Implémentation IOS de la fonctionnalité PE-CE iBGP

Contenu

Introduction

Ce document décrit comment la fonctionnalité iBGP (Internal Border Gateway Protocol) entre Provider Edge (PE) et Customer Edge (CE) est implémentée dans Cisco IOS®.

Informations générales

Jusqu'à la nouvelle fonctionnalité PE-CE iBGP, l'iBGP entre PE et CE (donc sur une interface VRF (Virtual Routing and Forwarding) sur le routeur PE) n'était pas officiellement pris en charge. Une exception est iBGP sur les interfaces VRF dans une configuration Multi-VRF CE (VRF-Lite). La motivation du déploiement de cette fonctionnalité est la suivante :

- Le client souhaite disposer d'un seul numéro de système autonome (ASN) sur les sites multiples du VRF, sans le déploiement du protocole eBGP (External Border Gateway Protocol) avec as-override.

- Le client souhaite fournir une réflexion de route interne vers les routeurs CE, agissant comme si le coeur du fournisseur de services (SP) était un réflecteur de route transparent (RR).

Avec cette fonctionnalité, les sites du VRF peuvent avoir le même ASN que le coeur du SP. Cependant, si l'ASN des sites VRF est différent de l'ASN du coeur du SP, il peut être fait qu'il

apparaisse de la même manière avec l'utilisation de la fonction système autonome local (AS).

Implémenter iBGP PE-CE

Voici les deux parties principales afin de faire fonctionner cette fonctionnalité :

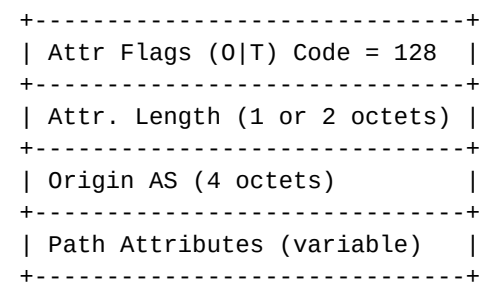
- Un nouvel attribut ATTR_SET a été ajouté au protocole BGP afin de transporter les attributs VPN BGP sur le coeur du SP de manière transparente.
- Faites du routeur PE un RR pour les sessions iBGP vers les routeurs CE dans le VRF et en tant que RR vers les voisins VPNv4 (autres routeurs PE ou RR).

Le nouvel attribut ATTR_SET permet au SP de transporter tous les attributs BGP du client de manière transparente et n'interfère pas avec les attributs SP et les stratégies BGP. Ces attributs sont la liste des clusters, les préférences locales, les communautés, etc.

Attribut de route client BGP

ATTR_SET est le nouvel attribut BGP utilisé pour transporter les attributs VPN BGP du client SP. Il s'agit d'un attribut transitoire facultatif. Dans cet attribut, tous les attributs BGP du client du message de mise à jour BGP, à l'exception des attributs MP_REACH et MP_UNREACH, peuvent être transportés.

L'attribut ATTR_SET a le format suivant :



Les indicateurs d'attribut sont les indicateurs d'attribut BGP standard (reportez-vous à la RFC 4271). La longueur d'attribut indique si la longueur d'attribut est d'un ou deux octets. L'objectif du champ AS d'origine est d'empêcher la fuite d'une route provenant d'un AS vers un autre AS sans la manipulation appropriée de l'AS_PATH. Le champ Path Attributes de longueur variable porte les attributs VPN BGP qui doivent être transportés sur le coeur du SP.

Sur le routeur PE de sortie, les attributs VPN BGP sont poussés dans cet attribut. Sur le routeur PE d'entrée, ces attributs sont affichés à partir de l'attribut, avant que le préfixe BGP ne soit envoyé au routeur CE. Cet attribut permet d'isoler les attributs BGP entre le réseau SP et le VPN client et vice versa. Par exemple, l'attribut de liste de cluster de réflexion de route SP n'est pas vu et pris en compte dans le réseau VPN. Mais également, l'attribut de liste de cluster de réflexion de route VPN n'est pas vu et pris en compte dans le réseau SP.

Regardez la Figure 1 afin de voir la propagation d'un préfixe BGP client sur le réseau SP.

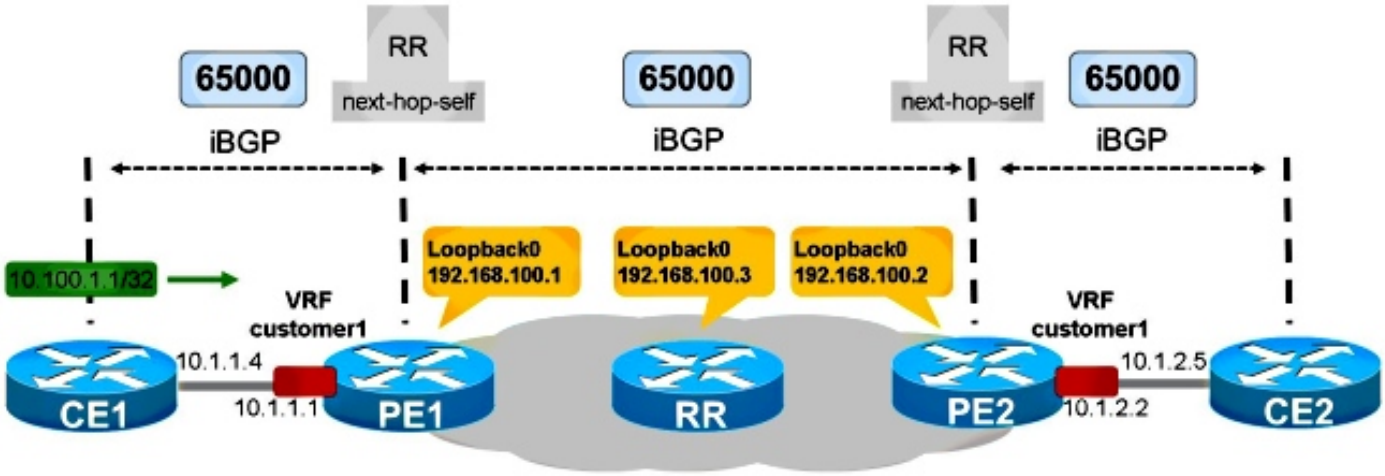


Figure 1

CE1 et CE2 se trouvent dans le même système autonome que le réseau SP : 65000. L'iBGP de PE1 est configuré vers CE1. PE1 reflète le chemin du préfixe 10.100.1.1/32 vers le RR dans le réseau SP. Le RR reflète le chemin iBGP vers les routeurs PE comme d'habitude. PE2 reflète le chemin vers CE2.

