F R A M E I P . C O M

Partage des connaissances du monde TCP/IP
(https://www.frameip.com/)

IMPLÉMENTATION DE MPLS AVEC CISCO

Sommaire [Masquer]

Commentaire et discussion
Laisser un commentaire

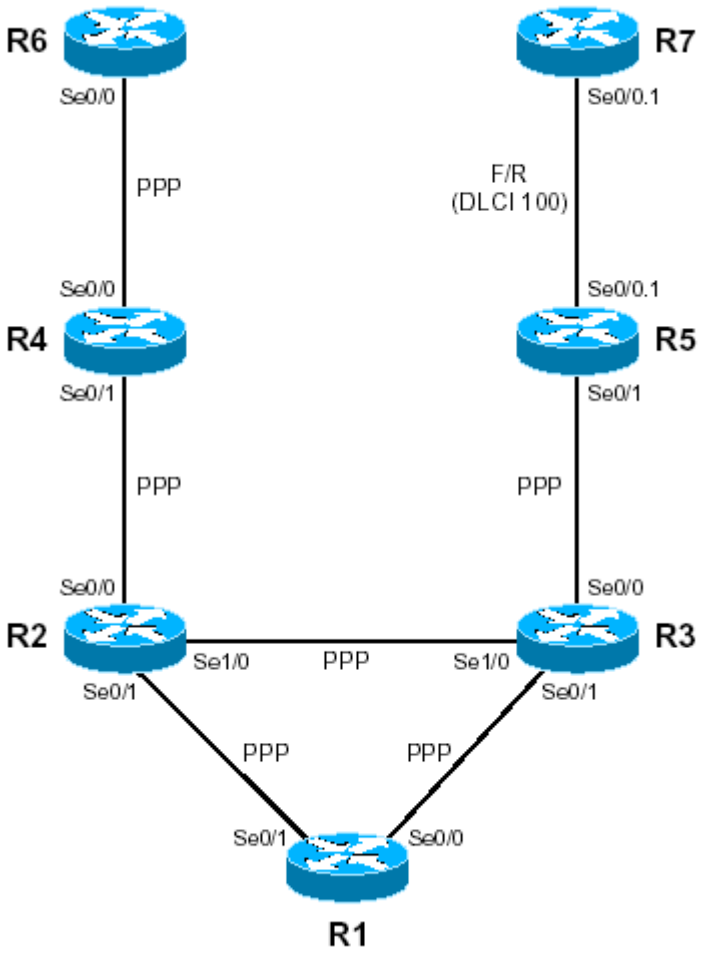
1 – Introduction au protocole MPLS

Dans les réseaux IP traditionnels, le routage des paquets s'effectue en fonction de l'adresse de destination contenue dans l'entête de niveau 3. Chaque routeur, pour déterminer le prochain saut (next-hop), consulte sa table de routage et détermine l'interface de sortie vers laquelle envoyer le paquet. Le mécanisme de recherche dans la table de routage est consommateur de temps CPU, et avec la croissance de la taille des réseaux ces dernières années, les tables de routage des routeurs ont constamment augmenté. Il était donc nécessaire de trouver une méthode plus efficace pour le routage des paquets. Le but de MPLS était à l'origine de donner aux routeurs IP une plus grande puissance de commutation, en basant la décision de routage sur une information de label (ou tag) inséré entre le niveau 2 (Data-Link Layer) et le niveau 3 (Network Layer). La transmission des paquets était ainsi réalisée en switchant les paquets en fonction du label, sans avoir à consulter l'entête de niveau 3 (http://www.frameip.com/osi/) et la table de routage.

Toutefois, avec le développement de techniques de commutation comme CEF (Cisco Express Forwarding) et la mise au point de nouveaux ASIC (Application Specific Interface Circuits), les routeurs IP ont vu leurs performances améliorées sans le recours à MPLS. L'intérêt de MPLS n'est actuellement plus la rapidité mais l'offre de services qu'il permet, avec notamment les réseaux privés virtuels (VPN) (http://www.frameip.com/vpn/) et le Traffic Engineering (TE), qui ne sont pas réalisables sur des infrastructures IP traditionnelles. Ce document se focalise principalement sur la présentation des principes de MPLS et une étude approfondie de MPLS/VPN. Des notions essentielles de Traffic Engineering sont également présentées en dernière partie.

2 – Matériel Cisco et configurations utilisées

Afin d'étayer d'exemples pratiques les différents principes de MPLS abordés dans ce document, deux « pods » (nommés L10 et L20) composés de 7 routeurs chacun ont été utilisés. Ces routeurs sont reliés au moyen de liaisons série, de la manière suivante :



Les deux pods sont connectés entre eux par les deux routeurs R1, au moyen d'une interface FastEthernet.

Les équipements mis en jeu sont des routeurs des familles 2600 et 3600 :

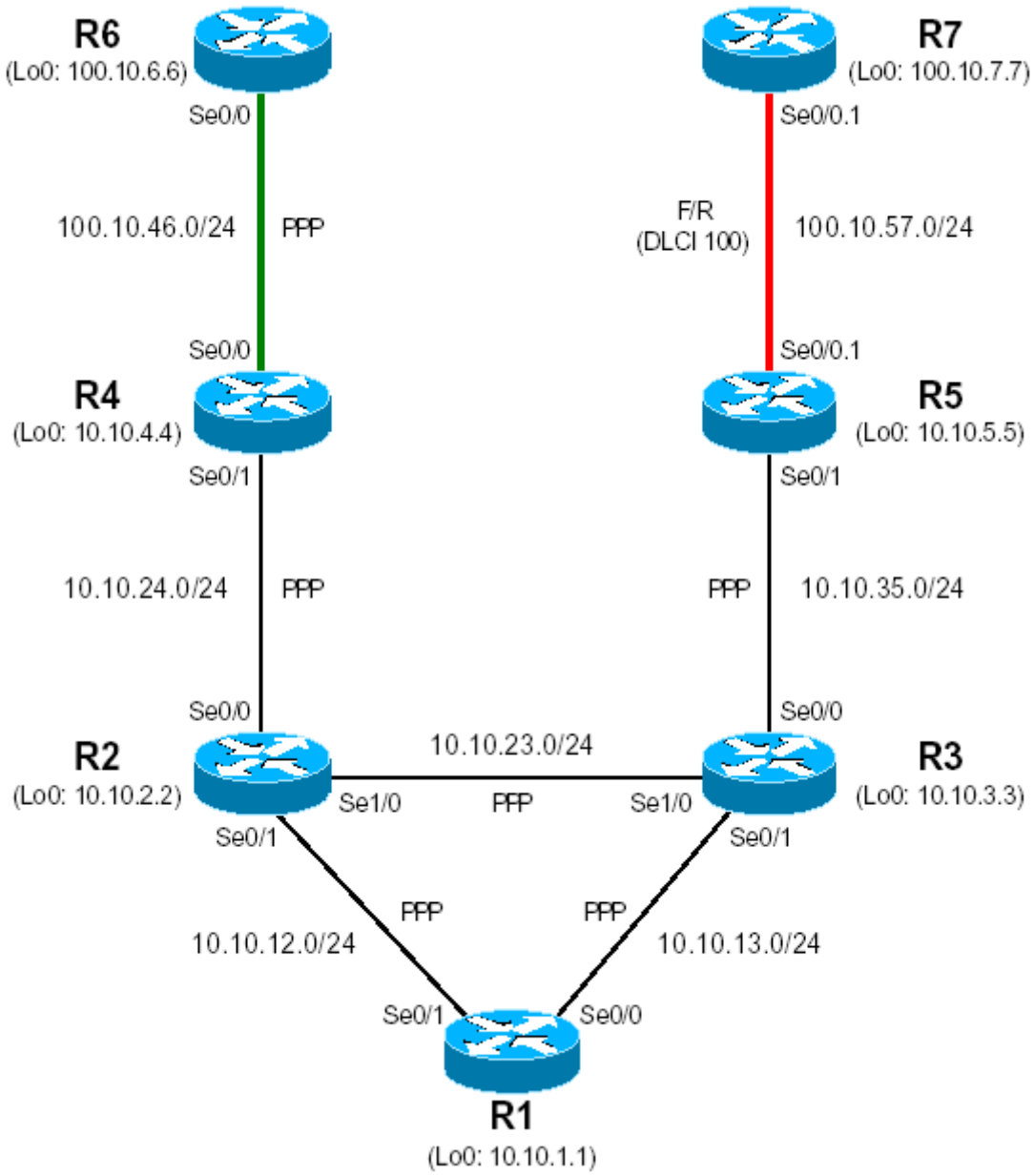
- Modèle 3640 : R2, R3
- Modèle 3620 : R1, R4, R5
- Modèle 2600 : R6, R7

La version de Cisco IOS tournant sur les routeurs est la 12.2(0.4).

3 – Principes généraux – Terminologie

3.1 – Réseau Cisco de démonstration

Pour cette partie, seul le pod L10 a été utilisé. Le schéma suivant résume les différents subnets et adresses IP configurés sur les routeurs Cisco :



Tous les routeurs utilisent OSPF comme protocole de routage interne (IGP) et toutes les interfaces séries ont été configurées pour fonctionner avec MPLS. Les configurations des routeurs pour cette partie sont fournies en Annexe 1.

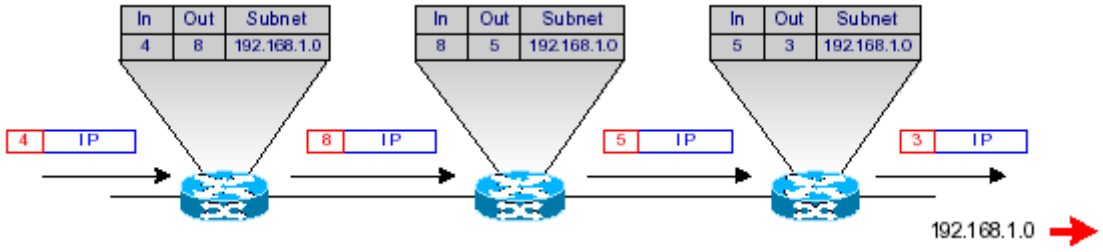
La convention utilisée pour les adresses IP est la suivante:

- Interface Loopback0 : 10.10.x.x pour routeur Rx ;
- Subnet entre deux routeurs Rx et Ry (x < y) : 10.10.xy.0/24 ;
- Adresse IP pour Rx : 10.10.xy.x et Ry : 10.10.xy.y.

3.2 – Commutation par labels

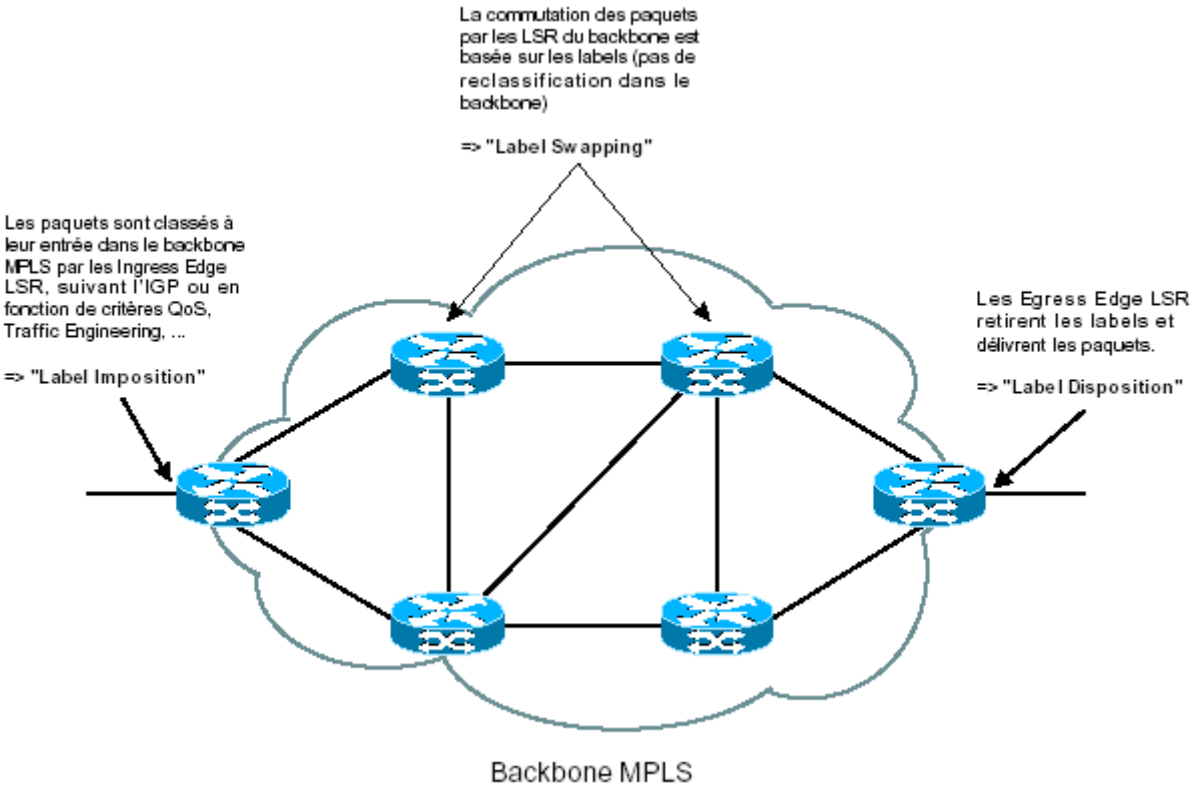
Comme il l'a été brièvement expliqué en introduction, le principe de base de MPLS est la commutation de labels. Ces labels, simples nombres entiers, sont insérés entre les entêtes de niveaux 2 et 3, les routeurs permutant ces labels tout au long du réseau jusqu'à destination, sans avoir besoin de consulter l'entête IP et leur table de routage. Cette technique, appelée Label Swapping, est similaire à la commutation de cellules sur ATM avec les informations de VPI/VCI ou à la commutation sur réseau Frame Relay avec les DLCI. Toutefois, MPLS permet de définir des piles de labels (label stack), dont l'intérêt apparaîtra avec le TE et les VPN. Les routeurs réalisant les opérations de label swapping sont appelés LSR pour Label Switch Routers.

Le schéma suivant montre un exemple de label swapping :



Les routeurs MPLS situés à la périphérie du réseau (Edge LSR), qui possèdent à la fois des interfaces IP traditionnelles et des interfaces connectées au backbone MPLS, sont chargés d'imposer ou de retirer les labels des paquets IP qui les traversent. Les routeurs d'entrées, qui imposent les labels, sont appelés Ingress LSR, tandis que les routeurs de sortie, qui retirent les labels, sont appelés Egress LSR.

Le schéma suivant montre le rôle des différents routeurs en fonction de leur emplacement dans le réseau MPLS:



3.3 – Classification des paquets

A l'entrée du réseau MPLS, les paquets IP sont classés dans des FEC (Forwarding Equivalent Classes). Des paquets appartenant à une même FEC suivront le même chemin et auront la même méthode de forwarding. Typiquement, les FEC sont des préfixes IP appris par l'IGP tournant sur le backbone MPLS, mais peuvent aussi être définies par des informations de QoS ou de TE. La classification des paquets s'effectue à l'entrée du backbone MPLS, par les Ingress LSR. A l'intérieur du backbone MPLS, les paquets sont label-switchés, et aucune reclassification des paquets n'a lieu. Chaque LSR affecte un label local, qui sera utilisé en entrée, pour chacune de ses FEC et le propage à ses voisins. Les LSR

voisins sont appris grâce à l'IGP. L'ensemble des LSR utilisés pour une FEC, constituant un chemin à travers le réseau, est appelé Label Switch Path (LSP). Il existe un LSP pour chaque FEC et les LSP sont unidirectionnels.

3.4 – Mode trame et mode cellule

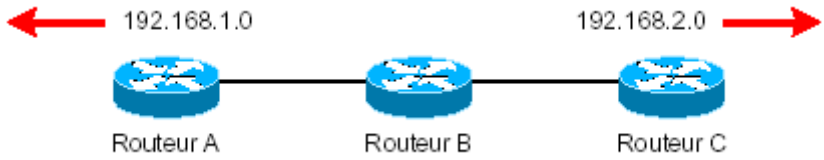
Il existe deux catégories d'interfaces MPLS sur les routeurs, dépendant de leur mode de fonctionnement. Le premier mode, appelé mode trame (framed mode), correspond aux interfaces traitant des paquets de taille variable, comme par exemple Ethernet, Frame-Relay, PPP, etc. Le second mode concerne les interfaces ATM et est appelé mode cellule (cell mode), la commutation étant basée sur la notion de circuit. Sur ATM, les circuits virtuels sont définis par les champs VPI/VCI de l'entête des cellules. Suivant le mode de fonctionnement d'une interface, les méthodes de propagation des labels aux routeurs voisins diffèrent.

3.5 – Distribution des labels (TDP / LDP)

Les LSR se basent sur l'information de label pour commuter les paquets au travers du backbone MPLS. Chaque routeur, lorsqu'il reçoit un paquet taggué, utilise le label pour déterminer l'interface et le label de sortie. Il est donc nécessaire de propager les informations sur ces labels à tous les LSR. Pour cela, des protocoles de distributions de labels sont utilisés. Suivant le type des FEC, différents protocoles sont employés pour l'échange de labels entre LSR :

- TDP/LDP (Tag/Label Distribution Protocol): Mapping des adresses IP unicast ;
- RSVP (Resource Reservation Protocol): utilisé en Traffic Engineering pour établir des LSP en fonction de critères de ressources et d'utilisation des liens ;
- MP-BGP (MultiProtocol Border Gateway Protocol) pour l'échange de routes VPN.

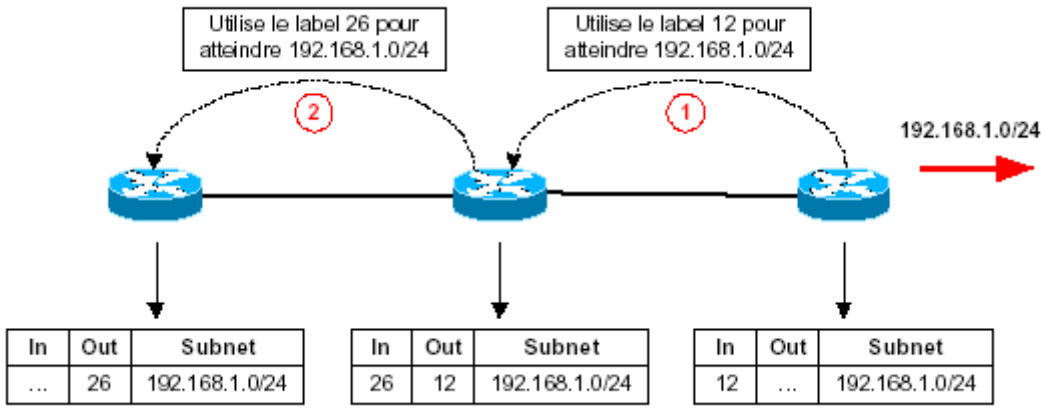
Les deux derniers protocoles seront abordés dans leurs sections respectives (Traffic Engineering et Virtual Private Networks). Remarque : aucun label n'est affecté pour les routes apprises par eBGP. Il existe deux méthodes pour propager les labels entre LSR: upstream et downstream. Le schéma suivant explicite la notion d'upstream neighbor et de downstream neighbor par rapport à un réseau IP donné:



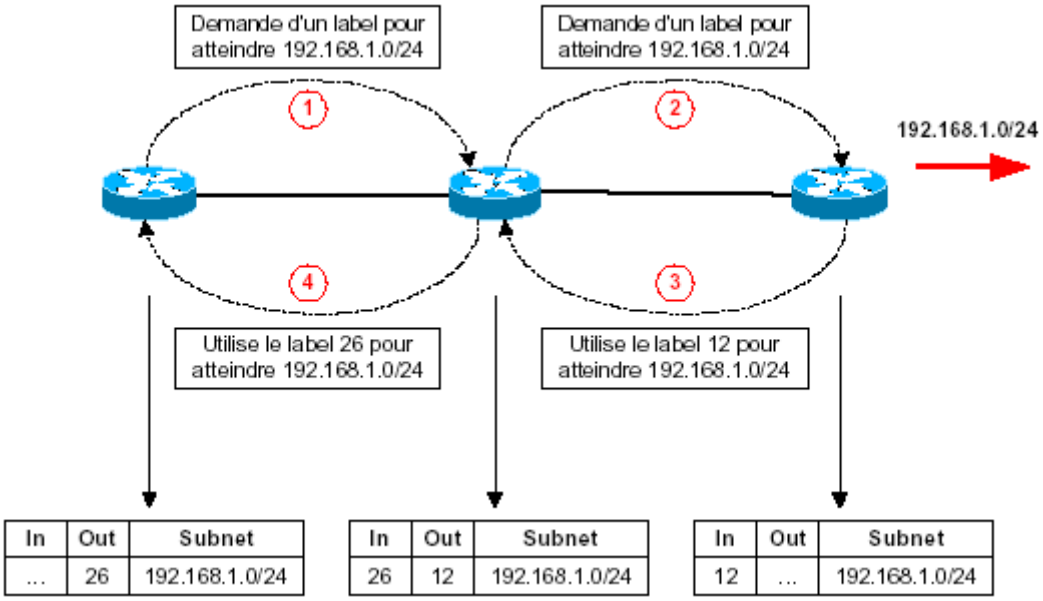
Sur le schéma ci-dessus, le routeur A est un upstream neighbor par rapport au routeur B pour le réseau 192.168.2.0. Le routeur A est aussi downstream neighbor par rapport au routeur B pour le réseau 192.168.1.0. Une méthode de distribution des labels dite « downstream » indique que la propagation des réseaux se fait du routeur le plus proche au routeur le plus éloigné (downstream vers upstream).

La méthode downstream, avec deux variantes: unsollicited downstream et downstream on demand. Dans la première variante, les LSR downstream propagent systématiquement tous leurs labels à leurs voisins. Dans la deuxième, les LSR upstreams demandent explicitement aux LSR downstreams de leur fournir un label pour le subnet IP demandé. Le mode non sollicité est utilisé dans le cas d'interfaces en mode trame, le downstream on demand étant utilisé par les LSR ATM (mode cellule).

