Realizacja styku międzyoperatorskiego dla usług L2/L3 VPN

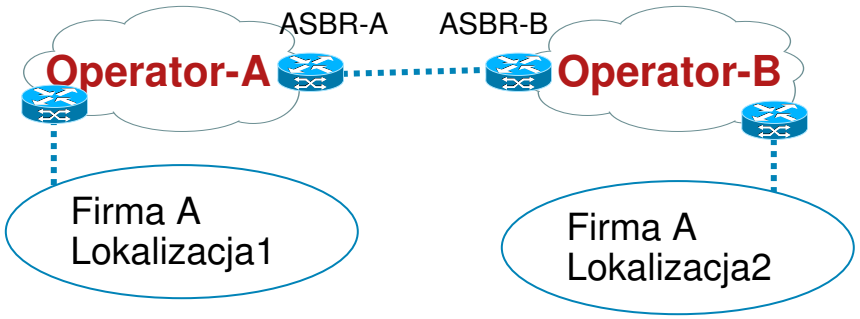
Piotr Jabłoński
Systems Engineer
pijablon@cisco.com

PLNOG, Kraków, 10.09.2009

Plan prezentacji

- Rozwiązania Inter-AS
  - Inter-AS L3 VPN
    - Metody połączenia
    - Bezpieczeństwo
  - Inter-AS L2VPN
    - Any Transport over MPLS (AToM)
  - Inter-AS MPLS TE
- Carrier Supporting Carrier
  - Metody połączeń CSC
  - MPLS IPv4 VPNs
  - MPLS L2 VPNs

Inter-AS vs. Carrier Supporting Carrier



Inter-AS

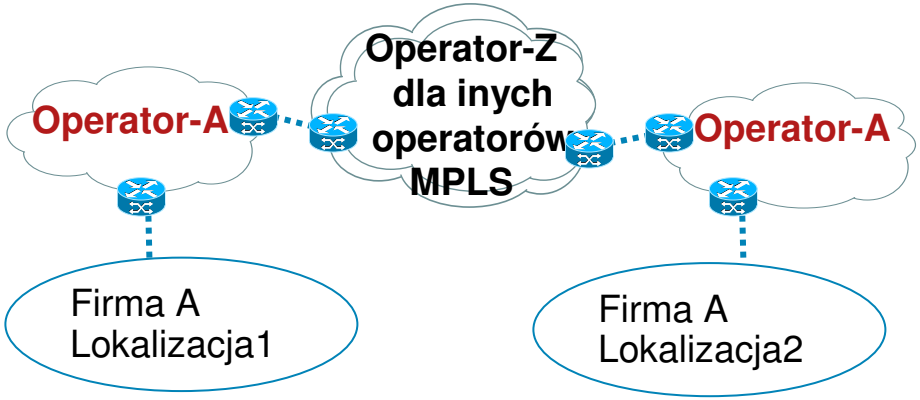
Model Peer-Peer

Operatorzy dostarczają usługę do tego samego klienta.

Pojedynczy POP niedostępny we wszystkich lokalizacjach wymaganych przez klienta.

Operatorzy muszą wspierać MPLS VPN.

Informacje VPN klienta są współdzielone między operatorów.



CSC

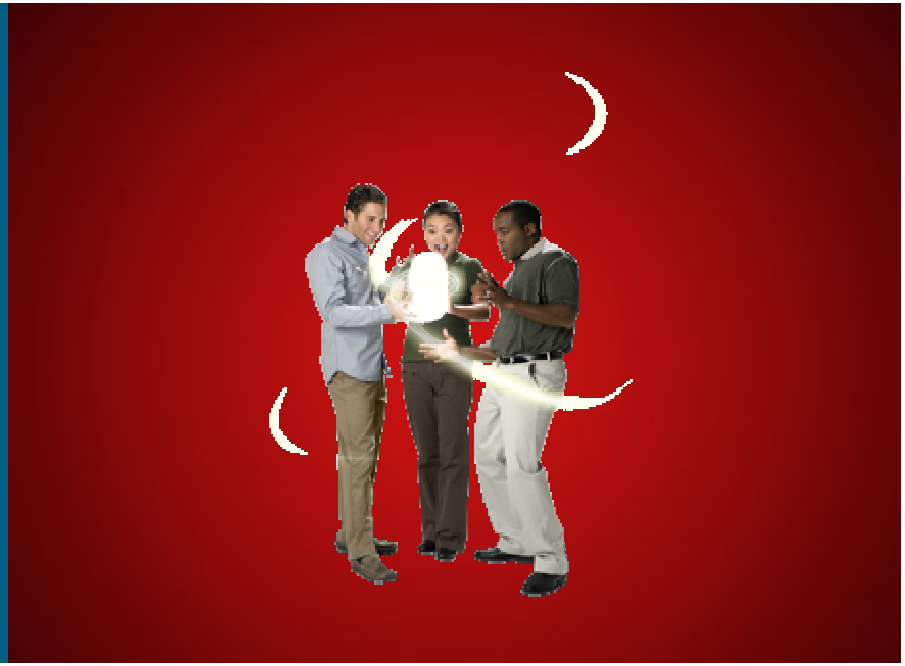
Model Client-Server

Operator-A jest klientem z perspektywy Operatora-Z.

Model biznesowy operatora-A nie uwzględnia budowy własnej, globalnej sieci szkieletowej.

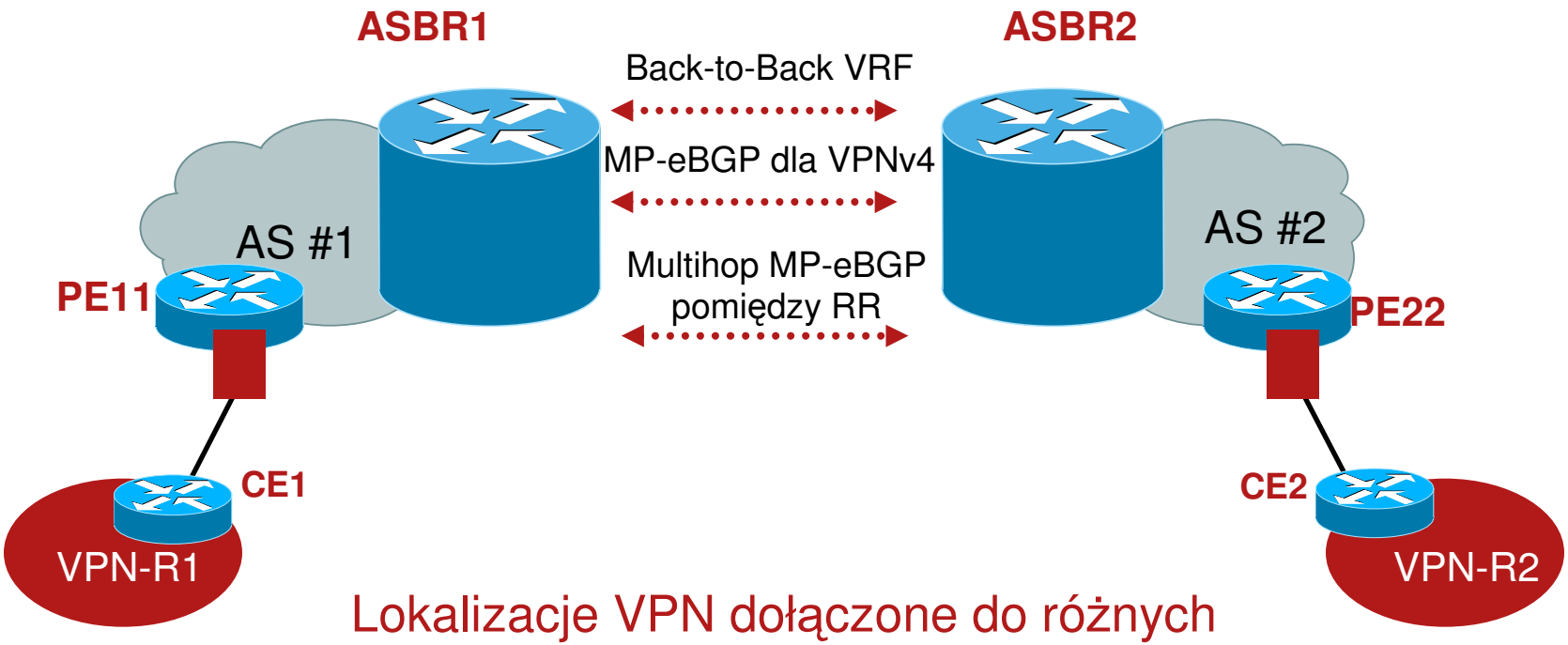
Informacje VPN klientów końcowych nie przenikają do sieci Operatora-Z.

Inter-AS L3 VPNs



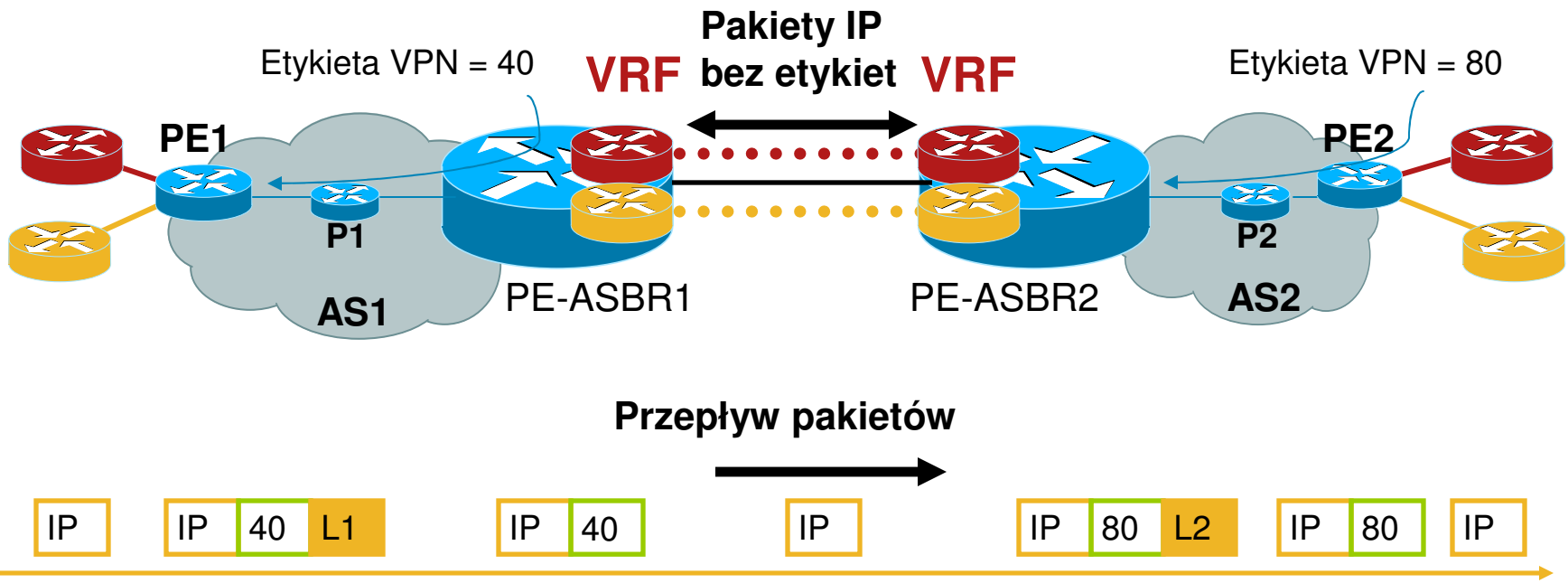
Dystrybucja danych poprzez Inter-AS VPNv4

Jak wymieniać informację o prefiksach VPN pomiędzy routerami ASBR?



Lokalizacje VPN dołączone do różnych operatorów.

Inter-AS VPN Opcja A



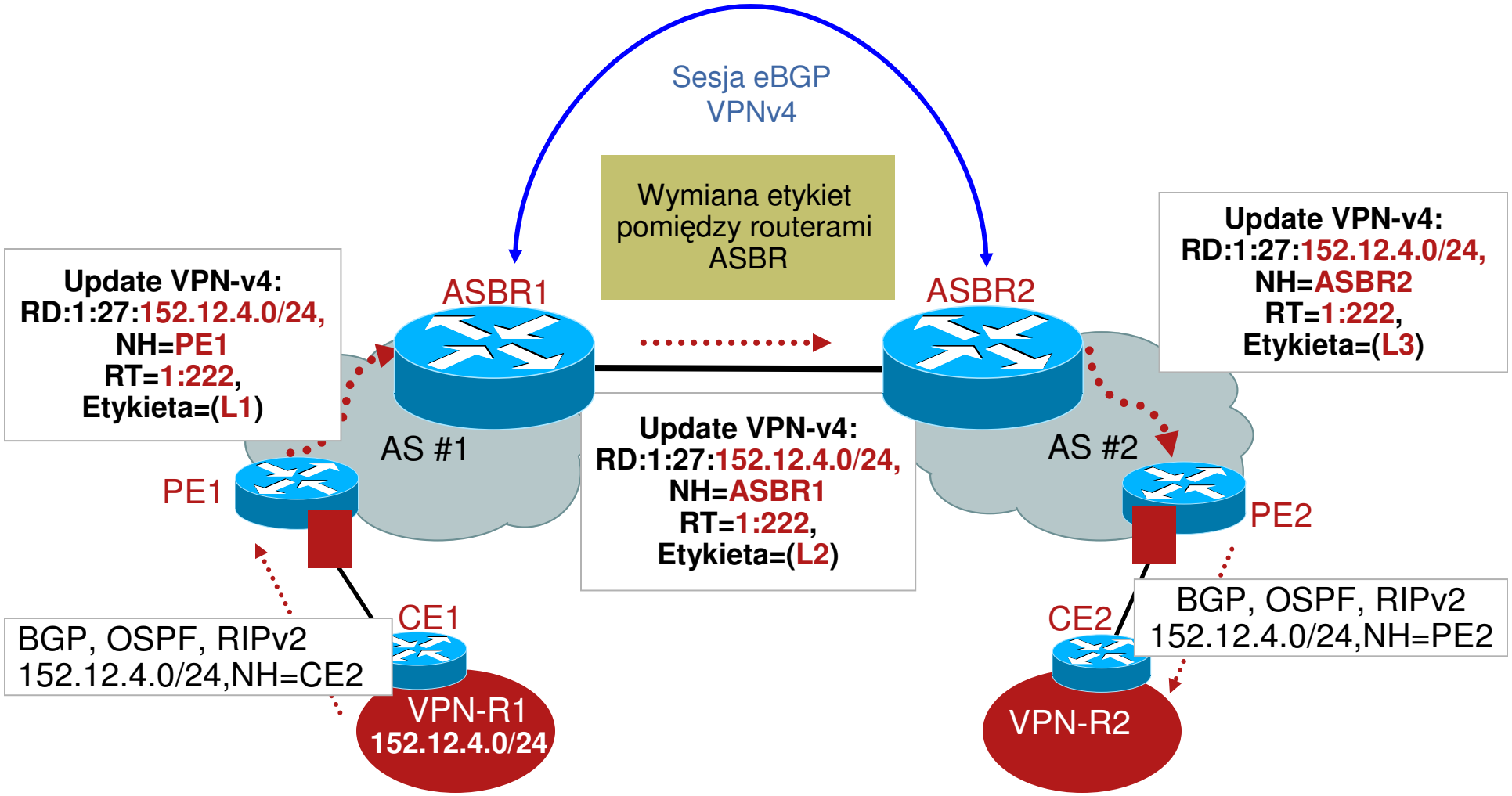
Jeden logiczny interfejs per VPN pomiędzy operatorami.

