

Border Gateway Protocol Optimal Route Reflection (BGP ORR)

Sommario

[Introduzione](#)

[Premesse](#)

[Esempio di rete](#)

[Teoria](#)

[Implementazione di IOS-XR](#)

[Configurazione](#)

[Esempio di configurazione](#)

[Progettazione del traffico MPLS sul router radice](#)

[Risoluzione dei problemi](#)

[Conclusioni](#)

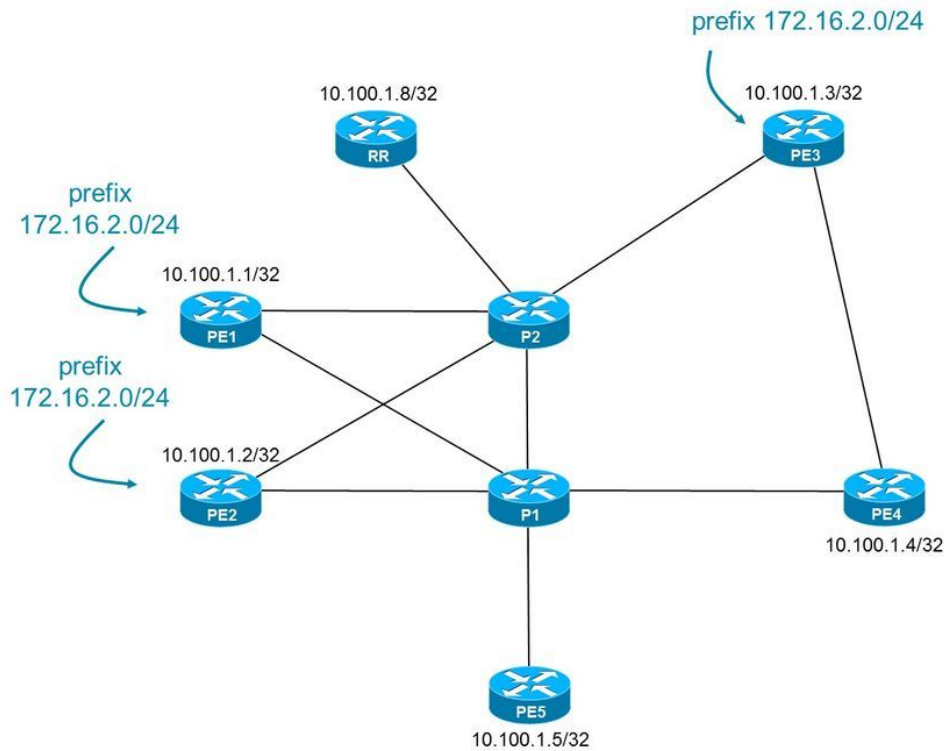
Introduzione

In questo documento viene descritto come influenzare il routing quando vi sono uno o più Route Reflector (RR) nella rete per evitare una mesh completa tra i router iBGP.

Premesse

Il passo 8 dell'[algoritmo di selezione del miglior percorso BGP](#) consiste nel preferire il percorso con la metrica IGP più bassa all'hop successivo BGP. Pertanto, se tutti i passaggi precedenti al passaggio 8 sono uguali, il passaggio 8 può essere il fattore decisivo per stabilire quale sia il percorso migliore per il record di risorse. Il costo IGP tra il router RR e il router iBGP pubblicitario viene quindi determinato dal posizionamento del router RR. Per impostazione predefinita, il record di risorse annuncia solo il miglior percorso ai propri client. A seconda della posizione in cui si trova il record di risorse, il costo IGP per il router pubblicitario può essere minore o maggiore. Se tutti i costi IGP dei percorsi sono gli stessi, è probabile che si verifichi un errore del router pubblicitario con l'ID del router BGP più basso.

Esempio di rete



I router PE1, PE2 e PE3 annunciano il prefisso 172.16.2.0/24. Se tutti i costi IGP dei collegamenti sono uguali, il record di risorse visualizzerà il percorso da PE1, PE2 e PE3 con un costo IGP pari a 2. Alla fine, il record di risorse sceglie il percorso da PE1 come migliore perché ha l'ID del router BGP più basso. Questo è il passo 11 nell'algoritmo di selezione del miglior percorso BGP. Di conseguenza, tutti i router PE, incluso PE4, sceglieranno PE1 come router PE in uscita per il prefisso 172.16.2.0/24. Dal punto di vista di PE4, il percorso IGP più breve per qualsiasi router PE in uscita è il percorso verso PE3, con un costo IGP di 1. Il costo IGP per qualsiasi altro router PE è 2. Per molte reti, è importante trasportare il traffico attraverso la rete di transito nel modo più breve possibile. Questo processo è noto come instradamento delle patate bollenti.

Esiste un altro motivo possibile per cui RR ha scelto il percorso migliore da PE1. Se nell'immagine il costo IGP (Interior Gateway Protocol) del collegamento P2-PE3 è 10 e tutti gli altri collegamenti hanno ancora un costo IGP pari a 1, RR non sceglierebbe il percorso da PE3 come il migliore, anche se PE3 aveva l'ID router BGP più basso.

