

IOS による iBGP PE-CE 機能の実装

内容

[概要](#)

[背景説明](#)

[iBGP PE-CE の実装](#)

[BGP カスタマー ルート属性](#)

[設定](#)

[新規コマンド](#)

[ATTR SET の詳細説明](#)

[ネクスト ホップ処理](#)

[RD](#)

[ローカル AS を使用した iBGP PE-CE 機能](#)

[異なる VRF サイト間のルート交換ルール](#)

[CE から CE への VRF-Lite の反映](#)

[PE ルータでの古い Cisco IOS](#)

[VRF での eBGP の next-hop-self](#)

概要

このドキュメントでは、Cisco IOS® でプロバイダー エッジ (PE) とカスタマー エッジ (CE) 間の内部ボーダー ゲートウェイ プロトコル (iBGP) 機能がどのように実装されているかについて説明します。

背景説明

新しい iBGP PE-CE 機能が搭載されるまで、PE と CE 間の iBGP (したがって、PE ルータ上の Virtual Routing and Forwarding (VRF) インターフェイス) は、公式にはサポートされていませんでした。唯一の例外は、Multi-VRF CE (VRF-Lite) セットアップでの VRF インターフェイスの iBGP です。この機能が導入された理由は次のとおりです。

- カスタマーは、AS オーバーライドを使用した外部 Border Gateway Protocol (eBGP) を展開せずに、VRF の複数のサイトで単一の自律システム番号 (ASN) を持ちたいと考えます。
- カスタマーは、サービス プロバイダー (SP) コアがまるで 1 つの透過ルート リフレクタ (RR) のように機能する、CE ルータへの内部ルート リフレクションを提供したいと考えます。

この機能を使用すると、VRF のサイトは、SP コアと同じ ASN を持つことができます。ただし、VRF サイトの ASN が SP コアの ASN と異なっている場合は、この機能のローカル自律システム (AS) を使用して、同じであるように表示できます。

iBGP PE-CE の実装

この機能を動作させるのは、次の 2 つの主要部分です。

- SP コアで VPN BGP 属性を透過的に伝送するために、新しい属性である ATTR_SET が BGP プロトコルに追加されました。
- PE ルータを VRF の CE ルータへの iBGP セッションの RR、および VPNv4 ネイバー (他の PE ルータまたは RR) への RR として機能させます。

新しい ATTR_SET 属性では SP がカスタマーの BGP 属性すべてを透過的に伝送でき、SP 属性や BGP ポリシーに干渉することがありません。こうした属性にはクラスタ リスト、ローカル設定、コミュニティなどがあります。

BGP カスタマー ルート属性

ATTR_SET は SP カスタマーの VPN BGP 属性を伝送するために使用される、新しい BGP 属性です。これは過渡的なオプション属性です。この属性では、MP_REACH 属性と MP_UNREACH 属性を除く、BGP 更新メッセージからのカスタマーの BGP 属性をすべて伝送できます。

ATTR_SET 属性の形式は次のとおりです。

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

属性フラグは通常の BGP 属性フラグです (RFC 4271 を参照)。属性の長さは、この属性の長さが 1 オクテットであるか 2 オクテットであることを示します。Origin AS フィールドの目的は、ある AS で発生するルートが、適切な AS_PATH 操作を行われずに、別の AS にリークされないようにすることです。可変長のパス属性フィールドには、SP コアで伝送されなければならない VPN BGP 属性が含まれます。

出力 PE ルータで、VPN の BGP 属性はこの属性にプッシュされます。入力 PE ルータでは、BGP プレフィクスが CE ルータに送信される前に、この属性からこれら VPN の BGP 属性がポップされます。この属性は、SP ネットワークとカスタマー VPN 間の BGP 属性の分離およびその逆を提供します。たとえば、SP ルート リフレクションのクラスタ リスト属性は VPN ネットワーク内では表示されず、考慮されません。また、VPN ルート リフレクションのクラスタ リスト属性も、SP ネットワーク内で表示されず、考慮されません。

