Implementazione IOS della funzionalità iBGP PE-CE

Sommario

Introduzione

In questo documento viene descritto come implementare il protocollo iBGP (Internal Border Gateway Protocol) tra Provider Edge (PE) e Customer Edge (CE) in Cisco IOS®.

Premesse

Fino alla nuova funzionalità iBGP PE-CE, iBGP tra PE e CE (quindi su un'interfaccia VRF (Virtual Routing and Forwarding) sul router PE) non era ufficialmente supportato. Un'eccezione è iBGP sulle interfacce VRF in una configurazione Multi-VRF CE (VRF-Lite). Motivo dell'installazione di questa funzionalità:

- Il cliente desidera avere un singolo ASN (Autonomous System Number) su più siti del VRF, senza la distribuzione di eBGP (External Border Gateway Protocol) con sostituzione automatica.

- Il cliente desidera fornire una riflessione interna del percorso verso i router CE, come se il core SP (Service Provider) fosse un riflettore di percorso trasparente (RR).

Grazie a questa funzionalità, i siti del VRF possono avere lo stesso ASN del core SP. Tuttavia, nel caso in cui l'ASN dei siti VRF sia diverso dall'ASN del core SP, è possibile fare in modo che

appaia uguale con l'utilizzo della funzionalità Local-Autonomous System (AS).

Implementazione di iBGP PE-CE

Di seguito sono riportate le due parti principali per il funzionamento di questa funzionalità:

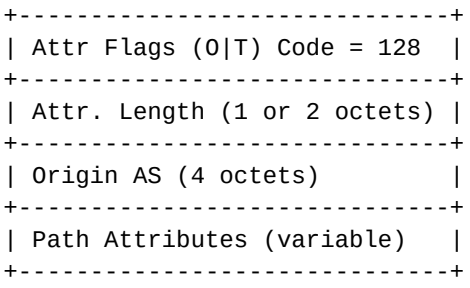
- Èstato aggiunto un nuovo attributo ATTR_SET al protocollo BGP per trasportare gli attributi VPN BGP nel core SP in modo trasparente.
- Rendere il router PE un RR per le sessioni iBGP verso i router CE nel VRF e come RR verso i router adiacenti VPNv4 (altri router PE o RR).

Il nuovo attributo ATTR_SET consente all'SP di trasportare tutti gli attributi BGP del cliente in modo trasparente e non interferisce con gli attributi SP e le policy BGP. Tali attributi sono l'elenco di cluster, le preferenze locali, le comunità e così via.

Attributo route cliente BGP

ATTR_SET è il nuovo attributo BGP utilizzato per trasportare gli attributi VPN BGP del cliente SP. Si tratta di un attributo transitivo facoltativo. In questo attributo è possibile trasportare tutti gli attributi BGP del cliente dal messaggio di aggiornamento BGP, ad eccezione degli attributi MP_REACH e MP_UNREACH.

Il formato dell'attributo ATTR_SET è il seguente:



I flag di attributo sono i normali flag di attributo BGP (fare riferimento alla RFC 4271). La lunghezza dell'attributo indica se la lunghezza dell'attributo è uno o due ottetti. Lo scopo del campo AS origine è impedire che la perdita di una route originata in un AS venga trasferita a un altro AS senza la corretta manipolazione di AS_PATH. Il campo Attributi percorso di lunghezza variabile contiene gli attributi VPN BGP che devono essere trasferiti nel core dell'SP.

Sul router PE in uscita, gli attributi VPN BGP vengono inseriti in questo attributo. Sul router PE in entrata, questi attributi vengono estratti dall'attributo prima che il prefisso BGP venga inviato al router CE. Questo attributo fornisce l'isolamento degli attributi BGP tra la rete SP e la VPN del cliente e viceversa. Ad esempio, l'attributo dell'elenco di cluster di riflessione route SP non viene visualizzato e considerato all'interno della rete VPN. Inoltre, l'attributo dell'elenco dei cluster di riflessione della route VPN non viene visualizzato e considerato all'interno della rete SP.

Nella Figura 1 è illustrata la propagazione di un prefisso BGP del cliente nella rete SP.

