

邊界閘道通訊協定 (BGP) 最佳路由反射

目錄

[簡介](#)

[背景資訊](#)

[網路圖表](#)

[理論](#)

[IOS-XR實作](#)

[設定](#)

[組態範例](#)

[根路由器上的MPLS流量工程](#)

[疑難排解](#)

[結論](#)

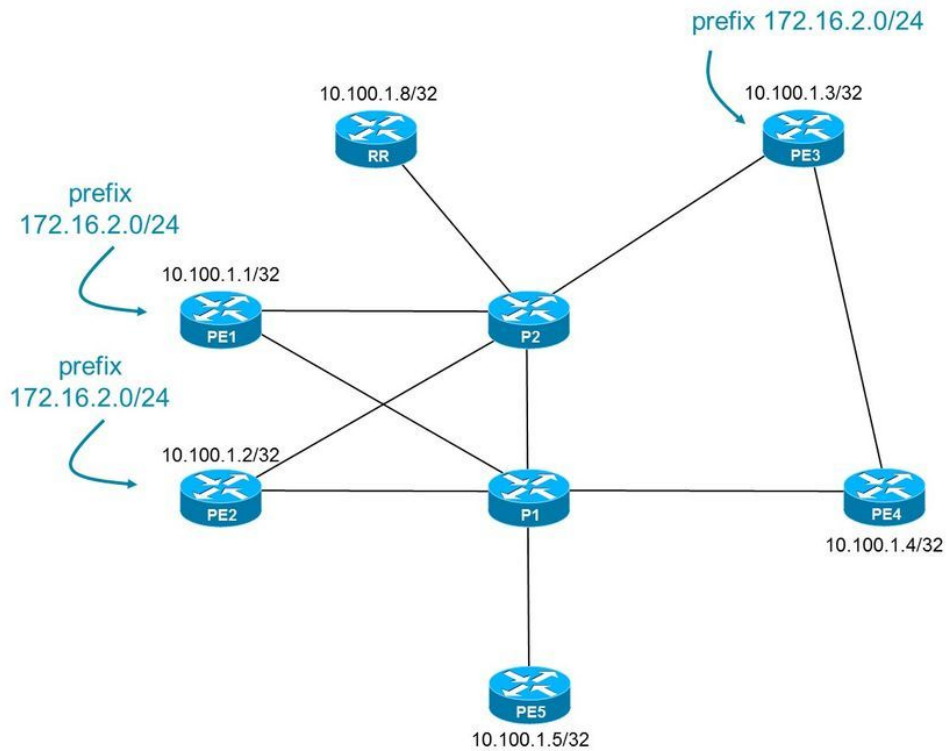
簡介

本檔案將說明當網路中有一個或多個路由反射器(RR)以避免在iBGP路由器之間產生全網狀網路時，如何影響路由。

背景資訊

[BGP最佳路徑選取演演算法](#)中的步驟8是優先使用具有最低IGP指標的路徑，而非BGP下一躍點。因此，如果步驟8之前的所有步驟相等，則步驟8可以成為決定哪個最佳路徑在RR上的決定因素。然後由RR的位置來確定從RR到廣告iBGP路由器的IGP開銷。預設情況下，RR只會向其使用者端通告最佳路徑。根據RR所在的位置，廣告路由器的IGP成本可以更小或更大。如果路徑的所有IGP成本相同，則最終可能會到達具有最低BGP路由器ID的廣告路由器的斷路器。