SP ネットワークにおけるカスタマー BGP プレフィクスの伝播を確認するには、図 1 を参照してください。

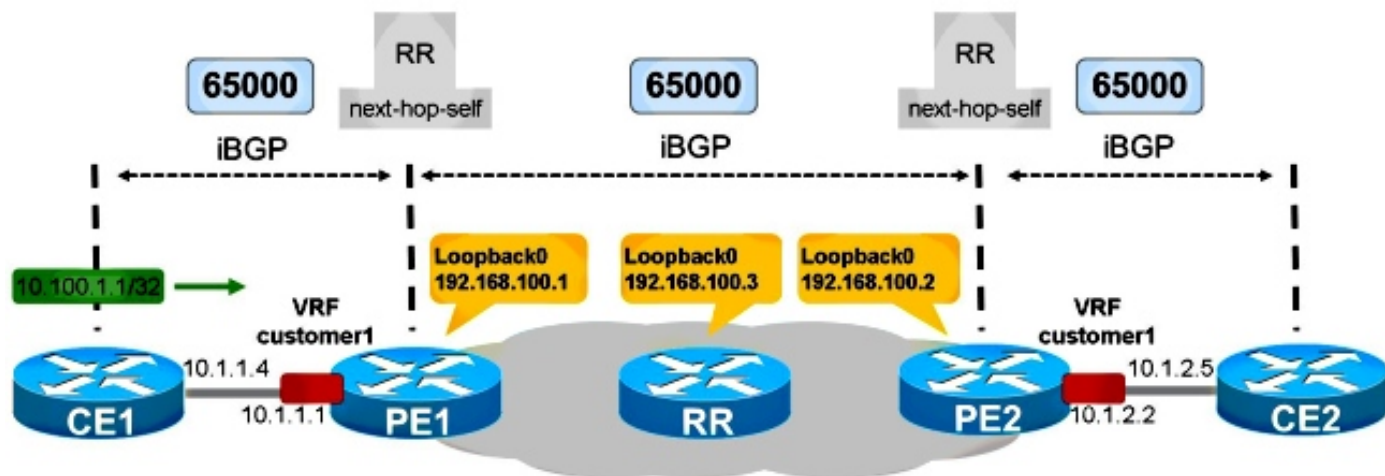


図 1 :

CE1 および CE2 は SP ネットワークと同じ AS にあります65000。PE1にはCE1に対してiBGPが設定されています。PE1は、プレフィックス10.100.1.1/32のSPネットワーク内のRRへのパスを反映します。RR は PE ルータへの iBGP パスを通常どおりに反映します。PE2 は、CE2 へのパスを反映します。

これが正しく動作するには、以下のことが必要です。

- PE1 および PE2 に、iBGP PE-CE 機能をサポートするコードがある
- それぞれの CE ルータの BGP セッションで、ルート リフレクションを実行するように PE1 および PE2 が設定されている
- PE ルータに、CE ルータへの BGP セッションの next-hop-self がある
- 各 VPN サイトが異なるルート識別子 (RD) を使用していることを確認する

設定

図 1 を参照してください。

PE1 および PE2 に必要な設定は次のとおりです。

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
```

注：PE に CE ネイバーの neighbor <internal-CE> internal-vpn-client コマンドがない場合、PE は CE のプレフィックスを SP の RR/PE ルータに伝播しません。

注：PE が VRF の RR ではない場合、PE は RR/PE ルータのプレフィックスを CE ルータに伝搬しません。

新規コマンド

この機能を動作させるための、新しいコマンド、neighbor <internal-CE> internal-vpn-client があります。これは、CE ルータへの iBGP セッションに対してのみ、PE ルータで設定する必要があります。

注：iBGP PE-CE Multi-VRF CE (VRF-Lite) 機能は、neighbor <internal-CE> internal-vpn-client コマンドがなくてもサポートされます。

注 : neighbor <internal-CE> internal-vpn-client コマンドを設定すると、neighbor <internal-CE> route-reflector-client コマンドと neighbor <internal-CE> next-hop-self コマンドも自動的に設定されます。neighbor <internal-CE> route-reflector-client コマンドか neighbor <internal-CE> next-hop-self コマンドのいずれか (または両方) が削除され、リロードが実行されると、再度、自動的に設定されます。

ATTR_SET の詳細説明

図 1 を参照してください。

これは CE1 によってアドバタイズされたプレフィックスです。

