

iBGP PE-CE功能的IOS實施

目錄

[簡介](#)

[背景資訊](#)

[實施iBGP PE-CE](#)

[BGP客戶路由屬性](#)

[設定](#)

[新指令](#)

[詳細檢視ATTR SET](#)

[下一跳處理](#)

[RD](#)

[使用本地AS的iBGP PE-CE功能](#)

[不同VRF站點之間的路由交換規則](#)

[CE-to-CE VRF-Lite反射](#)

[PE路由器上的舊版Cisco IOS](#)

[VRF上eBGP的Next-hop-self](#)

簡介

本檔案介紹如何在Cisco IOS®中實作提供者邊緣(PE)和客戶邊緣(CE)之間的內部邊界閘道通訊協定(iBGP)。

背景資訊

在新的iBGP PE-CE功能之前，PE和CE之間的iBGP(因此在PE路由器上的虛擬路由和轉發(VRF)介面上)沒有獲得正式支援。一個例外是多VRF CE(VRF-Lite)設定中的VRF介面上的iBGP。部署此功能的動機如下：

- 客戶希望在VRF的多個站點上擁有單個自治系統編號(ASN)，而不部署帶as-override的外部邊界網關協定(eBGP)。
- 客戶希望提供指向CE路由器的內部路由反射，就像服務提供商(SP)核心是一個透明路由反射器(RR)一樣。

通過此功能，VRF的站點可以擁有與SP核心相同的ASN。但是，如果VRF站點的ASN與SP核心的ASN不同，則使用功能本地 — 自治系統(AS)可以使其外觀相同。

實施iBGP PE-CE

以下是使此功能正常運作的兩個主要部分：

- 在BGP協定中新增了一個新屬性ATTR_SET，以便以透明方式在SP核心中承載VPN BGP屬性。
- 將PE路由器設定為指向VRF中CE路由器的iBGP會話的RR，以及指向VPNv4鄰居（其他PE路由器或RR）的RR。

新的ATTR_SET屬性允許SP以透明的方式承載客戶的所有BGP屬性，並且不會干擾SP屬性和BGP策略。這些屬性包括群集清單、本地首選項、社群等。

BGP客戶路由屬性

ATTR_SET是新的BGP屬性，用於承載SP客戶的VPN BGP屬性。它是可選傳遞屬性。在此屬性中，可以承載BGP更新消息中的所有客戶BGP屬性（MP_REACH和MP_UNREACH屬性除外）。

ATTR_SET屬性的格式如下：

```
+-----+
| Attr Flags (O|T) Code = 128 |
+-----+
| Attr. Length (1 or 2 octets) |
+-----+
| Origin AS (4 octets)         |
+-----+
| Path Attributes (variable)   |
+-----+
```

屬性標誌是常規BGP屬性標誌（請參閱RFC 4271）。屬性長度指示屬性長度是一個或兩個八位數。「原始AS」欄位的目的是在不正確處理AS_PATH的情況下，防止源自AS的一個路由洩漏到另一個AS。可變長度路徑屬性欄位包含必須通過SP核心傳輸的VPN BGP屬性。

在出口PE路由器上，VPN BGP屬性被推入此屬性。在輸入PE路由器上，在將BGP字首傳送到CE路由器之前，這些屬性將從屬性中彈出。此屬性提供SP網路和客戶VPN之間的BGP屬性的隔離，反之亦然。例如，在VPN網路內看不到並考慮SP路由反射群集清單屬性。但是，在SP網路內看不到並考慮VPN路由反射群集清單屬性。