Pour que cela fonctionne correctement, vous devez :

- Ont du code sur PE1 et PE2 qui a la prise en charge de la fonctionnalité PE-CE iBGP
- Configurez PE1 et PE2 afin d'effectuer une réflexion de route sur leur session BGP vers leurs routeurs CE respectifs
- Disposer du saut suivant sur les routeurs PE pour la session BGP vers leurs routeurs CE
- Assurez-vous que chaque site VPN utilise des identificateurs de route (RD) différents.

Configuration

Reportez-vous à la figure 1.

Voici la configuration requise pour PE1 et PE2 :

```
PE1

vrf definition customer1
rd 65000:1
route-target export 1:1
route-target import 1:1
!
address-family ipv4
exit-address-family

router bgp 65000
bgp log-neighbor-changes
neighbor 192.168.100.3 remote-as 65000
neighbor 192.168.100.3 update-source Loopback0
```

```
!
address-family vpnv4
 neighbor 192.168.100.3 activate
 neighbor 192.168.100.3 send-community extended
exit-address-family
!
address-family ipv4 vrf customer1
 neighbor 10.1.1.4 remote-as 65000
 neighbor 10.1.1.4 activate
 neighbor 10.1.1.4 internal-vpn-client
 neighbor 10.1.1.4 route-reflector-client
 neighbor 10.1.1.4 next-hop-self
exit-address-family

PE2

vrf definition customer1
 rd 65000:2
 route-target export 1:1
 route-target import 1:1
 !
 address-family ipv4
 exit-address-family

router bgp 65000
 bgp log-neighbor-changes
 neighbor 192.168.100.3 remote-as 65000
 neighbor 192.168.100.3 update-source Loopback0
 !
 address-family vpnv4
 neighbor 192.168.100.3 activate
 neighbor 192.168.100.3 send-community extended
 exit-address-family
 !
 address-family ipv4 vrf customer1
 neighbor 10.1.2.5 remote-as 65000
 neighbor 10.1.2.5 activate
 neighbor 10.1.2.5 internal-vpn-client
 neighbor 10.1.2.5 route-reflector-client
 neighbor 10.1.2.5 next-hop-self
 exit-address-family
```

> **Note**: Si le PE ne dispose pas de la commande **neighbor <internal-CE> internal-vpn-client** pour le voisin CE, il ne propage pas les préfixes du CE vers les routeurs SP RR/PE.

> **Note**: Si le PE n'est pas le RR dans le VRF, il ne propage pas les préfixes des routeurs RR/PE vers le routeur CE.

Nouvelle commande

Il existe une nouvelle commande, **neighbor <internal-CE> internal-vpn-client**, pour faire fonctionner cette fonctionnalité. Il doit être configuré sur le routeur PE uniquement pour la session iBGP vers les routeurs CE.

> **Note**: La fonctionnalité iBGP PE-CE Multi-VRF CE (VRF-Lite) est toujours prise en charge sans la commande **neighbor <internal-CE> internal-vpn-client**.

Note: Lorsque la commande **neighbor <internal-CE> internal-vpn-client** est configurée, les commandes **neighbor <internal-CE> route-reflector-client** et **neighbor <internal-CE> next-hop-self** sont automatiquement placées dans la configuration. Lorsque l'une des commandes **neighbor <internal-CE> route-reflector-client** et **neighbor <internal-CE> next-hop-self** (ou les deux) sont supprimées et qu'un rechargement est effectué, elles sont automatiquement replacées dans la configuration.

Détail de ATTR_SET

Reportez-vous à la figure 1.

Il s'agit du préfixe annoncé par CE1 :

```
CE1#show bgp ipv4 unicast 10.100.1.1/32
BGP routing table entry for 10.100.1.1/32, version 2
Paths: (1 available, best #1, table default)
 Advertised to update-groups:
    4
 Refresh Epoch 1
 Local
   0.0.0.0 from 0.0.0.0 (10.100.1.1)
     Origin IGP, metric 0, localpref 100, weight 32768, valid, sourced, local, best
     rx pathid: 0, tx pathid: 0x0
```

Lorsque PE1 reçoit le préfixe BGP 10.100.1.1/32 de CE1, il le stocke deux fois :

```
PE1#show bgp vpnv4 unicast all 10.100.1.1/32
BGP routing table entry for 65000:1:10.100.1.1/32, version 21
Paths: (2 available, best #1, table customer1)
 Advertised to update-groups:
    5
 Refresh Epoch 1
 Local, (Received from ibgp-pece RR-client)
   10.1.1.4 (via vrf customer1) from 10.1.1.4 (10.100.1.1)
     Origin IGP, metric 0, localpref 200, valid, internal, best
     mpls labels in/out 18/nolabel
     rx pathid: 0, tx pathid: 0x0
 Refresh Epoch 1
 Local, (Received from ibgp-pece RR-client), (ibgp sourced)
   10.1.1.4 (via vrf customer1) from 10.1.1.4 (10.100.1.1)
     Origin IGP, localpref 100, valid, internal
     Extended Community: RT:1:1
     mpls labels in/out 18/nolabel
     rx pathid: 0, tx pathid: 0
```

Le premier chemin est le chemin réel sur PE1, car il est reçu de CE1.

Le deuxième chemin est le chemin annoncé vers les routeurs RR/PE. Il est marqué avec **ibgp source**. Il contient l'attribut ATTR_SET. Notez que ce chemin comporte une ou plusieurs cibles de routage (RT).