- Unsollicited Downstream :



- Downstream on demand :



Pour échanger les labels correspondants aux routes unicast apprises par un IGP, les routeurs Cisco emploient TDP (Tag Distribution Protocol), utilisant TCP sur le port 711. Ce protocole est un protocole propriétaire défini par Cisco Systems. Le protocole défini par l'IETF est LDP (Label Distribution Protocol), qui utilise TCP sur le port 646. Bien que ces deux protocoles soient fonctionnellement identiques, ils sont incompatibles entre eux, à cause de différences dans le format des paquets. A l'avenir, Cisco IOS pourra utiliser soit TDP ou LDP, ou bien les deux simultanément.

Deux routeurs sont configurés pour échanger des labels par TDP avec la commande Cisco« tag-switching ip », spécifiée sur les interfaces qu'ils ont en commun.

Il est possible de connaître tous les voisins TDP d'un routeur en utilisant la commande Cisco« show tag-switching tdp neighbor » :

```
L10-R1# sh tag tdp neigh
 Peer TDP Ident: 10.10.3.3:0; Local TDP Ident 10.10.1.1:0
    TCP connection: 10.10.3.3.11004 - 10.10.1.1.711
    State: Oper; PIEs sent/rcvd: 1727/1740; ; Downstream
    Up time: 1d01h
    TDP discovery sources:
       Serial0/0
    Addresses bound to peer TDP Ident:
       10.10.3.3 10.10.35.3 10.10.13.3 10.10.23.3
 Peer TDP Ident: 10.10.2.2:0; Local TDP Ident 10.10.1.1:0
    TCP connection: 10.10.2.2.11006 - 10.10.1.1.711
    State: Oper; PIEs sent/rcvd: 1607/1616; ; Downstream
    Up time: 23:23:28
    TDP discovery sources:
       Serial0/1
    Addresses bound to peer TDP Ident:
       10.10.2.2 100.10.20.20 10.10.24.2 10.10.12.2
```

Chaque voisin est listé avec toutes les adresses IP qui lui appartiennent. La méthode d'allocation des labels (unsollicited downstream, downstream on demand) est également indiquée. Comme les interfaces des routeurs de cet exemple sont de type série, il s'agit d'interfaces en mode trame et le mode unsollicited downstream est employé.

Pour pouvoir établir correctement une adjacence TDP, les deux voisins doivent être convenablement configurés. La commande Cisco « show tag-switching tdp discovery » permet de s'assurer du bon établissement de l'adjacence :

```
L10-R1# sh tag tdp disc
 Local TDP Identifier:
    10.10.1.1:0
 TDP Discovery Sources:
    Interfaces:
       Serial0/0: xmit/recv
          TDP Id: 10.10.3.3:0
       Serial0/1: xmit/recv
          TDP Id: 10.10.2.2:0
```

Chaque voisin doit être marqué « xmit/recv » (émission / réception) pour que l'échange des labels puisse avoir lieu.

3.6 – Tables MPLS: TIB et TFIB

A partir des informations apprises par TDP / LDP, les LSR construisent deux tables, la TIB et la TFIB. De manière générale, la TIB contient tous les labels appris des LSR voisins, tandis que la TFIB, utilisée pour la commutation proprement dite des paquets, est un sous-ensemble de la TIB.

3.6.1 – Rôle de la TIB (Tag Information Base)

La première table construite par le routeur Cisco MPLS est la table TIB (Tag Information Base). Elle contient pour chaque subnet IP la liste des labels affectés par les LSR voisins. Il est possible de connaître les labels affectés à un subnet par chaque LSR voisin en utilisant la commande Cisco « show tag tdp bindings ». Un exemple de résultat de cette commande Cisco est donné ci-dessous:

```
L10-R1# sh tag tdp bind 10.10.4.4 255.255.255.255
    tib entry: 10.10.4.4/32, rev 31
       local binding: tag: 24
       remote binding: tsr: 10.10.3.3:0, tag: 20
       remote binding: tsr: 10.10.2.2:0, tag: 21
```

On remarque que le routeur Cisco a affecté le label local 24 pour atteindre le réseau 10.10.4.4/32, et que les routeurs L10-R2 (10.10.2.2) et L10-R3 (10.10.3.3) ont respectivement affecté les label 21 et 20 pour atteindre le subnet 10.10.4.4/32. Il est à noter qu'IOS emploie le terme TSR pour « Tag Switch Router », qui est équivalent à celui de LSR. Pour les interfaces ATM (fonctionnant en mode cellule), la commande Cisco à utiliser est « show tag atm-tdp bindings » :

```
ATM-LSR# show tag-switching atm-tdp bindings
 Destination: 193.12.161.1/32
    Tailend Switch XTagATM162 241/33 Active -> Terminating Active, VCD=2
 Destination: 194.16.16.4/32
    Transit XTagATM161 240/91 Active -> XTagATM162 241/276 Active
```

Les labels entre ATM LSR sont échangés au moyen d'un VC de contrôle MPLS, par défaut configuré sur VPI/VCI = 0/32.

3.6.2 – Rôle de la TFIB (Tag Forwarding Information Base)

A partir de la table TIB et de la table de routage IP, le routeur Cisco construit une table TFIB, qui sera utilisée pour commuter les paquets. Chaque réseau IP est appris par l'IGP, qui détermine le prochain saut (« next-hop ») pour atteindre ce réseau. Le LSR choisit ainsi l'entrée de la table TIB qui correspond au réseau IP et sélectionne comme label de sortie le label annoncé par le voisin déterminé par l'IGP (plus court chemin). Il est possible d'afficher le contenu de la table TFIB grâce à la commande Cisco « show tagswitching forwarding ». Le résultat de cette commande Cisco sur le routeur utilisé précédemment est donné ci-dessous :

```
L10-R1# sh tag for 10.10.4.4
 Local Outgoing Prefix Bytes tag Outgoing Next Hop
 tag tag or VC or Tunnel Id switched interface
 24 21 10.10.4.4/32 0 Se0/1 point2point
```

On remarque ainsi que le routeur Cisco L10-R1 a sélectionné pour le réseau 10.10.4.4/32 l'entrée de la TIB créée par le voisin 10.10.2.2 (connecté à L10-R1 par l'interface Serial0/1), qui a la meilleure métrique du point de vue de l'IGP (plus court chemin). Ainsi, pour chaque paquet reçu ayant comme

label 24, le routeur Cisco commutera le paquet sur l'interface de sortie Serial0/1, et en permutant le label 24 par 21. La sélection de L10-R2 comme next-hop est confirmée en consultant l'entrée 10.10.4.4/32 de la table de routage :

```
L10-R1# sh ip route 10.10.4.4
 Routing entry for 10.10.4.4/32
    Known via "ospf 10", distance 110, metric 1601, type intra area
    Last update from 10.10.12.2 on Serial0/1, 23:16:16 ago
    Routing Descriptor Blocks:
    * 10.10.12.2, from 10.10.4.4, 23:16:16 ago, via Serial0/1
       Route metric is 1601, traffic share count is 1
```

Le routeur, lorsqu'il reçoit un paquet taggué, se base sur la TFIB pour forwarder le paquet. A partir d'un label d'entrée (local tag), il en déduit l'interface et le label de sortie (Outgoing interface et Outgoing tag or VC). Pour pouvoir utiliser la TFIB, le routeur Cisco doit employer CEF comme technique de commutation, qui doit être activée globalement et pour chaque interface recevant des paquets taggués. CEF est en effet le seul mode de commutation capable d'utiliser la TFIB. Les anciens modes (fastswitching, optimum switching, etc.) ne sont pas conçus pour gérer cette table.

Il est possible de consulter la table CEF d'un routeur avec la commande Cisco « show ip cef ». Tous les préfixes IP connus seront alors affichés avec leur interface de sortie et l'adresse du next-hop. Il est possible d'obtenir des informations plus détaillées sur un subnet particulier avec la commande Cisco « show ip cef subnet netmask ». Il est ainsi aisé de connaître le(s) label(s) de sortie utilisés pour atteindre ce réseau, l'interface de sortie et le next-hop IP, comme le montre l'exemple suivant :

```
L10-R1# sh ip cef 10.10.4.4
 10.10.4.4/32, version 594, cached adjacency to Serial0/1
 0 packets, 0 bytes
    tag information set
       local tag: 24
       fast tag rewrite with Se0/1, point2point, tags imposed: {21}
    via 10.10.12.2, Serial0/1, 0 dependencies
       next hop 10.10.12.2, Serial0/1
    valid cached adjacency
    tag rewrite with Se0/1, point2point, tags imposed: {21}
```

L'interface de sortie à emprunter pour atteindre le subnet 10.10.4.4/32 est donc Serial0/1, avec comme adresse de next-hop 10.10.12.2 (routeur Cisco L10-R2). Le tag local affecté par le routeur Cisco L10-R1 est 24 et le tag utilisé en sortie est 21 (appris de L10-R2 par TDP).

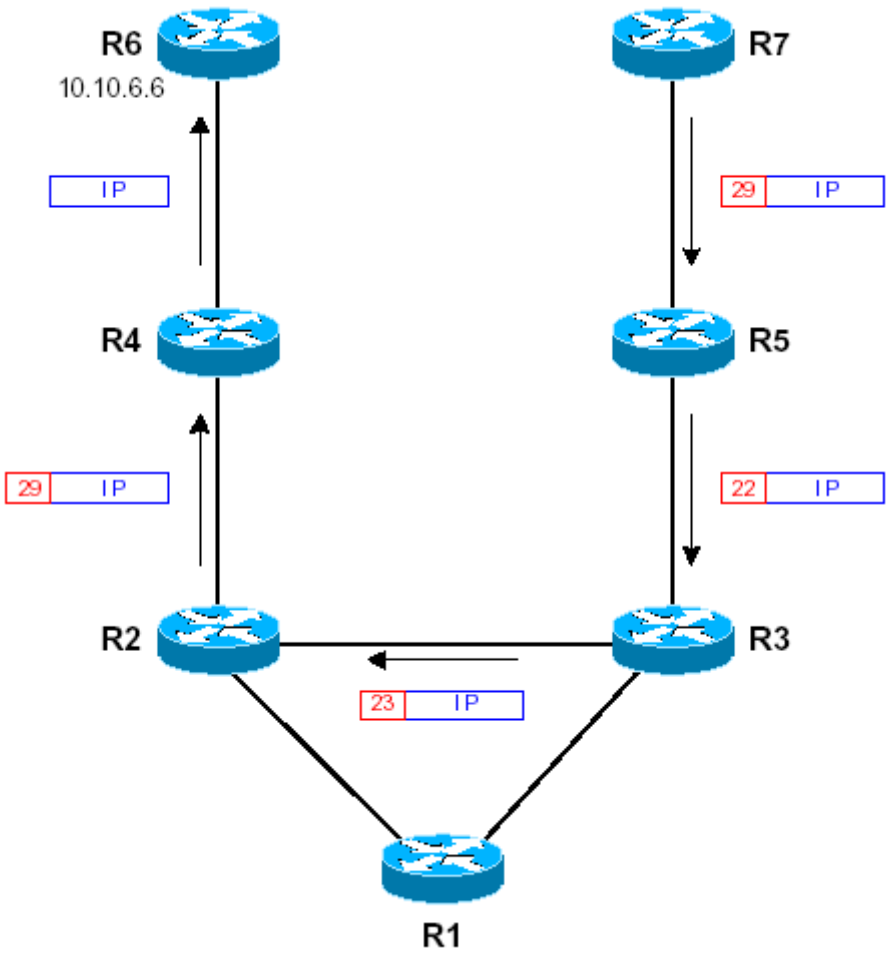
Sur L10-R2, le contenu de la TIB pour 10.10.4.4/32 est reproduit ci-dessous :

```
L10-R2# sh tag tdp bind 10.10.4.4 255.255.255.255
    tib entry: 10.10.4.4/32, rev 26
       local binding: tag: 21
       remote binding: tsr: 10.10.4.4:0, tag: imp-null
       remote binding: tsr: 10.10.1.1:0, tag: 24
```

On retrouve donc bien comme tag d'entrée le tag 21 pour atteindre le réseau 10.10.4.4. A partir de la version IOS 12.1(5)T, il est possible de connaître les labels d'un chemin servant à atteindre une destination précise, avec la commande Cisco « traceroute » :

```
L10-R7# trace 10.10.6.6

 Type escape sequence to abort.
 Tracing the route to 10.10.6.6

    1 10.10.57.5 [MPLS: Label 29 Exp 0] 120 msec 116 msec 116 msec
    2 10.10.35.3 [MPLS: Label 22 Exp 0] 105 msec 108 msec 104 msec
    3 10.10.23.2 [MPLS: Label 23 Exp 0] 92 msec 100 msec 96 msec
    4 10.10.24.4 [MPLS: Label 29 Exp 0] 89 msec 92 msec 84 msec
    5 10.10.46.6 40 msec * 40 msec
```

Le label MPLS affiché pour chaque hop correspond au label en entrée du routeur. Le champ « Exp » (codé sur 3 bits) est similaire au champ TOS de l'entête IP, mais n'est pas employé ici. Dans cet exemple, le chemin pour atteindre R6 à partir de R7 est { R5, R3, R2, R4, R6 }. Le schéma suivant montre comment les paquets sont acheminés de R7 jusqu'à R6 :

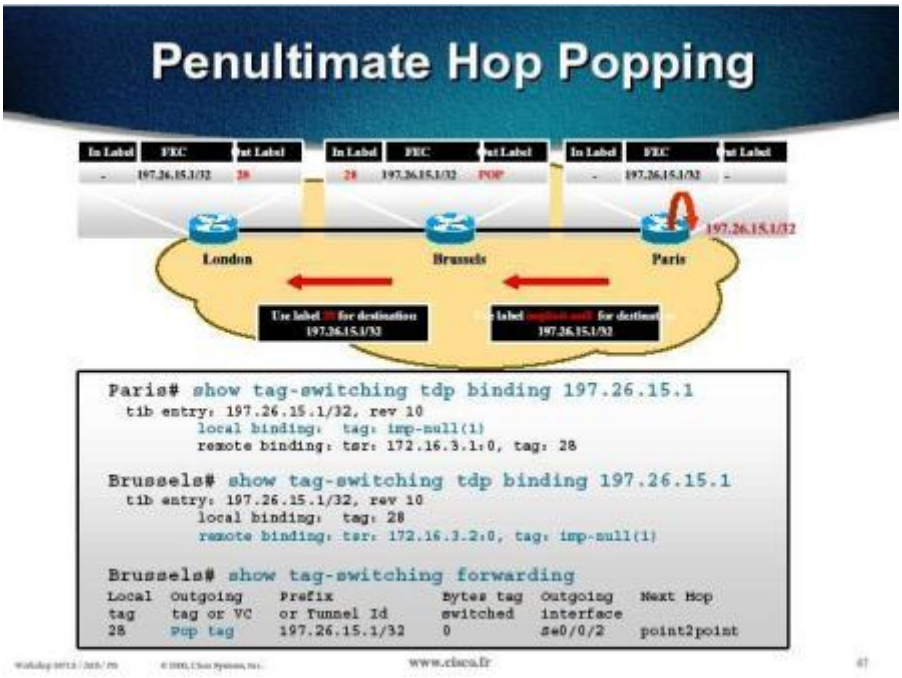


Les routeurs R5, R3, R2 et R4 commutent les paquets uniquement sur l'entête MPLS :

l'entête IP n'est pas examinée, et les routeurs ne consultent que leur table TFIB (leur table de routage n'est pas utilisée). On constate que les paquets arrivant sur le routeur Cisco R6 pour le réseau 10.10.6.6 ne sont pas taggués. Ce phénomène, appelé Penultimate Hop Popping, permet au routeur Cisco auquel sont rattachés des réseaux sur des interfaces non MPLS d'éviter un lookup dans la table TFIB. Le Penultimate Hop Popping est décrit plus précisement dans le paragraphe suivant.

3.7 – Penultimate Hop Popping

Un LSR « egress » annonçant un réseau, qui lui est soit directement connecté, soit rattaché (appris par IGP, EGP, routage statique…) par une interface non tagguée, n'a pas besoin de recevoir de paquets taggués pour atteindre ce réseau. En effet, si les paquets reçus étaient taggués, le routeur Cisco egress devrait d'abord déterminer l'interface de sortie grâce à la table TFIB, puis effectuer une recherche dans la table de routage IP. L'opération de recherche sur le label dans la TFIB est inutile, car dans tous les cas le routeur Cisco devra effectuer une recherche dans la table de routage. Le routeur Cisco egress annonce donc ces réseaux IP avec le label « implicit-null » à ses voisins. Un LSR ayant comme label de sortie « implicit-null » aura ainsi pour but de dépiler le premier label du paquet et de faire suivre le paquet sur l'interface de sortie spécifiée. Le routeur Cisco egress n'aura alors plus qu'une recherche à faire dans sa table de routage.



Un exemple de Penultimate Hop Popping entre L10-R1 et L10-R2 est donné ci-dessous :

```
L10-R1# sh tag for 10.10.2.2
 Local Outgoing Prefix Bytes tag Outgoing Next Hop
 tag tag or VC or Tunnel Id switched interface
 23 Pop tag 10.10.2.2/32 145198 Se0/1 point2point

 L10-R1# sh ip cef 10.10.2.2
 10.10.2.2/32, version 593, cached adjacency to Serial0/1
 0 packets, 0 bytes
    tag information set
       local tag: 23
    via 10.10.12.2, Serial0/1, 0 dependencies
       next hop 10.10.12.2, Serial0/1
       valid cached adjacency
       tag rewrite with Se0/1, point2point, tags imposed: {}
```

Le réseau 10.10.2.2/32 correspond à l'interface Loopback0 du routeur Cisco L10-R2, connecté par un lien série à L10-R1. On constate que le tag de sortie (Outgoing tag) pour 10.10.2.2/32 est déclaré sous le terme « pop tag » pour signaler l'action de dépilement du premier label.

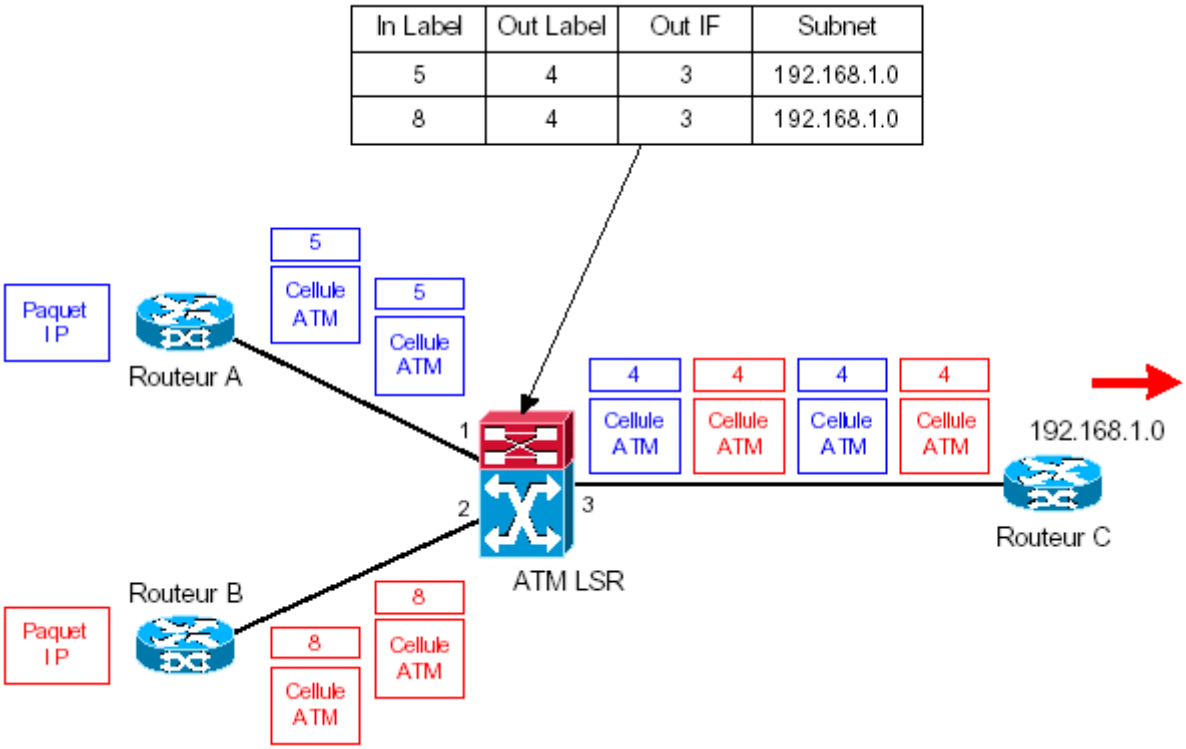
3.8 – Rétention des labels

Afin d'accélérer la convergence du réseau lors d'un changement de topologie (lien défectueux, dysfonctionnement d'un routeur), les LSR conservent dans leur table TIB la liste des labels annoncés pour chaque réseau IP par leurs voisins TDP, y compris de ceux n'étant pas les next-hops choisis par l'IGP. Ainsi, en cas de perte d'un lien ou d'un noeud, la sélection d'un nouveau label de sortie est immédiate : en effet, il suffit au routeur Cisco d'élire un nouveau next-hop et de sélectionner l'entrée correspondante dans la TIB, puis de mettre à jour la TFIB. Ce mode de fonctionnement est appelé mode libéral (liberal mode). L'avantage de ce procédé est naturellement une convergence plus rapide lorsque les informations de routage au niveau 3 changent, avec pour inconvénients que davantage de mémoire est allouée dans les routeurs et que des labels supplémentaires sont utilisés. Le mode libéral est appliqué dans le cas d'interfaces fonctionnant en mode trame.

Il existe un autre mode appelé mode conservatif, qui correspond au downstream on demand, utilisé par les LSR ATM. Pour atteindre un subnet donné au-delà d'une interface de type « cellule », les LSR ATM demandent à leurs voisins downstream de leur fournir un label pour chaque couple (interface d'entrée, subnet IP). La problématique des réseaux ATM est abordée dans le paragraphe suivant.

