Grands réseaux

Multi-services network

Cédric Llorens

Some references

Les cours sont disponibles à :

https://sites.google.com/site/courscedricllorens/home

Quelques références Internet:

http://www.ripe.net/
http://www.nanog.org/
http://www.radb.net/

Le Journal MISC :

http://www.miscmag.com/fr

Some references

Les Tableaux de bord de la sécurité réseau, 3ème édition, 562 pages, C.Llorens, L.Levier, D.Valois, B.Morin, Eyrolles, 2010, http://www.eyrolles.com

Mesure de la sécurité "logique" d'un réseau d'un opérateur de télécommunications, http://pastel.paristech.org/archive/00001492/

Agenda

- **Description of a multi-services network**

Protocols :

- MPLS
- LDP
- MPLS-TE (&& RSVP-TE && ISIS-TE)
- Diffserv (&& Diffserv-TE)

Services

- Ethernet over MPLS
- VPLS aka VPN layer 2
- VPN aka VPN layer 3
- etc.

Interconnections MPLS networks

Introduction

Introduction

1950 ~ Telex et routage des messages par des êtres humains.

1970 ~ Apparition d'une multitude de protocoles de services réseaux (type A, type B pour les compagnies aériennes :-)

1980 ~ Commutation de paquets avec routage statique. Le codage du système d'exploitation des noeuds réseaux est réalisé par l'opérateur de télécommunications.

1990 ~ Commutation de paquets avec routage dynamique (x25, frameRelay, etc.). Le codage du système d'exploitation des noeuds réseaux n'est plus réalisé par l'opérateur de télécommunications.

2000 ~ Commutation de paquets sur des labels et naissance des réseaux multi-services MPLS (Multi Protocol Label Switching).

2010 ~ Généralisation des réseaux multi-services MPLS (Multi Protocol Label Switching) et introduction du concept SDN (Software Defined Network).

Introduction

La maîtrise de la sécurité de l'architecture

La conception de l'architecture d'un réseau multi-services est primordiale pour assurer une évolution en toute sécurité des services offerts.

Cet effort de conception assure une pérennité de l'architecture (consistance) et limite grandement les risques.

En revanche, si les principes de départ étaient mauvais, l'architecture du réseau ne serait alors pas consistante et toute évolution ne sera pas sans risque.

Enfin, toute modification ultérieure apportera alors son lot de problèmes de sécurité ou d'effets de bord plus ou moins contrôlables.

Introduction

La maîtrise du routage

La multiplication des protocoles de routage supportés et qui doivent cohabités tous ensemble au sein du réseau : IGP (Interior Gateway Protocol), BGP (Border Gateway Protocol), MP-BGP (Multi Protocol), PIM (Protocol Independent Multicast), etc.

Les topologies de routage à multi-niveaux engendrées par ces protocoles et qui doivent vérifier des contraintes très précises.

L'augmentation vertigineuse du nombre de routes à gérer avec des tailles prévisionnelles de l'ordre du million de routes, ainsi que la convergence des algorithmes de routage concernés.

La protection du cœur du réseau

La maîtrise du périmètre d'un réseau multi-services est fondamentale afin de protéger le protocole de routage interne, de distribution de tag, etc.. de toutes attaques pouvant faire effondrer le réseau et ses services. Ce périmètre doit aussi prendre en compte l'interconnexion à d'autres réseaux tels que les réseaux DSLs d'accès, le réseau Internet, les réseaux VPN BGP/MPLS, etc.

Le contrôle de ce périmètre repose avant tout sur les mécanismes offerts par les équipements réseaux composant ce périmètre. Ces mécanismes nécessitent à la fois une bonne compréhension du fonctionnement interne d'un équipement réseau, mais aussi des avantages et des inconvénients des techniques offertes.

Le contrôle des configurations

Avec la concentration des réseaux et la diversité des services offerts, les configurations deviennent désormais un axe crucial de sécurité. La maîtrise des configurations devient réellement majeure si on prend en compte la dimension associée à la taille des configurations.

En effet et dans des grands réseaux, on peut recenser plusieurs millions de lignes de configuration si on tient compte de tous les équipements réseaux nécessaires au fonctionnement du réseau.

Nous décrivons dans le cours « analyse de risque » l'outil HDIFF qui permet de valider les configurations de tout type d'équipement basé sur un modèle "ligne à ligne" à partir de patrons d'expressions régulières.

Les autres problèmes

La gestion des événements réseau : la prise en compte de millions d'événements réseau par jour couplée avec des besoins de corrélation pose des difficultés.

La mise à jour des systèmes : mettre à jour de manière cohérente et efficace des milliers d'équipements réseau n'est pas un problème trivial.

L'administration : la gestion de tels réseaux nécessite des compétences de plus en plus pointues afin de faire face à la complexité de l'architecture et des services.

La diversité des contrats : la multiplication de contrats tant au niveau constructeur, système de gestion, etc. peut être un sérieux frein aux déploiements des services.

Description of a multi-services network

MPLS Terminology

LDP: Label Distribution Protocol

LSP: Label Switched Path

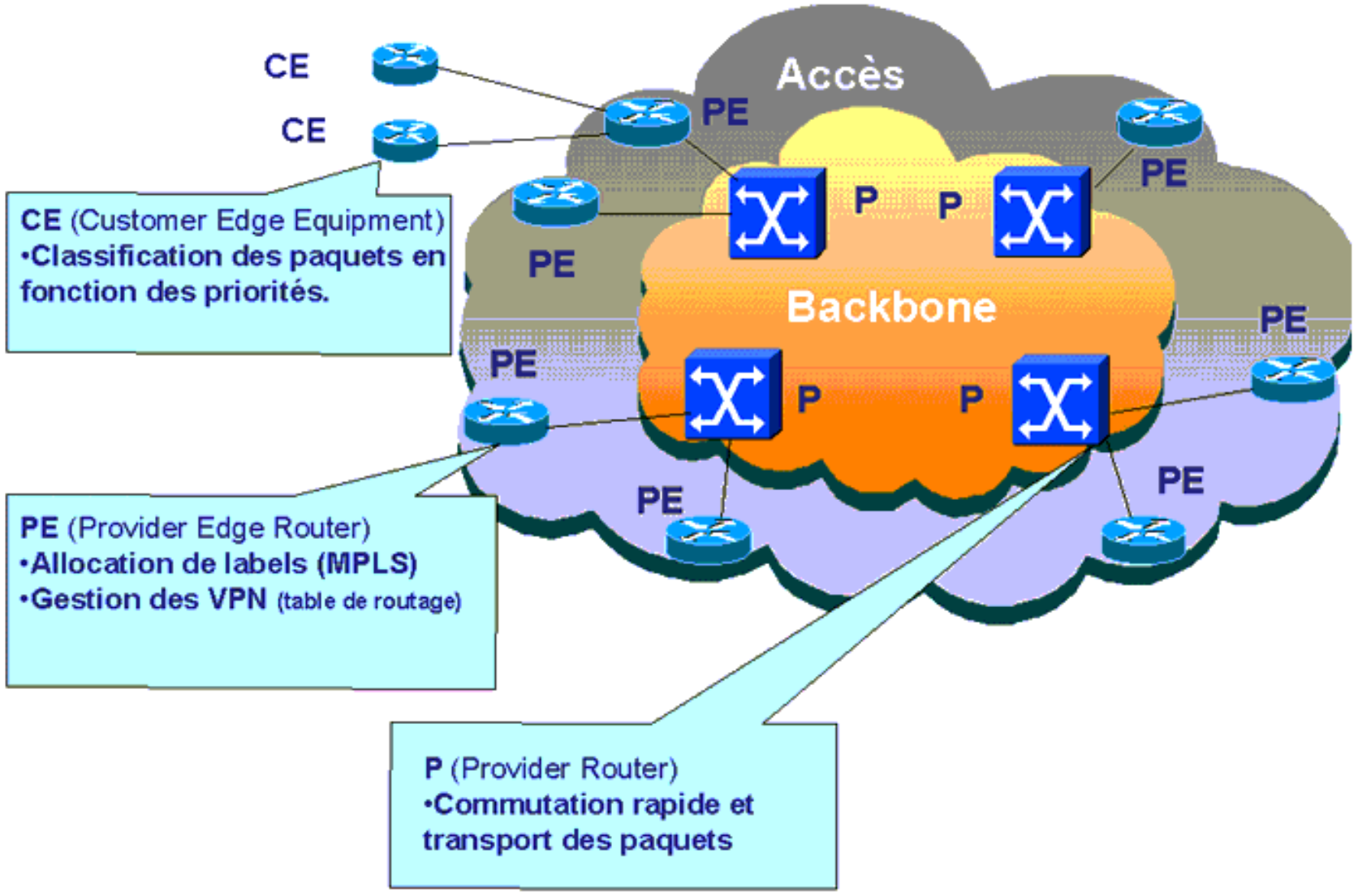
FEC: Forwarding Equivalence Class

LSR (or P router : Provider): Label Switching Router

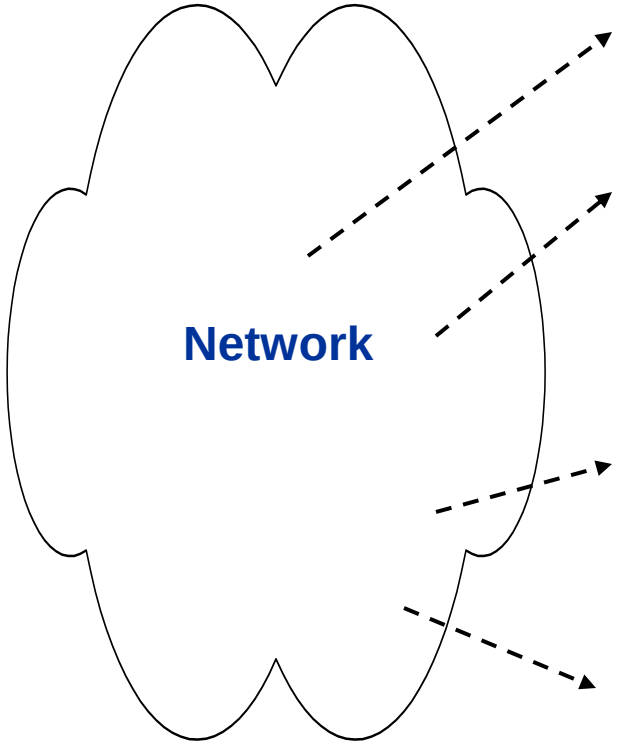
LER (or PE router : Provider Edge): Label Edge Router

(Useful term not in standards)

Description of a multi-services network (mpls backbone)



Description of a multi-services network (routing protocols)



IGP (interior gateway protocol)
-> IS-IS
- Fast convergence - Fast failure detection - …

LDP (Label Distribution Protocol).
La distribution implicite de labels aux LSR est réalisée grâce au protocole LDP (Label Distribution Protocol).

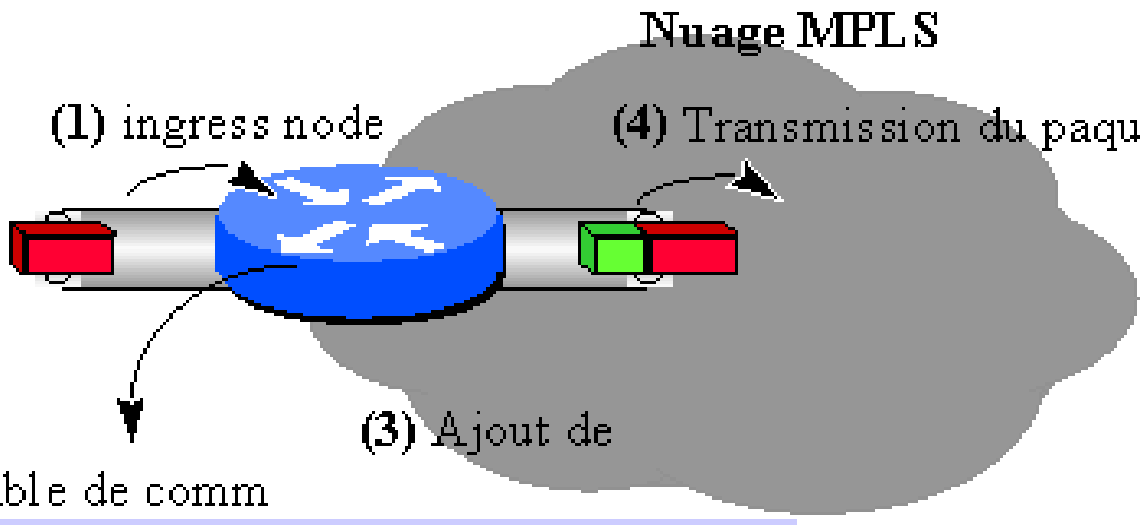
MP-BGP (BGP over MPLS to build IP VPN).
- MPi-BGP: pour les connections intra MPLS.
- MPe-BGP: pour les connections extra MPLS.

EGP (exterior gateway protocol)
-> BGPv4
- Flexible – Scalable - …

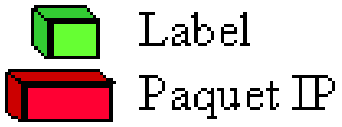
Agenda

MPLS

Description of a multi-services network (mpls backbone)

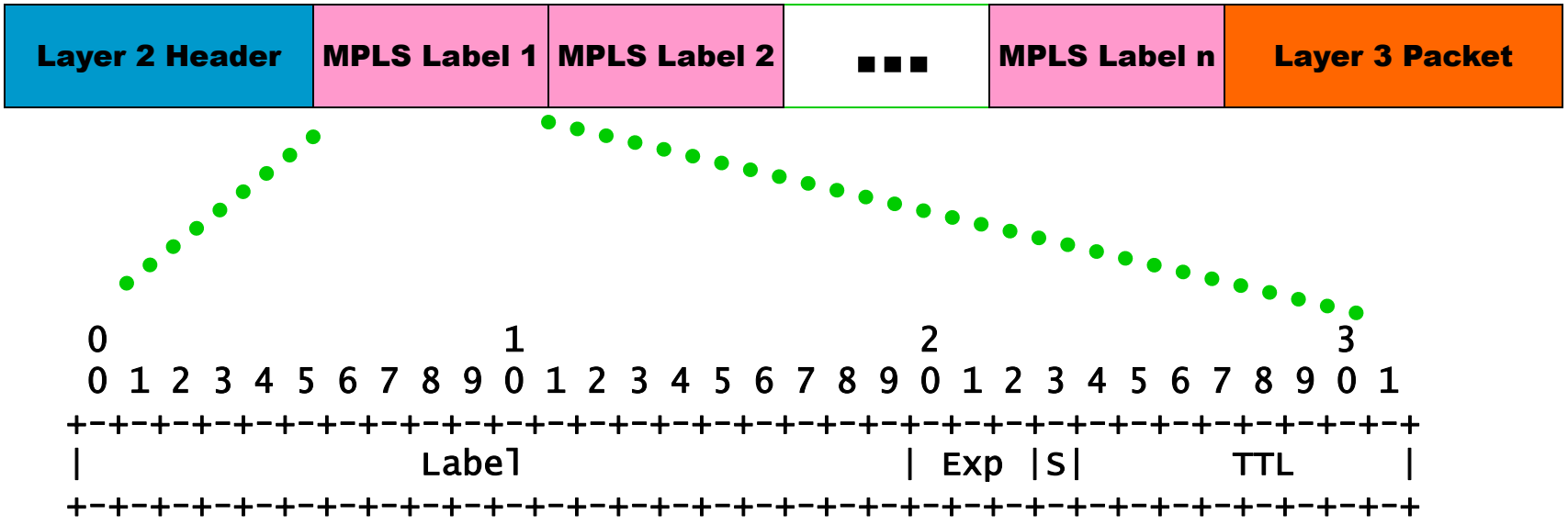


InPort	@Ipdst	Label	OuPor
P1	193.48.0.0/16	L5	P3
P2	193.48.0.0/16	L7	P4



1 - Le paquet IP arrive sur l'ingress node
2 - Le protocole de routage IP détermine, à partir de l'adresse IP de l'eggress node, la FEC, le label et le port de sortie.
3 - Ajout de l'en-tête
4 - Paquet IP + Label envoyé vers le noeud suivant

Generic MPLS Encapsulation



Often called a "shim" (or "sham") header

RFC 3032. MPLS Label Stack Encoding

- **Label: Label Value, 20 bits**
- **Exp: Experimental, 3 bits**
- **S: Bottom of Stack, 1 bit**
- **TTL: Time to Live, 8 bits**