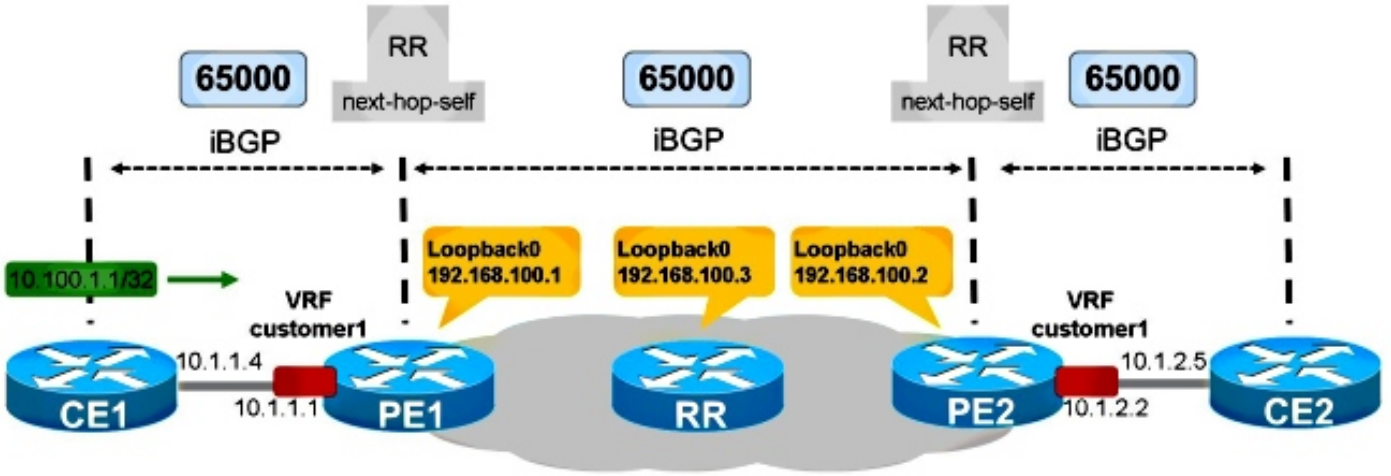


Figura 1

CE1 e CE2 si trovano nella stessa AS della rete SP: 65000. In PE1 è configurato iBGP verso CE1. PE1 riflette il percorso del prefisso 10.100.1.1/32 verso RR nella rete SP. Il record di risorse riflette il percorso iBGP verso i router PE come di consueto. PE2 riflette il percorso verso CE2.

Affinché questa operazione funzioni correttamente, è necessario:

- Disporre di codice in PE1 e PE2 con supporto per la funzionalità iBGP PE-CE
- Configurare PE1 e PE2 in modo da eseguire la riflessione delle route sulla sessione BGP verso i rispettivi router CE
- Eseguire l'hop successivo sui router PE per la sessione BGP verso i router CE
- Verificare che ogni sito VPN utilizzi distintori di route (RD) diversi

Configurazione

Fare riferimento alla Figura 1.

Ecco la configurazione necessaria per PE1 e PE2:

```
PE1

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

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

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

PE2

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

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

> **Nota:** Se il PE non dispone del comando **neighbors <internal-CE>internal-vpn-client** per il router adiacente CE, non propaga i prefissi dal CE ai router RR/PE SP.

> **Nota:** Se il tipo PE non è il tipo RR nel VRF, non propaga i prefissi dai router RR/PE al router CE.

Nuovo comando

Per il corretto funzionamento di questa funzione, è disponibile un nuovo comando **<internal-CE> internal-vpn-client**, Deve essere configurato sul router PE solo per la sessione iBGP verso i router CE.

> **Nota:** La funzionalità iBGP PE-CE Multi-VRF CE (VRF-Lite) è ancora supportata senza il comando **<internal-CE> internal-vpn-client** del **router adiacente**.

Nota: Quando si configura il comando **neighbors <internal-CE>internal-vpn-client**, anche i comandi **<internal-CE> route-reflector-client** e **next-hop-self <internal-CE>adiacenti** vengono inseriti automaticamente nella configurazione. Quando vengono rimossi uno dei comandi **<internal-CE>route-reflector-client** e **<internal-CE>next-hop-self** (o entrambi) del router **adiacente** e viene eseguito un ricaricamento, i comandi vengono automaticamente ripristinati nella configurazione.

Analisi dettagliata di ATTR_SET

Fare riferimento alla Figura 1.

Questo è il prefisso annunciato da CE1:

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

Quando PE1 riceve il prefisso BGP 10.100.1.1/32 da CE1, lo memorizza due volte:

```
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
```

Il primo percorso è il percorso effettivo in PE1, poiché viene ricevuto da CE1.

Il secondo percorso è il percorso annunciato verso i router R/PE. È contrassegnato con **origine ibgp**. Contiene l'attributo ATTR_SET. A questo percorso sono collegati uno o più oggetti Route (RT).

PE1 annuncia il prefisso come illustrato di seguito:

```
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
```

Ecco come l'RR vede il percorso:

```
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
```

Si noti che la preferenza locale di questo prefisso unicast VPNv4 nel core è 100. In ATTR_SET, viene memorizzata la preferenza locale originale di 200. Tuttavia, ciò è trasparente al RR nel core SP.

In PE2 è possibile visualizzare il prefisso come illustrato di seguito:

```
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
```

Il primo percorso è quello ricevuto dall'RR, con il parametro ATTR_SET. Notare che il RD è 65000:1, il RD di origine. Il secondo percorso è il percorso importato dalla tabella VRF con RD 6500:1. ATTR_SET è stato rimosso.

Questo è il percorso visualizzato in 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
```

Si noti che l'hop successivo è 10.1.2.2, ovvero PE2. L'elenco dei cluster contiene i router PE1 e PE2. Si tratta degli RR rilevanti all'interno della VPN. L'SP RR (10.100.1.3) non è presente nell'elenco dei cluster.

La preferenza locale di 200 è stata mantenuta all'interno della VPN attraverso la rete SP.

Il comando debug bgp vpnv4 unicast updates mostra l'aggiornamento propagato nella rete SP:

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

RR#
BGP(4): 192.168.100.1 rcvd UPDATE w/ attr: nexthop 192.168.100.1, origin i, localpref
100, originator 10.100.1.1, clusterlist 192.168.100.1, extended community RT:1:1,
[ATTR_SET attribute:  originator AS 65000, origin IGP, aspath , med 0, localpref 200,
cluster list 192.168.100.1 , originator 10.100.1.1]
BGP(4): 192.168.100.1 rcvd 65000:1:10.100.1.1/32, label 16
```

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

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

Nota: PE1 ha ricevuto un proprio aggiornamento da RR e quindi lo ha eliminato. PE1 e PE2 si trovano infatti nello stesso gruppo di aggiornamento di RR.

Nota: Se si desidera eseguire il dump del messaggio Update completo in formato esadecimale, utilizzare la parola chiave **detail** per il comando **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 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,
```

Gestione hop successivo

Per questa funzionalità è necessario configurare Next-Hop-Self sui router PE. Il motivo è che normalmente l'hop successivo viene trasportato senza modifiche con iBGP. Tuttavia, esistono due reti separate: la rete VPN e la rete SP, che eseguono IGP (Interior Gateway Protocol) separati. Pertanto, la metrica IGP non può essere facilmente confrontata e utilizzata per il calcolo del miglior percorso tra le due reti. L'approccio scelto dalla RFC 6368 prevede l'obbligatorietà dell'hop successivo per la sessione iBGP verso il CE, evitando così il problema descritto in precedenza. Un vantaggio è che i siti VRF possono eseguire diversi IGP con questo approccio.

RD

La RFC 6368 indica che è consigliabile che siti VRF diversi della stessa VPN utilizzino RD diversi

(univoci). In Cisco IOS, questa operazione è obbligatoria per questa funzione.

Funzione iBGP PE-CE con Local-AS

Fare riferimento alla Figura 2. Il cliente VPN 1 ha ASN 65001.