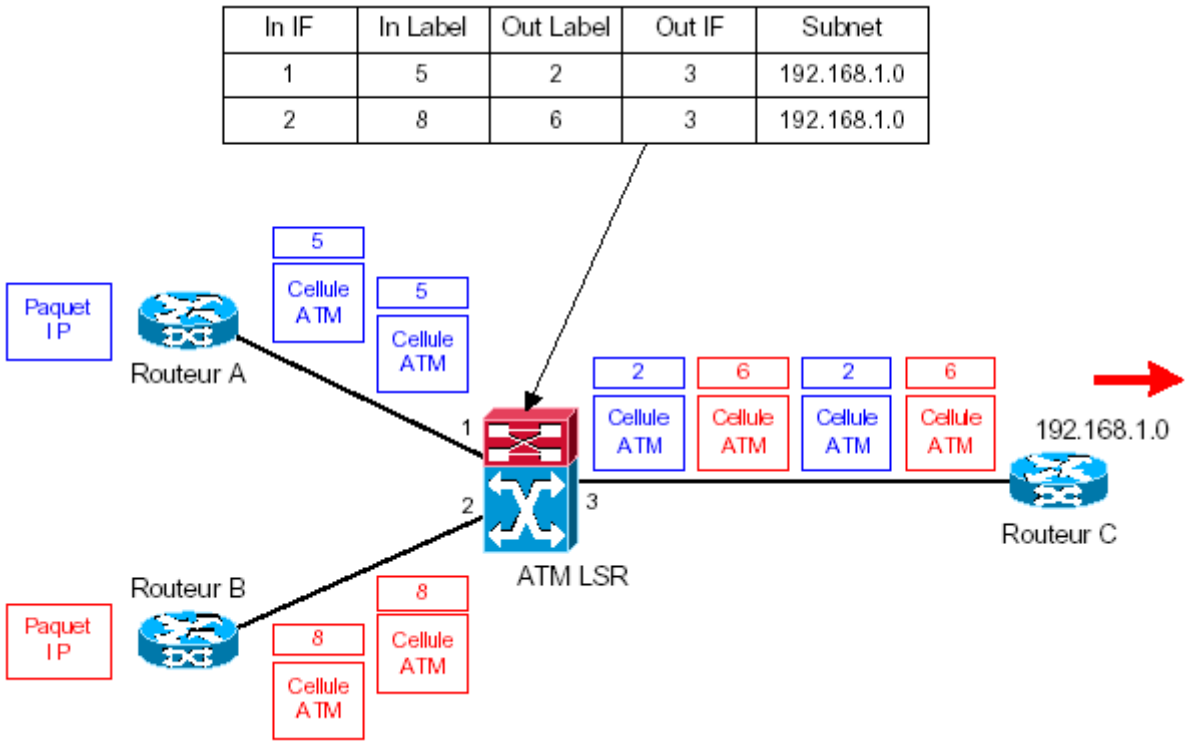
3.9 – MPLS sur ATM

Il existe deux manières d'implémenter MPLS sur des réseaux de type ATM. La première consiste à mettre en place un backbone constitué de switches purement ATM, c'est-à-dire sans aucune connaissance de MPLS ou du routage IP. Dans ce cas, des PVCs sont simplement établis entre les routeurs MPLS et les labels sont alors encapsulés entre l'entête LLC/SNAP et l'entête IP. La deuxième méthode consiste à mettre en oeuvre MPLS sur des switches ATM dits « IP-aware », c'est-à-dire ayant connaissance de la topologie IP grâce à un protocole de routage, et où l'information de label est encodée dans les champs VPI/VCI. Ces switches sont alors appelés ATM LSR. Ce paragraphe aborde les spécificités d'un backbone MPLS composé de LSR ATM par rapport à un backbone purement IP, notamment dans les mécanismes de distribution des labels. Le MPLS sur ATM natif ayant un fonctionnement similaire à des LSR « traditionnels », cette architecture ne sera pas étudiée ici. Pour distribuer des labels MPLS entre LSR ATM, les protocoles TDP / LDP en mode downstream on demand sont utilisés. Si le mode unsollicited downstream était employé, comme dans le cas de LSR non ATM, on aurait le scénario suivant :



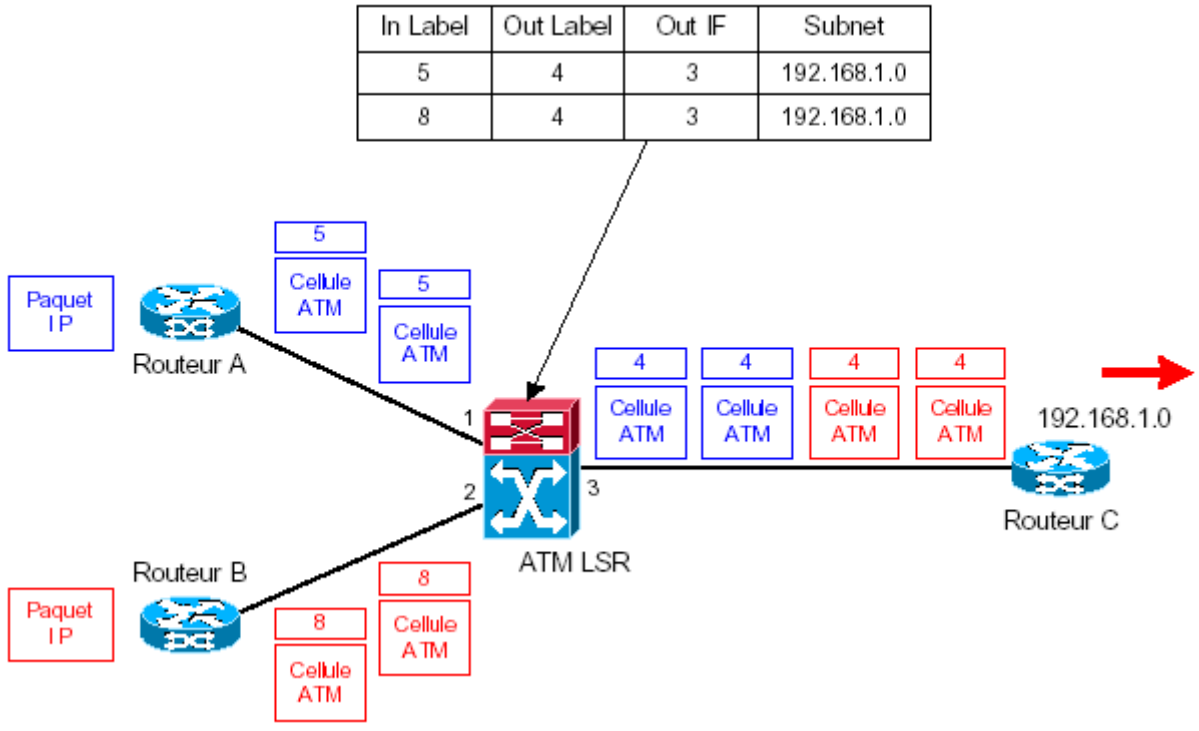
Dans cet exemple, le routeur C aurait fourni au switch ATM le label 4 pour atteindre le subnet 192.168.1.0/24. On remarque alors que si des paquets IP sont envoyés par les routeurs A et B à destination de ce subnet, les cellules ATM reçues par le routeur C ont toutes pour label 4. Le label étant encodé dans les champs VPI / VCI pour des LSR ATM, il y a mélange des cellules composant les paquets IP, sans moyen de resynchronisation (impossible de distinguer les cellules les unes des autres pour reformer les paquets). La solution mise en oeuvre pour éviter le mélange des cellules est d'affecter un label en fonction du subnet de destination et de l'interface d'entrée. Dans ce cas de figure, les LSR upstream demandent à leurs voisins downstream de leur fournir un label pour chaque subnet IP et pour chacune de leur interface d'entrée. Ce mode de fonctionnement est donc appelé « downstream on demand ». Il est à noter que le choix du mode de distribution des labels est fixé automatiquement de manière optimale par les routeurs (en fonction du type des interfaces), sans possibilité de modification au niveau de la configuration.

Le schéma ci-dessous montre le fonctionnement des LSR ATM avec un label de sortie défini pour chaque couple (interface d'entrée, subnet IP) :



Sur cet exemple, le switch ATM, fonctionnant en mode downstream on demand, a demandé au routeur C de lui fournir deux labels (2 et 6) pour atteindre le subnet 192.168.1.0 : un label différent est alors utilisé en fonction de l'interface d'entrée pour atteindre le même subnet.

L'allocation de plusieurs labels, mappés dans les champs VPI/VCI, peut rapidement dépasser les limites des équipements ATM. En effet, bien que les champs VPI/VCI soient codés sur 32 bits, il peut exister des limitations hardware, qui dans certains cas, ne permettent pas d'utiliser plus d'un certain nombre de VC par interface. Le VC Merge permet de réduire le nombre de labels utilisés sur une interface ATM, tout en gardant les paquets IP synchronisés. Le principe de cette méthode est de grouper les cellules composant un paquet IP dans un buffer et de ne les émettre sur l'interface de sortie que lorsque tout le paquet a été reçu. Les cellules sont émises dans l'ordre et le LSR downstream les recevant peut donc reconstituer le paquet dans risque de mélange, grâce au champ End Of Frame de l'entête AAL5.

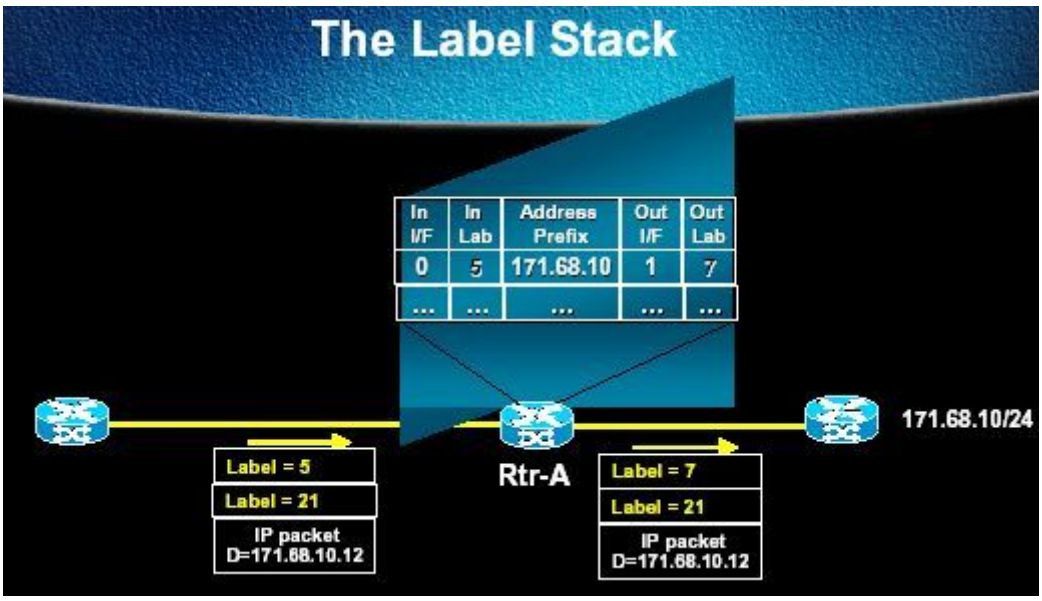


L'avantage de cette méthode est de pouvoir affecter un label unique pour chaque subnet IP traité. Toutefois, la bufferisation des paquets augmente la latence de transmission des paquets, et le débit de l'interface risque d'être limité.

3.10 – Pile de labels (label stacking)

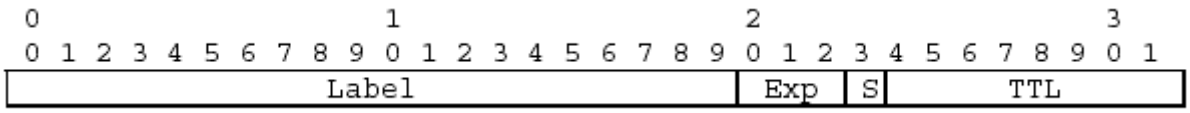
Chaque paquet MPLS est susceptible de transporter plusieurs labels, formant ainsi une pile de labels, qui sont empilés et dépilés par les LSR. Cette possibilité d'empiler des labels, désignée sous le terme de Label Stacking, est utilisée par le Traffic Engineering et MPLS / VPN.

Lorsqu'un LSR commute un paquet, seul le premier label est traité, comme le montre la figure suivante:



3.11 – Description de l'entête MPLS

Un label MPLS occupe 4 octets (32-bits) et se présente sous la forme:



La signification des différents champs est données ci-dessous:

- Label (20 bits)
- Exp (3 bits): Champ expérimental, utilisé pour la QoS. Equivalent au champ TOS de l'entête IP ;
- S (1 bit): Champ « bottom of stack ». Lorsque ce bit est à 1, le bas de la pile est atteint, et l'entête de niveau 3 est placé juste après.
- TTL (8 bits): Ce champ a le même rôle que le champ TTL de l'entête IP.

Le format des labels MPLS est générique et peut notamment être utilisé sur Ethernet, 802.3, PPP (http://www.frameip.com/l2tp-pppoe-ppp-ethernet/), Frame-Relay et sur des PVC ATM (backbone ATM natif). En cas d'emploi d'un médium non supporté (par ex. ISDN), des tunnels GRE peuvent être mis en place. L'adjacence TDP peut alors s'établir entre les deux extrémités du tunnel et les paquets labellisés sont encapsulés dans IP.

3.12 – Configuration d'un routeur Cisco LSR

Cette partie, axée sur la configuration IOS, indique la liste des différentes étapes devant être suivies pour configurer MPLS sur un backbone IP. Les configurations résultantes sont fonctionnelles bien que dénuées d'intérêt pratique, aucun service spécifique à MPLS n'étant mis en oeuvre. Elles permettent toutefois d'appréhender les changements induits par MPLS au niveau des commandes IOS par rapport à des routeurs purement IP. Les configurations minimales MPLS ainsi décrites peuvent être consultées en Annexe 1 de ce document.

3.12.1 – Configuration de CEF

La première opération à effectuer pour utiliser MPLS est d'activer CEF (Cisco Express Forwarding) comme méthode de commutation sur tous les routeurs du backbone. En effet, CEF est la seule méthode de routage capable d'utiliser la TFIB pour commuter les paquets. En cas d'oubli, MPLS ne sera pas fonctionnel. CEF se configure avec la commande Cisco globale: « ip cef [distributed] ». Le mot-clé optionnel « distributed » permet d'activer CEF de manière distribuée sur les routeurs disposant de

cartes de routage et de cartes filles comme les cartes VIP des routeurs 7500. Ce type de carte fait tourner une version réduite d'IOS et a une certaine autonomie de fonctionnement car disposant d'un processeur et de mémoire dédiée.

3.12.2 – Configuration d'un IGP

Un protocole de routage interne doit être utilisé sur le backbone pour pouvoir diffuser les labels MPLS. Il est conseillé d'utiliser un protocole « link-state », tel que OSPF ou IS-IS, qui sont les seuls à permettre le Traffic Engineering. Il est bien entendu nécessaire de s'assurer que la connectivité est établie partout sur le backbone avant de procéder à la configuration de MPLS.

3.12.3 – Configuration de TDP / LDP

Pour permettre à un routeur Cisco d'établir une adjacence TDP avec un voisin sur une interface donnée, cette interface doit être configurée avec la commande Cisco « tagswitching ip ». Bien que le protocole LDP (standard de l'IETF) ne soit pas encore supporté, la commande Cisco « mpls ip » (correspondant à LDP) existe dans la version 12.1(5)T. Toutefois, cette commande Cisco a le même effet que « tag-switching ip ».

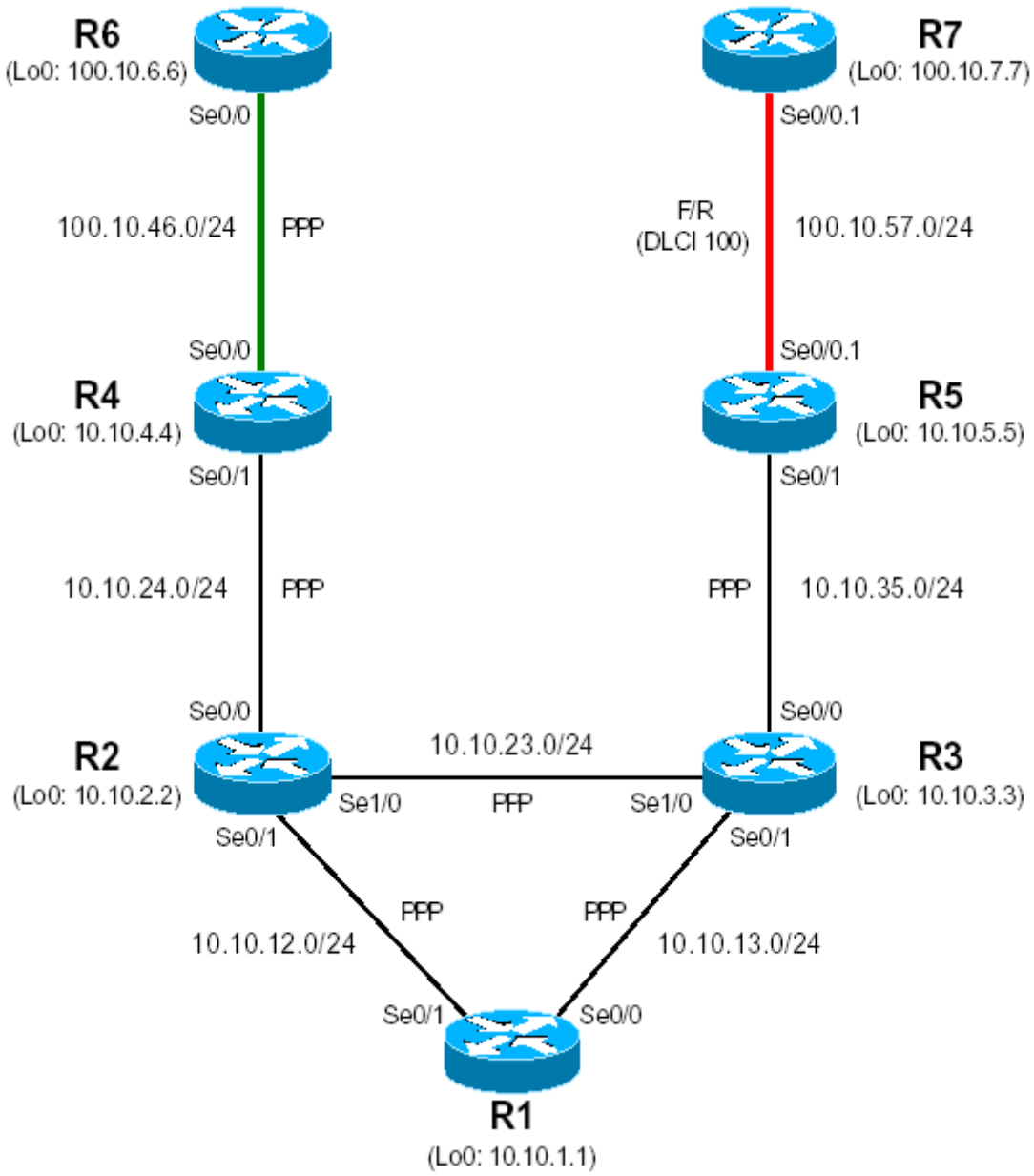
4 – Virtual Private Networks (VPN)

Actuellement, il est très courant qu'une entreprise soit constituée de plusieurs sites géographiques (parfois très éloignés) et dont elle souhaite interconnecter les réseaux informatiques à travers un WAN (Wide Area Network). La solution la plus connue et la plus employée consiste à relier les sites au moyen de liaisons spécialisées, dédiées à l'entreprise. Toutefois, le coût prohibitif de ces liaisons, et éventuellement la non aisabilité technique, par exemple avec des sites séparés de plusieurs centaines de km, amènent à rechercher des solutions plus abordables. Les fournisseurs d'accès Internet disposent de backbones étendus, et couvrant la plupart du temps une large portion de territoire. Il est donc plus simple pour une entreprise de relier ses sites aux points de présence (POP) de l'opérateur et mettre en place une solution VPN (Virtual Private Networks) (http://www.frameip.com/vpn/).

MPLS/VPN fournit une méthode de raccordement de sites appartenant à un ou plusieurs VPN, avec possibilité de recouvrement des plans d'adressage IP pour des VPN différents. En effet, l'adressage IP privé (voir RFC 1918) est très employé aujourd'hui, et rien ne s'oppose à ce que plusieurs entreprises utilisent les mêmes plages d'adresses (par exemple 172.16.1.0/24). MPLS/VPN permet d'isoler le trafic entre sites n'appartenant pas au même VPN, et en étant totalement transparent pour ces sites entre eux. Dans l'optique MPLS/VPN, un VPN est un ensemble de sites placés sous la même autorité administrative, ou groupés suivant un intérêt particulier. Cette partie aborde les concepts de MPLS/VPN, en particulier avec les notions de routeurs virtuels (VRF) et le protocole MP-BGP, dédié à l'échange de routes VPN. Dans les documents présentant MPLS/VPN, les VPN sont généralement définis avec des noms de couleurs (red, blue, etc.). Cette convention sera conservée dans ce rapport.

4.1 – Réseau Cisco de démonstration

Dans cette partie, le réseau de démonstration est constitué du pod L10, avec R6 et R7 comme placés en tant que routeurs clients. Afin de séparer le plan d'adressage du backbone MPLS (adresses en 10.x.y.z), les adresses utilisées par les VPN sont du type 100.x.y.z, avec les mêmes conventions que précédemment. Par exemple, l'interface Loopback0 du routeur Cisco L10-R6 sera 100.10.6.6.