網路圖表



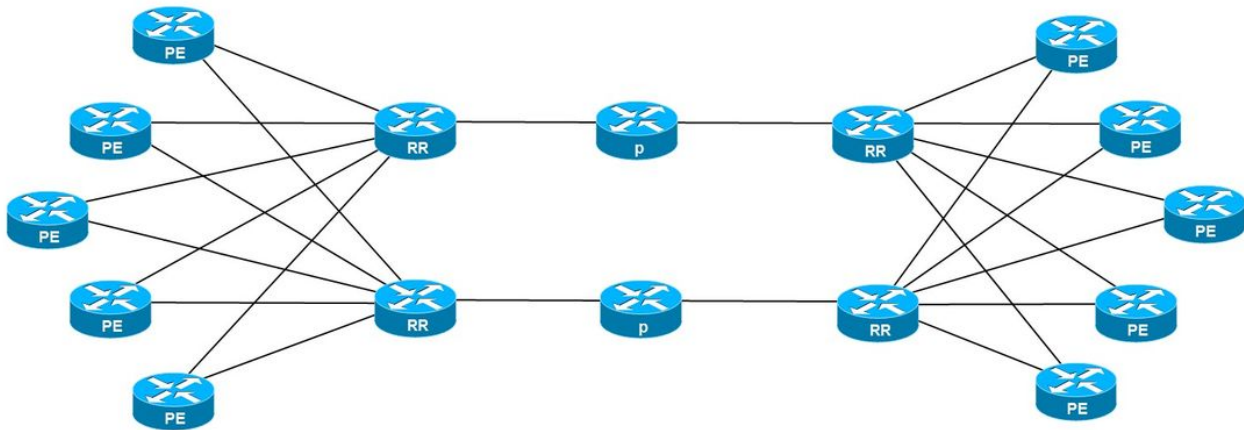
路由器PE1、PE2和PE3通告字首172.16.2.0/24。如果鏈路的所有IGP開銷相同，則RR將看到來自PE1、PE2和PE3的路徑，其IGP開銷為2。最後，RR從PE1中選擇最佳路徑，因為它具有較低的BGP路由器ID。這是BGP最佳路徑選取演演算法中的步驟11。結果是，包括PE4在內的所有PE路由器都會選擇PE1作為字首172.16.2.0/24的輸出PE路由器。從PE4的角度來看，通向任何輸出PE路由器的IGP路徑更短，即通往PE3的路徑，IGP開銷為1。通向任何其他PE路由器的IGP開銷為2。對於許多網路而言，以儘可能最短的方式傳輸流量這一事實很重要。這稱為熱薯路由。

RR從PE1中選擇最佳路徑還有另一個可能的原因。如果在影象中，鏈路P2-PE3的內部網關協定(IGP)開銷為10，而所有其他鏈路的IGP開銷仍為1，則RR不會從PE3中選擇最佳路徑，即使PE3的BGP路由器ID最低。

如果此網路的管理員想要熱薯路由，則必須建立一個機制，以便當網路中存在RR時，入口路由器仍然可以獲知到iBGP網路中最近的出口路由器的路徑。BGP功能Add Path可以達到此目的。但是，利用該功能，路由路由器和邊界路由器必須擁有瞭解該功能的較新代碼。由於BGP最佳路由反射的特性，這不是一項要求。此功能將允許RR根據從輸入BGP路由器的角度看來RR認為的最佳路徑來傳送最佳路徑到輸入BGP邊界路由器。

在部署RR時允許熱薯路由的另一個解決方案是RR的線上放置。這些RR不是專用的RR，它們僅運行BGP和IGP。這些內嵌RR位於轉送路徑中，並放在網路中，如此一來，這些內嵌RR具有其自己的RR使用者端組，因此它們可以反映到每個RR使用者端的最佳路徑，而從該RR使用者端的角度來看，這亦是最佳路徑。

如下圖所示，RR被置於網路中，以便它們可以服務的一小組鄰近的RR使用者端。由於網路設計，RR使用者端會從RR接收最佳路徑，此最佳路徑是從其視點開始的最佳路徑，如此一來，網路中便可能出現熱馬鈴薯路由。



理論

IETF draft-ietf-idr-bgp-optimal-route-reflection中描述了BGP最佳路由反射。

BGP最佳路由反射解決方案允許RR將特定最佳路徑傳送到特定BGP邊界路由器。RR可以選擇將不同的最佳路徑傳送到不同的BGP邊界路由器或一組邊界路由器。邊界路由器必須是RR的RR客戶端。目標是每個輸入BGP邊界路由器可以具有相同的字首的不同出口或輸出BGP路由器。如果入口邊界路由器總是將流量轉發到關閉的AS出口路由器，則允許熱傳送路由。

