使用BGP路由反射和多个集群ID

目录

简介

先决条件

- 要求
- 使用的组件

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
- 禁用了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
- 两个集群，无客户端到客户端反射

群集列表和环路预防

- 客户端与非客户端之间的思考
- 集群内反射
- 簇间反射
- MCID和环路预防

相关信息

简介

本文档介绍边界网关协议(BGP)路由反射和多个集群ID使用的不同场景。

先决条件

要求

假设预先了解了BGP概念，特别是集群和路由反射。

使用的组件

本文档不限于特定的软件和硬件版本。

本文档中的信息都是基于特定实验室环境中的设备编写的。本文档中使用的所有设备最初均采用原始（默认）配置。如果您的网络处于活动状态，请确保您了解所有命令的潜在影响。

BGP路由反射的说明

BGP扬声器是启用BGP的路由器。默认情况下，BGP扬声器不会向iBGP对等体通告iBGP学习的前缀 — 这样做是为了保持环路预防。RFC4456引入了路由反射功能，无需在iBGP扬声器之间使用全

网状。当路由反射器反映前缀时，它会创建/修改名为CLUSTER_LIST的可选非传递属性，方法是将自己的集群ID添加到其中。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，但可以设置为任意32位值。多集群ID(MCID)功能允许您为每个邻居分配集群ID。因此，路由反射场景有3种类型。