```
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
```

PE1 が CE1 から BGP プレフィックス 10.100.1.1/32 を受信すると、これを 2 回、保存します。

```
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
```

最初のパスは CE1 から受信したため、PE1 の実際のパスです。

第 2 のパスは RR/PE ルータへアドバタイズされたパスです。2 番目のパスには **ibgp sourced** というマークが付きます。これには ATTR_SET 属性が含まれています。このパスには、それに接続する 1 つ以上のルート ターゲット (RT) があります。

PE1 は次のようにプレフィックスをアドバタイズします。

```
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

これは、RR でパスが表示される方法です。

```
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
```

コアのこのVPNv4ユニキャストプレフィックスのローカルプリファレンスは100であることに注意してください。ATTR_SETでは、元のローカルプリファレンスの200が保存されます。ただし、これはSP コアのRR に対して透過的です。

PE2 では、次のようにプレフィックスが表示されます。

```
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

```

最初のパスは ATTR_SET により RR から受信したものです。RD は 65000:1 (元の RD) であることがわかります。2番目のパスは、RD 65000:1を使用してVRFテーブルからインポートされたパスです。ATTR_SETが削除されました。

これが CE2 に表示されるパスです。

```

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

```

ネクストホップは10.1.2.2(PE2)で、クラスタリストにはルータPE1とPE2が含まれています。これらはVPN内部で重要なRRです。SP RR (10.100.1.3) はクラスタ リストにありません。

ローカル設定の 200 は SP ネットワークにおいて VPN 内で維持されます。

debug bgp vpnv4 unicast updates コマンドにより、SP ネットワーク内で伝播された更新が表示されます。

```

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,
```

注：PE1 は RR から自身の更新を受信し、その後これを廃棄します。これは、PE1 と PE2 が RR 上の同じ更新グループにあるためです。

注：完全な更新メッセージを 16 進数でダンプするには、`debug BGP updates` コマンドに `detail` キーワードを使用します。

```
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,
```

ネクスト ホップ処理

この機能を使用するには、`next-hop-self` を PE ルータに設定する必要があります。これは、通常、ネクストホップは iBGP による変更なしで伝送されるからです。ただし、ここでは 2 つの独立したネットワークがあります。VPN ネットワークと SP ネットワークで、これらは別々の内部ゲートウェイプロトコル (IGP) を実行しています。したがって、IGP メトリックは簡単に比較できず、2 つのネットワーク間の最適パスの計算には使用できません。RFC 6368 によって選択されたアプローチは、CE への iBGP セッションで `next-hop-self` を必須とし、これにより、前述の問題をすべて回避するというものです。これには、VRF のサイトがこのアプローチにより、さまざまな IGP を実行できるという利点があります。

