrznych”)

Dwie sieci szkieletowe
ARPANET, NSFNET

Routing - stan obecny



3 poziomy ISP (Internet Service Provider), NAP to obecnie IXP (Internet Exchange Point)

Routing i systemy autonomiczne (AS-y)

AS = ang. **Autonomous System**, zbiór sieci fizycznych (intersieć),
zarządzana przez jedna instytucję, ze wspólną „polityką routingową”,
z tym samym algorytmem routingu wewn

AS-y mają spec nr ASN (2 lub 4 bajty), nadawane przez RIR (dawniej IANA)
RIR nadaje też (publiczne adr IP, bloki adresów), osobne RIRy dla kontynentów (?)

Potrzebne dwa rodzaje routingu:

- **wewnętrzny**, wewnątrz AS, typu DV (np. RIP) lub LS (np. OSPF(?))
- **zewnętrzny**, między AS-owy, dostarcza info o osiagalności sieci przez sąsiednie AS-y... (BGP-4, dawniej „EGP”, uwaga na EGP/ nazwę ogólną !!)

Tablica routingowa routera pochodzi z 2 źródeł: od routingu wewn ORAZ zewn...

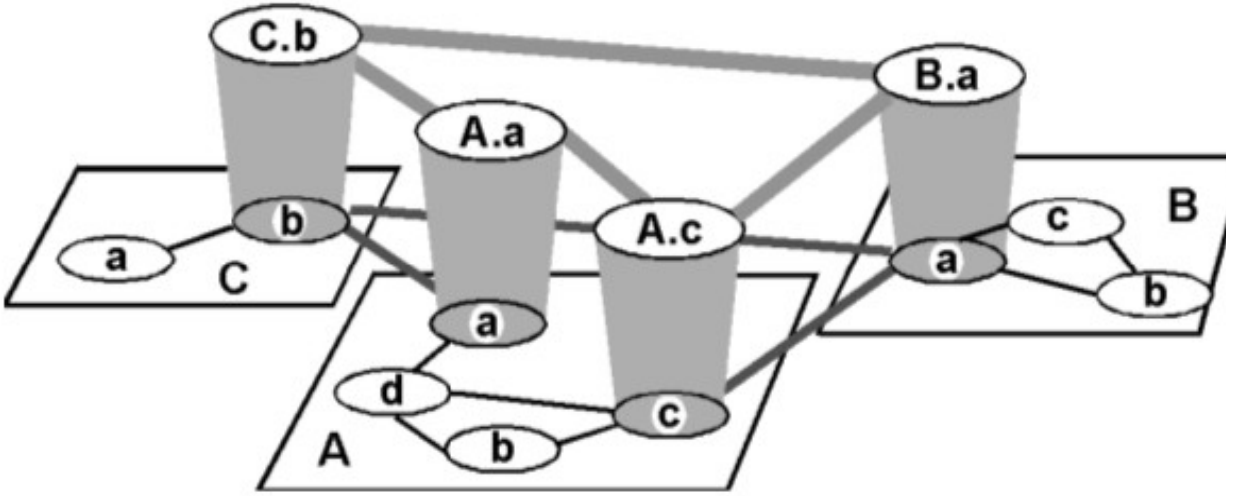


Figure 4.3-1: Intra-AS and Inter-AS routing.