問題在於RR通常只會將相同的最佳路徑傳送到每個BGP邊界路由器，這會防止熱播式馬鈴薯路由。為了解決此問題，您需要RR能夠根據輸入BGP邊界路由器為相同首碼計算不同的最佳路徑。RR上的最佳路徑計算是根據輸入BGP邊界路由器的位置完成的，因此RR會從輸入邊界路由器的角度執行BGP最佳路徑計算。僅能這樣做的RR是從RR和入口邊界路由器所在的IGP角度全面瞭解網路拓撲的RR。為了滿足此要求，IGP必須是鏈路狀態路由協定。

在這種情況下，RR可以運行以入口邊界路由器為樹根的最短路徑優先(SPF)計算，並計算其他每台路由器的開銷。這樣，從入口邊界路由器到所有其他出口邊界路由器的開銷是已知的。以另一台路由器為根的這種特殊的SPF計算稱為反向SPF(rSPF)。僅當RR從所有BGP邊界路由器獲取所有BGP路徑時，才能執行此操作。運行的rSPF可能與RR客戶端的數量相同。這將在一定程度上增加RR上的CPU負載。

此解決方案允許最佳路徑計算基於BGP最佳路徑選取演演算法，這會導致RR從RR傳送路徑的輸入邊界路由器的角度選取最佳路徑。這表示會根據到BGP下一躍點的最短IGP成本來選取最佳路徑。該解決方案還允許根據某些已配置的策略選擇最佳路徑。入口邊界路由器可以根據某些已配置的策略選擇其最佳路徑，而不是以最低的IGP成本來選擇。此解決方案允許RR在IGP成本（在網路上的位置）或某些已配置的策略上，或兩者上實施最佳路由反射。如果兩者都部署，則首先應用策略，然後在剩餘路徑上發生基於IGP的最佳路由反射。

IOS-XR實作

IOS-XR實現最多允許三個根節點進行rSPF計算。如果您在一個更新組中有許多RR客戶端，則如果

這些RR客戶端對於不同的輸出BGP邊界路由器將具有相同的策略和/或相同的IGP成本，則無需對每個RR客戶端進行一次rSPF計算。後者通常意味著RR客戶端位於同一位置（可能位於同一個POP中）。如果是這種情況，不需要將每個RR客戶端配置為根。IOS-XR實現允許為冗餘目的為每一組RR客戶端配置三個根，即主根、次根和第三根。要將BGP ORR功能應用於任何RR客戶端，該RR客戶端必須配置為屬於ORR策略組。

每個地址系列啟用BGP ORR功能。

鏈路狀態協定是必需的。可以是OSPF或IS-IS。

IOS XR僅根據BGP下一躍點的IGP成本實施BGP ORR功能，而不是根據某些已配置的策略。

具有相同出站策略的BGP對等體置於同一個更新組中。對RR上的iBGP通常是這種情況。啟用功能BGP ORR後，來自不同ORR組的對等體將位於不同的更新組中。這是符合邏輯的，因為從不同BGP ORR組中的RR客戶端傳送的更新將不同，因為BGP最佳路徑不同。

rSPF計算結果儲存在資料庫中。

ORRSPF是IOS-XR中的新元件，BGP ORR功能需要該元件。ORRSPF負責處理：

1. 收集鏈路狀態資訊並維護鏈路狀態資料庫
2. 按策略組運行rSPF並維護SPT
3. 使用度量將字首從SPT下載到RIB

資料庫直接從鏈路狀態IGP或BGP-LS獲取其鏈路狀態資訊。

rSPF計算得出的拓撲顯示從RR客戶端到區域/級別中任何其他路由器的最短路徑。

拓撲中每台路由器上掛起的路由都儲存在ORR組策略和每個AFI/SAFI的特殊RIB表中。此表由RSI建立。該表由以主根為根的rSPF計算的路由填充。如果主根不可用，則輔助根是根並填充ORR RIB表中的路由。這同樣適用於第三根。

設定

所需的最低配置：

1. 需要為特定BGP鄰居組的BGP地址系列啟用ORR
2. 對於每組BGP鄰居，至少需要配置一個根。或者，可以配置第二根和第三根。
3. 需要啟用ORR路由從IGP重分發到BGP。