RD

RFC 6368 では、同じ VPN の異なる VRF サイトが別々の (一意の) RD を使用することを推奨

しています。Cisco IOS では、これが、この機能のために必須となっています。

ローカル AS を使用した iBGP PE-CE 機能

図2を参照してください。VPN customer1にはASN 65001があります。

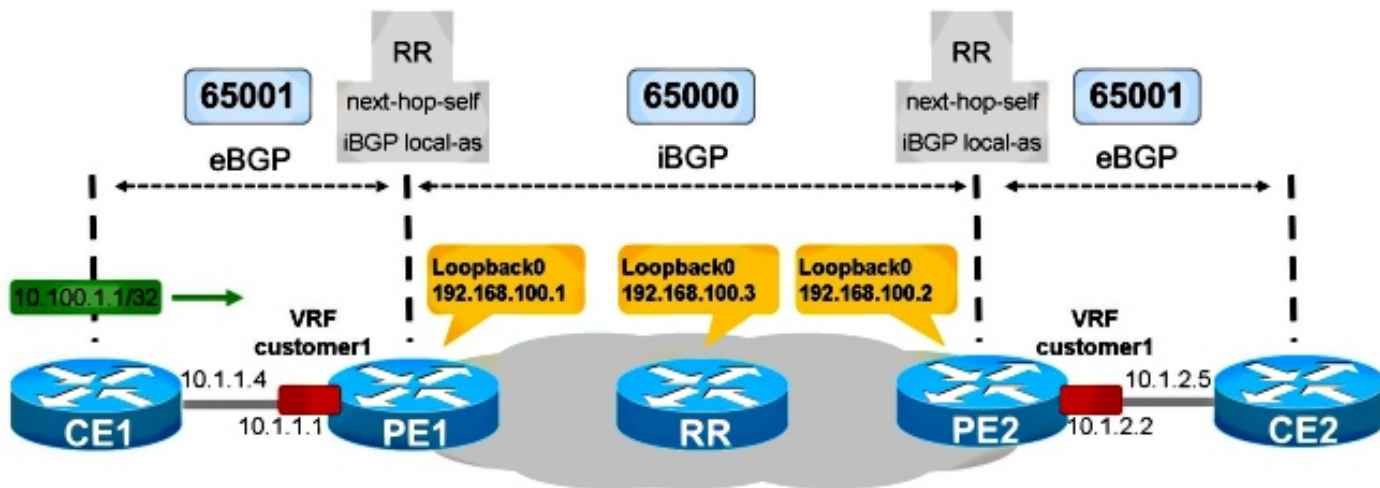


図 2

CE1はAS 65001にあります。この内部BGPをPE1の観点から作成するには、iBGP local-as機能が必要です。

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 と CE2 は同じように設定します。

PE1 には、次に示すように BGP プレフィクスが表示されます。

```
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
```

このプレフィックスは内部 BGP です。

PE2 には次が表示されます。

```
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
```

Originator ASは65001で、プレフィクスがPE2からCE2に送信される際に使用されるASです。したがって、ASは維持され、この例ではローカル設定も同様になります。

```
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
```

「AS Path」ではなく、「Local」となっていることがわかります。これは、AS 65001から発信された内部BGPルートであり、ルータCE2の設定済みASNでもあります。すべてのBGP属性はATTR_SET属性から取得されたものです。これは、次のセクションのケース1のルールに従っています。

異なる VRF サイト間のルート交換ルール

ATTR_SETには、起点VRFのオリジネータASが含まれています。この起点ASは、リモートPEがプレフィックスをCEルータに送信する前にATTR_SETを削除するとき、リモートPEによりチェックされます。

ケース1: 起点ASがCEルータに設定されたASに一致すると、PEが宛先VRFにパスをインポートするときに、ATTR_SET属性からBGP属性が取得されます。

ケース2: 起点ASがCEルータに設定されたASに一致しないと、構築されたパスの属性セットが、次に示すように取得されます。

1. パス属性はATTR_SET属性に含まれている属性に設定されます。
2. iBGP固有の属性は廃棄されます (LOCAL_PREF、ORIGINATOR、CLUSTER_LIST)。
3. ATTR_SET属性に含まれているOrigin AS番号がAS_PATHの前に付き、発信元と宛先AS間をピアする外部BGPに適用されるルールに従います。
4. VRFに関連付けられた自律システムがVPNプロバイダーの自律システムと同じであり、VRFルートのAS_PATH属性が空でない場合、このASがVPNルートのAS_PATH属性の前に付加されます。

図3を参照してください。CE1とPE1はAS 65000で、iBGP PE-CE機能が設定されています。CE2にはASN 65001があります。これは、PE2とCE2の間にeBGPがあることを意味します。

図 3

PE2には次のようにルートが表示されます。

```
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

```

これは、CE2 に表示されるプレフィックスです。

```

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

```

ケース2です。ATTR_SET属性に含まれるOrigin AS番号は、PE2によってAS_PATHの前に付加され、送信元と宛先AS間のeBGPピアリングに適用される規則に従います。iBGP固有の属性は、CE2にアドバタイズされるルートを作成するときにPE2によって無視されます。したがって、ローカルプリファレンスは100であり、200ではありません (ATTR_SET属性に示されます)。

CE から CE への VRF-Lite の反映

図 4 を参照してください。

図 4

図4は、PE1に接続された追加のCEルータCE3を示しています。CE1とCE3は両方とも同じVRFインスタンスのPE1に接続されています。customer1。これは、CE1とCE3がPE1のマルチVRF CEルータ (VRF-Liteとも呼ばれる) であることを意味します。PE1は、CE1からCE3にプレ

