

# iBGP PE-CE功能的IOS实施

## 目录

[简介](#)

[背景信息](#)

[实现iBGP PE-CE](#)

[BGP客户路由属性](#)

[配置](#)

[new命令](#)

[详细的看看ATTR SET](#)

[下一跳处理](#)

[RD](#)

[iBGP与Local-AS的PE-CE功能](#)

[路由交换的规则区别VRF站点之间](#)

[CE对CE VRF-Lite反射](#)

[在PE路由器的更旧的Cisco IOS](#)

[next-hop-self在VRF的eBGP](#)

## 简介

本文描述内部边界网关协议(iBGP)在服务商边缘和用户边缘(CE)功能之间如何在Cisco IOS实现。

## 背景信息

直到新的iBGP PE-CE功能，(因此在PE路由器的虚拟路由和转发(VRF)接口)不正式支持在PE之间的iBGP和CE。一例外是在VRF接口的iBGP在(VRF-Lite)设置的多VRF CE。设置此功能的动机是：

- 客户要有一自治系统编号(ASN)在VRF的多个站点，不用外部边界网关协议(eBGP)的部署与覆盖。
- 客户要提供往CE路由器的内部路由反射，操作，好象服务提供商(SP)核心一台透明路由反射器(RR)。

使用此功能，VRF的站点能有ASN和SP核心一样。然而，万一VRF站点的ASN跟SP核心的ASN不同，它可以做显现同样使用功能本地自治系统(AS)。

## 实现iBGP PE-CE

这是两大部分为了使得此功能运作：

- 新团体ATTR\_SET被添加到BGP协议为了运载在SP核心间的VPN BGP属性以透明方式。
- 做PE路由器iBGP会话的RR往CE路由器VRF的和作为往Vpvn4邻居的RR (其他PE路由器或RR)。

新的ATTR\_SET属性允许SP运载所有客户的BGP属性以透明方式，并且不干涉SP属性和BGP策略。这样属性是集群列表，本地首选，社区，等等。

## BGP客户路由属性

ATTR\_SET是用于的新的BGP属性为了运载SP客户的VPN BGP属性。它是可选传递属性。在此属性，所有从BGP更新消息的客户BGP属性，除了MP\_REACH和MP\_UNREACH属性，可以运载。

ATTR\_SET属性有此格式：

```

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

```

属性标志是正常BGP属性标志(参考的RFC 4271)。属性长度指示属性长度是否是一两个八位位组。没有AS\_PATH的适当的处理，始发地AS字段的目的是防止于一个发起的一个路由泄漏至于漏对另一个AS。可变长的路径属性字段运载必须在SP核心间运载的VPN BGP属性。

在出口PE路由器，VPN BGP属性推送到此属性。在入口PE路由器，在BGP前缀发送对CE路由器前，这些属性从属性弹出。此属性提供BGP属性的隔离在SP网络和客户VPN之间的反之亦然。例如，SP路由反映集群列表属性看不到并且考虑在VPN网络里面。而且，VPN路由反映集群列表属性看不到并且考虑在SP网络里面。

