

TP3 – VPN BPG-MPLS

➤ Topologie et configuration

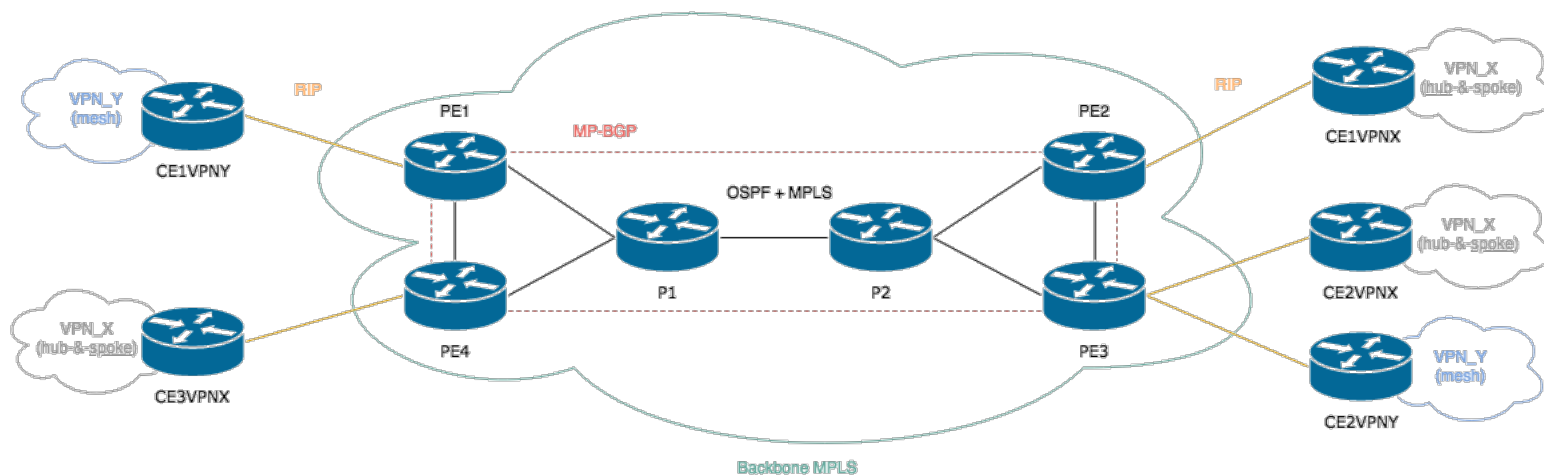


Figure 1 – Schéma des différents protocoles de routage et des VPN

L'architecture VPN BGP-MPLS, qui permet d'isoler le trafic entre sites n'appartenant pas au même VPN, est composée de plusieurs routeurs avec différents rôles :

- P (Provider) : ces routeurs n'ont aucune connaissance de la notion de VPN. Ils se contentent d'acheminer les paquets grâce à la commutation de labels ;
- PE (Provider Edge) : ces routeurs ont une ou plusieurs interfaces reliées à des routeurs clients (CE) et ont la capacité de gérer plusieurs tables de routage grâce à la notion de VRF (VPN Routing and Forwarding) ;
- CE (Customer Edge) : ces routeurs appartiennent au client et n'ont aucune connaissance des VPN ou même de la notion de label.

Différents protocoles de routage sont ainsi utilisés (Figure 1) :

- MPLS sur les routeurs du fournisseur (P et PE), c'est le backbone MPLS ;
- OSPF comme protocole de routage interne (IGP) dans ce backbone ;
- MP-BGP au niveau des PE, pour l'échange des routes multicast et des routes VPNv4 ;
- RIP au niveau des CE et PE, pour l'échange des routes entre ces deux types de routeur.