1. 客户端和非客户端之间
2. 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

本部分介绍一些路由器反射方案和相应的配置示例。

具有默认设置的单个集群

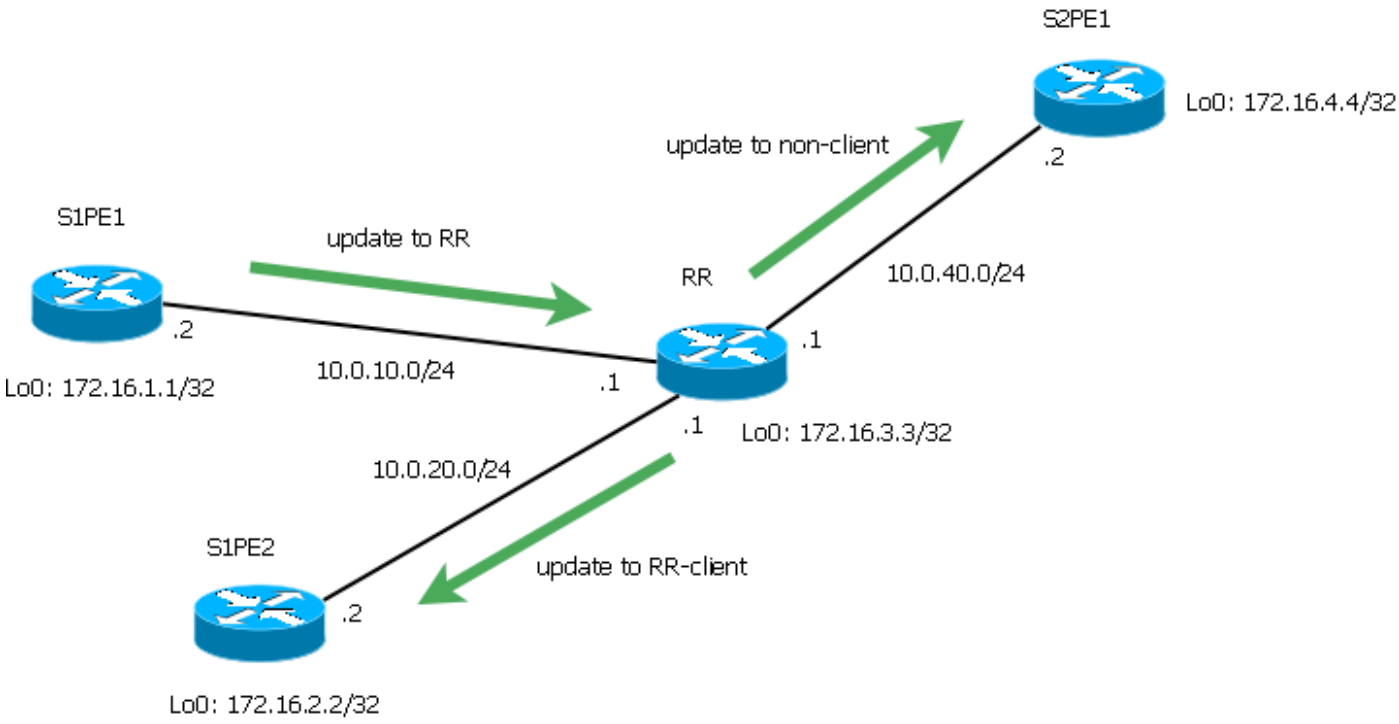


图 1

在充当路由反射器的路由器RR上完成了下一个配置。

<#root>

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
```

在这种情况下，S1PE1和S1PE2是RR的客户端，而S2PE1是非客户端。在传统设计中，非客户端路由器充当下一层级中路由器的路由反射器，但在本例中，为了简单起见，只使用另一个PE。

<#root>

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#

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)
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 NLRIs 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 NLRIs 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，更新也会发送回S1PE1，但这是因为S1PE1和S1PE2具有相同的路由策略，因此形成了同一个更新组。

禁用了客户端到客户端反射的单个集群

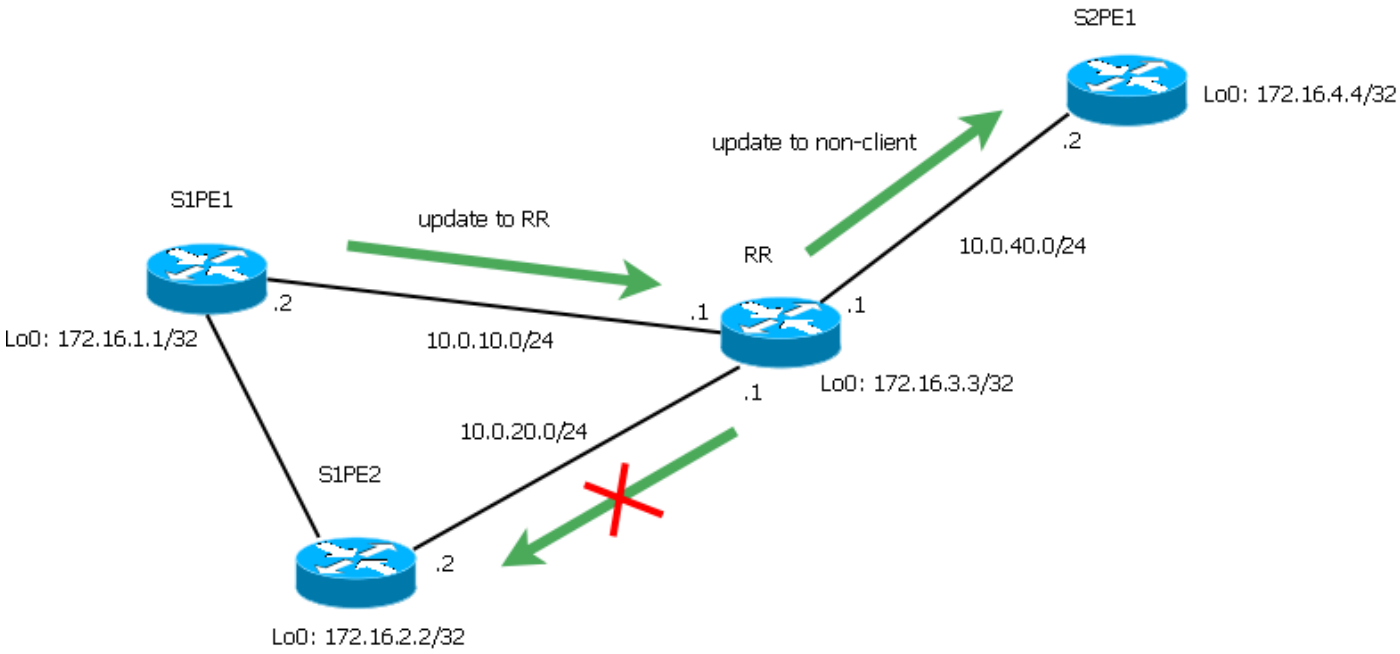


图 2

在充当路由反射器的路由器RR上完成了下一个配置。

<#root>

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
```

假设AS1部分网格化：S1PE1和S1PE2形成iBGP邻居关系（例如，它们位于同一站点，您希望优化网络处理更新的方式）。 在本例中，RR禁用了客户端到客户端的反射，它只将172.16.1.1/32从S1PE1反射到非客户端S2PE1。