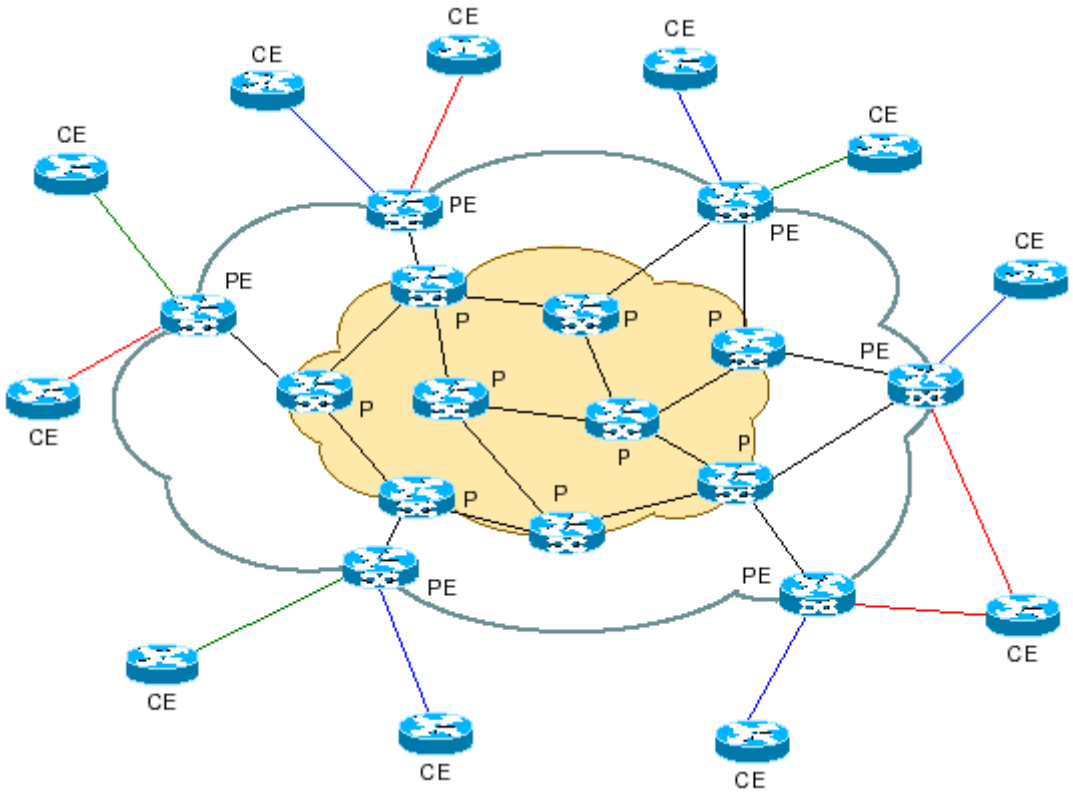
Les routeurs R6 et R7 sont placés respectivement dans les VPN « GREEN » et « RED ». Les routeurs R4 et R5 ont chacun trois interfaces Loopback placées dans les VRF « GREEN », « RED » et « BLUE ». Au niveau du backbone MPLS, les routeurs R1 à R5 emploient OSPF (aire 0) comme protocole de routage interne.

4.2 – Routeurs P, PE et CE

Une terminologie particulière est employée pour désigner les routeurs Cisco (en fonction de leur rôle) dans un environnement MPLS / VPN :

- P (Provider) : ces routeurs, composant le coeur du backbone MPLS, n'ont aucune connaissance de la notion de VPN. Ils se contentent d'acheminer les données grâce à la commutation de labels ;
- PE (Provider Edge) : ces routeurs sont situés à la frontière du backbone MPLS et ont par définition une ou plusieurs interfaces reliées à des routeurs clients ;
- CE (Customer Edge) : ces routeurs appartiennent au client et n'ont aucune connaissance des VPN ou même de la notion de label. Tout routeur Cisco « traditionnel » peut être un routeur CE, quelle que soit son type ou la version d'IOS utilisée.

Le schéma ci-dessous montre l'emplacement de ces routeurs dans une architecture MPLS :



4.3 – Routeurs Cisco virtuels : VRF

La notion même de VPN implique l'isolation du trafic entre sites clients n'appartenant pas aux mêmes VPN. Pour réaliser cette séparation, les routeurs PE ont la capacité de gérer plusieurs tables de routage grâce à la notion de VRF (VPN Routing and Forwarding). Une VRF est constituée d'une table de routage, d'une FIB (Forwarding Information Base) et d'une table CEF spécifiques, indépendantes des autres VRF et de la table de routage globale. Chaque VRF est désignée par un nom (par ex. RED, GREEN, etc.) sur les routeurs PE. Les noms sont affectés localement, et n'ont aucune signification vis-à-vis des autres routeurs. Chaque interface de PE reliée à un site client est rattachée à une VRF particulière. Lors de la réception de paquets IP sur une interfaces client, le routeur Cisco PE procède à un examen de la table de routage de la VRF à laquelle est rattachée l'interface, et donc ne consulte pas sa table de routage globale. Cette possibilité d'utiliser plusieurs tables de routage indépendantes permet de gérer un plan d'adressage par sites, même en cas de recouvrement d'adresses entre VPN différents.

Par exemple, L10-R4 est configuré de la manière suivante pour ses interfaces Loopback :

```
interface Loopback1
    ip vrf forwarding BLUE
    ip address 100.10.4.4 255.255.255.255
 !
 interface Loopback2
    ip vrf forwarding RED
    ip address 100.10.4.4 255.255.255.255
 !
 interface Loopback3
    ip vrf forwarding GREEN
    ip address 100.10.4.4 255.255.255.255
 !
```

La commande Cisco « ip vrf forwarding <vrf> » permet de placer une interface dans la VRF spécifiée. Comme le montre l'exemple ci-dessus, la même adresse IP peut être affectée plusieurs fois à différentes interfaces, car celles-ci sont placées dans des VRF différentes.

Pour construire leurs tables VRF, les PE doivent s'échanger les routes correspondant aux différents VPN. En effet, pour router convenablement les paquets destinés à un PE nommé PE-1, relié au site CE-1, le routeur Cisco PE-2 doit connaître les routes VPN de PE- 1. L'échange des routes VPN s'effectue grâce au protocole MP-BGP, décrit dans le paragraphe suivant. Les configurations des VRF ne comportant que des paramètres relatifs à MP-BGP (notamment pour l'export et l'import des routes), la syntaxe IOS et des exemples pratiques seront donc donnés dans les paragraphes suivants. Les VRF disposant de

tables de routage et de tables CEF spécifiques, il est possible de les consulter grâce à une extension des commandes classiques. Par exemple, pour consulter la table de routage de la VRF « RED » sur L10-R4, il suffit d'employer la commande Cisco « show ip route vrf <vrf> » :

```
L10-R4# sh ip route vrf RED
 Gateway of last resort is not set
     100.0.0.0/8 is variably subnetted, 6 subnets, 2 masks
 B 100.10.57.0/24 [200/0] via 10.10.5.5, 01:10:46
 B 100.10.7.7/32 [200/0] via 10.10.5.5, 01:10:46
 B 100.10.5.5/32 [200/0] via 10.10.5.5, 01:10:46
 C 100.10.4.4/32 is directly connected, Loopback2
 B 100.10.172.0/24 [200/0] via 10.10.5.5, 01:10:46
 B 100.10.171.0/24 [200/0] via 10.10.5.5, 01:10:46
```

Si l'on examine la table de routage globale de L10-R4, on constate qu'il s'agit bien d'une table complètement différente et indépendante :

```
L10-R4# sh ip route
 Gateway of last resort is not set
     10.0.0.0/8 is variably subnetted, 11 subnets, 2 masks
 O 10.10.5.5/32 [110/2401] via 10.10.24.2, 01:04:51, Serial0/1
 O 10.10.3.3/32 [110/1601] via 10.10.24.2, 01:04:51, Serial0/1
 O 10.10.1.1/32 [110/1601] via 10.10.24.2, 01:04:51, Serial0/1
 O 10.10.2.2/32 [110/801] via 10.10.24.2, 01:04:51, Serial0/1
 C 10.10.4.4/32 is directly connected, Loopback0
 O 10.10.12.0/24 [110/1600] via 10.10.24.2, 01:04:51, Serial0/1
 O 10.10.13.0/24 [110/2400] via 10.10.24.2, 01:04:51, Serial0/1
 O 10.10.23.0/24 [110/1600] via 10.10.24.2, 01:04:51, Serial0/1
 C 10.10.24.0/24 is directly connected, Serial0/1
 C 10.10.24.2/32 is directly connected, Serial0/1
 O 10.10.35.0/24 [110/2400] via 10.10.24.2, 01:04:51, Serial0/1
```

La table CEF d'une VRF peut également être examinée, au moyen de la commande Cisco « show ip cef vrf <vrf> <subnet> » :

```
L10-R4# sh ip cef vrf RED 100.10.7.7
 100.10.7.7/32, version 24, cached adjacency to Serial0/1
 0 packets, 0 bytes
     tag information set
         local tag: VPN-route-head
         fast tag rewrite with Se0/1, point2point, tags imposed: {27 30}
     via 10.10.5.5, 0 dependencies, recursive
         next hop 10.10.24.2, Serial0/1 via 10.10.5.5/32
         valid cached adjacency
         tag rewrite with Se0/1, point2point, tags imposed: {27 30}
```

La table CEF permet donc de déterminer le Next-Hop, l'interface de sortie et les labels utilisés pour atteindre un subnet particulier.

4.4 – Multiprotocol BGP (MP-BGP)

Le protocole MP-BGP est une extension du protocole BGP 4, et permettant d'échanger des routes Multicast et des routes VPNv4. MP-BGP adopte une terminologie similaire à BGP concernant le peering :

- MP- BGP : peering entre routeurs d'un même AS ;
- MP-eBGP : peering entre routeurs situés dans 2 AS différents.

4.4.1 – Notion de RD (Route Distinguisher)

Des sites appartenant à des VPN isolés ayant la possibilité d'utiliser des plans d'adressage recouvrants, les routes échangées entre PE doivent être rendues uniques au niveau des updates BGP. Pour cela, un identifiant appelé RD (Route Distinguisher), codé sur 64 bits, est accolé à chaque subnet IPv4 d'une VRF donnée. Le RD s'écrit sous la forme « ASN:nn » ou « IP-Address:nn ». Dans les exemples de configuration fournis avec ce document, le paramètre ASN a été fixé arbitrairement à 100, et « nn »

choisi en fonction de la VRF, quel que soit le routeur. Il est toutefois conseillé de choisir un RD unique par routeur et par VRF. Une route VPNv4, formé d'un RD et d'un préfixe IPv4, s'écrit ainsi sous la forme RD:Subnet/Masque. Exemple : 100:1:100.10.5.5/32. Lors de la création d'une VRF sur un PE, un RD doit être configuré. Les routes apprises soit localement (routes statiques, Loopback sur le PE), soit par les CE rattachés au PE seront ainsi exportées dans les updates MP-BGP avec ce RD. Les RD affectés aux différentes VRF existantes sur un PE peuvent être consultés au moyen de la commande Cisco « show ip vrf » :

```
L10-R4# sh ip vrf
  Name Default RD Interfaces
  BLUE 100:1 Loopback1
  GREEN 100:3 Serial0/0
                Loopback3
  RED 100:2 Loopback2
```

On constate que les interfaces connectées aux VRF sont également listées par cette commande.

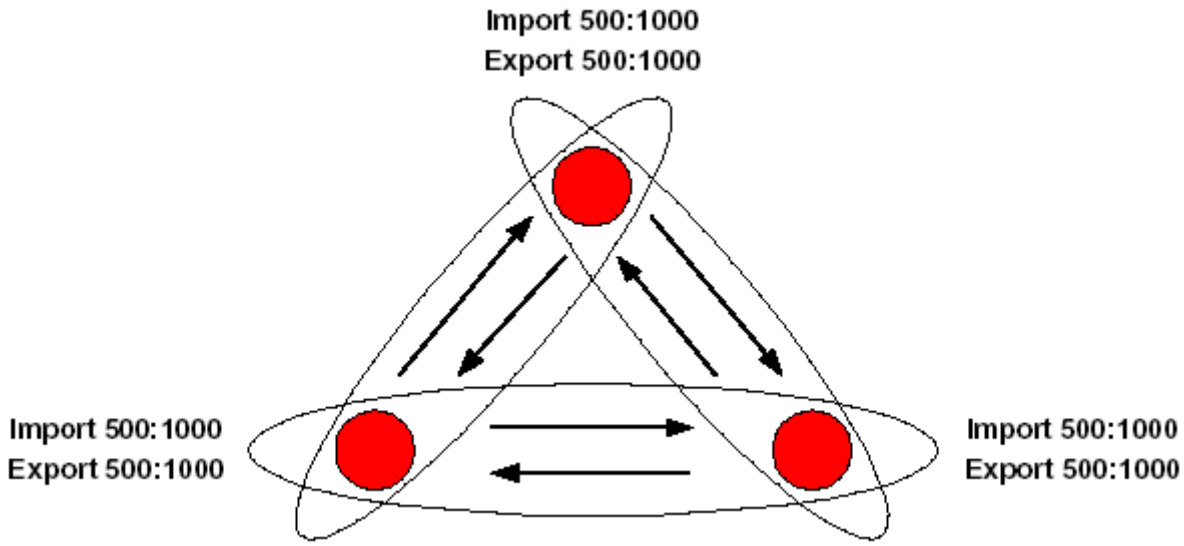
4.4.2 – Notion de RT (Route Target)

Le RD permet de garantir l'unicité des routes VPNv4 échangées entre PE, mais ne définit pas la manière dont les routes vont être insérées dans les VRF des routeurs Cisco PE. L'import et l'export de routes sont gérés grâce à une communauté étendue BGP (extended community) appelée RT (Route Target). Les RT ne sont rien de plus que des sortes de filtres appliqués sur les routes VPNv4. Chaque VRF définie sur un PE est configurée pour exporter ses routes suivant un certain nombre de RT. Une route VPN exportée avec un RT donné sera ajoutée dans les VRF des autres PE important ce RT. Par exemple, si la route VPN 2000:1:192.168.1.0/24 est exportée par un routeur PE avec comme liste de RT 2000:500 et 2000:501, tous les autres routeurs PE ayant une ou plusieurs VRF important au minimum un de ces deux RT ajouteront cette route dans leurs VRF concernées.

La configuration simple pour un VPN consiste à appliquer la règle :

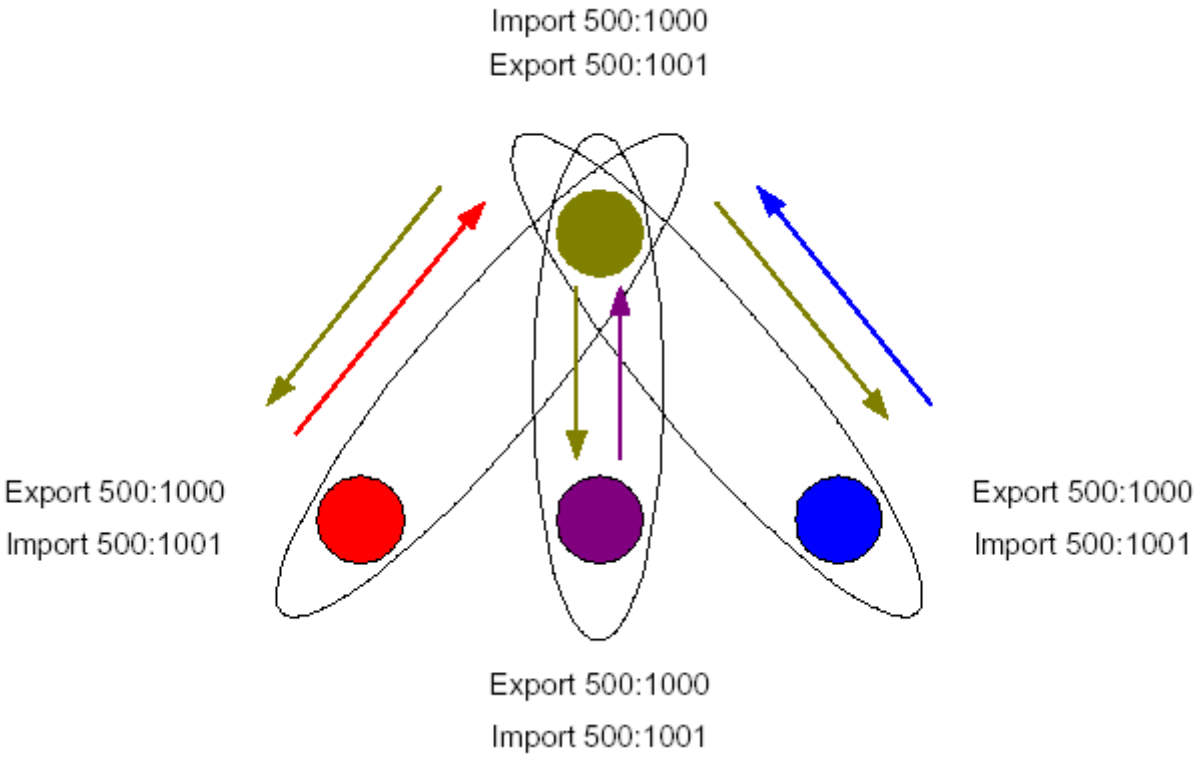
- RT_import = RT_export = RT_VPN

(Avec RT_VPN choisi comme identifiant spécifique au VPN).



Sur la figure ci-dessus, les 3 sites rouges appartiennent au même VPN. Pour échanger les routes entre tous les sites, chaque PE importe et exporte le RT 500:1000.

Le schéma suivant indique la marche à suivre pour créer une topologie de type « hub and spoke » :



Dans ce exemple, le site vert est un site central (par ex. pour l'administration des différents VPN). Chacun des 3 sites, appartenant à un VPN différent (Rouge, Violet et Bleu) importe les routes du site central (RT 500:1001). Réciproquement, le site central importe les routes de tous les sites clients (RT 500:1000). Bien que le site central ait accès à tous les sites clients, ceux-ci ne peuvent se voir entre eux. En effet, aucune relation de RT n'est définie entre les sites clients (aucun site n'importe ou n'exporte de route vers un autre).

Naturellement, comme le site vert « voit » les 3 sites clients, le plan d'adressage de ces sites doit être compatible (c'est-à-dire non recouvrant) au niveau des routes échangées pour garantir l'unicité des routes vis-à-vis du site central.

4.4.3 – Configuration d'une VRF

Les VRF sont configurées sur les routeurs Cisco PE avec les paramètres suivants :

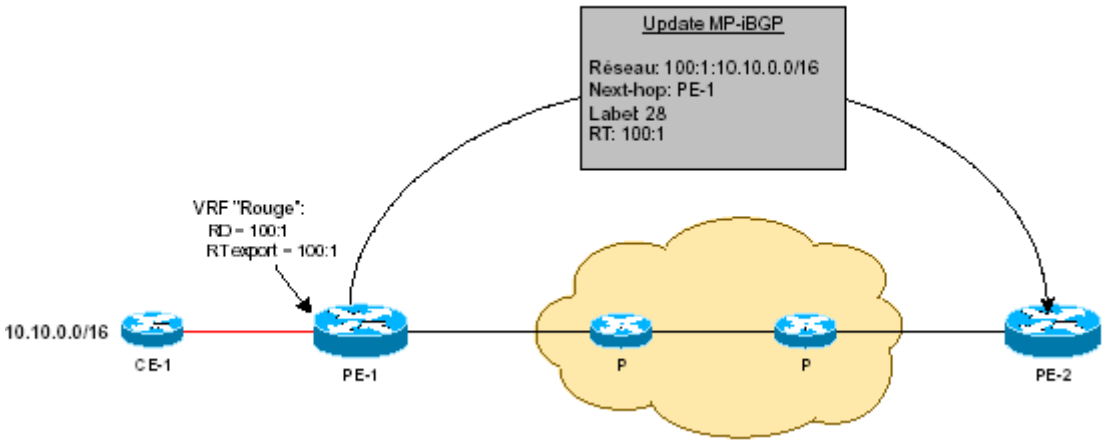
- Nom de VRF (case-sensitive) ;
- RD (Route Distinguisher) ;
- RT exportés ;
- RT importés ;
- Filtres sur l'import et l'export des routes (optionnel).

Le paramétrage d'une VRF « TEST » (avec un RD de 500:1000), exportant ses routes avec les RT 500:1 et 500:2 et important les routes avec les RT 500:1 et 500:3, serait le suivant :

```
ip vrf TEST
    rd 500:1000
    route-target export 500:1
    route-target export 500:2
    route-target import 500:1
    route-target import 500:3
 !
```

4.4.4 – Updates MP-BGP

En plus du RD et des RT, les updates MP-BGP contiennent d'autres informations, telles que le Site d'Origine (SOO), l'adresse IP du PE annonçant la route (PE nexthop) et le label VPN affecté par ce PE.



4.4.5 – Configuration MP-BGP d'un PE

La configuration d'un routeur Cisco PE pour échanger des routes VPNv4 se présente sous la forme suivante :

```
router bgp 10
    no synchronization
    bgp log-neighbor-changes
    neighbor 10.10.1.1 remote-as 10
    neighbor 10.10.1.1 update-source Loopback0
    no neighbor 10.10.1.1 activate
    no auto-summary
    !
    address-family vpnv4
    neighbor 10.10.1.1 activate
    neighbor 10.10.1.1 send-community extended
    no auto-summary
    exit-address-family
 !
```

On remarque la présence d'une section supplémentaire par rapport à une configuration BGP traditionnelle, introduite par la commande Cisco « address-family vpnv4 ». Cette partie de la configuration BGP contient tous les voisins tournant MP-BGP. Pour pouvoir ajouter un voisin dans la configuration VPNv4, ce voisin doit être préalablement déclaré dans la configuration globale de BGP (commande Cisco « remote- s » et autres). Pour éviter qu'un voisin ne soit actif à la fois pour BGP et MP-BGP, la ligne « no neighbor <neighbor> activate » doit être insérée globalement. Naturellement, il est tout à fait possible pour un routeur d'être actif pour BGP et MP-BGP simultanément. Par exemple, le BGP traditionnel servira à propager les routes Internet aux routeurs PE, tandis que MP-BGP servira à la propagation des routes VPN.

Pour propager les Route-Target (RT), qui définissent l'appartenance aux VPN et qui sont des communautés étendues BGP, la ligne « neighbor <neighbor> sendcommunity extended » doit être utilisée.

Dans le réseau de démonstration MPLS/VPN, le routeur Cisco L10-R1 fait office de Route Reflector BGP. Sa configuration Cisco est la suivante :

```
router bgp 10
    no synchronization
    no bgp default route-target filter
    bgp log-neighbor-changes
    bgp cluster-id 10
    neighbor iBGP peer-group
    no neighbor iBGP activate
    neighbor iBGP remote-as 10
    neighbor iBGP update-source Loopback0
    neighbor iBGP soft-reconfiguration inbound
    no auto-summary
!
    address-family vpnv4
    neighbor iBGP activate
    neighbor iBGP route-reflector-client
    neighbor iBGP send-community extended
    neighbor 10.10.2.2 peer-group iBGP
    neighbor 10.10.3.3 peer-group iBGP
    neighbor 10.10.4.4 peer-group iBGP
    neighbor 10.10.5.5 peer-group iBGP
    no auto-summary
    exit-address-family
!
```

Afin de faciliter l'ajout de nouveaux voisins, la notion de « peer-group » a été utilisée. Un peer group est défini par un nom, et il est possible de fixer certaines propriétés pour ce groupe : AS distant (remote-as), interface source pour les updates (update-source), etc. Chaque voisin est ensuite ajouté dans ce groupe avec un commande Cisco du type: « neighbor 10.10.5.5 peer-group iBGP ». Dans cet exemple, toutes les commandes appliquées au groupe « iBGP » le seront pour le voisin 10.10.5.5.