Pakiety pomiędzy routerami ASBR przekazywane są bez etykiet.

Na łączu można wykorzystać dowolny protokół PE-CE.

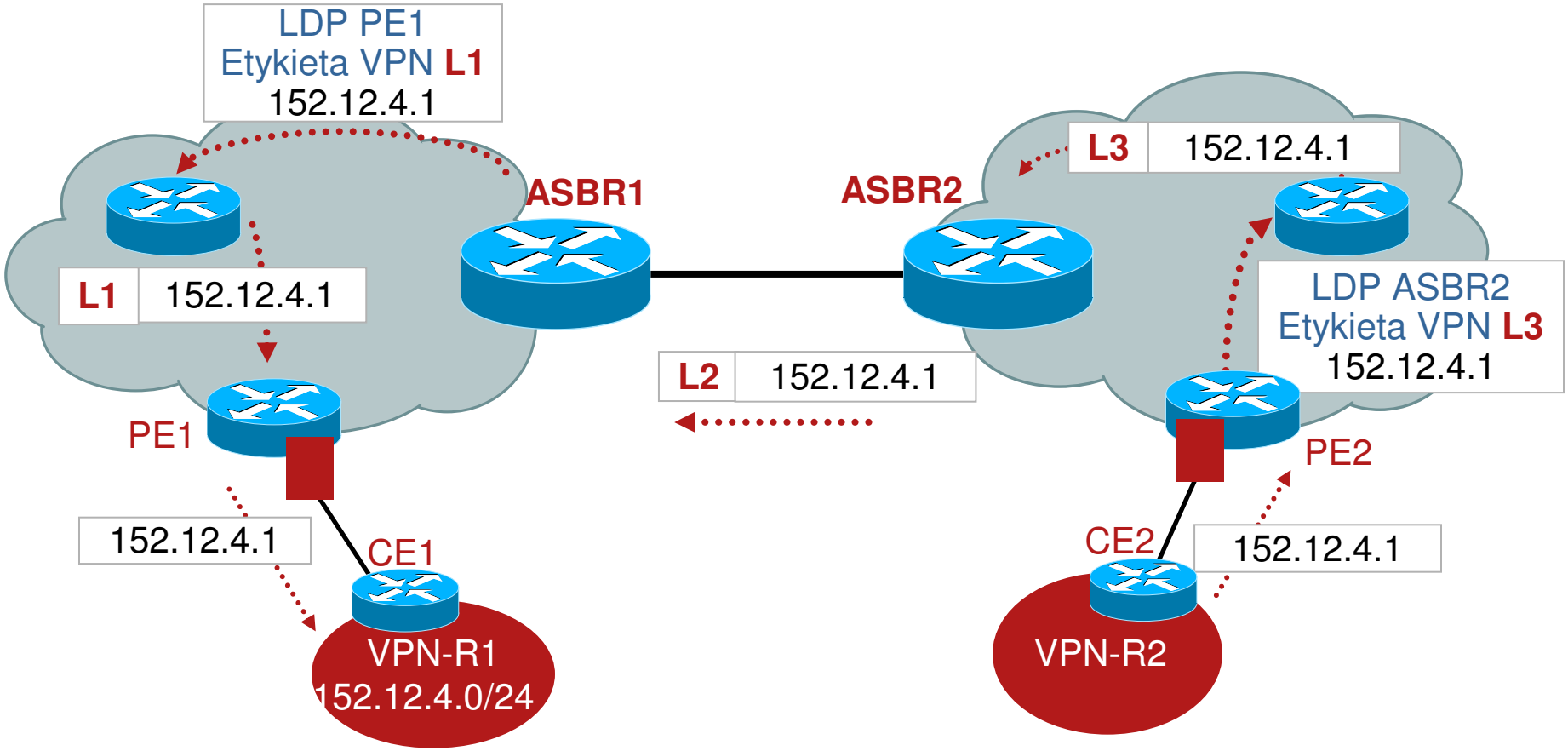
Polityki QoS ustalane niezależnie na każdym z routerów ASBR.

Inter-AS VPN Opcja B — Warstwa kontrolna

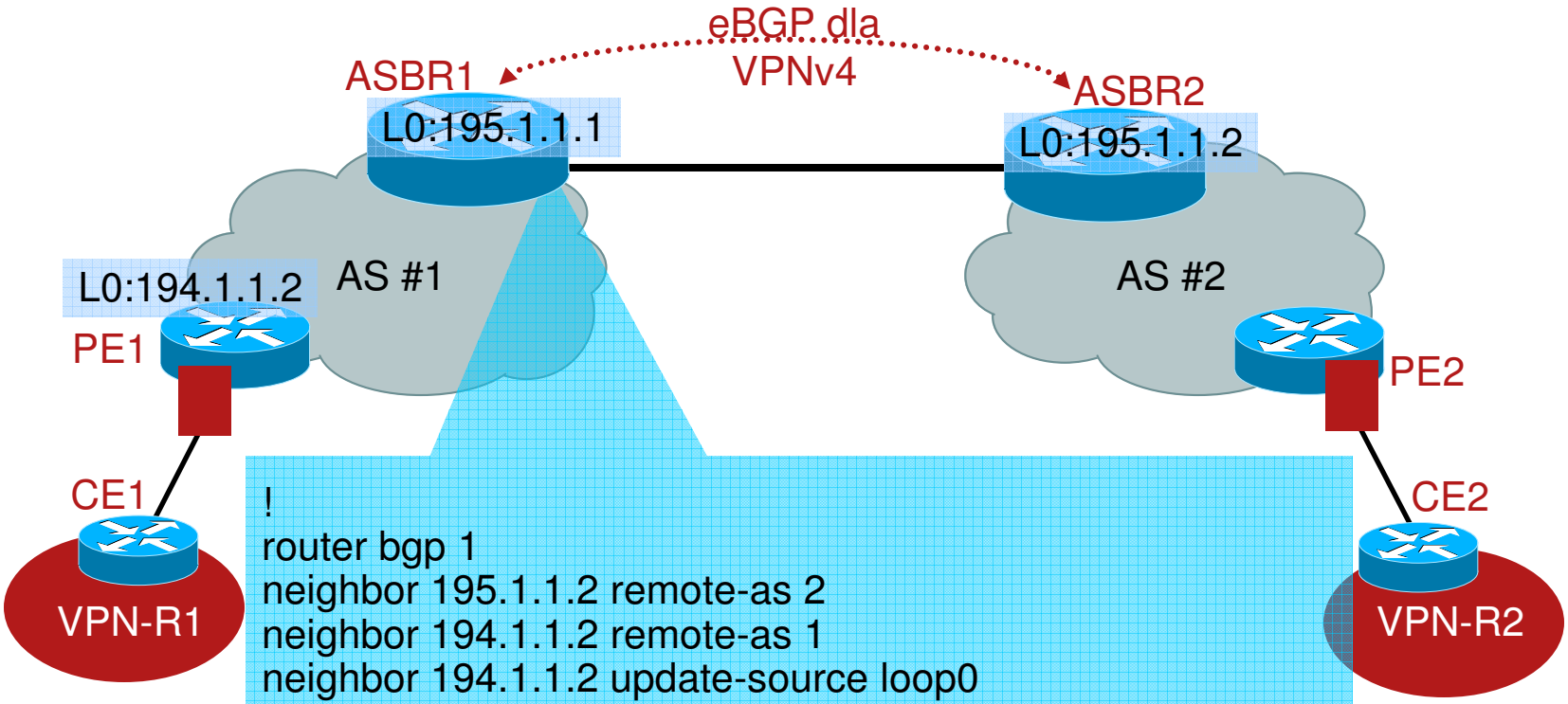


Dystrybucja prefiksów oraz etykiet pomiędzy CE1-CE2.

Inter-AS VPN Opcja B — Transmisja danych



Inter-AS VPN Opcja B – Przykład konfiguracji




```
!
router bgp 1
neighbor 195.1.1.2 remote-as 2
neighbor 194.1.1.2 remote-as 1
neighbor 194.1.1.2 update-source loop0
no bgp default route-target filter
!
address-family vpnv4
neighbor 194.1.1.2 remote-as 1 activate
neighbor 194.1.1.2 remote-as 1 next-hop-self
neighbor 195.1.1.2 remote-as 2 activate
neighbor 195.1.1.2 remote-as 2 send-community extended
```

Inter-AS VPN Opcja B

Routery ASBR wymieniają prefiksy poprzez eBGP

- Zewnętrzna sesja MP-BGP do wymiany prefiksów VPNv4.

Sesja MP-BGP z next-hop rozgłaszanym do sąsiedniego ASBR

- Adresy next-hop oraz etykiety są zmieniane na brzegu sieci operatora.

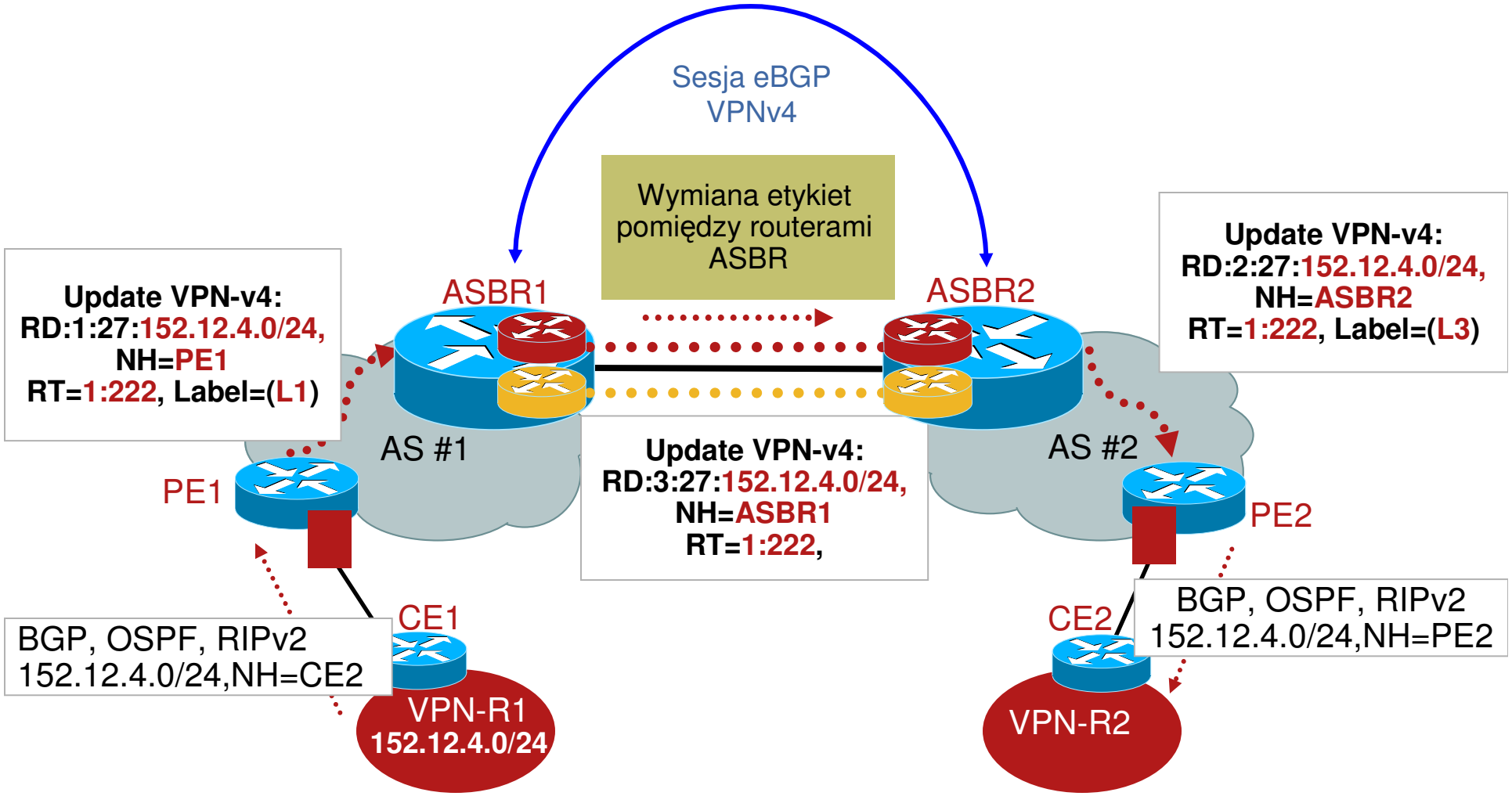
Router ASBR przechowuje i przekazuje dalej wszystkie prefiksy VPN

- Prefiksy trzymane są w tablicy BGP.
- Nie ma VRF na ASBR.

ASBR domyślnie przydziela nową etykietę VPN

- Można to wyłączyć poprzez konfigurację next-hop-unchanged.

Inter-AS VPN Opcja AB — Sterowanie



Dystrybucja prefiksów oraz etykiet pomiędzy CE1-CE2.