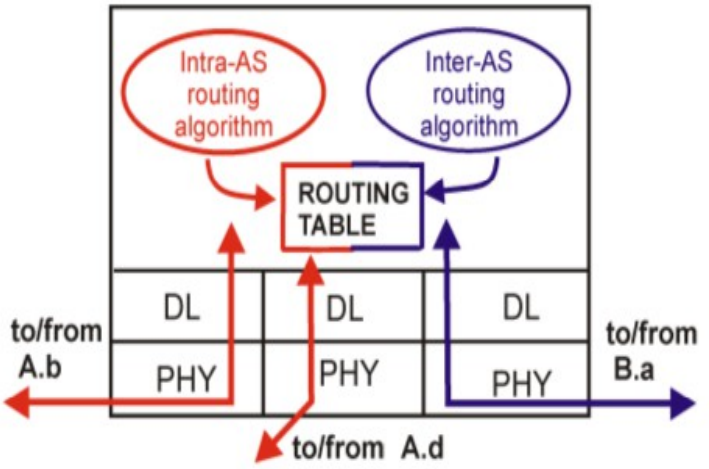
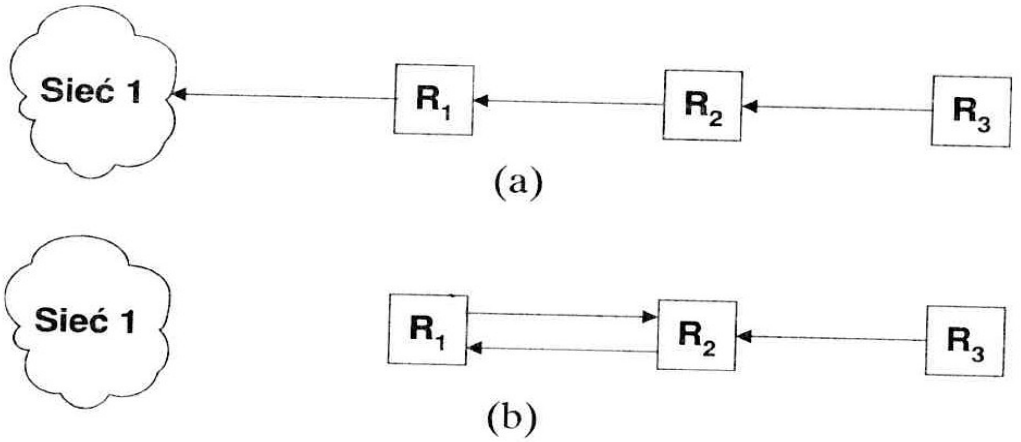


Figure 4.3-2: Internal architecture of gateway router A.c

Routing wewn/ RIP

Algorytm routingu typu DV; routing dynamiczny; używa UDP port 520
RIPv1 (IPv4 klasowe), RIPv2 (IPv4 bezklasowe), RIPng (Ipv6)
Węzły RIP akywne i pasywne
Węzły aktywne co 30sek wymieniają tabl routingową z sąsiadami
Sąsiad? Węzeł RIP w sieci bliskiej; broadcast (same „1” w adr ip dst)
Koszt trasy = liczba etapów, tzn waga połączenia = 1
Max średnica = 15, koszt 16 oznacza „nieskończoność”
Implementacje: routed, gated, **quagga** (nowsze!), gated i quagga zaw też inne prot!!
Problemy RIP...

1. nie wykrywa pętli
2. max średnica AS = 15
3. „powolna zbieżność” oraz powodowane przez nią kłopoty:



Co się dzieje dalej ???
R1 wprowadzony w błąd...

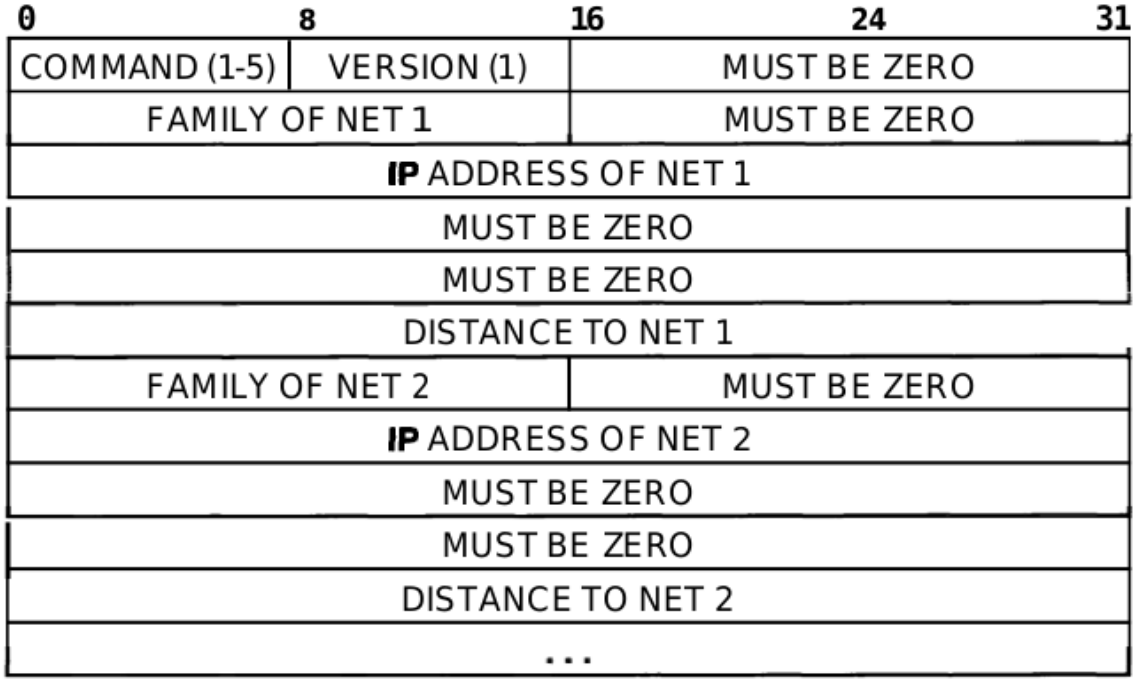
Rys. 16.4. Problem powolnej zbieżności. W przypadku (a) 3 rutery znają trasę do sieci *1*. W przypadku (b) znika połączenie z siecią *1*, ale *R2* nadal je oferuje, powodując powstanie pętli

