

# IOS-Implementierung der iBGP PE-CE-Funktion

## Inhalt

[Einführung](#)

[Hintergrundinformationen](#)

[Implementierung von iBGP PE-CE](#)

[Routenattribut für BGP-Kunden](#)

[Konfigurieren](#)

[Neuer Befehl](#)

[Detaillierte Übersicht über ATTR SET](#)

[Next-Hop-Verarbeitung](#)

[RD](#)

[iBGP-PE-CE-Funktion mit Local-AS](#)

[Regeln für den Routenaustausch zwischen verschiedenen VRF-Standorten](#)

[CE-to-CE VRF-Lite-Reflektion](#)

[Älteres Cisco IOS auf dem PE-Router](#)

[Next-Hop-Self für eBGP auf VRF](#)

## Einführung

In diesem Dokument wird beschrieben, wie das Internal Border Gateway Protocol (iBGP) zwischen Provider Edge (PE) und Customer Edge (CE) in Cisco IOS<sup>®</sup> implementiert wird.

## Hintergrundinformationen

Bis zur neuen iBGP-PE-CE-Funktion wurde iBGP zwischen PE und CE (also auf einer VRF-Schnittstelle (Virtual Routing and Forwarding) am PE-Router) nicht offiziell unterstützt. Eine Ausnahme bildet iBGP auf VRF-Schnittstellen in einer Multi-VRF-CE-Konfiguration (VRF-Lite). Die Motivation für die Bereitstellung dieser Funktion ist:

- Der Kunde möchte eine einzige autonome Systemnummer (ASN) an den mehreren Standorten der VRF haben, ohne dass eBGP (External Border Gateway Protocol) mit AS-override implementiert wird.
- Der Kunde möchte interne Routen-Reflektionen zu den CE-Routern bereitstellen, so als ob der SP-Core ein transparenter Routen-Reflektor (RR) wäre.

Mit dieser Funktion können die Standorte der VRF-Instanz dasselbe ASN wie der SP-Core aufweisen. Wenn sich das ASN der VRF-Standorte jedoch von dem ASN des SP-Core unterscheidet, kann es mit der Funktion Local-Autonomous System (AS) identisch dargestellt werden.

# Implementierung von iBGP PE-CE

Im Folgenden werden die beiden Hauptaspekte beschrieben, damit diese Funktion funktioniert:

- Dem BGP-Protokoll wurde ein neues Attribut ATTR\_SET hinzugefügt, um die VPN-BGP-Attribute im SP-Core transparent zu übertragen.
- RR für die iBGP-Sitzungen zu den CE-Routern in der VRF-Instanz und als RR zu den VPNv4-Nachbarn (andere PE-Router oder RRs)

Das neue ATTR\_SET-Attribut ermöglicht es dem SP, alle BGP-Attribute des Kunden transparent zu übertragen und stört nicht die SP-Attribute und BGP-Richtlinien. Solche Attribute sind die Cluster-Liste, lokale Präferenzen, Communitys usw.

## Routenattribut für BGP-Kunden

ATTR\_SET ist das neue BGP-Attribut, mit dem die VPN-BGP-Attribute des SP-Kunden übertragen werden. Es ist ein optionales transitives Attribut. In diesem Attribut können alle BGP-Attribute des Kunden aus der BGP Update-Nachricht mit Ausnahme der Attribute MP\_REACH und MP\_UNREACH übertragen werden.

Das ATTR\_SET-Attribut hat folgendes Format:

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

Die Attributflags sind die regulären BGP-Attributflags (siehe RFC 4271). Die Attributlänge gibt an, ob die Attributlänge ein oder zwei Oktette ist. Der Zweck des Origin AS-Felds besteht darin, das Leck einer Route, die von einem AS ausgeht, an ein anderes AS zu übertragen, ohne dass die AS\_PATH ordnungsgemäß geändert wird. Das Feld Path Attributes variabler Länge enthält die VPN-BGP-Attribute, die über den SP-Core übertragen werden müssen.

