MPLS

Co to jest?
Z czym to gryźć?
Jak i po co myśleć o mechanizmach MPLS we własnej sieci?

Bartłomiej Anszperger
B.Anszperger@cisco.com

PLNOG, Kraków, wrzesień 2011

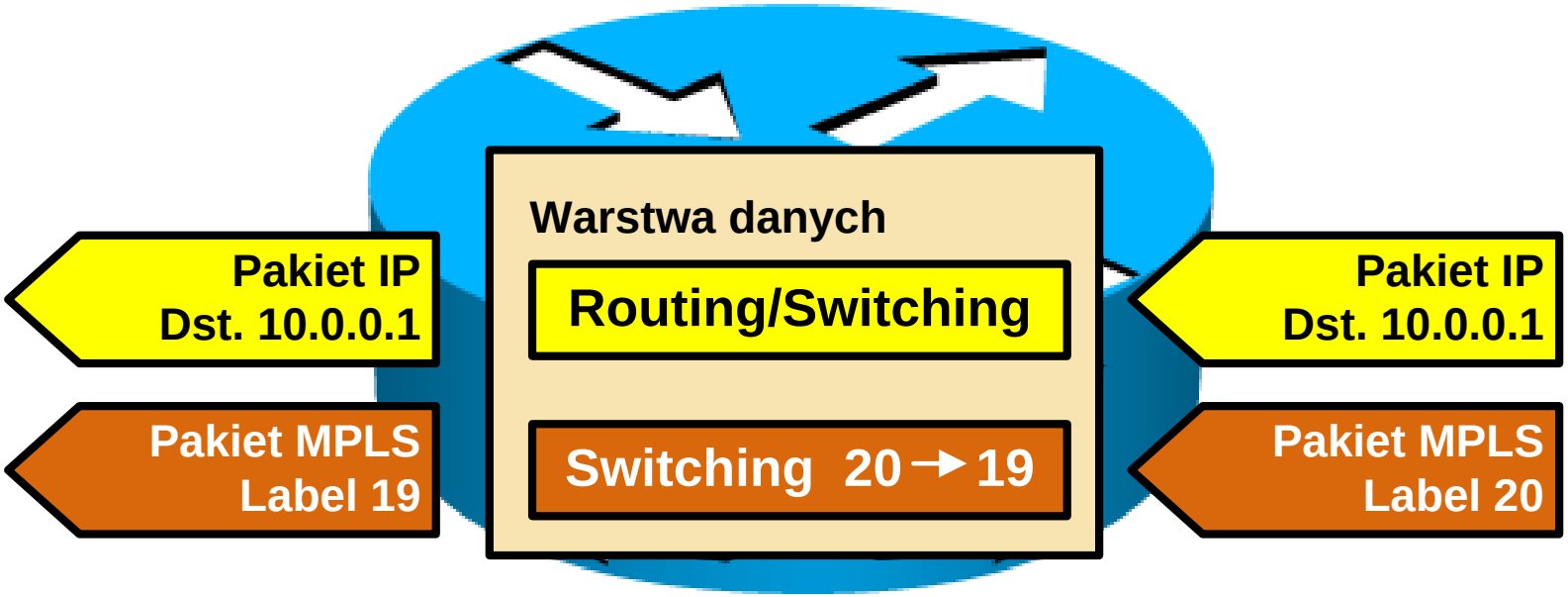
O czym będziemy dzisiaj rozmawiać?

- Co to ten MPLS?
- Mroki historii czyli MPLS 10 lat temu...
- Po co mi MPLS w sieci?
- Czy naprawdę nie ma alternatyw dla MPLS?
- Q&A

Co to ten MPLS?

Co to ten MPLS?

- **M**ulti **P**rotocol **L**abel **S**witching
- Przełączanie w możliwie najprostszy sposób (np. w oparciu o wartość czyli tzw. etykietę = najprostsza możliwa enkapsulacja)



Budowa etykiety MPLS

Etykieta MPLS

COS/EXP = Class of Service: 3 Bity; S = Bottom of Stack; TTL = Time to Live

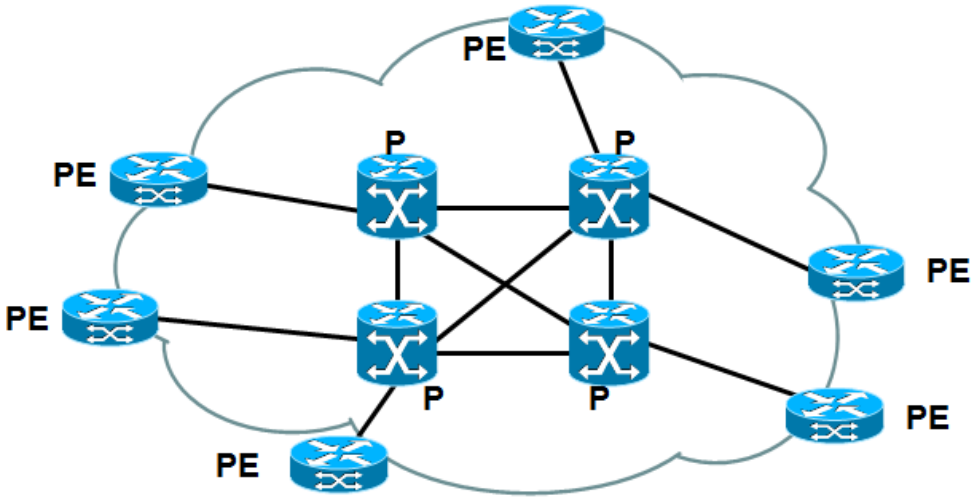
Enkapsulacja MPLS

Nagłówek PPP (Packet over SONET/SDH)

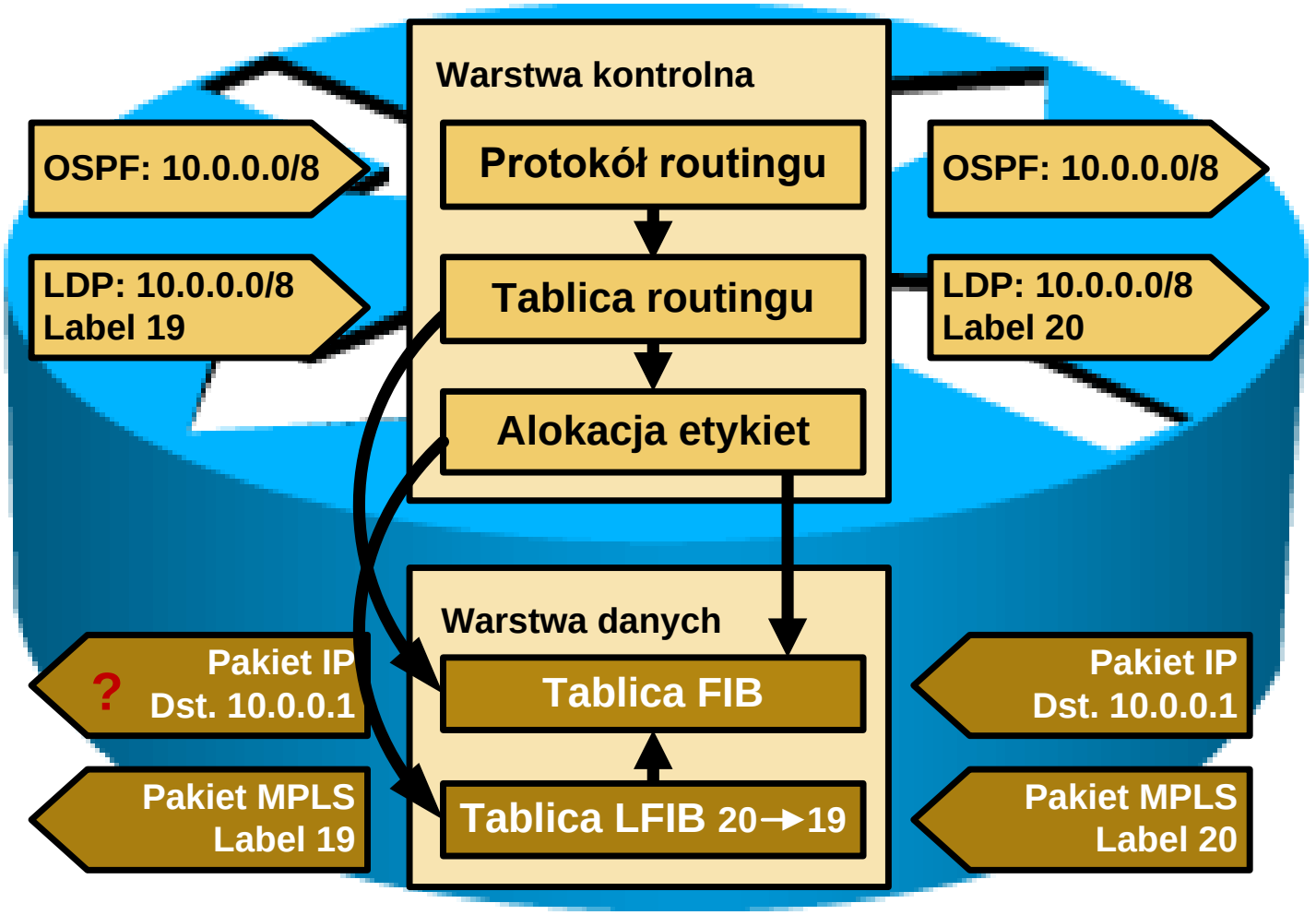
Nagłówek LAN MAC

Pojęcia wokół MPLS

- Typy urządzeń
  - Routery P (Provider) = LSR (Label Switching Routers)
  - Routery PE (Provider Edge) = Edge LSR
- Protokoły
  - IGP: OSPF, IS-IS, EIGRP
  - Label Distribution Protocol (LDP)
  - Multiprotocol BGP
  - Resource Reservation Protocol (RSVP)
- MPLS
  - Forwarding Equivalence Class (FEC)
  - Etykieta MPLS (ang. Label)
  - VRF (Virtual Routing and Forwarding)
- Płaszczyzny działania MPLS
  - Warstwa kontrolna
  - Warstwa danych

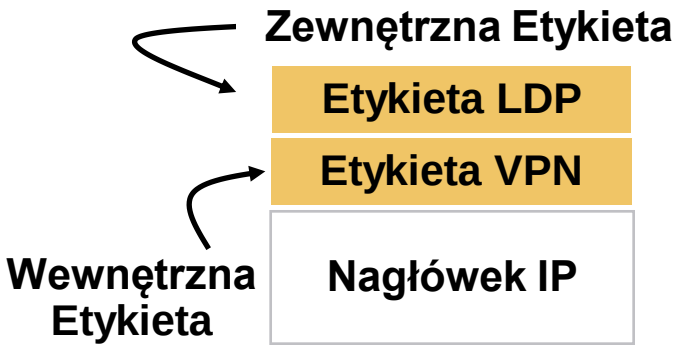


Przepływ danych w MPLS



Stos etykiet

- Może być więcej, niż jedna etykieta
- Zewnętrzna etykieta jest używana do przełączania w sieci MPLS.
- Ostatnia etykieta ma zaznaczony bit Bottom of Stack.
- Usługi z więcej, niż jedną etykietą:
  - MPLS VPN
  - MPLS Traffic Engineering oraz Fast Reroute
  - MPLS VPN poprzez MPLS-TE
  - AToM
  - CsC



Mroki historii czyli MPLS 10 lat temu...

Kiedy byłem młody...

