

BGP路由反射和多个集群ID

目录

[简介](#)

[BGP路由反射的说明](#)

[路由反射配置示例](#)

[具有默认设置的单个集群](#)

[禁用客户端到客户端反射的单个集群](#)

[两个集群，站点内和站点间路由反射](#)

[两个群集，无客户端到客户端反射](#)

[集群列表和环路防御](#)

[客户与非客户之间的思考](#)

[集群内反射](#)

[集群间反射](#)

[MCID和环路防御](#)

[参考](#)

[相关的思科支持社区讨论](#)

简介

本文介绍边界网关协议(BGP)路由反射和多个集群ID的使用的不同场景。假定先前对BGP概念（尤其是集群和路由反射）有所了解。

BGP路由反射的说明

BGP发言者是启用BGP的路由器。默认情况下，BGP扬声器不向iBGP对等体通告iBGP获知的前缀——这是为了维护环路防御。RFC4456引入了路由反射功能，无需iBGP扬声器之间的全网状网络。当路由反射器反映前缀时，它会通过向其添加自己的集群ID来创建/修改名为CLUSTER_LIST的可选非传递属性。此属性用于防止环路：当路由器收到包含路由器自己的集群ID的CLUSTER_LIST的更新时，此更新将被丢弃。

默认情况下，集群ID设置为BGP路由器ID值，但可设置为任意32位值。多集群ID(MCID)功能允许分配每个邻居的集群ID。因此，路由反射场景有3种类型。