I-AS VPN Opcja AB – Konfiguracja

```
!
ip vrf VPN-R1
   rd 3:27
   inter-as-hybrid next-hop 10.1.1.2
!
ip vrf VPN2
   rd 100:2
   inter-as-hybrid next-hop 10.2.2.2
!
interface gi1/0.1
 encapsulation dot1q 1
 ip address 10.0.0.1 255.255.255.0
!
interface gi1/0.27
 encapsulation dot1q 27
 ip vrf forwarding VPN-R1
 ip address 10.1.1.1 255.255.255.0
!
interface gi1/0.100
 encapsulation dot1q 100
 ip vrf forwarding VPN2
 ip address 10.2.2.1 255.255.255.0
!
router bgp 100
 neighbor 10.0.0.2 remote-as 200
 neighbor <PE1-loopback lub RR1-loopback>
 neighbor <PE1-loopback lub RR1-loopback> update-source Loopback_X
 !
 address-family vpnv4
  neighbor 10.0.0.2 activate
  neighbor 10.0.0.2 send-community extended
  neighbor 10.0.0.2 inter-as-hybrid
  neighbor <PE1-loopback lub RR1-loopback> activate
  neighbor <PE1-loopback lub RR1-loopback> send-community extended
!
```

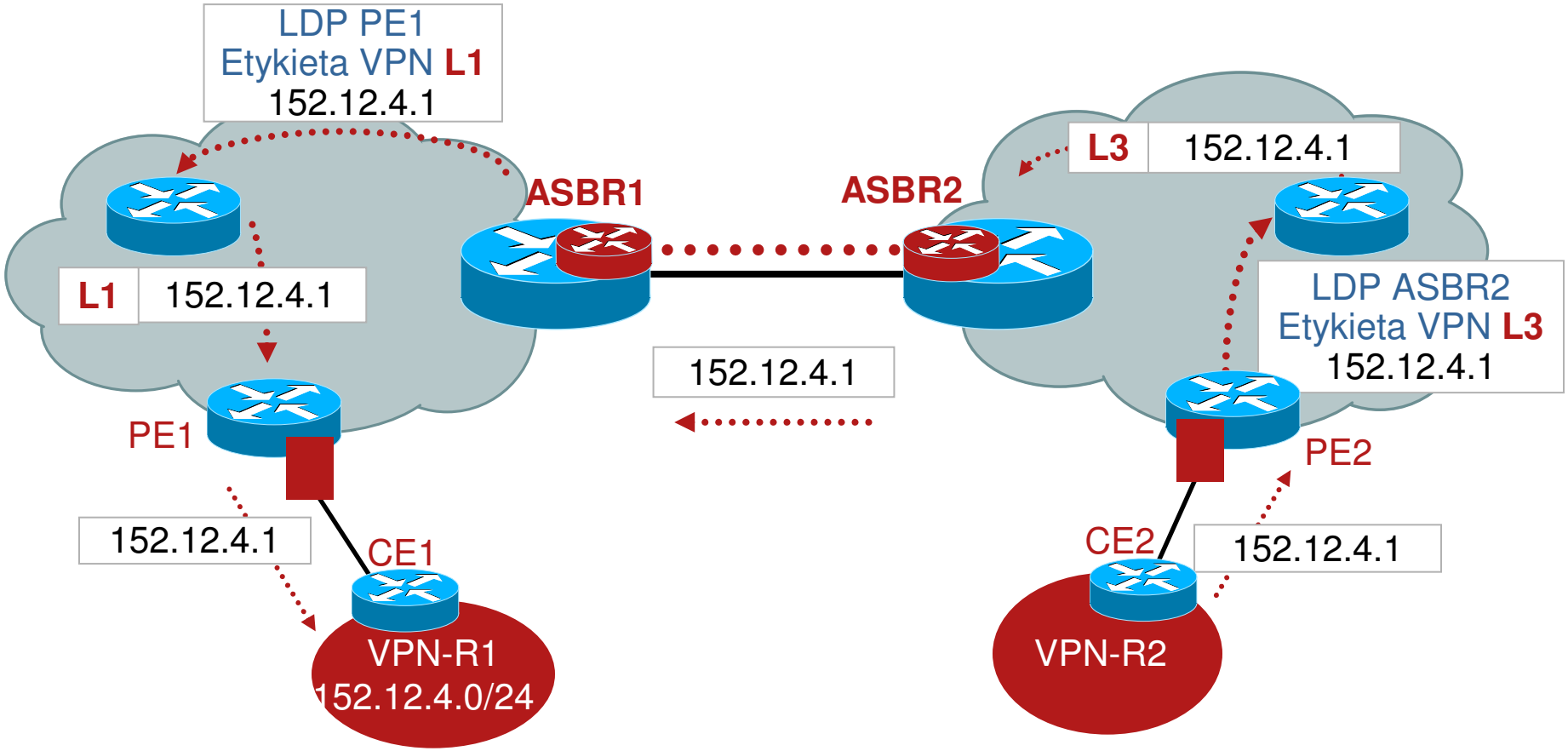
Włączenie inter-as-hybrid

Oddzielny interfejs dla sesji MP-BGP

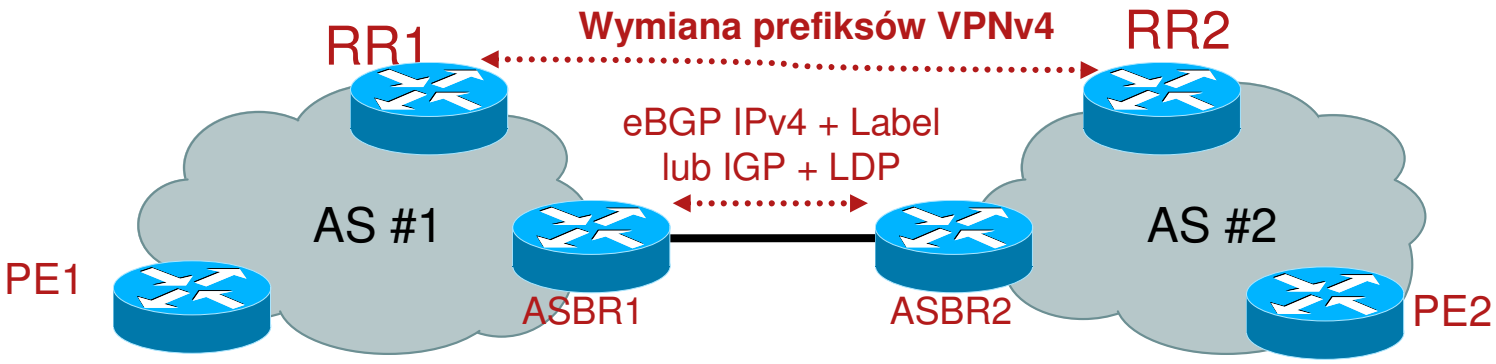
Dedykowany intf VPN-R1

Sesja BGP VPNv4 z ASBR drugiego operatora

Inter-AS VPN Opcja AB — Warstwa danych



Inter-AS VPN Opcja C
Multi-hop EBGP VPNv4 pomiędzy RR



Routery ASBR nie posiadają informacji o etykietach i prefiksach VPN. Prefiksy VPNv4 są wymieniane przez routery RR.

ASBR wymieniają między sobą adresy interfejsów Loopback routerów PE oraz etykiety do nich przydzielone.

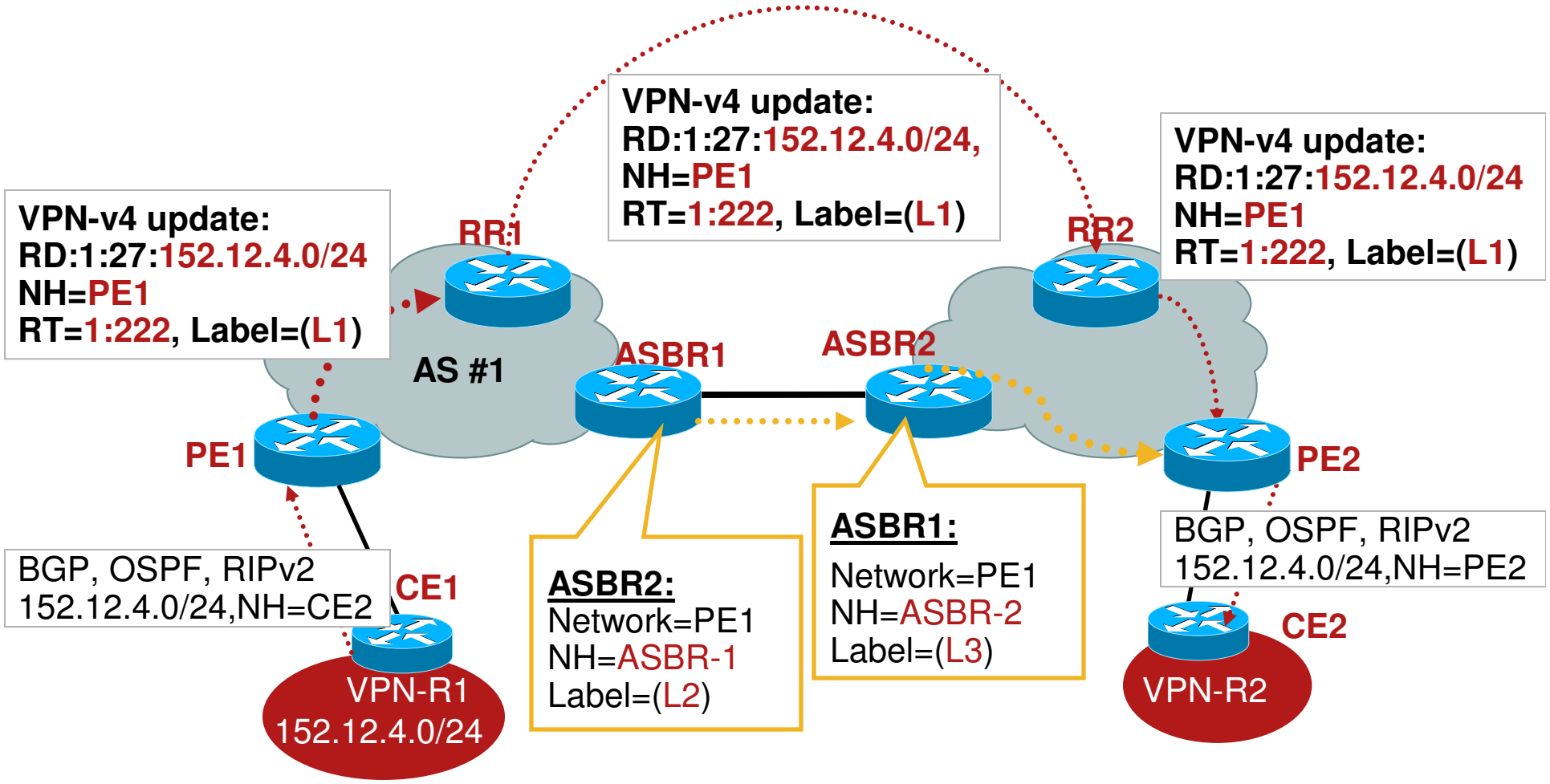
Dwie opcje dystrybucji etykiet dla adresu next-hop BGP:

IGP + LDP lub eBGP IPv4 + Label

Wymiana etykiet VPN pomiędzy routerami RR umożliwia zestawienie ścieżki LSP od początku do końca.

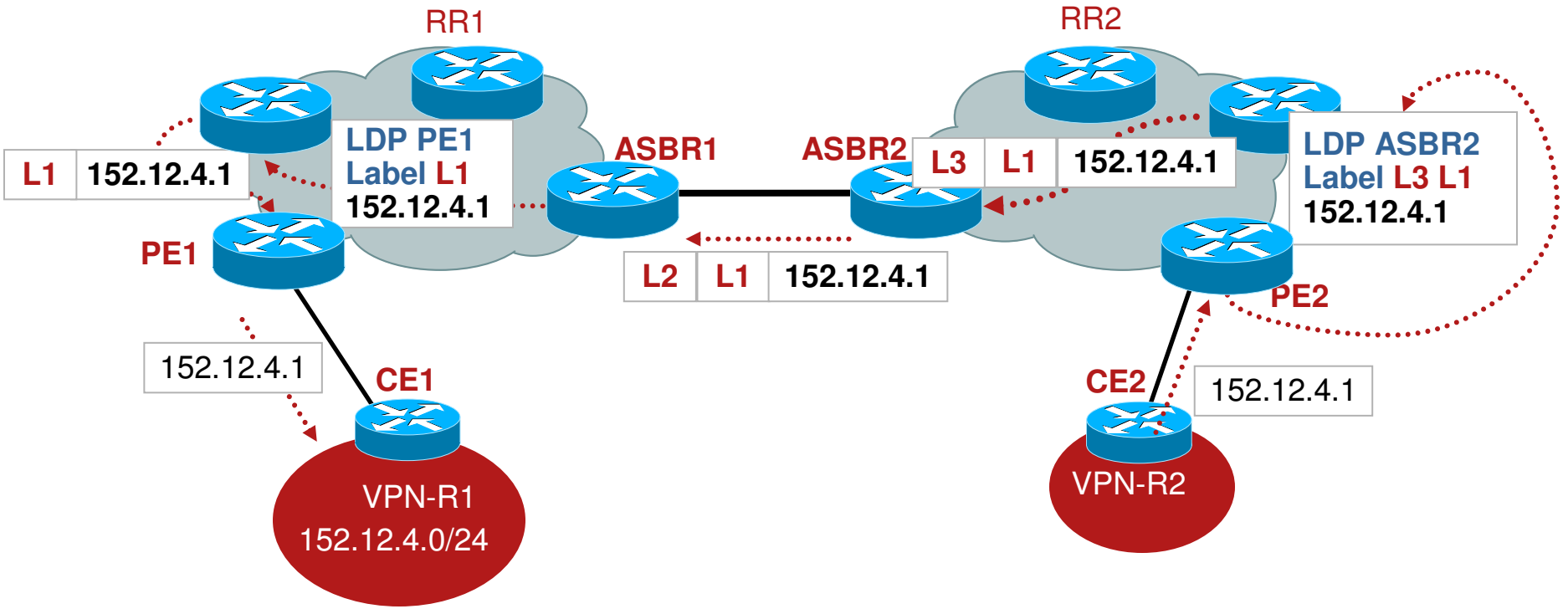
Next-Hop-Self powinien być wyłączony na routerach RR, gdyż pakiety danych nie powinny przechodzić przez RR.

I-AS VPN Opcja C — Warstwa kontrolna



Multihop eBGP dla VPNv4 z next-hop-unchanged na RR

I-AS VPN Opcja C — Warstwa danych



Inter-AS Bezpieczeńswo

Uwierzytelnianie MD5 sesji LDP, BGP, IGP

Ograniczyć ilość prefiksów per VRF

VPN ID

```
ip vrf PLNOG
 vpn id 1234EF:10
!
```

Identyfikuje jednoznacznie każdy VPN w celu prawidłowego zarządzania wewnątrz domeny, jak i pomiędzy operatorami.

vpn id ma postać **oui:vpn-index**

oui (adres LAN MAC o długości 3 oktetów przydzielany przez IEEE)

vpn-index (wartość o długości 4 oktetów, identyfikuje każdy VPN)

Statyczna alokacja etykiet

Sprawdzanie TTL w celu przeciwdziałania atakom DoS

Filtrowanie BGP ASPATH, sprawdzanie ext communities, RD.