Se l'amministratore di questa rete desidera utilizzare il routing tra patate a caldo, deve essere installato un meccanismo in modo che, quando nella rete sono presenti RR, il router in entrata possa continuare a conoscere il percorso al router in uscita più vicino nella rete iBGP. La funzione BGP Add Path può raggiungere questo obiettivo. Tuttavia, con questa funzionalità, i router RR e i router dei bordi devono avere codice più recente che la comprenda. Con la funzionalità BGP Optimal Route Reflection, questo non è un requisito. Questa funzione consente all'RR di inviare il miglior percorso al router di confine BGP in entrata, in base a ciò che l'RR ritiene essere il percorso migliore dal punto di vista del router BGP in entrata.

Un'altra soluzione che consentirebbe il routing delle patate bollenti durante l'installazione delle RR è il posizionamento in linea delle RR. Questi RR non sono RR dedicati, che eseguono solo BGP e IGP. Questi RR in linea si trovano nel percorso di inoltro e vengono collocati nella rete in modo che dispongano del proprio set di client RR, in modo che possano riflettere il miglior percorso verso ogni client RR, che è anche il miglior percorso dal punto di vista di quel client RR.

Come mostrato in questa immagine, i RR vengono posizionati nella rete in modo da poter servire un piccolo gruppo di client RR vicini. A causa della progettazione della rete, i client RR ricevono i

La soluzione permette di calcolare il miglior percorso basandosi sull'algoritmo di selezione del miglior percorso BGP, in modo che il record RR scelga il percorso migliore dal punto di vista del router di confine in entrata al quale invia il percorso. Ciò significa che verrà scelto il percorso migliore in base al costo IGP più breve per l'hop successivo BGP. La soluzione consente inoltre di scegliere il percorso migliore in base ad alcuni criteri configurati. I router di confine in entrata potrebbero scegliere i percorsi migliori in base ad alcune policy configurate, e non al costo IGP più basso. La soluzione consente all'RR di implementare la riflessione ottimale del percorso sia sul costo IGP (posizione sulla rete) o su alcuni criteri configurati, o entrambi. Se vengono distribuiti entrambi, il criterio viene applicato per primo, quindi la riflessione ottimale della route basata su IGP si verificherà sui percorsi rimanenti.

Implementazione di IOS-XR

L'implementazione IOS-XR consente di utilizzare fino a tre nodi radice per il calcolo rSPF. Se in un gruppo di aggiornamento sono presenti più client RR, non è necessario eseguire un calcolo rSPF per ogni client RR se tali client RR avranno la stessa policy e/o gli stessi costi IGP per i diversi router di confine BGP in uscita. Quest'ultimo significa in genere che i client di RR sono ubicati nello stesso punto (probabilmente nello stesso POP). In questo caso, non è necessario configurare ogni client RR come root. L'implementazione di IOS-XR consente di configurare tre, la principale, la secondaria e la terza radice, per set di client RR, a scopo di ridondanza. Affinché la funzionalità BGP OR venga applicata a qualsiasi client RR, è necessario che tale client sia configurato per far parte di un gruppo di criteri ORR.

La funzionalità BGP OR è abilitata per ciascuna famiglia di indirizzi.

È necessario un protocollo dello stato del collegamento. Può essere OSPF o IS-IS.

IOS XR implementa la funzionalità BGP OR solo in base al costo IGP per l'hop successivo BGP e non in base ad alcune policy configurate.

I peer BGP con lo stesso criterio in uscita vengono inseriti nello stesso gruppo di aggiornamento. Questo è generalmente il caso di iBGP su RR. Quando la funzionalità BGP OR è abilitata, i peer di gruppi ORR diversi si troveranno in gruppi di aggiornamento diversi. Ciò è logico, in quanto gli aggiornamenti inviati dal record di risorse ai client di risorse di risorse di risorse in gruppi BGP OR diversi saranno diversi, in quanto il miglior percorso BGP è diverso.

Il risultato dei calcoli rSPF viene memorizzato in un database.

ORSPF è il nuovo componente di IOS-XR necessario per la funzione BGP OR. ORRSPF si occupa di:

1. Raccolta delle informazioni sullo stato del collegamento e gestione del database sullo stato del collegamento
2. Esecuzione di rSPF e gestione degli SPT per gruppo di criteri
3. Scaricamento dei prefissi dall'SPT alla RIB con le metriche

Il database ottiene le informazioni sullo stato del collegamento direttamente dall'IGP o dal BGP-LS.

I calcoli rSPF generano una topologia che mostra il percorso più breve tra il client RR e qualsiasi altro router dell'area o del livello.

Le route che si trovano su ogni router della topologia vengono archiviate in una tabella RIB

speciale per i criteri di gruppo ORR e per AFI/SAFI. Questa tabella viene creata da RSI. Nella tabella vengono inserite le route calcolate dagli rSPF con la radice primaria come radice. Se la radice primaria diventa non disponibile, la radice secondaria è la radice e popola le route nella tabella RIB OR. Lo stesso vale per la radice terziaria.

Configurazione

Configurazione minima richiesta:

1. La funzione ORR deve essere abilitata per la famiglia di indirizzi BGP, per gruppi specifici di vicini BGP
2. Per ogni gruppo di vicini BGP, è necessario configurare almeno una radice. Facoltativamente, è possibile configurare una radice secondaria e terziaria.
3. La redistribuzione delle route ORR dall'IGP al BGP deve essere abilitata.

Esempio di configurazione

Come mostrato nella prima immagine, il router RR è un router IOS-XR con la funzionalità BGP ORR.