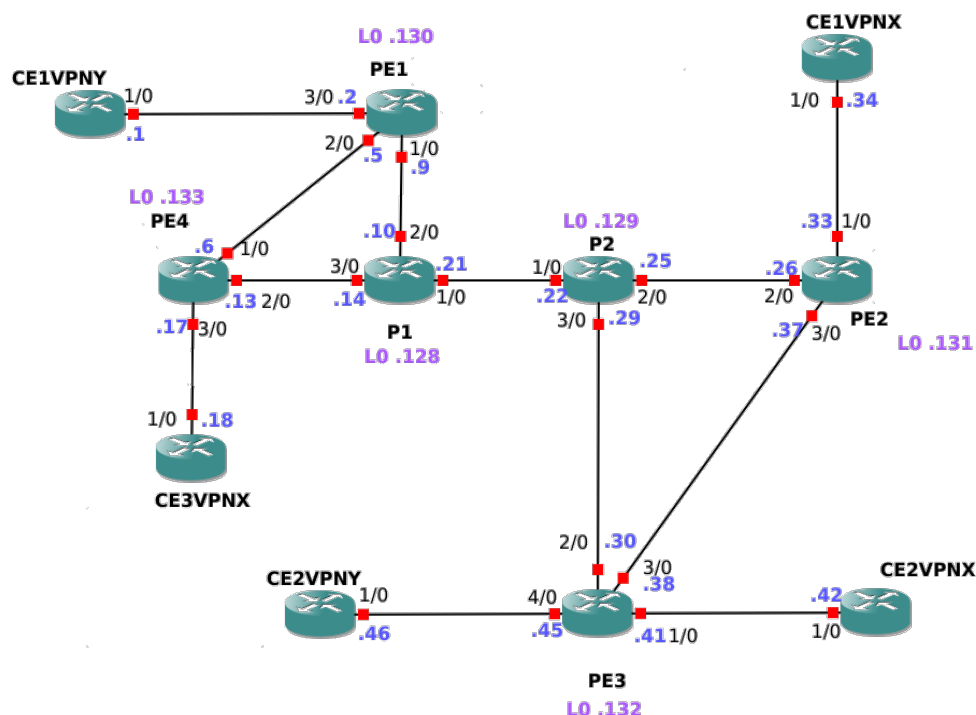


Figure 2 – Topologie du réseau

Le fichier `tp3_vpn.gns3` décrit la topologie (Figure 2).

Les fichiers de configuration des routeurs `<libellé_routeur>.cfg` sont joints à l'archive.

Les adresses IP sont de la forme `10.0.0.x/30`. Les routeurs du fournisseur (P et PE) possèdent une adresse de loopback sur l'interface `lo0` de type `10.0.0.x/32`.

Les routeurs ont pu être redémarrés entre les questions, la valeur des labels peut donc être différente.

➤ VPN sans ingénierie de trafic

Q0. Le premier VPN, `VPN_X`, est connecté via trois CE (`CE1VPNX`, `CE2VPNX` et `CE3VPNX`). Il utilise la topologie Hub-and-Spoke avec `PE2` comme hub et les autres PE comme stub.

On configure la VRF `VPN_X` de cette façon (à gauche sur le hub `PE2` et à droite sur les stubs `PE1`, `PE3` et `PE4`):

```
ip vrf VPN_X
rd 500:1
route-target export 500:1
route-target import 500:2
!

ip vrf VPN_X
rd 500:1
route-target export 500:2
route-target import 500:1
!
```

Le site central (`PE2`) importe les routes de tous les sites clients `RT 500:2`, et chaque site client importe celles du site central `RT:500:1`.

Le deuxième VPN, `VPN_Y`, est connecté via deux CE (`CE1VPNY` et `CE2VPNY`). Il utilise la topologie Mesh. On configure la VRF `VPN_Y` de cette façon (sur l'ensemble des PE) :

```
ip vrf VPN_Y
rd 500:3
route-target export 500:3
route-target import 500:3
!
```

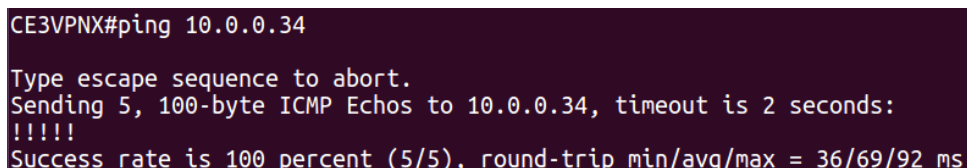
Pour échanger les routes entre tous les sites, chaque PE importe et exporte le `RT 500:3`.

Q1. Pour vérifier la connectivité des VPN, on utilise les commandes `ping` et `traceroute`.

Sur les PE, il faut adapter la commande pour forcer l'utilisation du VPN grâce à la VRF correspondante : `ping vrf <nom_vrf> <adresse_ip>` (remplacer `ping` par `traceroute` si besoin).