4.4.6 – Vérification du fonctionnement de MP-BGP

Plusieurs commandes existent pour s'assurer du bon fonctionnement de BGP sur les routeurs. Par exemple, pour connaître toutes les routes apprises par MP-BGP sur un routeur Cisco donné, la commande Cisco « show ip bgp vpnv4 all » peut être utilisée :

```
L10-R4# sh ip bgp vpnv4 all
 BGP table version is 129, local router ID is 10.10.4.4
 Status codes: s suppressed, d damped, h history, * valid, > best, i -
 internal
 Origin codes: i - IGP, e - EGP, ? - incomplete

    Network Next Hop Metric LocPrf Weight Path
 Route Distinguisher: 100:1 (default for vrf BLUE)
 *> 100.10.4.4/32 0.0.0.0 0 32768 ?
 *>i100.10.5.5/32 10.10.5.5 0 100 0 ?
 Route Distinguisher: 100:2 (default for vrf RED)
 *> 100.10.4.4/32 0.0.0.0 0 32768 ?
 *>i100.10.5.5/32 10.10.5.5 0 100 0 ?
 *>i100.10.7.7/32 10.10.5.5 0 100 0 102 i
 *>i100.10.57.0/24 10.10.5.5 0 100 0 ?
 *>i100.10.171.0/24 10.10.5.5 0 100 0 102 i
 *>i100.10.172.0/24 10.10.5.5 0 100 0 102 i
 Route Distinguisher: 100:3 (default for vrf GREEN)
 *>i100.10.3.3/32 10.10.3.3 0 100 0 ?
 *> 100.10.4.4/32 0.0.0.0 0 32768 ?
 *>i100.10.5.5/32 10.10.5.5 0 100 0 ?
 *> 100.10.6.6/32 100.10.46.6 1 32768 ?
 *> 100.10.46.0/24 0.0.0.0 0 32768 ?
 *> 100.10.46.6/32 0.0.0.0 0 32768 ?
 *> 100.10.161.0/24 100.10.46.6 1 32768 ?
 *> 100.10.162.0/24 100.10.46.6 1 32768 ?
 Route Distinguisher: 100:2000
 *>i100.10.3.3/32 10.10.3.3 0 100 0 ?
```

Si l'on souhaite se restreindre à une VRF donnée, la commande Cisco « show ip bgp vpnv4 vrf <vrf> » peut être employée :

```
L10-R4# sh ip bgp vpnv4 vrf RED
 BGP table version is 129, local router ID is 10.10.4.4
 Status codes: s suppressed, d damped, h history, * valid, > best, i -
 internal
 Origin codes: i - IGP, e - EGP, ? - incomplete

    Network Next Hop Metric LocPrf Weight Path
 Route Distinguisher: 100:2 (default for vrf RED)
 *> 100.10.4.4/32 0.0.0.0 0 32768 ?
 *>i100.10.5.5/32 10.10.5.5 0 100 0 ?
 *>i100.10.7.7/32 10.10.5.5 0 100 0 102 i
 *>i100.10.57.0/24 10.10.5.5 0 100 0 ?
 *>i100.10.171.0/24 10.10.5.5 0 100 0 102 i
 *>i100.10.172.0/24 10.10.5.5 0 100 0 102 i
```

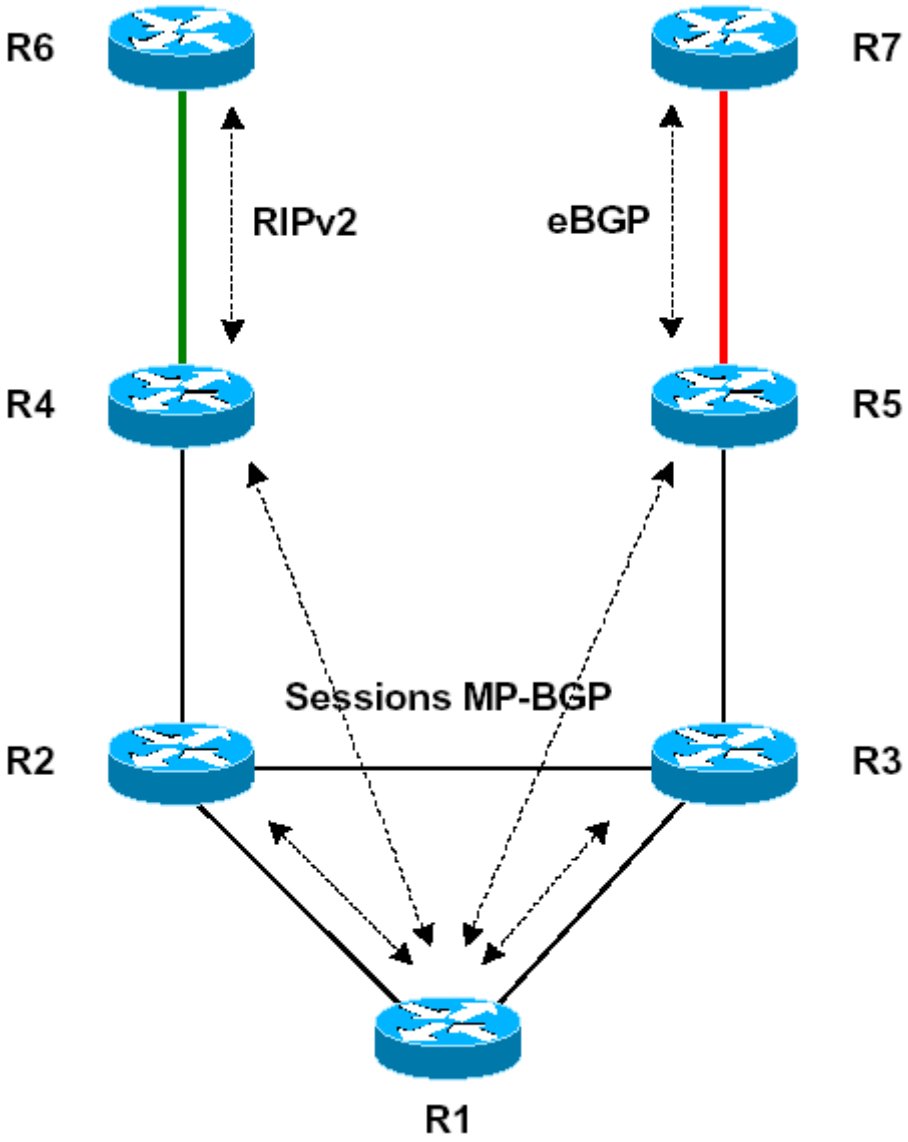
Les labels fournis dans les updates MP-BGP peuvent être affichés au moyen de la commande Cisco « sh ip bgp vpnv4 [vrf <vrf> | all] tags » :

```
L10-R4# sh ip bgp vpnv4 vrf RED tags
    Network Next Hop In tag/Out tag
 Route Distinguisher: 100:2 (RED)
    100.10.4.4/32 0.0.0.0 28/aggregate(RED)
    100.10.5.5/32 10.10.5.5 notag/26
    100.10.7.7/32 10.10.5.5 notag/29
    100.10.57.0/24 10.10.5.5 notag/28
    100.10.171.0/24 10.10.5.5 notag/33
    100.10.172.0/24 10.10.5.5 notag/34
```

4.5 – Echange des routes avec les CE

Les CE sont des routeurs Cisco clients traditionnels, n'ayant aucune connaissance de MPLS ou des VRF. Les CE doivent donc échanger leurs routes IP avec leur PE au moyen de protocoles de routages classiques. Les protocoles supportés par IOS sont eBGP (external BGP), RIPv2 et OSPF.

Les deux paragraphes suivants donnent des exemples de configuration avec eBGP et RIPv2. Dans le réseau de démonstration, eBGP est utilisé entre L10-R5 et L10-R7, tandis que RIPv2 est utilisé entre L10-R4 et L10-R6. Le routeur Cisco L10-R7 a été placé dans le VPN « RED », et le routeur Cisco L10-R6 dans le VPN « GREEN ».



4.5.1 – Configuration eBGP

La configuration BGP de L10-R5 (routeur PE) pour la VRF « RED » est listée ci-dessous :

```
router bgp 10
 !
    [...]
    address-family ipv4 vrf RED
    redistribute connected
    neighbor 100.10.57.7 remote-as 102
    neighbor 100.10.57.7 activate
    no auto-summary
    no synchronization
    exit-address-family
    [...]
 !
```

L'interface reliant L10-R5 et L10-R7 étant paramétrée comme suit (sur L10-R5) :

```
interface Serial0/0.1 point-to-point
    description Vers L10-R7
    bandwidth 125
    ip vrf forwarding RED
    ip address 100.10.57.5 255.255.255.0
    frame-relay interface-dlci 100
```

Chaque voisin devant être actif en eBGP dans une VRF donné doit donc être configuré dans la section « address-family ipv4 vrf <vrf> » correspondante. A titre indicatif, la configuration BGP de L10-R7 est fournie ci-dessous :

```
router bgp 102
    bgp log-neighbor-changes
    network 100.10.7.7 mask 255.255.255.255
    network 100.10.171.0 mask 255.255.255.0
    network 100.10.172.0 mask 255.255.255.0
    neighbor 100.10.57.5 remote-as 10
 !
```

On constate donc que la configuration du routeur CE est tout à fait classique, sans présence de VRF ou de toute autre notion de VPN.

4.5.2 – Configuration RIPv2

La configuration Cisco RIPv2 avec un routeur CE suit le même principe qu'une configuration eBGP, mais les routes apprises par RIP doivent être réinjectées dans MP-BGP et réciproquement :

```
router rip
    version 2
    !
    address-family ipv4 vrf GREEN
    version 2
    redistribute bgp 10 metric 1
    network 100.0.0.0
    no auto-summary
    exit-address-family
 !
 [...]

 router bgp 10
 [...]
    address-family ipv4 vrf GREEN
    redistribute connected
    redistribute rip
    no auto-summary
    no synchronization
    exit-address-family
    !
 [...]
```

L'interface reliant L10-R4 et L10-R6 est paramétrée de la manière suivante (sur L10-R4) :

```
interface Serial0/0
    description Vers L10-R6
    bandwidth 125
    ip vrf forwarding GREEN
    ip address 100.10.46.4 255.255.255.0
    encapsulation ppp
    no fair-queue
    clockrate 125000
 end
```

4.6 – Transmission des paquets IP

La transmission des paquets IP provenant des CE sur le backbone MPLS emploie la notion de label stacking. Pour atteindre un site donné, le PE source encapsule deux labels : le premier sert à atteindre le PE de destination, tandis que le second détermine l'interface de sortie sur le PE, à laquelle est reliée le CE. Le second label est appris grâce aux updates MP-BGP. Les tables CEF des routeurs peuvent être consultées pour déterminer les labels utilisées. Par exemple, supposons que l'on souhaite connaître les tags employés pour atteindre l'adresse de Loopback 100.10.5.5 configurée sur le routeur Cisco L10-R5, et placée dans la VRF « RED ». La consultation de la table de routage de la VRF « RED » sur L10-R4 montre que le next-hop est bien L10-R5 (10.10.5.5) :

```
L10-R4# sh ip route vrf RED

 Gateway of last resort is not set

     100.0.0.0/8 is variably subnetted, 6 subnets, 2 masks
 B 100.10.57.0/24 [200/0] via 10.10.5.5, 00:01:45
 B 100.10.7.7/32 [200/0] via 10.10.5.5, 00:01:46
 B 100.10.5.5/32 [200/0] via 10.10.5.5, 00:01:46
 C 100.10.4.4/32 is directly connected, Loopback2
 B 100.10.172.0/24 [200/0] via 10.10.5.5, 00:01:46
 B 100.10.171.0/24 [200/0] via 10.10.5.5, 00:01:46
```

Le label utilisé pour atteindre L10-R5 est déterminé grâce à la table CEF globale du routeur :

```
L10-R4# sh ip cef 10.10.5.5
 10.10.5.5/32, version 41, cached adjacency to Serial0/1
 0 packets, 0 bytes
    tag information set
       local tag: 17
       fast tag rewrite with Se0/1, point2point, tags imposed: {27}
    via 10.10.24.2, Serial0/1, 7 dependencies
       next hop 10.10.24.2, Serial0/1
       valid cached adjacency
       tag rewrite with Se0/1, point2point, tags imposed: {27}
```

L10-R4 imposera donc le label 27 pour atteindre L10-R5, le prochain saut étant le routeur Cisco L10-R2 (10.10.24.2). Le deuxième label (dit label VPN), servant à sélectionner l'interface de sortie sur L10-R5, peut être déterminé grâce à la table CEF de la VRF « RED », indépendante de la table CEF globale :

```
L10-R4# sh ip cef vrf RED 100.10.5.5
 100.10.5.5/32, version 23, cached adjacency to Serial0/1
 0 packets, 0 bytes
    tag information set
       local tag: VPN-route-head
       fast tag rewrite with Se0/1, point2point, tags imposed: {27 26}
    via 10.10.5.5, 0 dependencies, recursive
       next hop 10.10.24.2, Serial0/1 via 10.10.5.5/32
       valid cached adjacency
       tag rewrite with Se0/1, point2point, tags imposed: {27 26}
```

Le label VPN est donc 26. Pour joindre l'adresse IP 100.10.5.5 de la VRF « RED », le routeur Cisco L10-R4 imposera donc la pile de label { 27 26 }. Sur le backbone MPLS, la commutation se fera uniquement en fonction du premier label, comme le montre le résultat de la commande Cisco traceroute :

```
L10-R4# trace vrf RED 100.10.5.5

    1 10.10.24.2 [MPLS: Labels 27/26 Exp 0] 72 msec 72 msec 68 msec
    2 10.10.23.3 [MPLS: Labels 23/26 Exp 0] 56 msec 60 msec 60 msec
    3 100.10.57.5 28 msec * 28 msec
```

On remarque que seul le premier label a été modifié, le label VPN ayant été conservé intact pendant tout le cheminement sur le backbone. La copie d'écran suivante montre quel aurait été le résultat de la commande Cisco traceroute pour atteindre l'adresse de Loopback 10.10.5.5 (adresse globale) à partir de L10-R4 :

```
L10-R4# trace 10.10.5.5

    1 10.10.24.2 [MPLS: Label 27 Exp 0] 68 msec 68 msec 64 msec
    2 10.10.23.3 [MPLS: Label 23 Exp 0] 52 msec 56 msec 56 msec
    3 10.10.35.5 28 msec * 24 msec
```

Il apparaît clairement que la présence du label VPN n'a pas d'effet sur les routeurs P du backbone : ceux-ci switchent les paquets entre routeurs PE et n'ont aucune information sur les labels VPN.

Rappelons que les labels VPN, appris par MP-BGP, peuvent être affichés au moyen de la commande Cisco « show ip bgp vpnv4 vrf <vrf> tags » :

```
L10-R4# sh ip bgp vpnv4 vrf RED tags
     Network Next Hop In tag/Out tag
 Route Distinguisher: 100:2 (RED)
    100.10.4.4/32 0.0.0.0 28/aggregate(RED)
    100.10.5.5/32 10.10.5.5 notag/26
    100.10.7.7/32 10.10.5.5 notag/30
    100.10.57.0/24 10.10.5.5 notag/28
    100.10.171.0/24 10.10.5.5 notag/31
    100.10.172.0/24 10.10.5.5 notag/32
```

On retrouve bien le label 26 pour atteindre le subnet 100.10.5.5/32 sur le routeur Cisco L10-R5. Si l'on effectue un traceroute vers l'adresse 100.10.7.7, appartenant à L10-R7, on obtient le résultat suivant :

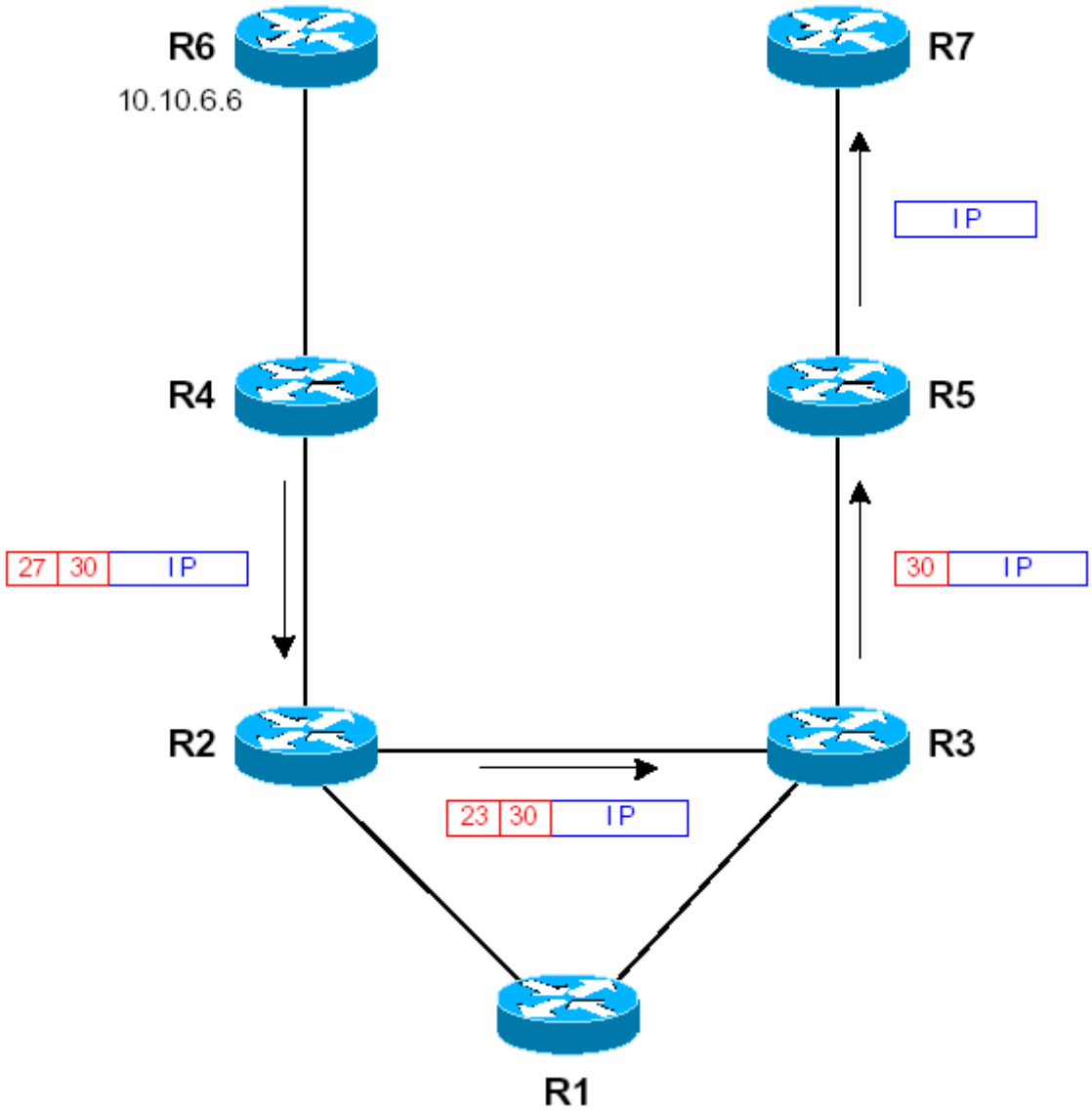
```
L10-R4# trace vrf RED 100.10.7.7

 Type escape sequence to abort.
 Tracing the route to 100.10.7.7

    1 10.10.24.2 [MPLS: Labels 27/30 Exp 0] 96 msec 92 msec 92 msec
    2 10.10.23.3 [MPLS: Labels 23/30 Exp 0] 84 msec 88 msec 84 msec
    3 100.10.57.5 [MPLS: Label 30 Exp 0] 76 msec 76 msec 72 msec
    4 100.10.57.7 36 msec * 36 msec
```

On constate la présence du Penultimate Hop Popping, entre les routeurs L10-R3 (10.10.24.3) et L10-R5. (100.10.57.5). En effet, L10-R3 a retiré le premier label 23, servant à atteindre L10-R5. Ce fonctionnement est confirmé en consultant la table TFIB de L10-R3 :

```
L10-R3# sh tag for
 Local Outgoing Prefix Bytes tag Outgoing Next Hop
 tag tag or VC or Tunnel Id switched interface
 16 Untagged 10.10.13.1/32 0 Se0/1 point2point
 17 Untagged 10.10.35.5/32 0 Se0/0 point2point
 18 Untagged 10.10.23.2/32 0 Se1/0 point2point
 20 Pop tag 10.10.1.1/32 1693 Se0/1 point2point
 21 Pop tag 10.10.2.2/32 0 Se1/0 point2point
 22 20 10.10.4.4/32 3020 Se1/0 point2point
 23 Pop tag 10.10.5.5/32 5821 Se0/0 point2point
 26 Pop tag 10.10.12.0/24 0 Se1/0 point2point
    Pop tag 10.10.12.0/24 0 Se0/1 point2point
 27 Pop tag 10.10.24.0/24 2304 Se1/0 point2point
 29 Aggregate 100.10.3.3/32[V] 0
```

```
L10-R3# sh ip cef 10.10.5.5
 10.10.5.5/32, version 34, cached adjacency to Serial0/0
 0 packets, 0 bytes
    tag information set
       local tag: 23
    via 10.10.35.5, Serial0/0, 0 dependencies
       next hop 10.10.35.5, Serial0/0
       valid cached adjacency
       tag rewrite with Se0/0, point2point, tags imposed: {}
```

Le schéma ci-dessous montre le trajet suivi par les paquets, de L10-R4 vers L10-R7 :



4.7 – Accès Internet

Le principe des VRF permet de concevoir aisément des routeurs virtuels, qui consultent leurs différentes tables de routage en fonction de l'interface d'entrée des paquets. Ces tables sont remplies avec les routes du VPN associé, et le backbone MPLS assure la transmission des paquets entre les routeurs PE. Il se pose alors le problème de l'accès à Internet, situé par définition à l'extérieur des différents VPN. De plus, les fournisseurs d'accès Internet disposant de plusieurs points de sortie, il est important que les sites clients puissent utiliser le meilleur chemin vers l'extérieur.

Différentes méthodes existent pour permettre un accès Internet. Les deux premières se situent dans la catégorie « Sub-Optimized Routing », du fait qu'elles ne permettent pas de sélectionner le meilleur chemin vers Internet. La dernière, nommée « Optimum Routing », permet de choisir le chemin optimal, tout en étant la plus « propre » techniquement.