查看图1为了发现一个客户BGP前缀的传播在间SP网络的。

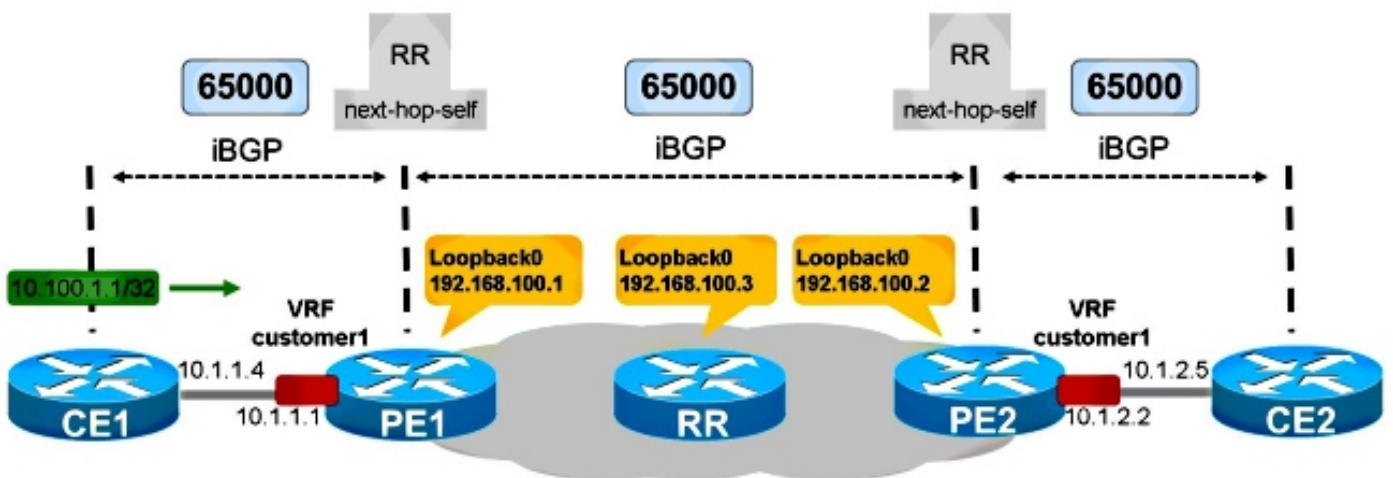


图 1

CE1与CE2在AS和SP网络一样：65000。PE1有iBGP配置往CE1。PE1反射前缀的10.100.1.1/32路径往在SP网络的RR。RR照常反射往PE路由器的iBGP路径。PE2反射往CE2的路径。

为使这能正常工作，您必须：

- 有iBGP PE-CE功能支持在PE1及PE2的代码
- 配置PE1及PE2为了执行在他们的BGP会话的路由反映往他们的各自CE路由器
- 有next-hop-self在BGP会话的PE路由器往他们的CE路由器
- 确保每个VPN站点使用不同的路由鉴别器(RD)

## 配置

参考图1。

这是PE1及PE2的需要的配置：

```
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

```

**注意：**如果PE没有CE邻居的邻接<internal-CE>内部VPN客户端命令，不传播从CE的前缀往SP RRs/PE路由器。

**注意：**如果PE不是在VRF的RR，不传播从RRs/PE路由器的前缀往CE路由器。

## new命令

有一个new命令，邻接<internal-CE>内部VPN客户端，做此feature工作。在仅PE路由器必须配置它往CE路由器的iBGP会话的。

**注意：**iBGP PE-CE多VRF CE (VRF-Lite)功能仍然支持，不用邻接<internal-CE>内部VPN客户端命令。

**注意：**当邻接<internal-CE>内部VPN客户端命令配置时，邻接<internal-CE>和route-reflector-client邻接<internal-CE> next-hop-self命令在配置里自动地放置。当route-reflector-client邻接<internal-CE>之一和邻接<internal-CE> next-hop-self命令(或两个)时删除和重新加载执行，然后他们是自动地放置的上一步在配置里。

## 详细的看看ATTR\_SET

参考的图1。

这是CE1通告的前缀：

```

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

```

当PE1接收从CE1时的BGP前缀10.100.1.1/32，两次存储它：

```

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

```

因为从CE1，接收第一个路径是PE1的实际路径。

第二个路径是通告往RRs/PE路由器的路径。它标记用来源的**ibgp**。它包含ATTR\_SET属性。注意此路径有一个或更多路由目标(Rts)附加对它。

PE1通告前缀如显示此处：

```

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

```

这是RR如何看到路径：

```

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

```

注意此Vpnv4单播前缀本地首选在核心的是100。在ATTR\_SET中，原始本地首选200存储。然而，这是透明对在SP核心的RR。

在PE2，您看到前缀如显示此处：

```
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
```

第一个路径是从RR接收的那个，与ATTR\_SET。注意RD是65000:1，始发地RD。第二个路径是从VRF表的已导入路径与RD 65000:1。ATTR\_SET删除。

这是路径如被看到在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
```

注意下一跳是10.1.2.2，是PE2。集群列表包含路由器PE1及PE2。这些是RR该问题在VPN里面。SP RR (10.100.1.3)不在集群列表。

本地首选200保留在间SP网络的VPN里面。

调试bgp vpnv4单播更新命令显示在SP网络传播的更新：

```

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,

```

**注意：**PE1接收其从RR的自己的更新然后丢弃了它。这是因为PE1及PE2在RR的同一更新组中。

**注意：**如果在十六进制要转存完整更新消息，请使用详细信息关键字调试BGP更新命令。