フィックスをアドバタイズするときにネクストホップとして自身を配置します。この動作が不要で1.これを設定するには、PE1でneighbor 10.1.3.6 next-hop-selfを削除する必要があります。その後、CE3は、CE1を持つCE1からのルートをそれらのBGPプレフィックスのネクストホップとして認識します。これを機能させるには、CE3のルーティングテーブル内のBGPネクストホップのルートが必要です。CE1、PE1、およびCE3のダイナミックルーティングプロトコル(IGP)またはスタティックルートを使用して、ルータが他のネクストホップIPアドレスのルートを持つようにします。ただし、この設定には問題があります。

PE1 の設定は次のとおりです。

```
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
```

CE1 からのプレフィックスは CE3 で正常に表示されます。

```
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
```

しかし、CE2 からのプレフィックスは CE3 では次のように表示されます。

```
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ネクストホップは192.168.100.2で、PE2のループバックIPアドレスです。PE1は、プレフィックス10.100.1.2/32をCE3にアドバタイズしたときに、BGPネクストホップを自身に書き換えませんでした。

したがって、MPLS-VPN の iBGP PE-CE 機能と iBGP VRF-Lite が混在する場合、PE ルータに next-hop-self が設定されていることを常に確認する必要があります。

PE ルータが、ある CE から別の CE への VRF インターフェイスの iBGP ルートを PE 上でローカルに反映する RR の場合は、ネクストホップを保存することはできません。MPLS VPN ネットワークで PE-CE iBGP を実行する場合は、CE ルータへの iBGP セッションに **internal-vpn-client** を使用する必要があります。PE ルータの VRF で複数のローカル CE があると、これらの BGP ピアの **next-hop-self** を保持する必要があります。

ローカル接続された他の CE ルータから反映されたプレフィックスではなく、他の PE ルータから受信したプレフィックスに対してネクストホップを自身に設定するために、ルートマップを調べることができません。ただし、発信ルート マップでの自身へのネクストホップの設定は、現在サポートされていません。この設定例を次に示します。

```
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
!
```

ただし、これはサポートされていません。

```
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 ルータでの古い Cisco IOS

PE1 が、iBGP PE-CE 機能のない古い Cisco IOS ソフトウェアを実行している場合は、PE1 は反映される iBGP プレフィックスのネクストホップとして、自身を設定することができません。これは、CE1 (10.100.1.1) から PE1 を経由して CE2 に反映される BGP プレフィックス (10.100.1.1/32) が、ネクストホップとして CE1 (10.1.1.4) を持つことを意味します。

```
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) からのプレフィックスにはネクストホップとして PE2 が表示されます。これは、PE1 がこのプレフィックスの next-hop-self も実行しないからです。