4.7.1 – Route par défaut statique (Static Default Route)

La première méthode (la plus ancienne) consiste à une utiliser une extension de la commande Cisco « ip route » pour définir une route par défaut dans les VRF des routeurs PE, au moyen de la commande Cisco :

```
ip route vrf GREEN 0.0.0.0 0.0.0.0 PE-Internet global
```

où PE-Internet est l'adresse globale du PE fournissant l'accès à Internet. Pour que le retour des paquets puisse être effectif vers le CE concerné, les routes VPN du CE doivent être déclarées globalement sur son PE de rattachement, et propagées au PE-Internet (IGP, iBGP, etc.). Par exemple, si un réseau 120.2.1.0/24 est connecté à un CE 120.1.1.1 appartenant au VPN « GREEN », le PE doit contenir les deux lignes suivantes :

```
ip route vrf GREEN 0.0.0.0 0.0.0.0 120.1.1.2 global
 ip route 120.2.1.0 255.255.255.0 120.1.1.1
```

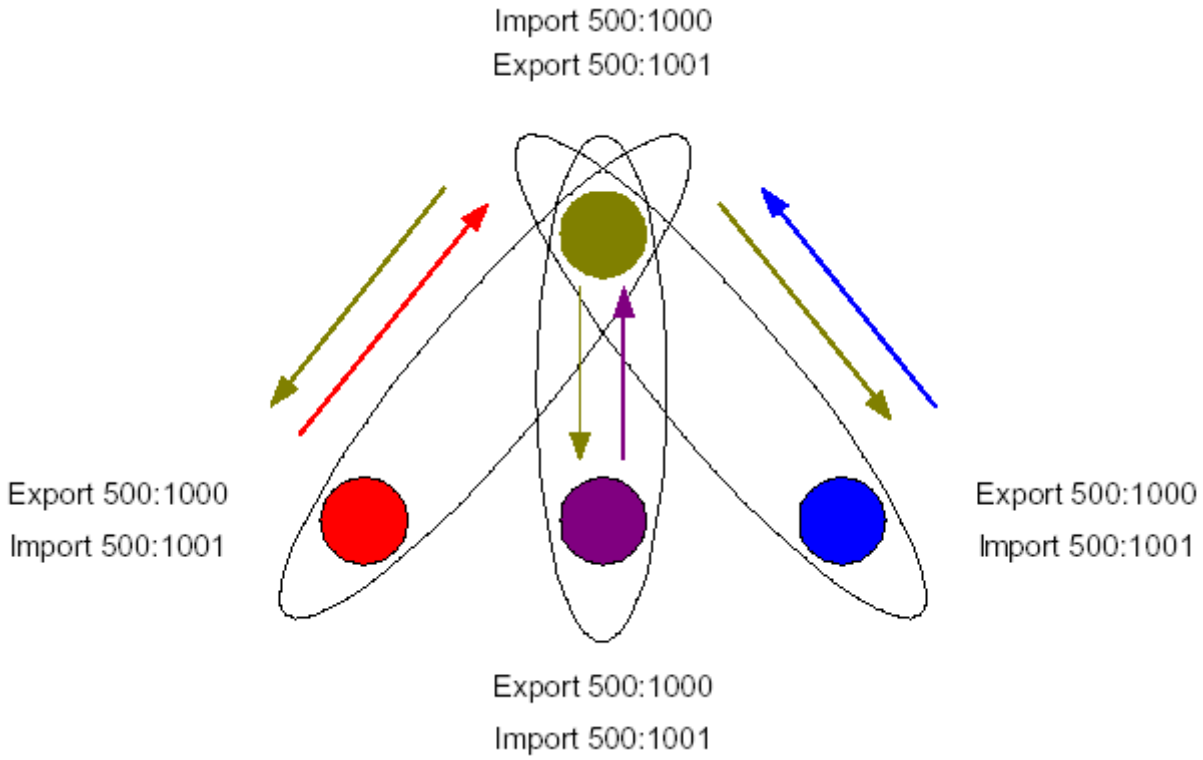
Dans cet exemple, on a supposé que PE-Internet a pour adresse 120.1.1.2.

Lorsque le PE recevra un paquet sur la VRF « GREEN », il effectuera un lookup dans la table de routage de cette VRF. Si aucune entrée n'est trouvée pour la destination IP, la route par défaut injectée au moyen de la commande Cisco « ip route » sera utilisée. Il est à noter que la table de routage globale du routeur est examinée pour atteindre PEInternet, et que les paquets traversent le backbone MPLS sans label VPN (seul le label de PE-Internet est accolé par le PE).

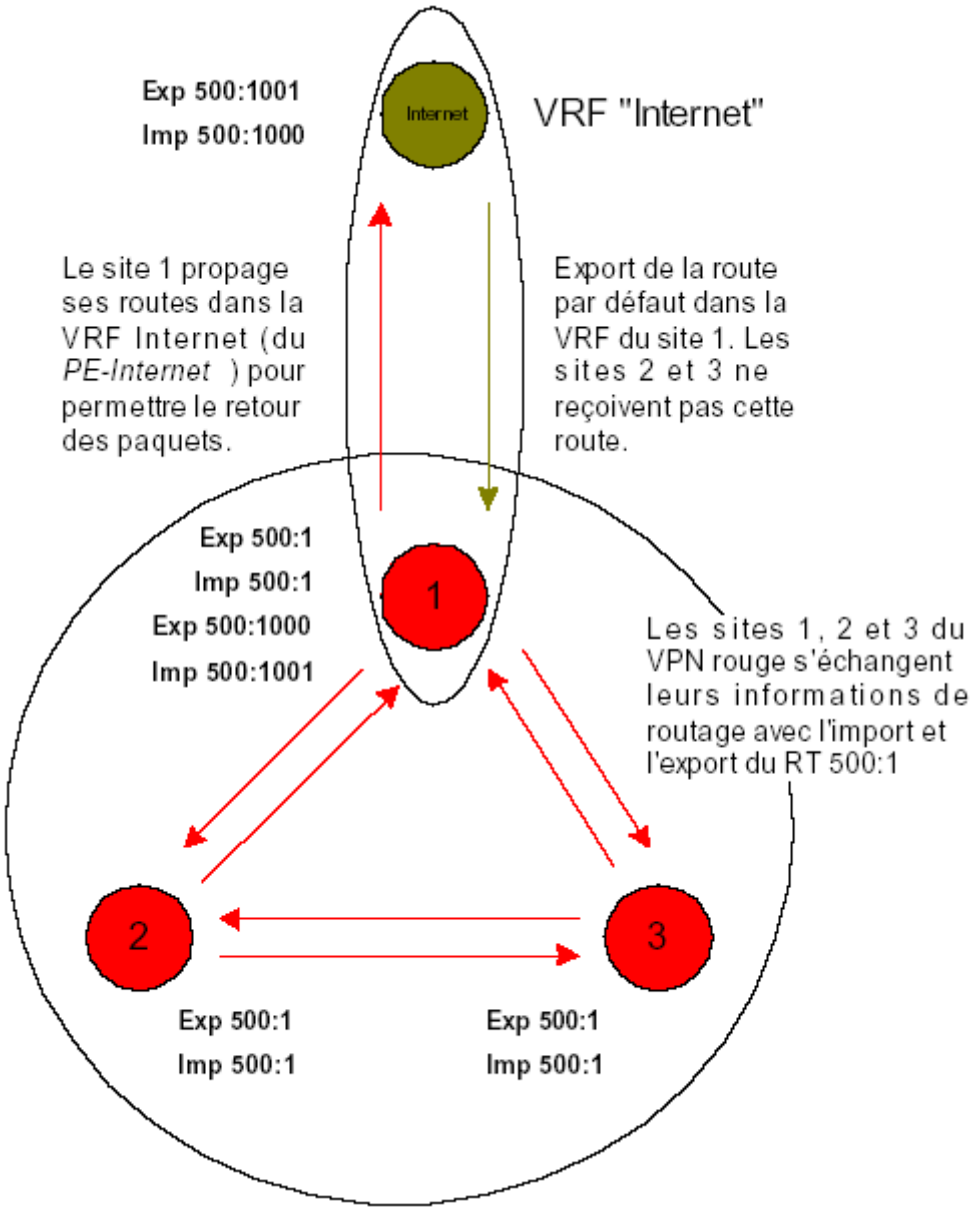
Naturellement, ce type de routage n'est pas optimal, car si plusieurs PE disposent d'un accès Internet, seul le PE déclaré dans la route par défaut sera employé. De plus, cette méthode « casse » la notion de VRF avec la déclaration des routes VPN de manière globale sur les PE. Enfin, tous les PE doivent être configurés de cette manière, avec pour chacun la route par défaut et la mise en place dans la table de routage globale des routes VPN.

4.7.2 – Route par défaut dynamique (Dynamic Default Route)

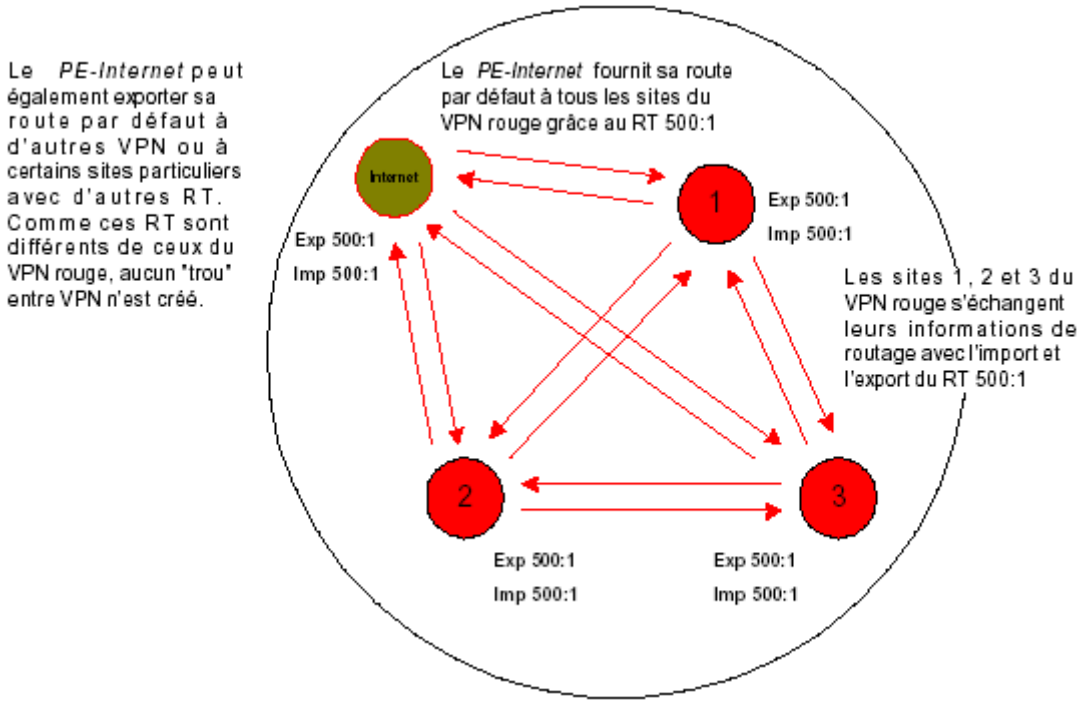
Une solution plus propre techniquement pour propager une route par défaut à tous les PE est d'utiliser la notion de VPN avec une topologie « Hub and Spoke ». Sur le routeur PE-Internet, une VRF particulière est configurée pour annoncer la route par défaut (apprise par un autre routeur, généralement avec eBGP). Si l'on souhaite propager manuellement cette route à un certains sites de différents VPN, il suffit d'employer la politique d'attribution des RT du schéma ci-dessous :



Chacun des sites clients reçoit la route par défaut provenant du site vert central grâce au RT 500:1001. Pour permettre le retour des paquets, chaque site doit exporter vers le site central ses propres routes (RT 500:1000). Chaque PE doit donc être configuré avec ces RT pour permettre la propagation de la route par défaut. Il est ainsi possible de ne propager la route par défaut qu'à certains sites d'un même VPN (les VRF de ces sites devant traiter les RT du VPN « Internet »), en configurant les PE adéquats et en ne changeant rien sur les autres :



Si l'on souhaite propager automatiquement la route par défaut à tous les sites d'un même VPN sans avoir à modifier la configuration des différents PE, il suffit d'importer et d'exporter le RT de ce VPN dans la VRF « Internet ». De cette manière, aucun changement dans la configuration des PE rattachés à ces sites n'est nécessaire.
Internet

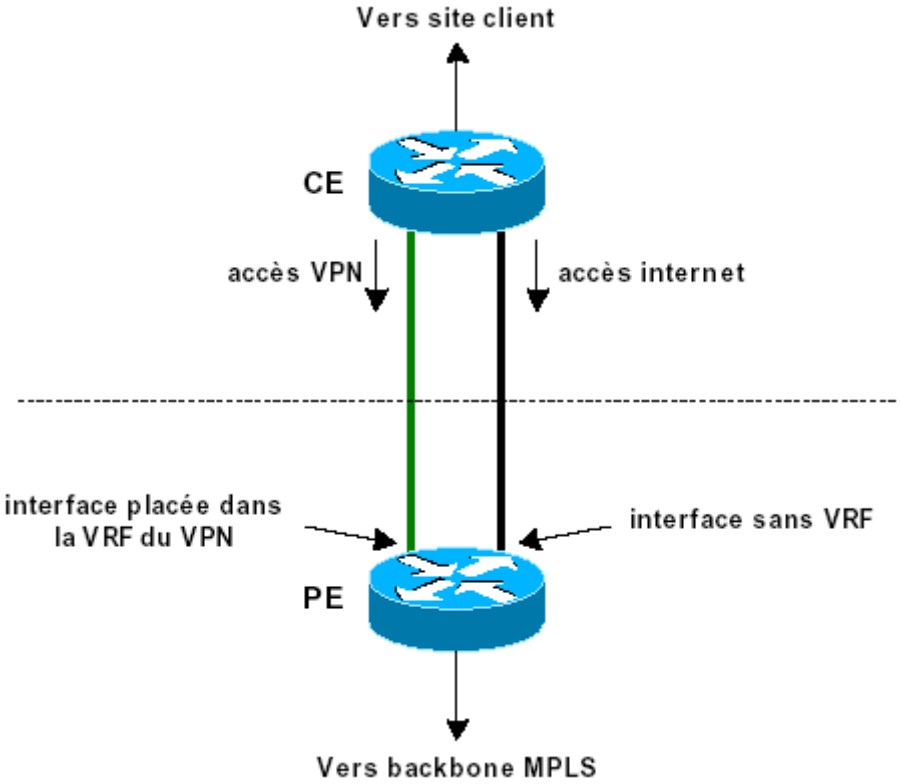


4.7.3 – Routage optimal (Optimum Routing)

La méthode dite « Optimum Routing » permet aux sites clients d'accéder à Internet, tout en sélectionnant le meilleur chemin si plusieurs PE sont passerelles vers l'extérieur.

Le concept clé de l'Optimum Routing est de propager l'ensemble des routes Internet sur tous les PE du backbone MPLS. Naturellement, il n'est pas possible de propager ces routes dans chaque VPN : en supposant que l'on compte 100000 routes Internet ; propager ces routes dans 100 VPN différents (soit 10 millions de routes au total) n'est évidemment pas réalisable. Les routes Internet sont donc échangées entre PE grâce au protocole BGP standard (sessions iBGP entre les PE), et ce sont les tables de routage globales des PE qui contiennent ces routes.

Pour permettre aux sites d'accéder à l'extérieur, les CE sont reliés de deux manières au backbone de l'opérateur : la première connexion permet l'accès aux routes Internet globales, tandis que la seconde permet l'accès aux autres sites du VPN, avec l'utilisation d'une VRF.



La mise en place d'une double liaison avec le CE (une avec VRF, l'autre sans) peut être réalisée au moyen de deux sous-interfaces (2 VLAN sur trunk Ethernet, 2 DLCI sur Frame Relay, 2 VC sur ATM, etc.) ou avec un tunnel GRE.

Exemple de configuration avec une interface Frame-Relay :

```
interface Serial1/0
    no ip address
    encapsulation frame-relay
!
interface Serial1/0.1
    description Interface pour accès Internet
    ip address 100.2.1.1 255.255.255.252
    frame-relay interface-dlci 100
!
interface Serial1/0.2
    description Interface pour accès VPN
    ip vrf forwarding RED
    ip address 10.2.1.1 255.255.255.252
    frame-relay interface-dlci 200
!
```

Exemple de configuration avec une interface Ethernet en mode trunk :

```
interface FastEthernet0/0
    no ip address
!
interface FastEthernet0/0.1
    description Interface pour accès Internet
    ip address 100.1.1.1 255.255.255.252
    encapsulation isl 1
!
interface FastEthernet0/0.2
    description Interface pour accès VPN
    ip vrf forwarding GREEN
    ip address 10.1.1.1 255.255.255.252
    encapsulation isl 2
!
```

Sur le CE, deux approches sont possibles pour accéder à la fois aux autres sites du VPN et à Internet. La première, la plus classique, consiste à sélectionner l'interface de sortie grâce au Policy Routing, au moyen des commandes route-map. La seconde consiste à employer la notion de VRF sur le routeur CE lui-même, mais sans utilisation de MP-BGP. Les VRF peuvent en effet être mises en place sur un routeur, indépendamment de MPLS/VPN. L'exemple suivant montre comment implémenter le Policy Routing sur un CE (en reprenant l'exemple du PE ci-dessus, connecté au CE par une interface FastEthernet) avec translation d'adresse (NAT) (http://www.frameip.com/nat/) :

```
interface FastEthernet0/0
    description Vers site client
    ip address 10.2.0.1 255.255.255.0
    ip policy route-map ACCESS
    ip nat inside
!
interface FastEthernet1/0
    no ip address
!
interface FastEthernet1/0.1
    description Interface pour accès Internet
    ip address 100.1.1.2 255.255.255.252
    encapsulation isl 1
    ip nat outside
!
interface FastEthernet1/0.2
    description Interface pour accès VPN
    ip address 10.1.1.2 255.255.255.252
    encapsulation isl 2
!
ip classless
ip route 0.0.0.0 0.0.0.0 10.1.1.1
!
route-map ACCESS permit 10
    match ip address 100
    set ip next-hop 100.1.1.1
!
ip nat inside source list 101 interface FastEthernet1/0.1 overload
access-list 100 deny ip any 10.0.0.0 0.255.255.255
access-list 100 permit ip any any
access-list 101 permit ip 10.0.0.0 0.255.255.255 any
!
```

Pour cet exemple, il est supposé que le VPN client emploie le réseau 10.0.0.0/8 dans son plan d'adressage IP interne. Le route-map « ACCESS » permet de rediriger le trafic destiné à Internet (c'est-à-dire n'étant pas dirigé vers une adresse en 10.x.y.z, voir liste d'accès 100) sur l'interface « Internet », FastEthernet1/0.1. Pour pouvoir communiquer avec l'extérieur, une translation des adresses source en 10.0.0.0/8 (voir liste d'accès 101) du site doit également avoir lieu. Cette action est réalisée avec les commandes « ip nat […] ».

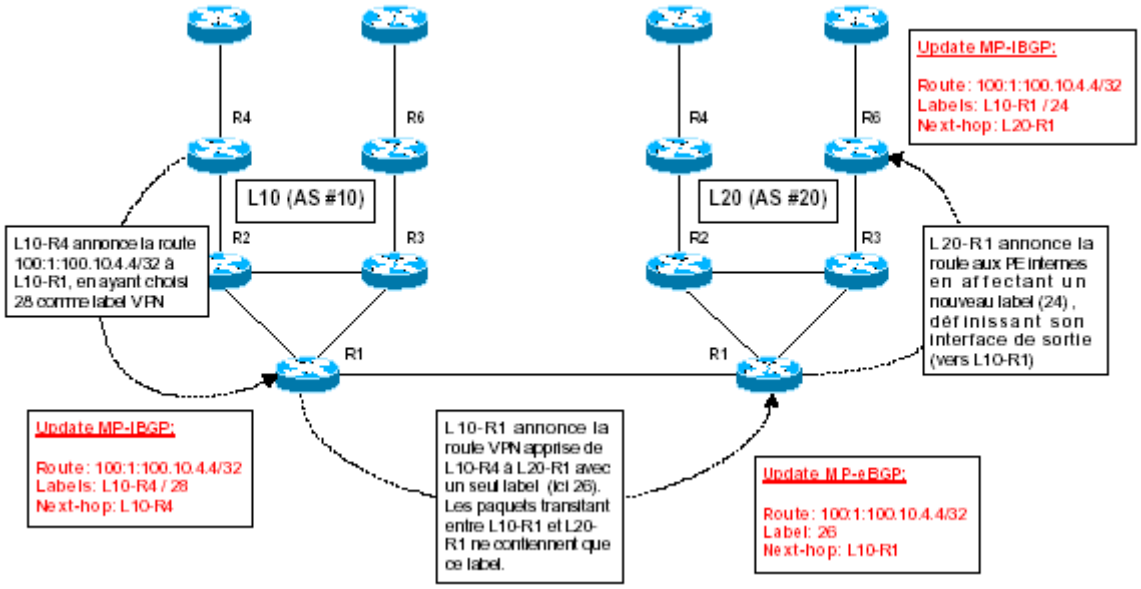
Remarque: en cas d'utilisation des VRF sur un routeur CE, il est important de garder à l'esprit que certaines fonctions (par exemple NAT, HSRP, etc.) peuvent ne pas être supportées avec des interfaces VRF.