1. 在客户端和非客户端之间
2. 在同一集群中的客户端之间（集群内）
3. 不同群集中的客户端之间（群集间）

路由反射配置示例

以下是一些路由器反射场景和相应的配置示例。

具有默认设置的单个集群

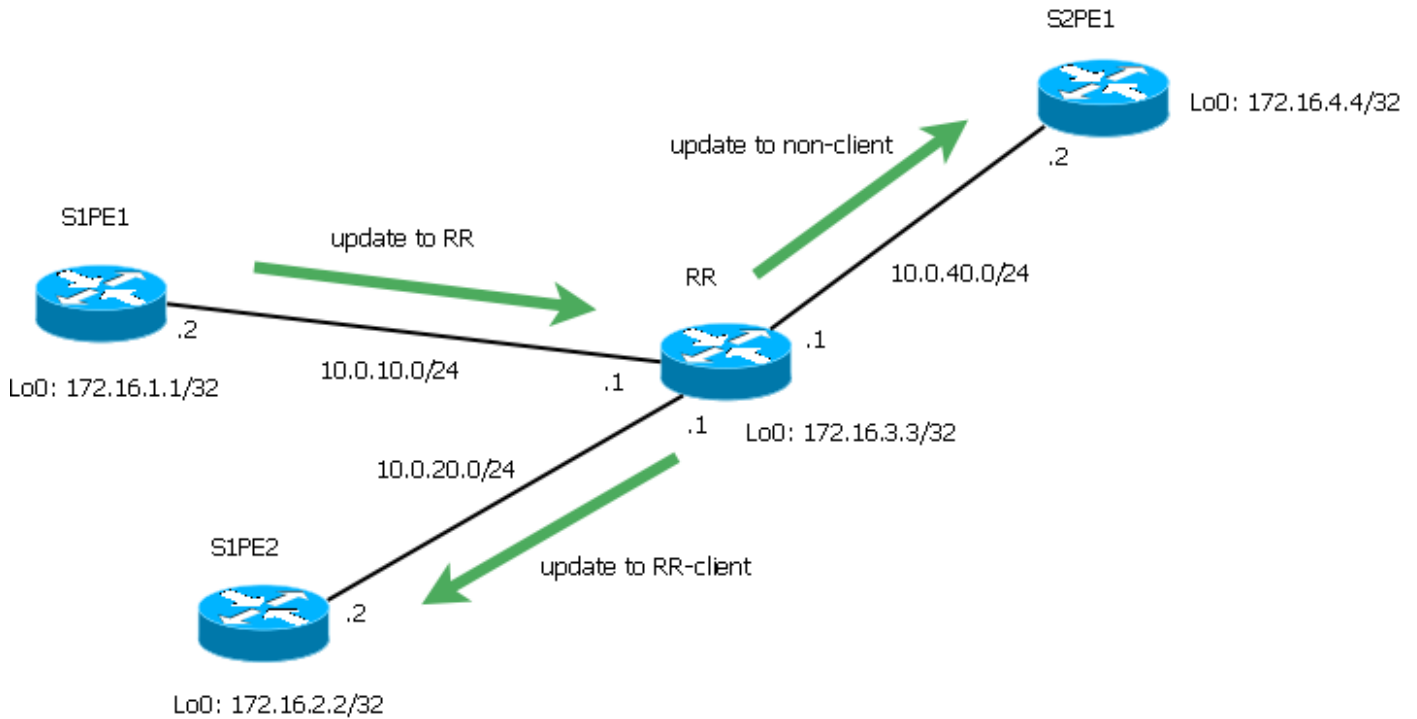


图 1

路由器RR充当路由反射器时已执行以下配置。

```
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
```

在本例中，S1PE1和S1PE2是RR的客户端，而S2PE1是非客户端。在传统设计中，非客户端路由器将成为下一层次级路由器的路由反射器，但在本例中，为简单起见，仅使用另一个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
```

这些输出显示，RR从S1PE1接收172.16.1.1/32前缀，并将其反映到客户端S1PE2和非客户端S2PE1。在这种情况下，更新也会发回到S1PE1，但是，由于S1PE1和S1PE2具有相同的路由策略，因此，从同一更新组。

禁用客户端到客户端反射的单个集群

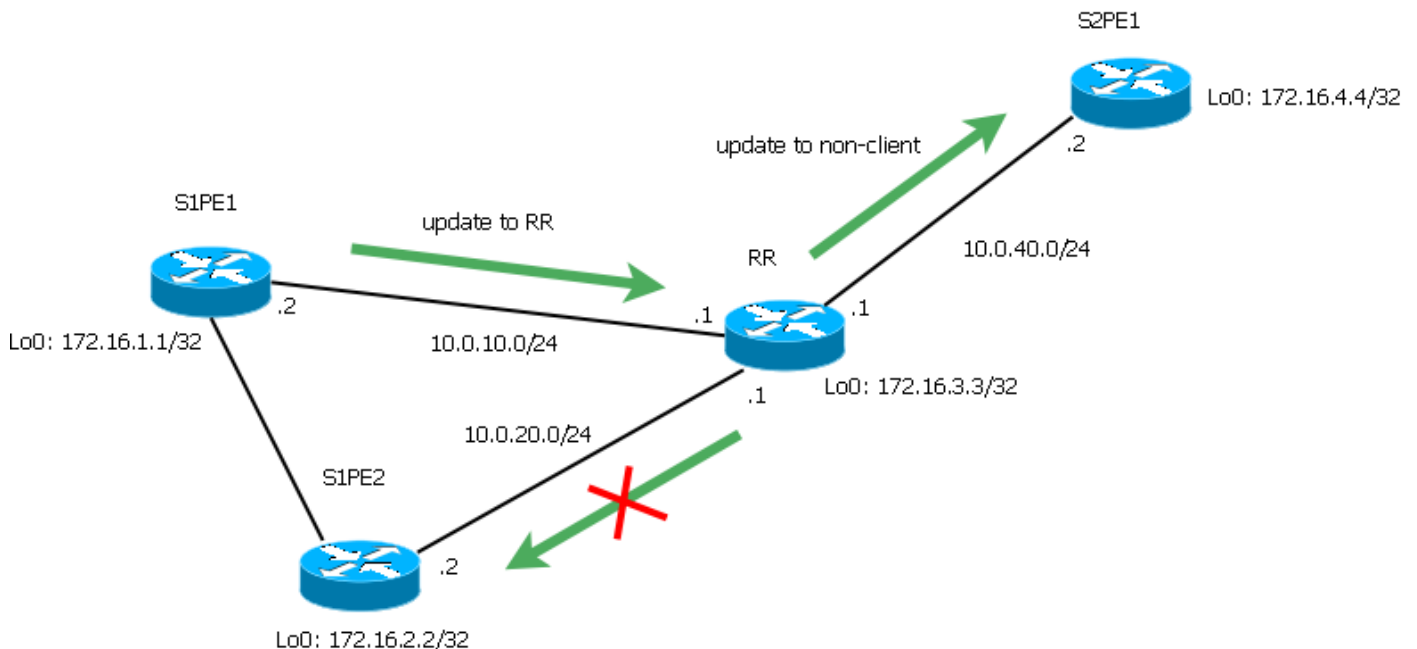


图 2

路由器RR充当路由反射器时已执行以下配置。