組態範例

如第一幅影象所示，RR是具備BGP ORR功能的IOS-XR路由器。

所有其他路由器都運行IOS。這些路由器沒有BGP ORR功能。

PE1、PE2和PE3在AFI/SAFI 1/1（IPv4單播）中通告字首172.16.2.0/24。RR與PE1和PE2的距離相等於PE3。所有鏈路的IGP開銷均為1。此字首的最佳路徑是將R1作為下一跳的路徑，因為BGP路由器ID最低。

```

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將接收帶PE1作為下一跳的路徑。因此，PE4沒有熱馬鈴薯路由。

如果您想在PE4上使用熱播路由，那麼對於PE1、PE2和PE3通告的字首(例如字首172.16.2.0/24),PE1應該將PE3作為出口點。這意味著PE4上的路徑應該是將PE3作為下一跳的路徑。我們可以使RR將具有下一跳PE3的路由傳送到具有此ORR配置的PE4。

```

router ospf 1
distribute bgp-ls
  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
!
!
!

```

如果IGP為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
!
!
!

```

附註：無需全域性配置地址系列鏈路狀態或在BGP鄰居下配置。

根路由器上的MPLS流量工程

RR需要在IGP資料庫中找到已配置的根地址，才能運行rSPF。在ISIS中，路由器ID存在於ISIS資料庫中。對於OSPF，OSPF LSA中不存在路由器ID。解決方案是讓根路由器通告與RR上配置的根地址匹配的多協定標籤交換(MPLS)TE路由器ID。

對於OSPF，根路由器需要額外的配置來使BGP OR正常工作。任何根路由器上都需要最少的MPLS TE配置才能通告此MPLS TE路由器ID。精確的最小命令集取決於根路由器的作業系統。根路由器上的MPLS TE配置需要啟用MPLS TE的最低配置，以便OSPF在不透明區域LSA (型別10) 中通告

MPLS TE路由器ID。

一旦RR具有不透明區域LSA，其MPLS TE路由器ID與配置的根路由器地址匹配，rSPF可以運行，RR上的BGP可以通告最佳路由。

如果根路由器是IOS路由器，則OSPF所需的最低配置為：

```
!  
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  
!
```

請注意：

- 在特定OSPF區域中啟用MPLS TE
- mpls TE路由器ID配置為與RR上配置的根地址匹配
- 在至少一個介面上配置MPLS TE
- 無需配置RSVP-TE
- 無需在區域內的任何其他路由器上配置MPLS TE

如果根路由器是IOS-XR路由器，則OSPF所需的最低配置為：

```
!  
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  
!
```

如果在根路由器上有上述配置，則RR應將MPLS TE路由器ID放在OSPF資料庫中。

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

請注意，MPLS TE路由器ID(10.100.1.4)和OSPF路由器ID不同。

PE4將PE3作為字首的下一跳 (具有到下一跳的正確IGP度量) :

```
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仍將PE1作為字首的下一跳 (具有到下一跳的正確IGP度量) :

```
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
```

疑難排解

驗證RR : 上的字首

```
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
```

請注意，add-path已新增到其他非最佳路徑，因此除了最佳路徑之外，還可以通告這些路徑。RR與其使用者端之間不使用新增路徑功能：不會使用路徑識別符號通告路徑。

驗證是否已向特定BGP鄰居通告路由（仍然）。

對於鄰居PE4，下一跳是字首172.16.2.0/24的PE3:

```
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

對於鄰居PE5，下一跳是字首172.16.2.0/24的PE1:

```
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
```

鄰居10.100.1.4位於自己的更新組中，因為存在了ORR策略：

```
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

```

show orrspf database命令顯示ORR組及其根，

```
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
```

帶有detail關鍵字的同一命令將rSPF的根的開銷提供給同一OSPF區域中的其它路由器/字首：

```
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
```

