

BGP 路由反射和多個叢集 ID

目錄

[簡介](#)

[BGP路由反射的說明](#)

[路由反射配置示例](#)

[具有預設設定的單個群集](#)

[禁用了客戶端到客戶端反射的單個群集](#)

[兩個集群，站點內和站點間路由反射](#)

[兩個群集，無客戶端到客戶端反射](#)

[叢集清單和回圈防止](#)

[客戶端與非客戶端之間的思考](#)

[簇內反射](#)

[叢集間反射](#)

[MCID和環路預防](#)

[參考資料](#)

[相關思科支援社群討論](#)

簡介

本文說明邊界閘道通訊協定 (BGP) 路由反射的不同情形，以及多個叢集 ID 的用途。假設先前已知 BGP 概念，特別是群集和路由反射。

BGP路由反射的說明

BGP 發言人是啟用 BGP 的路由器。預設情況下，BGP 揚聲器不會向 iBGP 對等體通告 iBGP 學習的字首 — 這樣做是為了保持環路預防。RFC 4456 引入了路由反射功能，無需在 iBGP 揚聲器之間實現全網狀。當路由反射器反射字首時，它通過新增自己的集群 ID 來建立/修改名為 CLUSTER_LIST 的可選非傳遞屬性。此屬性用於環路預防：當路由器收到包含路由器自己的群集 ID 的 CLUSTER_LIST 的更新時，將放棄此更新。

預設情況下，集群 ID 設定為 BGP 路由器 ID 值，但可以設定為任意 32 位值。多群集 ID (MCID) 功能允許為每個鄰居分配群集 ID。因此，路由反射方案有 3 種型別。

1. 在客戶端和非客戶端之間
2. 同一群集中的客戶端之間 (群集內)
3. 不同群集中的客戶端之間 (群集間)