Sur les autres routeurs, P et CE, la commande est la même qu'habituellement, c'est-à-dire sans spécifier la VRF à utiliser.

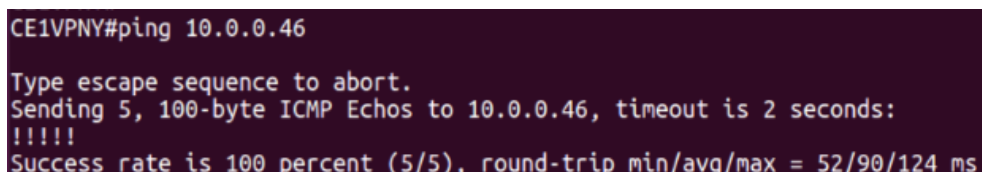
Dans le cas du VPN `VPN_X`, en topologie Hub-and-Spoke, chaque client Spoke, `CE2VPNX` et `CE3VPNX`, peut communiquer avec le client Hub (Figure 3), `CE1VPNX`. Ce dernier peut communiquer avec n'importe quel autre client, cependant, un client Spoke ne peut pas communiquer avec un autre client du même type.



```
CE3VPNX#ping 10.0.0.34
Type escape sequence to abort.
Sending 5, 100-byte ICMP Echos to 10.0.0.34, timeout is 2 seconds:
!!!!
Success rate is 100 percent (5/5), round-trip min/avg/max = 36/69/92 ms
```

Figure 3 – Ping de `CE3VPNX` (Spoke) à `CE1VPNX` (Hub)

Dans le cas du VPN `VPN_Y`, en topologie Mesh, on peut constater que les deux seuls clients connectés par ce VPN, `CE1VPNY` et `CE2VPNY`, peuvent communiquer entre eux. La connexion a donc bien été établie.



```
CE1VPNY#ping 10.0.0.46
Type escape sequence to abort.
Sending 5, 100-byte ICMP Echos to 10.0.0.46, timeout is 2 seconds:
!!!!
Success rate is 100 percent (5/5), round-trip min/avg/max = 52/90/124 ms
```

Figure 4 – Ping de `CE1VPNY` à `CE2VPNY`

Du fait que la topologie de ce VPN soit en Mesh, les clients utiliseront la route la plus courte pour communiquer entre eux, ie. CE2VPNY → PE3 → P2 → P1 → PE1 → CE1VPNY ici (Figure 5).

```
CE2VPNY#traceroute 10.0.0.1
Type escape sequence to abort.
Tracing the route to 10.0.0.1

 0 10.0.0.45 64 msec 72 msec 64 msec
 1 10.0.0.29 96 msec 100 msec 96 msec
 2 10.0.0.21 100 msec 100 msec 96 msec
 3 10.0.0.2 108 msec 100 msec 104 msec
 4 10.0.0.1 100 msec * 140 msec
```

Figure 5 – Traceroute de CE2VPNY à CE1VPNY

Les préfixes joignables par un client sont uniquement ceux correspondant aux autres clients connectés au même VPN, en effet la connaissance des clients est très limitée.

```
CE1VPNX#sh ip route
Codes: C - connected, S - static, R - RIP, M - mobile, B - BGP
       D - EIGRP, EX - EIGRP external, O - OSPF, IA - OSPF inter area
       N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
       E1 - OSPF external type 1, E2 - OSPF external type 2
       i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
       ia - IS-IS inter area, * - candidate default, U - per-user static route
       o - ODR, P - periodic downloaded static route

Gateway of last resort is not set

10.0.0.0/30 is subnetted, 3 subnets
R    10.0.0.16 [120/1] via 10.0.0.33, 00:00:22, GigabitEthernet1/0
R    10.0.0.40 [120/1] via 10.0.0.33, 00:00:22, GigabitEthernet1/0
C    10.0.0.32 is directly connected, GigabitEthernet1/0
```

Figure 6 – Table de routage de CE1VPNX

```
CE1VPNY#sh ip route
Codes: C - connected, S - static, R - RIP, M - mobile, B - BGP
       D - EIGRP, EX - EIGRP external, O - OSPF, IA - OSPF inter area
       N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
       E1 - OSPF external type 1, E2 - OSPF external type 2
       i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
       ia - IS-IS inter area, * - candidate default, U - per-user static route
       o - ODR, P - periodic downloaded static route

Gateway of last resort is not set

10.0.0.0/30 is subnetted, 2 subnets
C    10.0.0.0 is directly connected, GigabitEthernet1/0
R    10.0.0.44 [120/1] via 10.0.0.2, 00:00:16, GigabitEthernet1/0
```