檢視圖1，檢視客戶BGP字首在SP網路中的傳播。

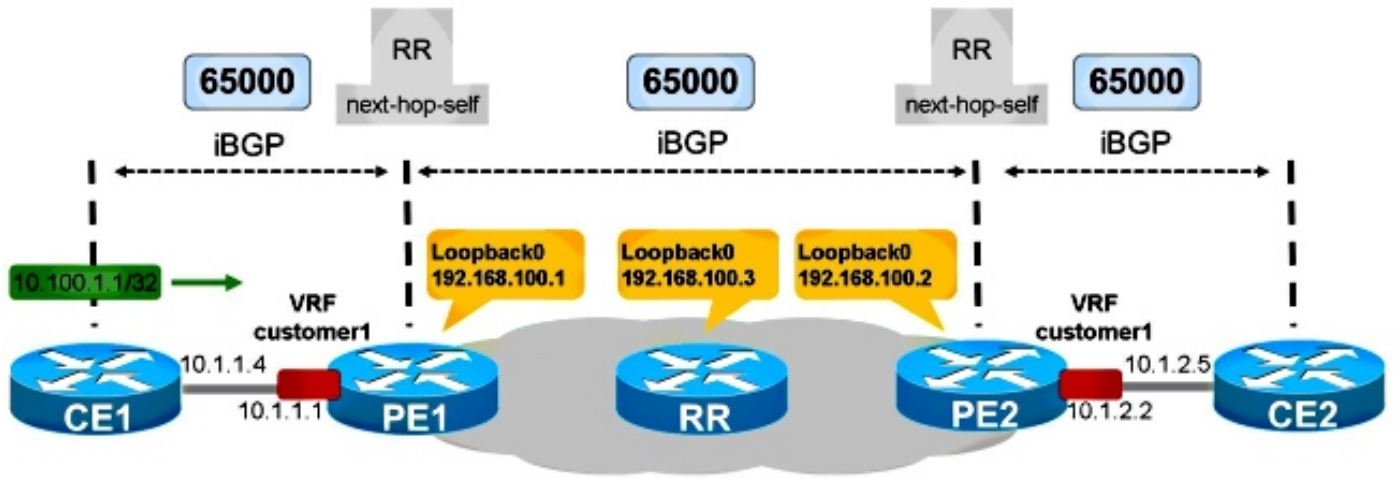


圖1

CE1和CE2與SP網路位於同一個AS中：65000. PE1為CE1配置了iBGP。PE1將字首10.100.1.1/32的路徑反映到SP網路中的RR。RR會照常反映前往PE路由器的iBGP路徑。PE2反映了通向CE2的路徑。

為了使該操作正常工作，您必須：

- 在PE1和PE2上具有支援iBGP PE-CE功能的代碼
- 配置PE1和PE2，以便在其通向各自的CE路由器的BGP會話上執行路由反射
- 在PE路由器上使用下一跳自身，用於通向其CE路由器的BGP會話
- 確保每個VPN站點使用不同的路由區分器(RD)

設定

請參閱圖1。

下面是PE1和PE2所需的配置：

```
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
```

附註：如果PE沒有用於CE鄰居的`neighbor <internal-CE> internal-vpn-client`命令，則它不會將字首從CE傳播到SP RR/PE路由器。

附註：如果PE不是VRF中的RR，它不會將字首從RR/PE路由器傳播到CE路由器。

新指令

有一個新命令`neighbor <internal-CE> internal-vpn-client`使此功能正常工作。它必須在PE路由器上配置，僅用於指向CE路由器的iBGP會話。

附註：不使用`neighbor <internal-CE> internal-vpn-client`命令，仍支援iBGP PE-CE多重VRF CE(VRF-Lite)功能。

附註：設定`neighbor <internal-CE> internal-vpn-client`命令時，`neighbor <internal-CE> route-reflector-client`和`neighbor <internal-CE> next-hop-self`命令也會自動加入設定中。當刪除了`neighbor <internal-CE> route-reflector-client`和`neighbor <internal-CE> next-hop-self`命令（或

同時刪除兩者)之一并執行重新載入時,它們會自動放回配置中。

詳細檢視ATTR_SET

請參閱圖1。

這是CE1通告的字首:

```
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
```

當PE1收到來自CE1的BGP字首10.100.1.1/32時,它會將其儲存兩次:

```
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
```