- Zdarzyło mi się pisać o MPLS (2000)

NIE

TAK/NIE

TAK/NIE

TAK/NIE

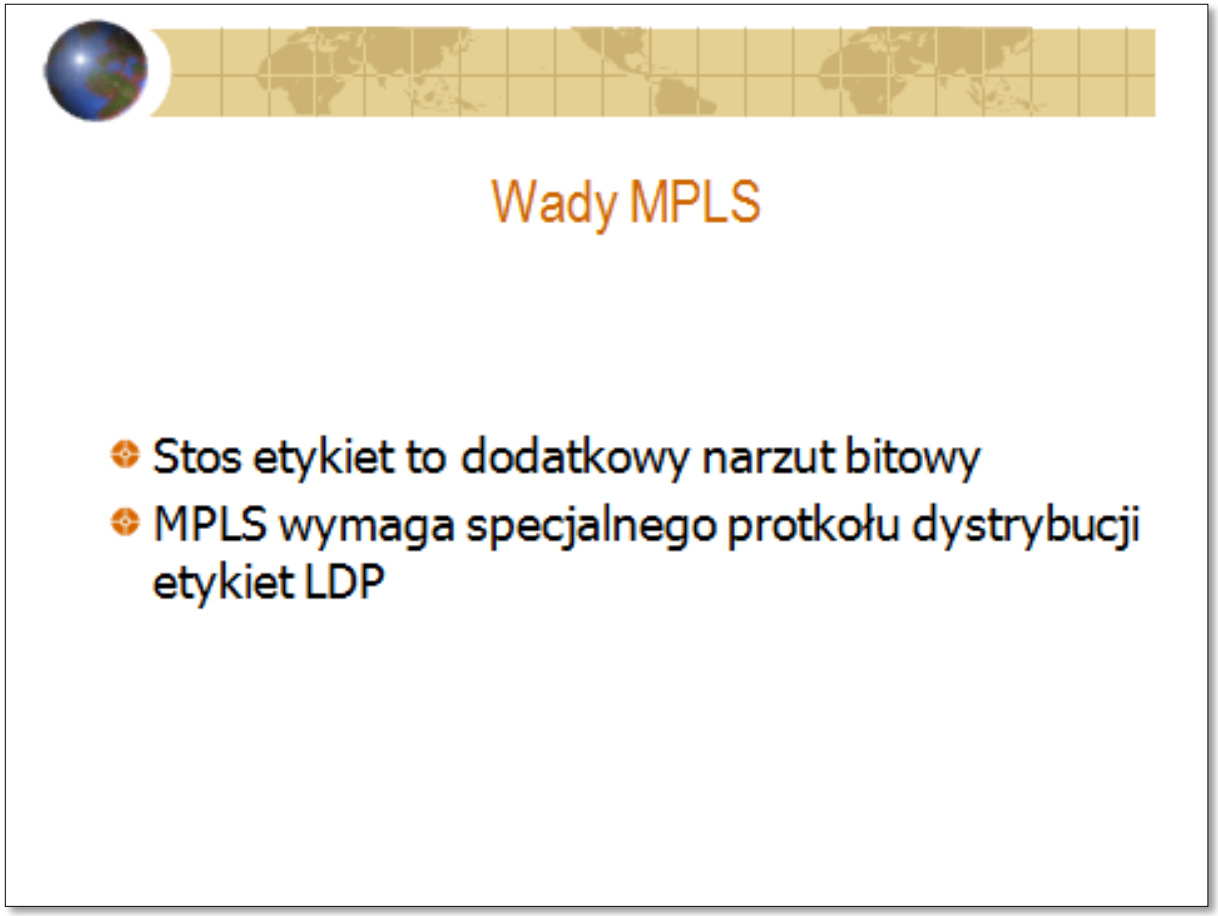
NIE

TAK

TAK

MPLS 10 lat temu

- Już wówczas mpls prawie nie miał wad ;-)



TAK

TAK

MPLS 10 lat temu (2)

- MPLS a ATM

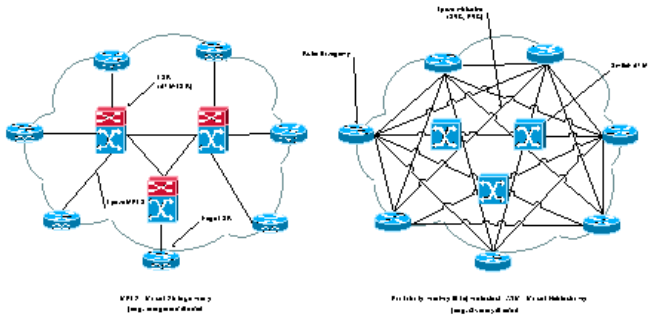
MPLS a sieci ATM

- Sieci ATM są wykorzystywane szeroko jako sieci szkieletowe u ISP (np. sieć WARMAN)
- Brak „naturalnej" metody przenoszenia bezpołączeniowego protokołu IP:
  - Classical IP over ATM
  - LANE
  - MultiProtocol Over ATM
- MPLS (IP+ATM) wydaje się być odpowiednim rozwiązaniem

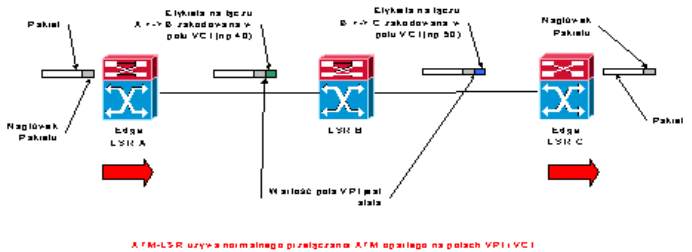
MPLS a sieci ATM (2)

- Połączenie wydajności i łatwości zarządzania ruchem (warstwa II modelu OSI) z skalowalnością i modyfikowalnością routingu warstwy III-ciej
- Możliwość „współistnienia" obu rozwiązań na pojedynczym urządzeniu
- Brak czasochłonnej procedury zestawiania połączeń
- Urządzenia ATM LSR biorą udział w wymianie informacji routingowej (model zintegrowany)

Model Zintegrowany a Nakładkowy



MPLS w ATM (2)



Po co mi MPLS w sieci?

Uzasadnienia do wdrożenia MPLS

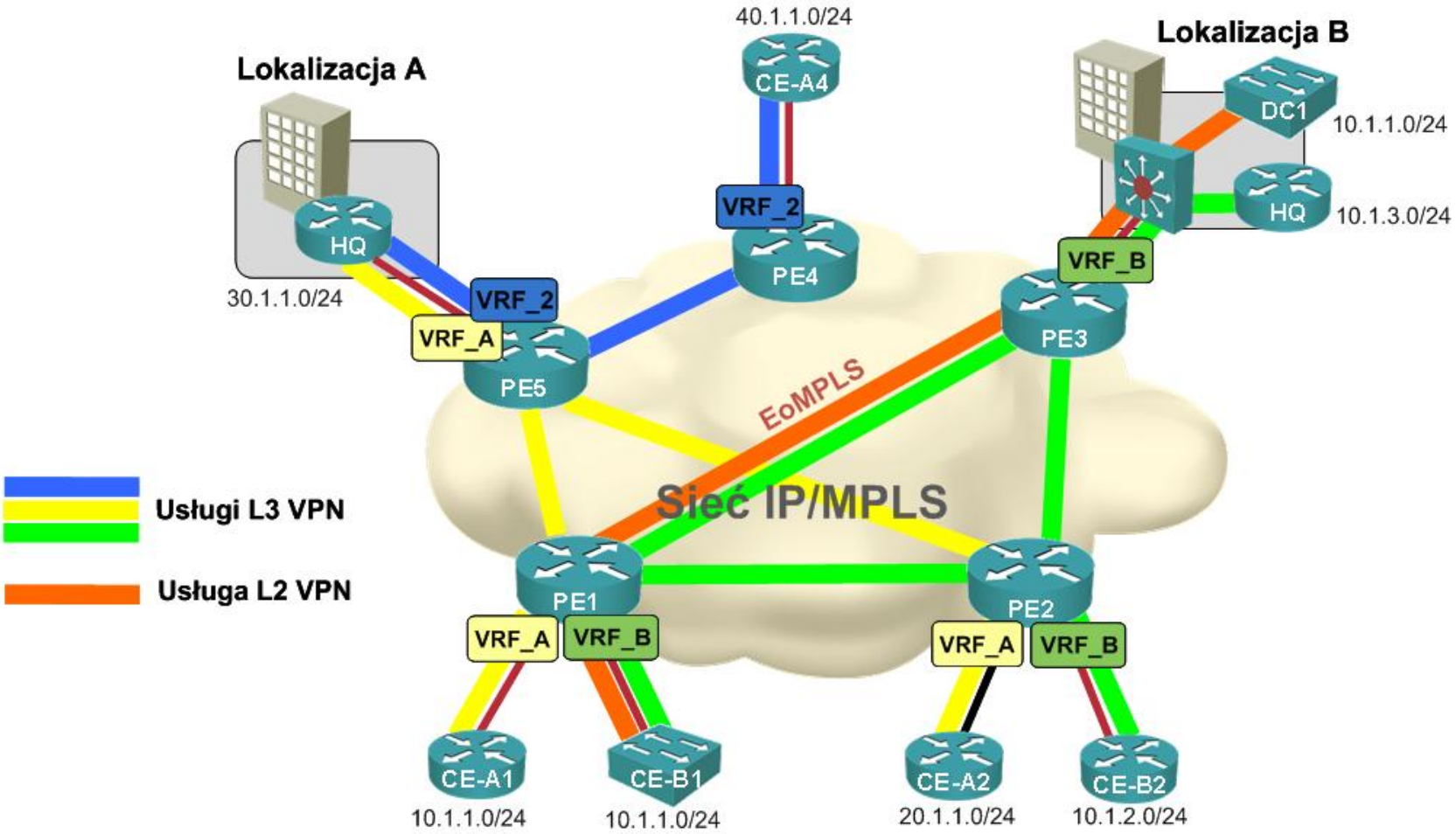
- Urządzenia szybciej przełączają pakiety MPLS, niż pakiety IP. **NIE**
- Lepsza realizacja QoS w sieci operatorskiej. **NIE**
- Świadczenie usługi p2p L2 VPN. **Niekoniecznie**
- Świadczenie usługi L3 VPN. **Niekoniecznie**
- Niski koszt implementacji usług. **NIE**
- Konsolidacja różnych sieci (ATM, FR, Ethernet) **TAK**
- Sieć wielousługowa L2/L3 VPN. **TAK**
- Szybka konwergencja sieci **TAK**
- Brak BGP na urządzeniach szkieletowych **TAK**

Inne zalety MPLS

- Większa elastyczność w zarządzaniu usługą L2.
- Brak konieczności uczenia się adresów MAC.
- Tworzenie skalowalnych sieci VPN pomiędzy operatorami z zachowaniem separacji danych.
- Zarządzanie ruchem w sieci za pomocą MPLS Traffic Engineering (MPLS-TE).
- Wysoka dostępność sieci: MPLS-TE, LDP FRR.
- Świadczenie usług IPv6 bez konieczności uruchamiania stosu IPv6 na routerach rdzeniowych.

Istnieją inne technologie umożliwiające realizację powyższych założeń, ale siłą MPLS jest to że łączy te zalety w jeden system.