Agenda

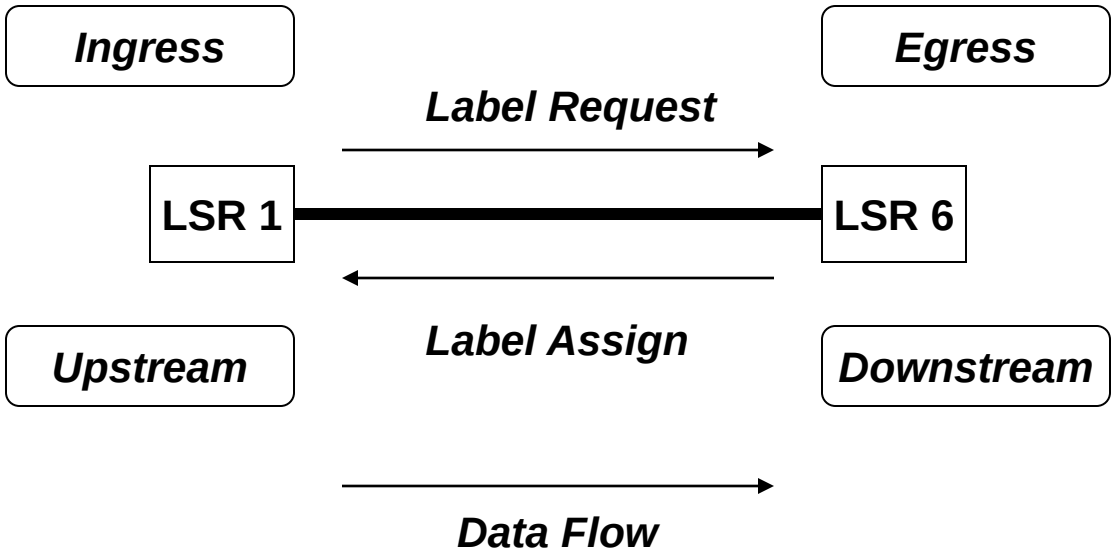
LDP

LDP Message Categories

- **Discovery messages: used to announce and maintain the presence of an LSR in a network.**
- **Session messages: used to establish, maintain, and terminate sessions between LDP peers.**
- **Advertisement messages: used to create, change, and delete label mappings for FECs.**
- **Notification messages: used to provide advisory information and to signal error information.**

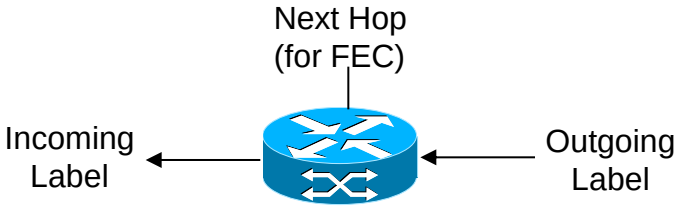
Downstream On Demand ordered

Ingress requests label from egress



Distribution Control: Ordered v. Independent

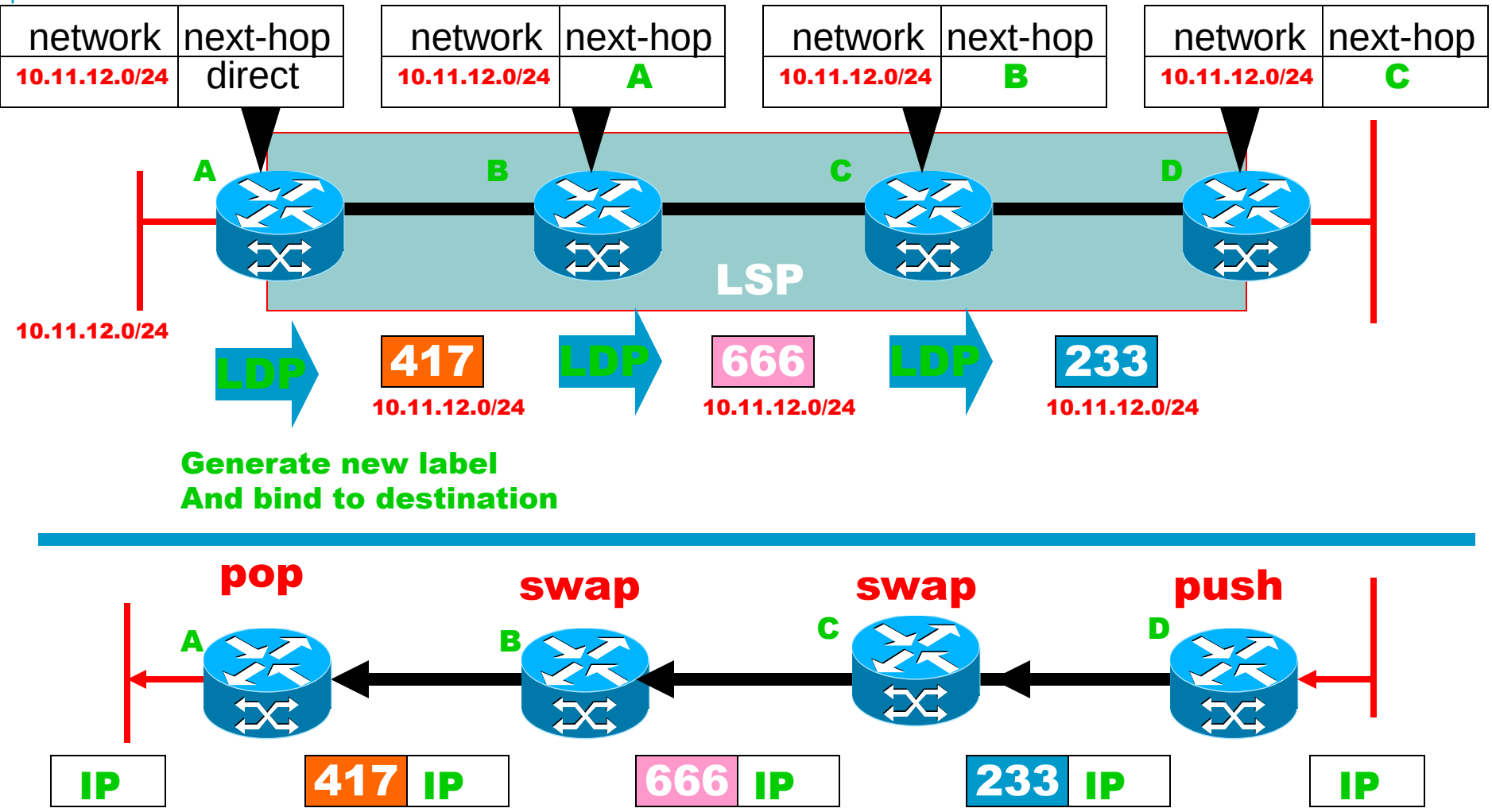
MPLS path forms as associations are made between FEC next-hops and incoming and outgoing labels



	Independent LSP Control	Ordered LSP Control
Definition	• Each LSR makes independent decision on when to generate labels and communicate them to upstream peers • Communicate label-FEC binding to peers once next-hop has been recognized • LSP is formed as incoming and outgoing labels are spliced together	• Label-FEC binding is communicated to peers if: - LSR is the 'egress' LSR to particular FEC - label binding has been received from upstream LSR • LSP formation 'flows' from egress to ingress
Comparison	• Labels can be exchanged with less delay • Does not depend on availability of egress node • Granularity may not be consistent across the nodes at the start • May require separate loop detection/mitigation method	• Requires more delay before packets can be forwarded along the LSP • Depends on availability of egress node • Mechanism for consistent granularity and freedom from loops • Used for explicit routing and multicast

Both methods are supported in the standard and can be fully interoperable

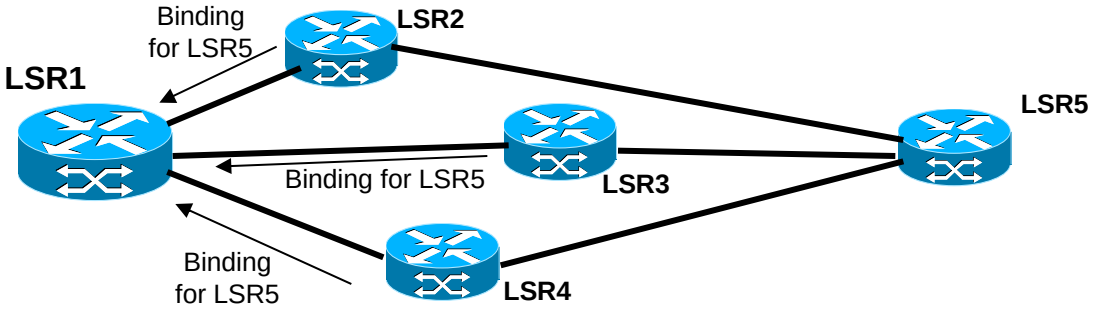
LDP and Hop-by-Hop routing



Label Retention Methods

An LSR may receive label bindings from multiple LSRs

Some bindings may come from LSRs that are not the valid next-hop for that FEC



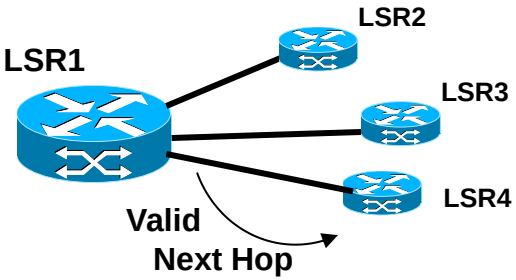
Liberal Label Retention

Label Bindings for LSR5

LSR4's Label

LSR3's Label

LSR2's Label



- LSR maintains bindings received from LSRs other than the valid next hop
- If the next-hop changes, it may begin using these bindings immediately
- May allow more rapid adaptation to routing changes
- Requires an LSR to maintain many more labels

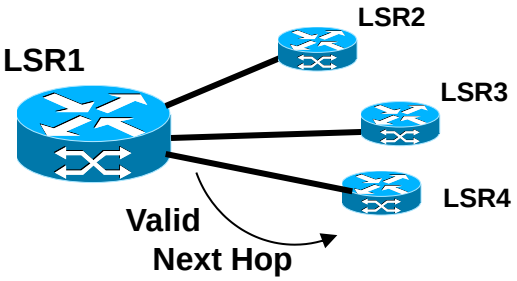
Conservative Label Retention

Label Bindings for LSR5

LSR4's Label

~~*LSR3's Label*~~

~~*LSR2's Label*~~



- LSR only maintains bindings received from valid next hop
- If the next-hop changes, binding must be requested from new next hop
- Restricts adaptation to changes in routing
- Fewer labels must be maintained by LSR

Agenda

MPLS-TE

MS-PW TE

- MPLS-TE means MPLS Traffic Engineering intends to provide a set of value added features such as :
  - Network resources optimization by offering a constraint-based routing. The use of RSVP-TE to setup constraints TE over the MPLS cloud.
  - Fast rerouting of tunnels among the MPLS cloud allowing to swap from a LSP to an another LSP.
  - QOS over the MPLS cloud allowing to deliver different set of services needed specific network constraints (video, voice, etc.). Use of Diffserv (differentiated serivces) on all MPLS devices and use of the EXP field of MPLS frame to set information required for Diffserv.

Constraint Based Routing

Basic components

1. **Specify path constraints**
2. **Extend topology database to include resource and constraint information** — Extend Link State Protocols (ISIS-TE, OSPF-TE)
3. **Find paths that do not violate constraints and optimize some metric**
4. **Signal to reserve resources along path**
5. **Set up LSP along path (with explicit route)** — Extend RSVP or LDP Protocols (RSVP-TE, CR-LDP)
6. **Map ingress traffic to the appropriate LSPs**

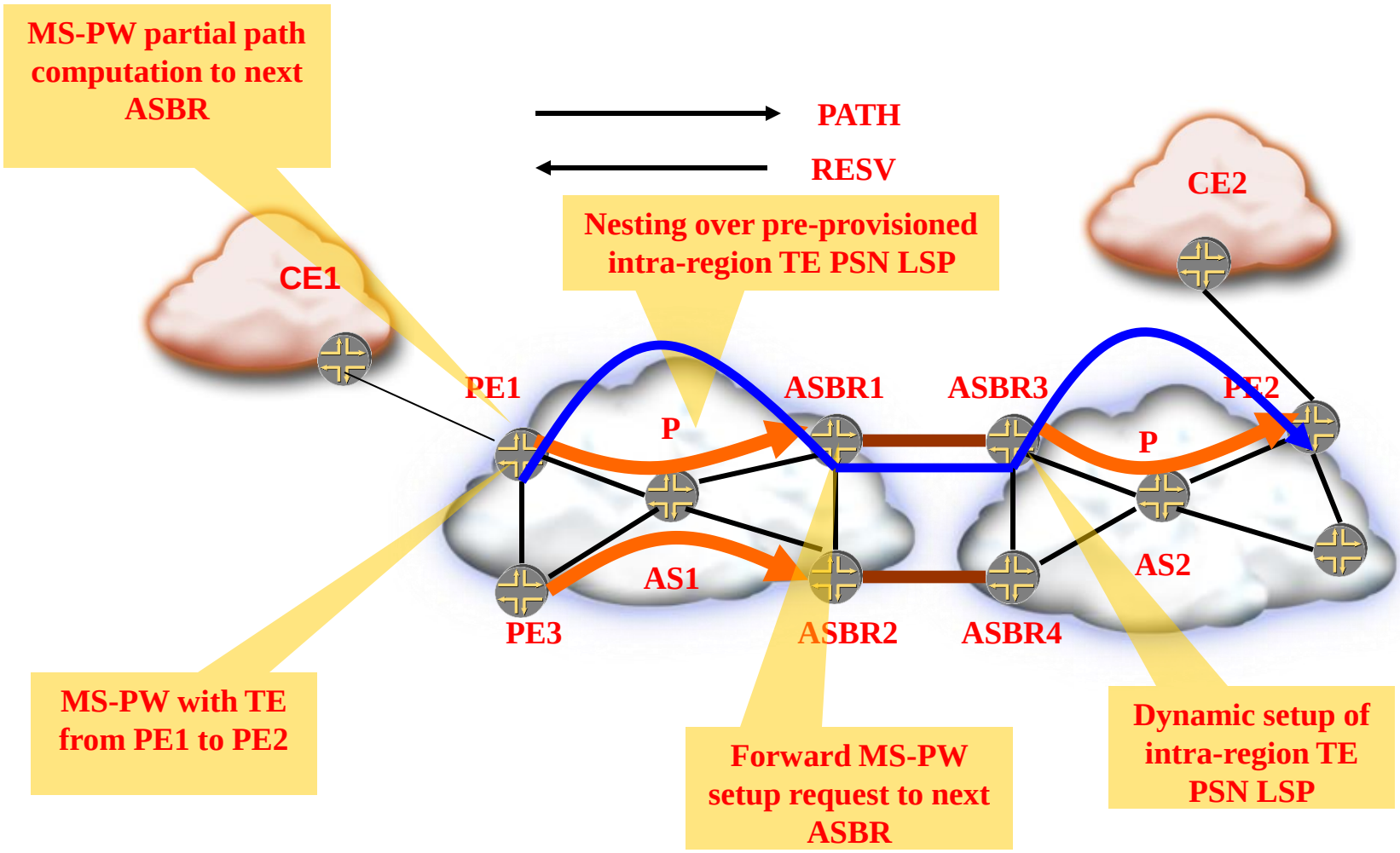
Agenda

MPLS-TE && RSVP-TE

MS-PW TE using RSVP-TE

- RSVP-TE protocol is an extension of the protocol RSVP.
- RSVP-TE permits to distribute labels in order to establish LSP with specific constraints. RSVP-TE is using the distribution mode named : "Downstream on Demand".
- When a LER intends to establish a LSP to the targeted LER, it sends a "PATH" message towards the targeted LER in order to establish the LSP. The message is propagated inside the MPLS domain.
- When the targeted LER receives the "PATH« message, then it will answer by a "RESV" message towards the LER. At this step and when the Ingress LER received the message, then the LSP is established and the resources needed to established the LSP are blocked.
- Le message "RESV" message contains the following information : FEC + LSP ID, etc.