Routing wewn/ RIP

Jak unikać oscylacji między trasami o tym samym koszcie?
zasada: zmieniać trasę tylko jeśli pojawi się trasa o „<” koszcie...

Rozwiązania problemu „powolnej zbieżności”:
1. „uaktualnianie z podzielonym horyzontem” nie propaguje się informacji przez interf, przez który się tą informację otrzymało
2. „wstrzymanie” zmian; przez 60sek ignoruje się informacje o nieosiągalności danej sieci
3. „metoda odtrutki” po informacji o zniknięciu połączenia zachowuję się informacje o nim ale z wagą 16 (nieskończoność)

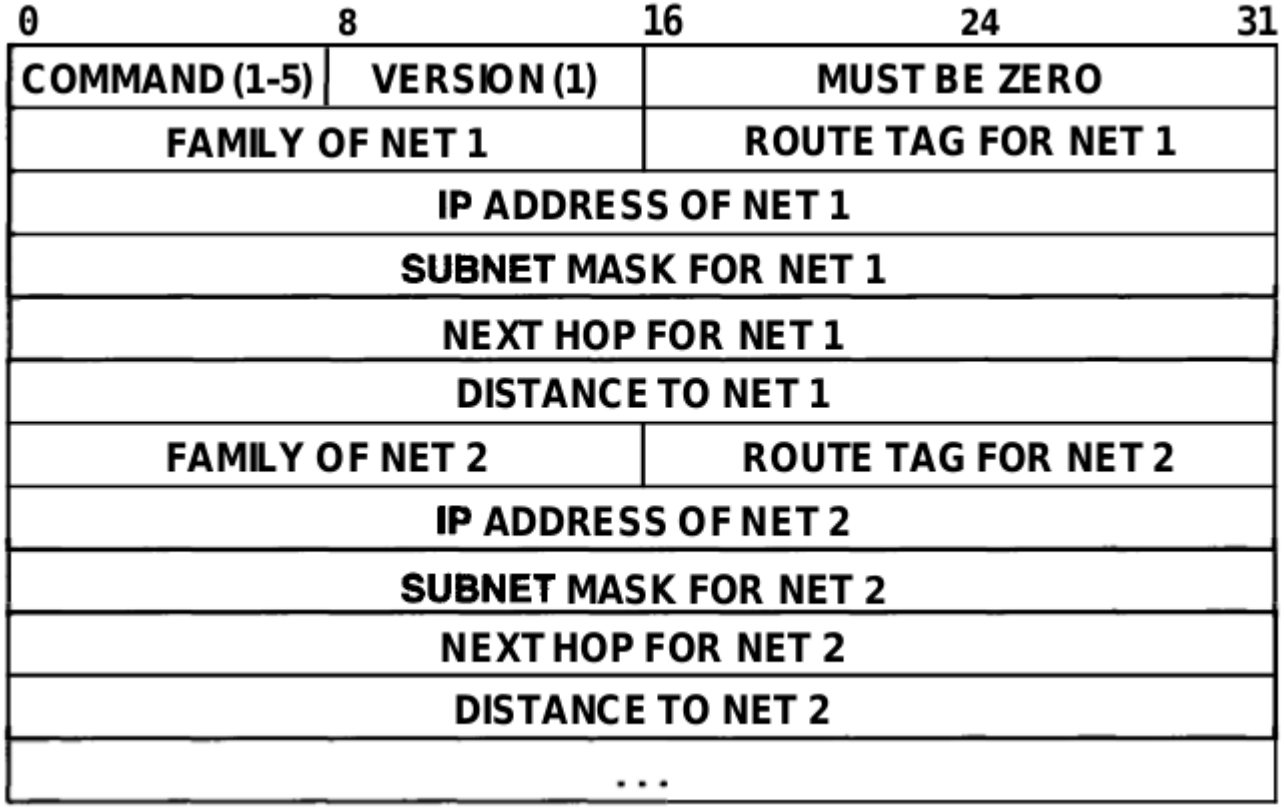
Format komunikatów RIP v1:



Command	Meaning
1	Request for partial or full routing information
2	Response containing network-distance pairs from sender's routing table
3	Turn on trace mode (obsolete)
4	Turn off trace mode (obsolete)
5	Reserved for Sun Microsystems internal use
9	Update Request (used with demand circuits)
10	Update Response (used with demand circuits)
11	Update Acknowledge (used with demand circuits)

Routing wewn/ RIP

Format komunikatu RIP v2:



Routing wewn/ OSPF

- działa nad IP, proto 89, RFC 2328, używany przez „dużych” ISP

- podział AS na obszary, w tym 1 obszar szkieletowy...

- w zasadzie LS (wewn obszarów), między obszarami DV...
LS: rozgłasza zmiany o stanie łaczy wsz routerom swojego obszaru...

- wiele tras do danego celu/sieci wybieranych na podst TOS w nagł ip! oraz adr ip dst (?)
- „load balancing” gdy jest wiele tras do tego samego celu

- komunikacja między routerami wymaga uwierzyteln (inaczej niż RIP)

- trasy do konkretnych maszyn są możliwe
- obsługuje CIDR (adr ip bezklasowe)
- wykorzystuje broadcast, o ile możliwe

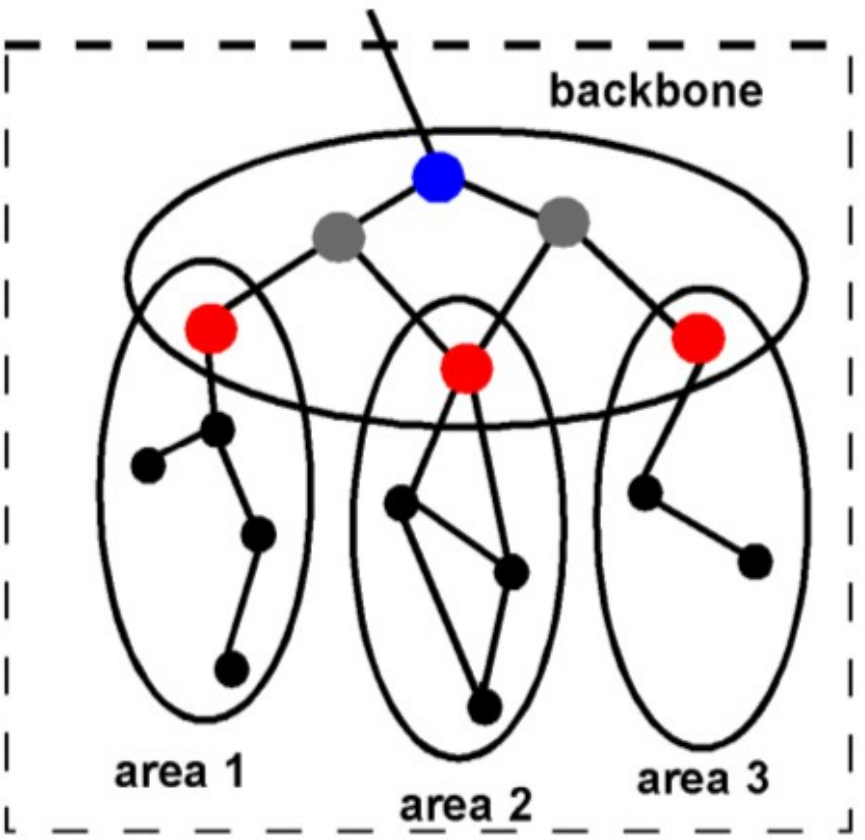


Figure 4.5-7: Hierarchically structured OSPF AS with four areas.

Niebieski – router graniczny do innego AS
Czerwony – router brzegowy obszaru (area)

Routing zewn/ EGP i BGP-4

Odcinki trasy między hostami h1 i h2 pochodzące od routingu wewn i zewn...