路由反射配置示例

以下是一些路由器反射方案和各自的配置示例。

具有預設設定的單個群集

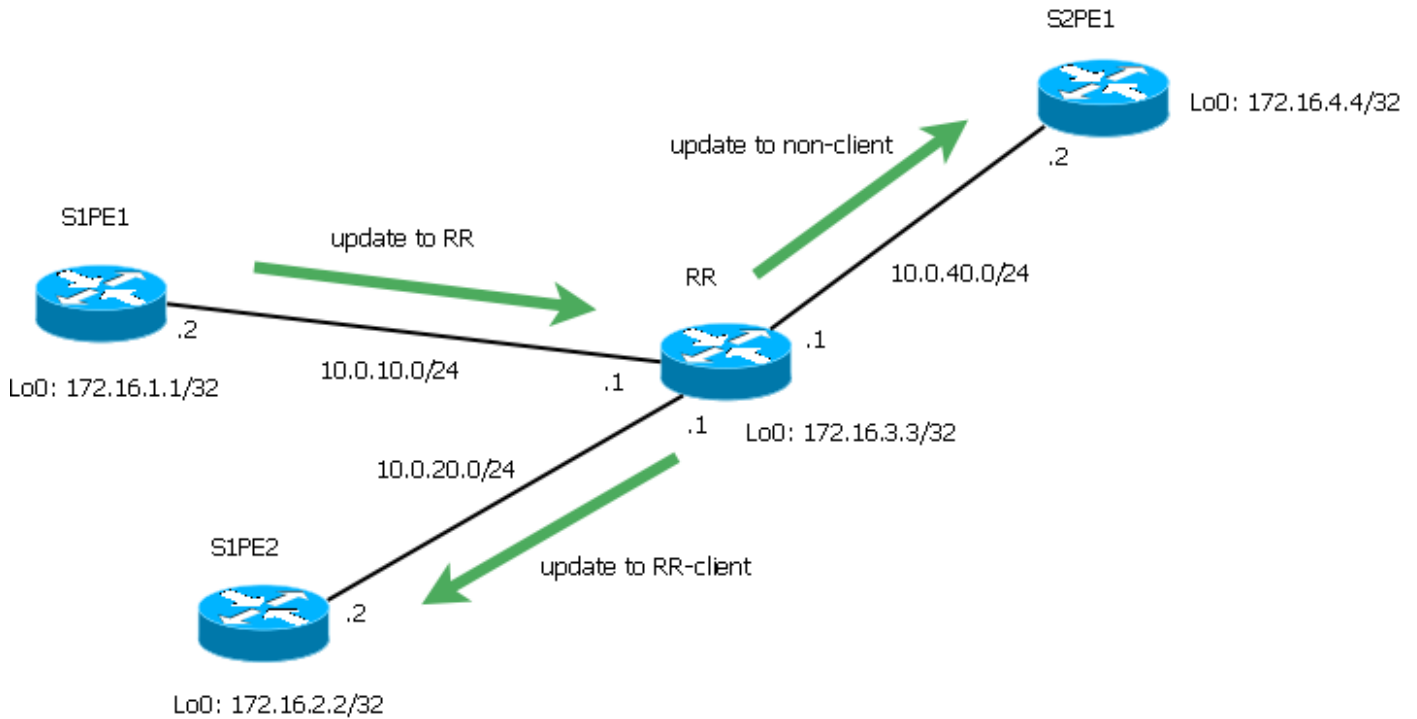


圖1

在充當路由反射器的路由器RR上完成了以下配置。

```
RR#show run | sec bgp
router bgp 1
  bgp log-neighbor-changes
  neighbor 10.0.10.2 remote-as 1
  neighbor 10.0.10.2 route-reflector-client
  neighbor 10.0.20.2 remote-as 1
  neighbor 10.0.20.2 route-reflector-client
  neighbor 10.0.40.2 remote-as 1
```

在這種情況下，S1PE1和S1PE2是RR的客戶端，而S2PE1是非客戶端。在常規設計中，非客戶端路由器將成為下一層級路由器的路由反射器，但在本例中，為了簡便起見，僅使用另一個PE。

```
RR#show ip bgp cluster-ids
Global cluster-id: 172.16.3.3 (configured: 0.0.0.0)
BGP client-to-client reflection:          Configured   Used
  all (inter-cluster and intra-cluster): ENABLED
  intra-cluster:                          ENABLED       ENABLED
```

```
List of cluster-ids:
Cluster-id      #-neighbors C2C-rfl-CFG C2C-rfl-USE
```

```
RR#sh ip bgp 172.16.1.1
BGP routing table entry for 172.16.1.1/32, version 2
Paths: (1 available, best #1, table default)
  Advertised to update-groups:
    1      2
  Refresh Epoch 2
  Local, (Received from a RR-client)
    10.0.10.2 from 10.0.10.2 (172.16.1.1)
      Origin IGP, metric 0, localpref 100, valid, internal, best
      rx pathid: 0, tx pathid: 0x0
```

```
RR#show ip bgp update-group 1
BGP version 4 update-group 1, internal, Address Family: IPv4 Unicast
  BGP Update version : 4/0, messages 0
  Topology: global, highest version: 4, tail marker: 4
  Format state: Current working (OK, last not in list)
                Refresh blocked (not in list, last not in list)
  Update messages formatted 2, replicated 2, current 0, refresh 0, limit 1000
  Number of NLRI in the update sent: max 1, min 0
  Minimum time between advertisement runs is 0 seconds
  Has 1 member:
    10.0.40.2
```

```
RR#show ip bgp update-group 2
BGP version 4 update-group 2, internal, Address Family: IPv4 Unicast
  BGP Update version : 4/0, messages 0
  Route-Reflector Client
  Topology: global, highest version: 4, tail marker: 4
  Format state: Current working (OK, last not in list)
                Refresh blocked (not in list, last not in list)
  Update messages formatted 3, replicated 6, current 0, refresh 0, limit 1000
  Number of NLRI in the update sent: max 1, min 0
  Minimum time between advertisement runs is 0 seconds
  Has 2 members:
    10.0.10.2      10.0.20.2
```

這些輸出顯示，RR從S1PE1收到172.16.1.1/32字首，並將其反映到客戶端S1PE2和非客戶端S2PE1。在此特定情況下，更新也會傳送回S1PE1，但這是因為S1PE1和S1PE2具有相同的路由策略，因此形成了相同的更新組。