```
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
```

我们假设AS1已部分网格化：S1PE1和S1PE2形成iBGP邻居（例如，它们位于同一站点，我们希望优化网络处理更新的方式）。在这种情况下，RR禁用了客户端到客户端反射，它反映172.16.1.1/32仅从S1PE1到非客户端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
```

两个集群，站点内和站点间路由反射

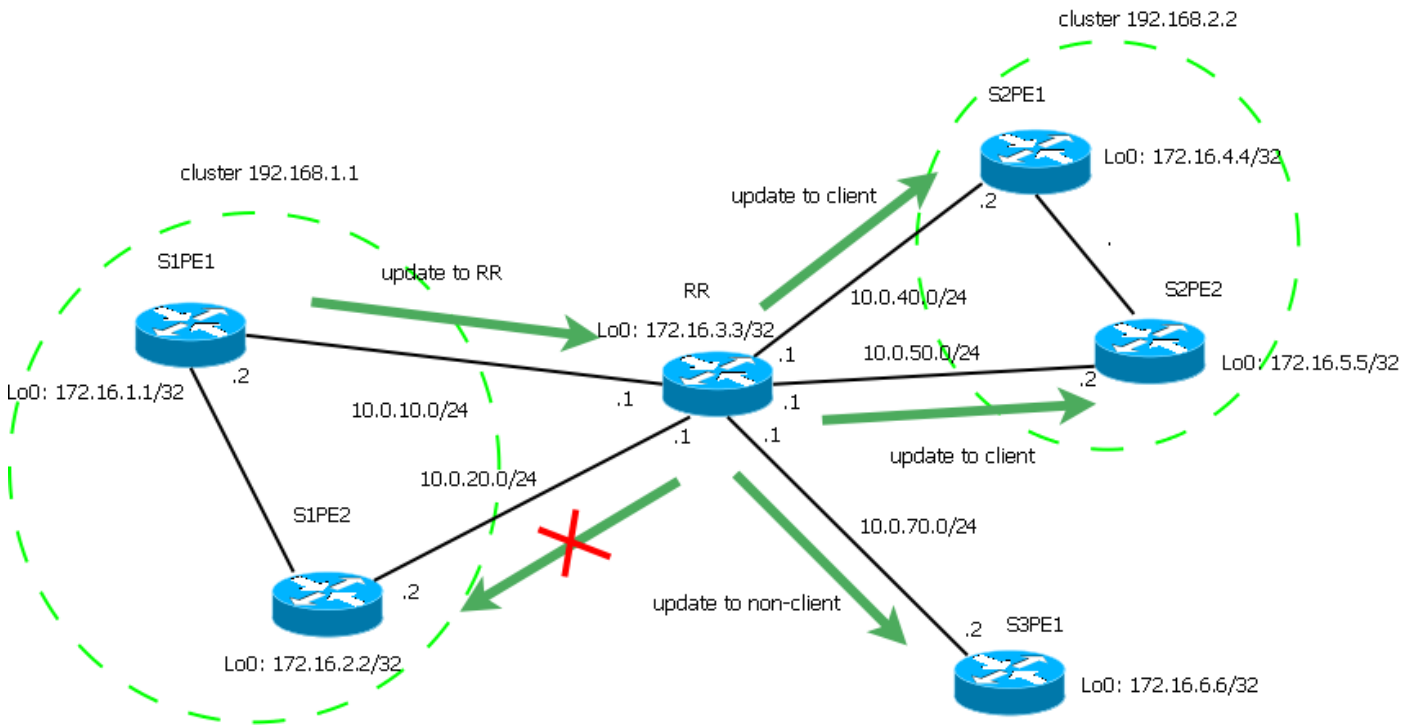


图 3

路由器RR充当路由反射器时已执行以下配置。

```
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
```

在本例中，站点1上的两个PE都形成集群192.168.1.1，而站点2上的两个PE都形成集群192.168.2.2。S3PE1是非客户端。站点1上的PE具有直接iBGP会话，集群192.168.1.1的集群内反射被禁用，但集群192.168.2.2的集群内反射仍然启用。集群间反射已启用。