```

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,

```

**下一跳处理**

next-hop-self在此功能的PE路由器必须配置。对此的原因是通常下一跳是传输的不可更改与iBGP。然而，在这里有两个独立的网络：VPN网络和SP网络，用分开的内部网关协议(IGP)。因此，IGP量度不可能容易地比较和用于在两网络之间的最佳路径计算。RFC 6368选择的方法将有next-hop-self必需对于往CE的iBGP会话，避免一起以前描述的问题全部。优点是VRF站点能运行不同的IGP以此方法。

## RD

RFC 6368提及推荐同样VPN的不同的VRF站点使用另外(唯一) RDs。在Cisco IOS，这对于此功能是必需的。

## iBGP与Local-AS的PE-CE功能

参考的图2。VPN customer1有ASN 65001。

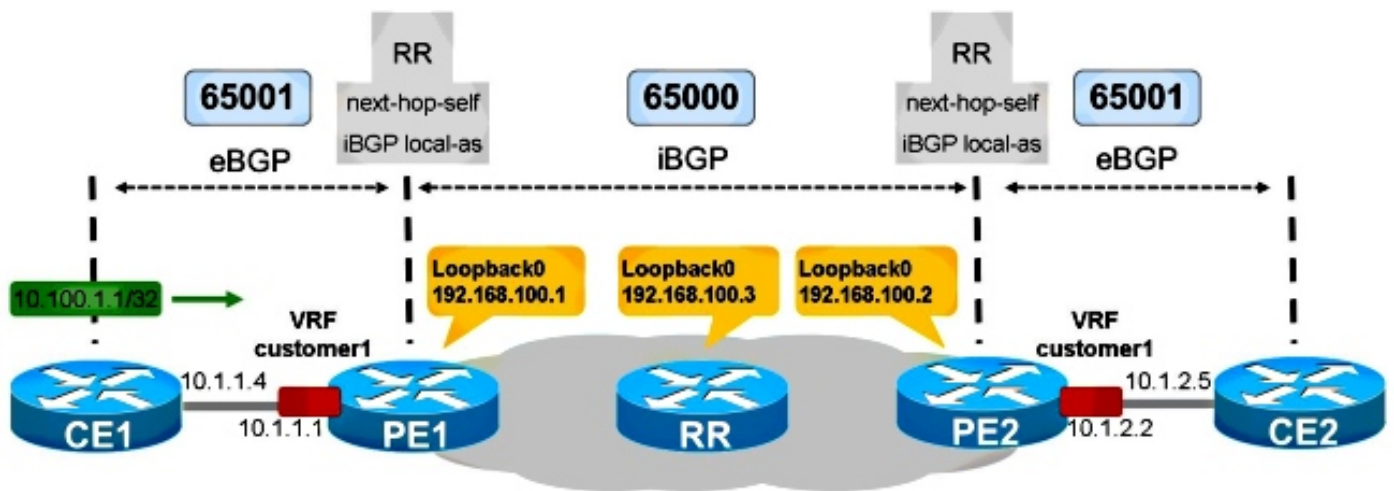


图 2

CE1是AS 65001。为了由PE1支架做此内部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类似配置。

PE1看到BGP前缀如显示此处：

```
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
```

前缀是内部BGP。

PE2看到此：

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

创建人AS是65001，是使用的AS，当前缀从PE2发送到CE2时。因此，AS保留，和，因此是在本例中的本地首选。

```
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
```

您看到本地而不是AS路径。这意味着它是内部BGP路由发起的于AS 65001，也是路由器CE2已配置的ASN。所有BGP属性从ATTR\_SET属性被采取了。这遵守案例的1规则在下一部分。

## 路由交换的规则区别VRF站点之间

ATTR\_SET自产生的VRF包含创建人。此产生的AS由远程PE检查，当删除ATTR\_SET时，在发送前缀对CE路由器前。

**第 1 种情况：**如果产生的AS匹配CE路由器的已配置的AS，则BGP属性从ATTR\_SET属性被采取，当PE导入路径到目的地VRF时。

**第 2 种情况：**如果产生的AS不匹配CE路由器的已配置的AS，则套被修建的路径的属性被采取如显示此处：

1. 路径属性设置为在ATTR\_SET属性包含的属性。
2. iBGP特定属性丢弃(LOCAL\_PREF、创建人和CLUSTER\_LIST)。
3. 在ATTR\_SET属性包含的始发地AS编号被加在前面对AS\_PATH并且遵从适用于在源和目的AS之间的外部BGP同位体的规则。
4. 如果自治系统关联与VRF是相同的象VPN供应商自治系统和VPN路由的AS\_PATH属性不空的，将被加在前面对VRF路由的AS\_PATH属性。

参考的图3. CE1和PE1有AS 65000和配置与iBGP PE-CE功能。CE2有ASN 65001。这意味着有在PE2和CE2之间的eBGP。

图 3

PE2看到路由如下：