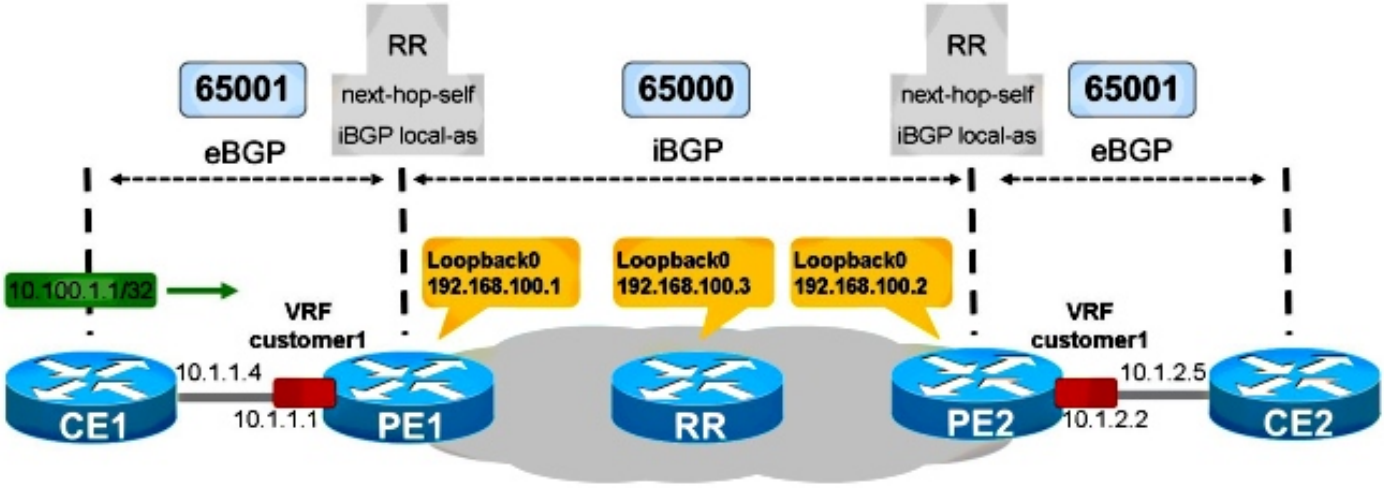


Figura 2

CE1 è in AS 65001. Per rendere questo BGP interno dal punto di vista di PE1, è necessario che iBGP disponga della funzionalità iBGP local-as.

```
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 e CE2 sono configurati in modo simile.

PE1 vede il prefisso BGP come mostrato di seguito:

```
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
```

Il prefisso è un BGP interno.

PE2 rileva quanto segue:

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

L'AS creatore è 65001, ovvero l'AS utilizzato quando il prefisso viene inviato da PE2 a CE2. In questo esempio, l'AS viene mantenuto, così come la preferenza locale.

```
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
```

Viene visualizzato **Locale** anziché un percorso AS. Ciò significa che si tratta di una route BGP interna originaria di AS 65001, che è anche l'ASN configurato del router CE2. Tutti gli attributi BGP sono stati presi dall'attributo ATTR_SET. In questo modo, vengono rispettate le regole relative al caso 1 nella sezione successiva.

Regole per lo scambio di route tra siti VRF diversi

Il parametro ATTR_SET contiene il valore AS di origine del VRF di origine. Questo SA di origine viene controllato dal PE remoto quando rimuove il valore ATTR_SET prima di inviare il prefisso al router CE.

Caso 1: Se l'AS di origine corrisponde all'AS configurato per il router CE, gli attributi BGP vengono ricavati dall'attributo ATTR_SET quando il PE importa il percorso nel VRF di destinazione.

Caso 2: Se l'AS di origine non corrisponde all'AS configurato per il router CE, il set di attributi per il percorso costruito viene preso come mostrato di seguito:

1. Gli attributi path vengono impostati sugli attributi contenuti nell'attributo ATTR_SET.

2. Gli attributi specifici di iBGP vengono eliminati (LOCAL_PREF, ORIGINATOR e CLUSTER_LIST).

3. Il numero **Origin AS** contenuto nell'attributo ATTR_SET viene anteposto a AS_PATH e segue le regole che si applicano a un peering BGP esterno tra il SA di origine e quello di destinazione.

4. Se il sistema autonomo associato al VRF è lo stesso del sistema autonomo del fornitore VPN e l'attributo AS_PATH della route VPN non è vuoto, deve essere anteposto all'attributo AS_PATH della route VRF.

   Fare riferimento alla Figura 3. CE1 e PE1 dispongono di AS 65000 e sono configurati con la funzione iBGP PE-CE. CE2 ha ASN 65001. Ciò significa che c'è eBGP tra PE2 e CE2.

Figura 3

PE2 vede la route come segue:

```
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
```

Questo è il prefisso visto su 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
```

Caso 2. Il numero **Origin AS** contenuto nell'attributo ATTR_SET viene anteposto a AS_PATH da PE2 e segue le regole applicabili a un peer eBGP tra l'attributo AS di origine e quello di destinazione. Gli attributi specifici di iBGP vengono ignorati da PE2 quando crea la route da annunciare a CE2. Pertanto, la preferenza locale è 100 e non 200 (come mostrato nell'attributo ATTR_SET).

Riflesso CE-to-CE VRF-Lite

Fare riferimento alla Figura 4.