第一個路徑是PE1上的實際路徑,因為它從CE1接收。

第二條路徑是通告給RR/PE路由器的路徑。標有**ibgp sourced**。它包含ATTR_SET屬性。請注意,此路徑連線有一個或多個路由目標(RT)。

PE1通告字首,如下所示:

```
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

RR看到的路徑如下所示：

```
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
```

請注意，核心中此VPNv4單播字首的本地優先順序為100。在ATTR_SET中，儲存原始本地優先順序200。但是，這對SP核心中的RR是透明的。

在PE2上，您會看到如下所示的字首：

```
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 no-label/18
rx pathid:0, tx pathid: 0x0
```

第一個路徑是從RR接收的路徑以及ATTR_SET。請注意，RD是65000:1，即原始RD。第二個路徑是從VRF表中匯入的路徑，RD 65000:1。已刪除ATTR_SET。

這是CE2上顯示的路徑：

```
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
```

請注意，下一跳是10.1.2.2，即PE2。集群清單包含路由器PE1和PE2。這些是影響VPN內部的RR。SP RR(10.100.1.3)不在群集清單中。

在SP網路的VPN內部保留了200的本地首選項。

debug bgp vpnv4 unicast updates命令顯示SP網路中傳播的更新：

```
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,

附註：PE1從RR收到自己的更新，然後丟棄該更新。這是因為PE1和PE2在RR上位於同一個更新組中。

附註：如果要以十六進位制轉儲完整的更新消息，請使用**debug BGP updates**命令的**detail**關鍵字。

```
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 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,
```

下一跳處理

必須在PE路由器上為此功能配置Next-hop-self。原因是使用iBGP傳輸下一躍點時，通常不會發生更改。但是這裡有兩個獨立的網路：運行單獨的內部網關協定(IGP)的VPN網路和SP網路。因此，IGP度量不能輕易比較並用於兩個網路之間的最佳路徑計算。RFC 6368選擇的方法是對CE的iBGP會話強制使用next-hop-self，這樣可以同時避免前面描述的問題。優點在於，VRF站點可使用此方法運行不同的IGP。