Tutti gli altri router eseguono IOS. Questi router non dispongono della funzionalità BGP ORR.

PE1, PE2 e PE3 annunciano il prefisso 172.16.2.0/24 in AFI/SAFI 1/1 (unicast IPv4). Il record di risorse è simile a PE1 e PE2 che a PE3. Il costo IGP di tutti i collegamenti è 1. Il percorso migliore per questo prefisso è quello con R1 come hop successivo a causa dell'ID del router BGP più basso.

```
RP/0/0/CPU0:RR#show bgp ipv4 unicast 172.16.2.0/24 bestpath-compare
BGP routing table entry for 172.16.2.0/24
Versions:
  Process          bRIB/RIB  SendTblVer
  Speaker          34        34
Last Modified: Mar  7 20:29:48.156 for 11:36:44
Paths: (3 available, best #1)
  Advertised to update-groups (with more than one peer):
    0.3
  Path #1: Received by speaker 0
  Advertised to update-groups (with more than one peer):
    0.3
  Local, (Received from a RR-client)
    10.100.1.1 (metric 3) from 10.100.1.1 (10.100.1.1)
      Origin IGP, metric 0, localpref 100, valid, internal, best, group-best
      Received Path ID 0, Local Path ID 1, version 34
      best of local AS, Overall best
  Path #2: Received by speaker 0
  Not advertised to any peer
  Local, (Received from a RR-client)
    10.100.1.2 (metric 3) from 10.100.1.2 (10.100.1.2)
      Origin IGP, metric 0, localpref 100, valid, internal, add-path
      Received Path ID 0, Local Path ID 6, version 33
      Higher router ID than best path (path #1)
  Path #3: Received by speaker 0
  ORR bestpath for update-groups (with more than one peer):
    0.1
  Local, (Received from a RR-client)
```

```
10.100.1.3 (metric 5) from 10.100.1.3 (10.100.1.3)
Origin IGP, metric 0, localpref 100, valid, internal, add-path
Received Path ID 0, Local Path ID 7, version 34
Higher IGP metric than best path (path #1)
```

PE4 riceverà il percorso con PE1 come hop successivo. Per questo motivo, non è previsto il routing delle patate bollenti per PE4.

Se si desidera che il routing delle patate a caldo venga eseguito su PE4, per i prefissi pubblicizzati da PE1, PE2 e PE3 (ad esempio il prefisso 172.16.2.0/24), PE1 deve avere PE3 come punto di uscita. Ciò significa che il percorso in PE4 deve essere quello con PE3 come hop successivo. Con questa configurazione ORR, è possibile fare in modo che l'RR invii il percorso con l'hop successivo da PE3 a PE4.

```
router ospf 1
distribute bgp-1s
area 0
interface Loopback0
!
interface GigabitEthernet0/0/0/0
network point-to-point
!
!
!

router bgp 1
address-family ipv4 unicast
optimal-route-reflection ipv4-orr-group 10.100.1.4
!
address-family vpnv4 unicast
!
neighbor 10.100.1.1
remote-as 1
update-source Loopback0
address-family ipv4 unicast
route-reflector-client
!
!
neighbor 10.100.1.2
remote-as 1
update-source Loopback0
address-family ipv4 unicast
route-reflector-client
!
!
neighbor 10.100.1.3
remote-as 1
update-source Loopback0
address-family ipv4 unicast
route-reflector-client
!
!
neighbor 10.100.1.4
remote-as 1
update-source Loopback0
address-family ipv4 unicast
optimal-route-reflection ipv4-orr-group
route-reflector-client
!
!
neighbor 10.100.1.5
remote-as 1
```

```
update-source Loopback0
address-family ipv4 unicast
  route-reflector-client
!
!
```

Se IGP è IS-IS:

```
router isis 1
net 49.0001.0000.0000.0008.00
  distribute bgp-ls
address-family ipv4 unicast
metric-style wide
!
interface Loopback0
address-family ipv4 unicast
!
!
interface GigabitEthernet0/0/0/0
address-family ipv4 unicast
!
!
!
```

Nota: Non è necessario configurare lo stato del collegamento della famiglia di indirizzi, a livello globale o nei router adiacenti BGP.

Progettazione del traffico MPLS sul router radice

Per eseguire l'rSPF, l'RR deve trovare l'indirizzo radice configurato nel database IGP. Nell'ISIS, l'ID del router è presente nel database ISIS. Per OSPF, non è presente alcun ID router negli LSA OSPF. La soluzione consiste nel fare in modo che i router radice dichiarino MPLS (Multi Protocol Label Switching), l'ID del router corrispondente all'indirizzo radice configurato sul record di risorse.

Per il protocollo OSPF, i router radice necessitano di una configurazione aggiuntiva per il funzionamento di BGP OR. Per annunciare questo ID router MPLS TE è necessaria una configurazione MPLS TE minima su qualsiasi router radice. L'esatto insieme minimo di comandi dipende dal sistema operativo del router principale. La configurazione MPLS TE sul router radice deve avere la configurazione minima per MPLS TE abilitata in modo che OSPF annunci l'ID del router MPLS TE in una LSA area opaca (tipo 10).