禁用了客戶端到客戶端反射的單個群集

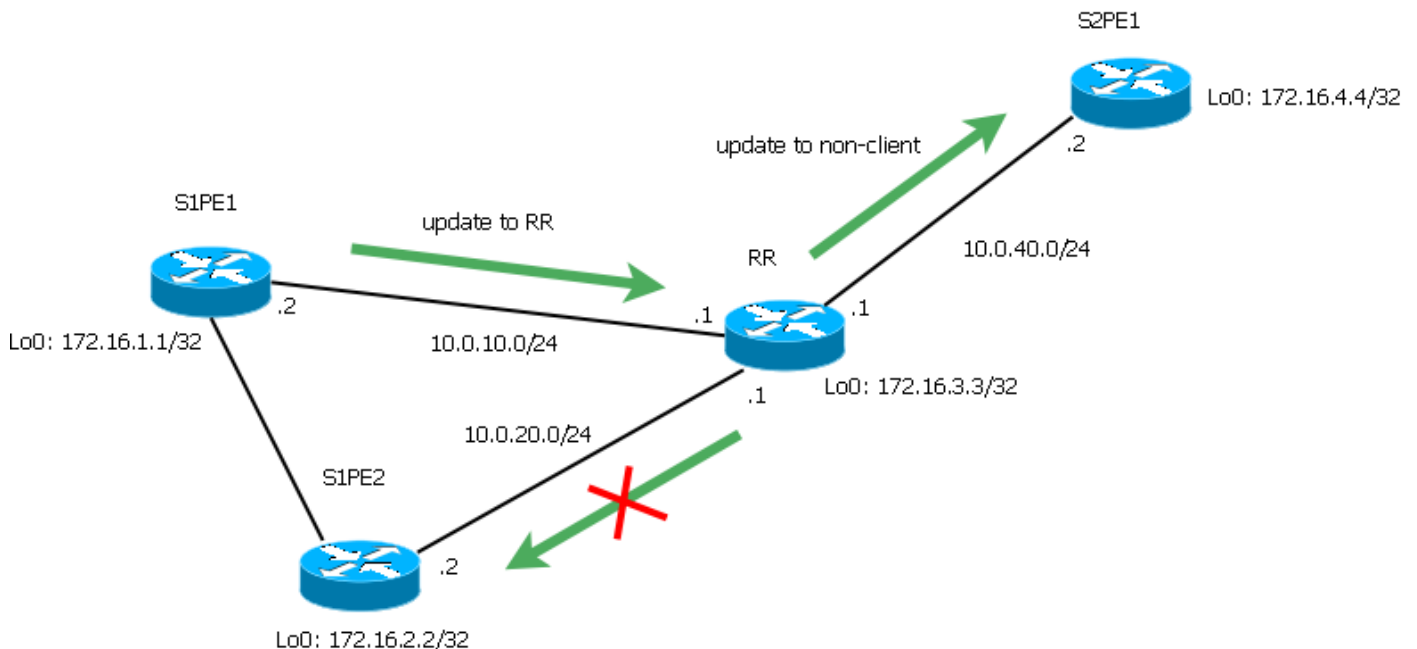


圖2

在充當路由反射器的路由器RR上完成了以下配置。

```
RR#show run | sec bgp
```

```
router bgp 1
  no bgp client-to-client reflection
  bgp log-neighbor-changes
  neighbor 10.0.10.2 remote-as 1
  neighbor 10.0.10.2 route-reflector-client
  neighbor 10.0.20.2 remote-as 1
  neighbor 10.0.20.2 route-reflector-client
  neighbor 10.0.40.2 remote-as 1
```

假設AS1部分網格化：S1PE1和S1PE2形成iBGP鄰居（例如，它們位於同一站點上，我們希望最佳化網路處理更新的方式）。在此案例中，RR已停用使用者端對使用者端的反射，它只會將172.16.1.1/32從S1PE1反射到非使用者端S2PE1。

```
RR#show ip bgp cluster-ids
Global cluster-id: 172.16.3.3 (configured: 0.0.0.0)
BGP client-to-client reflection:          Configured   Used
  all (inter-cluster and intra-cluster): DISABLED
  intra-cluster:                          ENABLED       DISABLED
```

```
List of cluster-ids:
Cluster-id      #-neighbors C2C-rfl-CFG C2C-rfl-USE
```

```
RR#show ip bgp 172.16.1.1
BGP routing table entry for 172.16.1.1/32, version 5
Paths: (1 available, best #1, table default, RIB-failure(17))
  Advertised to update-groups:
    1
  Refresh Epoch 2
  Local, (Received from a RR-client)
    10.0.10.2 from 10.0.10.2 (172.16.1.1)
      Origin IGP, metric 0, localpref 100, valid, internal, best
      rx pathid: 0, tx pathid: 0x0
```

```
RR#show ip bgp update-group 1 BGP version 4 update-group 1, internal, Address Family: IPv4
Unicast BGP Update version : 7/0, messages 0 Topology: global, highest version: 7, tail marker:
7 Format state: Current working (OK, last not in list) Refresh blocked (not in list, last not in
list) Update messages formatted 4, replicated 4, current 0, refresh 0, limit 1000 Number of
NLRIs in the update sent: max 1, min 0 Minimum time between advertisement runs is 0 seconds Has
1 member: 10.0.40.2
```