MS-PW TE using RSVP-TE



Agenda

DiffServ && DiffServ-TE

DiffServ

- **DiffServ Architecture – RFC 2475**
- **Scales well with large flows through aggregation**
- **Creates a means for traffic conditioning (TC)**
- **Defines per-hop behavior (PHB)**
- **Edge nodes perform TC**

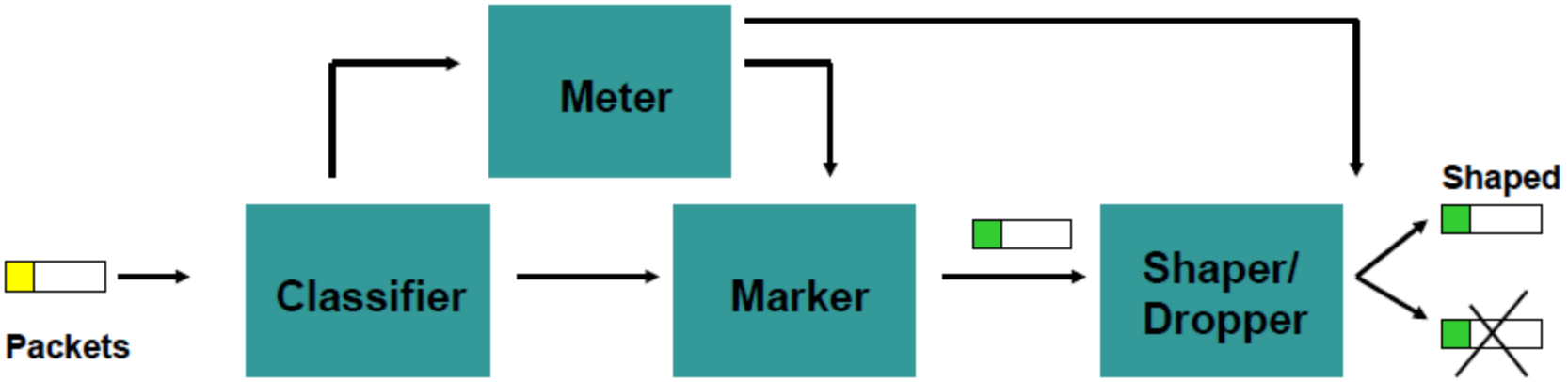
– **Allows core routers to do more important processing tasks**

- **Tough to predict end-to-end behavior**

– **Especially with multiple DiffServ Domains**

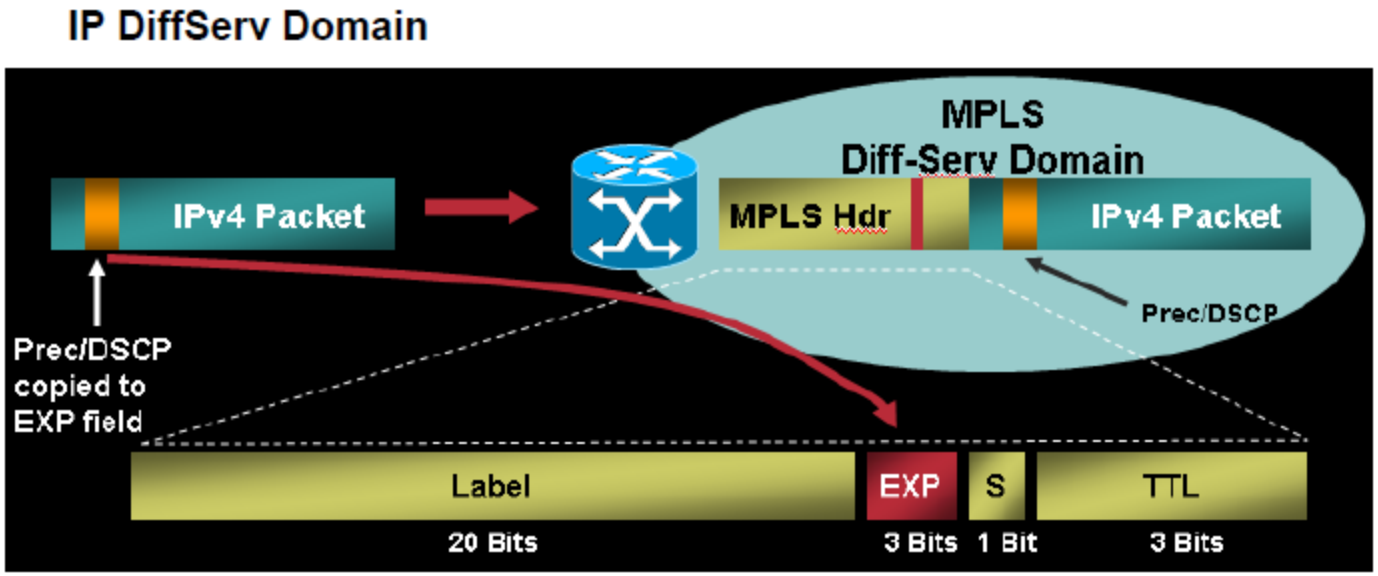
– **DiffServ implementation versus Capacity planning**

DiffServ



- Classifier: Selects a packet in a traffic stream based on the content of some portion of the packet header
- Meter: Checks compliance to traffic parameters (eg Token Bucket) and passes result to the marker and shaper/dropper to trigger a particular action for in/out of profile packets
- Marker: Writes/rewrites DSCP
- Shaper: Delays some packets to be compliant with a profile
- Dropper: Discards some or all of the packets in a traffic stream in order to bring the stream into compliance with a traffic profile

DiffServ && MPLS



- Prec/DSCP field is not directly visible to MPLS Label Switch Routers (they forward based on MPLS Header and EXP field)
- Information on DiffServ must be made visible to LSR in MPLS Header using EXP field / Label.
- How do we map DSCP into EXP ? Interaction between them.

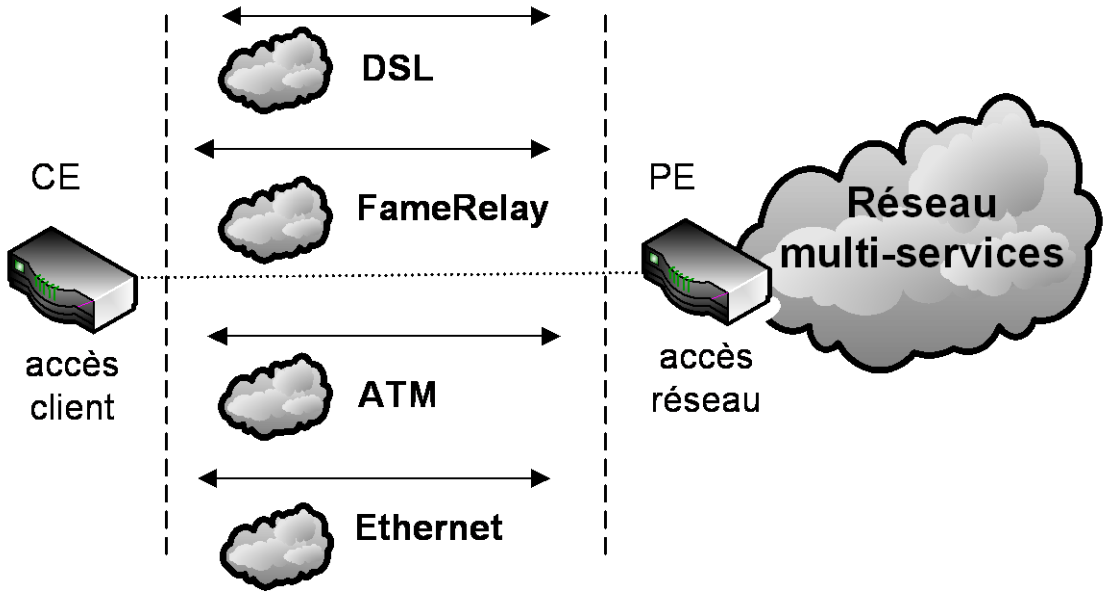
DiffServ-TE

- Regular TE allows for one reservable bandwidth amount per link
- DS-TE allows for more than one reservable bandwidth amount per link
- Brings per-class dimension to TE
- Basic idea: connect PHB class bandwidth to DS-TE bandwidth sub-pool

Examples of services

Access

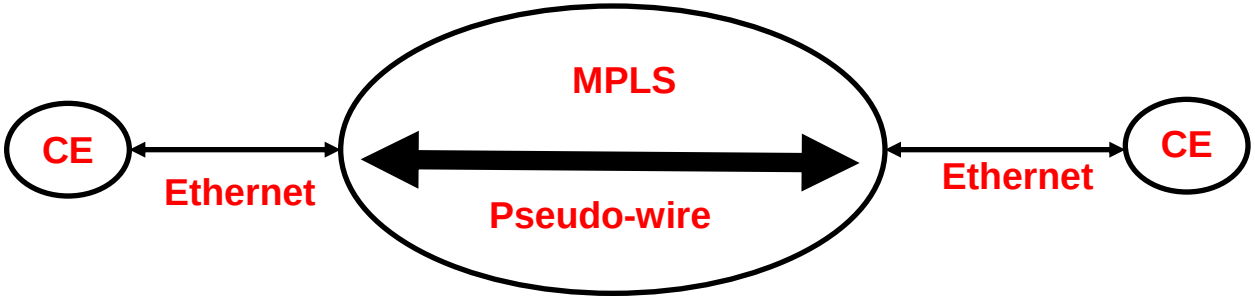
Description of a multi-services network (protocol access)



Quel que soit le service offert par le réseau multi-services, un équipement d'accès du réseau multi-services est capable d'agréger différents types de protocoles d'accès

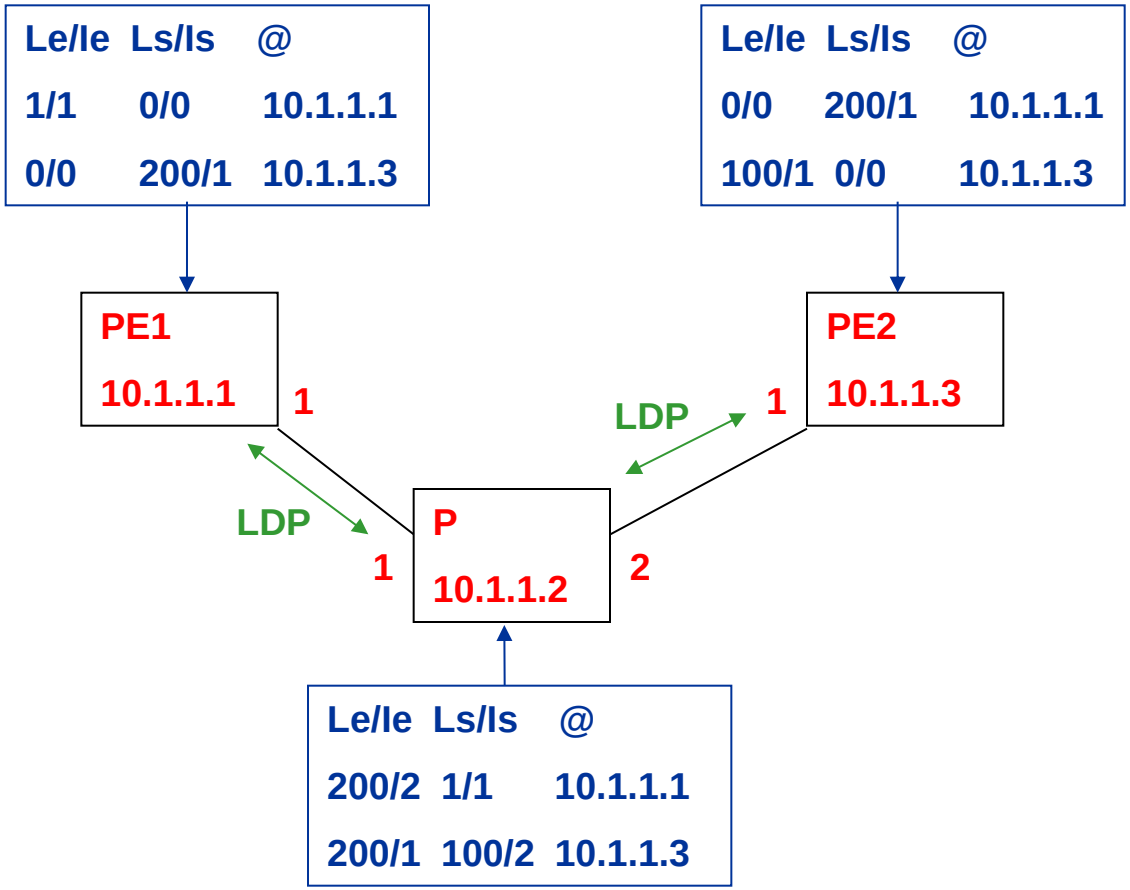
Ethernet connexion

Description of a multi-services network (L2VPN : Edge to Edge tunneling)

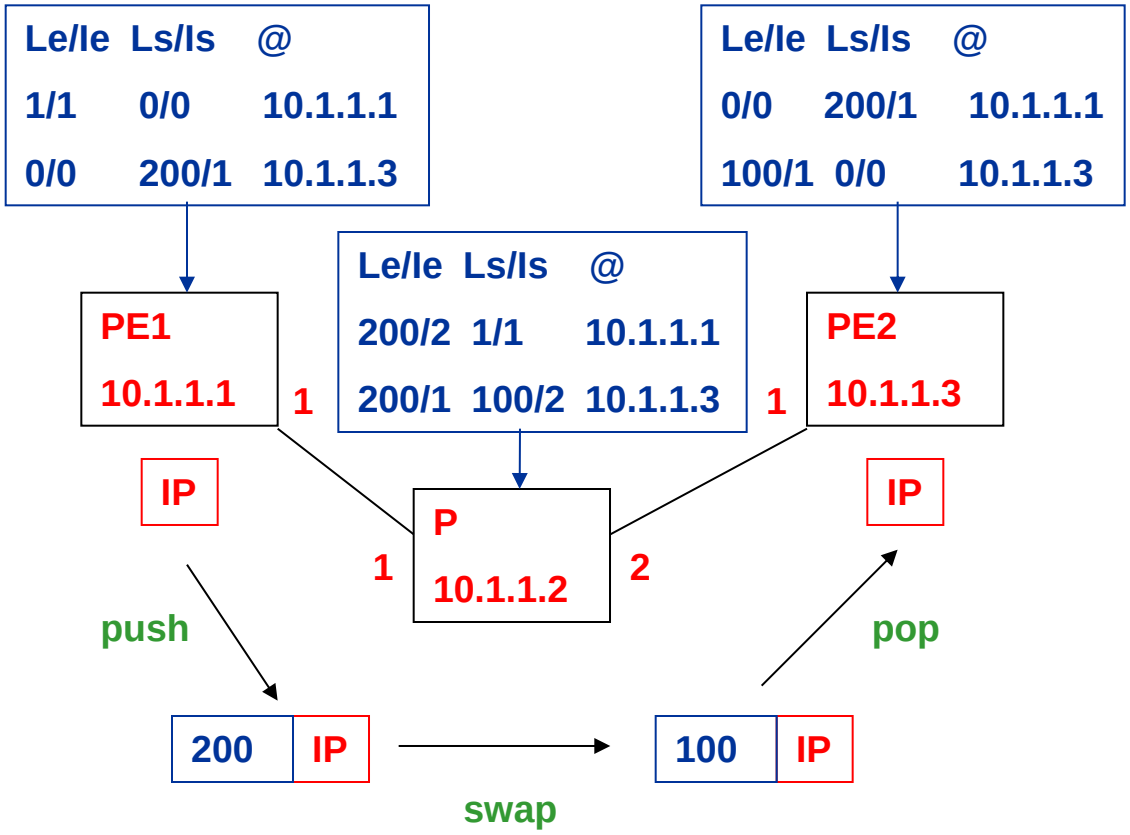


MPLS peut émuler une connexion Ethernet en faisant croire que le réseau est un câble dédié à chacun de ses clients.