Se il router dispone di un LSA a area opaca con MPLS TE router-ID corrispondente all'indirizzo del router radice configurato, il router rSPF può essere eseguito e il protocollo BGP sul router può annunciare il percorso ottimale.

La configurazione minima richiesta per OSPF sul router radice se è un router IOS è:

```
!
interface GigabitEthernet0/2
 ip address 10.1.34.4 255.255.255.0
 ip ospf network point-to-point
mpls traffic-eng tunnels
!

router ospf 1
```

```

mpls traffic-eng router-id Loopback0
mpls traffic-eng area 0
router-id 10.200.1.155
network 10.0.0.0 0.255.255.255 area 0
!

```

Si noti che:

- MPLS TE è abilitato nell'area OSPF specifica
- MPLS TE router-ID è configurato in modo corrispondente all'indirizzo radice configurato nell'RR
- MPLS TE è configurato su almeno un'interfaccia
- non è necessario configurare RSVP-TE
- non è necessario configurare MPLS TE su alcun altro router dell'area

La configurazione minima richiesta per OSPF sul router radice se è un router IOS-XR è:

```

!
router ospf 1
router-id 5.6.7.8
area 0
mpls traffic-eng
interface Loopback0
!
interface GigabitEthernet0/0/0/0
network point-to-point
!
!
mpls traffic-eng router-id 10.100.1.11
!
mpls traffic-eng
!

```

Se la configurazione precedente è presente sul router principale, RR deve avere MPLS TE router-ID nel database OSPF.

```
RP/0/0/CPU0:RR#show ospf 1 database
```

```
OSPF Router with ID (10.100.1.99) (Process ID 1)
```

```
Router Link States (Area 0)
```

Link ID	ADV Router	Age	Seq#	Checksum	Link count
10.1.12.1	10.1.12.1	1297	0x8000002b	0x006145	3
10.100.1.2	10.100.1.2	646	0x80000025	0x00fb6f	7
10.100.1.3	10.100.1.3	1693	0x80000031	0x003ba9	5
10.100.1.99	10.100.1.99	623	0x8000001e	0x00ade1	3
10.200.1.155	10.200.1.155	28	0x80000002	0x009b2e	5

```
Type-10 Opaque Link Area Link States (Area 0)
```

Link ID	ADV Router	Age	Seq#	Checksum	Opaque ID
1.0.0.0	10.200.1.155	34	0x80000001	0x00a1ad	0
1.0.0.3	10.200.1.155	34	0x80000001	0x0057ff	3

```
RP/0/0/CPU0:RR#show ospf 1 database opaque-area adv-router 10.200.1.155
```


OSPF Router with ID (10.100.1.99) (Process ID 1)

Type-10 Opaque Link Area Link States (Area 0)

LS age: 184
Options: (No TOS-capability, DC)
LS Type: Opaque Area Link
Link State ID: 1.0.0.0
Opaque Type: 1
Opaque ID: 0
Advertising Router: 10.200.1.155
LS Seq Number: 80000001
Checksum: 0xalad
Length: 28

MPLS TE router ID : 10.100.1.4

Number of Links : 0

LS age: 184
Options: (No TOS-capability, DC)
LS Type: Opaque Area Link
Link State ID: 1.0.0.3
Opaque Type: 1
Opaque ID: 3
Advertising Router: 10.200.1.155
LS Seq Number: 80000001
Checksum: 0x57ff
Length: 132

Link connected to Point-to-Point network

Link ID : 10.100.1.3 (all bandwidths in bytes/sec)
Interface Address : 10.1.34.4
Neighbor Address : 10.1.34.3
Admin Metric : 1
Maximum bandwidth : 125000000
Maximum reservable bandwidth global: 0
Number of Priority : 8

Priority 0 :	0	Priority 1 :	0
Priority 2 :	0	Priority 3 :	0
Priority 4 :	0	Priority 5 :	0
Priority 6 :	0	Priority 7 :	0

Affinity Bit : 0
IGP Metric : 1

Number of Links : 1

Si noti che l'ID router MPLS TE (10.100.1.4) e l'ID router OSPF sono diversi.

PE4 ha PE3 come hop successivo per il prefisso (con la metrica IGP corretta per l'hop successivo):

```
PE4#show bgp ipv4 unicast 172.16.2.0
BGP routing table entry for 172.16.2.0/24, version 37
Paths: (1 available, best #1, table default)
Not advertised to any peer
Refresh Epoch 1
Local
 10.100.1.3 (metric 2) from 10.100.1.8 (10.100.1.8)
  Origin IGP, metric 0, localpref 100, valid, internal, best
  Originator: 10.100.1.3, Cluster list: 10.100.1.8
  rx pathid: 0, tx pathid: 0x0
```

PE5 ha ancora PE1 come hop successivo per il prefisso (con la metrica IGP corretta per l'hop successivo):

```
PE5#show bgp ipv4 unicast 172.16.2.0/24
BGP routing table entry for 172.16.2.0/24, version 13
Paths: (1 available, best #1, table default)
  Not advertised to any peer
  Refresh Epoch 1
  Local
    10.100.1.1 (metric 3) from 10.100.1.8 (10.100.1.8)
      Origin IGP, metric 0, localpref 100, valid, internal, best
      Originator: 10.100.1.1, Cluster list: 10.100.1.8
      rx pathid: 0, tx pathid: 0x0
```