Figure 7 – Table de routage de CE1VPNY

Les PE et P, quant à eux, ont connaissance de tous les préfixes joignables dans le backbone MPLS (Figure 1). Par exemple, on peut atteindre le préfixe 10.0.0.12/30 (entre PE4 et P1) à partir de PE1 via 10.0.0.10 (P1) ou 10.0.0.6 (PE4).

```
PE1#sh ip route
Codes: C - connected, S - static, R - RIP, M - mobile, B - BGP
       D - EIGRP, EX - EIGRP external, O - OSPF, IA - OSPF inter area
       N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
       E1 - OSPF external type 1, E2 - OSPF external type 2
       i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
       ia - IS-IS inter area, * - candidate default, U - per-user static route
       o - ODR, P - periodic downloaded static route

Gateway of last resort is not set

10.0.0.0/8 is variably subnetted, 14 subnets, 2 masks
C    10.0.0.8/30 is directly connected, GigabitEthernet1/0
O    10.0.0.12/30 [110/2] via 10.0.0.10, 01:31:17, GigabitEthernet1/0
      [110/2] via 10.0.0.6, 01:31:17, GigabitEthernet2/0
C    10.0.0.4/30 is directly connected, GigabitEthernet2/0
O    10.0.0.24/30 [110/3] via 10.0.0.10, 01:31:17, GigabitEthernet1/0
O    10.0.0.28/30 [110/3] via 10.0.0.10, 01:31:17, GigabitEthernet1/0
O    10.0.0.20/30 [110/2] via 10.0.0.10, 01:31:17, GigabitEthernet1/0
O    10.0.0.36/30 [110/4] via 10.0.0.10, 01:31:17, GigabitEthernet1/0
C    10.0.0.48/30 is directly connected, GigabitEthernet4/0
C    10.0.0.130/32 is directly connected, Loopback0
O    10.0.0.131/32 [110/4] via 10.0.0.10, 01:31:17, GigabitEthernet1/0
O    10.0.0.128/32 [110/2] via 10.0.0.10, 01:31:17, GigabitEthernet1/0
O    10.0.0.129/32 [110/3] via 10.0.0.10, 01:31:17, GigabitEthernet1/0
O    10.0.0.132/32 [110/4] via 10.0.0.10, 01:31:18, GigabitEthernet1/0
O    10.0.0.133/32 [110/2] via 10.0.0.6, 01:31:18, GigabitEthernet2/0
```

Figure 8 – Table de routage de PE1

De plus, les PE ont connaissance des préfixes des clients connectés à un VPN si ceux-ci ont la VRF associée. Par exemple, PE1 a connaissance des préfixes des deux clients connectés au VPN VPN_Y (Figure 9).

```
PE1#sh ip route vrf VPN_Y

Routing Table: VPN_Y
Codes: C - connected, S - static, R - RIP, M - mobile, B - BGP
       D - EIGRP, EX - EIGRP external, O - OSPF, IA - OSPF inter area
       N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
       E1 - OSPF external type 1, E2 - OSPF external type 2
       i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
       ia - IS-IS inter area, * - candidate default, U - per-user static route
       o - ODR, P - periodic downloaded static route

Gateway of last resort is not set

10.0.0.0/30 is subnetted, 2 subnets
C    10.0.0.0 is directly connected, GigabitEthernet3/0
B    10.0.0.44 [200/0] via 10.0.0.132, 00:03:56
```

Figure 9 – Table de routage de la VRF VPN_Y sur PE1

Ils ont aussi connaissance des préfixes qui relient les PE et les CE ainsi que les VRF associés grâce à MP-BGP utilisé entre les PE (Figure 10).

```
PE2#sh ip bgp vpnv4 all
BGP table version is 11, local router ID is 10.0.0.131
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
               r RIB-failure, S Stale
Origin codes: i - IGP, e - EGP, ? - incomplete

   Network        Next Hop        Metric LocPrf Weight Path
Route Distinguisher: 500:1 (default for vrf VPN_X)
*>i10.0.0.16/30    10.0.0.133         0      100      0 ?
*> 10.0.0.32/30    0.0.0.0            0      100     32768 ?
*>i10.0.0.40/30    10.0.0.132         0      100      0 ?
Route Distinguisher: 500:3 (default for vrf VPN_Y)
*>i10.0.0.0/30     10.0.0.130         0      100      0 ?
*>i10.0.0.44/30    10.0.0.132         0      100      0 ?
```