兩個集群，站點內和站點間路由反射

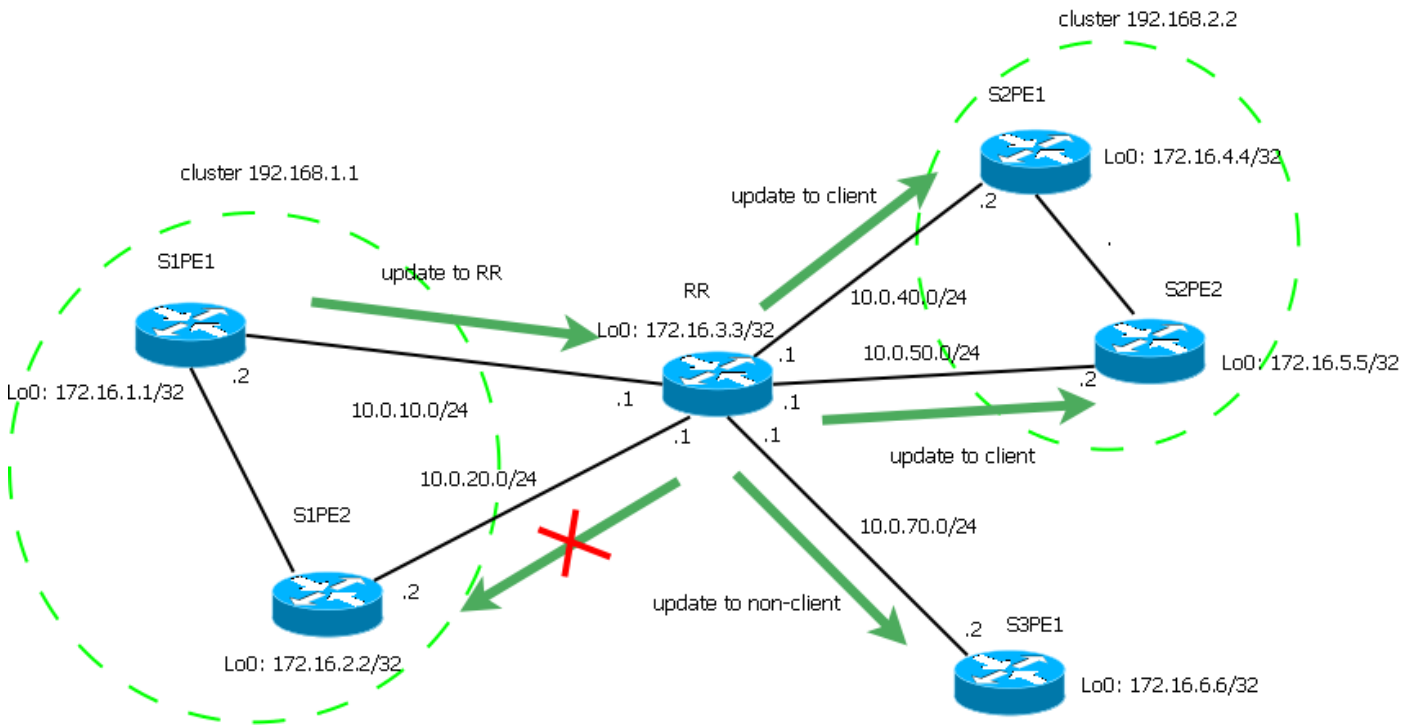


圖3

在充當路由反射器的路由器RR上完成了以下配置。

```
RR#sh run | sec bgp
router bgp 1
  no bgp client-to-client reflection intra-cluster cluster-id 192.168.1.1
  bgp log-neighbor-changes
  neighbor 10.0.10.2 remote-as 1
  neighbor 10.0.10.2 cluster-id 192.168.1.1
  neighbor 10.0.10.2 route-reflector-client
  neighbor 10.0.20.2 remote-as 1
  neighbor 10.0.20.2 cluster-id 192.168.1.1
  neighbor 10.0.20.2 route-reflector-client
  neighbor 10.0.40.2 remote-as 1
  neighbor 10.0.40.2 cluster-id 192.168.2.2
  neighbor 10.0.40.2 route-reflector-client
  neighbor 10.0.50.2 remote-as 1
  neighbor 10.0.50.2 cluster-id 192.168.2.2
  neighbor 10.0.50.2 route-reflector-client
  neighbor 10.0.70.2 remote-as 1
```