表ID由RSI分配給ORR組和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

rSPF的根(R4/10.100.1.4)到其他路由器的開銷與show ip route ospf在PE4上看到的開銷相同：

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
```

BGP ORR組的RIB:

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

show bgp neighbor命令會顯示對等點是否為ORR根：

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

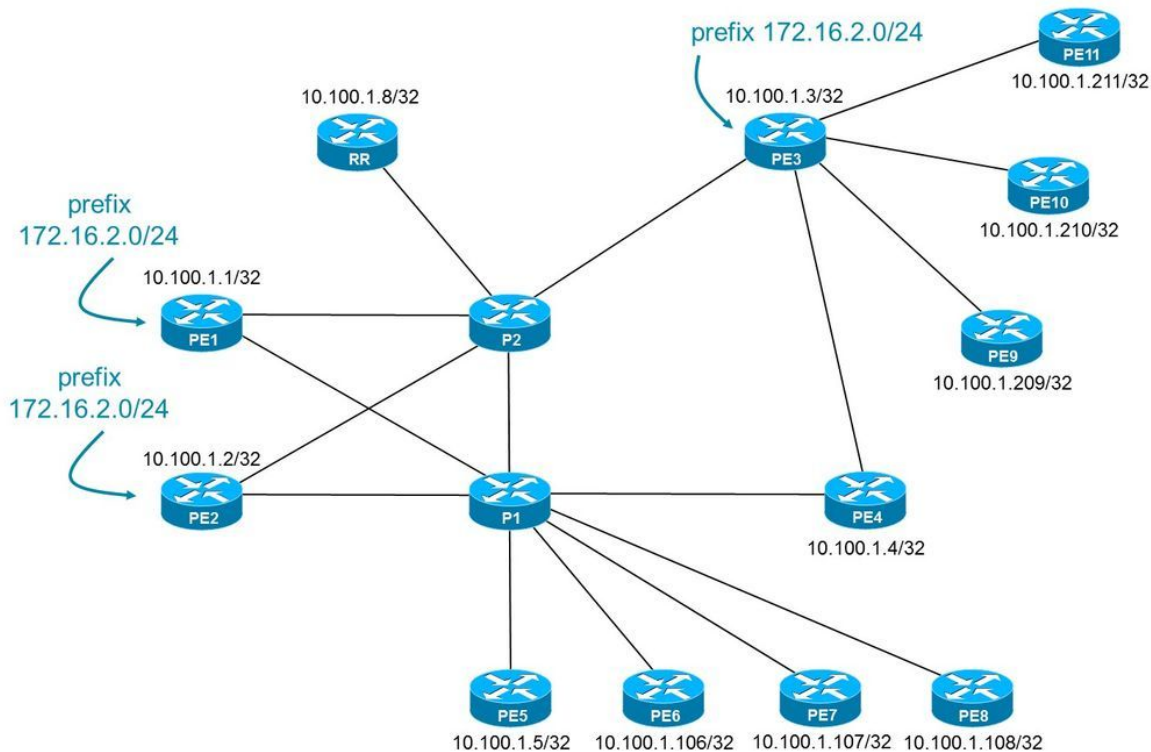
BGP neighbor is 10.100.1.4
Remote AS 1, local AS 1, internal link
Remote router ID 10.100.1.4
Cluster ID 10.100.1.8
BGP state = Established, up for 01:17:41
NSR State: None
Last read 00:00:52, Last read before reset 01:18:30
Hold time is 180, keepalive interval is 60 seconds
Configured hold time: 180, 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
```

如下圖所示，已配置多組RR客戶端



有一組RR客戶端連線到PE3，另一組連線到P1。每組中的RR客戶端與任何出口BGP邊界路由器相等距離。

```
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
  !
  !
!
```

兩個組的orrsfp資料庫：

```
RP/0/0/CPU0:RR#show orrsfp 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

如果組的主根已關閉或無法訪問，則輔助根將成為實際使用的根。在本例中，無法訪問組ipv4-orr-group-1的主根。次根成為實際的根：

```
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

結論

BGP最佳路由反射(ORR)是一種功能，當存在路由反射器時，它允許在iBGP網路中進行熱播式路由，而無需在邊界路由器上使用較新的作業系統軟體。前提條件是IGP是鏈路狀態路由協定。