Figura 4

La figura 4 mostra un router CE aggiuntivo, CE3, collegato a PE1. CE1 e CE3 sono entrambi

collegati a PE1 sulla stessa istanza VRF: cliente1. Ciò significa che CE1 e CE3 sono router CE multi-VRF (anche noti come VRF-Lite) di PE1. PE1 si pone come hop successivo quando annuncia i prefissi da CE1 a CE3. Se questo comportamento non è desiderato, è possibile configurare l'**hop successivo 10.1.3.6 senza modifiche** su PE1. Per configurare questo comportamento, è necessario rimuovere l'**hop successivo 10.1.3.6** su PE1. In seguito CE3 vede le route da CE1 con CE1 come hop successivo per tali router Prefissi GP. Per eseguire questa operazione, è necessario specificare le route per gli hop successivi BGP nella tabella di routing di CE3. È necessario specificare un protocollo IGP (Dynamic Routing Protocol) o route statiche su CE1, PE1 e CE3 per assicurarsi che i router dispongano di una route per gli altri indirizzi IP dell'hop successivo. Si è verificato tuttavia un problema con questa configurazione.

La configurazione in PE1 è la seguente:

```
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
```

Il prefisso di CE1 è visto correttamente su CE3:

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

Tuttavia, il prefisso CE2 è visibile su CE3 come mostrato di seguito:

```
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
```

L'hop successivo BGP è **192.168.100.2**, l'indirizzo IP di loopback di PE2. PE1 non ha riscritto l'hop successivo BGP su se stesso quando ha annunciato il prefisso 10.100.1.2/32 su CE3. Questo rende il prefisso inutilizzabile su CE3.

Pertanto, nel caso di una combinazione della funzionalità iBGP PE-CE tra MPLS-VPN e iBGP VRF-Lite, è necessario assicurarsi di disporre sempre di un hop successivo sui router PE.

Non è possibile mantenere l'hop successivo quando un router PE è un RR che riflette le route iBGP da un CE a un altro CE attraverso le interfacce VRF localmente sul PE. Quando si esegue iBGP PE-CE su una rete VPN MPLS, è necessario utilizzare **internal-vpn-client** per le sessioni iBGP verso i router CE. Se in un VRF su un router PE sono presenti più CE locali, è necessario mantenere l'**hop successivo** per i peer BGP.

Èpossibile esaminare le route map per impostare l'hop successivo su se stesso per i prefissi ricevuti da altri router PE, ma non per i prefissi riflessi da altri router CE connessi localmente. Tuttavia, non è attualmente supportato l'impostazione dell'hop successivo su se stesso in una route-map in uscita. La configurazione viene mostrata di seguito:

```
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
!
```

Tuttavia, questa operazione non è supportata:

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

Cisco IOS precedente sul router PE

Se PE1 esegue un software Cisco IOS precedente privo della funzionalità iBGP PE-CE, PE1 non si imposta mai come hop successivo per i prefissi iBGP riflessi. Ciò significa che il prefisso BGP riflesso (10.100.1.1/32) da CE1 (10.100.1.1) a CE2 -tramite PE1- avrà CE1 (10.1.1.4) come hop successivo.

```
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
```

Il prefisso di CE2 (10.100.1.2/32) viene visto con PE2 come hop successivo, perché PE1 non esegue l'hop successivo per questo prefisso:

```
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
```

Affinché la funzionalità iBGP PE-CE funzioni correttamente, tutti i router PE per la VPN in cui la funzionalità è abilitata devono disporre del codice per supportare la funzionalità e averla abilitata.

Next-hop-self per eBGP su VRF

Fare riferimento alla Figura 5.