在這種情況下，站點1上的兩個PE形成群集192.168.1.1，站點2上的兩個PE形成群集192.168.2.2。S3PE1是非客戶端。站點1上的PE具有直接iBGP會話，群集192.168.1.1禁用了群集內反射，但群集192.168.2.2仍啟用了該反射。啟用了群集間反射。

```
RR#show ip bgp cluster-ids
Global cluster-id: 172.16.3.3 (configured: 0.0.0.0)
BGP client-to-client reflection:           Configured   Used
  all (inter-cluster and intra-cluster): ENABLED
  intra-cluster:                          ENABLED       ENABLED
```

```
List of cluster-ids:
Cluster-id    #-neighbors  C2C-rfl-CFG  C2C-rfl-USE
192.168.1.1   2            DISABLED     DISABLED
192.168.2.2   2            ENABLED      ENABLED
```

```
RR#show ip bgp 172.16.1.1
BGP routing table entry for 172.16.1.1/32, version 5
Paths: (1 available, best #1, table default, RIB-failure(17))
  Advertised to update-groups:
    3          5
  Refresh Epoch 9
  Local, (Received from a RR-client)
    10.0.10.2 from 10.0.10.2 (172.16.1.1)
      Origin IGP, metric 0, localpref 100, valid, internal, best
      rx pathid: 0, tx pathid: 0x0
```

```
RR#show ip bgp update-group 3 BGP version 4 update-group 3, internal, Address Family: IPv4
Unicast BGP Update version : 11/0, messages 0 Topology: global, highest version: 11, tail
marker: 11 Format state: Current working (OK, last not in list) Refresh blocked (not in list,
last not in list) Update messages formatted 20, replicated 20, current 0, refresh 0, limit 1000
Number of NLRI in the update sent: max 1, min 0 Minimum time between advertisement runs is 0
seconds Has 1 member: 10.0.70.2
```

```
RR#show ip bgp update-group 5 BGP version 4 update-group 5, internal, Address Family: IPv4
Unicast BGP Update version : 11/0, messages 0 Route-Reflector Client Configured with cluster-id
192.168.2.2 Topology: global, highest version: 11, tail marker: 11 Format state: Current working
(OK, last not in list) Refresh blocked (not in list, last not in list) Update messages formatted
22, replicated 34, current 0, refresh 0, limit 1000 Number of NLRI in the update sent: max 3,
min 0 Minimum time between advertisement runs is 0 seconds Has 2 members: 10.0.40.2 10.0.50.2
```

從S1PE1收到的字首172.16.1.1/32將反映到群集192.168.2.2中的客戶端和非客戶端。同時，從S2PE1收到的字首172.16.4.4/32將反映到所有客戶端和非客戶端。

```
RR#show ip bgp 172.16.4.4
BGP routing table entry for 172.16.4.4/32, version 9
Paths: (1 available, best #1, table default, RIB-failure(17))
  Advertised to update-groups:
    3          4          5
  Refresh Epoch 6
  Local, (Received from a RR-client)
    10.0.40.2 from 10.0.40.2 (172.16.4.4)
      Origin IGP, metric 0, localpref 100, valid, internal, best
      rx pathid: 0, tx pathid: 0x0
```

```
RR#show ip bgp update-group 3
BGP version 4 update-group 3, internal, Address Family: IPv4 Unicast
  BGP Update version : 11/0, messages 0
  Topology: global, highest version: 11, tail marker: 11
  Format state: Current working (OK, last not in list)
    Refresh blocked (not in list, last not in list)
  Update messages formatted 20, replicated 20, current 0, refresh 0, limit 1000
  Number of NLRI in the update sent: max 1, min 0
  Minimum time between advertisement runs is 0 seconds
  Has 1 member:
    10.0.70.2
```

```
RR#show ip bgp update-group 4
BGP version 4 update-group 4, internal, Address Family: IPv4 Unicast
  BGP Update version : 11/0, messages 0
  Route-Reflector Client
  Configured with cluster-id 192.168.1.1
  Topology: global, highest version: 11, tail marker: 11
  Format state: Current working (OK, last not in list)
    Refresh blocked (not in list, last not in list)
  Update messages formatted 26, replicated 47, current 0, refresh 0, limit 1000
  Number of NLRI in the update sent: max 3, min 0
  Minimum time between advertisement runs is 0 seconds
  Has 2 members:
```

10.0.10.2 10.0.20.2

```
RR#show ip bgp update-group 5
BGP version 4 update-group 5, internal, Address Family: IPv4 Unicast
  BGP Update version : 11/0, messages 0
  Route-Reflector Client
  Configured with cluster-id 192.168.2.2
  Topology: global, highest version: 11, tail marker: 11
  Format state: Current working (OK, last not in list)
                 Refresh blocked (not in list, last not in list)
  Update messages formatted 22, replicated 34, current 0, refresh 0, limit 1000
  Number of NLRI's in the update sent: max 3, min 0
  Minimum time between advertisement runs is 0 seconds
  Has 2 members:
    10.0.40.2            10.0.50.2
```