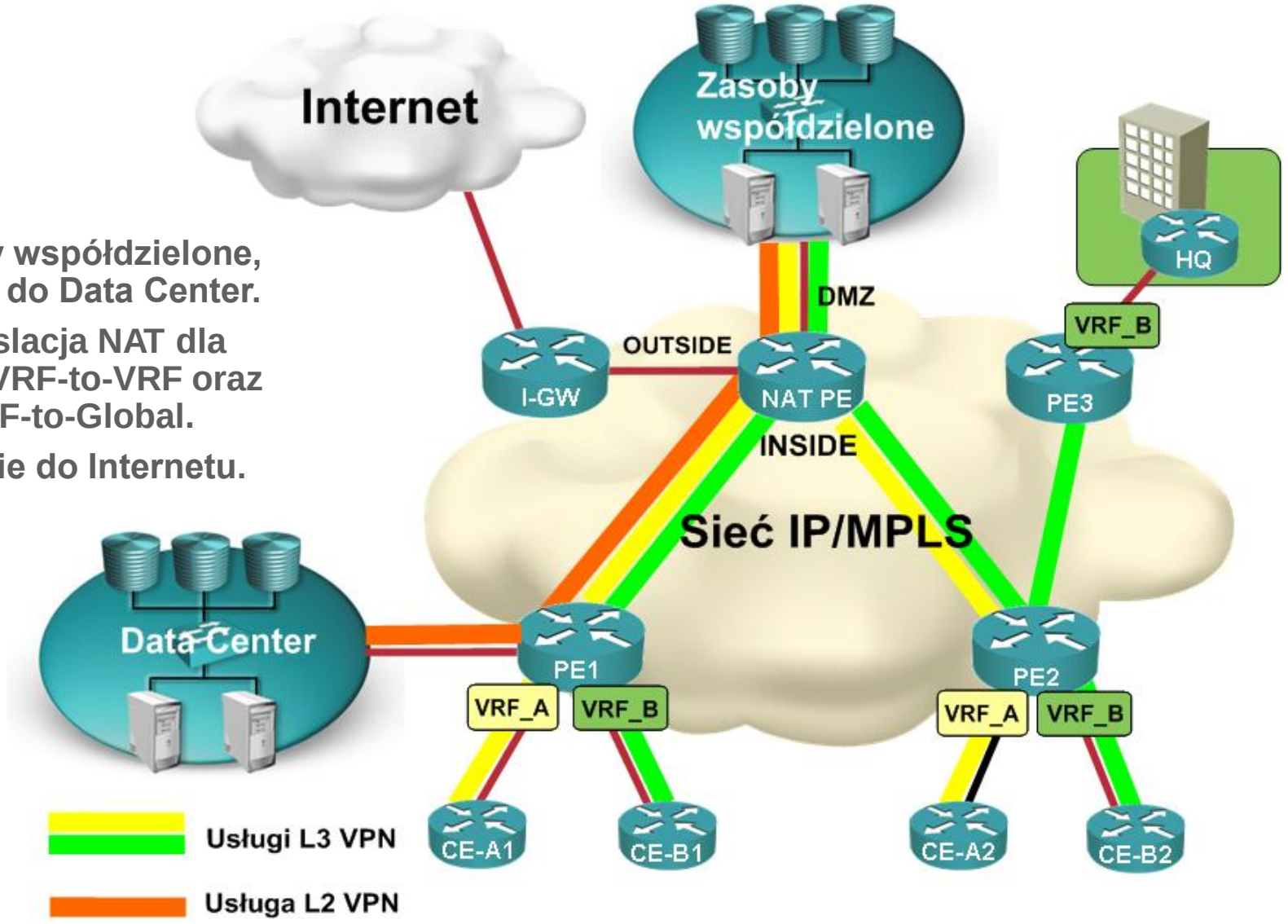
Zaleta 1: Konsolidacja usług



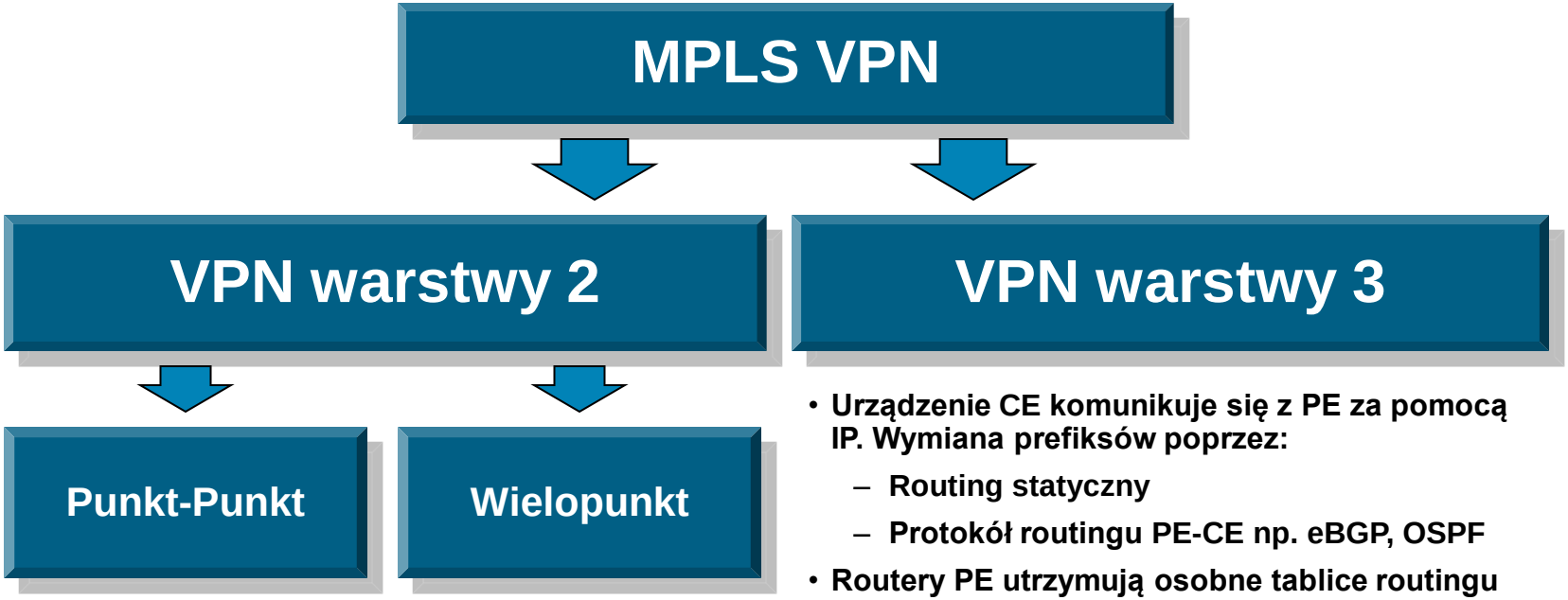
1. Jednoczesne świadczenie usług VPN warstwy 2 oraz 3 (p2p, m2m).
2. Separacja, wirtualizacja, elastyczność.

Zaleta 2: Zasoby współdzielone

1. Zasoby współdzielone, dostęp do Data Center.
2. Translacja NAT dla ruchu VRF-to-VRF oraz VRF-to-Global.
3. Wyjście do Internetu.



Model sieci MPLS VPN



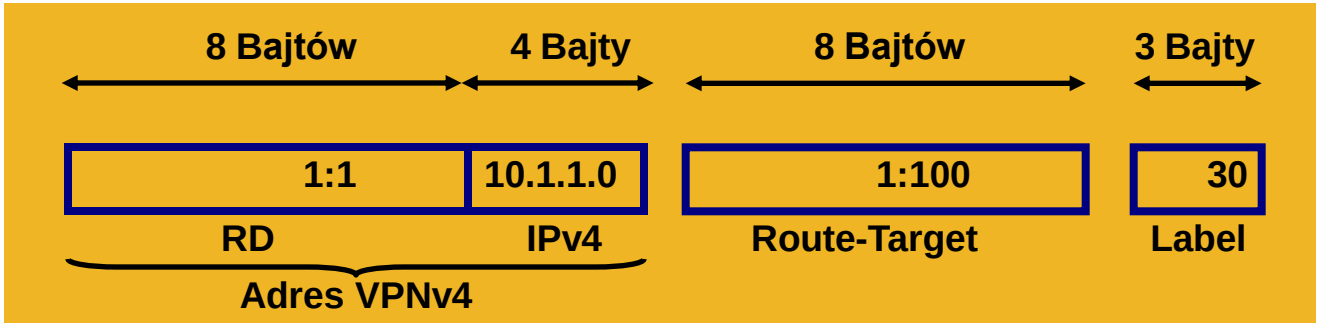
- Połączenie dwóch CE w ramach tunelu p2p.
- Technologia dostępowa ETH, FR, ATM
- Routing CE-CE. Urządzenia PE są transparentne dla ruchu między CE.

- Urządzenia CE mogą uzyskać bezpośrednie połączenie full-mesh między sobą w ramach VPN.
- Routing CE-CE. Urządzenia PE są transparentne dla ruchu między CE.

- Urządzenie CE komunikuje się z PE za pomocą IP. Wymiana prefiksów poprzez:
  - Routing statyczny
  - Protokół routingu PE-CE np. eBGP, OSPF
- Routery PE utrzymują osobne tablice routingu dla tras rozsyłanych przez CE.
- Routery PE operatora biorą udział w routingu, nie są transparentne dla przesyłanego ruchu.

Warstwa kontrolna MPLS VPN

Do obsługi MPLS VPN potrzebny jest Multi-Protocol BGP



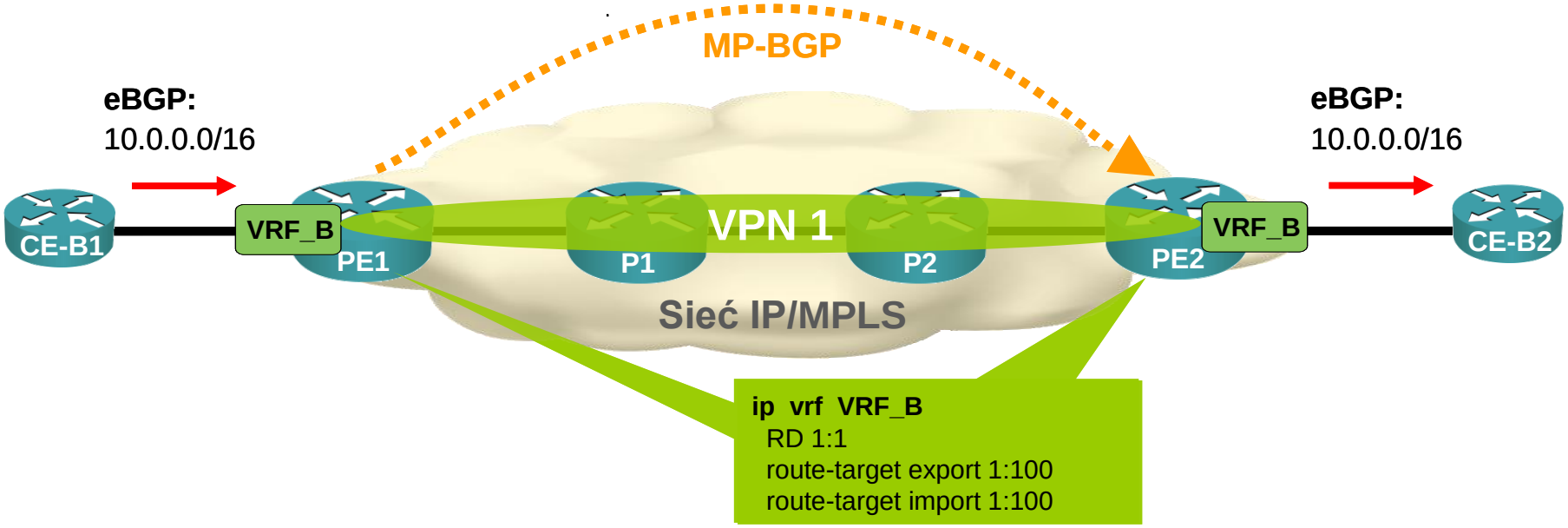
Komunikat MP-BGP typu UPDATE

Prefiksy rozgłaszane przez PE mają przydzielone:

- Route Distinguisher (RD)
- Route Target (RT)
- Etykietę

Informacja rozgłaszana przez MP-BGP w tym przypadku nie wskazuje mechanizmu użytego do przysłania ruchu w sieci (może to być ale nie musi MPLS)

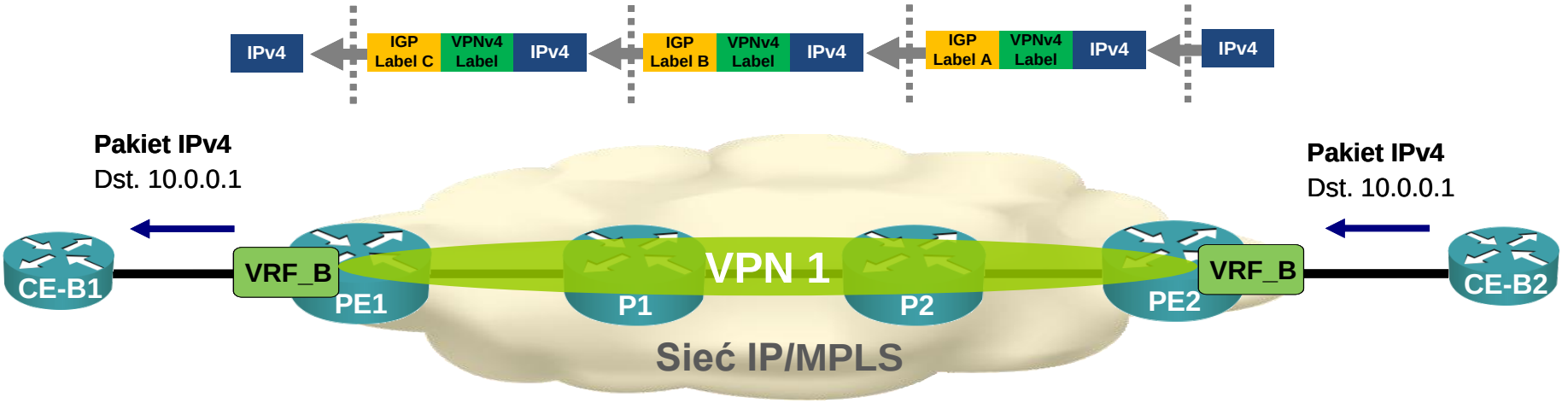
Warstwa kontrolna MPLS VPN



Kolejność przetwarzania:

- CE-B1 rozgłasza trasę IPv4 do PE1 poprzez eBGP.
- PE1 przydziela etykietę VPN do prefiksu IPv4 tworząc wpis VPNv4.
- PE1 wykonuje redystrybucję trasy VPNv4 do MP-iBGP, PE1 ustawia siebie samego jako next hop. MP-BGP przekazuje trasę do PE2.
- PE2 otrzymuje wpis VPNv4, identyfikuje VPN i przeprowadza redystrybucję do VRF-B skąd trasa jest przekazywana do CE-B2.

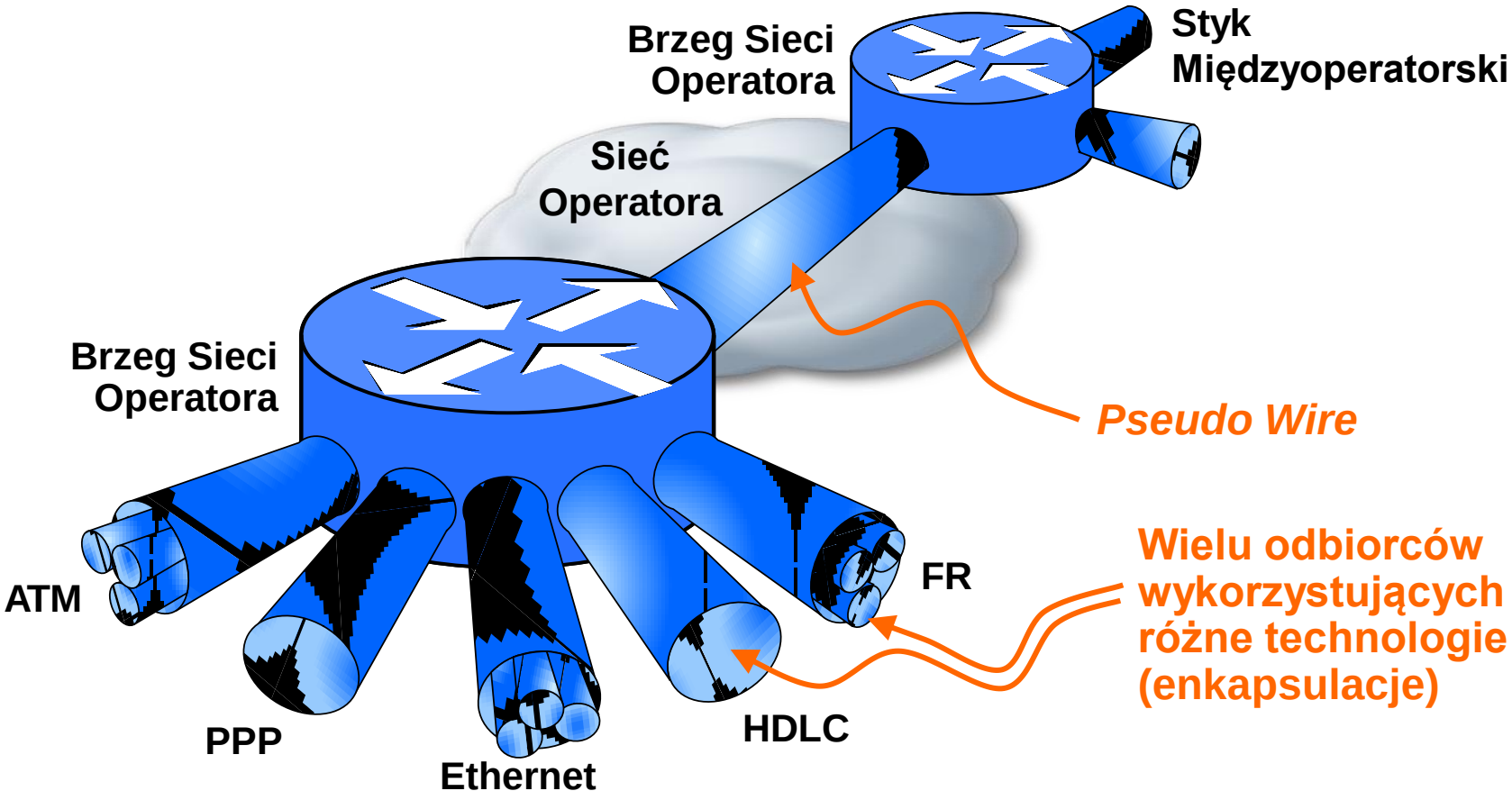
Warstwa danych MPLS VPN



Kolejność przetwarzania:

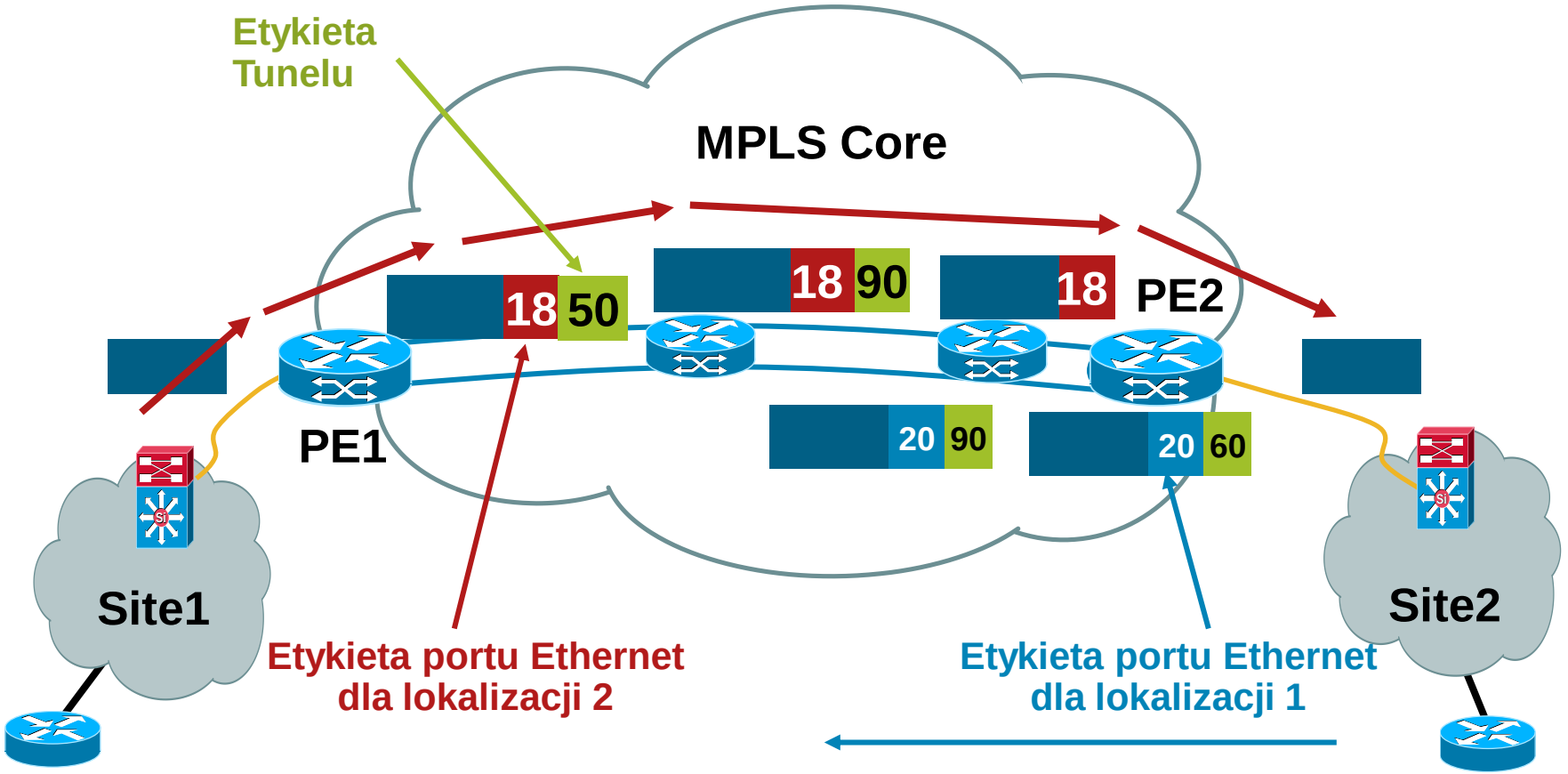
1. **CE-B2 przesyła pakiet IPv4 do PE2.**
2. **PE2 dokłada etykietę VPN rozgłoszoną przez MP-BGP do pakietu IPv4.**
3. **PE2 dokłada dodatkowo etykietę IGP rozgłoszoną przez LDP i przesyła pakiet dalej do routera P2.**
4. **Routery P2/P1 zamieniają zewnętrzną etykietę IGP na etykietę, która była rozgłoszona przez P1/PE1. Pakiet MPLS wysłany jest kolejno do P1/PE1.**
5. **Router PE1 zdejmuje etykietę IGP oraz VPN i przesyła pakiet IPv4 do CE-B1.**