Figure 10 – Liste des VPN

Q2. Les labels sont utilisés pour la transmission de paquets provenant des clients sur le backbone MPLS. Par exemple, pour atteindre le client CE2VPNY depuis CE1VPNY, PE1 va encapsuler deux labels : le premier sert à atteindre le PE de destination, ici PE3, et le deuxième sert à déterminer l'interface de sortie à laquelle est reliée le client CE2VPNY. Ce label est appris grâce aux messages Updates de MP-BGP.

La table de routage de la VRF VPN_Y sur PE1 (Figure 9) montre que le préfixe 10.0.0.44/30 (CE2VPNY) est bien accessible via 10.0.0.132 (PE3).

Le premier label, pour atteindre CE2VPNY, est déterminé par la table des labels de PE1 (Figure 11), ici 23. Il est utilisé pour la commutation dans le backbone MPLS.

```
PE1#sh mpls forwarding-table
Local  Outgoing  Prefix      Bytes tag  Outgoing   Next Hop
tag    tag or VC  or Tunnel Id  switched   interface
24     23        10.0.0.132/32  0          Gi1/0      10.0.0.10
```

Figure 11 – Table des labels partielle de PE1

Le deuxième label, pour déterminer l'interface de sortie sur PE3, est appris par MP-BGP (Figure 12), ici 27.

```
PE1#sh ip bgp vpnv4 vrf VPN_Y labels
   Network        Next Hop        In label/Out label
Route Distinguisher: 500:3 (VPN_Y)
10.0.0.0/30       0.0.0.0         26/aggregate(VPN_Y)
10.0.0.44/30      10.0.0.132      nlabel/27
```

Figure 12 – Liste des labels VPN sur PE1

Un `traceroute` depuis PE1 vers PE3 et l'adresse IP 10.0.0.132 (Figure 13) et un autre vers 10.0.0.45 (Figure 14) montrent que seul le premier label, de valeur 23, est utilisé.

```
PE1#traceroute 10.0.0.132

 1 10.0.0.10 [MPLS: Label 23 Exp 0] 108 msec 32 msec 72 msec
 2 10.0.0.22 [MPLS: Label 23 Exp 0] 76 msec 28 msec 44 msec
 3 10.0.0.30 80 msec * 52 msec
```

Figure 13 – Traceroute de PE1 vers PE3 (loopback)

```

PE1#traceroute vrf VPN_Y 10.0.0.45
 1 10.0.0.10 [MPLS: Labels 23/27 Exp 0] 40 msec 36 msec 32 msec
 2 10.0.0.22 [MPLS: Labels 23/27 Exp 0] 36 msec 32 msec 72 msec
 3 10.0.0.45 36 msec * 20 msec

```

Figure 14 – Traceroute de PE1 vers PE3 (VPN_Y)

Q3. Pour simplifier, on décide de retirer la ligne `ip vrf forwarding VPN_Y` de PE1 et PE3, associée à l'interface liant le PE au CE, de lancer la capture de paquets avec Wireshark puis de rajouter cette ligne. On observe des messages RIPv2 (voire question suivante **Q4**). On observe aussi des messages BGP de type Update échangés entre les PE (Figure 15).

10.0.0.130	10.0.0.131	BGP	149	UPDATE	Message
10.0.0.130	10.0.0.132	BGP	149	UPDATE	Message
10.0.0.132	10.0.0.130	BGP	145	UPDATE	Message
10.0.0.130	10.0.0.131	BGP	149	UPDATE	Message
10.0.0.130	10.0.0.132	BGP	149	UPDATE	Message
10.0.0.132	10.0.0.130	BGP	145	UPDATE	Message

Figure 15 – Capture de paquets sur le lien PE1-P1 avec Wireshark

Q4. Pour ajouter un préfixe dans un VPN existant, on retire puis on rajoute l'annonce d'un préfixe dans RIP sur un des clients. Par exemple, sur CE2VPNY, on utilise les commandes `router rip` puis `no network 10.0.0.44`, on lance la capture de paquets avec Wireshark puis on rajoute l'annonce du préfixe avec `router rip` puis `network 10.0.0.44`.