```
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
```

iBGP PE-CE 機能が正しく動作するには、この機能が有効な VPN のすべての PE ルータに、機能をサポートするコードがあり、この機能を有効にしている必要があります。

VRF での eBGP の next-hop-self

図 5 を参照してください。

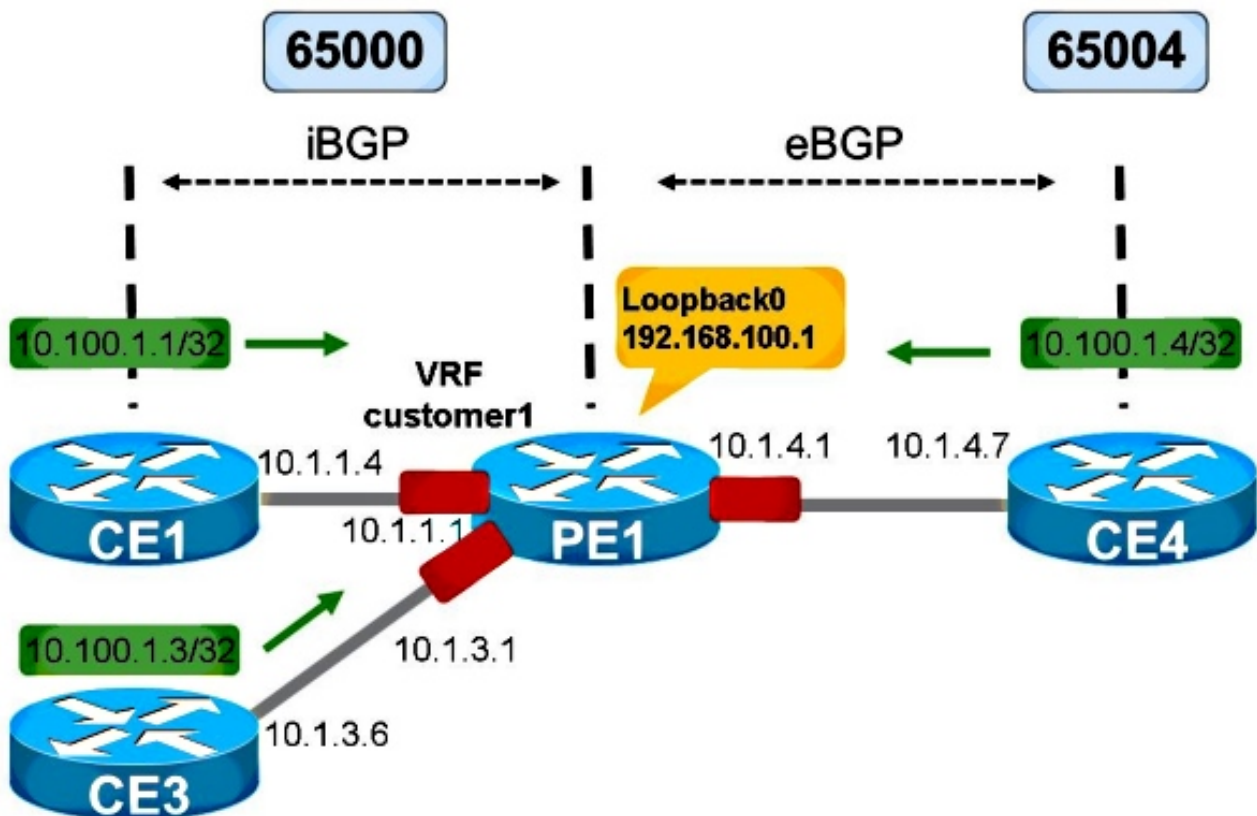


図 5 :

図 5 に、VRF-Lite のセットアップを図示します。PE1 から CE4 へのセッションは eBGP です。PE1 から CE3 へのセッションはまだ iBGP です。

eBGP プレフィクスでは、VRF の iBGP ネイバーへプレフィクスがアドバタイズされる時、ネクストホップは常に自身に設定されます。これには、VRF での iBGP ネイバーへのセッションで next-hop-self が設定されているかどうかは関係ありません。

図 5 で、CE3 にはネクストホップが PE1 である CE4 からのプレフィクスが表示されます。

```
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
```

これは、CE3 に対して PE1 に next-hop-self があってもなくても発生します。

CE3 と CE4 への PE1 のインターフェイスが VRF になくても、グローバル コンテキストにある場合は、CE3 への next-hop-self に違いが発生します。

CE3 への PE1 の next-hop-self がない場合は次のようになります。

```
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)
```

next-hop-self が暗黙的に有効になっていても、出力にはそれが示されません。

CE3 への PE1 の next-hop-self がある場合は次のようになります。

```
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
```

ただし、CE3 と CE4 へのインターフェイスがグローバル コンテキストにあるときには、CE4 のプレフィックスのネクスト ホップは、next-hop-self が設定されていない場合、CE4 自身になります。

```
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
```

CE3 への PE1 の next-hop-self の場合

```
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
```

これは、RFC 4364 に基づいて実行されました。

VRF インターフェイスで iBGP セッションへの eBGP プレフィックスの next-hop-self を設定しないようにするには、next-hop-unchanged を設定する必要があります。これについてのサポートは、Cisco Bug ID [CSCuj11720](#) のみです。

```
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
```

現在では、CE3 には CE4 によりアドバタイズされるプレフィックスのネクストホップとして CE4 が表示されます:

```
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
```

Cisco Bug ID [CSCUj11720](#) 以前の Cisco IOS コードで、CE3 への iBGP セッションに [next-hop-unchanged](#) キーワードを設定しようとする、次のエラーが発生します。

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

Cisco Bug ID [CSCUj11720](#) 以降、[next-hop-unchanged](#) キーワードはマルチホップ eBGP ネイバーと iBGP VRF-Lite ネイバーで有効です。