MPLS VPN – L2



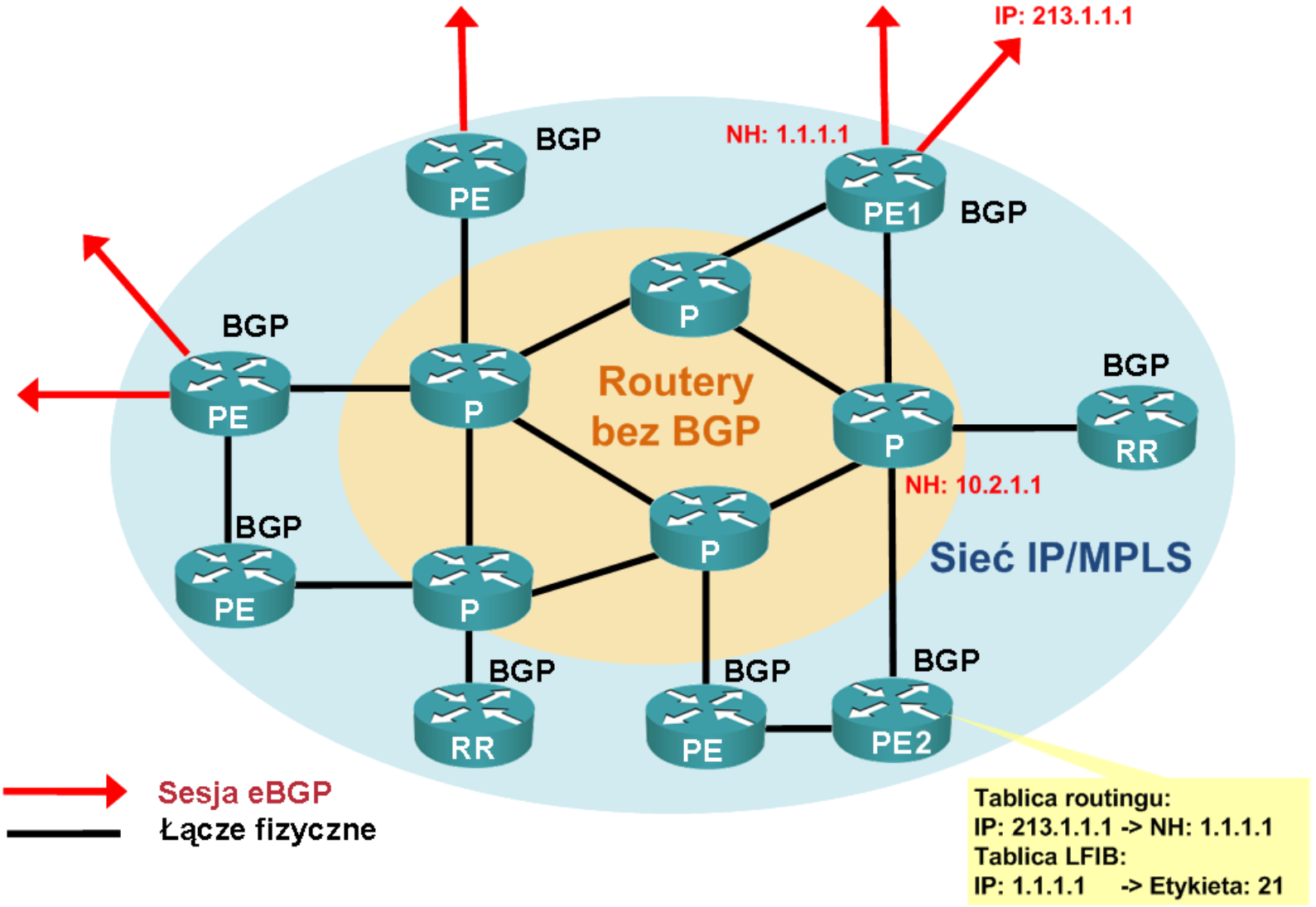
- L2VPN to połączenie pomiędzy odbiorcami przełączane poprzez wspólną sieć

„Any Transport over MPLS”



- AToM nie wymaga protokołu BGP – za alokację etykiet odpowiada LDP

Zaleta 3: BGP nie jest wszędzie potrzebne



W sieci MPLS routery P przełączają na bazie etykiet. Nie trzeba na nich uruchamiać protokołu BGP.

Zaleta 4: MPLS *Traffic Engineering*

- Lepsze wykorzystanie zasobów sieci

  Można łatwo przesłać ruch po ścieżce, która nie jest ścieżką wybieraną przez protokół routingu

- Szybka konwergencja (najczęstszy powód implementacji TE)

  Błyskawiczne przełączenie ruchu na przygotowane wcześniej, zapasowe ścieżki (protokół routingu wylicza takie trasy dopiero w momencie zaistnienia problemu) omijające uszkodzone miejsce w sieci (łącze, węzeł, port źródłowy/docelowy)

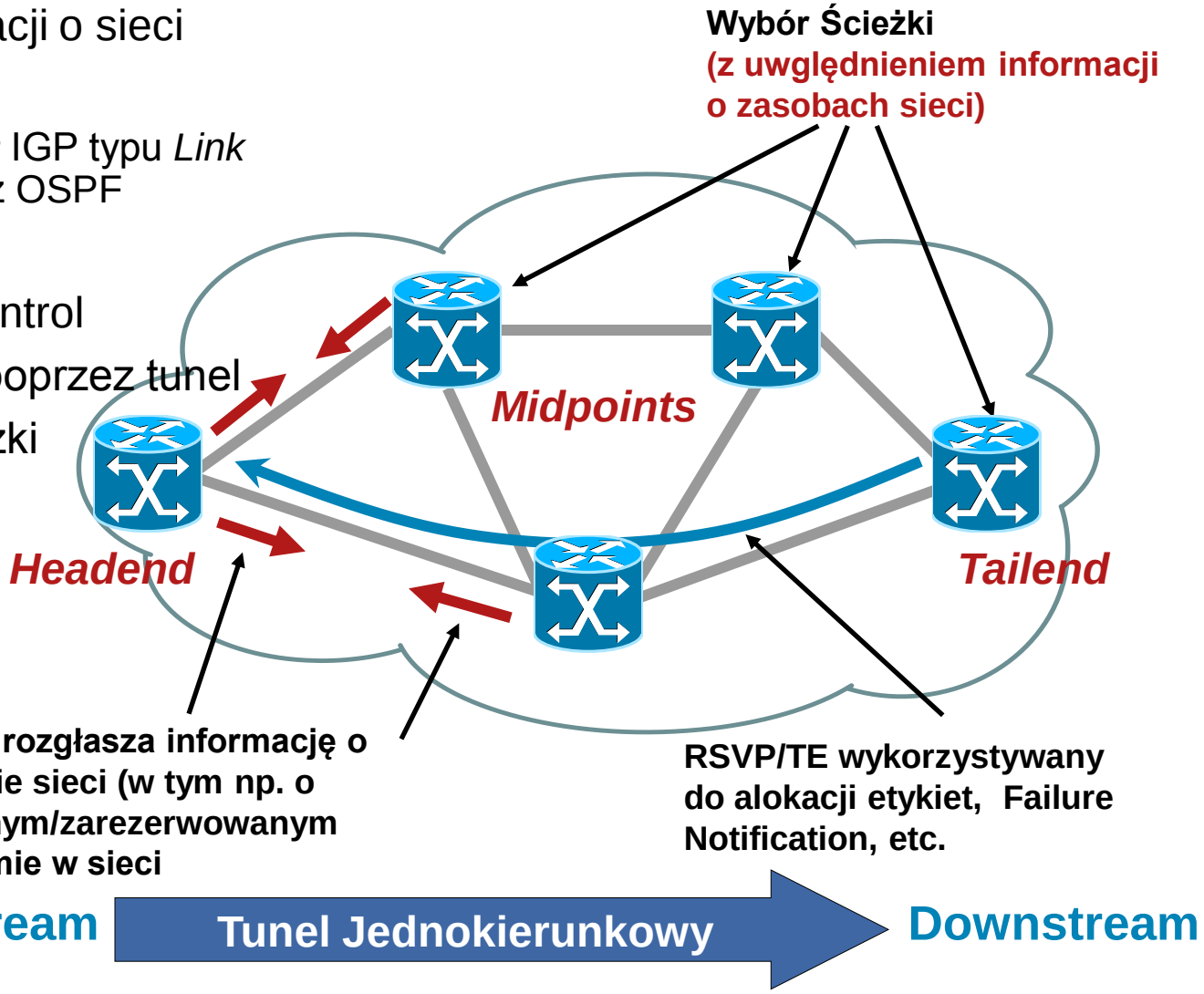
- Możliwość wykorzystania TE do alokacji pasma w sieci

  Wybrane usługi mogą mieć zaalokowane wcześniej zasoby sieci (pasmo, kolejki itp)

MPLS TE

1. Dystrybucja informacji o sieci
2. Wybór ścieżki

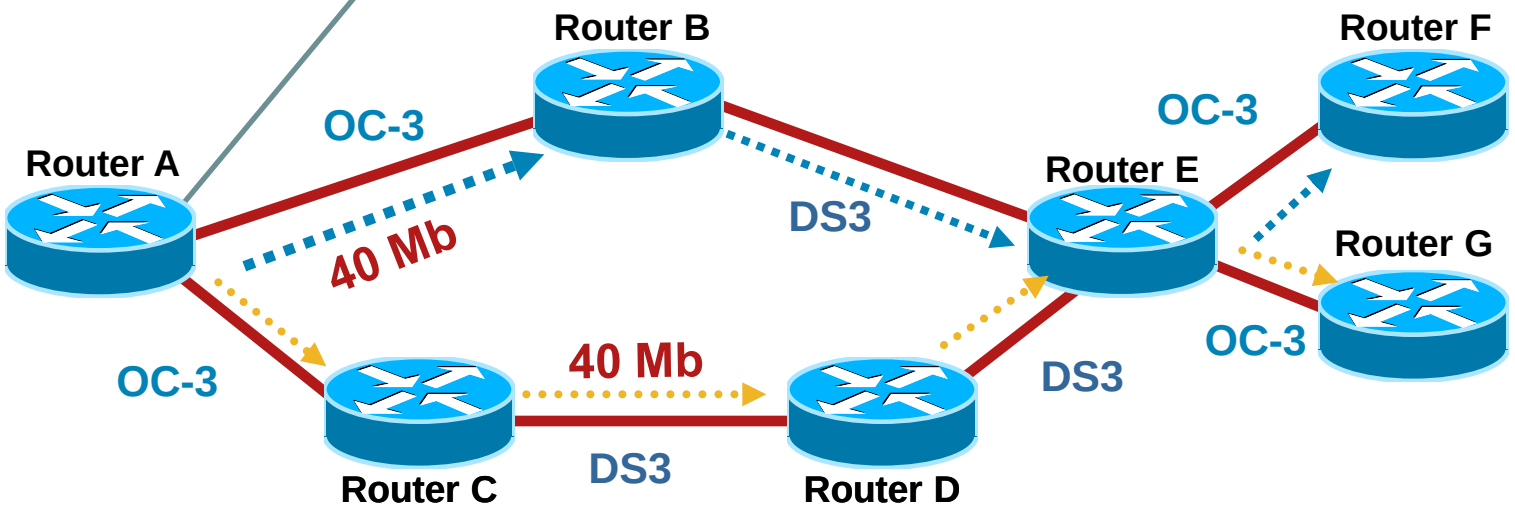
   Wymagany protokoł IGP typu *Link State* np. IS-IS bądź OSPF
3. Zestawianie ścieżki
4. Trunk admission control
5. Przesyłanie ruchu poprzez tunel
6. Utrzymywanie ścieżki i rezerwacji w sieci