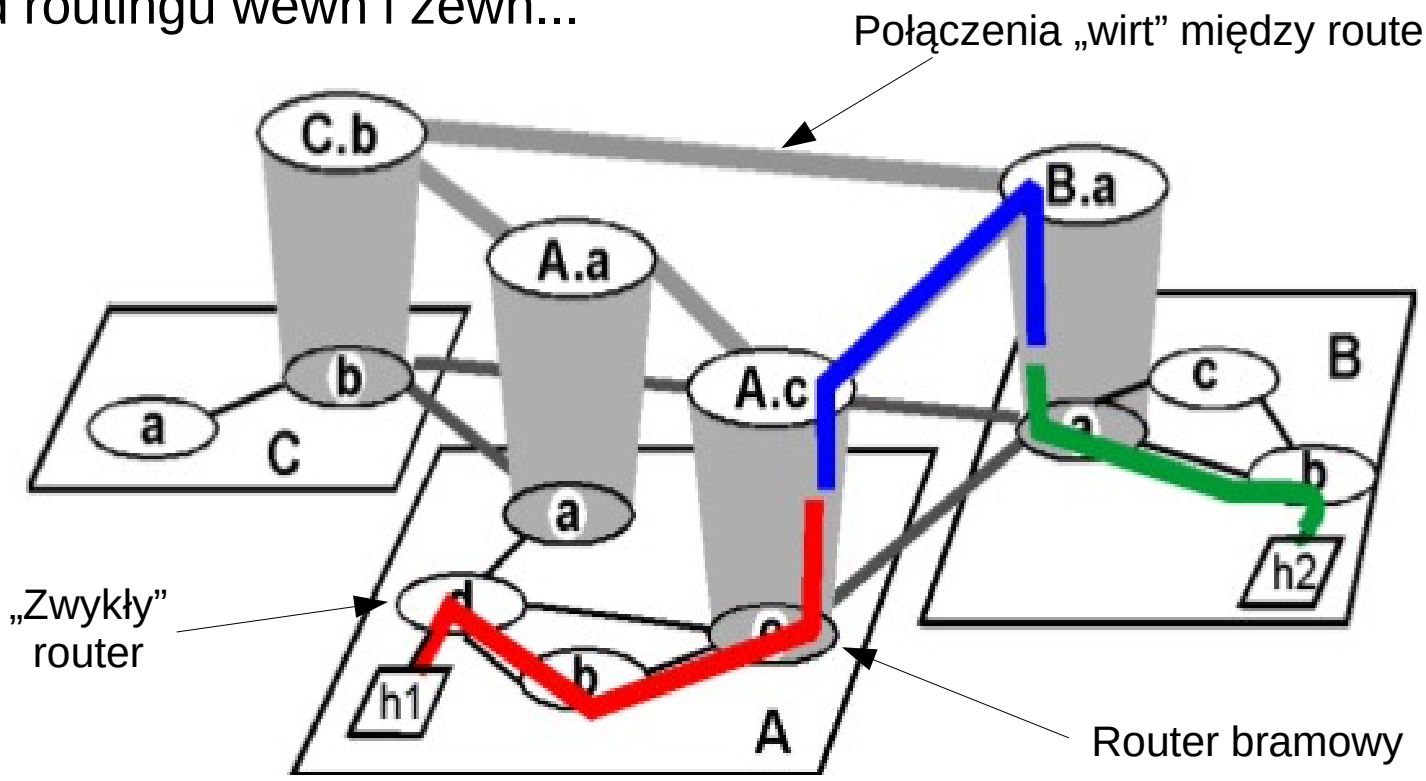
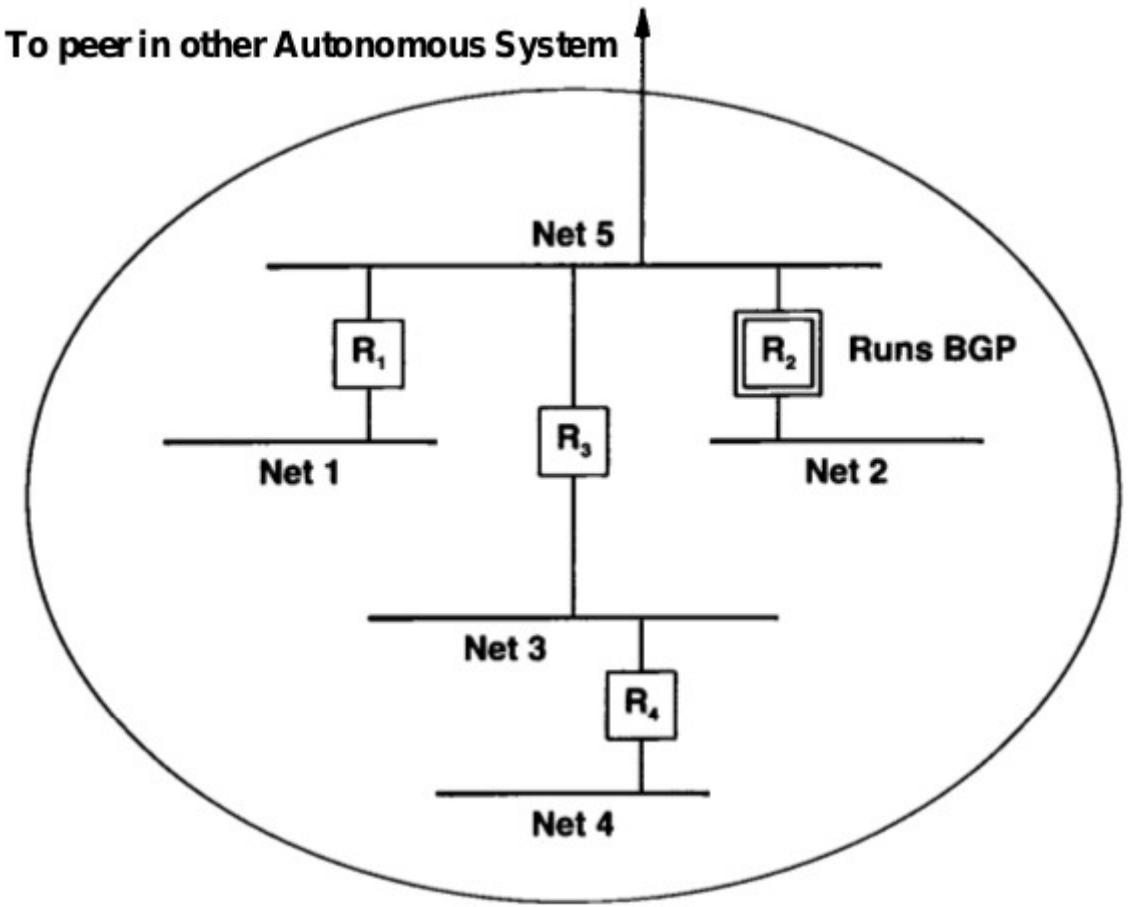