Risoluzione dei problemi

Verificare il prefisso sul record di risorse:

```
RP/0/0/CPU0:RR#show bgp ipv4 unicast 172.16.2.0
BGP routing table entry for 172.16.2.0/24
Versions:
  Process          bRIB/RIB  SendTblVer
  Speaker          19        19
Last Modified: Mar  7 16:41:20.156 for 03:07:40
Paths: (3 available, best #1)
  Advertised to update-groups (with more than one peer):
    0.3
  Path #1: Received by speaker 0
  Advertised to update-groups (with more than one peer):
    0.3
  Local, (Received from a RR-client)
    10.100.1.1 (metric 3) from 10.100.1.1 (10.100.1.1)
      Origin IGP, metric 0, localpref 100, valid, internal, best, group-best
      Received Path ID 0, Local Path ID 1, version 14
  Path #2: Received by speaker 0
  Not advertised to any peer
  Local, (Received from a RR-client)
    10.100.1.2 (metric 3) from 10.100.1.2 (10.100.1.2)
      Origin IGP, metric 0, localpref 100, valid, internal, add-path
      Received Path ID 0, Local Path ID 4, version 14
  Path #3: Received by speaker 0
  ORR bestpath for update-groups (with more than one peer):
    0.1
  Local, (Received from a RR-client)
    10.100.1.3 (metric 5) from 10.100.1.3 (10.100.1.3)
      Origin IGP, metric 0, localpref 100, valid, internal, add-path
      Received Path ID 0, Local Path ID 5, version 19
```

Si noti che add-path è stato aggiunto agli altri percorsi non ottimali, in modo che possano essere annunciati, oltre al percorso migliore. La funzione di aggiunta del percorso non viene utilizzata tra RR e i relativi client: i percorsi non vengono annunciati con un identificatore di percorso.

Verificare che le route siano (ancora) annunciate agli specifici router BGP adiacenti.

Per il sistema PE4 adiacente, l'hop successivo è PE3 per il prefisso 172.16.2.0/24:

```
RP/0/0/CPU0:RR#show bgp ipv4 unicast neighbors 10.100.1.4 advertised-routes
```

Network	Next Hop	From	AS Path
172.16.1.0/24	10.100.1.5	10.100.1.5	i
172.16.2.0/24	10.100.1.3	10.100.1.3	i

Processed 2 prefixes, 2 paths

Per il router adiacente PE5, l'hop successivo è PE1 per il prefisso 172.16.2.0/24:

```
RP/0/0/CPU0:RR#show bgp ipv4 unicast neighbors 10.100.1.5 advertised-routes
```

Network	Next Hop	From	AS Path
172.16.1.0/24	10.100.1.8	10.100.1.5	i
172.16.2.0/24	10.100.1.1	10.100.1.1	i

Il router adiacente 10.100.1.4 si trova nel proprio gruppo di aggiornamento a causa dei criteri ORR in vigore:

```
RP/0/0/CPU0:RR#show bgp ipv4 unicast update-group
```

Update group for IPv4 Unicast, index 0.1:

Attributes:

- Neighbor sessions are IPv4
- Internal
- Common admin
- First neighbor AS: 1
- Send communities
- Send GSHUT community if originated
- Send extended communities
- Route Reflector Client
- ORR root (configured): ipv4-orr-group; Index: 0
- 4-byte AS capable
- Non-labeled address-family capable
- Send AIGP
- Send multicast attributes
- Minimum advertisement interval: 0 secs

Update group desynchronized: 0

Sub-groups merged: 0

Number of refresh subgroups: 0

Messages formatted: 8, replicated: 8

All neighbors are assigned to sub-group(s)

Neighbors in sub-group: 0.1, Filter-Groups num:1

Neighbors in filter-group: 0.3(RT num: 0)

10.100.1.4

Update group for IPv4 Unicast, index 0.3:

Attributes:

- Neighbor sessions are IPv4
- Internal
- Common admin
- First neighbor AS: 1
- Send communities
- Send GSHUT community if originated
- Send extended communities
- Route Reflector Client
- 4-byte AS capable
- Non-labeled address-family capable
- Send AIGP
- Send multicast attributes
- Minimum advertisement interval: 0 secs

Update group desynchronized: 0

Sub-groups merged: 1

Number of refresh subgroups: 0

```
Messages formatted: 12, replicated: 42
All neighbors are assigned to sub-group(s)
  Neighbors in sub-group: 0.3, Filter-Groups num:1
    Neighbors in filter-group: 0.1(RT num: 0)
      10.100.1.1          10.100.1.2          10.100.1.3
10.100.1.5
```

Il comando `show orrspf database` mostra il gruppo ORR e le relative radici,

```
RP/0/0/CPU0:RR#show orrspf database
```

```
ORR policy: ipv4-orr-group, IPv4, RIB tableid: 0xe0000012
Configured root: primary: 10.100.1.4, secondary: NULL, tertiary: NULL
Actual Root: 10.100.1.4
```

```
Number of mapping entries: 1
```

Lo stesso comando con la parola chiave `detail` fornisce il costo della radice dell'rSPF per ciascun router/prefisso nella stessa area OSPF:

```
RP/0/0/CPU0:RR#show orrspf database detail
```

```
ORR policy: ipv4-orr-group, IPv4, RIB tableid: 0xe0000012
Configured root: primary: 10.100.1.4, secondary: NULL, tertiary: NULL
Actual Root: 10.100.1.4
```