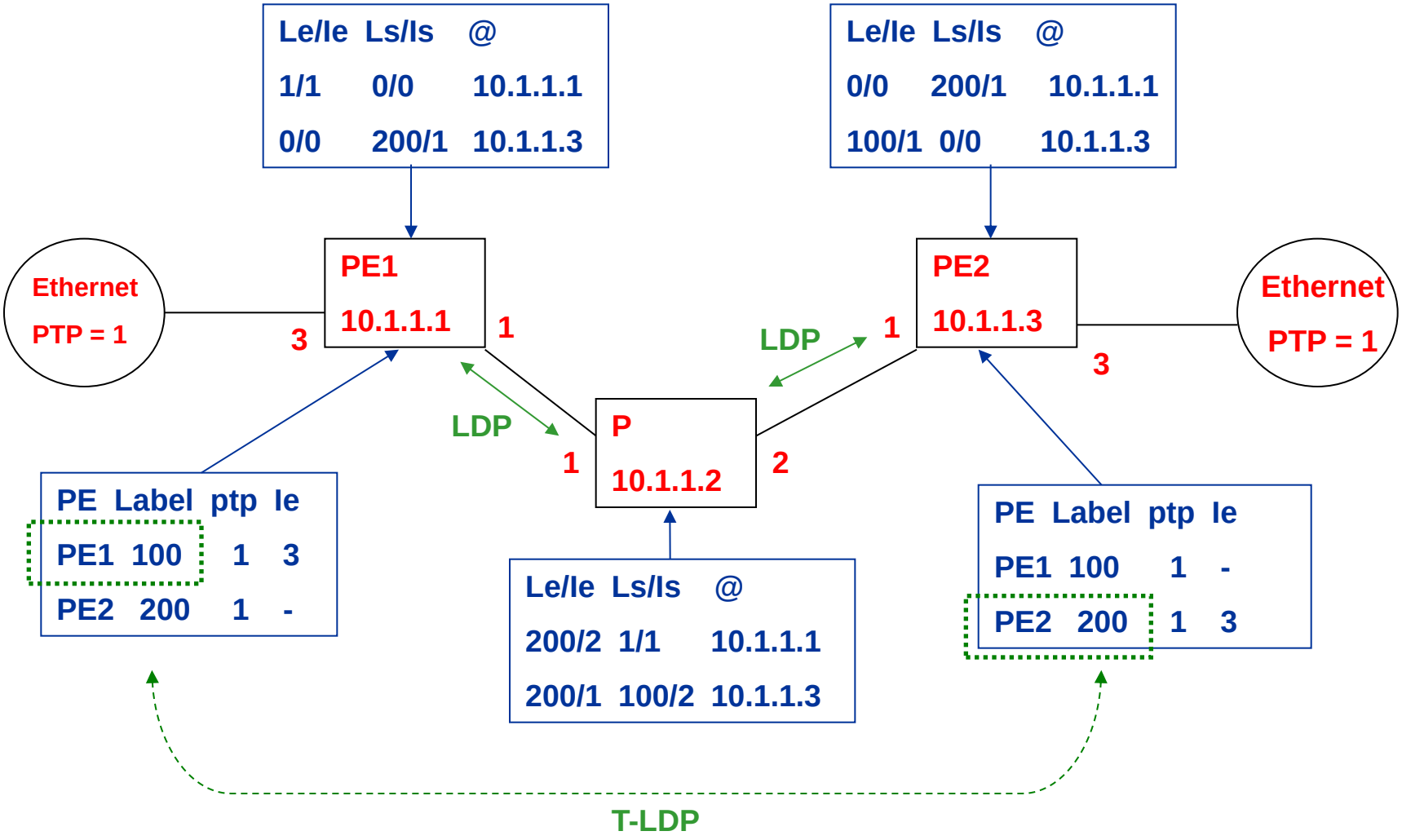
Example : label distribution



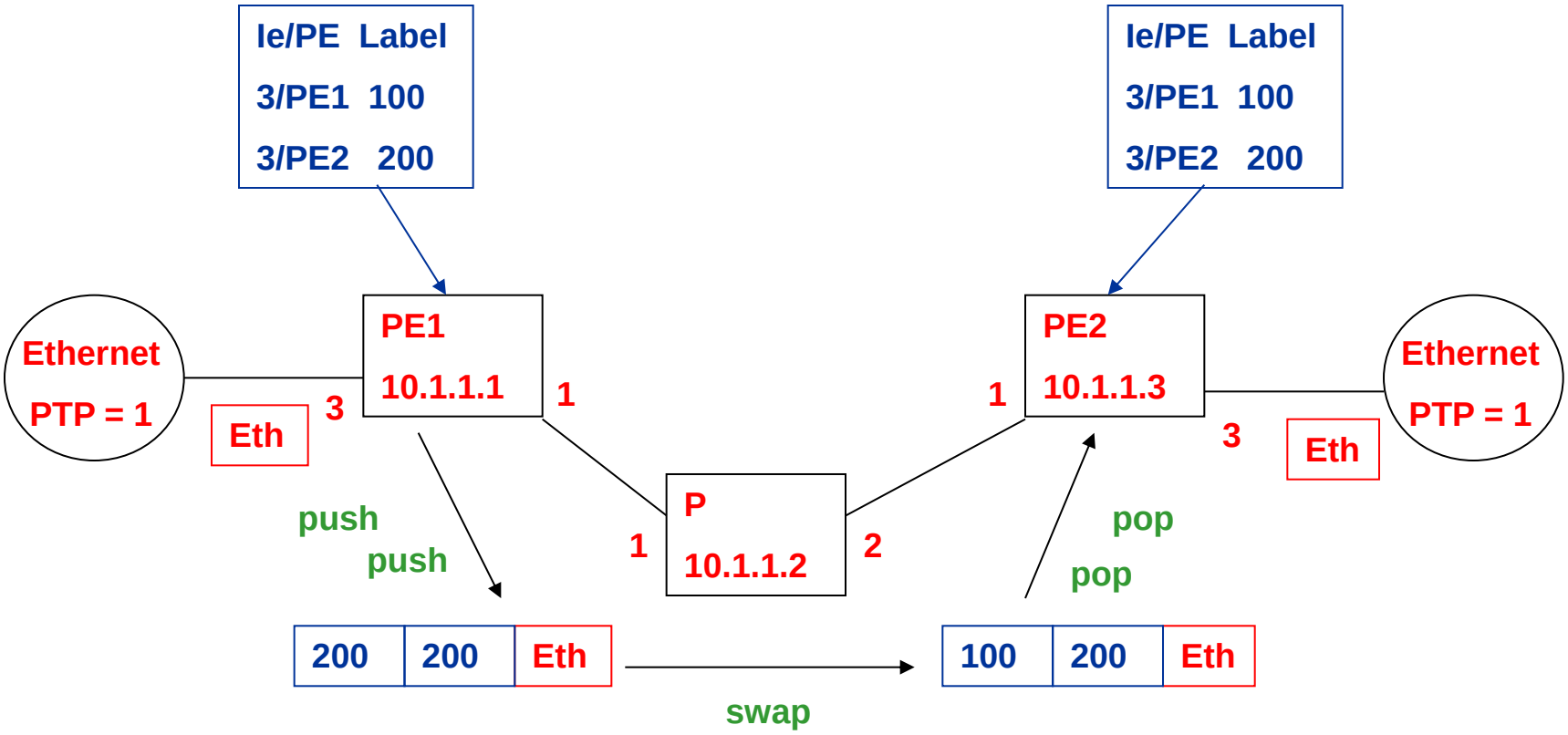
Exemple : packet switching between PEs



Example : label distribution

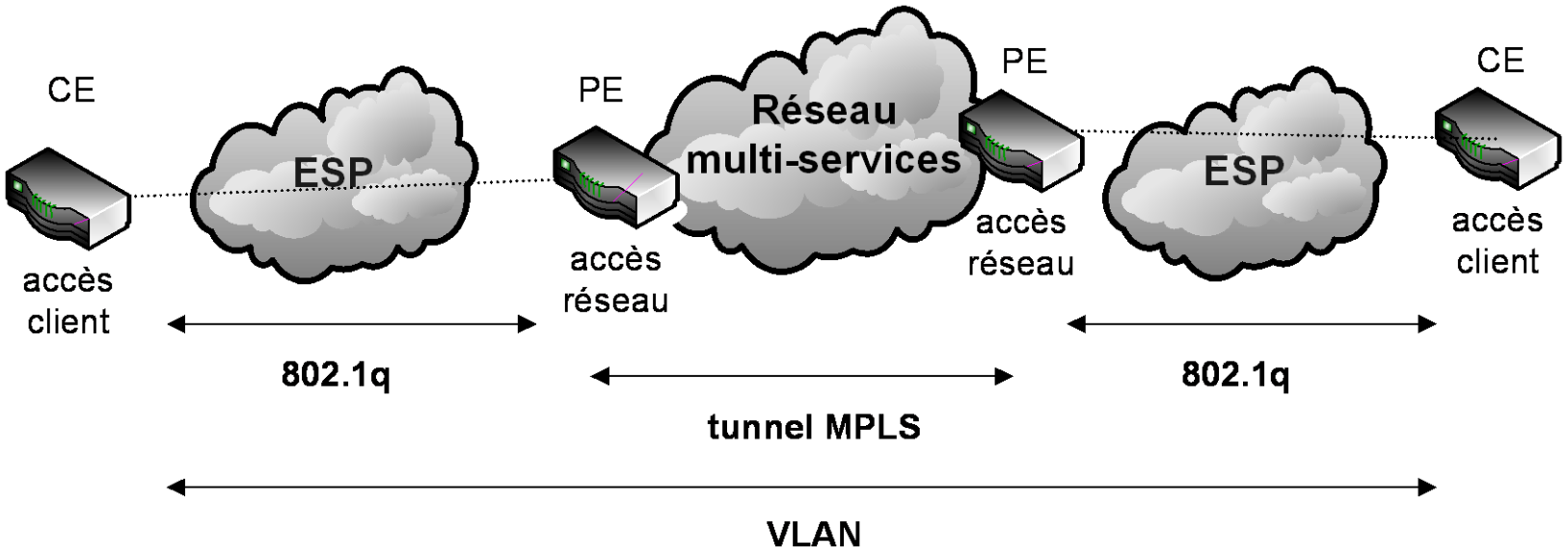


Exemple : packet switching between VPN sites



VPN layer 2

Description of a multi-services network (L2VPN : Edge to Edge tunneling)



Les réseaux Ethernet ont été pendant longtemps limités à des réseaux locaux. Cependant et avec le concept de VLAN (Virtual LAN) [IEEE 802.1], il est possible de les étendre sur une portion MAN (Metropolitan Area Network) ou WAN (Wide Area Network). En effet, la technique de tagging des trames Ethernet (norme 802.1q) couplée avec la création de tunnels MPLS permet d'étendre un LAN partout dans le monde

Description of a multi-services network (L2VPN : Edge to Edge tunneling)

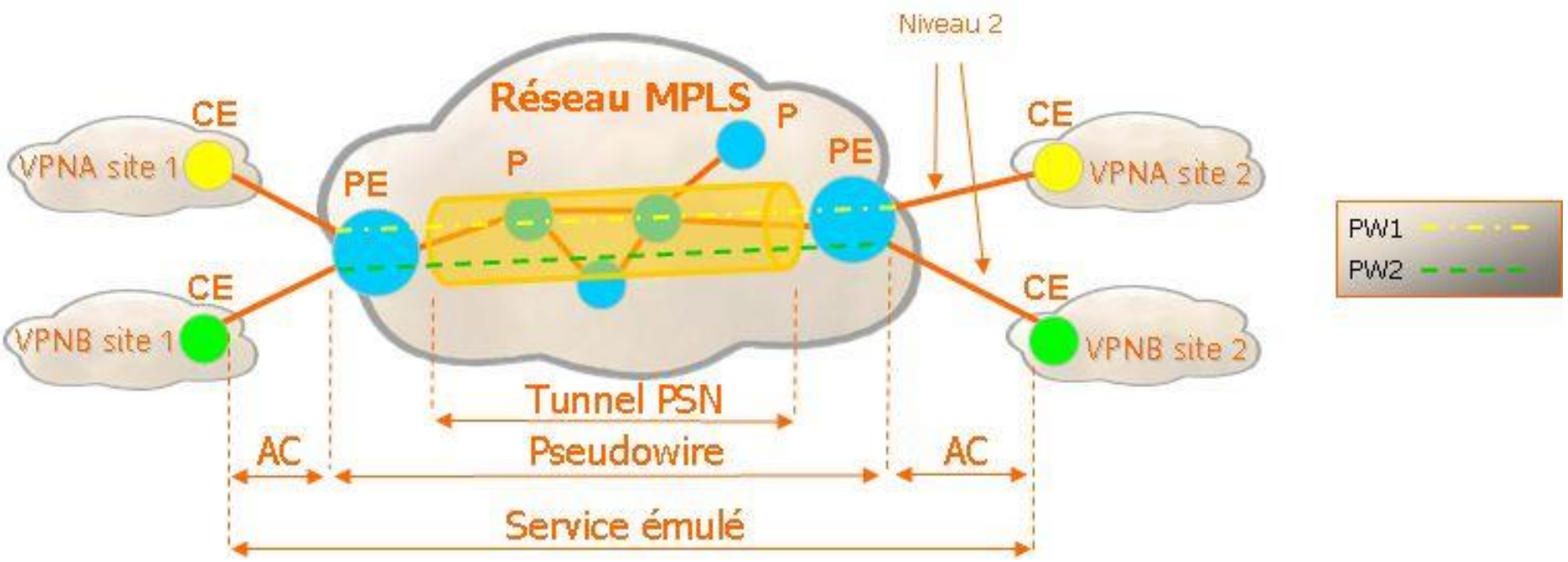
Packet Switched Network (PSN)

- network that forwards packets
- IPv4, IPv6, MPLS, Ethernet (although IETF does not touch)

Pseudowire (PW): A mechanism that emulates the essential attributes of a native service while transporting over a packet switched network (PSN)

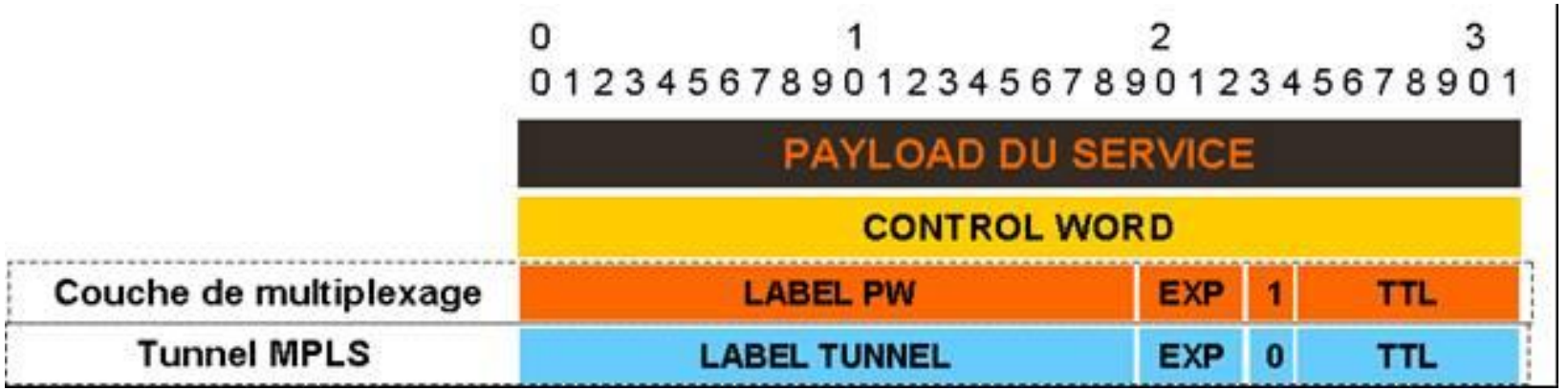
PWs are bidirectional (unlike MPLS LSPs)

Description of a multi-services network (L2VPN : Edge to Edge tunneling)



Attachment Circuit : AC : Ethernet, ATM, FR, etc.

Description of a multi-services network (L2VPN : Edge to Edge tunneling)



En-tête tunnel MPLS : ce niveau est aussi appelé l'en-tête du PSN. On y trouve le label du tunnel MPLS permettant de relier les deux PE

En-tête PW : ce niveau est également un en-tête MPLS et contient le label du pseudowire.

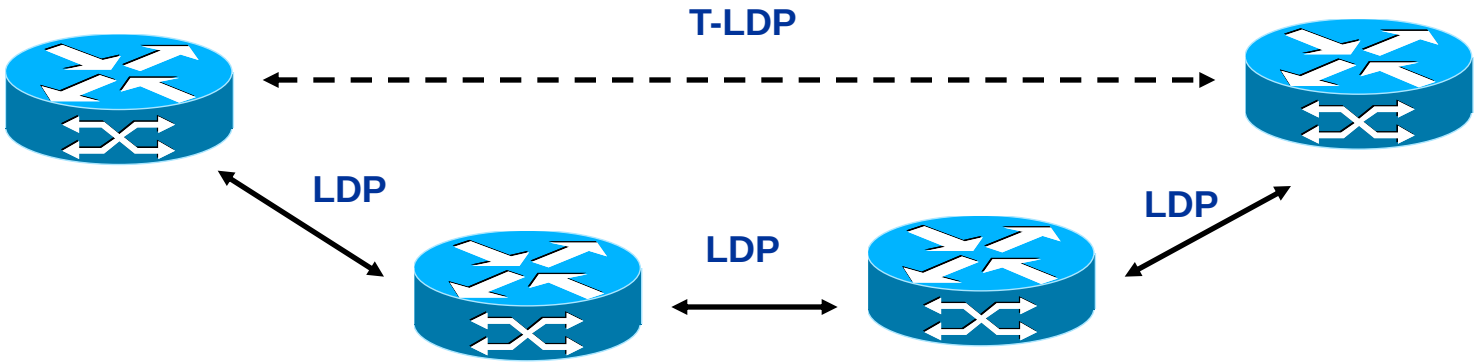
Control Word : suivant le type de PW, c'est-à-dire le trafic à transporter, ce niveau est optionnel. Le «Control Word» est une information spécifique ajoutée avant la payload [RFC4385] (renseigne des notions de fragmentation, séquencement, etc.).