Import/export prefiksów do/z VRF tylko wyznaczonych prefiksów.

Inter-AS L3VPN Podsumowanie

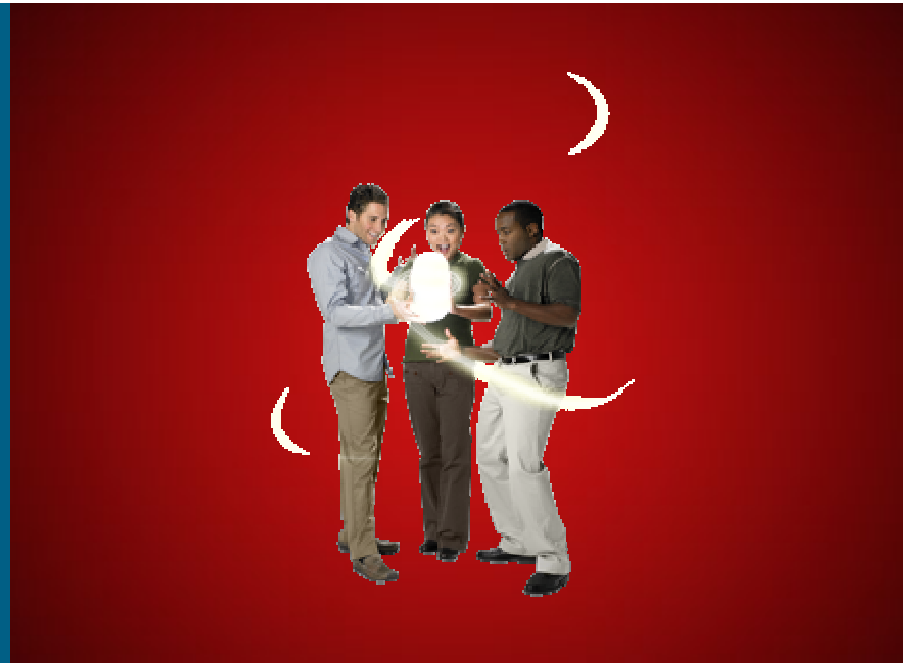
Opcja A - bezpieczna i najczęściej stosowana. Większa granularność polityk QoS, niż opcja B i C.

Opcja AB - najbardziej bezpieczna, a przy tym bardziej skalowalna, niż opcja A. Polityka QoS jak powyżej.

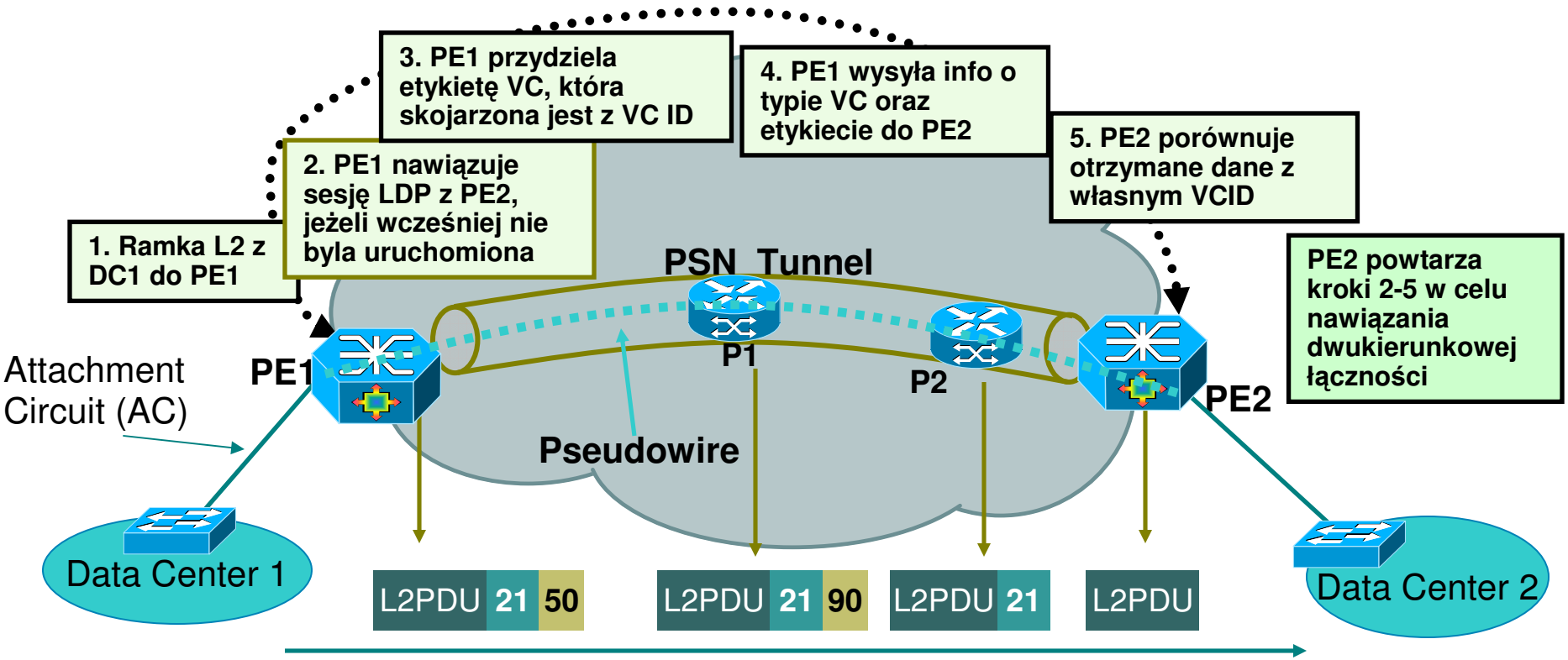
Opcja B - umiarkowanie inwazyjna, metoda bardziej skalowalna, niż opcja A i AB, ale też mniej bezpieczna.

Opcja C - najbardziej skalowalna, najbardziej inwazyjna metoda, stosowana głównie przez jednego operatora posiadającego wiele obszarów AS.

Inter-AS L2 VPN: Any Transport over MPLS (AToM)



AToM w ramach jednego AS

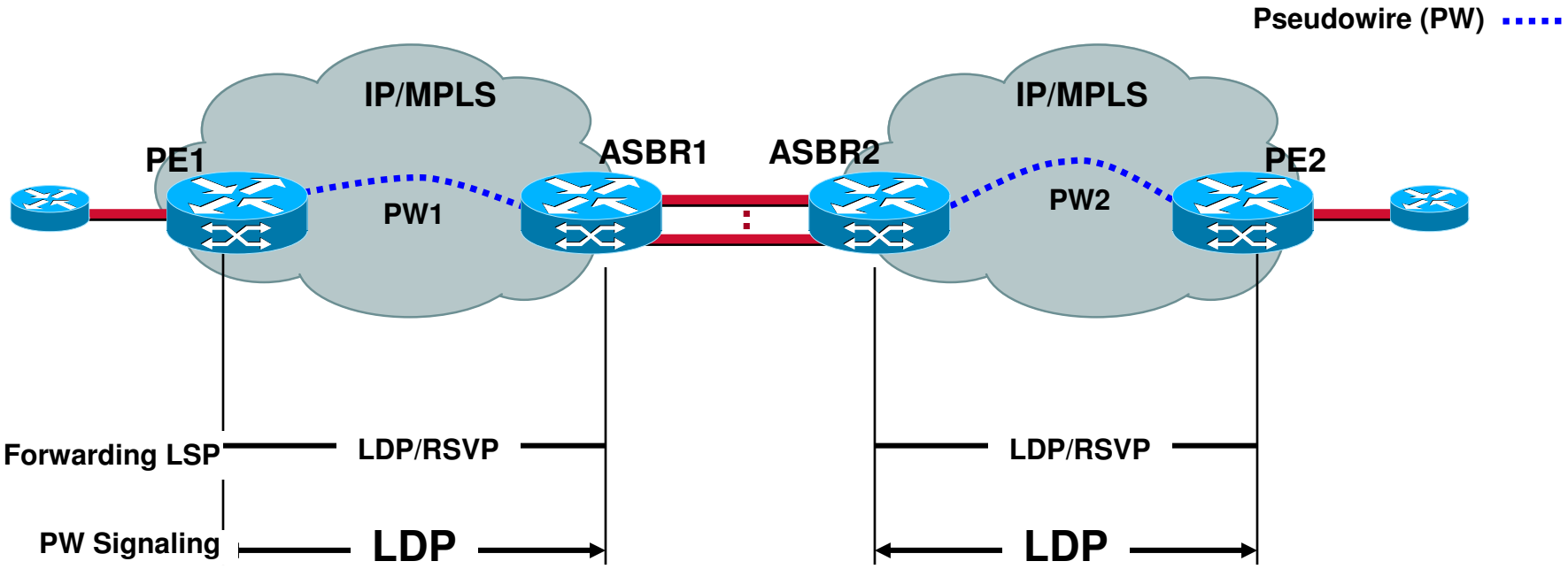


Sesja typu targeted LDP zapewnia wymianę informacji między PE.

VC ID identyfikuje połączenie L2. Etykieta VC identyfikuje tunel PW.

Wsparcie dla enkapsulacji HDLC, PPP, Ethernet, ATM, Frame Relay

Inter-AS AToM
Sąsiedztwo w warstwie 2 — Opcja A



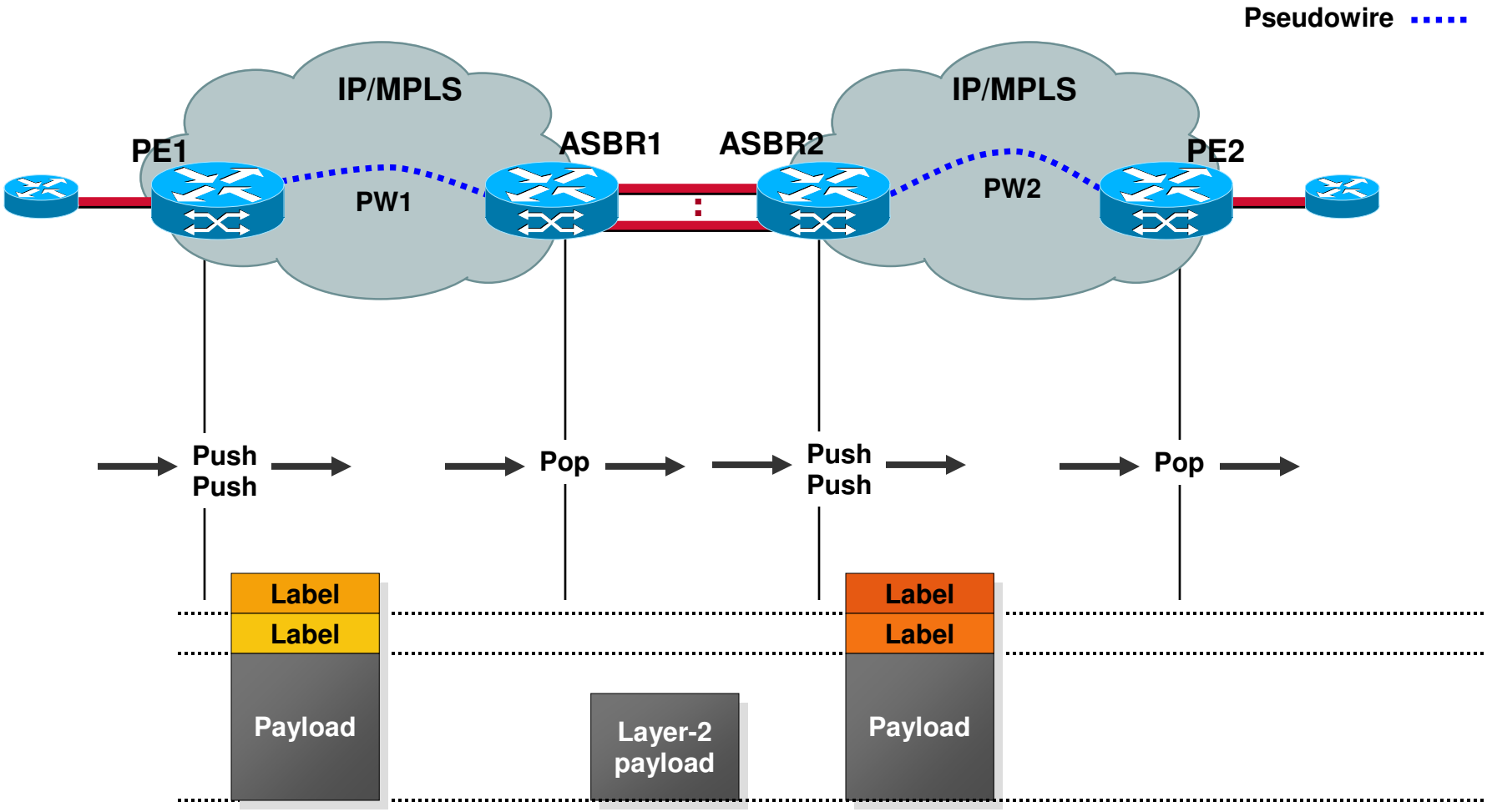
Jeden interfejs L2 dla każdego z tuneli PW z osobna.

Klarowny podział administracyjny między AS.