Prefix	Cost
10.100.1.6	2
10.100.1.1	3
10.100.1.2	3
10.100.1.3	2
10.100.1.4	0
10.100.1.5	3
10.100.1.7	3
10.100.1.8	4

```
Number of mapping entries: 9
```

L'ID tabella è stato assegnato da RSI per il gruppo ORR e per AFI/SAFI:

```
RP/0/0/CPU0:RR#show rsi table-id 0xe0000012
```

```
TBL_NAME=ipv4-orr-group, AFI=IPv4, SAFI=Ucast TBL_ID=0xe0000012 in VRF=default/0x60000000 in
VR=default/0x20000000
Refcnt=1
VRF Index=4 TCM Index=1
Flags=0x0 LST Flags=(0x0) NULL
```

```
RP/0/0/CPU0:RR#show rib tables
```

```
Codes: N - Prefix Limit Notified, F - Forward Referenced
       D - Table Deleted, C - Table Reached Convergence
```

VRF/Table	SAFI	Table ID	PrfxLmt	PrfxCnt	TblVersion	N	F	D	C
default/default	uni	0xe0000000	5000000	22	128	N	N	N	Y
**nVSatellite/default	uni	0xe0000010	5000000	2	4	N	N	N	Y
default/ipv4-orr-grou	uni	0xe0000012	5000000	9	27	N	N	N	Y
default/default	multi	0xe0100000	5000000	0	0	N	N	N	Y

Il costo della radice (R4/10.100.1.4) dell'rSPF per il router è lo stesso del costo rilevato con `show ip route ospf` su PE4:

PE4#show ip route ospf

Codes: L - local, C - connected, S - static, R - RIP, M - mobile, B - BGP
D - EIGRP, EX - EIGRP external, O - OSPF, IA - OSPF inter area
N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
E1 - OSPF external type 1, E2 - OSPF external type 2
i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
ia - IS-IS inter area, * - candidate default, U - per-user static route
o - ODR, P - periodic downloaded static route, H - NHRP, l - LISP
a - application route
+ - replicated route, % - next hop override, p - overrides from PfR

Gateway of last resort is not set

10.0.0.0/8 is variably subnetted, 20 subnets, 2 masks
O 10.100.1.1/32 [110/3] via 10.1.7.6, 4d05h, GigabitEthernet0/1
O 10.100.1.2/32 [110/3] via 10.1.7.6, 4d05h, GigabitEthernet0/1
O 10.100.1.3/32 [110/2] via 10.1.8.3, 4d06h, GigabitEthernet0/2
O 10.100.1.5/32 [110/3] via 10.1.7.6, 4d05h, GigabitEthernet0/1
O 10.100.1.6/32 [110/2] via 10.1.7.6, 4d05h, GigabitEthernet0/1
O 10.100.1.7/32 [110/3] via 10.1.8.3, 4d06h, GigabitEthernet0/2
[110/3] via 10.1.7.6, 4d05h, GigabitEthernet0/1
O 10.100.1.8/32 [110/4] via 10.1.8.3, 4d05h, GigabitEthernet0/2
[110/4] via 10.1.7.6, 4d05h, GigabitEthernet0/1

RIB per il gruppo BGP OR:

RP/0/0/CPU0:RR#show route afi-all safi-all topology ipv4-orr-group

IPv4 Unicast Topology ipv4-orr-group:

Codes: C - connected, S - static, R - RIP, B - BGP, (>) - Diversion path
D - EIGRP, EX - EIGRP external, O - OSPF, IA - OSPF inter area
N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
E1 - OSPF external type 1, E2 - OSPF external type 2, E - EGP
i - ISIS, L1 - IS-IS level-1, L2 - IS-IS level-2
ia - IS-IS inter area, su - IS-IS summary null, * - candidate default
U - per-user static route, o - ODR, L - local, G - DAGR, l - LISP
A - access/subscriber, a - Application route
M - mobile route, r - RPL, (!) - FRR Backup path

Gateway of last resort is not set

o 10.100.1.1/32 [255/3] via 0.0.0.0, 14:14:52, Unknown
o 10.100.1.2/32 [255/3] via 0.0.0.0, 14:14:52, Unknown
o 10.100.1.3/32 [255/2] via 0.0.0.0, 00:04:53, Unknown
o 10.100.1.4/32 [255/0] via 0.0.0.0, 14:14:52, Unknown
o 10.100.1.5/32 [255/3] via 0.0.0.0, 14:14:52, Unknown
o 10.100.1.6/32 [255/2] via 0.0.0.0, 14:14:52, Unknown
o 10.100.1.7/32 [255/3] via 0.0.0.0, 14:14:52, Unknown
o 10.100.1.8/32 [255/4] via 0.0.0.0, 14:14:52, Unknown

RP/0/0/CPU0:RR#show rsi table name ipv4-orr-group

VR=default:

TBL_NAME=ipv4-orr-group, AFI=IPv4, SAFI=Ucast TBL_ID=0xe0000012 in VRF=default/0x60000000 in
VR=default/0x20000000
Refcnt=1
VRF Index=4 TCM Index=1
Flags=0x0 LST Flags=(0x0) NULL