Auf dem Egress-PE-Router werden die VPN-BGP-Attribute in dieses Attribut übertragen. Auf dem Eingangs-PE-Router werden diese Attribute aus dem Attribut entfernt, bevor das BGP-Präfix an den CE-Router gesendet wird. Dieses Attribut ermöglicht die Isolierung von BGP-Attributen zwischen dem SP-Netzwerk und dem Kunden-VPN und umgekehrt. Beispielsweise wird das Cluster-Listenattribut für die SP-Route-Reflektion nicht im VPN-Netzwerk angezeigt und berücksichtigt. Das Attribut der VPN-Route-Reflection-Cluster-Liste wird jedoch nicht im SP-Netzwerk angezeigt und berücksichtigt.

In Abbildung 1 sehen Sie die Verbreitung eines Kunden-BGP-Präfixes im SP-Netzwerk.

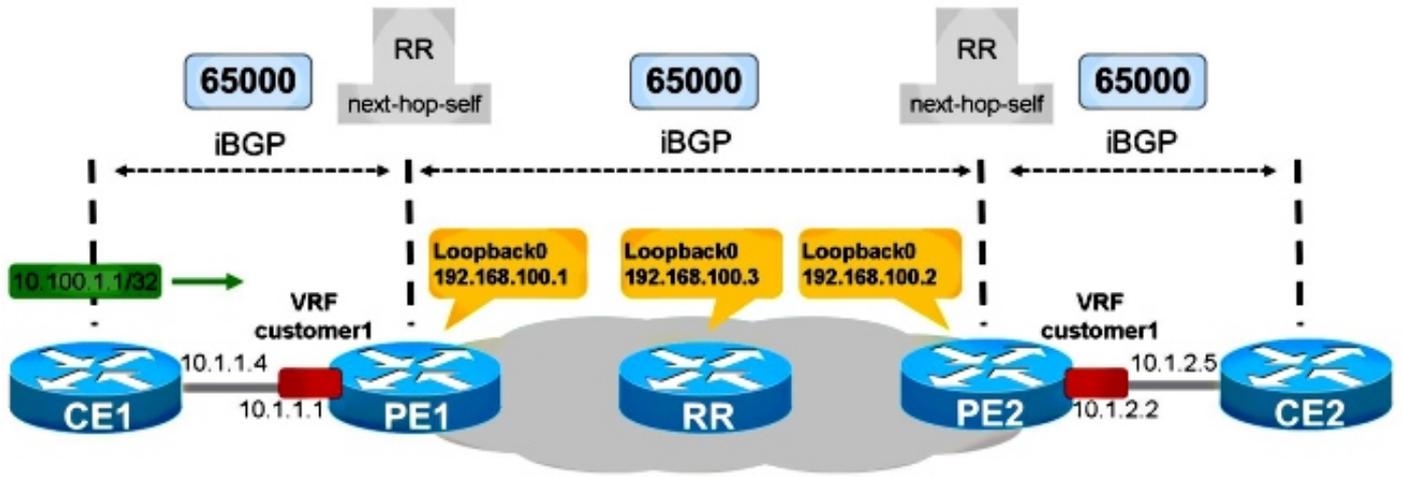


Abbildung 1

CE1 und CE2 sind mit dem SP-Netzwerk identisch: 65000 Für PE1 wurde iBGP für CE1 konfiguriert. PE1 gibt den Pfad für das Präfix 10.100.1.1/32 zum RR im SP-Netzwerk an. Der RR gibt wie gewohnt den iBGP-Pfad zu den PE-Routern wieder. PE2 spiegelt den Pfad zu CE2 wider.

Damit dies ordnungsgemäß funktioniert, müssen Sie:

- Code auf PE1 und PE2 mit Unterstützung für iBGP-PE-CE-Funktionen
- Konfigurieren von PE1 und PE2 zur Durchführung von Routen-Reflexionen in der BGP-Sitzung mit den entsprechenden CE-Routern
- Next-Hop-Self auf den PE-Routern für die BGP-Sitzung zu ihren CE-Routern
- Stellen Sie sicher, dass jeder VPN-Standort unterschiedliche Route Distinguisher (RD) verwendet.

## Konfigurieren

Siehe Abbildung 1.

Die erforderliche Konfiguration für PE1 und PE2 ist wie folgt:

```
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

```

**Hinweis:** Verfügt der PE nicht über den Befehl `neighbor <internal-CE> internal-vpn-client` für den CE-Nachbarn, werden die Präfixe vom CE nicht an die SP-RRs/PE-Router weitergeleitet.

**Hinweis:** Wenn der PE nicht der RR in der VRF-Instanz ist, werden die Präfixe von den RRs/PE-Routern nicht an den CE-Router weitergeleitet.