```
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

```

这是前缀如被看到在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

```

这是案例2。在ATTR\_SET属性包含的始发地AS编号被加在前面对AS\_PATH由PE2并且遵从适用于eBGP对等体在源和目的AS之间的规则。当它创建将通告的路由对CE2时，iBGP特定属性由PE2忽略。因此，本地首选是100和没有200 (如在ATTR\_SET属性中看到)。

## CE对CE VRF-Lite反射

请参阅图 4。

图 4

图4显示一个另外的CE路由器，CE3，连接对PE1。CE1和CE3两个连接对在同一个VRF实例的PE1 : customer1。这意味着CE1和CE3是多VRF CE路由器(亦称VRF-Lite) PE1。当通告前缀从CE1到CE3时，PE1放置自己作为下一跳。在案件中此行为没有希望，您可能配置邻居10.1.3.6 next-hop-unchanged在PE1。为了配置此，您必须删除PE1的邻居10.1.3.6 next-hop-self。然后CE3看到从CE1的路由与CE1是那些BGP前缀的下一跳。为了做此工作，您需要那些BGP next-hop in的路由CE3路由表。您需要动态路由协议(IGP)或在CE1、PE1和CE3的静态路由为了确保路由器

有每其他的一个路由下一跳IP地址。然而，有与此配置的一问题。

在PE1的配置是：

```
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
```

从CE1的前缀在CE3优良被看到：

```
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
```

然而，从CE2的前缀在CE3被看到如显示此处：

```
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下一跳是**192.168.100.2**，PE2环回IP地址。当通告前缀10.100.1.2/32对CE3，PE1没有重写BGP下一跳对本身。这在CE3使此前缀不能使用。

因此，一旦iBGP PE-CE功能的混合在MPLS-VPN和iBGP VRF-Lite间的，您必须确保，您总是有next-hop-self在PE路由器。

反射从一个CE的iBGP路由到在VRF接口间的另一个CE本地在PE的您不能保留下一跳，当PE路由器是RR时。当您运行在间MPLS VPN网络时的iBGP PE-CE，您必须使用内部VPN客户端往CE路由器的iBGP会话。当您有超过一个本地CE在PE路由器时的VRF，然后您必须为那些BGP对等体保持next-hop-self。

您可能查看route-map为了设置下一跳为自己从其他PE路由器接收的前缀的，但是不从其他本地连接的CE路由器的反射的前缀的。然而，当前不支持它设置下一跳为在出局路由映射的自己。该配置显示此处：

```
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
```

然而，不支持这：

```
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
```

## 在PE路由器的更旧的Cisco IOS

如果PE1运行缺乏功能iBGP PE-CE的更旧的Cisco IOS软件，则PE1从未设置作为反射的iBGP前缀的下一跳。这意味着反射的BGP前缀(10.100.1.1/32)从CE1 (10.100.1.1)对CE2 -通过PE1-将有CE1 (10.1.1.4)作为下一跳。

```
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
```

从CE2 (10.100.1.2/32)的前缀在PE2看到作为下一跳，因为PE1不为此前缀执行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
```