Payload du service (ou charge utile): ce niveau présente le flux des données (TDM, ATM, FR, ATM, Ethernet…) à transporter dans un paquet PW.

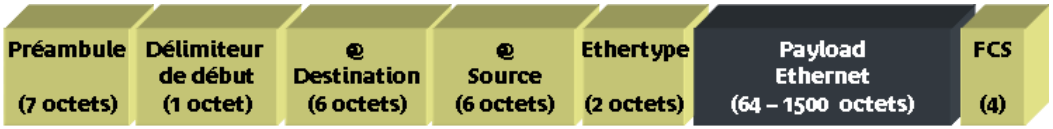
Description of a multi-services network (L2VPN : Edge to Edge tunneling)

Le PE utilise le plan de contrôle MPLS (pour établir les tunnels MPLS vers le PE distant et LDP pour signaler les PW. A la différence du LSP où une session LDP est utilisée entre chaque nœud IP/MPLS, la session LDP pour le PW est généralement établie entre deux PE non voisins.

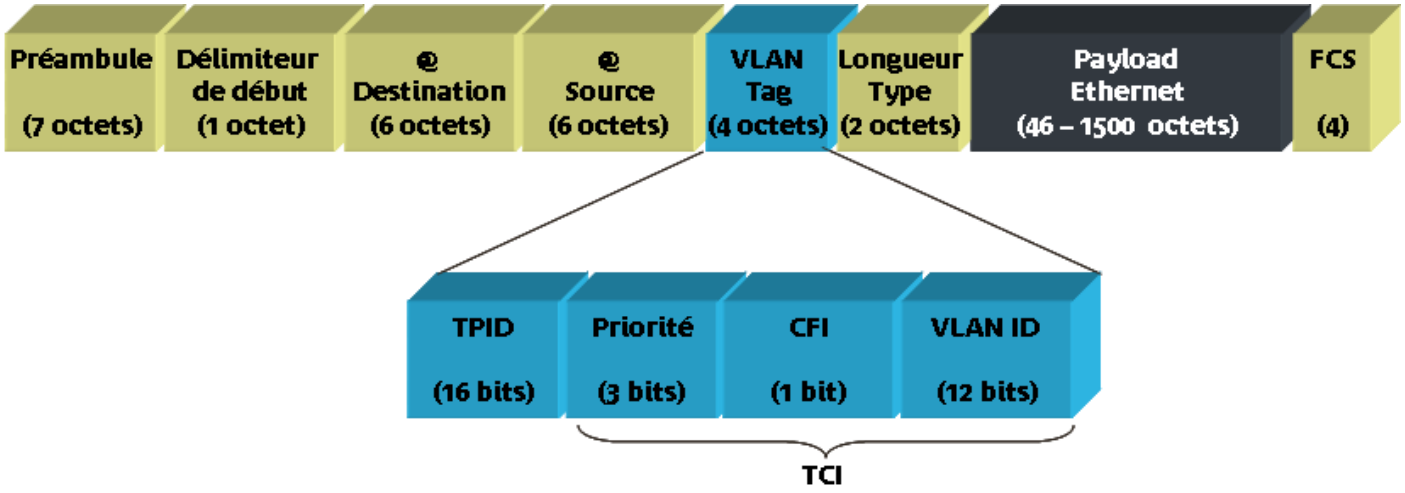
On parlera donc de session Targeted LDP (T-LDP). Sur chacun des PE on devra alors configurer l'adresse du PE distant (destination ou "targeted") avec lequel on désire établir une session LDP.



Description of a multi-services network (L2VPN : Edge to Edge tunneling)

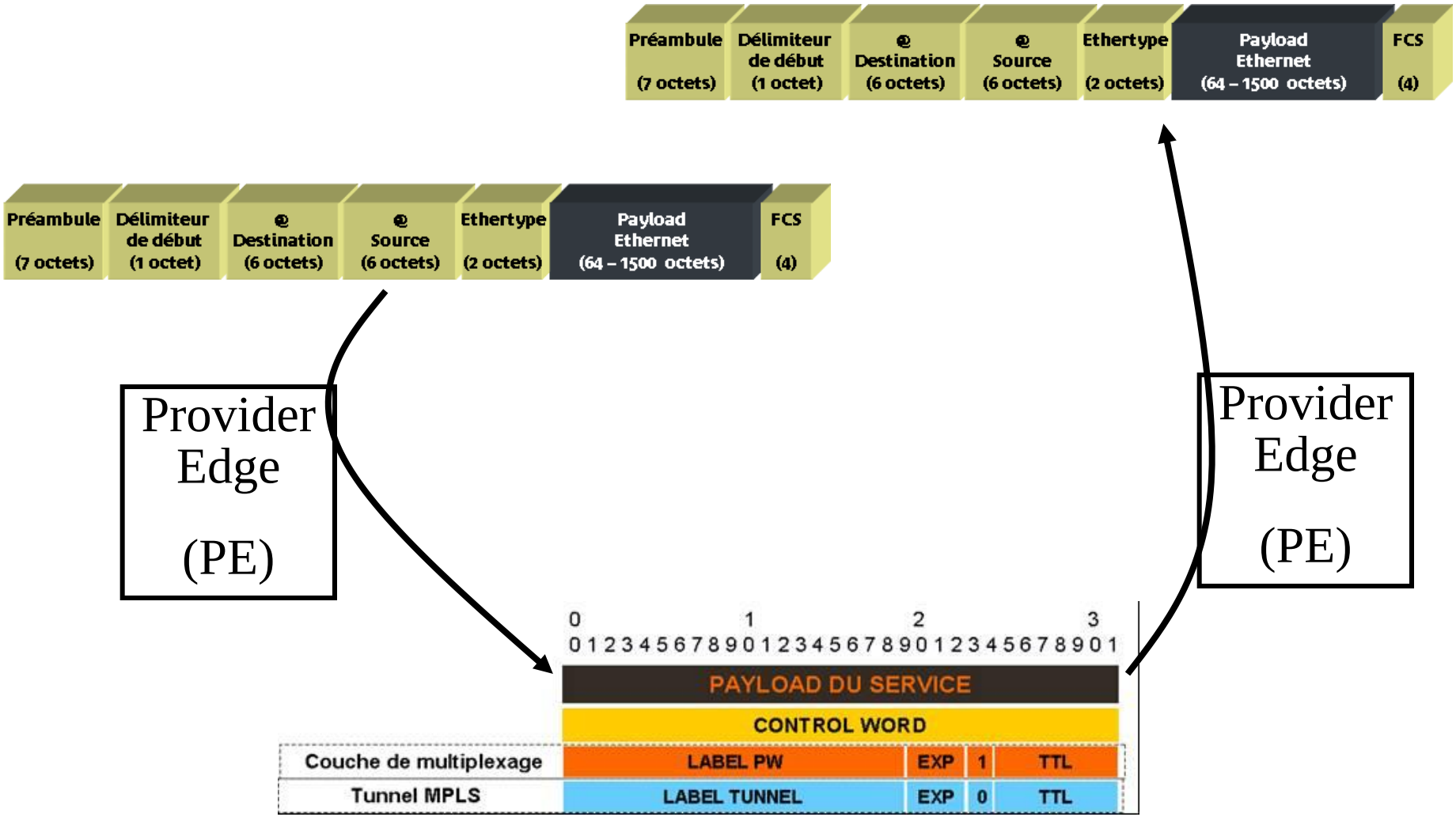


Standard Ethernet packet

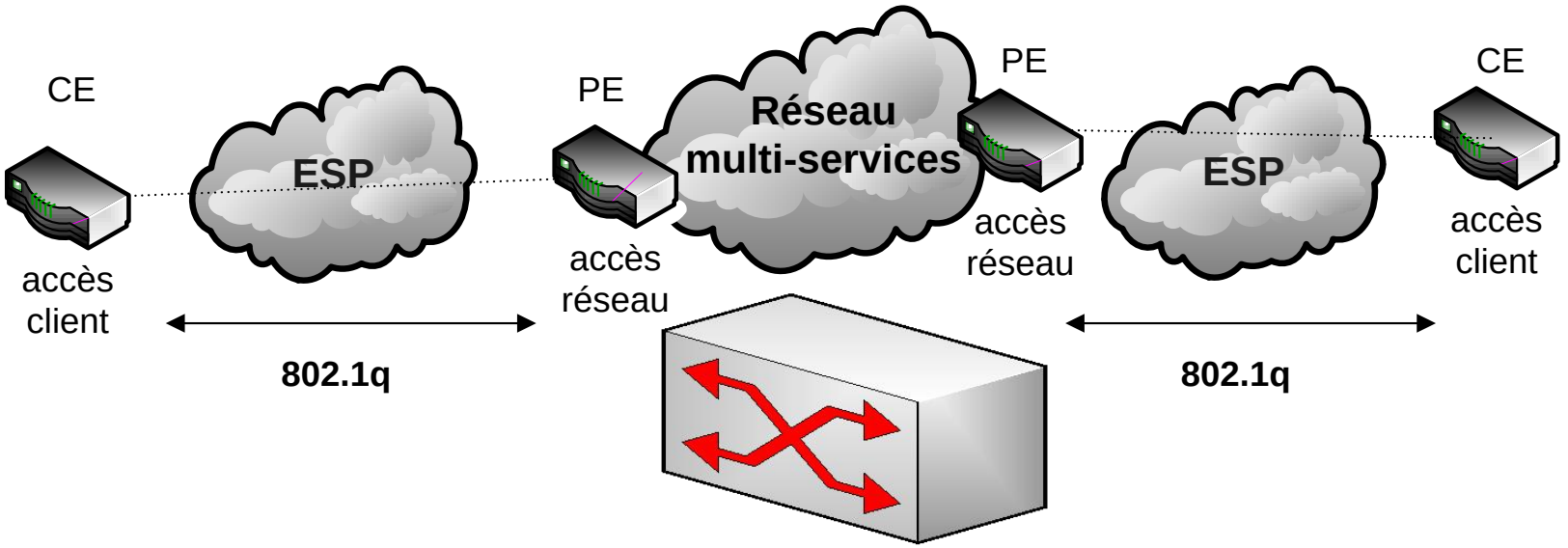


Ethernet packet modified (802.1q) allowing to build Virtual VLANs

Description of a multi-services network (L2VPN : Edge to Edge tunneling)



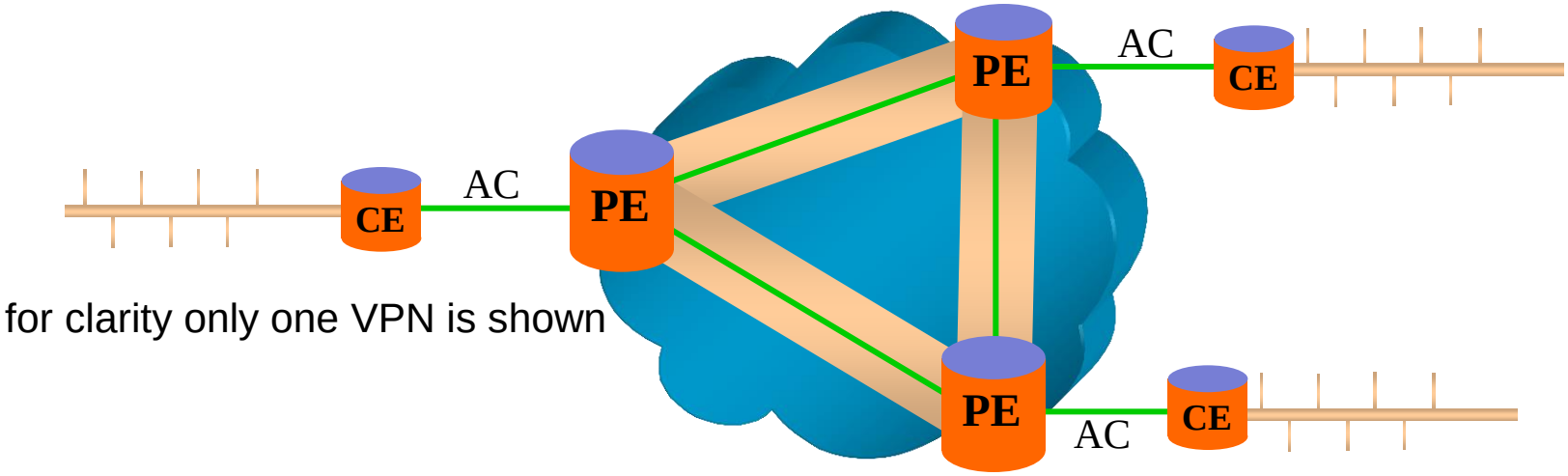
Description of a multi-services network (L2VPN : Edge to Edge tunneling)



Le réseau MPLS est vu du client comme un unique commutateur reliant tous ces sites.

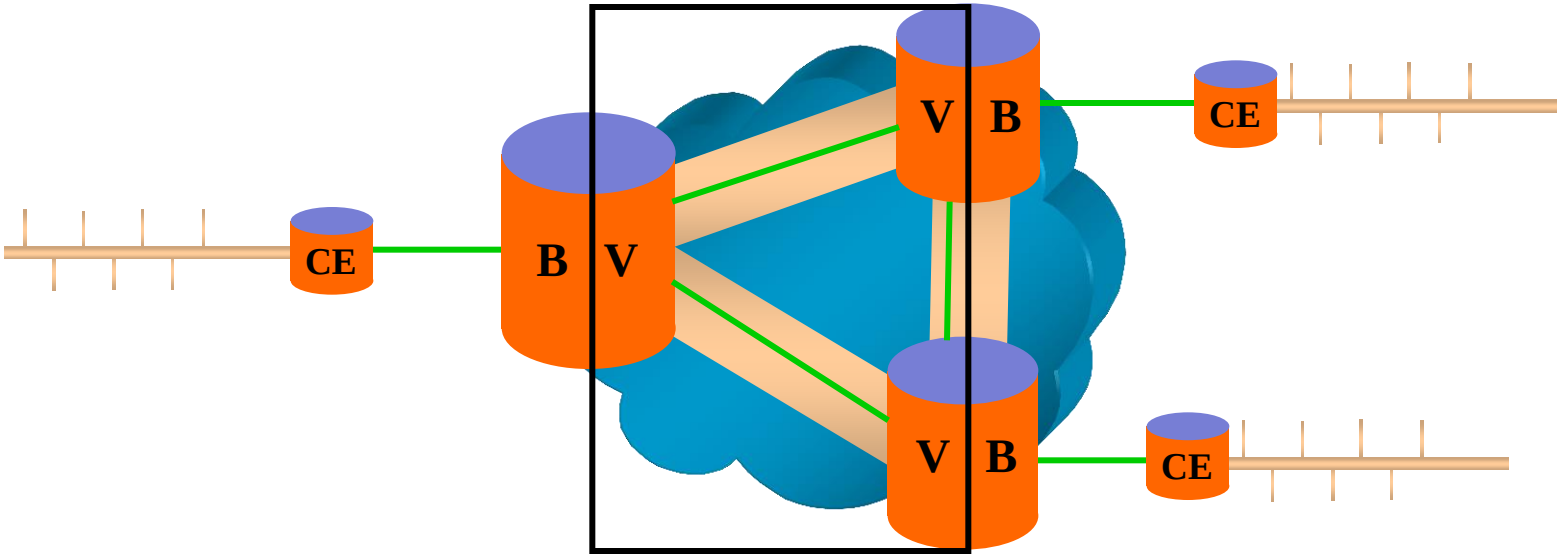
Le commutateur apprend et distribue les adresses Ethernet/MAC comme un commutateur classique.