## Neuer Befehl

Der neue Befehl `neighbor <internal-CE> internal-vpn-client` soll diese Funktion ermöglichen. Sie muss auf dem PE-Router nur für die iBGP-Sitzung mit den CE-Routern konfiguriert werden.

**Hinweis:** Die Funktion iBGP PE-CE Multi-VRF CE (VRF-Lite) wird weiterhin ohne den Befehl `neighbor <internal-CE> internal-vpn-client` unterstützt.

**Hinweis:** Wenn der Befehl `neighbor <internal-CE> internal-vpn-client` konfiguriert ist, werden die Befehle `neighbor <internal-CE> route-reflektor-client` und `neighbor <internal-CE> next-hop-self` ebenfalls automatisch in die Konfiguration übernommen. Wenn einer der **Nachbarn <internal-CE> Route-Reflector-Client** und **Nachbarn <internal-CE> Next-Hop-Self-Befehle** (oder beide) entfernt werden und ein Neuladen durchgeführt wird, werden sie automatisch wieder in die Konfiguration aufgenommen.

## Detaillierte Übersicht über ATTR\_SET

Siehe Abbildung 1.

Dies ist das von CE1 angegebene Präfix:

```
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
```

Wenn PE1 das BGP-Präfix 10.100.1.1/32 von CE1 empfängt, wird es zweimal gespeichert:

```
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
```

Der erste Pfad ist der tatsächliche Pfad auf PE1, da er von CE1 empfangen wird.

Der zweite Pfad ist der Pfad, der den RRs/PE-Routern angekündigt wird. Er ist mit **ibgp-sourcing** markiert. Sie enthält das ATTR\_SET-Attribut. Beachten Sie, dass diesem Pfad mindestens ein Route Targets (RTs) zugeordnet ist.

PE1 gibt das Präfix wie folgt an:

```
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

So sieht der RR den Pfad:

```
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
```

Beachten Sie, dass die lokale Präferenz dieses VPNv4-Unicast-Präfix im Core 100 beträgt. Im ATTR\_SET wird die ursprüngliche lokale Voreinstellung von 200 gespeichert. Dies ist jedoch für den RR im SP-Core transparent.

Auf PE2 wird das Präfix wie folgt angezeigt:

```
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

```

Der erste Pfad ist der, der vom RR mit ATTR\_SET empfangen wurde. Beachten Sie, dass der RD 65000:1, der Ursprungs-RD, ist. Der zweite Pfad ist der importierte Pfad aus der VRF-Tabelle mit RD 65000:1. ATTR\_SET wurde entfernt.

Dies ist der Pfad, der auf CE2 angezeigt wird:

```

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

```

Beachten Sie, dass der Next-Hop **10.1.2.2** ist, d. h. PE2. Die Cluster-Liste enthält Router PE1 und PE2. Dies sind die RRs, die innerhalb des VPNs wichtig sind. Der SP-RR (10.100.1.3) ist nicht in der Cluster-Liste enthalten.

Die lokale Präferenz von 200 wurde innerhalb des VPN im gesamten SP-Netzwerk beibehalten.

Der Befehl **debug bgp vpnv4 unicast updates** zeigt die im SP-Netzwerk weitergeleitete Aktualisierung an:

```

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,
```

**Hinweis:** PE1 erhielt sein eigenes Update von RR und verwarf es dann. Dies liegt daran, dass PE1 und PE2 in derselben Aktualisierungsgruppe auf dem RR sind.

**Hinweis:** Wenn Sie die vollständige Aktualisierungsmeldung im Hexadezimalformat auslesen möchten, verwenden Sie das **detail**-Schlüsselwort für den Befehl **debug BGP updates**.

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,
```

## Next-Hop-Verarbeitung

Für diese Funktion muss Next-Hop-Self auf den PE-Routern konfiguriert werden. Der Grund dafür ist, dass der Next-Hop normalerweise unverändert mit iBGP transportiert wird. Es gibt jedoch zwei separate Netzwerke: VPN-Netzwerk und SP-Netzwerk, die separate Interior Gateway Protocols (IGPs) ausführen. Daher kann die IGP-Metrik nicht einfach verglichen und für die Berechnung des besten Pfads zwischen den beiden Netzwerken verwendet werden. Der RFC 6368 hat sich für eine Next-Hop-Selbstverpflichtung für die iBGP-Sitzung zum CE entschieden, wodurch das zuvor beschriebene Problem zusammen vermieden wird. Ein Vorteil ist, dass an den VRF-Standorten verschiedene IGPs mit diesem Ansatz ausgeführt werden können.