4.8 – Signalisation Inter-AS (MP-eBGP)

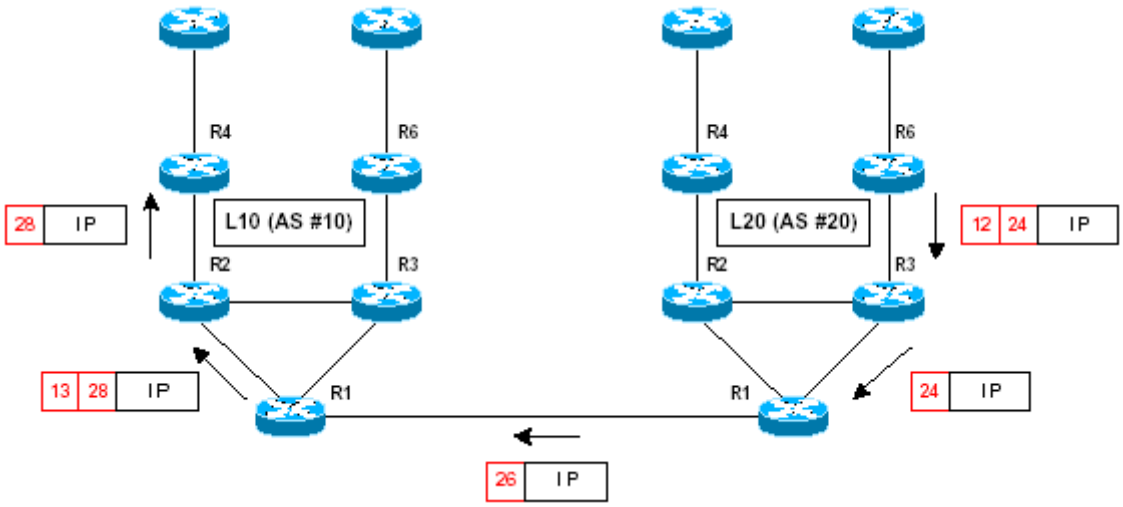
A partir de la version 12.1(5)T, il devient possible de créer relier des sites d'un 3ème VPN à travers des Autonomous Systems différents. En effet, l'annonce des routes VPNv4 inter-AS est fonctionnelle, ainsi que la transmission des paquets labellisés entre plusieurs backbones.

Le peering MP-eBGP se fait de manière similaire au peering MP-iBGP, à l'exception naturellement du numéro d'AS et du format des updates MP-eBGP, qui subit quelques légères adaptations. Chaque routeur d'interconnexion entre deux AS va annoncer les routes VPNv4 à ses voisins externes avec un seul label. Ce label sera utilisé entre les 2 AS pour la transmission des paquets.

Des tests ont été effectués lors du Workshop MPLS avec les versions 12.1(5)T et 12.2(0.4) pour étudier le fonctionnement de MP-eBGP. Pour cela, les deux pods, L10 et L20, ont été raccordés au moyen d'une liaison FastEthernet entre les deux routeurs R1. La signalisation MP-BGP est alors la suivante :



La transmission des paquets suivrait alors le schéma ci-dessous :



A l'intérieur du backbone MPLS, la méthode de transmission n'est en aucun cas modifiée. Seuls les routeurs MPLS en bordure d'AS ont un comportement différent : par exemple, sur le schéma ci-dessus, le routeur Cisco L10-R1 reçoit des paquets de l'AS #20 ne contenant qu'un 3eul label. Cependant, celui-ci est « converti » en deux labels (label PE + label VPN) lors de l'émission du paquet dans le backbone. Pour fonctionner correctement, certaines précautions doivent être prises au niveau des routeurs reliant les AS :

- Aucune session TDP ne doit être établie entre ces routeurs : la commande Cisco « tag-switching ip » ne doit donc pas être passée sur les interfaces interconnectant les voisins MP-eBGP. L'absence de cette commande Cisco n'empêche pas les routeurs de pouvoir traiter des paquets taggués arrivant sur ces interfaces ;
- Les routeurs doivent annoncer les routes apprises par MP-eBGP avec leur propre adresse IP (next-hop-self). Il est conseillé de se reporter à configurations du routeur Cisco L10-R1 fournie en Annexe II pour disposer d'un exemple complet et opérationnel. Le pod L20 étant configuré de manière similaire au pod L10, les configurations des routeurs de ce pod ne sont pas incluses dans ce document.

5 – Traffic Engineering (TE)

5.1 – Introduction

La plupart des gros réseaux IP, en particulier ceux des opérateurs, disposent de liens de secours en cas de panne. Toutefois, il est assez difficile d'obtenir une répartition du trafic sur ces liens qui ne sont traditionnellement pas utilisés, car n'étant pas sélectionnés comme chemins optimaux par l'IGP.

Le Trafic Engineering permet un meilleur emploi des liaisons, puisqu'il permet aux administrateurs réseau d'établir des tunnels LSP à travers le backbone MPLS, indépendamment de l'IGP.

5.2 – Types de tunnels

Les tunnels MPLS (appelés également trunks) peuvent être créés en indiquant la liste des routeurs à emprunter (méthode explicite) ou bien en utilisant la notion d'affinité (méthode dynamique). La notion d'affinité est simplement une valeur sur 32 bits spécifiée sur les interfaces des routeurs MPLS. La sélection du chemin s'effectue alors en indiquant (sur le routeur initiant le tunnel) une affinité et un masque.

5.3 – Critères de bande passante

Pour permettre une gestion plus souple du trafic, chaque interface MPLS susceptible d'être un point de transit pour des tunnels MPLS dispose d'une notion de priorité, définie sur 8 niveaux. Lors de l'établissement d'un nouveau tunnel, si celui-ci a une priorité plus grande que les autres tunnels et que la bande passante totale utilisable pour le TE est insuffisante, alors un tunnel moins prioritaire sera fermé. Ce mode de fonctionnement est appelé préemption.

Les niveaux de priorité sont codés avec valeur de 0 à 8, 0 correspondant à la priorité plus élevée et 8 à la priorité la plus basse. Chaque interface MPLS/TE doit être configurée avec un débit maximum alloué pour le Traffic Engineering. Par exemple, il est possible de n'autoriser que 50 Kb/s pour les tunnels sur une interface ayant un débit de 128 Kb/s.

Pour bien comprendre la notion de préemption, considérons l'exemple suivant : un tunnel de priorité 3 avec une bande passante (BW) de 50 Kb/s est déjà établie sur une interface autorisant 100 Kb/s de tunnels MPLS. Le routeur autorisera l'établissement d'un tunnel de priorité inférieure (>= 3) jusqu'à un débit de 50 Kb/s, c'est-à-dire la bande passante disponible. Par contre, si une demande pour l'établissement d'un tunnel de priorité supérieure survient (< 3), alors le routeur considère que la bande passante disponible est de 100 Kb/s. Le nouveau tunnel sera ainsi établi, et des tunnels de priorité moindre seront fermés si besoin est (préemption). Il est important de garder à l'esprit que créer un tunnel MPLS ne garantit pas la présence de la bande passante tout au long du réseau. En effet, les interfaces des routeurs sont configurées pour allouer un certain débit au TE ; mais en cas de congestion d'un lien avec du trafic hors tunnel, les tunnels n'auront pas de bande passante garantie.

5.4 – Etablissement d'un tunnel

Pour permettre au Trafic Engineering de fonctionner, le protocole de routage interne (IGP) doit être un protocole à état de liens (link-state). En effet, pour déterminer le chemin à emprunter par un tunnel, les routeurs doivent avoir la connaissance complète de la topologie du réseau. Les seuls protocoles supportant le TE sont donc OSPF et ISIS. Des extensions ont été rajoutés à ces deux protocoles pour gérer les critères de bande passante et de priorité sur les liens du réseau. Pour OSPF, des Opaque LSA ont été mis en place et pour IS-IS de nouveaux enregistrements TLV ont été créés.

Pour choisir le meilleur chemin correspondant aux critères de bande passante spécifiés, l'algorithme de SPF de ces protocoles a été modifié pour tenir compte de ces contraintes. Cet algorithme de calcul du meilleur chemin pour les tunnels LSP est appelé PCALC et permet donc le Constrained Based Routing.

L'algorithme PCALC est le suivant :

- Supprimer les liens qui ne disposent pas de la bande passante suffisante ;
- Supprimer les liens qui ne correspondent pas à l'affinité demandé ;
- Exécuter l'algorithme de Dijkstra sur la topologie restante (avec les métriques de l'IGP) ;
- Si plusieurs chemins subsistent, maximiser la valeur V, définie comme étant le minimum des valeurs de bande passante disponible sur chaque lien du chemin. Cela permet de load-balancer les tunnels de mêmes critères sur plusieurs chemins différents ;
- S'il existe encore plusieurs chemins, sélectionner celui-ci avec le nombre minimal de sauts (hops) ;
- Enfin, si plusieurs chemins sont encore présents, en choisir un aléatoirement. L'établissement d'un tunnel, après exécution de l'algorithme PCALC, est réalisé au moyen du protocole RSVP, auquel des extensions ont été rajoutées pour permettre le TE. Chaque noeud du chemin vérifie si les critères demandés sont compatibles avec son état actuel ; dans le cas contraire, la signalisation RSVP déclenche une annulation et le noeud « floode » son état pour en informer ses voisins.

5.5 – Réoptimisation

Un mécanisme important du Traffic Engineering est le fait de pouvoir réoptimiser les chemins empruntés par les tunnels LSP. Pour éviter une rupture dans le routage, un nouveau tunnel est établi parallèlement de celui déjà ouvert. Dès que l'établissement du tunnel de remplacement est réussie, le premier tunnel est fermé.

5.6 – Configuration IOS

Ce paragraphe présente les différentes étapes de configurations permettant d'activer le Traffic Engineering sur un backbone MPLS déjà en place.

5.6.1 – Activation globale de Traffic Engineering

Pour qu'un LSR puisse gérer le TE, la commande Cisco « mpls traffic-eng tunnels » doit être saisie en mode de configuration globale.

5.6.2 – Configuration de l'IGP

Suivant l'IGP (OSPF ou IS-IS), la méthode de configuration diffère :

- OSPF :

```
router ospf 10
    network 10.0.0.0 0.255.255.255 area 0
    mpls traffic-eng router-id Loopback0
    mpls traffic-eng area 0
 !
```

- IS-IS :

```
router isis
    net 49.0020.4444.4444.4444.00
    metric-style wide
    mpls traffic-eng router-id Loopback0
    mpls traffic-eng level-1
 !
```

5.6.3 – Configurations des interfaces Cisco

Chaque interface MPLS devant permettre le Traffic Engineering doit être configurée de la manière analogue à l'exemple ci-dessous :

```
interface serial 0/0
    ip address 10.20.24.4 255.255.255.0
    no ip directed-broadcast
    bandwidth 125
    encapsulation ppp
    tag-switching ip
    mpls traffic-eng tunnels
    ip rsvp bandwidth 100 100
 !
```

La commande Cisco « mpls traffic-eng tunnels » permet d'autoriser le passage de tunnels LSP à travers cette interface.

La commande Cisco « ip rsvp bandwidth » indique quel débit en Kb/s peut être utilisé pour les tunnels.

5.6.4 – Création d'un tunnel explicite

En prenant comme exemple le pod L20 avec l'établissement d'un tunnel entre L20-R4 et L20-R3, la configuration de L20-R4 serait la suivante :

```
interface Tunnel1
    ip unnumbered Loopback0
    no ip directed-broadcast
    no ip route-cache cef
    tunnel destination 10.20.3.3
    tunnel mode mpls traffic-eng
    tunnel mpls traffic-eng autoroute announce
    tunnel mpls traffic-eng bandwidth 10
    tunnel mpls traffic-eng path-option 1 explicit name WAY1
 !
 ip explicit-path name WAY1 enable
    next-address 10.20.24.2
    next-address 10.20.12.1
    next-address 10.20.13.3
 !
```

La liste des hops, définie par le chemin explicite « WAY1 » est ainsi :

- L20-R2 (10.20.24.2)
- L20-R1 (10.20.12.1)
- L20-R3 (10.20.13.3)

5.6.5 – Création d'un tunnel dynamique

La création d'un tunnel dynamique est relativement similaire à la création d'un tunnel explicite. Ici, l'affinité est fixée à 0x10, avec un masque de 0x11.

```
interface Tunnel2
    ip unnumbered Loopback0
    no ip directed-broadcast
    no ip route-cache cef
    tunnel destination 10.20.3.3
    tunnel mode mpls traffic-eng
    tunnel mpls traffic-eng autoroute announce
    tunnel mpls traffic-eng bandwidth 10
    tunnel mpls traffic-eng affinity 0x10 mask 0x11
    tunnel mpls traffic-eng path-option 1 dynamic
```

5.7 – Utilisation avec MPLS/VPN

Pour assurer un bon fonctionnement du Traffic Engineering avec MPLS/VPN, les deux extrémités des tunnels doivent établir une adjacence TDP. Pour cela, deux tunnels LSP doivent être établis (un pour chaque direction) et la commande Cisco « tagswitching ip » passée sur les interfaces Tunnels.

Au niveau de la transmission des paquets, 3 labels sont utilisés : un label correspondant au tunnel sera placé en première position, les deux labels MPLS/VPN (label PE + label VPN) étant placés après. Le label de tunnel est naturellement retiré par l'extrémité terminale du tunnel, et les paquets sont ensuite forwardés
normalement.

6 – Conclusion

A l'origine développé pour la rapidité, la commutation de paquets par tag switching a permis la mise en oeuvre de solutions de plus haut niveau, comme les VPN ou le Traffic Engineering. MPLS rassemble en une seule entité l'efficacité des protocoles de niveau 3, alliée à la vitesse de commutation de niveau 2.

7 – Annexe I: Configurations MPLS simples

Configuration Cisco de L10-R1 :

```
!
version 12.2
no service single-slot-reload-enable
service timestamps debug uptime
service timestamps log uptime
no service password-encryption
!
hostname L10-R1
!
boot system flash slot0:c3620-js-mz.122-0.4.bin
logging rate-limit console 10 except errors
enable secret 5 $1$AKNE$ZjooLX5ZFP7nbGr6E/ejh/
!
ip subnet-zero
!
no ip finger
no ip domain-lookup
!
ip cef
no ip dhcp-client network-discovery
call rsvp-sync
!
interface Loopback0
ip address 10.10.1.1 255.255.255.255
!
interface Serial0/0
description Vers L10-R3
bandwidth 125
ip address 10.10.13.1 255.255.255.0
encapsulation ppp
clockrate 125000
!
interface Serial0/1
description Vers L10-R2
bandwidth 125
ip address 10.10.12.1 255.255.255.0
encapsulation ppp
clockrate 125000
!
router ospf 10
log-adjacency-changes
passive-interface Ethernet0/0
network 10.0.0.0 0.255.255.255 area 0
!
ip classless
no ip http server
!
line con 0
exec-timeout 0 0
transport input none
line aux 0
line vty 0 4
exec-timeout 0 0
password cisco
login
!
end
```

Configuration Cisco de L10-R2 :

```
!
 version 12.2
 no service single-slot-reload-enable
 service timestamps debug uptime
 service timestamps log uptime
 no service password-encryption
 !
 hostname L10-R2
 !
 boot system flash flash:c3640-js-mz.122-0.4.bin
 logging rate-limit console 10 except errors
 enable secret 5 $1$M7Ig$NBWKag8D2u7Q9sOU9xDfm/
 !
 ip subnet-zero
 !
 no ip finger
 no ip domain-lookup
 !
 ip cef
 no ip dhcp-client network-discovery
 call rsvp-sync
 !
 interface Loopback0
 ip address 10.10.2.2 255.255.255.255
 !
 interface Serial0/0
 description Vers L10-R4
 bandwidth 125
 ip address 10.10.24.2 255.255.255.0
 encapsulation ppp
 clockrate 125000
 !
 interface Serial0/1
 description Vers L10-R1
 bandwidth 125
 ip address 10.10.12.2 255.255.255.0
 encapsulation ppp
 !
 interface Serial1/0
 description Vers L10-R3
 bandwidth 125
 ip address 10.10.23.2 255.255.255.0
 encapsulation ppp
 clockrate 125000
 !
 router ospf 10
 log-adjacency-changes
 network 10.0.0.0 0.255.255.255 area 0
 !
 ip classless
 no ip http server
 !
 line con 0
 exec-timeout 0 0
 transport input none
 line aux 0
 line vty 0 4
 exec-timeout 0 0
 password cisco
 login
```

```
!
end
```

Configuration Cisco de L10-R3 :

```
!
version 12.2
no service single-slot-reload-enable
service timestamps debug uptime
service timestamps log uptime
no service password-encryption
!
hostname L10-R3
!
boot system flash slot0:c3640-js-mz.122-0.4.bin
logging rate-limit console 10 except errors
enable secret 5 $1$3Paa$QoFQfhYLZLCidMokHBanf1
!
ip subnet-zero
!
no ip finger
no ip domain-lookup
!
ip cef
no ip dhcp-client network-discovery
call rsvp-sync
!
interface Loopback0
ip address 10.10.3.3 255.255.255.255
!
interface Serial0/0
description Vers L10-R5
bandwidth 125
ip address 10.10.35.3 255.255.255.0
encapsulation ppp
clockrate 125000
!
interface Serial0/1
description Vers L10-R1
bandwidth 125
ip address 10.10.13.3 255.255.255.0
encapsulation ppp
!
interface Serial1/0
description Vers L10-R2
bandwidth 125
ip address 10.10.23.3 255.255.255.0
encapsulation ppp
!
router ospf 10
log-adjacency-changes
network 10.0.0.0 0.255.255.255 area 0
!
ip classless
no ip http server
!
line con 0
exec-timeout 0 0
transport input none
line aux 0
line vty 0 4
exec-timeout 0 0
password cisco
login
```

```
!
end
```

Configuration Cisco de L10-R4 :

```
!
version 12.2
no service single-slot-reload-enable
service timestamps debug uptime
service timestamps log uptime
no service password-encryption
!
hostname L10-R4
!
boot system flash slot0:c3620-js-mz.122-0.4.bin
logging rate-limit console 10 except errors
enable secret 5 $1$mLZz$9KLAmd1tA023Bd5Z5mV.s1
!
ip subnet-zero
!
no ip finger
no ip domain-lookup
!
ip cef
no ip dhcp-client network-discovery
call rsvp-sync
!
interface Loopback0
ip address 10.10.4.4 255.255.255.255
!
interface Serial0/0
description Vers L10-R6
bandwidth 125
ip address 10.10.46.4 255.255.255.0
encapsulation ppp
no fair-queue
clockrate 125000
!
interface Serial0/1
description Vers L10-R2
bandwidth 125
ip address 10.10.24.4 255.255.255.0
encapsulation ppp
!
router ospf 10
log-adjacency-changes
network 10.0.0.0 0.255.255.255 area 0
!
ip classless
no ip http server
!
line con 0
exec-timeout 0 0
transport input none
line aux 0
line vty 0 4
exec-timeout 0 0
password cisco
login
!
end
```

Configuration Cisco de L10-R5 :

```
!
version 12.2
no service single-slot-reload-enable
service timestamps debug uptime
service timestamps log uptime
no service password-encryption
!
hostname L10-R5
!
boot system flash flash:c3620-js-mz.122-0.4.bin
logging rate-limit console 10 except errors
enable secret 5 $1$m80i$NykiaSFf0D9EdTDAjJozu.
!
ip subnet-zero
!
no ip finger
no ip domain-lookup
!
ip cef
no ip dhcp-client network-discovery
frame-relay switching
call rsvp-sync
!
interface Loopback0
ip address 10.10.5.5 255.255.255.255
!
interface Serial0/0
no ip address
encapsulation frame-relay
clockrate 125000
frame-relay lmi-type cisco
frame-relay intf-type dce
!
interface Serial0/0.1 point-to-point
description Vers L10-R7
bandwidth 125
ip address 10.10.57.5 255.255.255.0
frame-relay interface-dlci 100
!
interface Serial0/1
description Vers L10-R3
bandwidth 125
ip address 10.10.35.5 255.255.255.0
encapsulation ppp
!
router ospf 10
log-adjacency-changes
network 10.0.0.0 0.255.255.255 area 0
!
ip classless
no ip http server
!
line con 0
exec-timeout 0 0
transport input none
line aux 0
line vty 0 4
exec-timeout 0 0
password cisco
login
```

```
!
end
```

Configuration Cisco de L10-R6 :

```
!
version 12.2
no service single-slot-reload-enable
service timestamps debug uptime
service timestamps log uptime
no service password-encryption
!
hostname L10-R6
!
boot system flash flash:c2600-js-mz.122-0.4.bin
logging rate-limit console 10 except errors
enable secret 5 $1$bZ2c$wdF872FRcLXaObuYJD90u1
!
ip subnet-zero
!
no ip finger
no ip domain-lookup
!
ip cef
no ip dhcp-client network-discovery
call rsvp-sync
!
interface Loopback0
ip address 10.10.6.6 255.255.255.255
!
interface Serial0/0
description Vers L10-R4
bandwidth 125
ip address 10.10.46.6 255.255.255.0
encapsulation ppp
no fair-queue
!
router ospf 10
log-adjacency-changes
network 10.0.0.0 0.255.255.255 area 0
!
ip classless
no ip http server
!
line con 0
exec-timeout 0 0
transport input none
line aux 0
line vty 0 4
exec-timeout 0 0
password cisco
login
line vty 5 15
login
!
end
```

Configuration Cisco de L10-R7 :

```
!
 version 12.2
 no service single-slot-reload-enable
 service timestamps debug uptime
 service timestamps log uptime
 no service password-encryption
 !
 hostname L10-R7
 !
 boot system flash flash:c2600-js-mz.122-0.4.bin
 logging rate-limit console 10 except errors
 enable secret 5 $1$Fgps$hTMFn1J6vXX9P3Dh9JDaK/
 !
 ip subnet-zero
 !
 no ip finger
 no ip domain-lookup
 !
 ip cef
 no ip dhcp-client network-discovery
 call rsvp-sync
 !
 interface Loopback0
 ip address 10.10.7.7 255.255.255.255
 !
 interface Serial0/0
 no ip address
 encapsulation frame-relay
 frame-relay lmi-type cisco
 !
 interface Serial0/0.1 point-to-point
 description Vers L10-R5
 bandwidth 125
 ip address 10.10.57.7 255.255.255.0
 frame-relay interface-dlci 100
 !
 router ospf 10
 log-adjacency-changes
 network 10.0.0.0 0.255.255.255 area 0
 !
 ip classless
 no ip http server
 !
 line con 0
 exec-timeout 0 0
 transport input none
 line aux 0
 line vty 0 4
 password cisco
 login
 line vty 5 15
 login
 !
 end
```