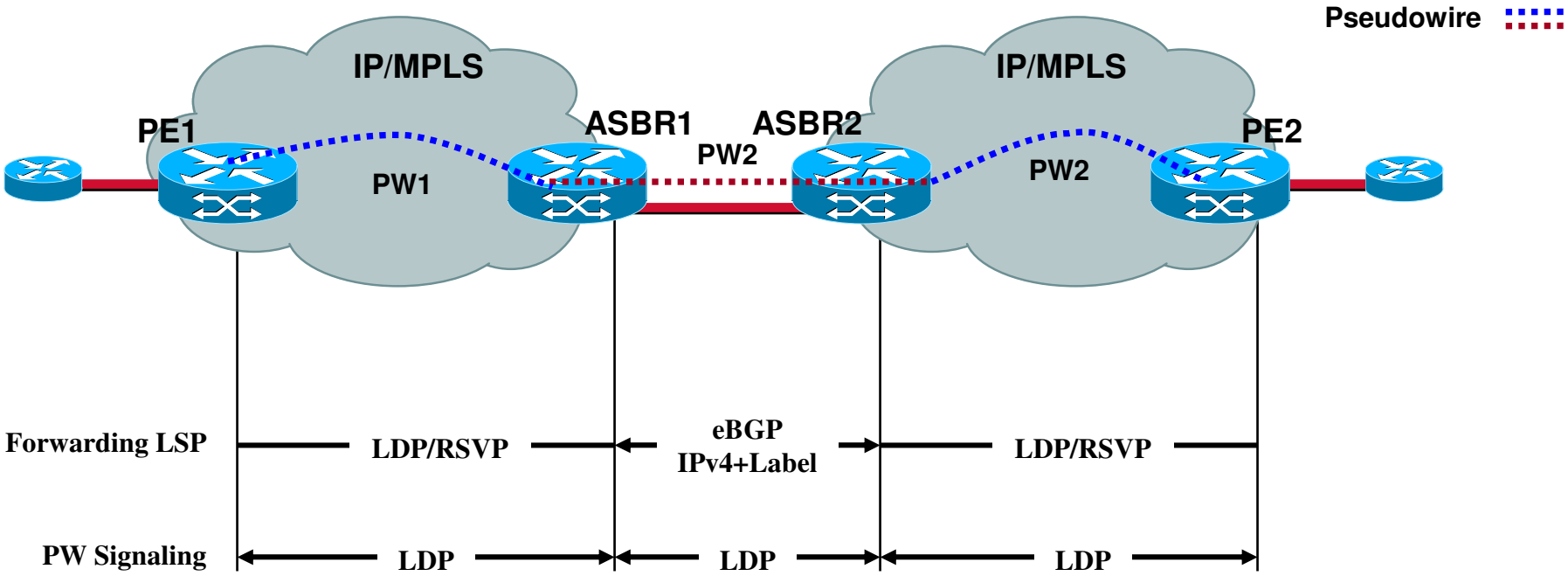
Zachowana autonomiczność protokołów sterujących.

Duże możliwości określania polityk QoS na styku między ISP.

Inter-AS AToM
Sąsiedztwo w warstwie 2 — Opcja A



Inter-AS AToM
Multi-Hop PW – Opcja B

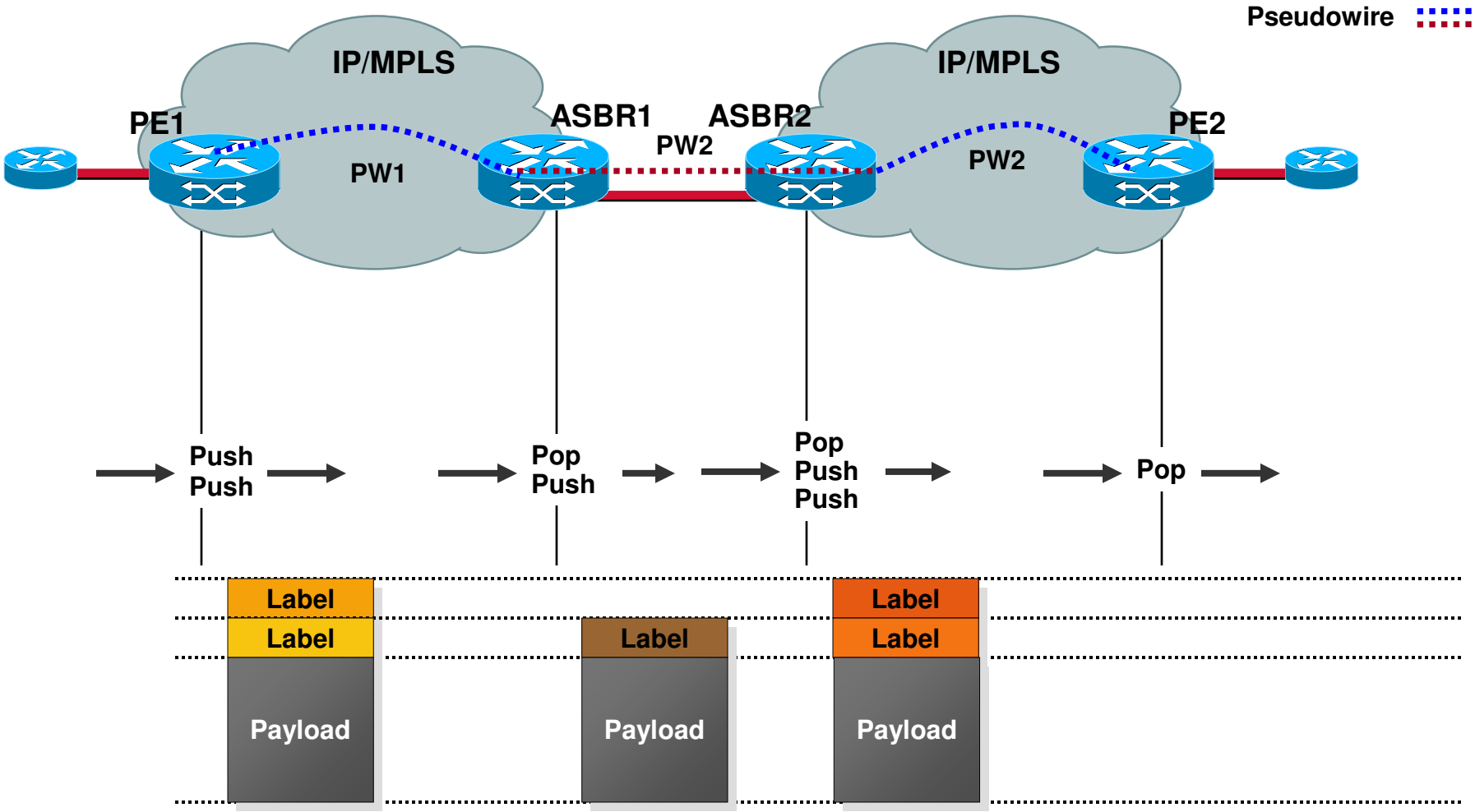


Pojedynczy interfejs pomiędzy ASBR, ruch etykietowany.

Operatorzy muszą wymienić między sobą jeden adress końca tunelu PW2

Routery P oraz PE nie znają adresów końcowych tunelu w sąsiednim AS.

Inter-AS L2VPN
Multi-Hop PW — Opcja B

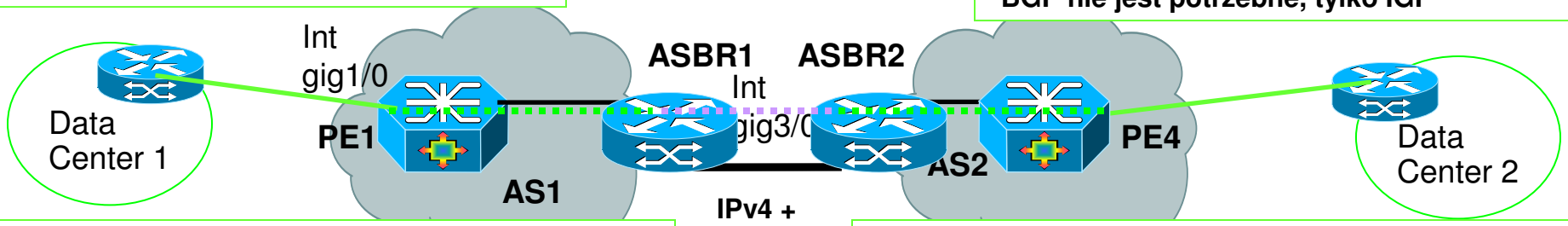


Inter-AS AToM Opcja B – Konfiguracja

Łączenie tuneli L2 (PW) na ASBR

```
!
HOSTNAME PE1
!
interface giga1/0
 xconnect <ASBR1> 10  encapsulation mpls
 !
*BGP nie jest potrzebne, tylko IGP
```

```
!
HOSTNAME PE4
!
interface giga1/0
 xconnect <ASBR2> 20  encapsulation mpls
!
*BGP nie jest potrzebne, tylko IGP
```

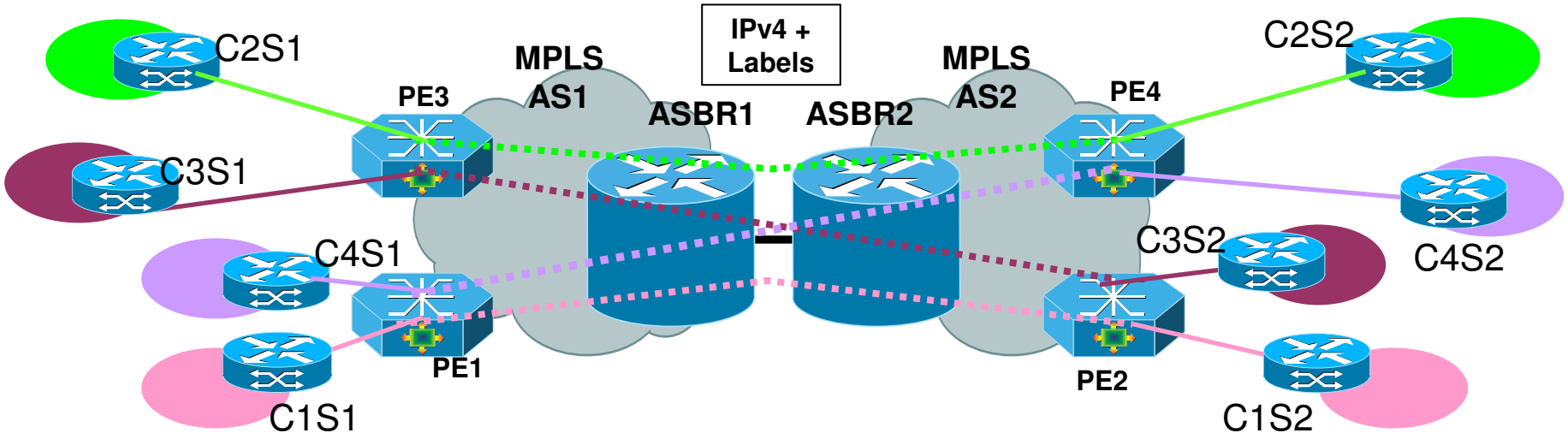


IPv4 + Labels między ASBR

```
HOSTNAME ASBR1
!
pseudowire-class pw-switch
encapsulation mpls
!
l2 vfi pw-switch point-to-point
neighbor <ASBR2>  100 pw-class pw-switch
neighbor <PE3> 10 pw-class pw-switch
!
Interface giga3/0
mpls bgp forwarding
!
! router bgp 1
neighbor <ASBR2-WAN> remote-as 2
neighbor <ASBR2-WAN> send-label
exit-address-family
 !
*Trzeba również rozgłosić adres Loopback (xconnect ID)
ASBR1 poprzez IGP (AS1) oraz eBGP do ASBR2.
```

```
HOSTNAME ASBR2
!
pseudowire-class pw-switch
encapsulation mpls
!
L2 vfi pw-switch point-to-point
neighbor <ASBR1> 100  pw-class pw-switch
neighbor <PE4> 20  pw-class pw-switch
!
Interface giga3/0
mpls bgp forwarding
!
router bgp 1
neighbor <ASBR1-WAN> remote-as1
neighbor <ASBR1-WAN> send-label
exit-address-family
 !
*Trzeba również rozgłosić adre Loopback ASBR2 do AS2 oraz
ASBR1 poprzez eBGP.
```

Inter-AS AToM Opcja C — IPv4+Label

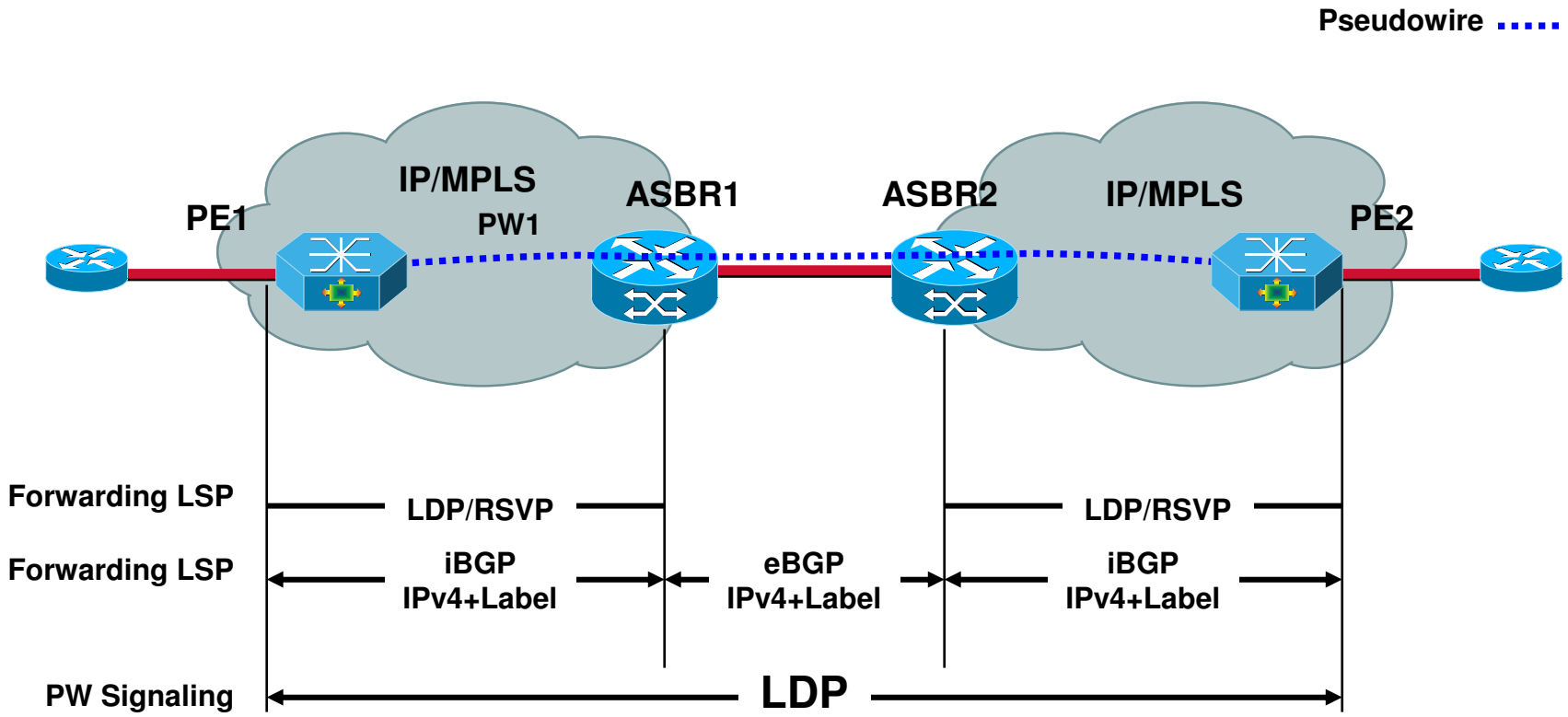


Sygnalizacja tunelu L2 (pseudowire - PW) jest przesyłana między operatorami.