为了iBGP PE-CE功能适当地工作，功能启用的VPN的所有PE路由器必须有安排的代码支持功能和功能启用。

## next-hop-self在VRF的eBGP

请参阅图 5。

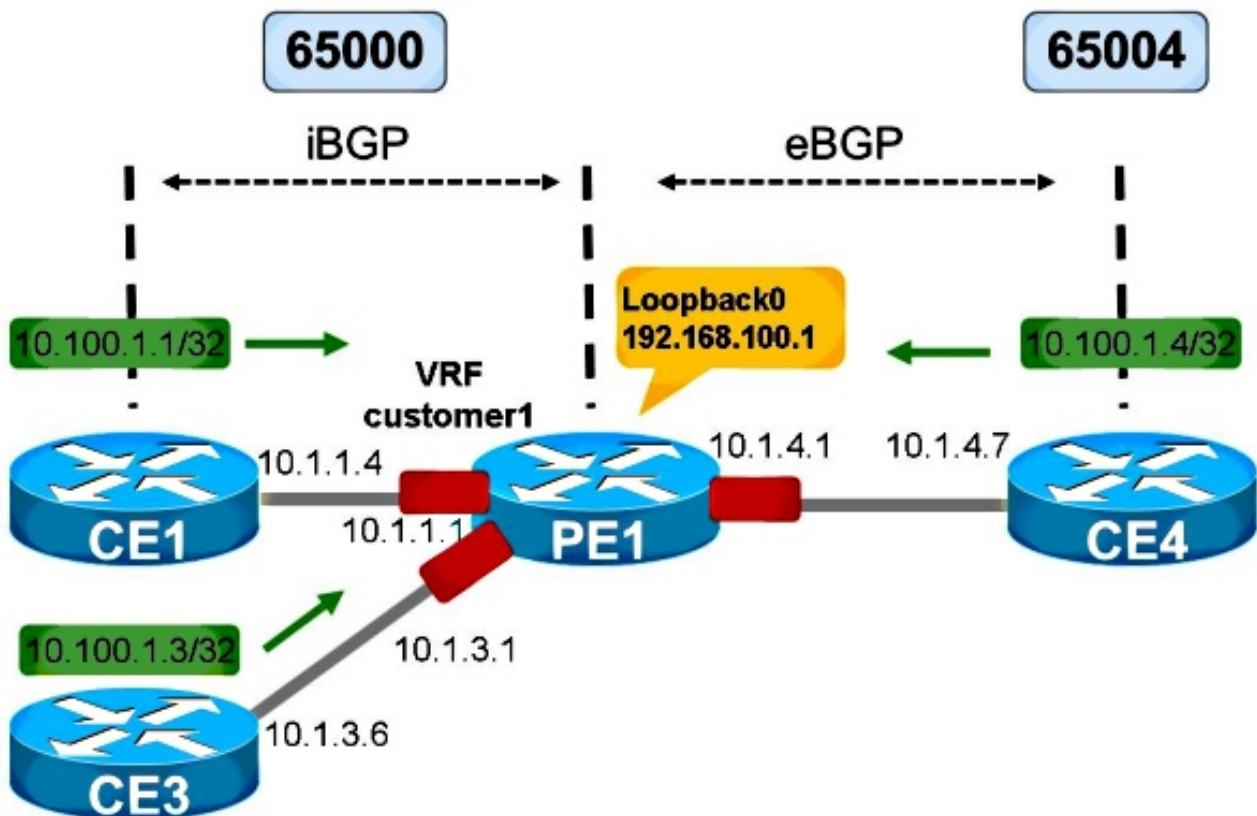


图 5

图5表示VRF-Lite设置。从PE1的会话往CE4是eBGP。从PE1的会话往CE3仍然是iBGP。

对于eBGP前缀，当通告往一个iBGP邻居的前缀VRF的时，下一跳总是设置为自已。这是不管事实往iBGP邻居的会话在VRF间next-hop-self是否设置。

在表5， CE3看到从CE4的前缀与PE1作为下一跳。

```
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
```

这发生在next-hop-self往CE3的无PE1或。

如果在PE1的接口往CE3和CE4没有在VRF，然而在全局上下文， next-hop-self往CE3产生变化。

没有next-hop-self往CE3的PE1，您看到：

```
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)
```

虽然next-hop-self隐含地启用，输出不指示此。

使用next-hop-self往CE3的PE1，您看到：

```
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
```

而，如果往CE3和CE4的接口在全局上下文，前缀的下一跳从CE4是CE4，当next-hop-self没有配置：

```
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
```

next-hop-self往CE3的PE1：

```
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
```

这根据RFC 4364执行。

如果要不为往一iBGP会话的eBGP前缀设置next-hop-self在VRF接口间，您必须配置next-hop-unchanged。此的支持只发生在Cisco Bug ID [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
```

现在，CE3看到CE4作为下一跳为CE4通告的前缀：

```
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
```

如果设法在Cisco Bug ID [CSCuj11720之前](#)配置iBGP会话的**next-hop-unchanged**关键字往在Cisco IOS代码的CE3，您遇到此错误：

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

在Cisco Bug ID [CSCuj11720以后](#)，**next-hop-unchanged**关键字为多跳跃EBGP邻居和iBGP VRF-Lite邻居是有效。