<#root>

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

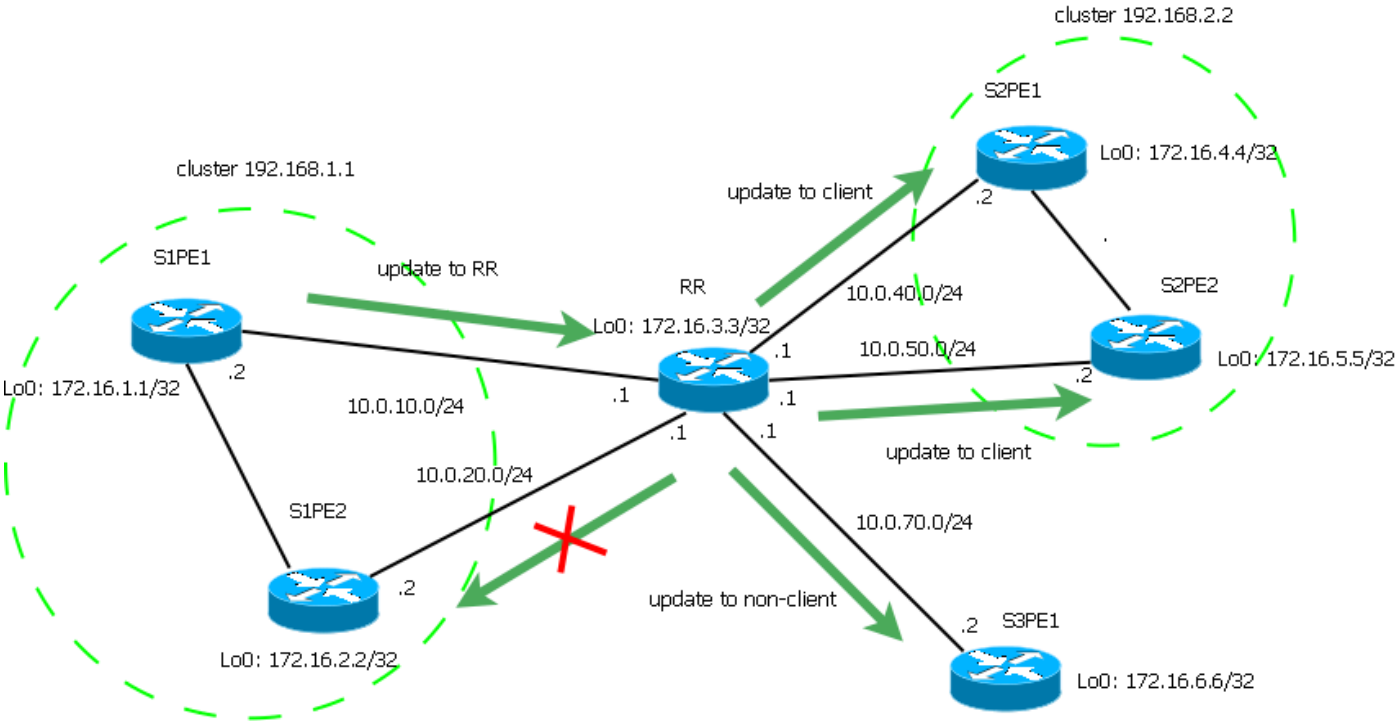


图 3

在充当路由反射器的路由器RR上完成了下一个配置。

<#root>

```
RR#
show 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
```

在这种情况下，站点1上的两个PE形成集群192.168.1.1，而站点2上的两个PE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，对集群192.168.1.1禁用集群内反射，但对集群192.168.2.2仍启用集群内反射。启用集群间反射。

<#root>

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 NLRIs 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 NLRIs 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接收的前缀172.16.1.1/32将反映到集群192.168.2.2中的客户端和非客户端。同时，从S2PE1接收的前缀172.16.4.4/32会反映到所有客户端和非客户端。

<#root>

```
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 NLRIs 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 NLRIs 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 NLRIs 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，但在此情况下，该集群中的客户端必须全网状iBGP会话：

<#root>