Il comando **show bgp neighbors** mostra se il peer è una radice OR:

```
RP/0/0/CPU0:RR#show bgp neighbor 10.100.1.4
```

```
BGP neighbor is 10.100.1.4
Remote AS 1, local AS 1, internal link
Remote router ID 10.100.1.4
Cluster ID 10.100.1.8
BGP state = Established, up for 01:17:41
NSR State: None
Last read 00:00:52, Last read before reset 01:18:30
Hold time is 180, keepalive interval is 60 seconds
Configured hold time: 180, keepalive: 60, min acceptable hold time: 3
Last write 00:00:34, attempted 19, written 19
Second last write 00:01:34, attempted 19, written 19
Last write before reset 01:17:43, attempted 19, written 19
Second last write before reset 01:18:43, attempted 19, written 19
Last write pulse rcvd Mar  8 10:20:13.779 last full not set pulse count 12091
Last write pulse rcvd before reset 01:17:42
Socket not armed for io, armed for read, armed for write
Last write thread event before reset 01:17:42, second last 01:17:42
Last KA expiry before reset 01:17:43, second last 01:18:43
Last KA error before reset 00:00:00, KA not sent 00:00:00
Last KA start before reset 01:17:43, second last 01:18:43
Precedence: internet
Non-stop routing is enabled
Multi-protocol capability received
Neighbor capabilities:
  Route refresh: advertised (old + new) and received (old + new)
  4-byte AS: advertised and received
  Address family IPv4 Unicast: advertised and received
Received 6322 messages, 0 notifications, 0 in queue
Sent 5782 messages, 4 notifications, 0 in queue
Minimum time between advertisement runs is 0 secs
Inbound message logging enabled, 3 messages buffered
Outbound message logging enabled, 3 messages buffered

For Address Family: IPv4 Unicast
BGP neighbor version 41
Update group: 0.1 Filter-group: 0.1 No Refresh request being processed
Route-Reflector Client
ORR root (configured): ipv4-orr-group; Index: 0
  Route refresh request: received 0, sent 0
  0 accepted prefixes, 0 are bestpaths
  Cumulative no. of prefixes denied: 0.
  Prefix advertised 2, suppressed 0, withdrawn 0
  Maximum prefixes allowed 1048576
  Threshold for warning message 75%, restart interval 0 min
  AIGP is enabled
  An EoR was received during read-only mode
  Last ack version 41, Last synced ack version 0
  Outstanding version objects: current 0, max 2
  Additional-paths operation: None
  Send Multicast Attributes
  Advertise VPNv4 routes enabled with option
  Advertise VPNv6 routes is enabled with Local with stitching-RT option

Connections established 6; dropped 5
Local host: 10.100.1.8, Local port: 25176, IF Handle: 0x00000000
Foreign host: 10.100.1.4, Foreign port: 179
Last reset 01:17:42, due to User clear requested (CEASE notification sent - administrative
reset)
```

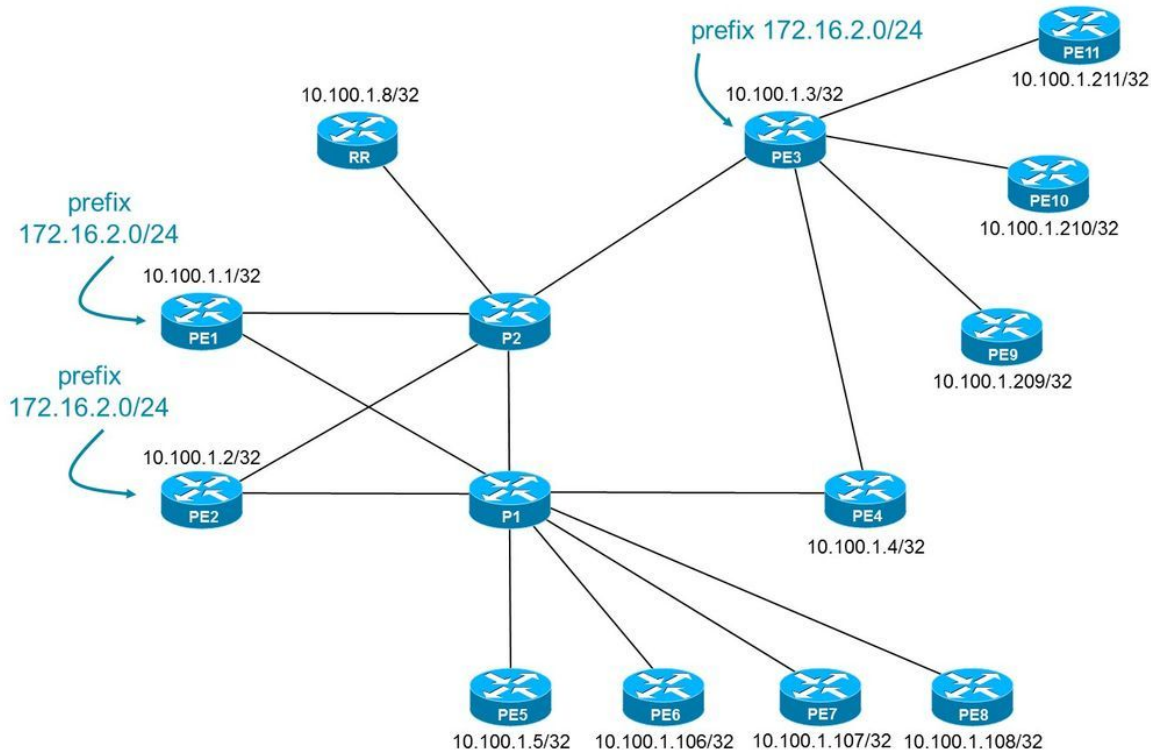
Time since last notification sent to neighbor: 01:17:42