RD

Laut RFC 6368 sollten verschiedene VRF-Standorte desselben VPN unterschiedliche (eindeutige) RDs verwenden. In Cisco IOS ist dies für diese Funktion obligatorisch.

## iBGP-PE-CE-Funktion mit Local-AS

Siehe Abbildung 2. Der VPN-Kunde1 hat ASN 65001.

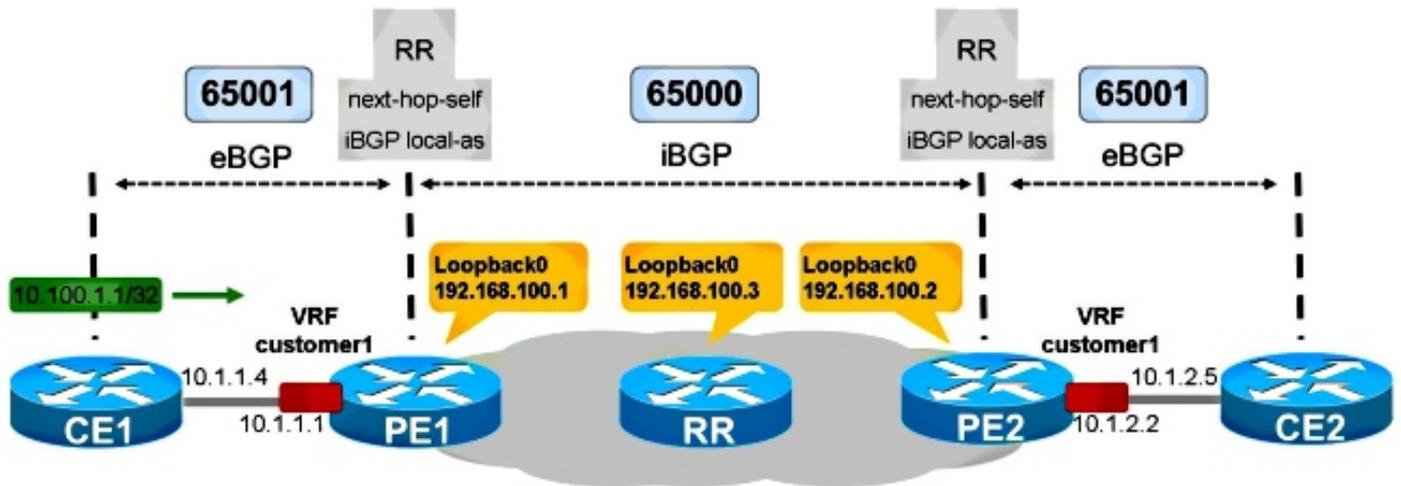


Abbildung 2

CE1 ist in AS 65001. Um dieses interne BGP vom Standpunkt von PE1 aus zu erstellen, ist die lokale iBGP-Funktion erforderlich.

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 und CE2 werden ähnlich konfiguriert.

PE1 sieht das BGP-Präfix wie folgt:

```
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
```

Das Präfix ist ein internes BGP.

PE2 sieht Folgendes:

```
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
```

Der Originator AS ist 65001. Dies ist das AS, das beim Versenden des Präfixes vom PE2 an CE2 verwendet wird. Das AS wird also beibehalten, und die lokale Präferenz in diesem Beispiel.

```

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

```

Sie sehen **Lokal** anstelle eines AS-Pfads. Dies bedeutet, dass es sich um eine interne BGP-Route handelt, die vom AS 65001 ausgeht und ebenfalls das konfigurierte ASN des Router-CE2 ist. Alle BGP-Attribute wurden aus dem ATTR\_SET-Attribut übernommen. Dies entspricht den Regeln für Fall 1 im nächsten Abschnitt.

## Regeln für den Routenaustausch zwischen verschiedenen VRF-Standorten

Der ATTR\_SET enthält das Originator-AS der ursprünglichen VRF-Instanz. Dieses Ursprungs-AS wird vom Remote-PE überprüft, wenn es den ATTR\_SET entfernt, bevor es das Präfix an den CE-Router sendet.