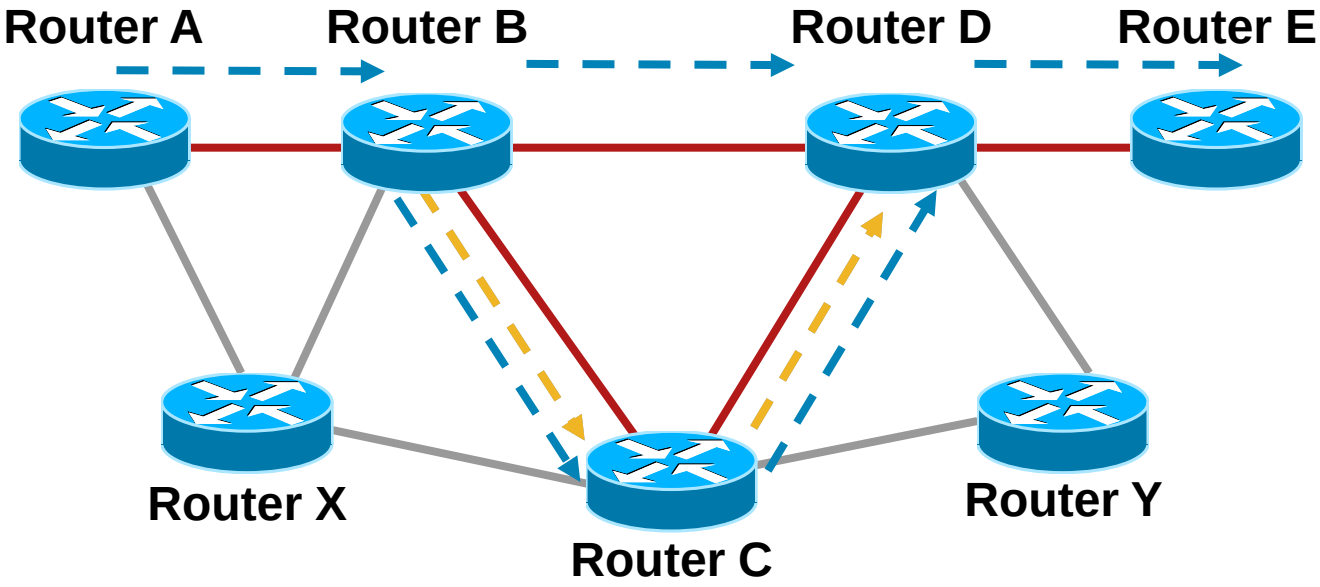
Problem najkrótszej ścieżki

Węzeł	Next-Hop	Koszt
B	B	10
C	C	10
D	C	20
E	B	20
F	Tunnel 0	30
G	Tunnel 1	30

- Router A widzi wszystkie łącza
- Router A wybiera ścieżkę poprzez sieć w oparciu o informacji o dodatkowych zasobach (np. paśmie)
- Ruch przechodzi w całości poprzez sieć!



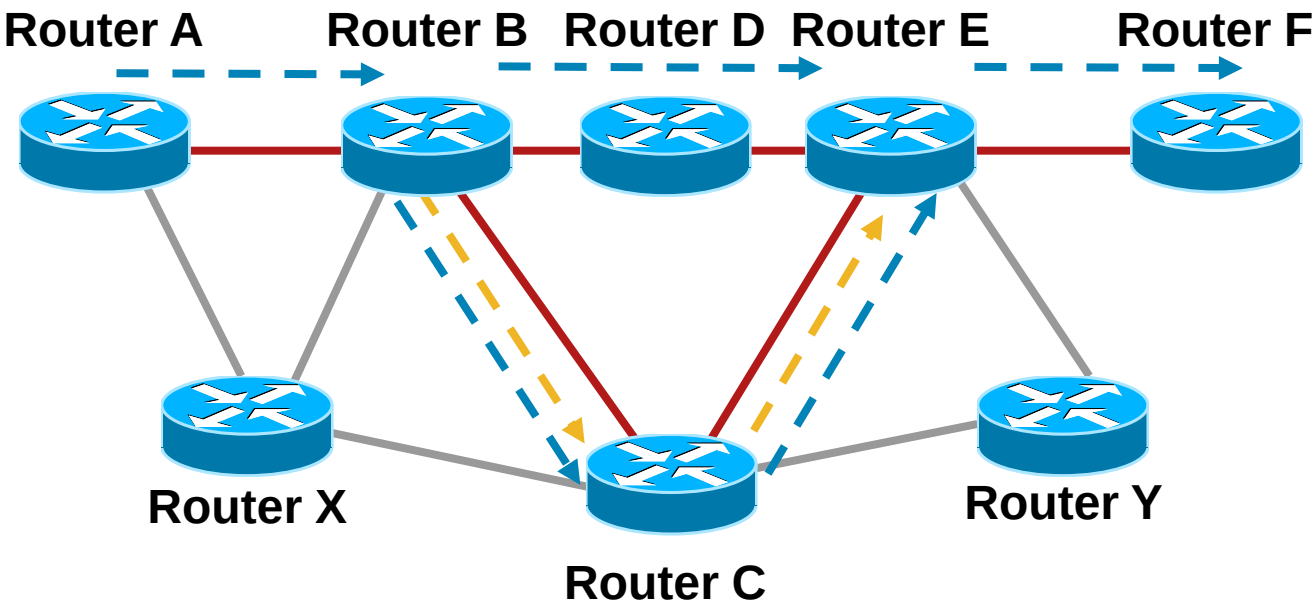
MPLS Fast Reroute
Protekcja łącza



- *Primary tunnel*: A → B → D → E
- *Backup tunnel:* B → C → D (zest. wcześniej)
- Konwergencja = ~50* ms

***Średnio**

MPLS Fast Reroute
Protekcja węzła



- *Primary tunnel*: A → B → D → E → F
- *Backup tunnel*: B → C → E (zest. wcześniej)
- Konwergencja = ~100 ms

Czy naprawdę nie ma alternatyw dla MPLS?

Alternatywy dla MPLS

- Istnieją technologie realizujące wybrane usługi dostępne w MPLS w sposób prostszy niż jest to zrealizowane w samym MPLSie

  np. L2VPN – L2TPv3, L3VPN – GRE

- Siłą MPLS jest fakt, iż stanowi uniwersalną platformę dzięki, której można osiągnąć stawiane założenia
- Ostatnio pojawiło się szereg pomysłów jak sieci IP/MPLS uczynić prostszymi i wydajniejszymi np.

  IP Fast Reroute – wykorzystanie predefiniowanych tras do szybkiej konwergencji (idea zaczerpnięta z MPLS FRR)

  VPNs over mGRE – przesyłanie ruchu VPN w sieci jedynie za pomocą IP (GRE), brak konieczności stosowania MPLS/LDP

  MPLS-TP – podzbiór funkcjonalności MPLS do budowy usług typowych dla sieci transportowych

Alternatywa #1 - IP Fast Reroute

- Zestaw mechanizmów i algorytmów umożliwiających konwergencje sieci IP w czasach zbliżonych do tych osiąganych w sieciach MPLS z FRR
- Zaproponowano kilka algorytmów efektywnego poszukiwania zapasowej trasy w sieci m.in *Loop-Free Alternates*, *U-Turns*, *Not-Via Addresses* - większość z nich jest jeszcze dyskutowana w *IETF Routing Area Working Group*
- Założenia:

  Łatwość stosowania, konfiguracji i diagnozy problemów związanych z używaniem mechanizmu IP FRR w sieci

  Zdolność do działania w każdej topologii sieci

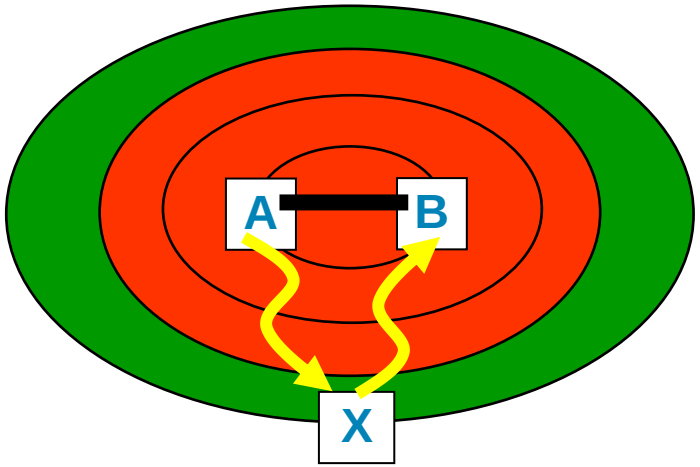
  Szybka detekcja awarii łączy

  Protekcja każdego typu ruchu: IPv4/v6, MPLS, *unicast*/*multicast*

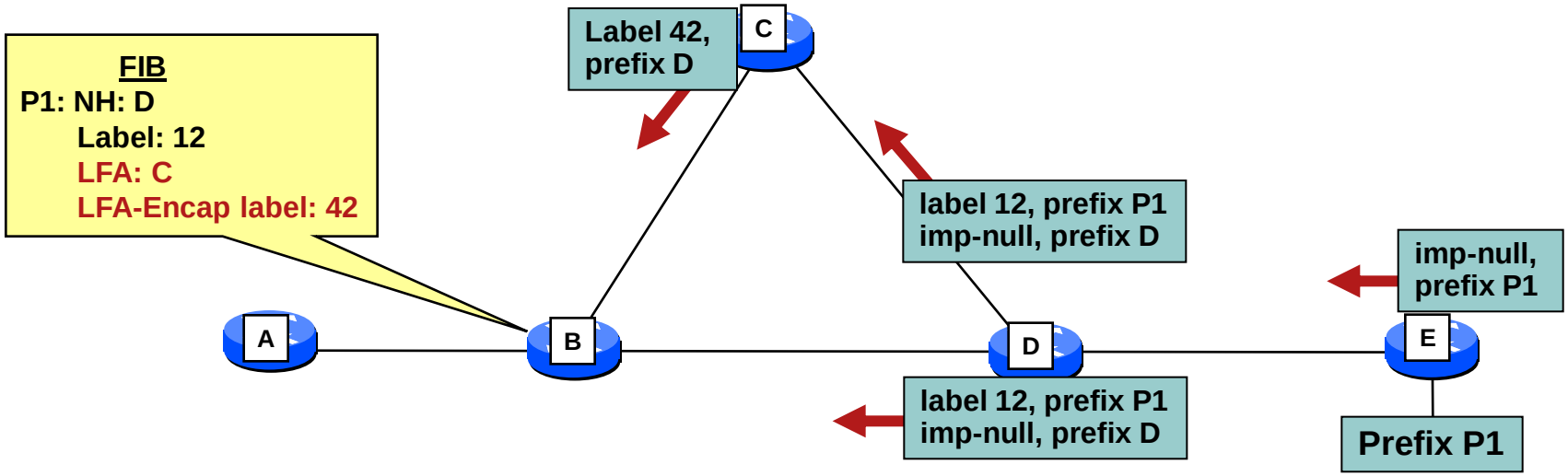
  Brak konieczności stosowania sygnalizacji w sieci
- Interesująca opcja, jeśli w swojej sieci planowałeś zastosować MPLS by uzyskać szybką konwergencję.

Ciekawe obserwacje...