```
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
```

从S1PE1收到的前缀172.16.1.1/32反映到集群192.168.2.2中的客户端和非客户端。同时，从S2PE1收到的前缀172.16.4.4/32反映给所有客户端和非客户端。

```
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 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的站点内路由反射，但在这种情况下，该集群中的客户端应具有iBGP会话的全网状：

```
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
```

还可禁用所有集群的站点内反射：

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

两个群集，无客户端到客户端反射

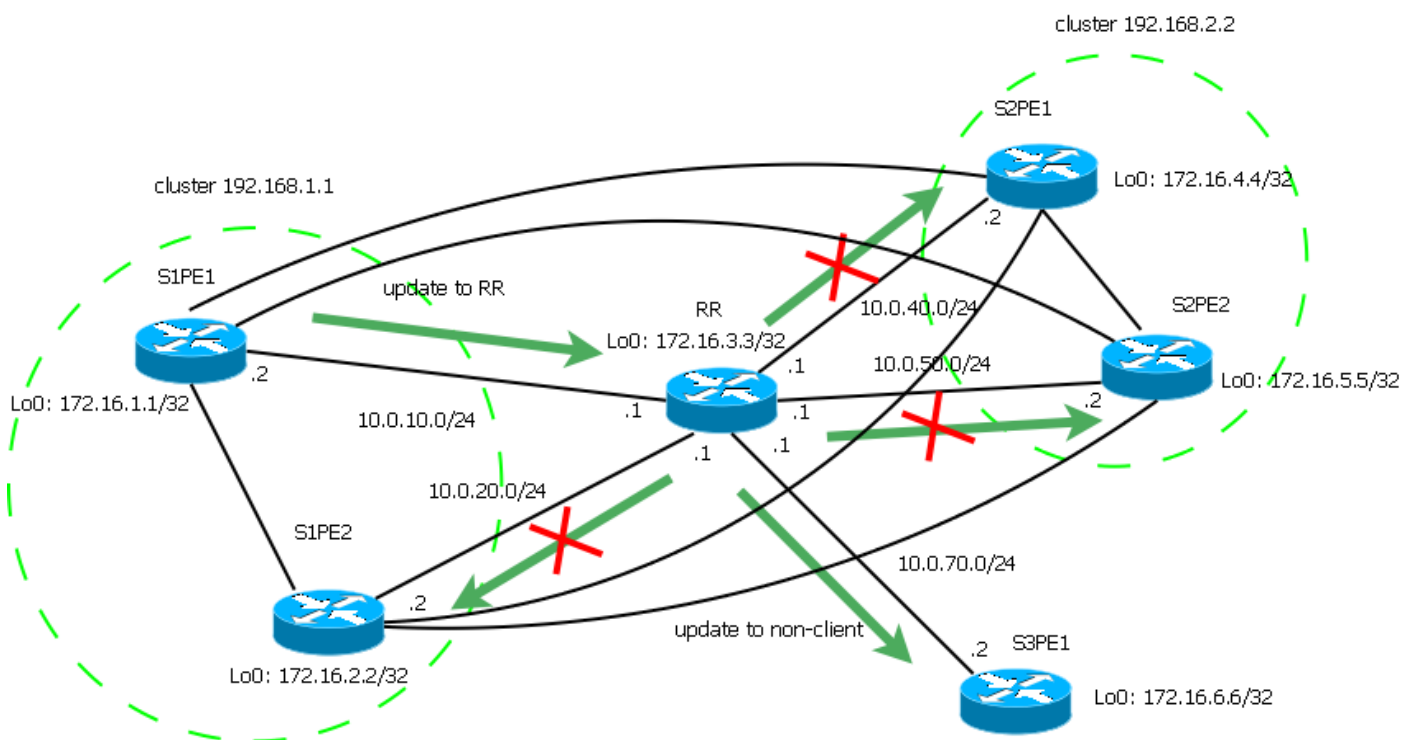


图 4

路由器RR充当路由反射器时已执行以下配置。