**Fall 1:** Wenn das ursprüngliche AS mit dem für den CE-Router konfigurierten AS übereinstimmt, werden die BGP-Attribute aus dem ATTR\_SET-Attribut übernommen, wenn der PE den Pfad in die Ziel-VRF importiert.

**Fall 2:** Wenn das ursprüngliche AS nicht mit dem für den CE-Router konfigurierten AS übereinstimmt, werden die Attribute für den konstruierten Pfad wie folgt verwendet:

1. Die Pfadattribute werden auf die im ATTR\_SET-Attribut enthaltenen Attribute festgelegt.
2. Die iBGP-spezifischen Attribute werden verworfen (LOCAL\_PREF, ORIGINATOR und CLUSTER\_LIST).
3. Die im ATTR\_SET-Attribut enthaltene **Origin AS**-Nummer wird dem AS\_PATH vorangestellt und befolgt die Regeln, die für ein externes BGP-Peering zwischen den Quell- und Ziel-AS gelten.
4. Wenn das mit der VRF-Instanz verbundene autonome System mit dem autonomen VPN-Provider-System identisch ist und das AS\_PATH-Attribut der VPN-Route nicht leer ist, MUSS dem AS\_PATH-Attribut der VRF-Route vorangestellt werden.

Siehe Abbildung 3. CE1 und PE1 verfügen über das AS 65000 und sind mit der iBGP-PE-CE-Funktion konfiguriert. CE2 hat ASN 65001. Dies bedeutet, dass ein eBGP zwischen PE2 und CE2 vorhanden ist.

Abbildung 3

PE2 sieht die Route wie folgt:

```

PE2#show bgp vpv4 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

```

Dies ist das Präfix, das auf CE2 angezeigt wird:

```

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

```

Dies ist Fall 2. Die im ATTR\_SET-Attribut enthaltene **Origin AS**-Nummer wird dem AS\_PATH von PE2 vorangestellt und befolgt die Regeln, die für ein eBGP-Peering zwischen dem Quell- und dem Ziel-AS gelten. Die iBGP-spezifischen Attribute werden von PE2 ignoriert, wenn die Route erstellt wird, die CE2 mitgeteilt werden soll. Die lokale Voreinstellung ist also 100 und nicht 200 (wie im ATTR\_SET-Attribut dargestellt).

## CE-to-CE VRF-Lite-Reflektion

Siehe Abbildung 4.

Abbildung 4 zeigt einen zusätzlichen CE-Router, CE3, der mit PE1 verbunden ist. CE1 und CE3 sind beide mit PE1 auf derselben VRF-Instanz verbunden: Kunde1. Dies bedeutet, dass CE1 und CE3 Multi-VRF-CE-Router (auch bekannt als VRF-Lite) von PE1 sind. PE1 stellt sich selbst als Next-Hop dar, wenn es Präfixe von CE1 an CE3 meldet. Falls dieses Verhalten nicht erwünscht ist, können Sie **Next-Hop-unveränderten Nachbar 10.1.3.6** auf PE1 konfigurieren. Um dies zu konfigurieren, müssen Sie **Next-Hop-Self (Next-Hop-Self)** aus dem **Nachbarn 10.1.3.6** auf PE1 entfernen. Anschließend wird von CE3 erkannt, dass die Routen von CE1 mit CE1 der nächste Hop für diese BGP-Präfixe sind. Damit dies funktioniert, müssen die Routen für diese BGP Next-Hops in der Routing-Tabelle von CE3 angegeben werden. Sie benötigen ein dynamisches Routing-Protokoll (IGP) oder statische Routen auf CE1, PE1 und CE3, um sicherzustellen, dass die Router über eine Route für die anderen Next-Hop-IP-Adressen verfügen. Es besteht jedoch ein Problem mit dieser Konfiguration.

Konfiguration für PE1:

```
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
```

Das Präfix von CE1 wird auf CE3 einwandfrei angezeigt:

```
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
```

Das Präfix von CE2 wird jedoch auf CE3 angezeigt, wie hier gezeigt:

```
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
```

Der BGP Next-Hop ist **192.168.100.2**, die Loopback-IP-Adresse von PE2. PE1 hat den BGP Next-Hop nicht selbst umgeschrieben, als es CE3 das Präfix 10.100.1.2/32 bekannt gab. Dadurch ist

dieses Präfix auf CE3 unbrauchbar.