RD

RFC 6368提到建議相同VPN的不同VRF站點使用不同（唯一）的RD。在Cisco IOS中，此功能必須如此。

使用本地AS的iBGP PE-CE功能

請參閱圖2。VPN客戶1具有ASN功65001。

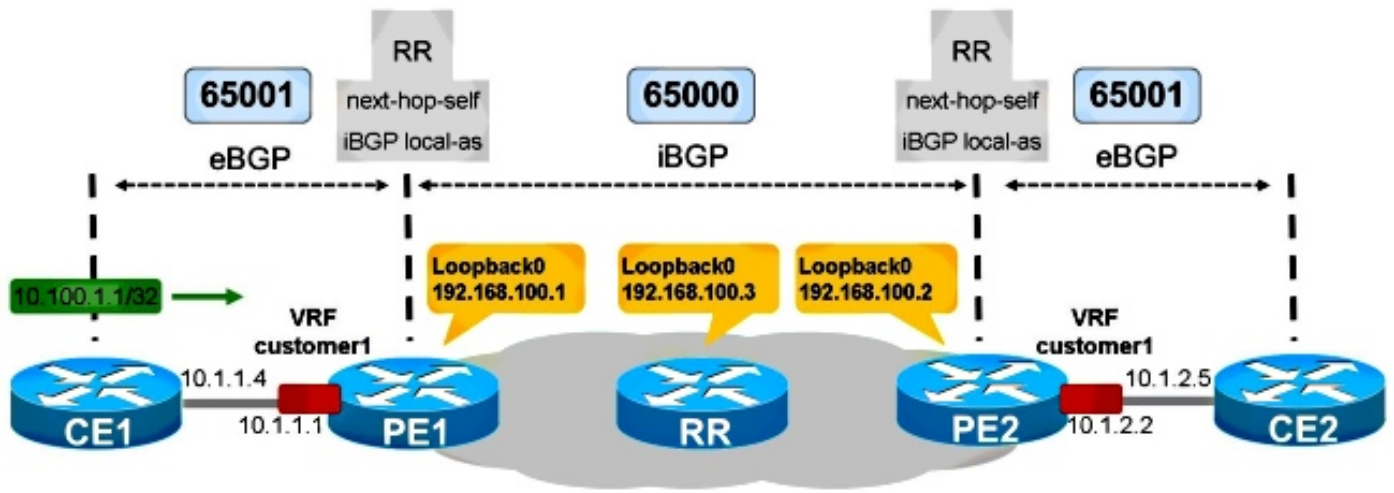


圖2

CE1位於AS 65001中。為了從PE1的角度建立此內部BGP，它需要iBGP local-as功能。

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和CE2的配置類似。

PE1看到BGP字首，如下所示：

```
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
```

首碼是內部BGP。

PE2看到以下內容：

```
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
```

建立者AS是65001，它是將字首從PE2傳送到CE2時使用的AS。因此，保留了AS，在本例中還保留本地首選項。

```
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
```

您會看到**Local**而不是AS路徑。這表示它是源自AS 65001的內部BGP路由，該路由也是路由器CE2的配置的ASN。所有BGP屬性均取自ATTR_SET屬性。這符合下一節中案例1的規則。

不同VRF站點之間的路由交換規則

ATTR_SET包含始發VRF的始發者AS。當遠端PE在將ATTR_SET字首傳送到CE路由器之前刪除ATTR_SET時，會檢查此Originating AS。

案例1:如果始發AS與為CE路由器配置的AS匹配，則當PE將路徑匯入目標VRF時，會從ATTR_SET屬性獲取BGP屬性。

案例2:如果始發AS與為CE路由器配置的AS不匹配，則採用所構建路徑的屬性集，如下所示：

1. 路徑屬性設定為ATTR_SET屬性中包含的屬性。
2. iBGP特定的屬性被丟棄 (LOCAL_PREF、ORIGINATOR和CLUSTER_LIST)。
3. ATTR_SET屬性中包含的**Origin AS**編號在AS_PATH之前並遵循應用於源和目標AS之間的外部BGP對等體的規則。
4. 如果與VRF關聯的自治系統與VPN提供商自治系統相同，並且VPN路由的AS_PATH屬性不為空，則它應優先於VRF路由的AS_PATH屬性。

請參閱圖3。CE1和PE1具有AS 65000並配置了iBGP PE-CE功能。CE2具有ASN 65001。這意味著PE2和CE2之間存在eBGP。

圖3

PE2看到該路由如下：

```
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

```

這是CE2上顯示的字首：

```

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

```

這是案例2。PE2將ATTR_SET屬性中包含的Origin AS編號附加到AS_PATH並遵循應用於源和目標AS之間的eBGP對等體的規則。當iBGP特定屬性建立要通告到CE2的路由時，PE2會將其忽略。因此，本地優先順序為100而不是200（如ATTR_SET屬性中所示）。

CE-to-CE VRF-Lite反射

請參閱圖4。

圖4

圖4顯示了連線到PE1的額外CE路由器CE3。CE1和CE3都連線到同一VRF例項上的PE1:customer1。這表示CE1和CE3是PE1的多VRF CE路由器（也稱為VRF-Lite）。PE1在從CE1通告字首到CE3時將自身設定為下一跳。如果不想發生此行為，您可以在PE1上配置neighbor 10.1.3.6 next-hop-unchanged。要配置此行為，您必須在PE1上刪除neighbor 10.1.3.6 next-hop-self。然後CE3從CE1到CE1的路由是這些BGP字首的下一跳。為了實現這一目的，您需要在CE3的路由表中為這些BGP下一跳提供路由。您需要在CE1、PE1和CE3上使用動態路由協定(IGP)或靜態路由，以確保路由器具有用於彼此下一跳IP地址的路由。但是此組態存在問題。