- Kiedy link AB ma awarię dotyka one jedynie niewielkiego obszaru topologii sieci (rejon czerwony)
  - Rozmiar tego obszaru zazwyczaj nie przekracza 5 przeskoków sieci IP (*hops*)
  - Dla pozostałej części sieci sytuacja związana z routingiem nie ulega zmianie
- IP Fast Reroute musi znaleźć punkt w sieci (węzeł) który:
  - Nie będzie dotknięty przez daną awarię
  - Może się do niego dostać niezależnie do tego czy występuje awaria
  - Będzie można przekazać ruch do pozostałej części sieci nie korzystając z łacza AB

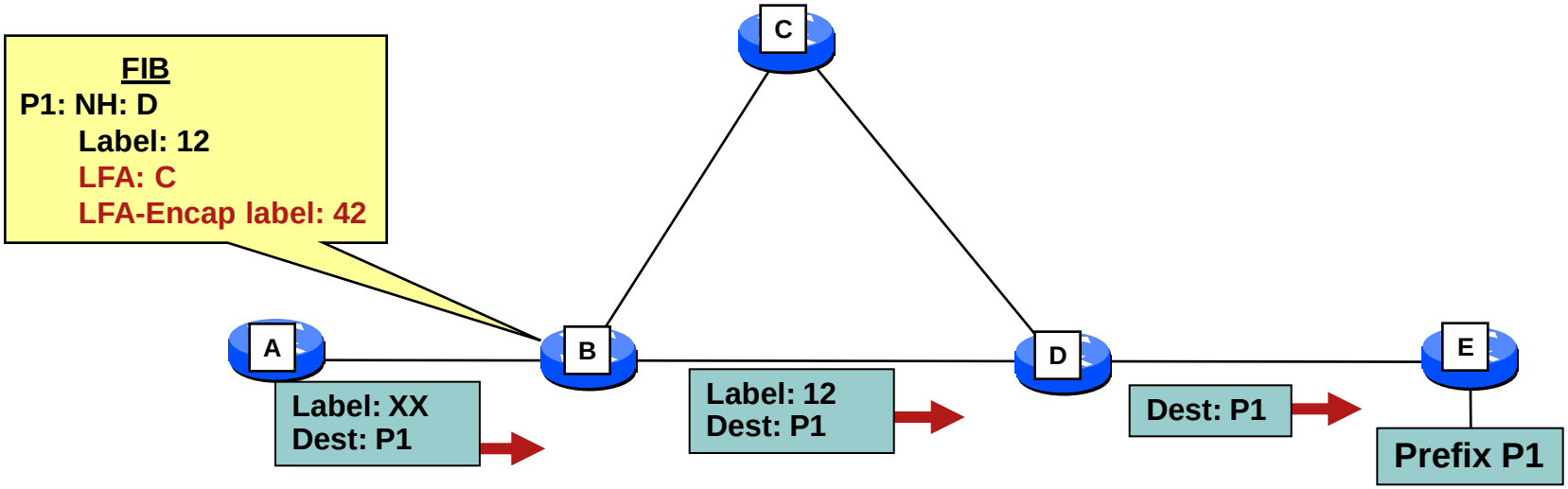


Trasy *Loop Free Alternate* (LFA) 1/3



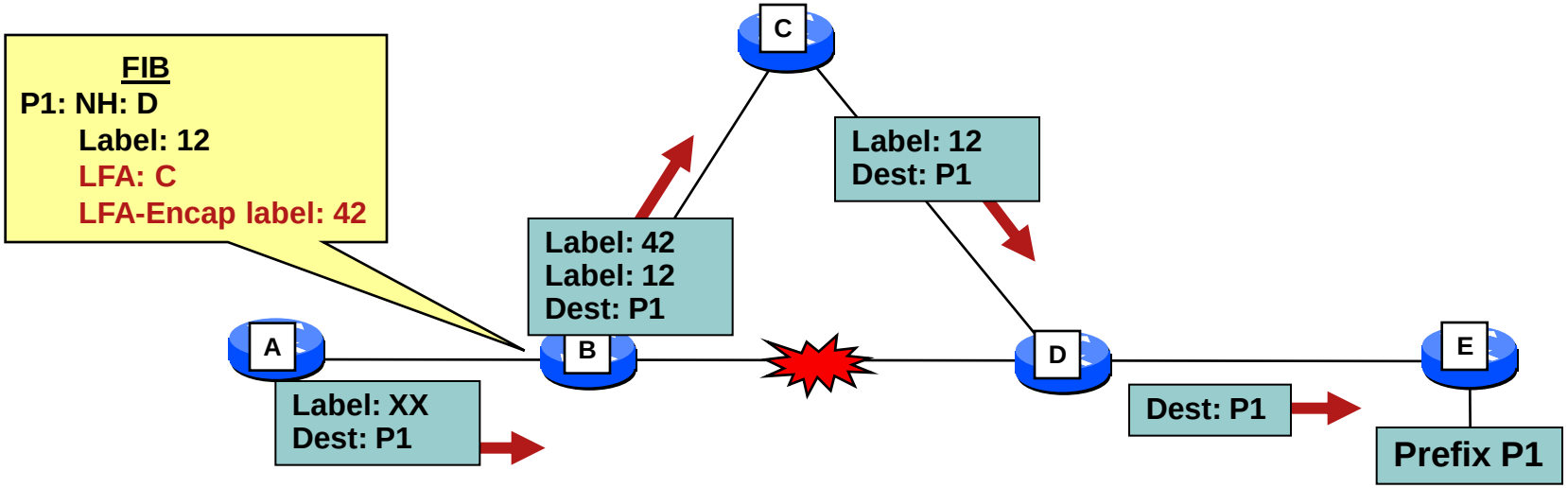
- B oblicza trasę zapasową (LFA) oraz związane z nimi informacje o prefiksie IP i etykiecie MPLS
  - informacja o prefiksie IP z tablicy protokołu routingowego *link state LSDB*
  - informacje o etykiecie z LDP/LIB
  - użyj etykiety która została rozgłoszona przez sąsiada LFAI

Trasy *LFA 2/3*



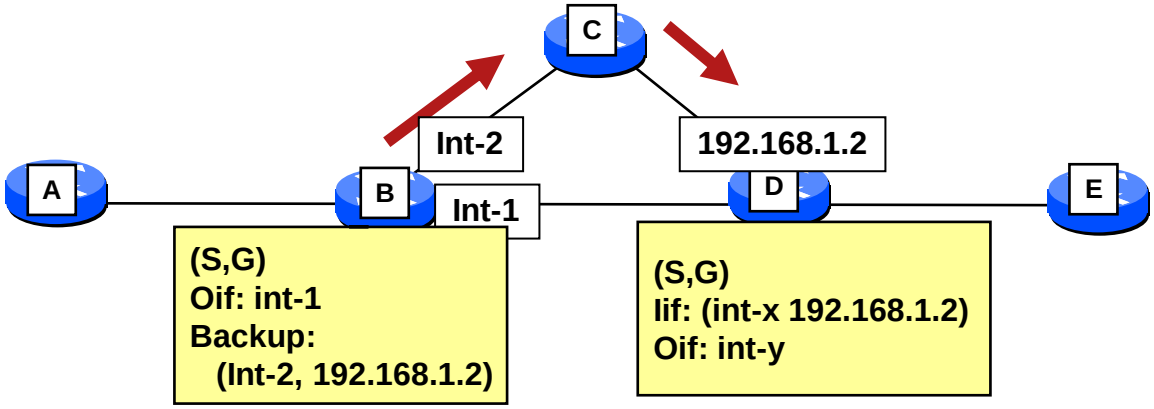
- B wykorzystuje zebrane informacje do normalnego przesyłania ruchu IP i ruchu MPLS

Trasy *LFA* 3/3



- Podczas awarii, pakiety są dodatkowo enkapsulowane w etykietę LFA
- Zachowanie podobne do MPLS-FRR, ale bez wykorzystanie żadnej dodatkowej sygnalizacji!

Trasy *LFA* a ruch typu Multicast



- Ruch typu *multicast* musi zostać poddany dodatkowej enkapsulacji
- Węzeł otrzymujący ruch enkapsulowany (u nas D) wykonuje RPF w oparciu o adres źródłowy przenoszony w zewnętrznym nagłówku

IP Fast Reroute

- Zalety

  Wsparcie dla IP (v4/v6), MPLS, multicast

  Mechanizm ścieżki zapasowej – dobrze znany z MPLS FRR

  Nie wymaga zdolności do współpracy (*interoperability)* pomiędzy platformami różnych dostawców

  Nie wymaga sygnalizacji ani rozszerzeń w IGP

  Łatwość wprowadzania do sieci, nie wymaga przebudowy istniejących w sieci mechanizmów szybkiej konwergencji

- Wady

  Wymaga topologii o dużej ilości połączeń - w przypadku LFA wg badań bez problemu daje się pokryć 70-85% topologii

  Wymaga wsparcia w sprzęcie do enkapsulacji ruchu (multicast)

Alternatywa #2 – MPLS VPN over mGRE

- Patrząc na usługi MPLS VPN widać że MPLS jest wykorzystywany jedynie do budowy ścieżki LSP pomiędzy routerami PE
- Możliwość wyeliminowania konieczności posiadania MPLS w sieci daje wiele zalet:

  Uproszczenie sieci - cały ruch w naszej sieci jest ruchem IPv4/IPv6

  Brak LDP i enkapsulacji MPLS – w sieci mamy jedynie protokół MP-BGP odpowiedzialny za dystrybucję tras VPN

  Możliwość zestawiania usług VPN poprzez sieci zbudowanej jedynie w oparciu o IP (np Internet)

  Łatwość realizacji styków międzyoperatorskich

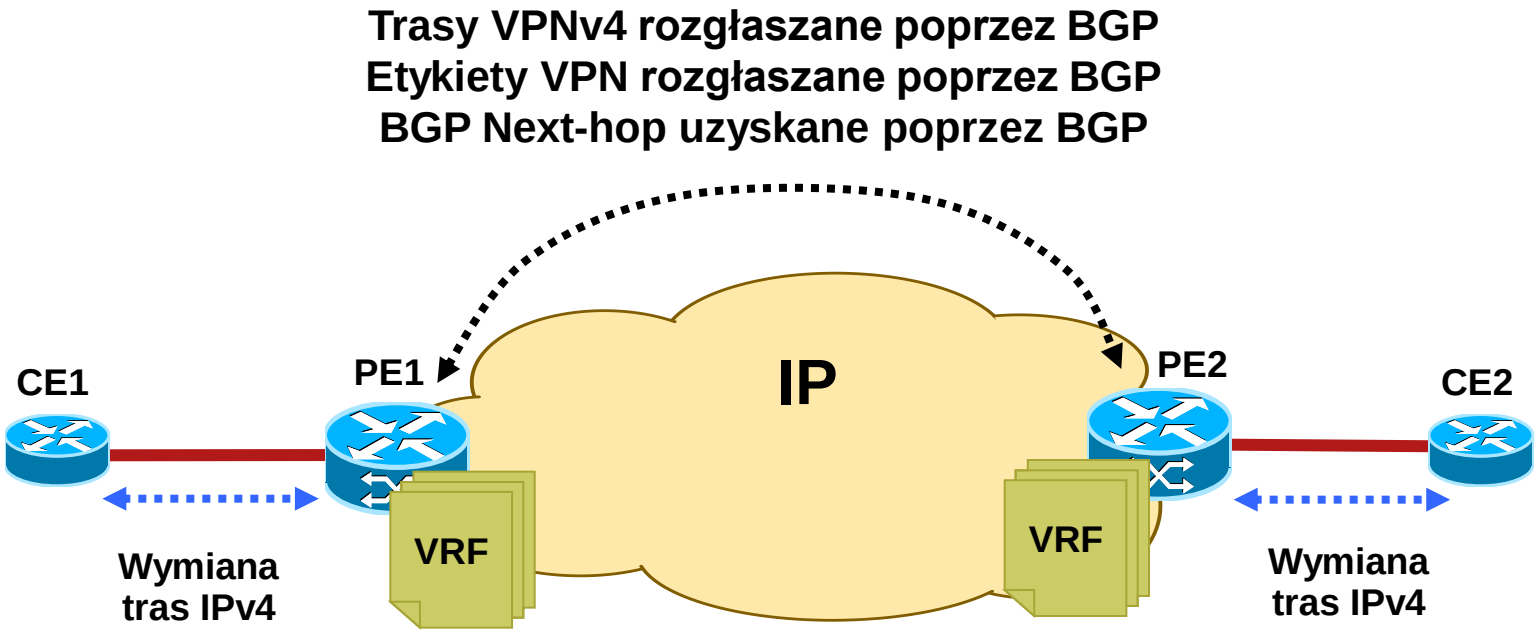
- Wady

  Wymaga sprzętowego wsparcia dla GRE

  Dodatkowy narzut bitowy związany z wielkością nagłowka GRE

- Interesująca opcja, jeśli w swojej sieci planowałeś zastosować MPLS by móc łatwo tworzyć sieci VPN