ASBR nie uczestniczą w sygnalizacji PW. Sygnalizacja LSP między operatorami może być zestawiona za pomocą BGP, BGP/LDP lub I-AS TE.

Urządzenia ASBR mogą służyć za routery agregacyjne, które będą ukrywać poszczególne adresy routerów PE należących do tego samego AS.

Inter-AS AToM — Opcja C

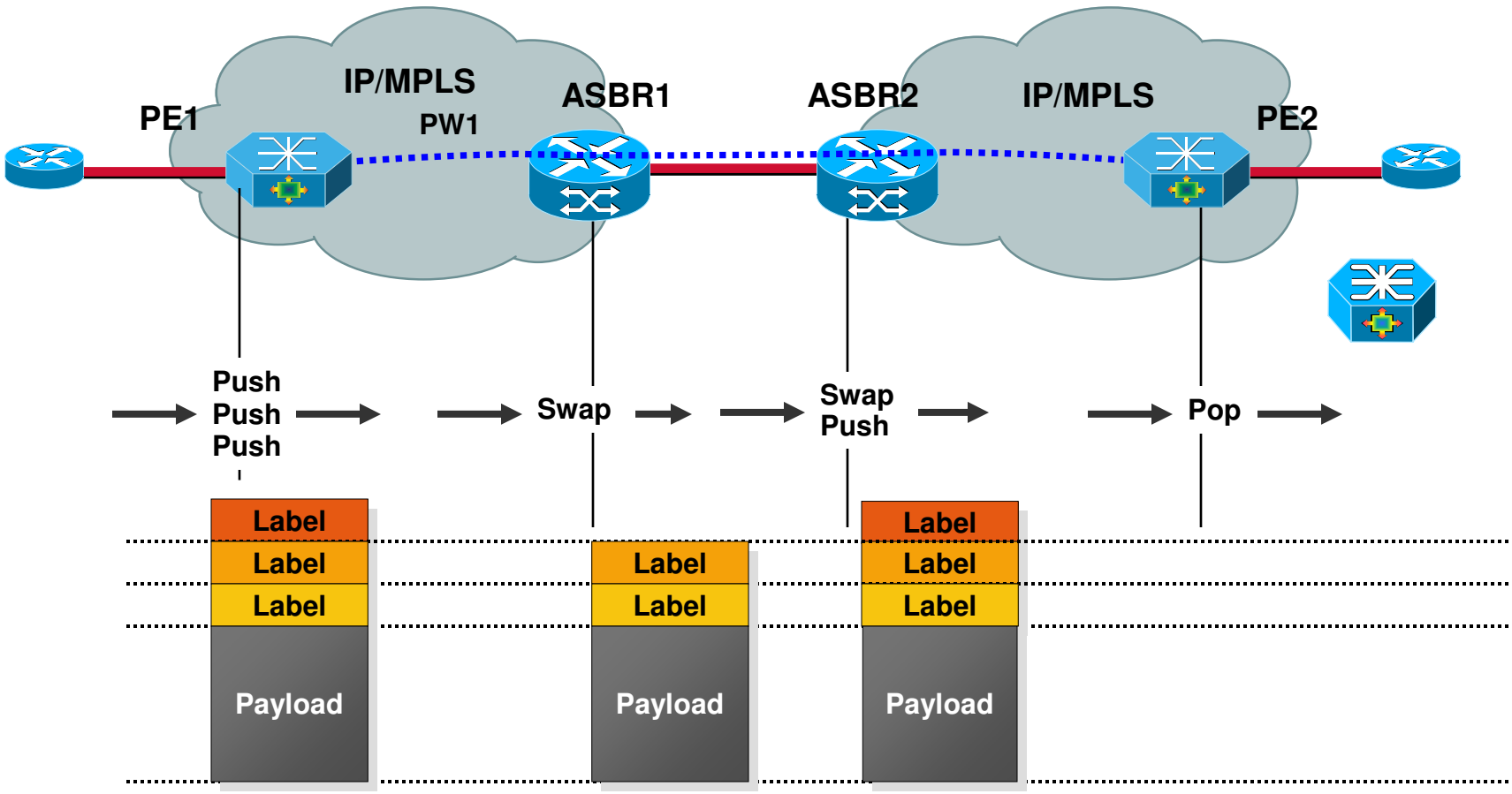


Pojedynczy, etykietowany interfejs pomiędzy ASBR.

Adresy końców tunelu PW są przekazywane za pomocą eBGP IPv4+label między operatorami oraz iBGP IPv4+label wewnątrz AS.

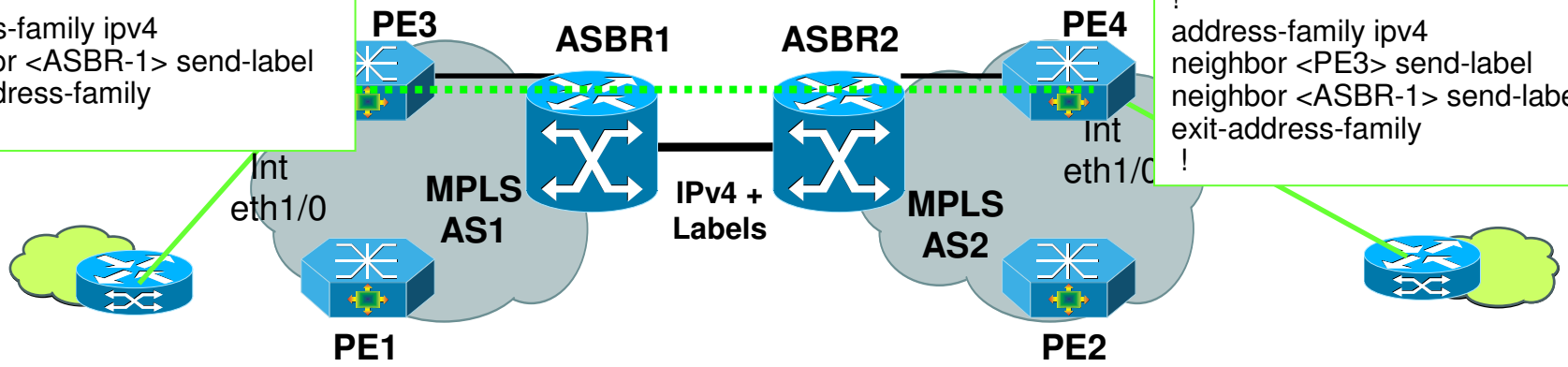
Tylko routery PE oraz ASBR znają adresy tuneli PW.

Inter-AS AToM — Opcja C



Inter-AS AToM Opcja C — konfiguracja

```
!
HOSTNAME PE3
!
interface Ethernet1/0
 xconnect <PE4> 100
encapsulation mpls
 !
! Włączenie rozgłaszania etykiet.
router bgp 1
!
address-family ipv4
neighbor <ASBR-1> send-label
exit-address-family
 !
```

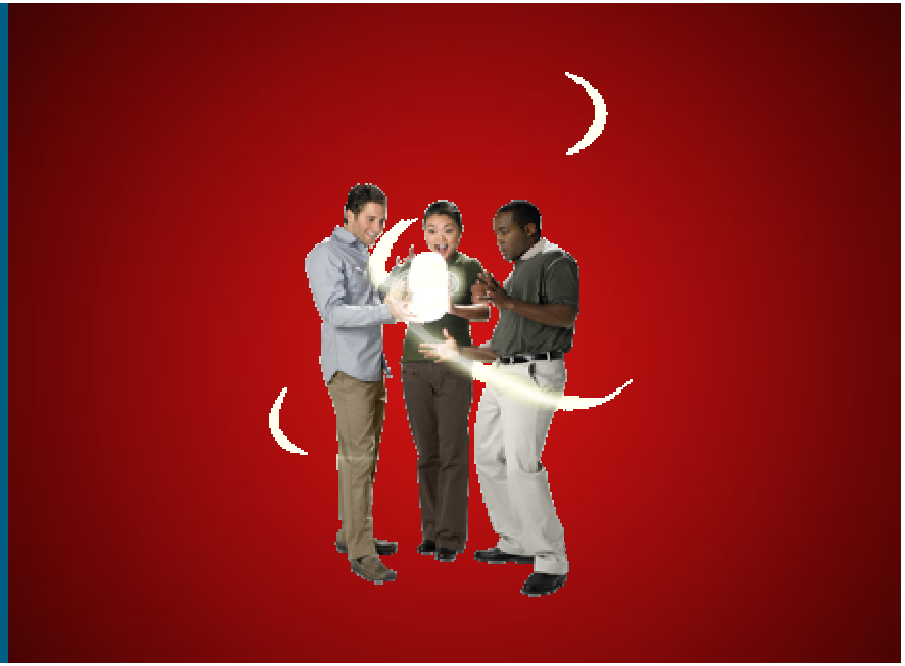



```
!
HOSTNAME PE4
!
interface Ethernet1/0
 xconnect <PE3> 100
encapsulation mpls
!
!Włączenie rozgłaszania etykiet
router bgp 1
!
address-family ipv4
neighbor <PE3> send-label
neighbor <ASBR-1> send-label
exit-address-family
 !
```

```
!
HOSTNAME ASBR1
!  Włączenie rozgłaszania etykiet.
router bgp 1
!
address-family ipv4
neighbor <PE1> send-label
neighbor <ASBR-1> send-label
exit-address-family
 !
```

```
!
HOSTNAME ASBR2
!  Włączenie rozgłaszania etykiet.
router bgp 1
!
address-family ipv4
neighbor <PE4> send-label
neighbor <ASBR-1> send-label
exit-address-family
 !
```

Inter-AS MPLS TE



Zestawienie tunelu MPLS TE w jednym AS

Router zestawiający tunel uczy się topologii sieci korzystając z:

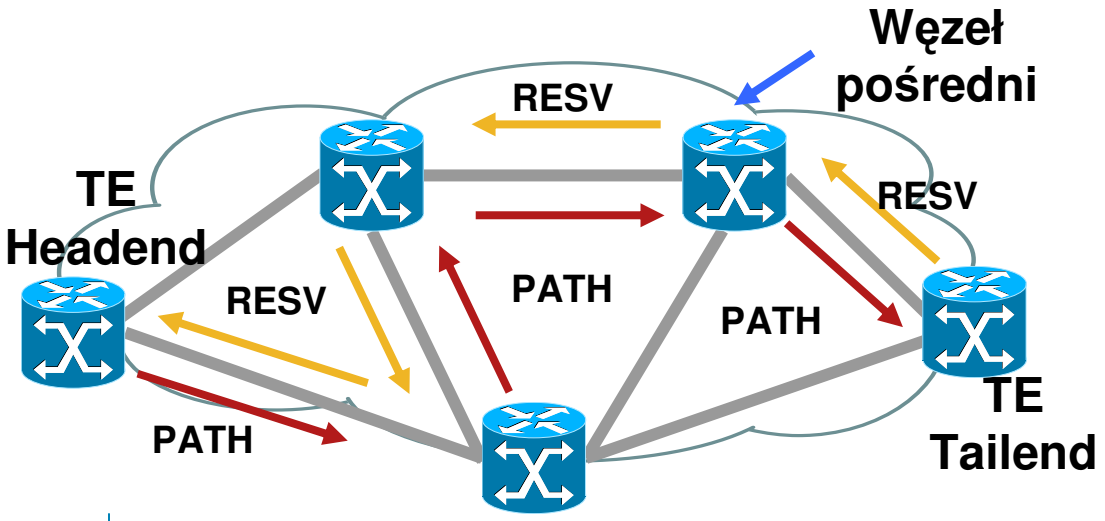
- ISIS-TE
- OSPF-TE

Przeliczenie ścieżek (CSPF)

Path Setup (RSVP-TE):

- Label_Request (PATH)
- Label (RESV)
- Explicit_Route Object
- Record_Route (Path/RESV)
- Session_Attribute (Path)

Propagacja tablicy LFIB poprzez etykiety RSVP



Pakiety wysłane w tunel za pomocą:

- Przekierowania statycznego
- Autoroute
- PBR
- CBTS
- Tunnel Select
- Forwarding Adjacency

Pakiety danych przechodzą wyznaczoną ścieżką LSP dla tunelu TE, nie IGP LSP.

Inter-Domain Traffic Engineering

Problem:

Routery początkowy i końcowy tunelu TE znajdują się w różnych domenach administracyjnych.

Informacje protokołu IGP nie są wymienianie między operatorami. Router początkowy tunelu nie zna pełnej topologii, która ma posłużyć do wyznaczenia prawidłowej ścieżki.

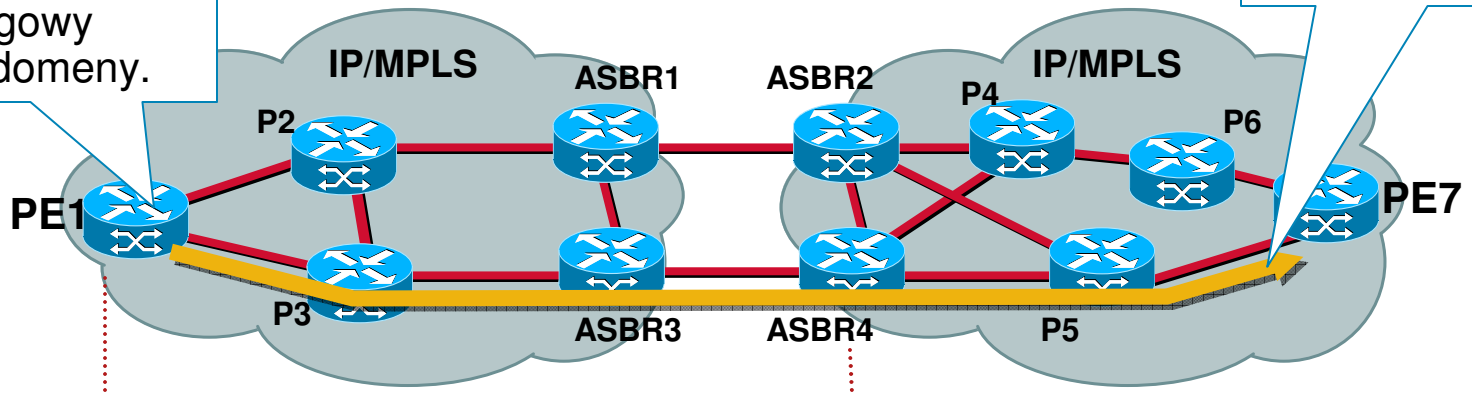
Rozwiązanie:

Rozszerzenie Explicit Route Object (ERO) o tryb „Loose”. Przekazywany parametr Node-id, który pomaga ustalić węzeł do ominięcia oraz punkt styku z operatorem.