PE1 annonce le préfixe comme indiqué ici :

```
PE1#show bgp vpnv4 unicast all neighbors 192.168.100.3 advertised-routes
BGP table version is 7, local router ID is 192.168.100.1
```

```
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
              r RIB-failure, S Stale, m multipath, b backup-path, f RT-Filter,
              x best-external, a additional-path, c RIB-compressed,
Origin codes: i - IGP, e - EGP, ? - incomplete
RPKI validation codes: V valid, I invalid, N Not found

    Network          Next Hop            Metric LocPrf Weight Path
Route Distinguisher: 65000:1 (default for vrf customer1)
 *>i 10.100.1.1/32    10.1.1.4                 0    200      0 i

Total number of prefixes 1
```

Voici comment le RR voit le chemin :

```
RR#show bgp vpnv4 un all 10.100.1.1/32
BGP routing table entry for 65000:1:10.100.1.1/32, version 10
Paths: (1 available, best #1, no table)
 Advertised to update-groups:
    3
 Refresh Epoch 1
 Local, (Received from a RR-client)
   192.168.100.1 (metric 11) (via default) from 192.168.100.1 (192.168.100.1)
     Origin IGP, localpref 100, valid, internal, best
     Extended Community: RT:1:1
     Originator: 10.100.1.1, Cluster list: 192.168.100.1
      ATTR_SET Attribute:
       Originator AS 65000
       Origin IGP
       Aspath
       Med 0
       LocalPref 200
       Cluster list
       192.168.100.1,
       Originator 10.100.1.1
     mpls labels in/out nolabel/18
     rx pathid: 0, tx pathid: 0x0
```

Notez que la préférence locale de ce préfixe de monodiffusion VPNv4 dans le coeur est 100. Dans ATTR_SET, la préférence locale d'origine de 200 est stockée. Toutefois, cela est transparent pour le RR dans le coeur du SP.

Sur PE2, vous voyez le préfixe comme indiqué ici :

```
PE2#show bgp vpnv4 unicast all 10.100.1.1/32
BGP routing table entry for 65000:1:10.100.1.1/32, version 5
Paths: (1 available, best #1, no table)
 Not advertised to any peer
 Refresh Epoch 2
 Local
   192.168.100.1 (metric 21) (via default) from 192.168.100.3 (192.168.100.3)
     Origin IGP, localpref 100, valid, internal, best
     Extended Community: RT:1:1
     Originator: 10.100.1.1, Cluster list: 192.168.100.3, 192.168.100.1
     ATTR_SET Attribute:
       Originator AS 65000
       Origin IGP
       Aspath
       Med 0
       LocalPref 200
       Cluster list
       192.168.100.1,
```

```
       Originator 10.100.1.1
     mpls labels in/out nolabel/18
     rx pathid: 0, tx pathid: 0x0
BGP routing table entry for 65000:2:10.100.1.1/32, version 6
Paths: (1 available, best #1, table customer1)
 Advertised to update-groups:
    1
 Refresh Epoch 2
 Local, imported path from 65000:1:10.100.1.1/32 (global)
   192.168.100.1 (metric 21) (via default) from 192.168.100.3 (192.168.100.3)
     Origin IGP, metric 0, localpref 200, valid, internal, best
     Originator AS(ibgp-pece): 65000
     Originator: 10.100.1.1, Cluster list: 192.168.100.1
     mpls labels in/out nolabel/18
     rx pathid:0, tx pathid: 0x0
```

Le premier chemin est celui reçu du RR, avec ATTR_SET. Notez que la RD est 65000:1, la RD d'origine. Le deuxième chemin est le chemin importé de la table VRF avec RD 65000:1. ATTR_SET a été supprimé.

Voici le chemin tel qu'il apparaît sur CE2 :

```
CE2#show bgp ipv4 unicast 10.100.1.1/32
BGP routing table entry for 10.100.1.1/32, version 10
Paths: (1 available, best #1, table default)
 Not advertised to any peer
 Refresh Epoch 1
 Local
   10.1.2.2 from 10.1.2.2 (192.168.100.2)
     Origin IGP, metric 0, localpref 200, valid, internal, best
     Originator: 10.100.1.1, Cluster list: 192.168.100.2, 192.168.100.1
     rx pathid: 0, tx pathid: 0x0
```

Notez que le saut suivant est 10.1.2.2, qui est PE2. La liste des clusters contient les routeurs PE1 et PE2. Voici les RR qui comptent dans le VPN. Le RR SP (10.100.1.3) ne figure pas dans la liste des clusters.

La préférence locale de 200 a été conservée à l'intérieur du VPN sur le réseau SP.

La commande debug bgp vpnv4 unicast update affiche la mise à jour propagée dans le réseau SP :

```
PE1#
BGP(4): Revise route installing 1 of 1 routes for 10.100.1.1/32 -> 10.1.1.4
(customer1) to customer1 IP table
BGP(4): 192.168.100.3 NEXT_HOP changed SELF for ibgp rr-client pe-ce net
65000:1:10.100.1.1/32,
BGP(4): 192.168.100.3 Net 65000:1:10.100.1.1/32 from ibgp-pece 10.1.1.4 format
ATTR_SET
BGP(4): (base) 192.168.100.3 send UPDATE (format) 65000:1:10.100.1.1/32, next
192.168.100.1, label 16, metric 0, path Local, extended community RT:1:1
BGP: 192.168.100.3 Next hop is our own address 192.168.100.1
BGP: 192.168.100.3 Route Reflector cluster loop; Received cluster-id 192.168.100.1
BGP: 192.168.100.3 RR in same cluster. Reflected update dropped

RR#
BGP(4): 192.168.100.1 rcvd UPDATE w/ attr: nexthop 192.168.100.1, origin i, localpref
100, originator 10.100.1.1, clusterlist 192.168.100.1, extended community RT:1:1,
[ATTR_SET attribute:  originator AS 65000, origin IGP, aspath , med 0, localpref 200,
cluster list 192.168.100.1 , originator 10.100.1.1]
```

```
BGP(4): 192.168.100.1 rcvd 65000:1:10.100.1.1/32, label 16
RT address family is not configured. Can't create RTC route
BGP(4): (base) 192.168.100.1 send UPDATE (format) 65000:1:10.100.1.1/32, next
192.168.100.1, label 16, metric 0, path Local, extended community RT:1:1

PE2#
BGP(4): 192.168.100.3 rcvd UPDATE w/ attr: nexthop 192.168.100.1, origin i, localpref
100, originator 10.100.1.1, clusterlist 192.168.100.3 192.168.100.1, extended community
RT:1:1, [ATTR_SET attribute:  originator AS 65000, origin IGP, aspath , med 0, localpref
200, cluster list 192.168.100.1 , originator 10.100.1.1]
BGP(4): 192.168.100.3 rcvd 65000:1:10.100.1.1/32, label 16
RT address family is not configured. Can't create RTC route
BGP(4): Revise route installing 1 of 1 routes for 10.100.1.1/32 -> 192.168.100.1
(customer1) to customer1 IP table
BGP(4): 10.1.2.5 NEXT_HOP is set to self for net 65000:2:10.100.1.1/32,
```

Note: PE1 a reçu sa propre mise à jour de RR, puis l'a supprimée. En effet, PE1 et PE2 font partie du même groupe de mise à jour sur RR.