```
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
```

可以同时禁用集群内和集群间反射。在这种情况下，将只执行客户端和非客户端之间的反射。

```
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

```

前缀172.16.1.1/32和172.16.4.4/32分别由集群192.168.1.1和192.168.2.2发起。这两个前缀仅反映到非客户端S3PE1。在这种情况下，所有客户端必须全网状。通常，在此特定场景中，MCID并不真正有意义（使用单个集群可以实现相同的行为），但是，如果您想为来自不同邻居的路由使用不同的集群列表，则仍可使用它们。

注意：在禁用集群间反射时，无法启用集群内反射（针对特定集群或所有集群）。

集群列表和环路防御

当RR反映前缀时，它会将集群ID添加到可选非传递属性CLUSTER_LIST。此外，它还将可选非传递属性ORIGINATOR_ID设置为已将前缀通告给RR的对等体的路由器ID。

当使用MCID且RR反映前缀时，它使用为已将前缀通告给RR的对等体配置的集群ID。如果该对等体未配置特定集群ID，则使用全局集群ID。

让我们看一些示例。RR启用了所有形式的路由反射。全局群集ID为172.16.3.3，群集ID 192.168.1.1和192.168.2.2分别设置为站点1和站点2上的PE（请参阅上面的拓扑图）。

```

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

```

客户与非客户之间的思考

```

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
```

非客户端S2PE3接收由集群192.168.1.1发起的前缀172.16.1.1/32 — 集群ID 192.168.1.1添加到集群列表。它还接收由集群192.168.2.2发起的前缀172.16.4.4/32 — 集群ID 192.168.2.2添加到集群列表。

```
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
```

客户端S1PE1接收由非客户端发起的前缀172.16.6.6/32 — 全局群集ID 172.16.3.3添加到群集列表

。

集群内反射

```
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属于集群192.168.1.1，并接收由S1PE1发起的前缀172.16.1.1/32，该前缀也属于集群192.168.1.1。集群ID 192.168.1.1添加到集群列表。

集群间反射

```
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属于集群192.168.2.2，并接收由集群192.168.1.1发起的前缀172.16.1.1/32 — 集群ID设置为192.168.1.1。

S1PE1属于集群192.168.1.1，并接收由集群192.168.2.2发起的前缀172.16.4.4/32 — 集群ID设置为192.168.2.2。

MCID和环路防御

如果路由器收到包含路由器自己的集群ID的前缀的更新，则会丢弃更新。如果使用MCID，将丢弃包含任何已配置集群ID（全局或每邻居）的更新。

参考

[RFC 4456](#)

[BGP多集群ID](#)