Description of a multi-services network
(L2VPN : VPLS)



- VPLS emulates a LAN over an MPLS network
- set up MPLS tunnel between every pair of PEs (full mesh), set up Ethernet -
- PW inside tunnels, for each VPN instance
- CEs appear to be connected by a single LAN
- PE must know where to send Ethernet frames

Description of a multi-services network
(L2VPN : VPLS)



a VPLS-enabled PE has, in addition to its MPLS functions:

- VPLS code module (IETF drafts)
- Bridging module (standard IEEE 802.1D learning bridge)

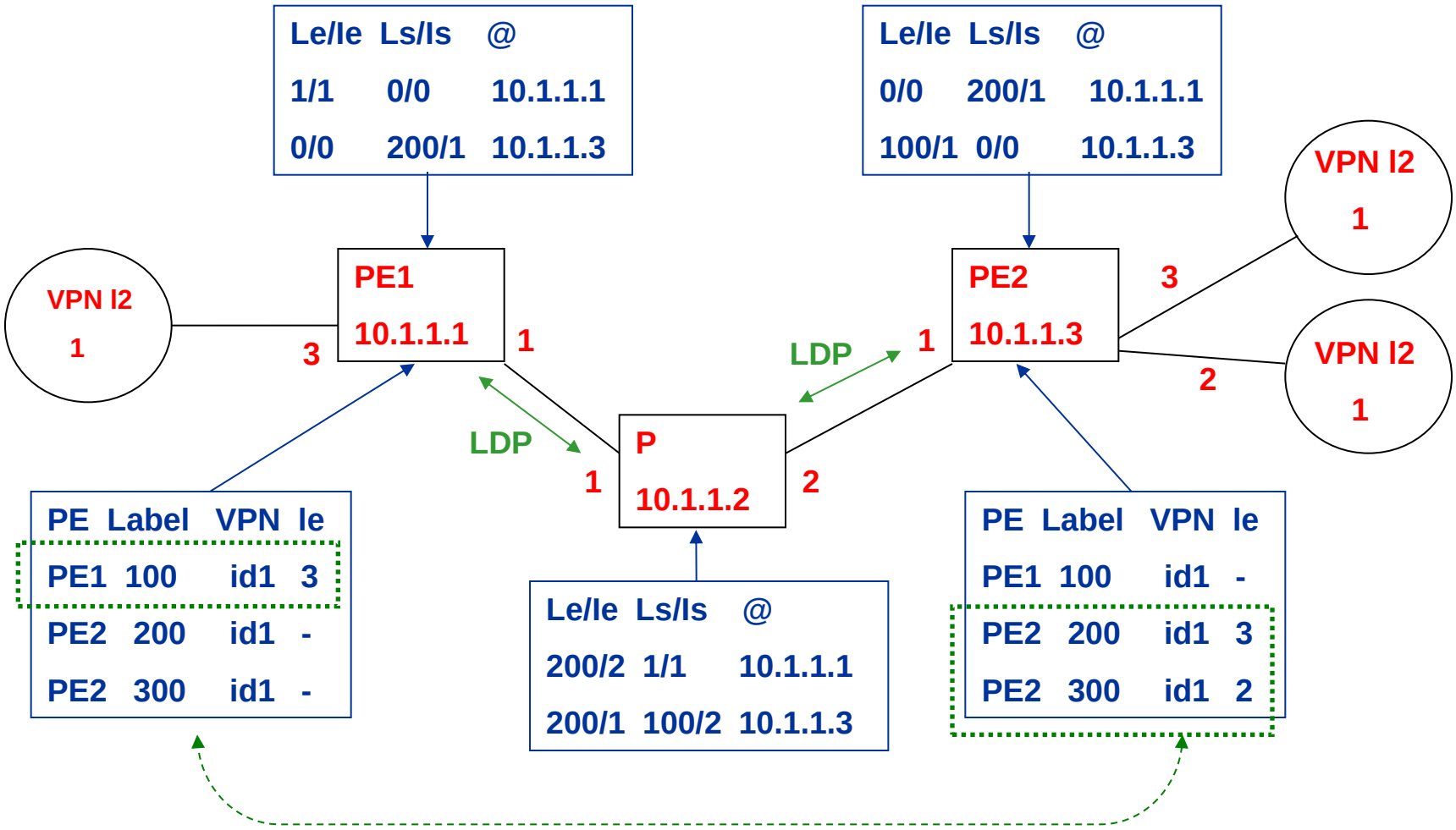
SP network (inside rectangle) looks like a single Ethernet bridge!

Description of a multi-services network (L2VPN : Edge to Edge tunneling)

Configuration d'un routeur PE, connecté à CE de niveau 2

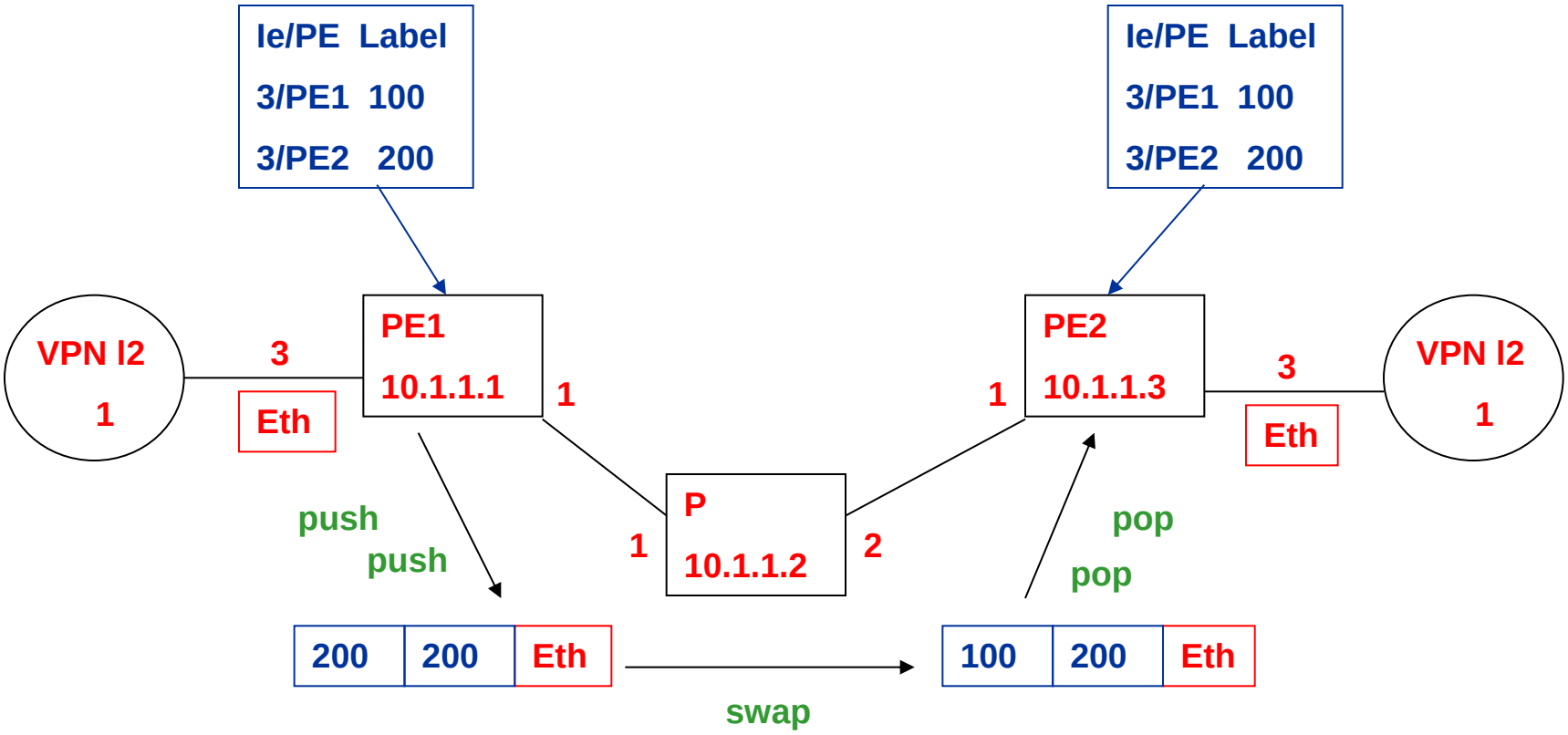
```
l2 vfiPE1-VPLS-A manual
    vpn id 100
        neighbor 2.2.2.2 encapsulation mpls → tunnel towards an other PE
        neighbor 3.3.3.3 encapsulation mpls → tunnel towards an other PE
    !
    Interface loopback 0
        ipaddress 1.1.1.1 255.255.255.255 → address used for tunnel setup
!
Interface fastethernet0/0          /* dedicated customer interface */
    switchport
        switchport mode dot1q tunnel → authorised 802.1Q
        switchport access vlan 100   → vlan number authorised
!
Interface vlan100
    no ip address
    xconnect vfiPE1-VPLS-A
!
vlan100
        state actives
    !
```

Example : label distribution



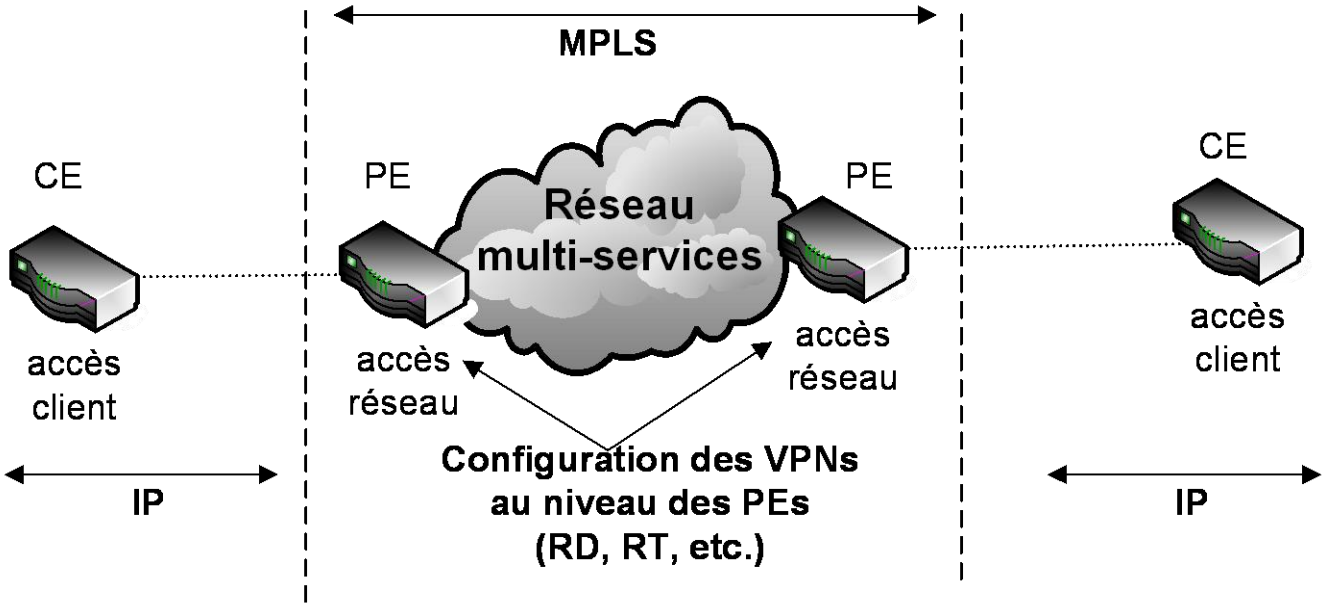
L'utilisation de T-LDP (ou MP-BGP) ne permet pas d'échanger d'adresse MAC (qui sont apprises via le mécanisme d'apprentissage MAC par la duplication et l'inondation).

Exemple : packet switching between VPN sites



VPN layer 3

Description of a multi-services network
(L3VPN : mpls/vpn bgp service)



Le service VPN BGP/MPLS permet de créer des VPNs sur un réseau mutualisé tout en permettant à chaque VPN d'avoir son propre plan d'adressage.

MPLS VPN Connection Model

(L3VPN : mpls/vpn bgp service) **mpls/vpn bgp service)**
MP-BGP Update / dedicated VRF (Virtual Routing and Forwarding)

VPN-IPV4 address

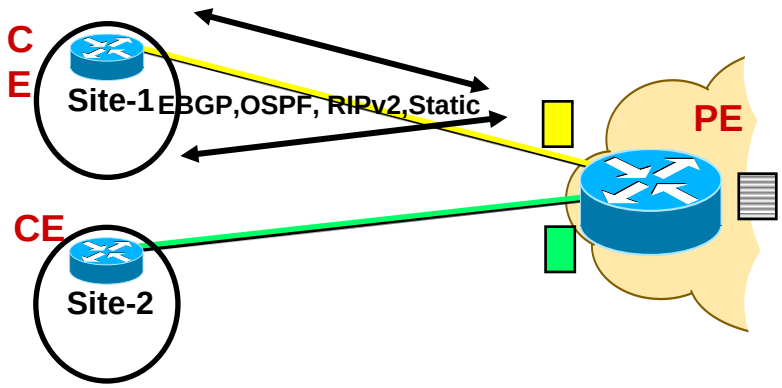
- Route Distinguisher
  - 64 bits
  - Makes the IPv4 route globally unique
  - RD is configured in the PE for each VRF
  - RD may or may not be related to a site or a VPN
- IPv4 address (32bits)

Extended Community attribute (64 bits)

- Route-target (RT): identifies the set of sites the route has to be advertised to

MPLS VPN Connection Model

(L3VPN : mpls/vpn bgp service)



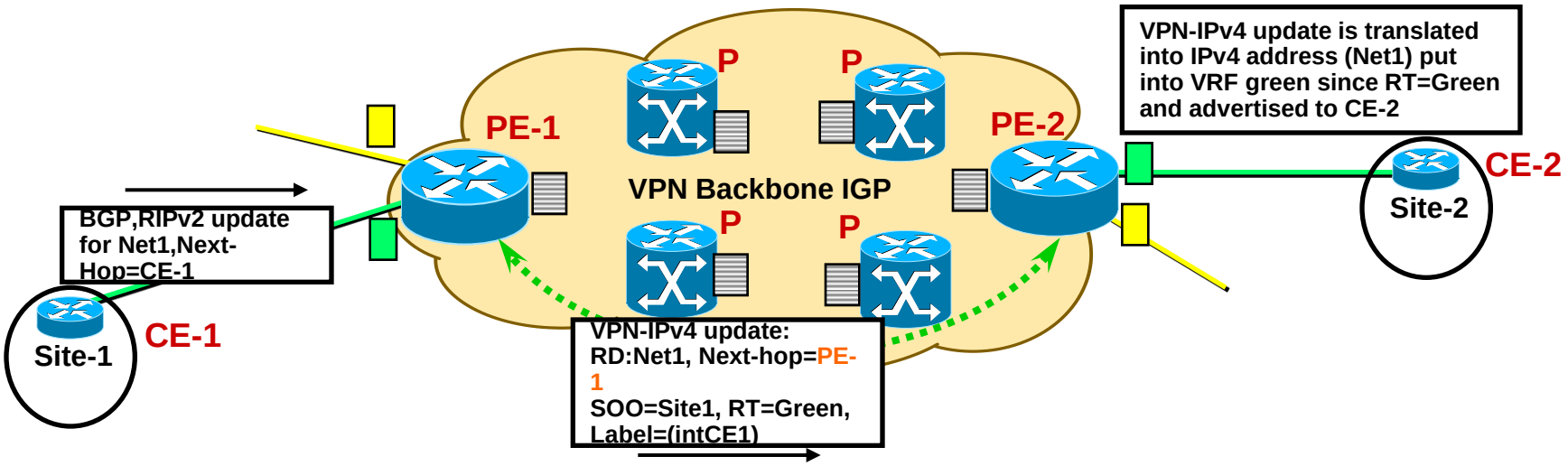
PE and CE routers exchange routing information through:

- EBGP, OSPF, RIPv2, Static routing

CE router run standard routing software

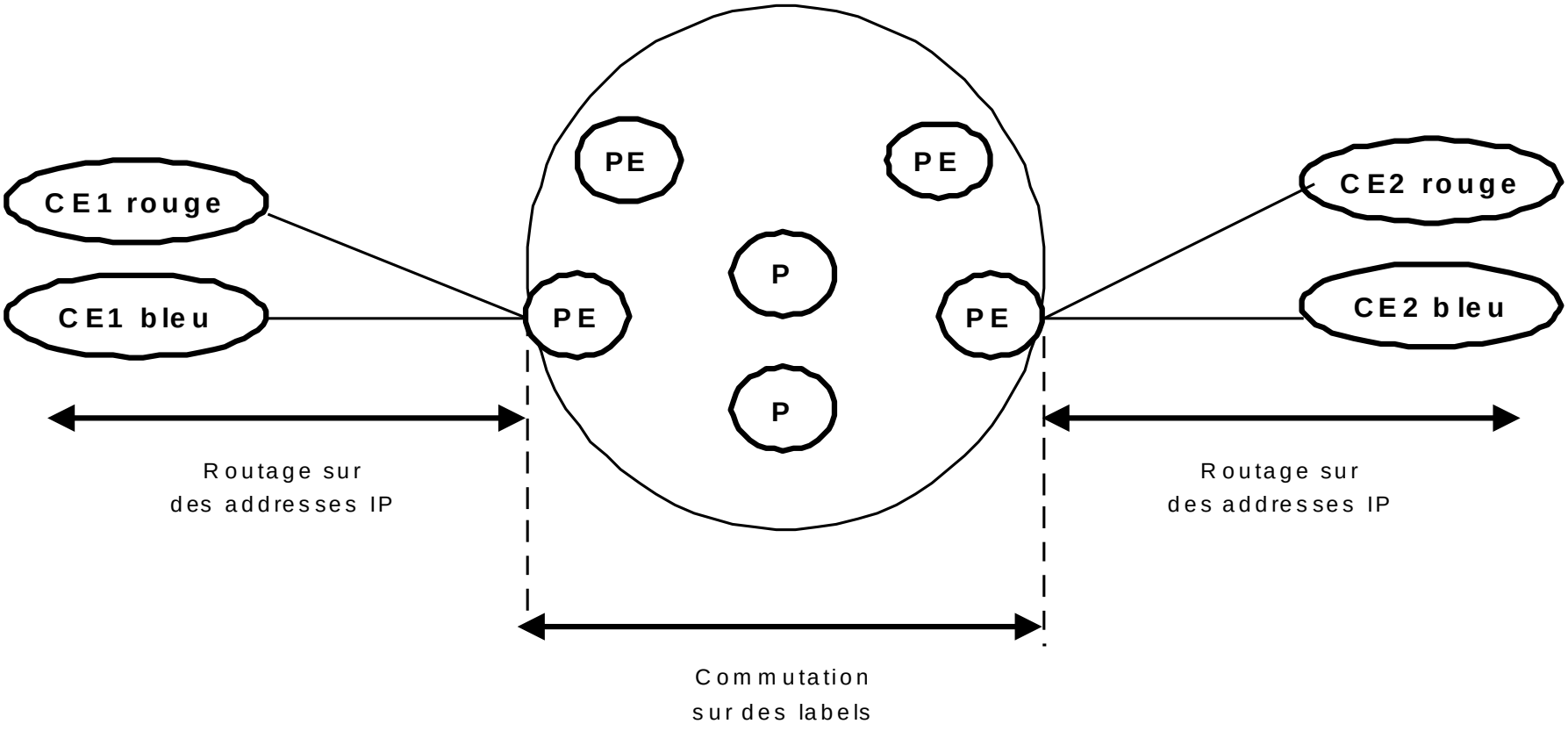
MPLS VPN Connection Model

(L3VPN : mpls/vpn bgp service)



- **PE routers receive IPv4 updates (EBGP, RIPv2, Static)**
- **PE routers translate into VPN-IPv4**
  - Assign a SOO and RT based on configuration
  - Re-write Next-Hop attribute
  - Assign a label based on VRF and/or interface
  - Send MP-iBGP update to all PE neighbors

Description of a multi-services network
(L3VPN : mpls/vpn bgp service)



Description of a multi-services network
(L3VPN : mpls/vpn bgp service)

Configuration du routeur PE, connecté à CE1 rouge et CE1 bleu :

Définition du MPLS/VPN rouge :

```
ip vrf rouge
        rd x1
        route-target import 100 : 1
        route-target export 100 : 1
```

Définition du MPLS/VPN bleu :

```
ip vrf bleu
  rd x2
  route-target import 100 : 2
  route-target export 100 : 2
```

Connexion de CE1 rouge au PE :

```
interface …
        ip vrf forwarding rouge
          …
```

Connexion de CE1 bleu au PE :

```
interface …
        ip vrf forwarding bleu
          …
```

Description of a multi-services network
(L3VPN : mpls/vpn bgp service)

Configuration du routeur PE, connecté à CE2 rouge et CE2 bleu :

Définition du MPLS/VPN rouge :

```
ip vrf rouge
        rd x3
        route-target import 100 : 1
        route-target export 100 : 1
```

Définition du MPLS/VPN bleu :

```
ip vrf bleu
   rd x4
   route-target import 100 : 2
   route-target export 100 : 2
```

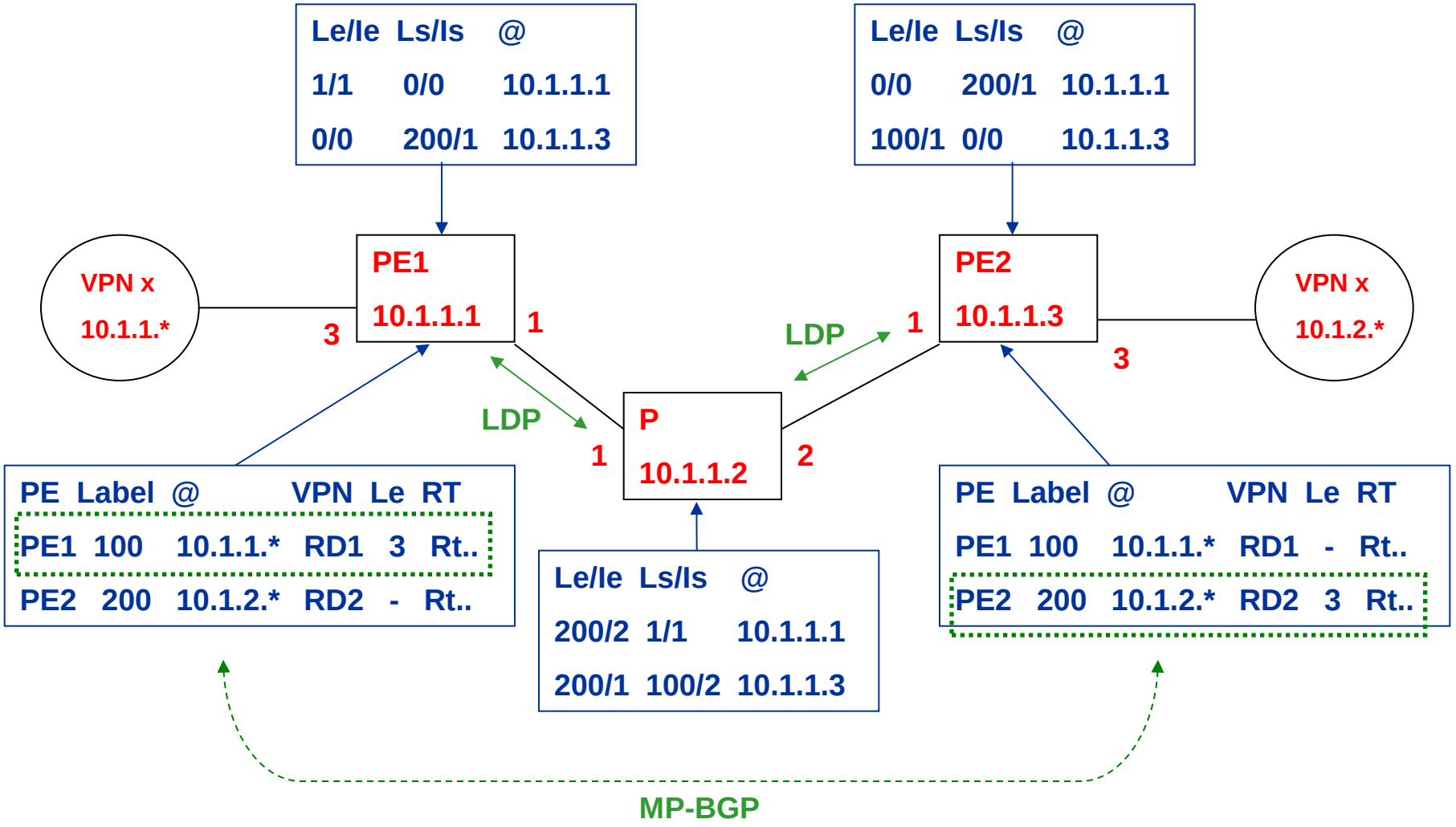
Connexion de CE2 rouge au PE :

```
interface …
        ip vrf forwarding rouge
           …
```

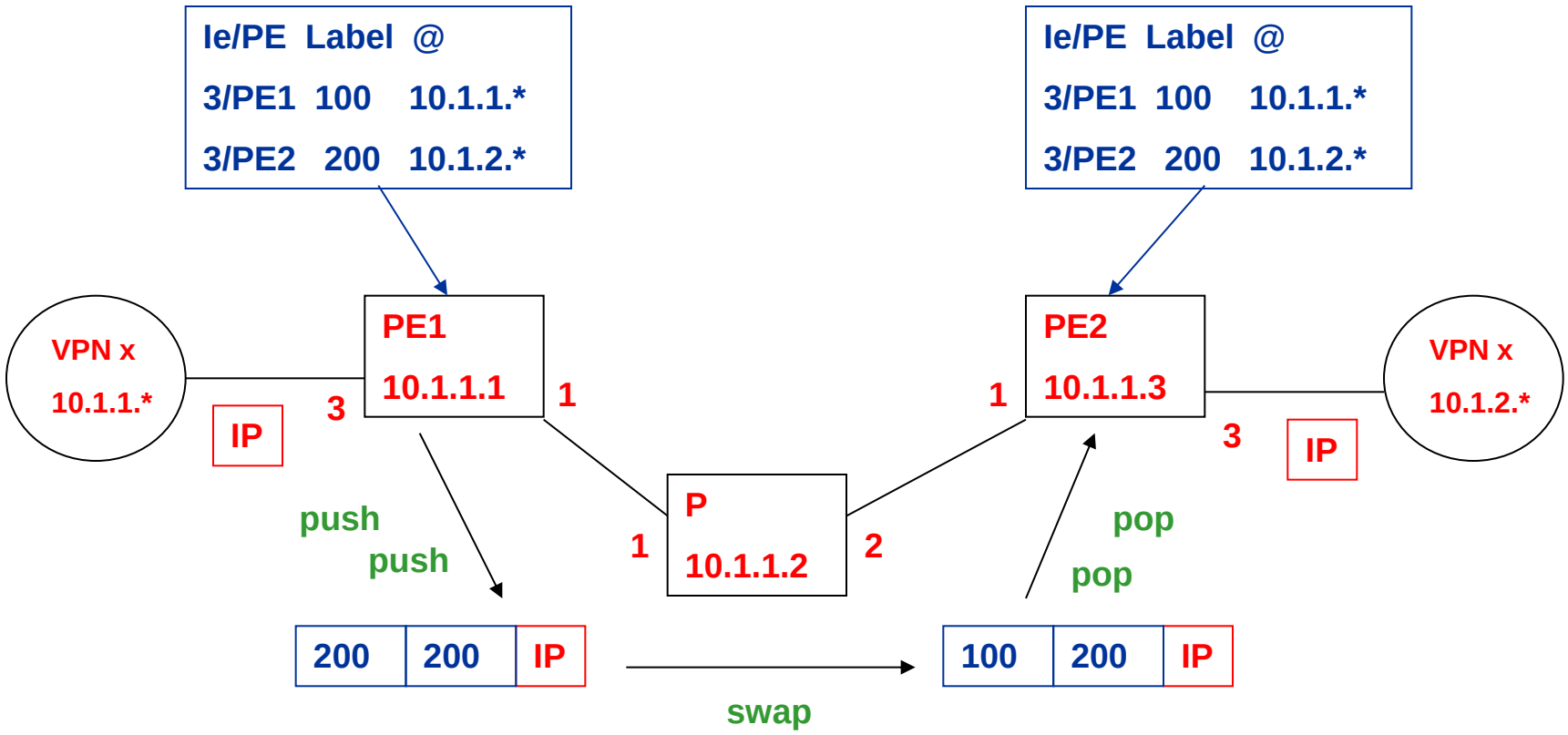
Connexion de CE2 bleu au PE :

```
interface …
        ip vrf forwarding bleu
           …
```

Example : label distribution



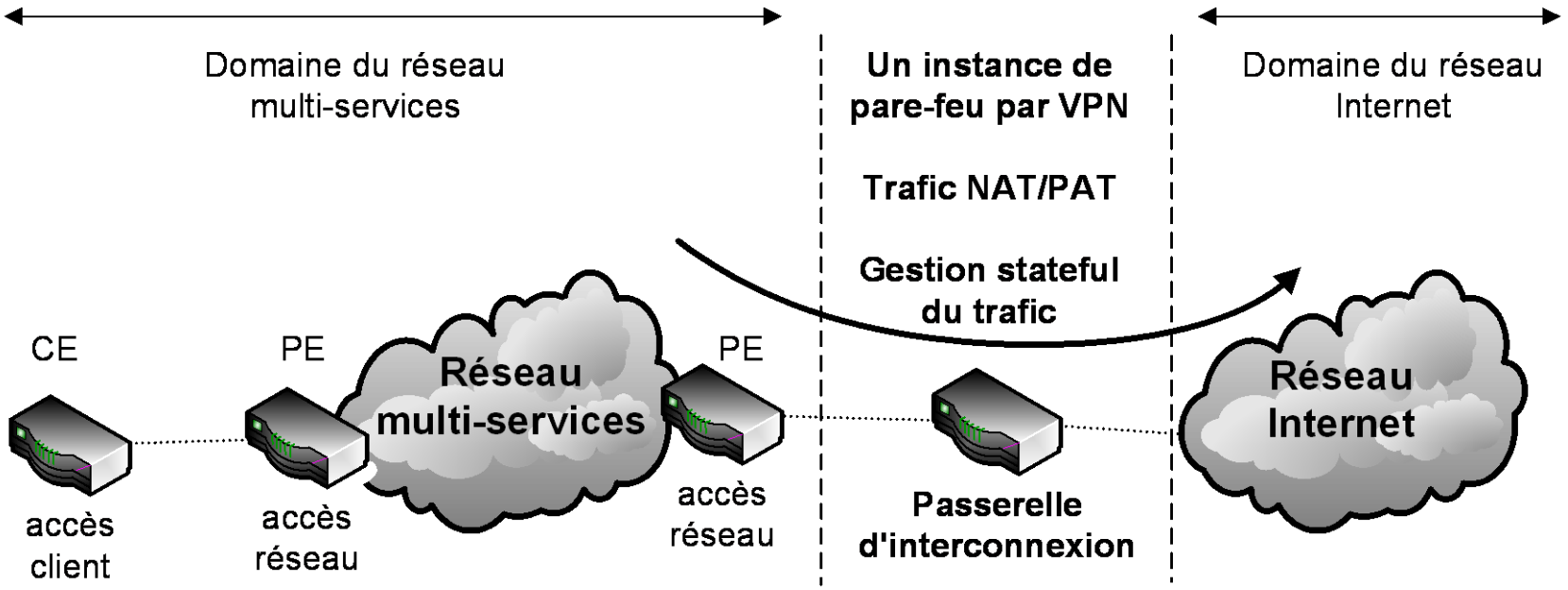
Exemple : packet switching between VPN sites



Agenda

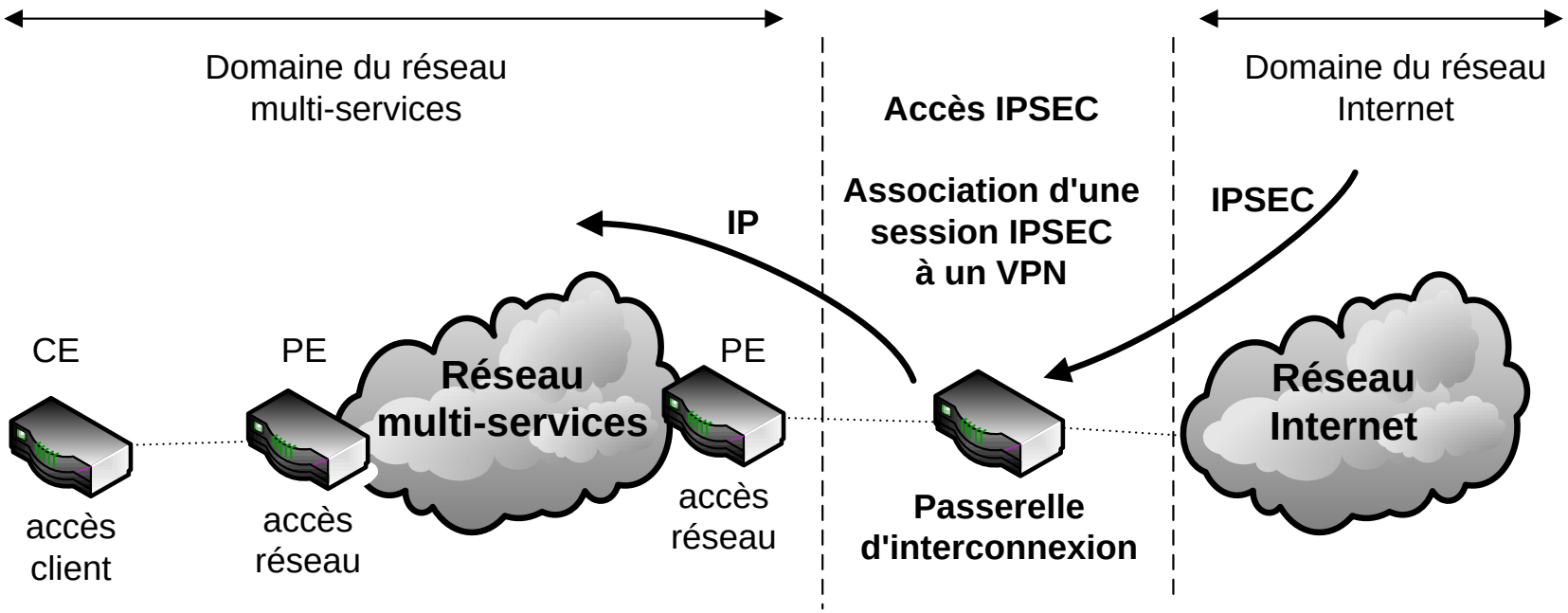
VPN layer 3
and
Internet access

Description of a multi-services network
(L3VPN : mpls/vpn bgp service / Internet access)



Ce service permet de prolonger un VPN BPG/MPLS à Internet.