On observe que le nouveau préfixe est ajouté à la table de routage des clients grâce au protocole RIPv2. Ce protocole a deux types de message (Figure 16) : le premier de type Request est envoyé en multicast pour demander à un routeur voisin d'envoyer sa table de routage, le second de type Response correspond à l'envoi de la table de routage suite à un message de type Request.

Time	Source	Destination	Protocol	Length	Info
112.023179	10.0.0.46	224.0.0.9	RIPv2	66	Request
112.035378	10.0.0.45	10.0.0.46	RIPv2	66	Response

Figure 16 – Capture de paquets sur le lien CE2VPNY-PE3 avec Wireshark

Grâce au message de type Response (Figure 17), le nouveau préfixe 10.0.0.44/30 est ajouté à la table de routage de CE1VPNY.

```

► Internet Protocol Version 4, Src: 10.0.0.2, Dst: 224.0.0.9
► User Datagram Protocol, Src Port: 520, Dst Port: 520
▼ Routing Information Protocol
  Command: Response (2)
  Version: RIPv2 (2)
  ▼ IP Address: 10.0.0.44, Metric: 1
    Address Family: IP (2)
    Route Tag: 0
    IP Address: 10.0.0.44
    Netmask: 255.255.255.252
    Next Hop: 0.0.0.0
    Metric: 1

```

Figure 17 – Capture de paquets sur le lien CE1VPNY-PE1 avec Wireshark

Ces messages transitent par le backbone MPLS. Cependant, nous n'avons pas réussi à retrouver ces messages.

Q5. La différence majeure entre la topologie Hub-and-Spoke et Mesh est que, pour la topologie Hub, chaque site client doit passer par le Hub (PE2 ici). Il y a donc plus de risques pour les clients de ne plus pouvoir communiquer entre eux si des liens tombent. Pour la topologie Mesh, au contraire, tous les routeurs sont reliés entre eux et un client n'est pas obligé de passer par un routeur en particulier donc il existe beaucoup plus de chemins.

Par exemple, si l'on désactive le lien entre PE1 et P1 en utilisant la commande `shutdown` sur l'interface `GigabitEthernet 2/0` de P1, on observe qu'un autre chemin, le plus court possible, est choisi automatiquement pour que CE1VPNY puisse communiquer avec CE2VPNY.


```

CE1VPNY#traceroute 10.0.0.46

Type escape sequence to abort.
Tracing the route to 10.0.0.46

 1 10.0.0.2 40 msec 68 msec 68 msec
 2 10.0.0.10 100 msec 100 msec 100 msec
 3 10.0.0.22 100 msec 104 msec 100 msec
 4 10.0.0.45 100 msec 100 msec 100 msec
 5 10.0.0.46 132 msec * 72 msec

```

```

CE1VPNY#traceroute 10.0.0.46

Type escape sequence to abort.
Tracing the route to 10.0.0.46

 1 10.0.0.2 68 msec 72 msec 68 msec
 2 10.0.0.6 132 msec 100 msec 100 msec
 3 10.0.0.14 88 msec 136 msec 132 msec
 4 10.0.0.22 128 msec 100 msec 132 msec
 5 10.0.0.45 100 msec 132 msec 132 msec
 6 10.0.0.46 160 msec * 84 msec

```

Figure 18 – Traceroute de CE1VPNY vers CE2VPNY (avant et après le shutdown)

La topologie Mesh est plus optimale et plus tolérante aux pannes mais elle est aussi plus coûteuse contrairement à la topologie Hub-and-Spoke. En effet, si un point d'accès tombe en panne, le trafic peut être acheminé via beaucoup d'autres chemins qui peuvent également être mis en place rapidement.

Cependant, les réseaux qui utilisent la topologie Mesh ne sont pas aussi transparents que ceux utilisant la topologie Hub-and-Spoke puisque cette dernière a un avantage : on peut connaître tout le trafic qui passe en analysant uniquement les routeurs Hub, ce qui accroît la sécurité.