Note: Si vous voulez vider le message Update complet au format hexadécimal, utilisez le mot clé **detail** pour la commande **debug BGP update**.

```
PE2#   debug bgp vpnv4 unicast updates detail
BGP updates debugging is on with detail for address family: VPNv4 Unicast

PE2#
BGP(4): 192.168.100.3 rcvd UPDATE w/ attr: nexthop 192.168.100.1, origin i,
localpref 100, originator 10.100.1.1, clusterlist 192.168.100.3 192.168.100.1,
extended community RT:1:1, [ATTR_SET attribute:  originator AS 65000, origin IGP,
aspath , med 0, localpref 200, cluster list 192.168.100.1 , originator 10.100.1.1]
BGP(4): 192.168.100.3 rcvd 65000:1:10.100.1.1/32, label 17
RT address family is not configured. Can't create RTC route
BGP: 192.168.100.3 rcv update length 125
BGP: 192.168.100.3 rcv update dump: FFFF FFFF FFFF FFFF FFFF FFFF FFFF FFFF
0090 0200 00
PE2#00 7980 0E21 0001 800C 0000 0000 0000 0000 C0A8 6401 0078 0001 1100 00FD E800
0000 010A 6401 0140 0101 0040 0200 4005 0400 0000 64C0 1008 0002 0001 0000 0001 800A
08C0 A864 03C0 A864 0180 0904 0A64 0101 C080 2700 00FD E840 0101 0040 0200 8004 0400
0000 0040 0504 0000 00C8 800A 04C0 A864 0180 0904 0A64 0101
BGP(4): Revise route installing 1 of 1 routes for 10.100.1.1/32 -> 192.168.100.1
(customer1) to customer1 IP table
BGP(4): 10.1.2.5 NEXT_HOP is set to self for net 65000:2:10.100.1.1/32,
```

Gestion du tronçon suivant

Le saut suivant doit être configuré sur les routeurs PE pour cette fonctionnalité. La raison en est que normalement le tronçon suivant est transporté sans modification avec iBGP. Cependant, il existe deux réseaux distincts : le réseau VPN et le réseau SP, qui exécutent des protocoles IGP (Interior Gateway Protocol) distincts. Par conséquent, la métrique IGP ne peut pas être facilement comparée et utilisée pour le calcul du meilleur chemin entre les deux réseaux. L'approche choisie par la RFC 6368 consiste à rendre obligatoire le prochain saut pour la session iBGP vers le CE, ce qui évite le problème décrit précédemment dans son ensemble. Un avantage est que les sites VRF peuvent exécuter différents IGP avec cette approche.

RD

Le document RFC 6368 indique qu'il est recommandé que différents sites VRF du même VPN utilisent des RD différentes (uniques). Dans Cisco IOS, ceci est obligatoire pour cette fonctionnalité.

Fonctionnalité PE-CE iBGP avec Local-AS

Reportez-vous à la figure 2. Le client VPN 1 a ASN 65001.

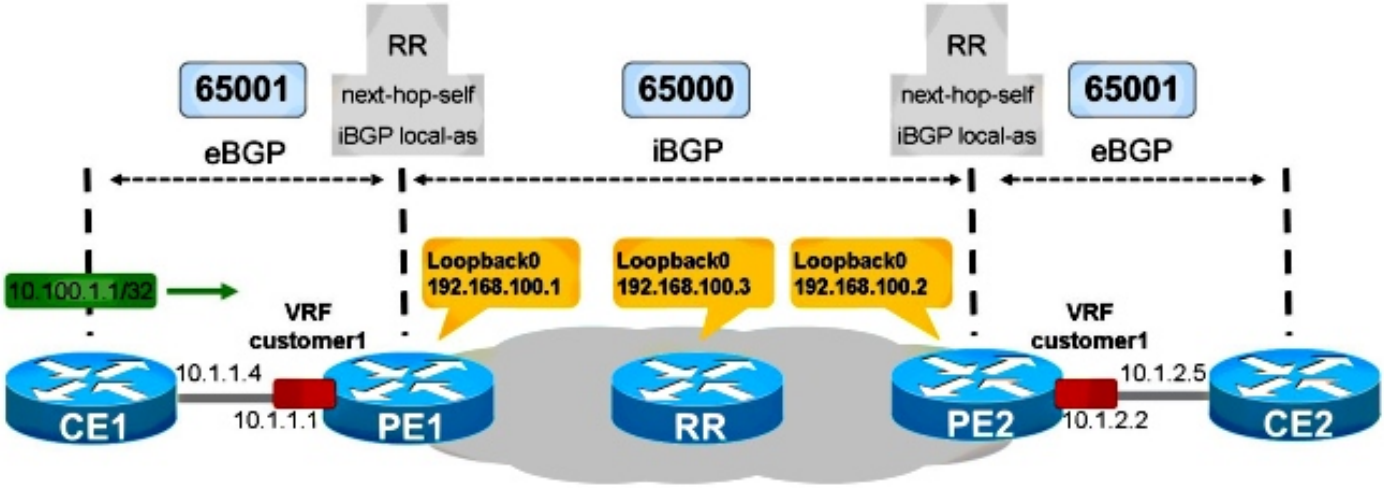


Figure 2

CE1 est dans AS 65001. Pour faire de ce BGP interne du point de vue de PE1, il a besoin de la fonctionnalité locale-as iBGP.

```
CE1

router bgp 65001
 bgp log-neighbor-changes
 network 10.100.1.1 mask 255.255.255.255
 neighbor 10.1.1.1 remote-as 65001

PE1

router bgp 65000
 bgp log-neighbor-changes
 neighbor 192.168.100.3 remote-as 65000
 neighbor 192.168.100.3 update-source Loopback0
 !
 address-family vpnv4
 neighbor 192.168.100.3 activate
 neighbor 192.168.100.3 send-community extended
 exit-address-family
 !
 address-family ipv4 vrf customer1
 neighbor 10.1.1.4 remote-as 65001
 neighbor 10.1.1.4 local-as 65001
 neighbor 10.1.1.4 activate
 neighbor 10.1.1.4 internal-vpn-client
 neighbor 10.1.1.4 route-reflector-client
 neighbor 10.1.1.4 next-hop-self
 exit-address-family
```

PE2 et CE2 sont configurés de la même manière.