Bei einer Kombination der iBGP-PE-CE-Funktion mit MPLS-VPN und iBGP-VRF-Lite müssen Sie also sicherstellen, dass die PE-Router stets über Next-Hop-Self verfügen.

Der Next-Hop kann nicht beibehalten werden, wenn es sich bei einem PE-Router um einen RR handelt, der iBGP-Routen von einem CE zu einem anderen CE über lokale VRF-Schnittstellen auf dem PE widerspiegelt. Wenn Sie iBGP PE-CE in einem MPLS-VPN-Netzwerk ausführen, müssen Sie den **internen VPN-Client** für die iBGP-Sitzungen zu den CE-Routern verwenden. Wenn sich in einem VRF auf einem PE-Router mehr als ein lokaler CE befindet, müssen Sie **Next-Hop-Self** für diese BGP-Peers beibehalten.

Sie können Routing-Maps anzeigen, um den Next-Hop für Präfixe festzulegen, die von anderen PE-Routern empfangen wurden, jedoch nicht für reflektierte Präfixe anderer lokal verbundener CE-Router. Es wird jedoch derzeit nicht unterstützt, den Next-Hop in einer ausgehenden route-map auf sich selbst festzulegen. Diese Konfiguration wird hier angezeigt:

```
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
!
```

Dies wird jedoch nicht unterstützt:

```
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
```

## Älteres Cisco IOS auf dem PE-Router

Wenn PE1 ältere Cisco IOS-Software ausführt, die nicht über die Funktion iBGP PE-CE verfügt, legt sich PE1 nie als Next-Hop für die reflektierten iBGP-Präfixe fest. Das bedeutet, dass das

reflektierte BGP-Präfix (10.100.1.1/32) von CE1 (10.100.1.1) zu CE2 (über PE1) CE1 (10.1.1.4) als Next-Hop fungieren würde.

```
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
```

Das Präfix von CE2 (10.100.1.2/32) wird mit PE2 als Next-Hop betrachtet, da PE1 für dieses Präfix auch kein Next-Hop-Self ausführt:

```
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
```

Damit die iBGP-PE-CE-Funktion ordnungsgemäß funktioniert, müssen bei allen PE-Routern für das VPN, für das die Funktion aktiviert ist, der Code zur Unterstützung der Funktion vorhanden sein und die Funktion aktiviert sein.

## Next-Hop-Self für eBGP auf VRF

Siehe Abbildung 5.