您也可以禁用集群192.168.2.2的站點內路由反射，但是在這種情況下，該集群中的客戶端應具有全網狀iBGP會話：

```
RR(config-router)#no bgp client-to-client reflection intra-cluster cluster-id 192.168.2.2
```

```
RR#sh ip bgp cluster-ids
Global cluster-id: 172.16.3.3 (configured: 0.0.0.0)
BGP client-to-client reflection:            Configured    Used
all (inter-cluster and intra-cluster):  ENABLED
intra-cluster:                            ENABLED        ENABLED
```

```
List of cluster-ids:
Cluster-id    #-neighbors  C2C-rfl-CFG  C2C-rfl-USE
192.168.1.1        2  DISABLED    DISABLED
192.168.2.2        2  DISABLED    DISABLED
```

還可以對所有群集禁用站點內反射：

```
RR(config-router)#no bgp client-to-client reflection intra-cluster cluster-id any
```

兩個群集，無客戶端到客戶端反射

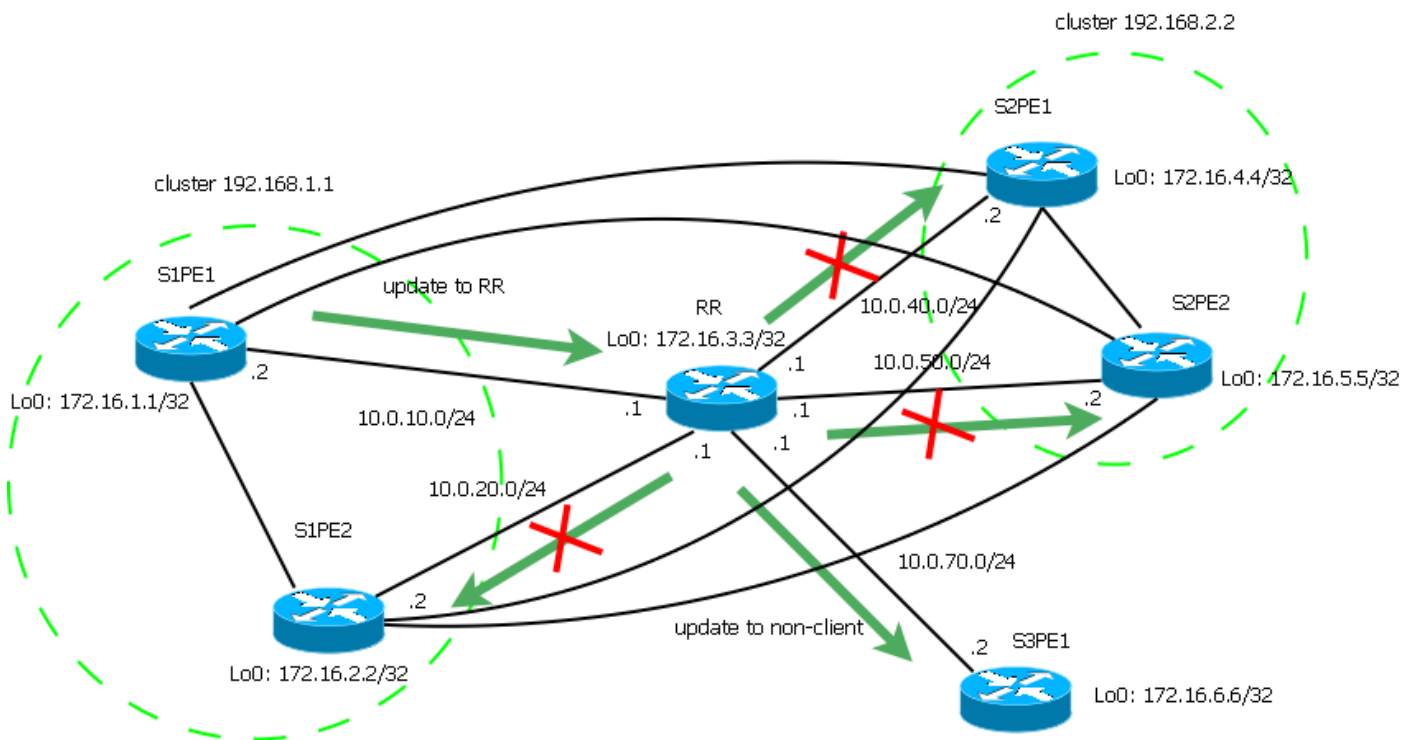


圖4

在充當路由反射器的路由器RR上完成了以下配置。

```
RR#show run | sec bgp
router bgp 1
  no bgp client-to-client reflection
  bgp log-neighbor-changes
  neighbor 10.0.10.2 remote-as 1
  neighbor 10.0.10.2 cluster-id 192.168.1.1
  neighbor 10.0.10.2 route-reflector-client
  neighbor 10.0.20.2 remote-as 1
  neighbor 10.0.20.2 cluster-id 192.168.1.1
  neighbor 10.0.20.2 route-reflector-client
  neighbor 10.0.40.2 remote-as 1
  neighbor 10.0.40.2 cluster-id 192.168.2.2
  neighbor 10.0.40.2 route-reflector-client
  neighbor 10.0.50.2 remote-as 1
  neighbor 10.0.50.2 cluster-id 192.168.2.2
  neighbor 10.0.50.2 route-reflector-client
  neighbor 10.0.70.2 remote-as 1
```

可以禁用集群內和集群間反射。在這種情況下，將只執行客戶端和非客戶端之間的反射。

```
RR#show ip bgp cluster-ids
Global cluster-id: 172.16.3.3 (configured: 0.0.0.0)
BGP client-to-client reflection:          Configured    Used
  all (inter-cluster and intra-cluster): DISABLED
  intra-cluster:                          ENABLED         DISABLED

List of cluster-ids:
Cluster-id    #-neighbors  C2C-rfl-CFG  C2C-rfl-USE
192.168.1.1   2           ENABLED      DISABLED
192.168.2.2   2           ENABLED      DISABLED
```

```
RR#show ip bgp 172.16.1.1
BGP routing table entry for 172.16.1.1/32, version 5
Paths: (1 available, best #1, table default, RIB-failure(17))
  Advertised to update-groups:
    3
  Refresh Epoch 9
  Local, (Received from a RR-client)
    10.0.10.2 from 10.0.10.2 (172.16.1.1)
      Origin IGP, metric 0, localpref 100, valid, internal, best
      rx pathid: 0, tx pathid: 0x0
```

```
RR#show ip bgp 172.16.4.4
BGP routing table entry for 172.16.4.4/32, version 9
Paths: (1 available, best #1, table default, RIB-failure(17))
  Advertised to update-groups:
    3
  Refresh Epoch 6
  Local, (Received from a RR-client)
    10.0.40.2 from 10.0.40.2 (172.16.4.4)
      Origin IGP, metric 0, localpref 100, valid, internal, best
      rx pathid: 0, tx pathid: 0x0
```

```
RR#show ip bgp update-group 3
```



```
BGP version 4 update-group 3, internal, Address Family: IPv4 Unicast
BGP Update version : 11/0, messages 0
Topology: global, highest version: 11, tail marker: 11
Format state: Current working (OK, last not in list)
                Refresh blocked (not in list, last not in list)
Update messages formatted 20, replicated 20, current 0, refresh 0, limit 1000
Number of NLRI's in the update sent: max 1, min 0
Minimum time between advertisement runs is 0 seconds
Has 1 member:
    10.0.70.2
```

字首172.16.1.1/32和172.16.4.4/32分別源自群集192.168.1.1和192.168.2.2。這兩個字首只反映到非客戶端S3PE1。在這種情況下，所有客戶端都必須完全網格化。通常，在此特定場景中MCID實際上沒有意義（對於單個群集，可以實現相同的行為），但是，如果您希望針對來自不同鄰居的路由使用不同的群集清單，則仍可以使用它們。

附註：禁用群集間反射時，無法啟用群集內反射（無論是針對特定群集還是針對所有群集）。

叢集清單和回圈防止

當RR反映字首時，它將群集ID新增到可選的非傳遞屬性CLUSTER_LIST。此外，它還將可選的非傳遞屬性ORIGINATOR_ID設定為對等體的路由器ID，該對等體已經將字首通告給RR。

當使用MCID且RR反映字首時，它使用為向RR通告該字首的對等體配置的群集ID。如果該對等體未配置特定群集ID，則使用全域性群集ID。

我們來看一些例子。RR啟用所有形式的路由反射。全域性群集ID為172.16.3.3，群集ID 192.168.1.1和192.168.2.2分別設定為站點1和站點2上的PE（請參閱上述拓撲圖）。

```
RR#show ip bgp cluster-ids
Global cluster-id: 172.16.3.3 (configured: 0.0.0.0)
BGP client-to-client reflection:          Configured      Used
  all (inter-cluster and intra-cluster): ENABLED
  intra-cluster:                          ENABLED          ENABLED
```

```
List of cluster-ids:
Cluster-id      #-neighbors C2C-rfl-CFG C2C-rfl-USE
192.168.1.1      2 ENABLED   ENABLED
192.168.2.2      2 ENABLED   ENABLED
```