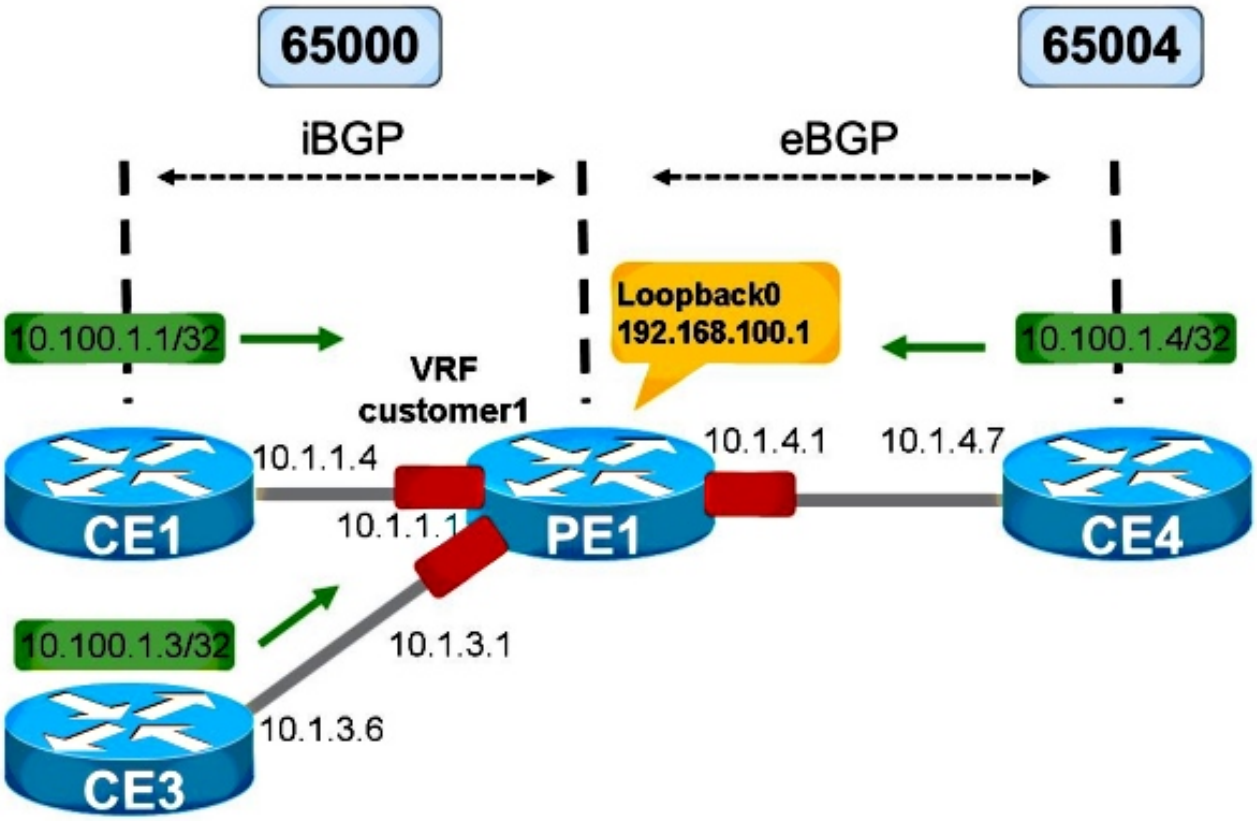


Figura 5

La Figura 5 mostra una configurazione VRF-Lite. La sessione da PE1 a CE4 è eBGP. La sessione da PE1 a CE3 è ancora iBGP.

Per i prefissi eBGP, l'hop successivo è sempre impostato su se stesso quando annuncia i prefissi verso un router adiacente iBGP su VRF. Ciò avviene indipendentemente dal fatto che la sessione verso il router adiacente iBGP attraverso il VRF abbia o meno l'impostazione hop-self successivo.

Nella Figura 5, CE3 vede i prefissi di CE4 con PE1 come hop successivo.

```
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
```

Ciò si verifica con l'hop successivo su PE1 verso CE3 o senza.

Se le interfacce di PE1 verso CE3 e CE4 non sono in un VRF, ma nel contesto globale, il passaggio successivo verso CE3 fa la differenza.

Senza l'hop successivo su PE1 verso CE3, è possibile visualizzare:

```
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)
```

Anche se l'opzione next-hop-self è implicitamente abilitata, l'output non lo indica.

Con l'hop successivo su PE1 verso CE3, è possibile visualizzare:

```
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
```

Mentre, se le interfacce verso CE3 e CE4 sono in un contesto globale, l'hop successivo per i prefissi da CE4 è CE4 stesso quando l'hop successivo non è configurato:

```
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
```

Per l'hop successivo su PE1 verso 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
```

Questa operazione è stata eseguita in base alla RFC 4364.

Se non si desidera impostare next-hop-self per i prefissi eBGP verso una sessione iBGP su un'interfaccia VRF, è necessario configurare **next-hop-unchanged**. Il supporto per questa condizione si è verificato solo con l'ID bug Cisco CSCuj11720.

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

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

Ora, CE3 vede CE4 come l'hop successivo per i prefissi pubblicizzati da CE4:

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

Se si cerca di configurare la parola chiave **next-hop-unchanged** per la sessione iBGP verso CE3 sul codice Cisco IOS prima dell'ID bug Cisco CSCuj1720, viene visualizzato questo errore:

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

Dopo l'ID bug Cisco CSCuj11720, la parola chiave **next-hop-unchanged** è valida per i vicini eBGP e iBGP VRF-Lite multi-hop.