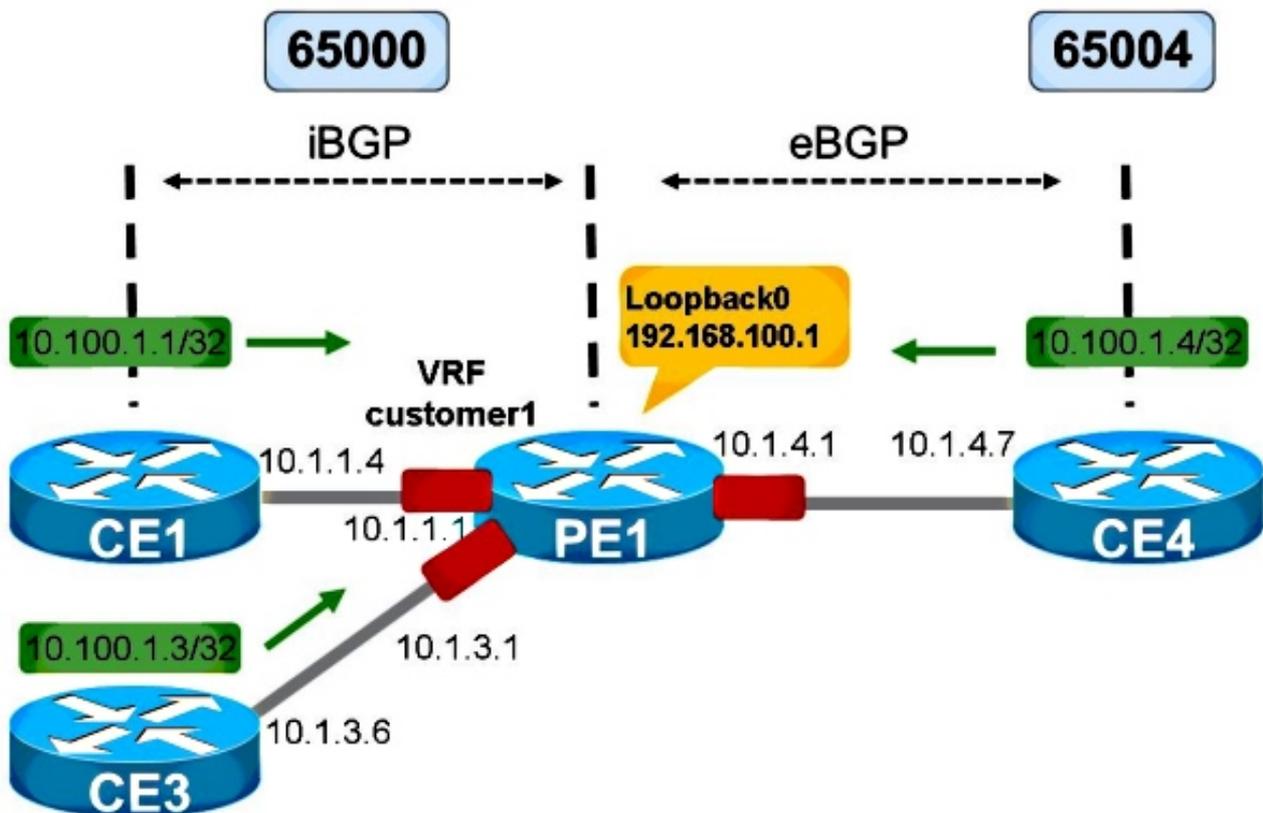


Abbildung 5

Abbildung 5 zeigt eine VRF-Lite-Konfiguration. Die Sitzung von PE1 zu CE4 ist eBGP. Die Sitzung von PE1 zu CE3 ist immer noch iBGP.

Für eBGP-Präfixe wird der Next-Hop immer auf sich selbst festgelegt, wenn er die Präfixe einem iBGP-Nachbarn auf VRF meldet. Dies gilt unabhängig davon, ob die Sitzung zum iBGP-Nachbarn über VRF den Next-Hop-Self-Set aufweist oder nicht.

In Abbildung 5 sieht CE3 die Präfixe von CE4 mit PE1 als Next-Hop.

```
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
```

Dies geschieht bei Next-Hop-Self auf PE1 zu CE3 oder ohne.

Wenn sich die Schnittstellen auf PE1 zu CE3 und CE4 nicht in einer VRF-Instanz befinden, aber im globalen Kontext, macht das Next-Hop-Self-To-CE3 einen Unterschied.

Ohne Next-Hop-Self auf PE1 zu CE3 sehen Sie:

```
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)
```

Obwohl Next-Hop-Self implizit aktiviert ist, gibt die Ausgabe dies nicht an.

Bei Next-Hop-Self auf PE1 zu CE3 sehen Sie:

```
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
```

Wenn sich die Schnittstellen zu CE3 und CE4 in einem globalen Kontext befinden, ist der nächste Hop für Präfixe von CE4 selbst CE4, wenn Next-Hop-Self nicht konfiguriert ist:

```
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
```

Für Next-Hop-Self auf PE1 zu 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
```

Dies wurde auf Basis von RFC 4364 durchgeführt.

Wenn Sie das Next-Hop-Self für eBGP-Präfixe für eine iBGP-Sitzung über eine VRF-Schnittstelle nicht festlegen möchten, müssen Sie **Next-Hop-unveränderlich** konfigurieren. Die Unterstützung dafür fand nur mit der Cisco Bug-ID [CSCuj11720](#) statt.

```
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
```

Nun sieht CE3 CE4 als Next-Hop für die von CE4 angekündigten Präfixe:

```
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
```

Wenn Sie versuchen, das **Next-Hop-unveränderte** Schlüsselwort für die iBGP-Sitzung zu CE3 im Cisco IOS-Code vor der Cisco Bug-ID [CSCuj11720](#) zu konfigurieren, wird folgender Fehler angezeigt:

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

Nach der Cisco Bug-ID [CSCuj11720](#) ist das **Next-Hop-unveränderte** Schlüsselwort für Multi-Hop eBGP-Nachbarn und iBGP VRF-Lite-Nachbarn gültig.