客戶端與非客戶端之間的思考

```
S2PE3#show ip bgp 172.16.1.1
BGP routing table entry for 172.16.1.1/32, version 2
Paths: (1 available, best #1, table default, RIB-failure(17))
  Not advertised to any peer
  Refresh Epoch 1
  Local
    10.0.10.2 (metric 20) from 10.0.70.1 (172.16.3.3)
      Origin IGP, metric 0, localpref 100, valid, internal, best
      Originator: 172.16.1.1, Cluster list: 192.168.1.1
      rx pathid: 0, tx pathid: 0x0
```

```
S2PE3#show ip bgp 172.16.4.4 BGP routing table entry for 172.16.4.4/32, version 4 Paths: (1
available, best #1, table default, RIB-failure(17)) Not advertised to any peer Refresh Epoch 1
Local 10.0.40.2 (metric 20) from 10.0.70.1 (172.16.3.3) Origin IGP, metric 0, localpref 100,
```

```
valid, internal, best Originator: 172.16.4.4, Cluster list: 192.168.2.2
  rx pathid: 0, tx pathid: 0x0
```

非客戶端S2PE3接收由群集192.168.1.1產生的字首172.16.1.1/32 — 群集ID 192.168.1.1已新增到群集清單中。它還會接收由群集192.168.2.2產生的字首172.16.4.4/32 — 群集ID 192.168.2.2已新增到群集清單中。

```
S1PE1#show ip bgp 172.16.6.6
BGP routing table entry for 172.16.6.6/32, version 5
Paths: (1 available, best #1, table default, RIB-failure(17))
  Not advertised to any peer
  Refresh Epoch 1
  Local
    10.0.70.2 (metric 20) from 10.0.10.1 (172.16.3.3)
      Origin IGP, metric 0, localpref 100, valid, internal, best
      Originator: 172.16.6.6, Cluster list: 172.16.3.3
      rx pathid: 0, tx pathid: 0x0
```

客戶端S1PE1接收由非客戶端發起的字首172.16.6.6/32 — 將全域性群集ID 172.16.3.3新增到群集清單中。

簇內反射

```
S1PE2#show ip bgp 172.16.1.1/32
BGP routing table entry for 172.16.1.1/32, version 8
Paths: (1 available, best #1, table default, RIB-failure(17))
  Not advertised to any peer
  Refresh Epoch 1
  Local
    10.0.10.2 (metric 20) from 10.0.20.1 (172.16.3.3)
      Origin IGP, metric 0, localpref 100, valid, internal
      Originator: 172.16.1.1, Cluster list: 192.168.1.1
      rx pathid: 0, tx pathid: 0
```

S1PE2屬於群集192.168.1.1，接收由S1PE1發起的字首172.16.1.1/32，該字首也屬於群集192.168.1.1。群集ID 192.168.1.1將新增到群集清單中。

叢集間反射

```
S2PE1#show ip bgp 172.16.1.1/32
BGP routing table entry for 172.16.1.1/32, version 4
Paths: (1 available, best #1, table default, RIB-failure(17))
  Not advertised to any peer
  Refresh Epoch 1
  Local
    10.0.10.2 (metric 20) from 10.0.40.1 (172.16.3.3)
      Origin IGP, metric 0, localpref 100, valid, internal, best
      Originator: 172.16.1.1, Cluster list: 192.168.1.1
      rx pathid: 0, tx pathid: 0x0
```

```
S1PE1#sh ip bgp 172.16.4.4/32
BGP routing table entry for 172.16.4.4/32, version 4
Paths: (1 available, best #1, table default, RIB-failure(17))
  Not advertised to any peer
  Refresh Epoch 1
  Local
    10.0.40.2 (metric 20) from 10.0.10.1 (172.16.3.3)
      Origin IGP, metric 0, localpref 100, valid, internal, best
```

Originator: 172.16.4.4, Cluster list: **192.168.2.2**

rx pathid: 0, tx pathid: 0x0

S2PE1屬於群集192.168.2.2，接收由群集192.168.1.1產生的字首172.16.1.1/32 — 群集ID設定為192.168.1.1。

S1PE1屬於群集192.168.1.1，並接收由群集192.168.2.2產生的字首172.16.4.4/32 — 群集ID設定為192.168.2.2。

MCID和環路預防

如果路由器收到包含路由器自己的群集ID的群集清單字首的更新，則丟棄該更新。如果使用MCID，將丟棄包含任何已配置群集ID（全域性或每個鄰居）的更新。

參考資料

[RFC 4456](#)

[BGP — 多個群集ID](#)