PE1 voit le préfixe BGP comme indiqué ici :

```
PE1#show bgp vpnv4 unicast all 10.100.1.1/32
BGP routing table entry for 65000:1:10.100.1.1/32, version 41
Paths: (2 available, best #1, table customer1)
 Advertised to update-groups:
    5
 Refresh Epoch 1
 Local, (Received from ibgp-pece RR-client)
   10.1.1.4 (via vrf customer1) from 10.1.1.4 (10.100.1.1)
     Origin IGP, metric 0, localpref 200, valid, internal, best
     mpls labels in/out 18/nolabel
     rx pathid: 0, tx pathid: 0x0
 Refresh Epoch 1
 Local, (Received from ibgp-pece RR-client), (ibgp sourced)
   10.1.1.4 (via vrf customer1) from 10.1.1.4 (10.100.1.1)
     Origin IGP, localpref 100, valid, internal
     Extended Community: RT:1:1
     mpls labels in/out 18/nolabel
     rx pathid: 0, tx pathid: 0
```

Le préfixe est un BGP interne.

PE2 voit ceci :

```
PE2#show bgp vpnv4 unicast all 10.100.1.1/32
BGP routing table entry for 65000:1:10.100.1.1/32, version 33
Paths: (1 available, best #1, no table)
 Not advertised to any peer
 Refresh Epoch 5
 Local
   192.168.100.1 (metric 21) (via default) from 192.168.100.3 (192.168.100.3)
     Origin IGP, localpref 100, valid, internal, best
     Extended Community: RT:1:1
     Originator: 10.100.1.1, Cluster list: 192.168.100.3, 192.168.100.1
     ATTR_SET Attribute:
       Originator AS 65001
       Origin IGP
       Aspath
       Med 0
       LocalPref 200
       Cluster list
       192.168.100.1,
       Originator 10.100.1.1
     mpls labels in/out nolabel/18
     rx pathid: 0, tx pathid: 0x0
BGP routing table entry for 65000:2:10.100.1.1/32, version 34
Paths: (1 available, best #1, table customer1)
 Advertised to update-groups:
    5
 Refresh Epoch 2
 Local, imported path from 65000:1:10.100.1.1/32 (global)
   192.168.100.1 (metric 21) (via default) from 192.168.100.3 (192.168.100.3)
     Origin IGP, metric 0, localpref 200, valid, internal, best
     Originator AS(ibgp-pece): 65001
     Originator: 10.100.1.1, Cluster list: 192.168.100.1
     mpls labels in/out nolabel/18
     rx pathid: 0, tx pathid: 0x0
```

L'AS d'origine est 65001, qui est le AS utilisé lorsque le préfixe est envoyé de PE2 à CE2. Ainsi, le système autonome est préservé, tout comme la préférence locale dans cet exemple.

```
CE2#show bgp ipv4 unicast 10.100.1.1/32
BGP routing table entry for 10.100.1.1/32, version 3
Paths: (1 available, best #1, table default)
 Not advertised to any peer
 Refresh Epoch 1
 Local
   10.1.2.2 from 10.1.2.2 (192.168.100.2)
     Origin IGP, metric 0, localpref 200, valid, internal, best
     Originator: 10.100.1.1, Cluster list: 192.168.100.2, 192.168.100.1
     rx pathid: 0, tx pathid: 0x0
```

Vous voyez **Local** au lieu d'un chemin AS. Cela signifie qu'il s'agit d'une route BGP interne provenant de l'AS 65001, qui est également l'ASN configuré du routeur CE2. Tous les attributs BGP ont été extraits de l'attribut ATTR_SET. Cela est conforme aux règles de l'affaire 1 dans la section suivante.

Règles pour l'échange de routes entre différents sites VRF

ATTR_SET contient l'AS d'origine du VRF d'origine. Cet AS d'origine est vérifié par le PE distant, lorsqu'il supprime l'ATTR_SET avant d'envoyer le préfixe au routeur CE.

Cas 1 : Si l'AS d'origine correspond à l'AS configuré pour le routeur CE, les attributs BGP sont extraits de l'attribut ATTR_SET lorsque le PE importe le chemin dans le VRF de destination.

Cas 2 : Si le système autonome d'origine ne correspond pas au système autonome configuré pour le routeur CE, l'ensemble des attributs du chemin construit est pris comme indiqué ici :

1. Les attributs de chemin sont définis sur les attributs contenus dans l'attribut ATTR_SET.

2. Les attributs spécifiques à iBGP sont ignorés (LOCAL_PREF, ORIGINATOR et CLUSTER_LIST).

3. Le numéro **AS** d'**origine** contenu dans l'attribut ATTR_SET est précédé de AS_PATH et suit les règles qui s'appliquent à un appairage BGP externe entre les AS source et de destination.

4. Si le système autonome associé au VRF est le même que le système autonome du fournisseur VPN et que l'attribut AS_PATH de la route VPN n'est pas vide, il DOIT être précédé de l'attribut AS_PATH de la route VRF.

   Reportez-vous à la figure 3. CE1 et PE1 ont l'AS 65000 et sont configurés avec la fonctionnalité PE-CE iBGP. CE2 a ASN 65001. Cela signifie qu'il existe eBGP entre PE2 et CE2.