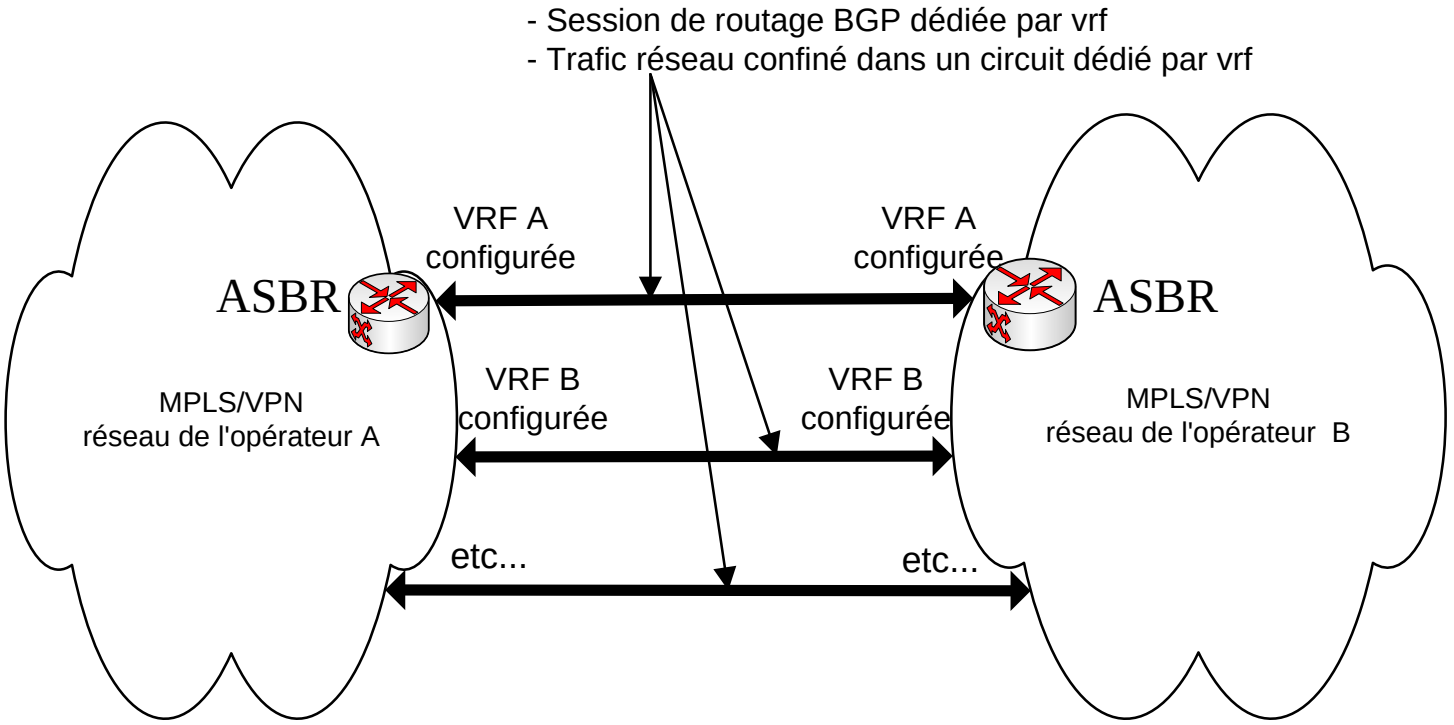
Description of a multi-services network
(L3VPN : mpls/vpn bgp service / Access from Internet)



Ce service permet d'accéder un VPN BGP/MPLS à partir d'Internet.

MPLS BGP/VPNs Inter-AS Interco

Description of a multi-services network
(MPLS/VPN BGP interconnection – Back to Back vrf – Option A)



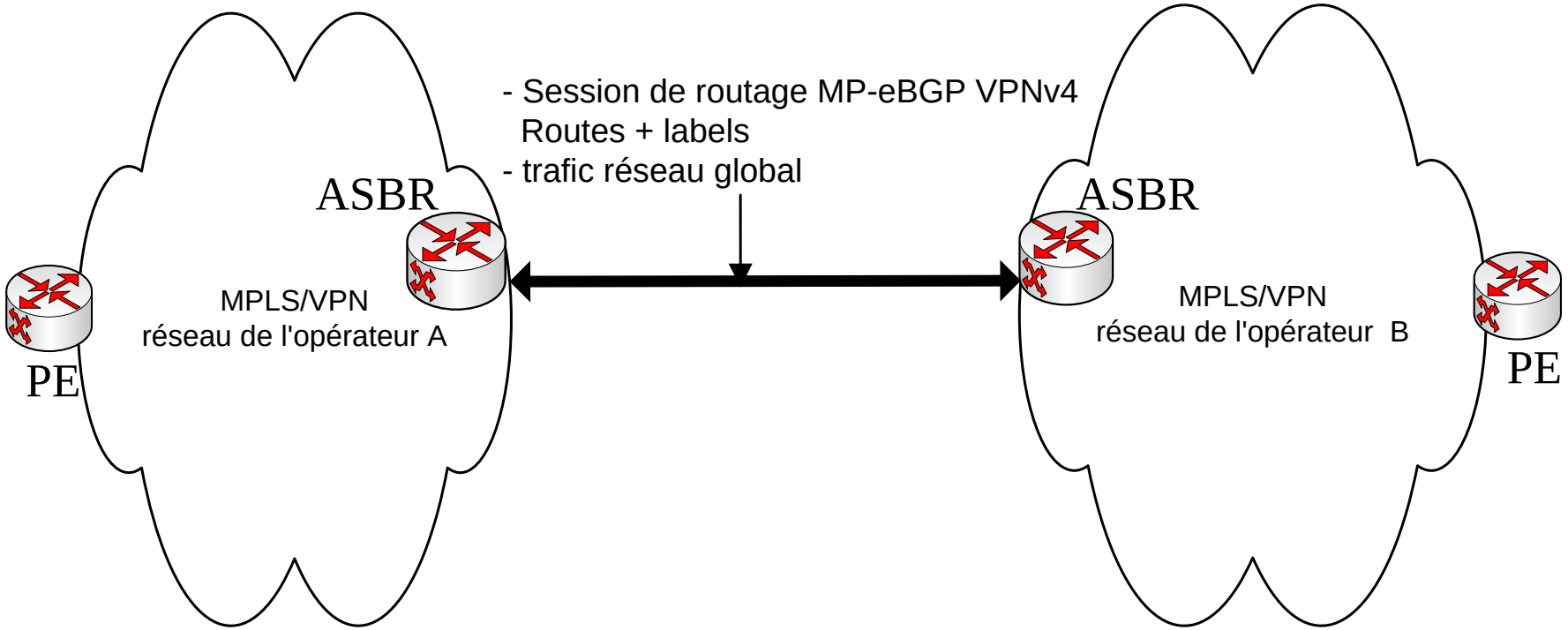
Les avantages de ce modèle sont les suivants:

- Le confinement des VPNs dans des tunnels dédiés en point à point.
- La granularité d'analyse en cas de problème réseau est fine par l'isolation de configuration des VPNs.
- Il est possible de filtrer le trafic IP par VPN.
- Il est possible de filtrer le routage par VPN.

Les désavantages de ce modèle sont les suivants:

- La configuration des VPNs devient consommatrice en terme de ressources mémoire.
- L'architecture d'interconnexion devient complexe.

Description of a multi-services network
MPLS/VPN BGP interconnection – MP-eBGP – Option B)



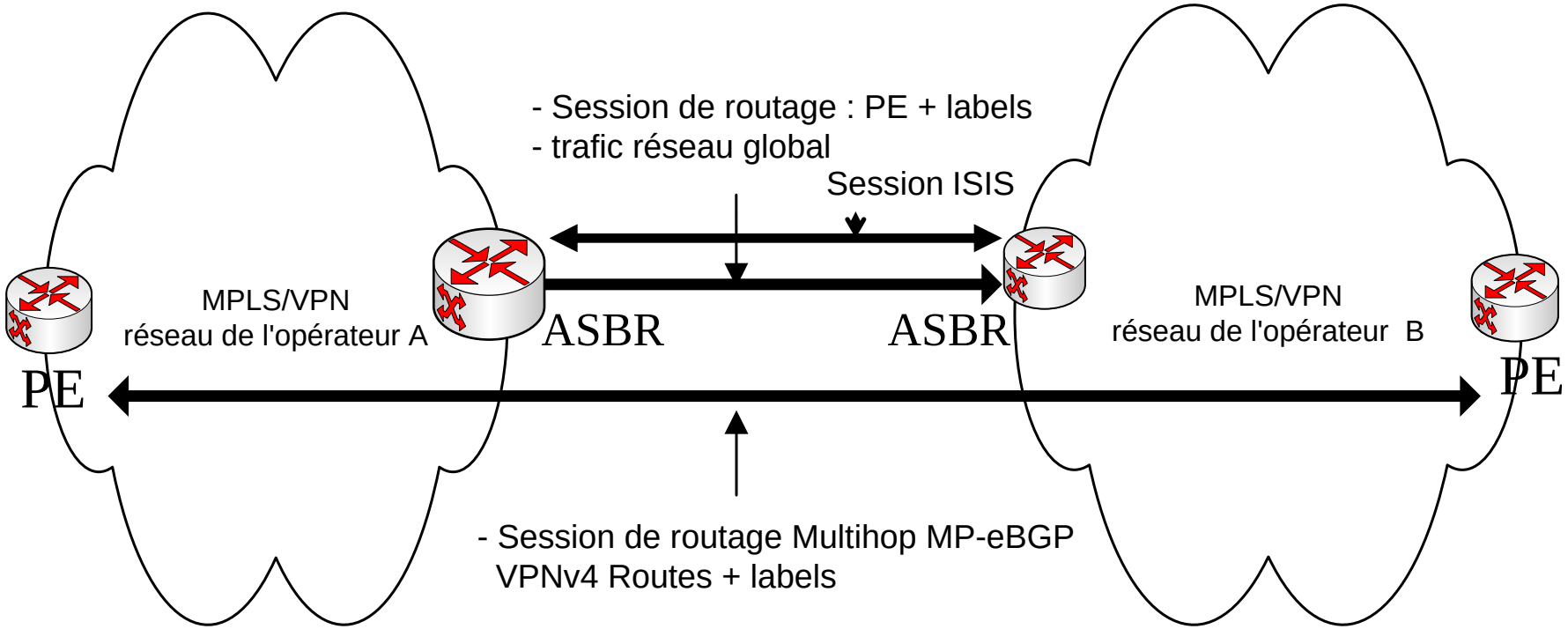
Les avantages de ce modèle sont les suivants:

- La configuration des VPNs est simplifiée.
- Le modèle est extensible.
- L'architecture réseau est simplifiée.

Les désavantages de ce modèle sont les suivants:

- La granularité d'analyse en cas de problème réseau n'est plus fine.
- Il n'y a pas de possibilité de filtrer le trafic IP par VPN.
- Il n'y a pas de possibilité de filtrer le routage des adresses IP par VPN.

Description of a multi-services network
MPLS/VPN BGP interconnection – MP-eBGP / ISIS – Option C)



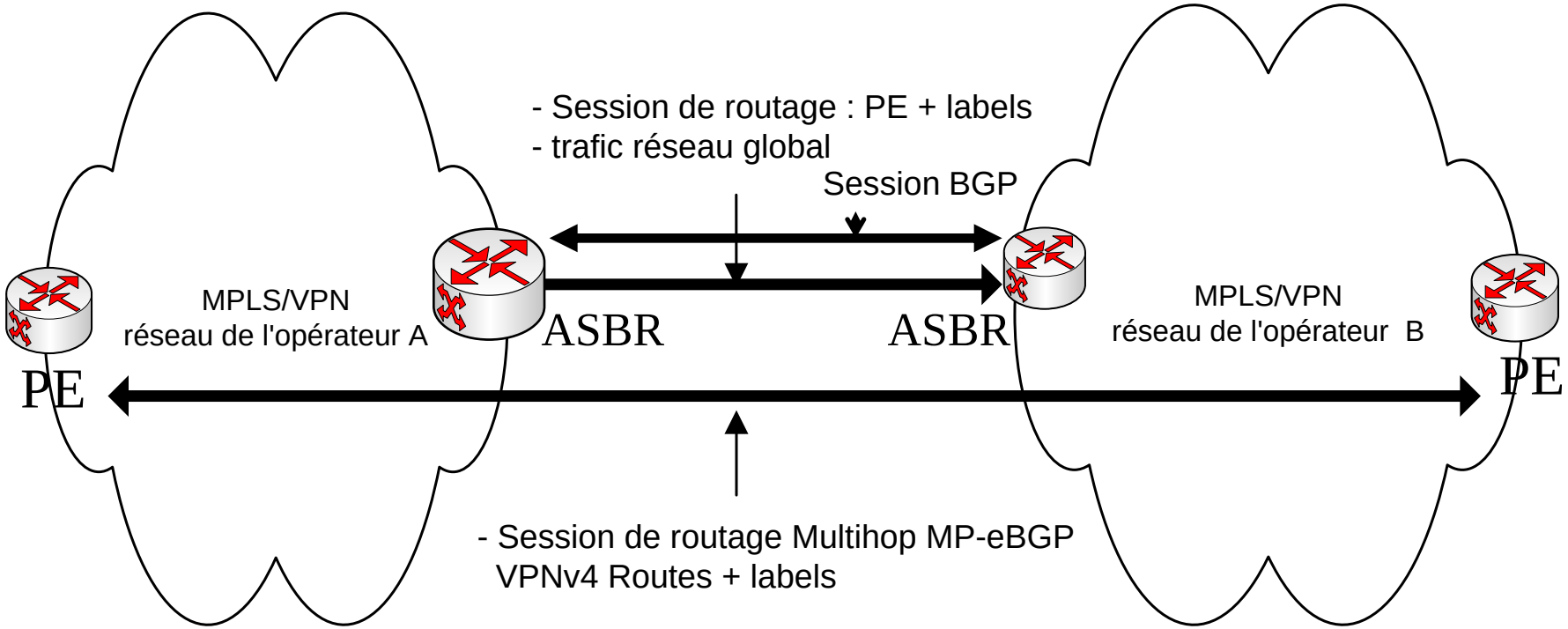
Les avantages de ce modèle sont les suivants:

- La configuration des VPNs est simplifiée.
- Le modèle est extensible.
- L'architecture réseau est simplifiée.

Les désavantages de ce modèle sont les suivants:

- La granularité d'analyse en cas de problème réseau n'est plus fine.
- Il n'y a pas de possibilité de filtrer le trafic IP par VPN.
- Il n'y a pas de possibilité de filtrer le routage des adresses IP par VPN.

Description of a multi-services network
MPLS/VPN BGP interconnection – MP-eBGP / BGP 3701 - Option C)



Les avantages de ce modèle sont les suivants:

- La configuration des VPNs est simplifiée.
- Le modèle est extensible.
- L'architecture réseau est simplifiée.

Les désavantages de ce modèle sont les suivants:

- La granularité d'analyse en cas de problème réseau n'est plus fine.
- Il n'y a pas de possibilité de filtrer le trafic IP par VPN.
- Il n'y a pas de possibilité de filtrer le routage des adresses IP par VPN.

Agenda

Questions ?