MPLS VPN over mGRE – warstwa kontrolna

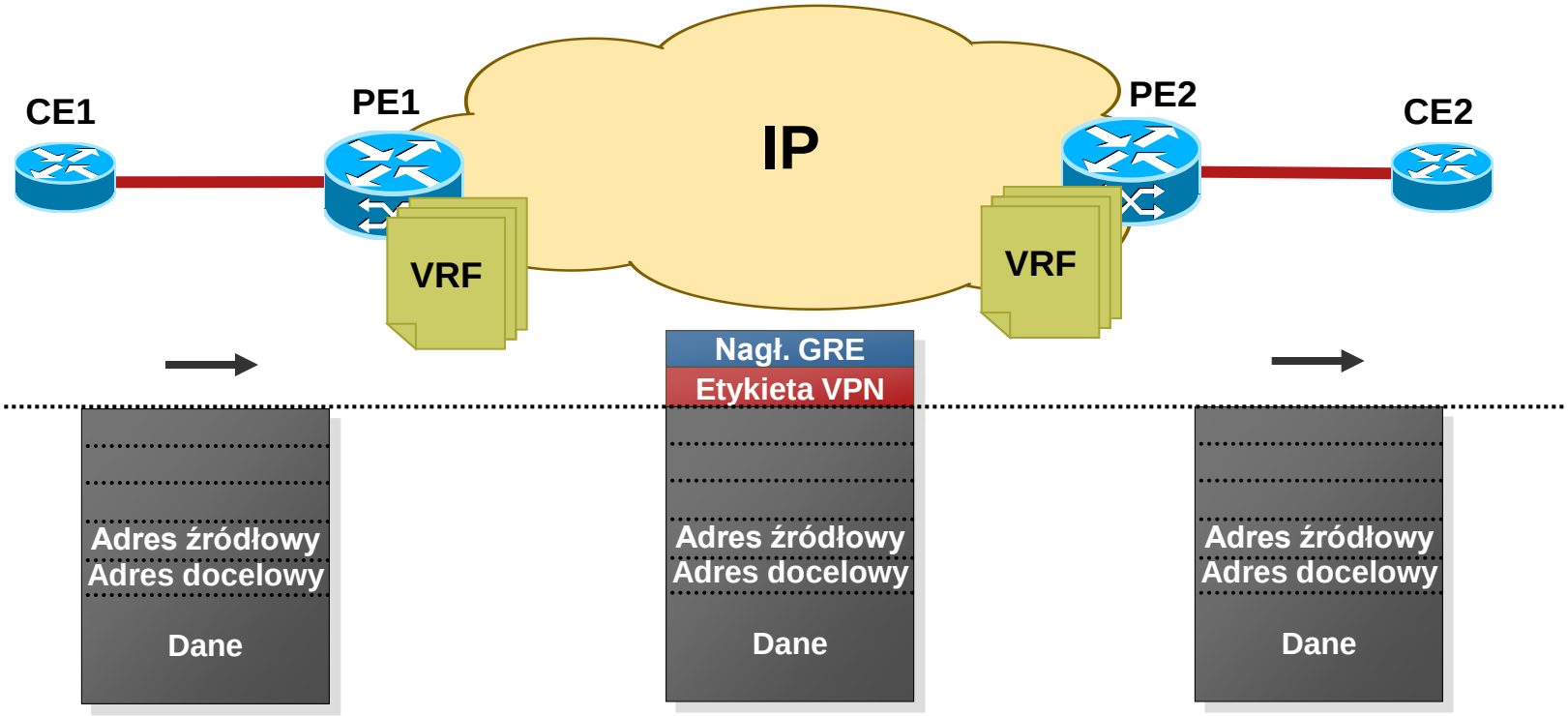


- Warstwa kontrolna taka sama jak w przypadku MPLS (VPNv4 add family, RD, RT, etc.)
- Automatyczna enkapsulacja GRE:

  jeśl *next-hop* dla trasy VPNv4 nie jest osiągalny poprzez sieć MPLS (LSP) dodaj nagłówek GRE z adresem docelowym takim jak rozpatrywany *next-hop*

  GRE enkapsuluje cały pakiet MPLS wraz z etykietą VPN i pełni rolę tunelu prowadzącego do zdalnego PE (dokładnie tak samo jak LSP MPLS)

MPLS VPN over mGRE – warstwa danych

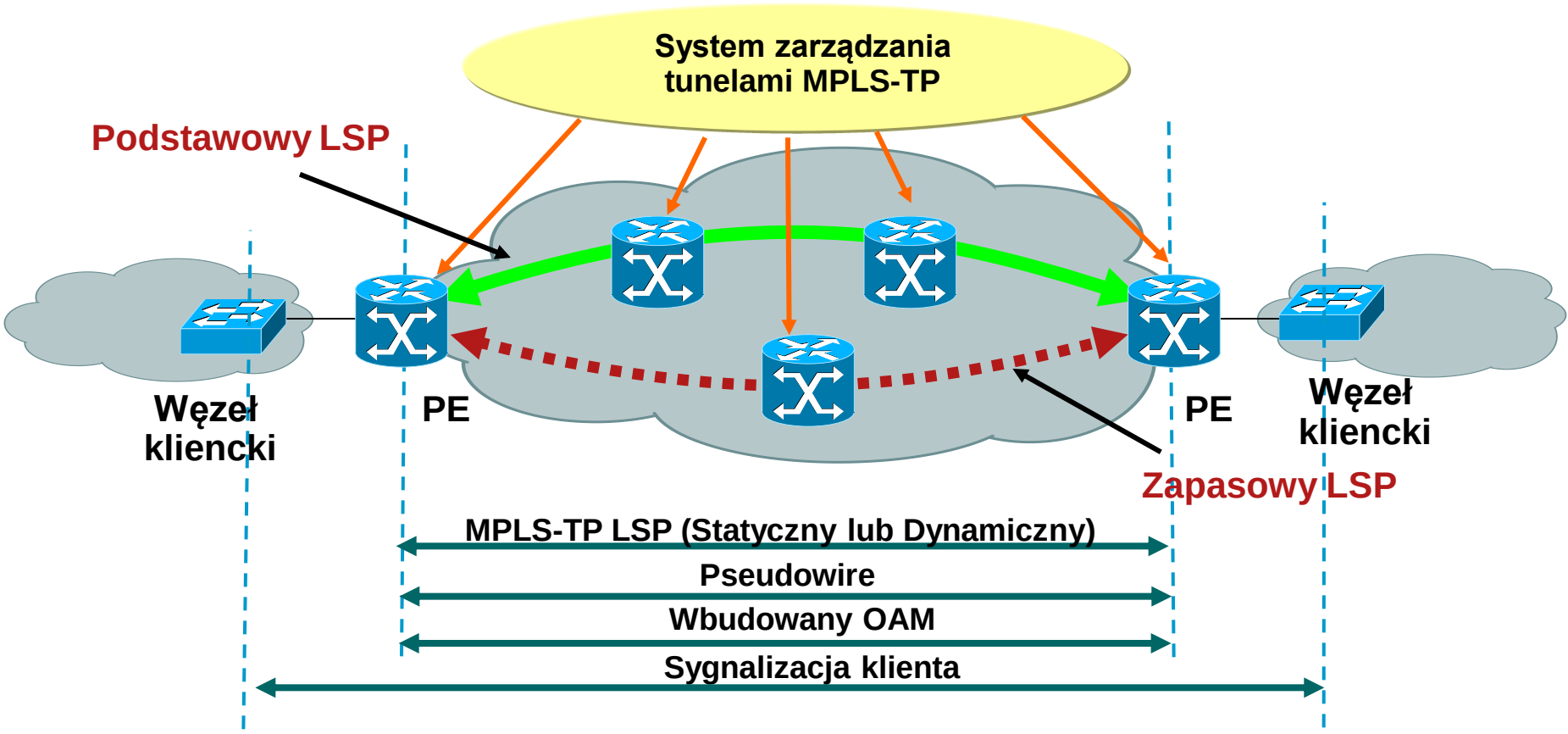


- Nagłowek GRE + etykieta VPN dodawane do każdego pakietu
- Przełączenia pakietu w sieci do docelowego PE w oparciu o informację w nagłówku GRE
- Docelowy PE wykorzystuje etykietę VPN do przesłania ruchu do odpowiedniego VRF

Alternatywa #3 – MPLS-TP

- MPLS Transport Profile – próba adaptacji MPLS do wymogów stawianych sieciom transportowym (przenoszącym ruch SDH, ATM itp)
  - Wykorzystanie MPLS-TE wraz z rozszerzeniami (np. kontrola stanu ścieżki LSP za pomocą BFD)
  - Brak PHP
  - Nacisk na OAM i *pseudowires*
  - Dwukierunkowe (*bidirectional*) ścieżki LSP
  - Deterministyczne zestawianie ścieżek LSP poprzez system NMS
- Wady
  - Wymaga NMS
- Interesująca opcja, jeśli twoja sieć spełnia głównie funkcje transportowe (przenosi ruch SDH, ATM itp)

Koncepcja MPLS-TP



Połączenia postawowe, jak i zapasowe są pre-konfigurowane, protekcja 1:1, przełączenie na bazie sygnalizacji OAM, wskazany system zarządzania do zarządzania tunelami.

Podsumowanie

Podsumowanie

- Sieci MPLS zapewniają obecnie największą elastyczność w budowie usług w warstwie 2 oraz 3.

- Jest to technologia dojrzała, szeroko stosowana w sieciach na całym świecie – nie pozbawiona jednak problemów (skomplikowanie niektórych mechanizmów np TE, trudność diagnozy problemów zwłaszcza związanych z *warstwą danych* itp)

- Coraz częściej pojawiają się technologie (IP FRR, MPLS VPN over GRE, MPLS-TP) które umożliwiają znaczne uproszczenie sieci i jej elementów przy zachowaniu porównywalnych parametrów (np. czasu konwergencji, funkcjonalności itp.). Wydaje się, że wkrótce będziemy świadkami ciekawych zmian, albowiem nic tak nie dobrze nie wpływa na rozwój technologii jak konkurencja.

Pytania?


CLAUDE
PRESBYTERIAN CHURCH
THERE ARE SOME
QUESTIONS THAT
CAN'T BE
ANSWERED BY
GOOGLE
Citytv


PLNOG