Figure 3

PE2 voit la route comme suit :

```
PE2#show bgp vpnv4 unicast all 10.100.1.1/32
BGP routing table entry for 65000:1:10.100.1.1/32, version 43
Paths: (1 available, best #1, no table)
 Not advertised to any peer
 Refresh Epoch 6
 Local
   192.168.100.1 (metric 21) (via default) from 192.168.100.3 (192.168.100.3)
     Origin IGP, localpref 100, valid, internal, best
     Extended Community: RT:1:1
     Originator: 10.100.1.1, Cluster list: 192.168.100.3, 192.168.100.1
     ATTR_SET Attribute:
       Originator AS 65000
       Origin IGP
       Aspath
       Med 0
       LocalPref 200
       Cluster list
       192.168.100.1,
       Originator 10.100.1.1
     mpls labels in/out nolabel/17
     rx pathid: 0, tx pathid: 0x0
BGP routing table entry for 65000:2:10.100.1.1/32, version 44
Paths: (1 available, best #1, table customer1)
 Advertised to update-groups:
    6
 Refresh Epoch 6
 Local, imported path from 65000:1:10.100.1.1/32 (global)
   192.168.100.1 (metric 21) (via default) from 192.168.100.3 (192.168.100.3)
     Origin IGP, metric 0, localpref 200, valid, internal, best
     Originator AS(ibgp-pece): 65000
     Originator: 10.100.1.1, Cluster list: 192.168.100.1
     mpls labels in/out nolabel/17
     rx pathid: 0, tx pathid: 0x0
```

Voici le préfixe tel qu'il apparaît sur CE2 :

```
CE2#show bgp ipv4 unicast 10.100.1.1/32
BGP routing table entry for 10.100.1.1/32, version 5
Paths: (1 available, best #1, table default)
 Not advertised to any peer
 Refresh Epoch 1
 65000
   10.1.2.2 from 10.1.2.2 (192.168.100.2)
     Origin IGP, localpref 100, valid, external, best
     rx pathid: 0, tx pathid: 0x0
```

C'est le cas 2. Le numéro **AS d'origine** contenu dans l'attribut ATTR_SET est précédé de AS_PATH par PE2 et suit les règles qui s'appliquent à un appairage eBGP entre l'AS source et le AS de destination. Les attributs spécifiques à iBGP sont ignorés par PE2 lorsqu'il crée la route à annoncer à CE2. Ainsi, la préférence locale est 100 et non 200 (comme le montre l'attribut ATTR_SET).

Réflexion VRF-Lite CE à CE

Reportez-vous à la figure 4.

Figure 4

La Figure 4 présente un routeur CE supplémentaire, CE3, connecté à PE1. CE1 et CE3 sont tous deux connectés à PE1 sur la même instance VRF : client1. Cela signifie que CE1 et CE3 sont des routeurs CE Multi-VRF (également appelés VRF-Lite) de PE1. PE1 se place comme tronçon suivant lorsqu'il annonce les préfixes de CE1 à CE3. Si ce comportement n'est pas souhaité, vous pouvez configurer le **voisin 10.1.3.6 prochain saut inchangé** sur PE1. Pour configurer ceci, vous devez supprimer **le voisin 10.1.3.6 prochain saut-self** sur PE1. Ensuite, CE3 voit les routes de CE1 avec CE1 comme tronçon suivant pour ces préfixes BGP. Pour que cela fonctionne, vous avez besoin des routes pour ces prochains sauts BGP dans la table de routage de CE3. Vous avez besoin d'un protocole de routage dynamique (IGP) ou de routes statiques sur CE1, PE1 et CE3 afin de vous assurer que les routeurs ont une route pour les adresses IP de tronçon suivant les uns des autres. Cependant, cette configuration présente un problème.

La configuration de PE1 est la suivante :

```
router bgp 65000
 !
 address-family ipv4 vrf customer1
 neighbor 10.1.1.4 remote-as 65000
 neighbor 10.1.1.4 activate
 neighbor 10.1.1.4 internal-vpn-client
 neighbor 10.1.1.4 route-reflector-client
 neighbor 10.1.1.4 next-hop-self
 neighbor 10.1.3.6 remote-as 65000
 neighbor 10.1.3.6 activate
 neighbor 10.1.3.6 internal-vpn-client
 neighbor 10.1.3.6 route-reflector-client
 neighbor 10.1.3.6 next-hop-unchanged
 exit-address-family
```

Le préfixe de CE1 s'affiche correctement sur CE3 :

```
CE3#show bgp ipv4 unicast 10.100.1.1
BGP routing table entry for 10.100.1.1/32, version 9
Paths: (1 available, best #1, table default)
 Not advertised to any peer
 Refresh Epoch 1
 Local
   10.1.1.4 from 10.1.3.1 (192.168.100.1)
     Origin IGP, metric 0, localpref 200, valid, internal, best
     Originator: 10.100.1.1, Cluster list: 192.168.100.1
     rx pathid: 0, tx pathid: 0x0
```

Cependant, le préfixe de CE2 est visible sur CE3 comme indiqué ici :

```
CE3#show bgp ipv4 unicast 10.100.1.2
BGP routing table entry for 10.100.1.2/32, version 0
Paths: (1 available, no best path)
 Not advertised to any peer
 Refresh Epoch 1
 Local
   192.168.100.2 (inaccessible) from 10.1.3.1 (192.168.100.1)
     Origin IGP, metric 0, localpref 100, valid, internal
     Originator: 10.100.1.2, Cluster list: 192.168.100.1, 192.168.100.2
     rx pathid: 0, tx pathid: 0
```

Le tronçon suivant BGP est **192.168.100.2**, l'adresse IP de bouclage de PE2. PE1 n'a pas réécrit

le tronçon suivant BGP à lui-même lorsqu'il a annoncé le préfixe 10.100.1.2/32 à CE3. Ce préfixe est donc inutilisable sur CE3.

Ainsi, dans le cas d'une combinaison de la fonctionnalité PE-CE iBGP sur MPLS-VPN et iBGP VRF-Lite, vous devez vous assurer que vous avez toujours le saut suivant sur les routeurs PE.

Vous ne pouvez pas conserver le saut suivant lorsqu'un routeur PE est un RR qui reflète les routes iBGP d'un CE à un autre CE à travers les interfaces VRF localement sur le PE. Lorsque vous exécutez iBGP PE-CE sur un réseau VPN MPLS, vous devez utiliser **internal-vpn-client** pour les sessions iBGP vers les routeurs CE. Lorsque vous avez plusieurs CE locaux dans un VRF sur un routeur PE, vous devez conserver le **saut suivant** pour ces homologues BGP.