Przeliczenie ścieżki z wykorzystaniem Explicit Route Object

Router PE określa ścieżkę poprzez router brzegowy sąsiedniej domeny.

Przeliczenie ścieżki zakończone podczas zestawienia TE LSP.

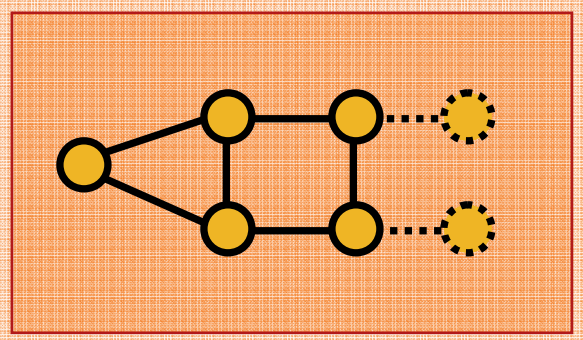


Loose ERO
ASBR4
PE7

rozwinięcie

Strict ERO
P3, ASBR3, ASBR4
PE7 (Loose)

Topologia na R1

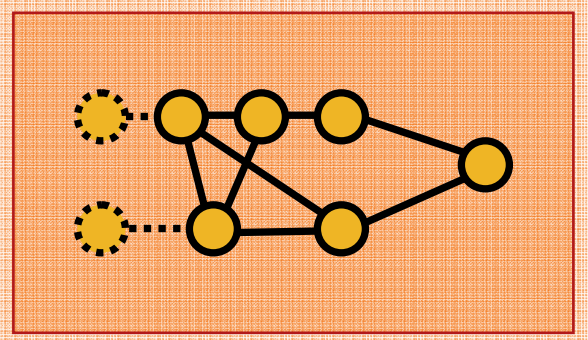


Loose ERO
PE7

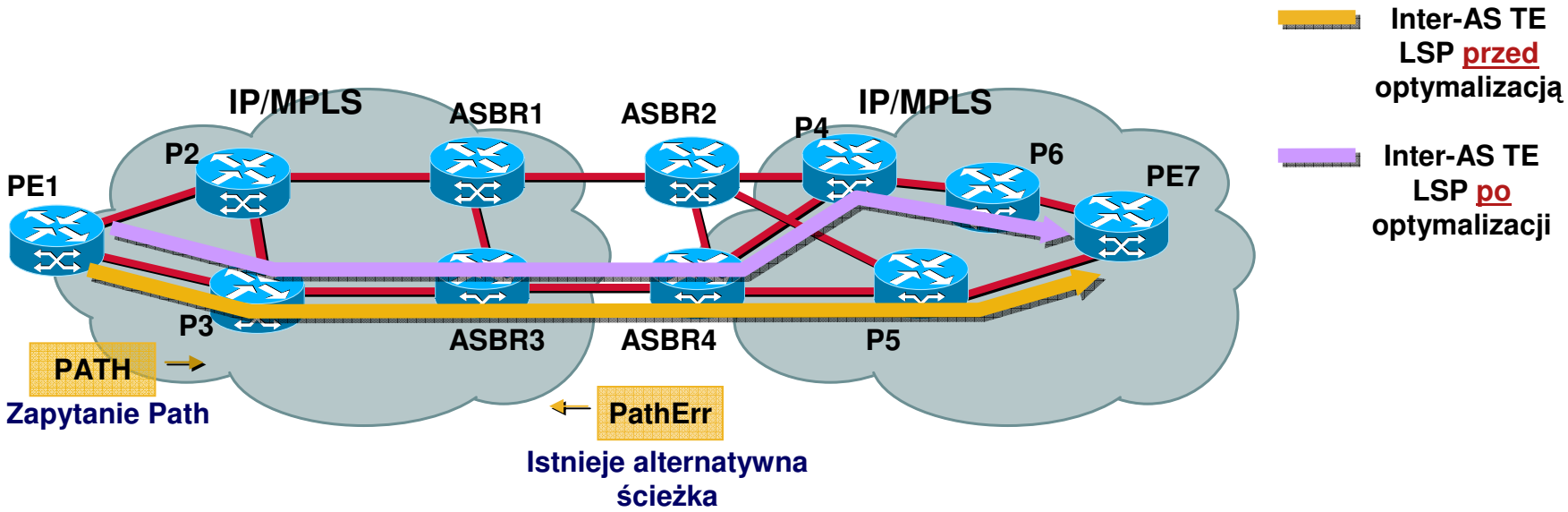
rozwinięcie

Strict ERO
P5, PE7

Topologia widoczna z ASBR4



Inter-Domain TE — Optymalizacja ścieżki TE LSP

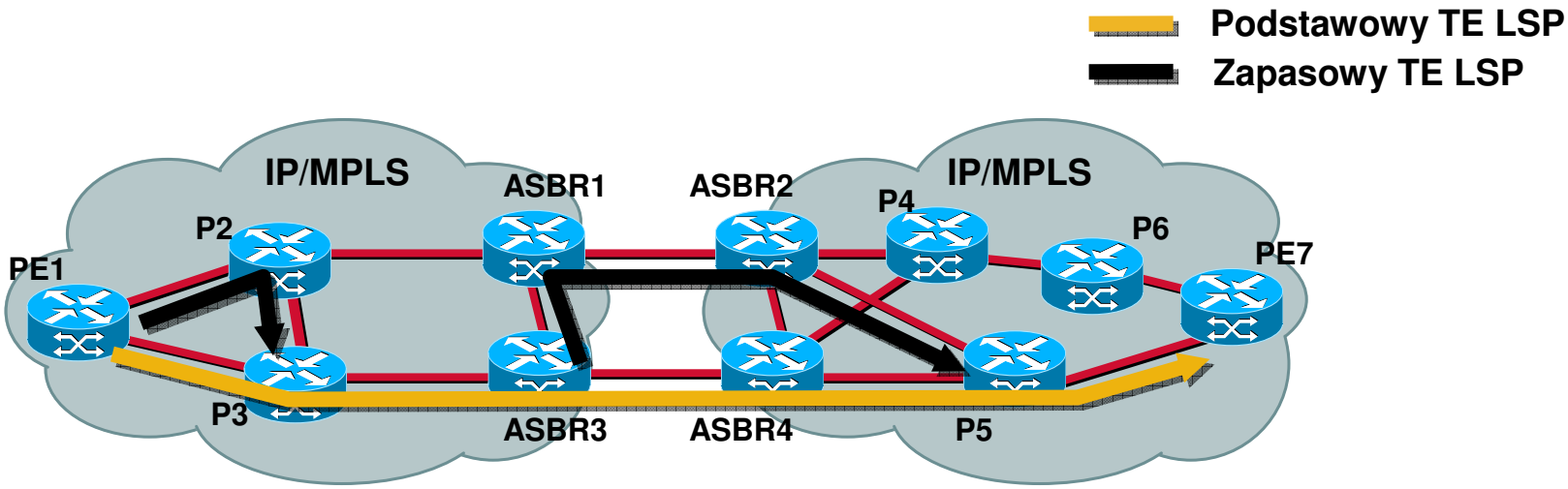


Optymalizacja włącza się według ustawień liczników czasu/ wydarzeń w sieci/ uruchomienia przez administratora.

Router początkowy ustawia flagę „path re-evaluation request” (SESSION_ATTRIBUTE).

Router PE1 otrzymuje komunikat PathErr jeżeli istnieją inne ścieżki.

Inter-Domain TE — Fast Re-Route



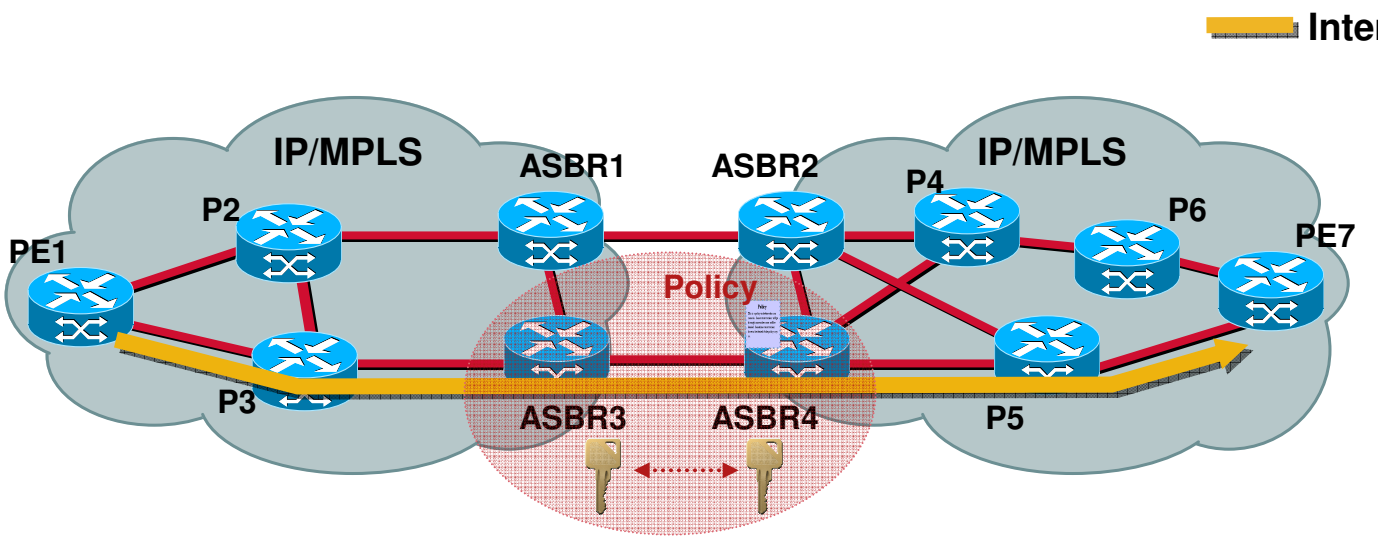
Konfiguracja taka, jak dla scenariusza MPLS TE w jednej domenie.

Link oraz Node protection dla ASBR oraz połączenia między AS.

Wsparcie dla obiektu Node-id jest wymagane w celu implementacji protekcji węzła ABR/ASBR.

Flaga Node-id zdefiniowana w draft-ietf-nodeid-subobject.

Inter-Domain TE — Bezpieczeństwo



ASBR może wymagać określonych parametrów w polityce lokalnej podczas zestawiania tunelu MPLS TE z innym ASBR.

Route Recording może być ograniczony.

ASBR może modyfikować adres źródłowy pakietów PathErr wysyłanych z własnego AS.

ASBR wspierają uwierzytelnienie RSVP (MD5/SHA-1)

Konfiguracja Tunelu Inter-AS MPLS TE

```
mpls traffic-eng tunnels
!
interface Tunnel1
 ip unnumbered Loopback0
 no ip directed-broadcast
 tunnel destination 172.31.255.7
 tunnel mode mpls traffic-eng
 tunnel mpls traffic-eng priority 7 7
 tunnel mpls traffic-eng bandwidth  1000
 tunnel mpls traffic-eng path-option 10 explicit name LOOSE-PATH
!
ip route 172.31.255.5 255.255.255.255 Tunnel1
!
ip explicit-path name LOOSE-PATH enable
 next-address loose 172.16.255.3
 next-address loose 172.31.255.4
!
```

Trasa statyczna do drugiego AS

Statyczne mapowanie ruchu do tunelu `Tunnel1`

Routery ASBR jako węzły typu „loose”

Konfiguracja Inter-AS TE na ASBR

```
mpls traffic-eng tunnels
!
key chain A-ASBR1-key
 key 1
  key-string 7 151E0E18092F222A
!
interface Serial1/0
 ip address 192.168.0.1 255.255.255.252
 mpls traffic-eng tunnels
 mpls traffic-eng passive-interface nbr-te-id 172.16.255.4 nbr-igp-id ospf 172.16.255.4
 ip rsvp bandwidth
 ip rsvp authentication key-chain A-ASBR1-key
 ip rsvp authentication type sha-1
 ip rsvp authentication
!
router bgp 65024
 no synchronization
 bgp log-neighbor-changes
 neighbor 172.24.255.3 remote-as 65024
 neighbor 172.24.255.3 update-source Loopback0
 neighbor 192.168.0.2 remote-as 65016
 no auto-summary
!
ip rsvp policy local origin-as 65016
 no fast-reroute
 maximum bandwidth single 10000
 forward all
!
```

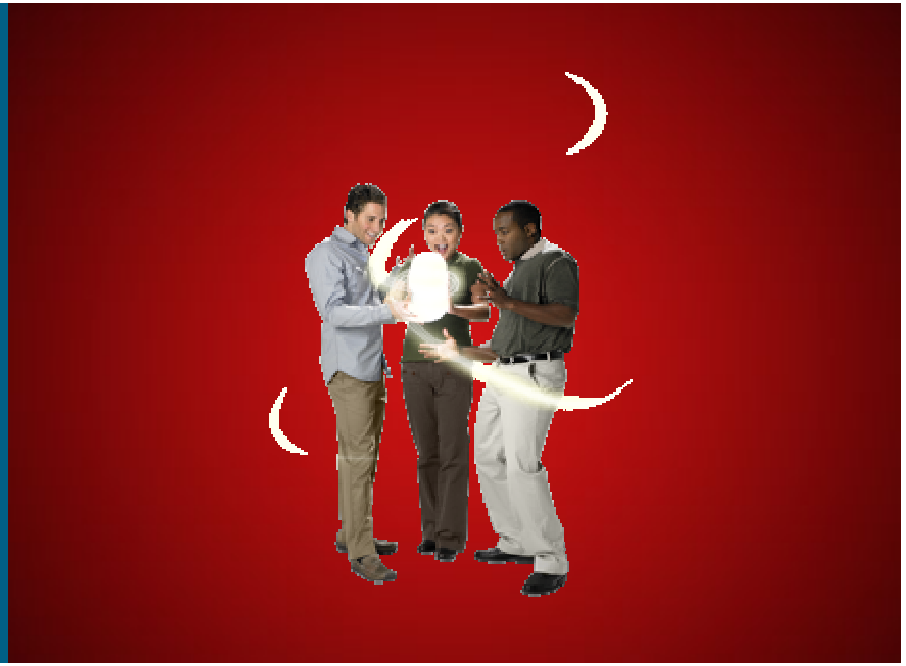
Klucz Auth do RSVP

Dodanie linku ASBR do topologii TE

Włączenie uwierzytelniania

Włączenie procesu sygnalizacji dla AS 65016

Inter-Provider MPLS Solutions: CSC



Modele usługowe Carrier Supporting Carrier

Customer Carrier oparty wyłącznie o protokół IP.