Figure 4.3-3: The route from A.d to B.b: intra-AS and inter-AS path segments.

Routing zewn/ EGP i BGP-4



- „routing hierarchiczny”:
podział na AS vs krótki prefiks nr sieci

- „EGP” vs EGP (nazwa ogólna!)

- „EGP” przestarzały prot rout zewn;
router EGP informuje tylko
  o sieciach w swoim AS !!!
  „zakaz informowania o obcych”
wymagany jest „syst routerów podst”,
  które znają wszystkie sieci...
brak info o odległości do obcych sieci
  które są znane w AS...

- komunikty „EGP” z aktualnymi trasami
zawieriaja odległości między podstwową
siecią (Net 5) a innymi sieciami AS-u...

- ”BGP-4” likwiduje te ograniczenia...
informuje sąsiednie routery BGP
o sieciach **osiągalnych** przez własny AS...

Routing zewn/ BGP-4

- opisany w RFC 1771, 1772, 1773; działa nad prot TCP, port 179

- cele BGP:
  1. uzyskanie z sąsiednich AS informacji o osiągalności sieci
  2. przekazanie tych info wszystkim routerom własnego AS
  3. wyznaczanie dobrych tras do sieci, na podstawie info o osiągalności sieci oraz zasady zdef dla AS („polityka”)

- sesje BGP: iBGP (wewn 1 AS), eBGP (między różnymi AS-ami)
sesje eBGP służą do wymiany info o osiągalnośći sieci poprzez sąsiedni AS...

- BGP jest prot „wektor ścieżka” (vs „wektor odległość” w RIP)
dla docelowej sieci podaje się nie tylko odległość ale trasę/ścieżkę...
router BGP przekazują sobie trasy do sieci (nie tylko info o osiągalności)...
są to m.in. numery ASN przez które przebiega trasa...
jest więcej informacji o różnych trasach prowadzących do celu...
można lepiej podejmować decyzję...