Vous pouvez regarder les routes-maps afin de définir le saut suivant en auto pour les préfixes reçus des autres routeurs PE, mais pas pour les préfixes réfléchis des autres routeurs CE connectés localement. Cependant, il n'est pas actuellement pris en charge de définir le saut suivant sur auto dans une route-map sortante. Cette configuration est présentée ici :

```
router bgp 65000

 address-family ipv4 vrf customer1
 neighbor 10.1.1.4 remote-as 65000
 neighbor 10.1.1.4 activate
 neighbor 10.1.1.4 internal-vpn-client
 neighbor 10.1.1.4 route-reflector-client
 neighbor 10.1.1.4 next-hop-self
 neighbor 10.1.3.6 remote-as 65000
 neighbor 10.1.3.6 activate
 neighbor 10.1.3.6 internal-vpn-client
 neighbor 10.1.3.6 route-reflector-client
 neighbor 10.1.3.6 route-map NH-setting out
 exit-address-family

ip prefix-list PE-loopbacks seq 10 permit 192.168.100.0/24 ge 32
!

route-map NH-setting permit 10
 description set next-hop to self for prefixes from other PE routers
 match ip route-source prefix-list PE-loopbacks
 set ip next-hop self
!

route-map NH-setting permit 20
 description advertise prefixes with next-hop other than the prefix-list in
route-map entry 10 above
!
```

Cependant, ceci n'est pas pris en charge :

```
PE1(config)#route-map NH-setting permit 10
PE1(config-route-map)# set ip next-hop self
% "NH-setting" used as BGP outbound route-map, set use own IP/IPv6 address for the nexthop not
supported
```

Cisco IOS plus ancien sur le routeur PE

Si PE1 exécute un logiciel Cisco IOS plus ancien qui ne possède pas la fonctionnalité iBGP PE-CE, PE1 ne se définit jamais comme le tronçon suivant des préfixes iBGP reflétés. Cela signifie

que le préfixe BGP réfléchi (10.100.1.1/32) de CE1 (10.100.1.1) à CE2 -via PE1- aurait le tronçon suivant CE1 (10.1.1.4).

```
CE3#show bgp ipv4 unicast 10.100.1.1
BGP routing table entry for 10.100.1.1/32, version 32
Paths: (1 available, best #1, table default)
 Not advertised to any peer
 Refresh Epoch 1
 Local
   10.1.1.4 from 10.1.3.1 (192.168.100.1)
     Origin IGP, metric 0, localpref 200, valid, internal, best
     Originator: 10.100.1.1, Cluster list: 192.168.100.1
     rx pathid: 0, tx pathid: 0x0
```

Le préfixe de CE2 (10.100.1.2/32) est vu avec PE2 comme tronçon suivant, car PE1 ne fait pas de tronçon suivant pour ce préfixe non plus :

```
CE3#show bgp ipv4 unicast 10.100.1.2
BGP routing table entry for 10.100.1.2/32, version 0
Paths: (1 available, no best path)
 Not advertised to any peer
 Refresh Epoch 1
 Local
   192.168.100.2 (inaccessible) from 10.1.3.1 (192.168.100.1)
     Origin IGP, localpref 100, valid, internal
     Originator: 10.100.1.2, Cluster list: 192.168.100.1, 192.168.100.3, 192.168.100.2
     ATTR_SET Attribute:
       Originator AS 65000
       Origin IGP
       Aspath
       Med 0
       LocalPref 100
       Cluster list
       192.168.100.2,
       Originator 10.100.1.2
     rx pathid: 0, tx pathid: 0
```

Pour que la fonctionnalité PE-CE iBGP fonctionne correctement, tous les routeurs PE pour le VPN sur lequel la fonctionnalité est activée doivent avoir le code pour prendre en charge la fonctionnalité et avoir cette fonctionnalité activée.

Prochain saut pour eBGP sur VRF

Reportez-vous à la figure 5.