PE1上的配置為：

```

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
```

在CE3上，可以看到CE1的字首正常：

```
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
```

但是，在CE3上可以看到CE2的字首，如下所示：

```
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
```

BGP下一躍點是**192.168.100.2**，即PE2的環回IP地址。PE1在向CE3通告字首10.100.1.2/32時，沒有將BGP下一躍點重寫為自身。這使得該字首在CE3上不可用。

因此，在跨MPLS-VPN和iBGP VRF-Lite混合使用iBGP PE-CE功能的情況下，您必須確保在PE路由器上始終具有下一跳自我。

當PE路由器是反映跨PE本地VRF介面從一個CE到另一個CE的iBGP路由的RR時，您無法保留下一跳。當您在MPLS VPN網路中運行iBGP PE-CE時，必須為指向CE路由器的iBGP會話使用**internal-vpn-client**。當您在PE路由器的VRF中有多個本地CE時，您必須為這些BGP對等體保留**next-hop-self**。

您可以檢視路由對映，以將從其他PE路由器收到的字首的下一跳設定為self，但不會將其他本地連線的CE路由器的反射字首設定為self。但是，當前不支援在出站路由對映中將下一跳設定為self。該配置如下所示：

```
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
!

```

但是不支援：

```

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

```

PE路由器上的舊版Cisco IOS

如果PE1運行舊版Cisco IOS軟體，但缺少iBGP PE-CE功能，則PE1從不將自身設定為所反映的iBGP字首的下一跳。這表示從CE1(10.100.1.1)到CE2 -via PE1— 的反射的BGP字首(10.100.1.1/32)會將CE1(10.1.1.4)用作下一跳。

```

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

```

CE2(10.100.1.2/32)中的字首與PE2一起被視為下一跳，因為PE1也不對此字首執行next-hop-self:

```

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
```