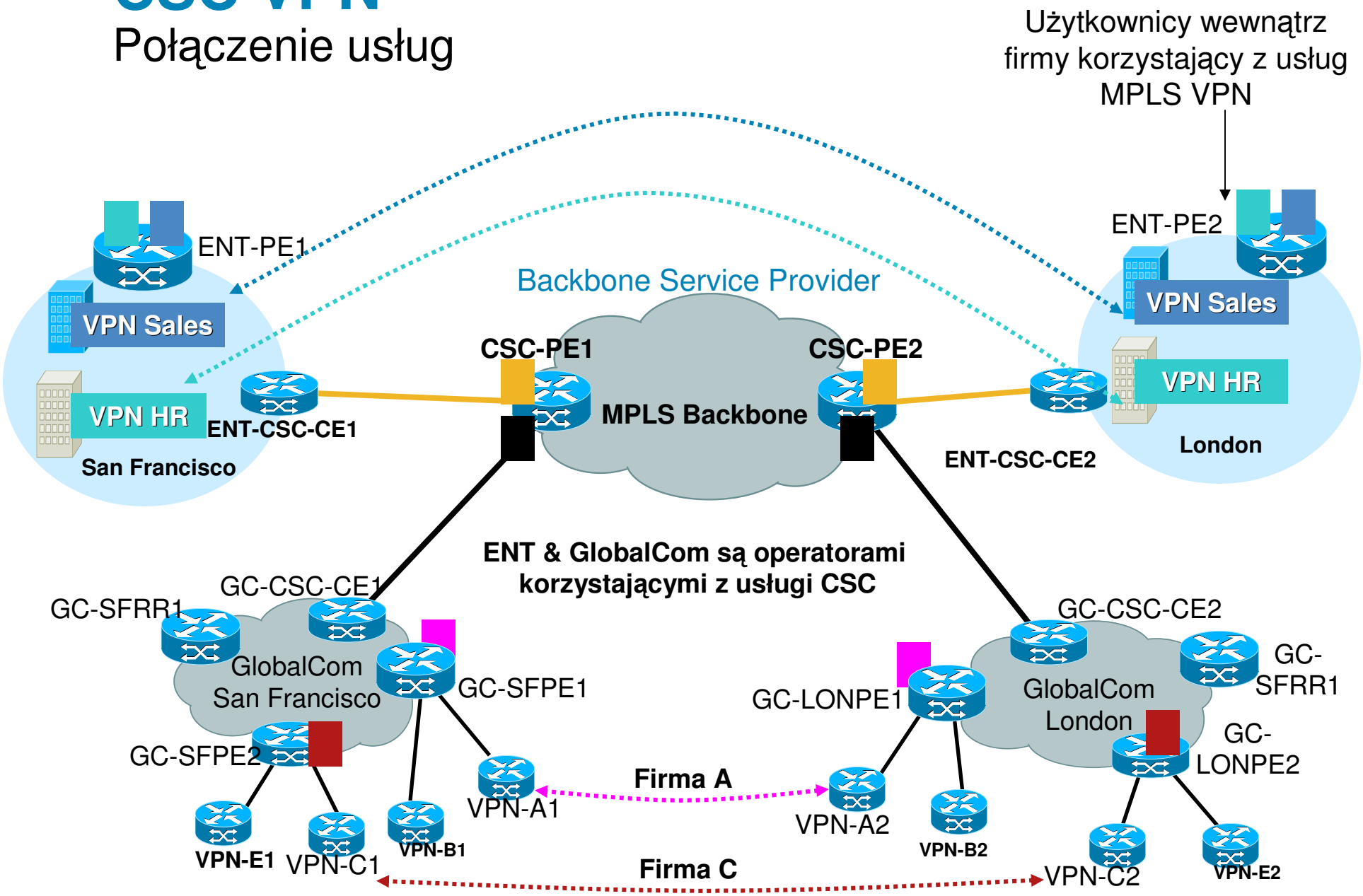
Customer Carrier z siecią MPLS.

Customer Carrier wspiera MPLS VPN.

Customer Carrier jest operatorem korzystającym z usługi CSC świadczonej przez innego operatora.

CSC VPN

Połączenie usług



CSC Warstwa kontrolna

Konfiguracja sygnalizacji i sterowania oparta na tych samych zasadach, co konfiguracja sieci MPLS VPN w ramach domeny.

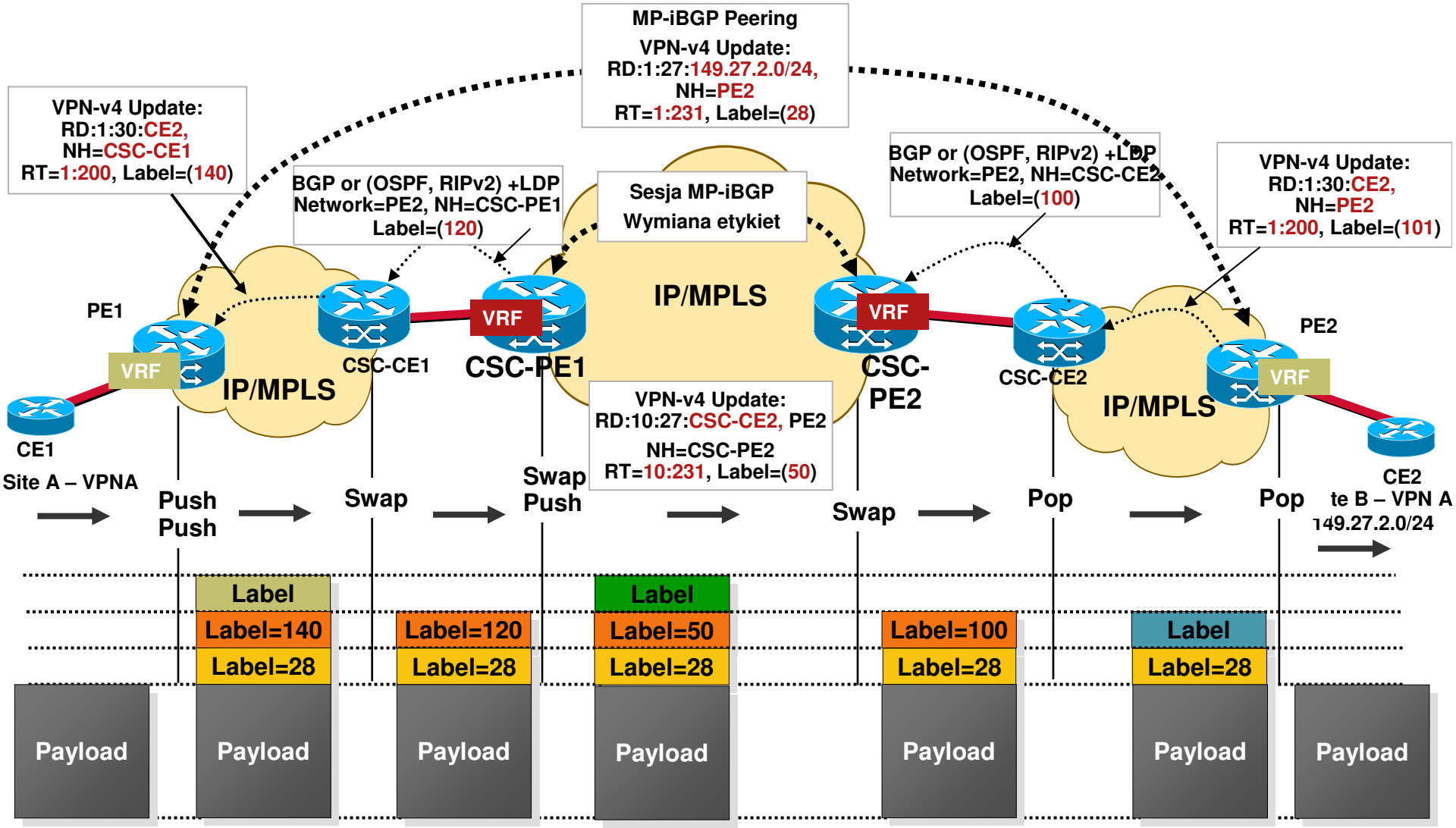
Link CSC-CE - CSC-PE jest połączeniem do wymiany wymiany prefiksów operatorów Customer Carrier. Informacje wymieniane są przy użyciu:

1. Trasy statyczne lub 2. Protokołu IGP lub 3. Sesji eBGP

Link CSC-PE - CSCE-CE jest etykietowany.

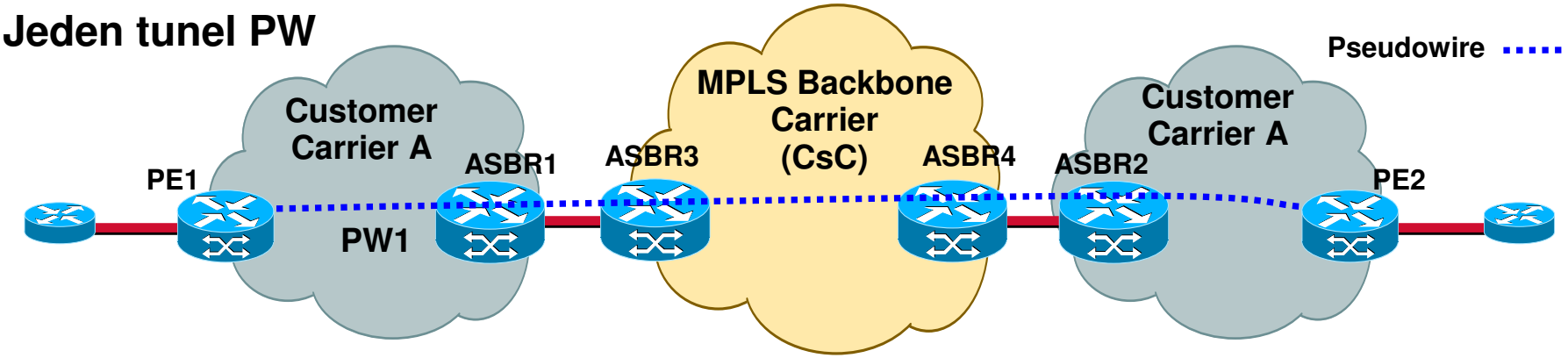
Operator CSC nie zna prefiksów przenoszonych przez operatorów Customer Carrier.

CSC Model nr 3 — Połączenie MPLS VPNs

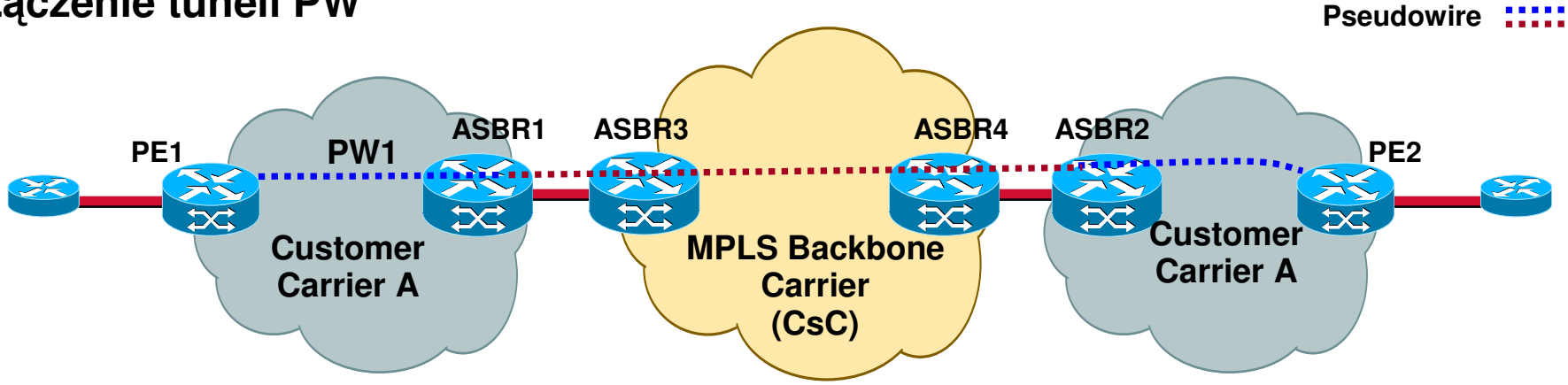


MPLS L2VPN poprzez sieć CSC

Jeden tunel PW



Łączenie tuneli PW



CSC Podsumowanie

CSC wspiera hierachiczny VPN

Informacje o prefiksach i routingu wewnątrz VPN pozostają transparentne dla operatora nadrzędnego CSC.

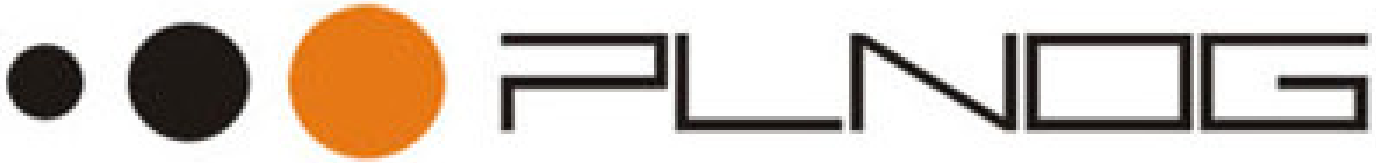
Można wykorzystać statyczną alokację etykiet MPLS do prefiksów IPv4 określających next-hop routera CSC-CE.

Polityka QoS jest oparta na znacznikach EXP. Powinna być uzgodniona między operatorami.

Mechanizmy dostępne w ramach CSC:

- MPLS IPv6 VPN
- Multicast VPN
- MPLS L2 VPN
- MPLS TE

Pytania?


PLNOG