Error Code: administrative reset

Notification data sent:

None

Come mostrato nell'immagine, sono stati configurati più set di client di ripristino



A PE3 è connesso un gruppo di client di ripristino e a P1 un altro gruppo. Ogni client di ripristino di ogni gruppo si trova alla stessa distanza da qualsiasi router di confine BGP in uscita.

```
router bgp 1
  address-family ipv4 unicast
    optimal-route-reflection ipv4-orr-group-1 10.100.1.4 10.100.1.209 10.100.1.210
    optimal-route-reflection ipv4-orr-group-2 10.100.1.5 10.100.1.106 10.100.1.107
  !
...
neighbor 10.100.1.4
  remote-as 1
  update-source Loopback0
  address-family ipv4 unicast
    optimal-route-reflection ipv4-orr-group-1
    route-reflector-client
  !
  !
neighbor 10.100.1.5
  remote-as 1
  update-source Loopback0
  address-family ipv4 unicast
    optimal-route-reflection ipv4-orr-group-2
    route-reflector-client
  !
  !
neighbor 10.100.1.106
  remote-as 1
  update-source Loopback0
  address-family ipv4 unicast
    optimal-route-reflection ipv4-orr-group-2
```

```

    route-reflector-client
!
!
neighbor 10.100.1.107
remote-as 1
update-source Loopback0
address-family ipv4 unicast
    optimal-route-reflection ipv4-orr-group-2
    route-reflector-client
!
!
neighbor 10.100.1.108
remote-as 1
update-source Loopback0
address-family ipv4 unicast
    optimal-route-reflection ipv4-orr-group-2
    route-reflector-client
!
!
neighbor 10.100.1.209
remote-as 1
update-source Loopback0
address-family ipv4 unicast
    optimal-route-reflection ipv4-orr-group-1
    route-reflector-client
!
!
neighbor 10.100.1.210
remote-as 1
update-source Loopback0
address-family ipv4 unicast
    optimal-route-reflection ipv4-orr-group-1    route-reflector-client
!
!
neighbor 10.100.1.211
remote-as 1
update-source Loopback0
address-family ipv4 unicast
    optimal-route-reflection ipv4-orr-group-1
    route-reflector-client
!
!
!

```

Database orrspf per entrambi i gruppi:

RP/0/0/CPU0:RR#**show orrspf database detail**

ORR policy: ipv4-orr-group-1, IPv4, RIB tableid: 0xe0000012
 Configured root: primary: 10.100.1.4, secondary: 10.100.1.209, tertiary: 10.100.1.210
 Actual Root: 10.100.1.4

Prefix	Cost
10.100.1.1	3
10.100.1.2	3
10.100.1.3	2
10.100.1.4	0
10.100.1.5	3
10.100.1.6	2
10.100.1.7	3
10.100.1.8	4
10.100.1.106	3
10.100.1.107	3


```
10.100.1.108          3
10.100.1.209          3
10.100.1.210          3
10.100.1.211          3
ORR policy: ipv4-orr-group-2, IPv4, RIB tableid: 0xe0000013
Configured root: primary: 10.100.1.5, secondary: 10.100.1.106, tertiary: 10.100.1.107
Actual Root: 10.100.1.5
```

Prefix	Cost
10.100.1.1	3
10.100.1.2	3
10.100.1.3	4
10.100.1.4	3
10.100.1.5	0
10.100.1.6	2
10.100.1.7	3
10.100.1.8	4
10.100.1.106	3
10.100.1.107	3
10.100.1.108	3
10.100.1.209	5
10.100.1.210	5
10.100.1.211	5

Number of mapping entries: 30

Se per un gruppo la radice primaria è inattiva o non raggiungibile, la radice secondaria sarà la radice effettivamente utilizzata. Nell'esempio, la radice primaria del gruppo ipv4-orr-group-1 non è raggiungibile. La radice secondaria diventa la radice effettiva:

```
RP/0/0/CPU0:RR#show orrspf database ipv4-orr-group-1
```

```
ORR policy: ipv4-orr-group-1, IPv4, RIB tableid: 0xe0000012
Configured root: primary: 10.100.1.4, secondary: 10.100.1.209, tertiary: 10.100.1.210
Actual Root: 10.100.1.209
```

Prefix	Cost
10.100.1.1	4
10.100.1.2	5
10.100.1.3	2
10.100.1.5	5
10.100.1.6	4
10.100.1.7	3
10.100.1.8	4
10.100.1.106	5
10.100.1.107	5
10.100.1.108	5
10.100.1.209	0
10.100.1.210	3
10.100.1.211	3

Number of mapping entries: 14

Conclusioni

La funzionalità BGP Optimal Route Reflection (ORR) consente il routing delle patate a caldo in una rete iBGP quando sono presenti riflettori di percorso, senza la necessità di nuovi software di sistema operativo sui router di confine. Il prerequisito è che l'IGP sia un protocollo di routing allo stato del collegamento.