8 – Annexe II: Configurations MPLS/VPN

Configuration Cisco de L10-R1 :

```
!
version 12.2
no service single-slot-reload-enable
service timestamps debug uptime
service timestamps log uptime
no service password-encryption
!
hostname L10-R1
!
boot system flash slot0:c3620-js-mz.122-0.4.bin
logging rate-limit console 10 except errors
enable secret 5 $1$AKNE$ZjooLX5ZFP7nbGr6E/ejh/
!
ip subnet-zero
!
no ip finger
no ip domain-lookup
!
ip cef
no ip dhcp-client network-discovery
call rsvp-sync
!
interface Loopback0
ip address 10.10.1.1 255.255.255.255
!
interface Ethernet0/0
ip address 200.1.112.1 255.255.255.0
half-duplex
!
interface Serial0/0
description Vers L10-R3
bandwidth 125
ip address 10.10.13.1 255.255.255.0
encapsulation ppp
clockrate 125000
tag-switching ip
!
interface Serial0/1
description Vers L10-R2
bandwidth 125
ip address 10.10.12.1 255.255.255.0
encapsulation ppp
clockrate 125000
tag-switching ip
!
interface FastEthernet1/0
ip address 50.50.50.1 255.255.255.0
duplex auto
speed auto
!
router ospf 10
log-adjacency-changes
passive-interface FastEthernet1/0
network 10.0.0.0 0.255.255.255 area 0
!
router bgp 10
no synchronization
no bgp default route-target filter
bgp log-neighbor-changes
bgp cluster-id 10
```

```
neighbor iBGP peer-group
no neighbor iBGP activate
neighbor iBGP remote-as 10
neighbor iBGP update-source Loopback0
neighbor iBGP soft-reconfiguration inbound
neighbor 50.50.50.2 remote-as 20
no neighbor 50.50.50.2 activate
no auto-summary
!
address-family vpnv4
neighbor iBGP activate
neighbor iBGP route-reflector-client
neighbor iBGP send-community extended
neighbor iBGP route-map NH_Rewrite out
neighbor 10.10.2.2 peer-group iBGP
neighbor 10.10.3.3 peer-group iBGP
neighbor 10.10.4.4 peer-group iBGP
neighbor 10.10.5.5 peer-group iBGP
neighbor 50.50.50.2 activate
neighbor 50.50.50.2 send-community extended
no auto-summary
exit-address-family
!
route-map NH_Rewrite permit 10
match as-path 1
set ip next-hop peer-address
!
route-map NH_Rewrite permit 20
!
ip as-path access-list 1 deny ^$
ip as-path access-list 1 permit .*
!
ip classless
no ip http server
!
line con 0
exec-timeout 0 0
transport input none
line aux 0
line vty 0 4
exec-timeout 0 0
password cisco
login
!
end
```

Configuration Cisco de L10-R2 :

```
!
version 12.2
no service single-slot-reload-enable
service timestamps debug uptime
service timestamps log uptime
no service password-encryption
!
hostname L10-R2
!
boot system flash flash:c3640-js-mz.122-0.4.bin
logging rate-limit console 10 except errors
enable secret 5 $1$M7Ig$NBWKag8D2u7Q9sOU9xDfm/
!
ip subnet-zero
!
no ip finger
no ip domain-lookup
!
ip cef
no ip dhcp-client network-discovery
call rsvp-sync
!
interface Loopback0
ip address 10.10.2.2 255.255.255.255
!
interface Ethernet0/0
no ip address
shutdown
half-duplex
!
interface Serial0/0
description Vers L10-R4
bandwidth 125
ip address 10.10.24.2 255.255.255.0
encapsulation ppp
clockrate 125000
tag-switching ip
!
interface Serial0/1
description Vers L10-R1
bandwidth 125
ip address 10.10.12.2 255.255.255.0
encapsulation ppp
tag-switching ip
!
interface Ethernet1/0
no ip address
shutdown
half-duplex
!
interface Serial1/0
description Vers L10-R3
bandwidth 125
ip address 10.10.23.2 255.255.255.0
encapsulation ppp
clockrate 125000
tag-switching ip
!
router ospf 10
log-adjacency-changes
```

```
network 10.0.0.0 0.255.255.255 area 0
!
router bgp 10
no synchronization
bgp log-neighbor-changes
neighbor 10.10.1.1 remote-as 10
neighbor 10.10.1.1 update-source Loopback0
no auto-summary
!
address-family vpnv4
neighbor 10.10.1.1 activate
neighbor 10.10.1.1 send-community extended
no auto-summary
exit-address-family
!
ip classless
no ip http server
!
line con 0
exec-timeout 0 0
transport input none
line aux 0
line vty 0 4
exec-timeout 0 0
password cisco
login
!
end
```

Configuration Cisco de L10-R3 :

```
!
 version 12.2
 no service single-slot-reload-enable
 service timestamps debug uptime
 service timestamps log uptime
 no service password-encryption
 !
 hostname L10-R3
 !
 boot system flash slot0:c3640-js-mz.122-0.4.bin
 logging rate-limit console 10 except errors
 enable secret 5 $1$3Paa$QoFQfhYLZLCidMokHBanf1
 !
 ip subnet-zero
 !
 no ip finger
 no ip domain-lookup
 !
 ip cef
 no ip dhcp-client network-discovery
 call rsvp-sync
 !
 ip vrf ADMIN
 rd 100:2000
 route-target export 100:2000
 route-target import 100:2001
 !
 interface Loopback0
 ip address 10.10.3.3 255.255.255.255
 !
 interface Loopback10
 description Interface de Management
 ip vrf forwarding ADMIN
 ip address 100.10.3.3 255.255.255.255
 !
 interface Ethernet0/0
 no ip address
 shutdown
 half-duplex
 !
 interface Serial0/0
 description Vers L10-R5
 bandwidth 125
 ip address 10.10.35.3 255.255.255.0
 encapsulation ppp
 clockrate 125000
 tag-switching ip
 !
 interface Serial0/1
 description Vers L10-R1
 bandwidth 125
 ip address 10.10.13.3 255.255.255.0
 encapsulation ppp
 tag-switching ip
 !
 interface Ethernet1/0
 no ip address
 shutdown
 half-duplex
 !
```

```
interface Serial1/0
description Vers L10-R2
bandwidth 125
ip address 10.10.23.3 255.255.255.0
encapsulation ppp
tag-switching ip
!
router ospf 10
log-adjacency-changes
network 10.0.0.0 0.255.255.255 area 0
!
router bgp 10
no synchronization
neighbor 10.10.1.1 remote-as 10
neighbor 10.10.1.1 update-source Loopback0
no neighbor 10.10.1.1 activate
no auto-summary
!
address-family ipv4 vrf ADMIN
redistribute connected
no auto-summary
no synchronization
exit-address-family
!
address-family vpnv4
neighbor 10.10.1.1 activate
neighbor 10.10.1.1 send-community extended
no auto-summary
exit-address-family
!
ip classless
no ip http server
!
line con 0
exec-timeout 0 0
transport input none
line aux 0
line vty 0 4
exec-timeout 0 0
password cisco
login
!
end
```

Configuration Cisco de L10-R4 :

```
!
 version 12.2
 no service single-slot-reload-enable
 service timestamps debug uptime
 service timestamps log uptime
 no service password-encryption
 !
 hostname L10-R4
 !
 boot system flash slot0:c3620-js-mz.122-0.4.bin
 logging rate-limit console 10 except errors
 enable secret 5 $1$mLZz$9KLAmd1tA023Bd5Z5mV.s1
 !
 memory-size iomem 10
 ip subnet-zero
 !
 no ip finger
 no ip domain-lookup
 !
 ip cef
 no ip dhcp-client network-discovery
 call rsvp-sync
 !
 ip vrf BLUE
 rd 100:1
 route-target import 100:1
 route-target export 100:1
 !
 ip vrf RED
 rd 100:2
 route-target import 100:2
 route-target export 100:2
 !
 ip vrf GREEN
 rd 100:3
 route-target import 100:3
 route-target export 100:3
 route-target import 100:2000
 route-target export 100:2001
 !
 interface Loopback0
 ip address 10.10.4.4 255.255.255.255
 !
 interface Loopback1
 ip vrf forwarding BLUE
 ip address 100.10.4.4 255.255.255.255
 !
 interface Loopback2
 ip vrf forwarding RED
 ip address 100.10.4.4 255.255.255.255
 !
 interface Loopback3
 ip vrf forwarding GREEN
 ip address 100.10.4.4 255.255.255.255
 !
 interface Ethernet0/0
 no ip address
 shutdown
 !
 interface Serial0/0
```

```
description Vers L10-R6
bandwidth 125
ip vrf forwarding GREEN
ip address 100.10.46.4 255.255.255.0
encapsulation ppp
no fair-queue
clockrate 125000
!
interface Serial0/1
description Vers L10-R2
bandwidth 125
ip address 10.10.24.4 255.255.255.0
encapsulation ppp
tag-switching ip
!
router ospf 10
log-adjacency-changes
network 10.0.0.0 0.255.255.255 area 0
!
router rip
version 2
!
address-family ipv4 vrf GREEN
version 2
redistribute bgp 10 metric 1
network 100.0.0.0
no auto-summary
exit-address-family
!
router bgp 10
no synchronization
bgp log-neighbor-changes
neighbor 10.10.1.1 remote-as 10
neighbor 10.10.1.1 update-source Loopback0
no neighbor 10.10.1.1 activate
no auto-summary
!
address-family ipv4 vrf BLUE
redistribute connected
no auto-summary
no synchronization
exit-address-family
!
address-family ipv4 vrf RED
redistribute connected
no auto-summary
no synchronization
exit-address-family
!
address-family ipv4 vrf GREEN
redistribute connected
redistribute rip
no auto-summary
no synchronization
exit-address-family
!
address-family vpnv4
neighbor 10.10.1.1 activate
neighbor 10.10.1.1 send-community extended
no auto-summary
exit-address-family
```

```
!
ip classless
no ip http server
!
line con 0
exec-timeout 0 0
transport input none
line aux 0
line vty 0 4
exec-timeout 0 0
password cisco
login
!
end
```

Configuration Cisco de L10-R5 :

```
!
version 12.2
no service single-slot-reload-enable
service timestamps debug uptime
service timestamps log uptime
no service password-encryption
!
hostname L10-R5
!
boot system flash flash:c3620-js-mz.122-0.4.bin
logging rate-limit console 10 except errors
enable secret 5 $1$m80i$NykiaSFf0D9EdTDAjJozu.
!
ip subnet-zero
!
no ip finger
no ip domain-lookup
!
ip cef
no ip dhcp-client network-discovery
frame-relay switching
call rsvp-sync
!
ip vrf BLUE
rd 100:1
route-target import 100:1
route-target export 100:1
!
ip vrf RED
rd 100:2
route-target import 100:2
route-target export 100:2
route-target import 100:2000
route-target export 100:2001
!
ip vrf GREEN
rd 100:3
route-target import 100:3
route-target export 100:3
!
interface Loopback0
ip address 10.10.5.5 255.255.255.255
!
interface Loopback1
ip vrf forwarding BLUE
ip address 100.10.5.5 255.255.255.255
!
interface Loopback2
ip vrf forwarding RED
ip address 100.10.5.5 255.255.255.255
!
interface Loopback3
ip vrf forwarding GREEN
ip address 100.10.5.5 255.255.255.255
!
interface Ethernet0/0
no ip address
half-duplex
!
interface Serial0/0
```

```
no ip address
encapsulation frame-relay
clockrate 125000
frame-relay lmi-type cisco
frame-relay intf-type dce
!
interface Serial0/0.1 point-to-point
description Vers L10-R7
bandwidth 125
ip vrf forwarding RED
ip address 100.10.57.5 255.255.255.0
frame-relay interface-dlci 100
!
interface Serial0/1
description Vers L10-R3
bandwidth 125
ip address 10.10.35.5 255.255.255.0
encapsulation ppp
tag-switching ip
!
router ospf 10
log-adjacency-changes
network 10.0.0.0 0.255.255.255 area 0
!
router bgp 10
no synchronization
bgp log-neighbor-changes
neighbor 10.10.1.1 remote-as 10
neighbor 10.10.1.1 update-source Loopback0
no neighbor 10.10.1.1 activate
no auto-summary
!
address-family ipv4 vrf BLUE
redistribute connected
no auto-summary
no synchronization
exit-address-family
!
address-family ipv4 vrf RED
redistribute connected
neighbor 100.10.57.7 remote-as 102
neighbor 100.10.57.7 activate
no auto-summary
no synchronization
exit-address-family
!
address-family ipv4 vrf GREEN
redistribute connected
no auto-summary
no synchronization
exit-address-family
!
address-family vpnv4
neighbor 10.10.1.1 activate
neighbor 10.10.1.1 send-community extended
no auto-summary
exit-address-family
!
ip classless
no ip http server
!
```

```
line con 0
exec-timeout 0 0
transport input none
line aux 0
line vty 0 4
exec-timeout 0 0
password cisco
login
!
end
```

Configuration Cisco de L10-R6 :

```
!
version 12.2
no service single-slot-reload-enable
service timestamps debug uptime
service timestamps log uptime
no service password-encryption
!
hostname L10-R6
!
boot system flash flash:c2600-js-mz.122-0.4.bin
logging rate-limit console 10 except errors
enable secret 5 $1$bZ2c$wdF872FRcLXaObuYJD90u1
!
ip subnet-zero
!
no ip finger
no ip domain-lookup
!
ip cef
no ip dhcp-client network-discovery
call rsvp-sync
!
interface Loopback0
ip address 100.10.6.6 255.255.255.255
!
interface Ethernet0/0
no ip address
half-duplex
!
interface Serial0/0
description Vers L10-R4
bandwidth 125
ip address 100.10.46.6 255.255.255.0
encapsulation ppp
no fair-queue
!
interface FastEthernet1/0
no ip address
duplex auto
speed auto
!
interface FastEthernet1/0.1
encapsulation dot1Q 161
ip address 100.10.161.6 255.255.255.0
no ip redirects
!
interface FastEthernet1/0.2
encapsulation dot1Q 162
ip address 100.10.162.6 255.255.255.0
no ip redirects
!
router rip
version 2
network 100.0.0.0
!
ip classless
no ip http server
!
line con 0
exec-timeout 0 0
```

```
transport input none
line aux 0
line vty 0 4
exec-timeout 0 0
password cisco
login
line vty 5 15
login
!
end
```

Configuration Cisco de L10-R7 :

```
!
version 12.2
no service single-slot-reload-enable
service timestamps debug uptime
service timestamps log uptime
no service password-encryption
!
hostname L10-R7
!
boot system flash flash:c2600-js-mz.122-0.4.bin
logging rate-limit console 10 except errors
enable secret 5 $1$Fgps$hTMFn1J6vXX9P3Dh9JDaK/
!
ip subnet-zero
!
no ip finger
no ip domain-lookup
!
ip cef
no ip dhcp-client network-discovery
call rsvp-sync
!
interface Loopback0
ip address 100.10.7.7 255.255.255.255
!
interface Ethernet0/0
no ip address
half-duplex
!
interface Serial0/0
no ip address
encapsulation frame-relay
frame-relay lmi-type cisco
!
interface Serial0/0.1 point-to-point
description Vers L10-R5
bandwidth 125
ip address 100.10.57.7 255.255.255.0
frame-relay interface-dlci 100
!
interface FastEthernet1/0
no ip address
duplex auto
speed auto
!
interface FastEthernet1/0.1
encapsulation dot1Q 171
ip address 100.10.171.7 255.255.255.0
no ip redirects
!
interface FastEthernet1/0.2
encapsulation dot1Q 172
ip address 100.10.172.7 255.255.255.0
no ip redirects
!
router bgp 102
bgp log-neighbor-changes
network 100.10.7.7 mask 255.255.255.255
network 100.10.171.0 mask 255.255.255.0
network 100.10.172.0 mask 255.255.255.0
```

```
neighbor 100.10.57.5 remote-as 10
!
ip classless
no ip http server
!
line con 0
exec-timeout 0 0
transport input none
line aux 0
line vty 0 4
password cisco
login
line vty 5 15
login
!
no scheduler allocate
end
```

9 – Les vidéos

- **9.1 - What is an autonomous system ? (https://www.frameip.com/mpls-cisco/?video=318#video-318)**

  Cette video en anglais vous présente ce qu'est un AS (Autonomous System). The internet is an internetwork—a network of computer networks. When discussing routing, we sometimes refer to these independent networks asautonomous systems. Each autonomous system is responsible for routing packets internally. To travel across the broader internet, packets must traverse multiple autonomous systems. So it is more accurate to think of internet routing as happening across autonomous systems, rather than between individual computers. So yes: the internet is an internetwork a network of computer networks !

  (https://www.frameip.com/mpls-cisco/?video=318#video-318)

- **9.2 - What is the border gateway protocol (BGP) ? (https://www.frameip.com/mpls-cisco/?video=317#video-317)**

  Cette video en anglais vous présente le protocole de routage BGP (Border Gateway Protoocol). Routing on the internet is actually performed between autonomous systems (ASes). Those autonomous systems need a way to determine how to send data to ASes that they are not directly connected with. To do this, the use the border gateway protocol, or BGP. By exchanging routing announcements using BGP routers located in each autonomous system can eventually build up a routing table allowing them to determine the best way to reach any other AS connected to the internet.

(https://www.frameip.com/mpls-cisco/?video=317#video-317)

- **9.3 - Configuration MPLS Cisco - 6/6 - Redistribution de route (https://www.frameip.com/mpls-cisco/?video=158#video-158)**

  Cette vidéo vous présente comment configurer la redistribution de route dans l'environnement MPLS Cisco.

(https://www.frameip.com/mpls-cisco/?video=158#video-158)

- **9.4 - Configuration MPLS Cisco - 5/6 - BGP (https://www.frameip.com/mpls-cisco/?video=157#video-157)**

  Cette vidéo vous présente comment configurer BGP (Border Gateway Protocol) dans l'environnement MPLS Cisco.

(https://www.frameip.com/mpls-cisco/?video=157#video-157)

- **9.5 - Configuration MPLS Cisco - 4/6 - MPLS (https://www.frameip.com/mpls-cisco/?video=156#video-156)**

Cette vidéo vous présente comment configurer MPLS (MultiProtocol Label Switching) dans l'environnement MPLS Cisco.

(https://www.frameip.com/mpls-cisco/?video=156#video-156)

- **9.6 - Configuration MPLS Cisco - 3/6 - IGP RIP (https://www.frameip.com/mpls-cisco/?video=155#video-155)**

Cette vidéo vous présente comment configurer RIP (Routing Information Protocol) dans l'environnement MPLS Cisco.

(https://www.frameip.com/mpls-cisco/?video=155#video-155)

- **9.7 - Configuration MPLS Cisco - 2/6 - IGP OSPF (https://www.frameip.com/mpls-cisco/?video=154#video-154)**

  Cette vidéo vous présente comment configurer OSPF (Open Shortest Path First) dans l'environnement MPLS Cisco.

(https://www.frameip.com/mpls-cisco/?video=154#video-154)

- **9.8 - Configuration MPLS Cisco - 1/6 - Adressage (https://www.frameip.com/mpls-cisco/?video=153#video-153)**

  Cette vidéo vous présente comment configurer MPLS (MultiProtocol Label Switching) sur un routeur Cisco.

(https://www.frameip.com/mpls-cisco/?video=153#video-153)

- **9.9 - Configuration de VRF sur un routeur Cisco (https://www.frameip.com/mpls-cisco/?video=146#video-146)**

Cette vidéo vous présente la notion de VRF (Virtual Routing & Forwarding) et procède à la configuration simple de 2 VRFs sur un routeur Cisco.

(https://www.frameip.com/mpls-cisco/?video=146#video-146)

- **9.10 - MultiProtocol Label Switching par Mr Cisco (https://www.frameip.com/mpls-cisco/?video=103#video-103)**

Vidéo en anglais de Mr Cisco présentant le protocole MPLS (MultiProtocol Label Switching).

(https://www.frameip.com/mpls-cisco/?video=103#video-103)

10 – Suivi du document

Création et suivi de la documentation par Christophe Fillot et _SebF

11 – Discussion autour de l'mplémentation de MPLS avec Cisco

Vous pouvez poser toutes vos questions, faire part de vos remarques et partager vos expériences à propos de l'mplémentation de MPLS avec Cisco. Pour cela, n'hésitez pas à laisser un commentaire ci-dessous :

Commentaire et discussion

11 commentaires sur la page : "Implémentation de MPLS avec Cisco"

Dospard dit :
16 juin 2022 à 14 h 36 min (https://www.frameip.com/mpls-cisco/#comment-235053)
Je voulais dire que le CE connecté au spoke1 ne parvient pas à faire un ping au CE connecté au spoke2.

Dospard dit :
16 juin 2022 à 14 h 34 min (https://www.frameip.com/mpls-cisco/#comment-235051)
Bonjour!
J'ai trouvé le document très édifiant.
J'ai conçu un réseau mpls-vpn hub and spoke sous gns3. Le hub a importé les réseaux liés aux différents spokes, cependant il ne parvient à exporter les réseaux du spoke1 vers le spoke2 et vice-versa. Conséquence, le spoke1 ne peut communiquer avec le spoke2. J'ai utilisé la commande » allowas-in » au niveau de l'une des interfaces du hub pour éviter le rejet des paquets par lecture de l'as-path mais cela ne marche pas.
Dois je aller dans les CE faire du » redistribute connected » ?

BILLY dit :
18 août 2021 à 14 h 34 min (https://www.frameip.com/mpls-cisco/#comment-105407)
MERCI

ARIFI Mouloud dit :
7 juillet 2021 à 12 h 44 min (https://www.frameip.com/mpls-cisco/#comment-93846)
Bonjour, ce que je ne comprends pas , c'est comment relier le router custemer edge au router provider edge . est ce par ligne téléphonique et un modém ?

Merci.

ARIFI Mouloud dit :
7 juillet 2021 à 12 h 37 min (https://www.frameip.com/mpls-cisco/#comment-93842)
Merci pour tout. C'est génial.

Anthony dit :
15 avril 2021 à 16 h 15 min (https://www.frameip.com/mpls-cisco/#comment-66081)
Merci pour la présentation et les explications très claires.

ZRELLI dit :
14 juin 2020 à 11 h 51 min (https://www.frameip.com/mpls-cisco/#comment-19100)
Merci pour ce support

Clovis dit :
23 août 2018 à 1 h 22 min (https://www.frameip.com/mpls-cisco/#comment-9118)
Je veux une configuration simple de MPLS sans VPN avec une description des différente commandes utilisé a chaque étape

sebastien.fontaine dit :
27 août 2018 à 20 h 02 min (https://www.frameip.com/mpls-cisco/#comment-9183)
Lu Clovis,

Il faudrait déjà connaître ton architecture, tes équipements et etc … 🙂

@+
Sebastien FONTAINE

alisawii dit :
9 octobre 2017 à 14 h 07 min (https://www.frameip.com/mpls-cisco/#comment-1200)
Merci votre documentation a plein des ressources.

serge dit :
13 août 2017 à 2 h 45 min (https://www.frameip.com/mpls-cisco/#comment-6)
merci pour votre document ça m'a vraiment aider.