要使iBGP PE-CE功能正常工作，啟用該功能的VPN的所有PE路由器都必須具有支援該功能的代碼並啟用該功能。

VRF上eBGP的Next-hop-self

請參閱圖5。

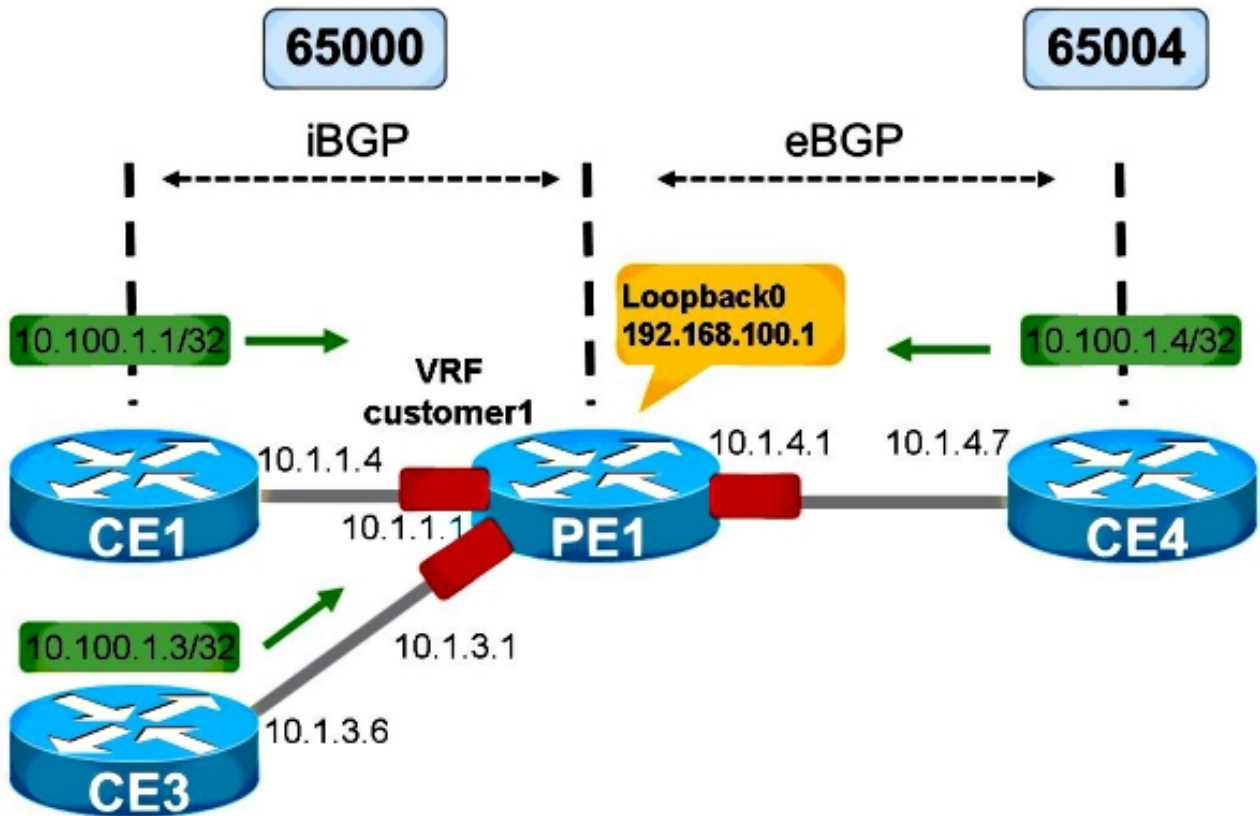


圖5

圖5顯示了VRF-Lite設定。從PE1到CE4的會話是eBGP。從PE1到CE3的會話仍然是iBGP。

對於eBGP字首，當下一躍點向VRF上的iBGP鄰居通告字首時，它始終設定為self。如果透過VRF對iBGP鄰居的作業階段已設定next-hop-self，則會發生這種情況。

在圖5中，CE3將來自CE4的字首視為PE1。

```
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
```

這與PE1上指向CE3或不指向CE3的下一跳自身有關。

如果PE1上指向CE3和CE4的介面不在VRF中，但在全域性環境中，指向CE3的下一跳自會起到作用。

如果沒有PE1上指向CE3的下一跳自我，您將看到：

```
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)
```

雖然隱式啟用了next-hop-self，但輸出並未指出這一點。

在PE1到CE3的next-hop-self中，您將看到：

```
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
```

但是，如果指向CE3和CE4的介面在全域性環境中，則如果未配置next-hop-self，來自CE4的字首的下一跳為CE4本身：

```
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
```

對於PE1到CE3的下一跳自我：

```
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
```

這是根據RFC 4364完成的。

如果您不想為通過VRF介面的iBGP會話設定eBGP字首的next-hop-self，則必須配置next-hop-

unchanged。僅思科錯誤ID [CSCuj11720](#)支援此功能。

```
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
```

現在，CE3將CE4視為CE4通告的字首的下一跳：

```
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
```

如果您嘗試在Cisco錯誤ID [CSCuj11720](#)之前在Cisco IOS代碼上為CE3的iBGP會話配置**next-hop-unchanged**關鍵字，則會遇到以下錯誤：

```
PE1(config-router-af)# neighbor 10.1.3.6 next-hop-unchanged
%BGP: Can propagate the nexthop only to multi-hop EBGP neighbor
思科錯誤ID CSCuj11720後，next-hop-unchanged關鍵字對多躍點eBGP鄰居和iBGP VRF-Lite鄰居有效。
```