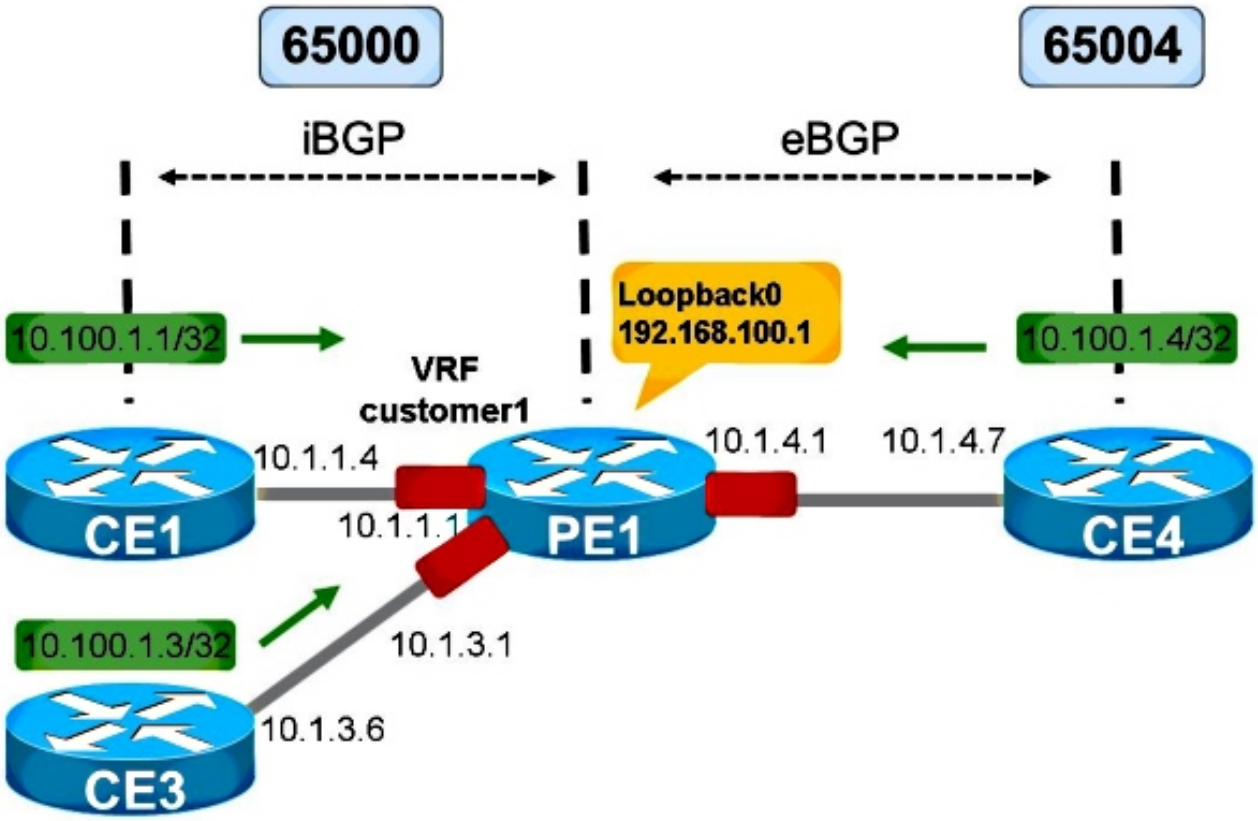


Figure 5

La Figure 5 illustre une configuration VRF-Lite. La session de PE1 vers CE4 est eBGP. La session de PE1 vers CE3 est toujours iBGP.

Pour les préfixes eBGP, le saut suivant est toujours défini sur auto lorsqu'il annonce les préfixes vers un voisin iBGP sur VRF. Ceci indépendamment du fait que la session vers le voisin iBGP à travers VRF a le saut suivant auto défini ou non.

Dans la Figure 5, CE3 voit les préfixes de CE4 avec PE1 comme tronçon suivant.

```
CE3#show bgp ipv4 unicast 10.100.1.4
BGP routing table entry for 10.100.1.4/32, version 103
Paths: (1 available, best #1, table default)
 Not advertised to any peer
 Refresh Epoch 1
 65004
   10.1.3.1 from 10.1.3.1 (192.168.100.1)
     Origin IGP, metric 0, localpref 100, valid, internal, best
     rx pathid: 0, tx pathid: 0x0
```

Ceci se produit avec le saut suivant sur PE1 vers CE3 ou sans.

Si les interfaces de PE1 vers CE3 et CE4 ne sont pas dans un VRF, mais dans le contexte global, le saut suivant vers CE3 fait une différence.

Sans le saut suivant sur PE1 vers CE3, vous voyez :

```
PE1#show bgp vrf customer1 vpnv4 unicast neighbors 10.1.3.6
```

```
BGP neighbor is 10.1.3.6,  vrf customer1,  remote AS 65000, internal link
...
 For address family: VPNv4 Unicast
 Translates address family IPv4 Unicast for VRF customer1
 Session: 10.1.3.6
 BGP table version 1, neighbor version 1/0
 Output queue size : 0
 Index 12, Advertise bit 0
 Route-Reflector Client
 12 update-group member
 Slow-peer detection is disabled
 Slow-peer split-update-group dynamic is disabled
 Interface associated: (none)
```

Bien que le prochain saut auto soit implicitement activé, le résultat n'indique pas ceci.

Avec le saut suivant sur PE1 vers CE3, vous voyez :

```
PE1#show bgp vrf customer1 vpnv4 unicast neighbors 10.1.3.6
BGP neighbor is 10.1.3.6,  vrf customer1,  remote AS 65000, internal link
..
 For address family: VPNv4 Unicast
...
 NEXT_HOP is always this router for eBGP paths
```

Alors que, si les interfaces vers CE3 et CE4 sont dans un contexte global, le tronçon suivant pour les préfixes de CE4 est le CE4 lui-même lorsque le tronçon suivant n'est pas configuré :

```
CE3#show bgp ipv4 unicast 10.100.1.4
BGP routing table entry for 10.100.1.4/32, version 124
Paths: (1 available, best #1, table default)
 Not advertised to any peer
 Refresh Epoch 1
 65004
   10.1.4.7 from 10.1.3.1 (192.168.100.1)
     Origin IGP, metric 0, localpref 100, valid, internal, best
     rx pathid: 0, tx pathid: 0x0
```

Pour le saut suivant sur PE1 vers CE3 :

```
CE3#show bgp ipv4 unicast 10.100.1.4
BGP routing table entry for 10.100.1.4/32, version 125
Paths: (1 available, best #1, table default)
 Not advertised to any peer
 Refresh Epoch 1
 65004
   10.1.3.1 from 10.1.3.1 (192.168.100.1)
     Origin IGP, metric 0, localpref 100, valid, internal, best
     rx pathid: 0, tx pathid: 0x0
```

Cela a été fait sur la base de la RFC 4364.

Si vous ne voulez pas définir le prochain saut pour les préfixes eBGP vers une session iBGP sur une interface VRF, vous devez configurer le **saut suivant inchangé**. La prise en charge de ce problème s'est produite uniquement avec l'ID de bogue Cisco CSCuj11720.

```
router bgp 65000
...
 address-family ipv4 vrf customer1
```

```
 neighbor 10.1.1.4 remote-as 65000
 neighbor 10.1.1.4 activate
 neighbor 10.1.1.4 route-reflector-client
 neighbor 10.1.3.6 remote-as 65000
 neighbor 10.1.3.6 activate
 neighbor 10.1.3.6 route-reflector-client
 neighbor 10.1.3.6 next-hop-unchanged
 neighbor 10.1.4.7 remote-as 65004
 neighbor 10.1.4.7 activate
 exit-address-family
```

Àprésent, CE3 voit CE4 comme le tronçon suivant des préfixes annoncés par CE4 :

```
CE3#show bgp ipv4 unicast 10.100.1.4
BGP routing table entry for 10.100.1.4/32, version 130
Paths: (1 available, best #1, table default)
 Not advertised to any peer
 Refresh Epoch 3
 65004
   10.1.4.7 from 10.1.3.1 (192.168.100.1)
     Origin IGP, metric 0, localpref 100, valid, internal, best
     rx pathid: 0, tx pathid: 0x0
```

Si vous essayez de configurer le mot clé du **prochain saut inchangé** pour la session iBGP vers CE3 sur le code Cisco IOS avant l'ID de bogue Cisco CSCuj11720, vous rencontrez cette erreur :

```
PE1(config-router-af)# neighbor 10.1.3.6 next-hop-unchanged
%BGP: Can propagate the nexthop only to multi-hop EBGP neighbor
```

Après l'ID de bogue Cisco CSCuj11720, le mot clé **next-hop-constant** est valide pour les voisins eBGP à sauts multiples et les voisins iBGP VRF-Lite.