```
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

也可以为所有群集禁用站点内反射：

<#root>

```
RR(config-router)#
```

no bgp client-to-client reflection intra-cluster cluster-id any

两个集群，无客户端到客户端反射

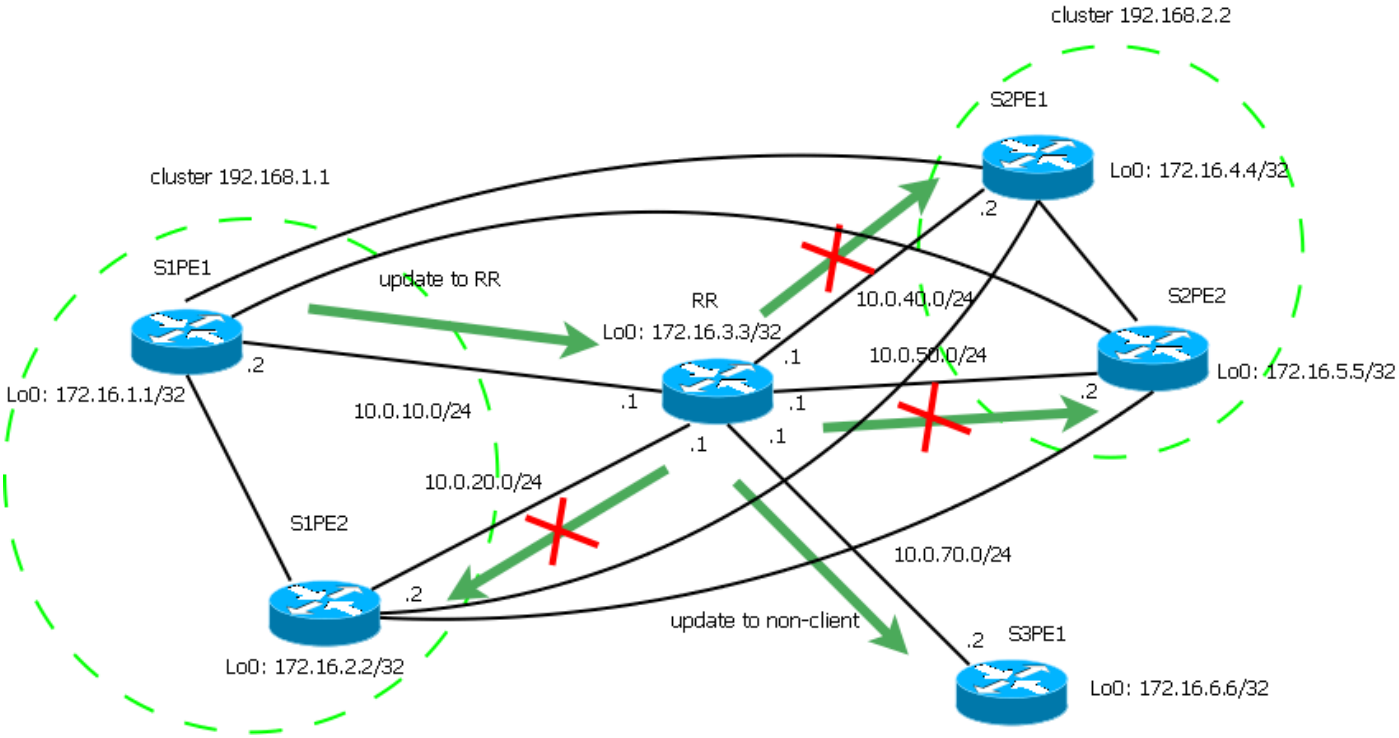


图 4

在充当路由反射器的路由器RR上完成了下一个配置。

<#root>

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
```

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

<#root>

```
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 NLRIs 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产生。这两个前缀都只反映到非客户端S3PE1。在这种情况下，所有客户端都必须完全网格化。通常，在此特定场景中，MCID实际上没有意义（同一行为可在单个集群中实现），但是，如果您希望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：禁用集群间反射时，无法启用集群内反射（无论是针对特定集群还是针对所有集群）。

群集列表和环路预防

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

当使用MCID且RR反映前缀时，它使用为向RR通告该前缀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。

让我们看一些例子。RR启用所有形式的路由反射。全局集群ID为172.16.3.3，集群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（请参阅上述拓扑图）。

```
<#root>
```

```
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
```

客户端与非客户端之间的思考

<#root>

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生成的前缀172.16.1.1/32 — 集群ID 192.168.1.1已添加到集群列表中。它还会接收由集群192.168.2.2发起的前缀172.16.4.4/32 — 集群ID 192.168.2.2已添加到集群列表中。

<#root>

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 — 全局集群ID 172.16.3.3将添加到集群列表中。

集群内反射

<#root>

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将添加到集群列表中。

簇间反射

<#root>

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#
```

`show 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，接收由集群192.168.1.1始发的前缀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始发的前缀172.16.4.4/32 — 集群ID设置为192.168.2.2。

MCID和环路预防

如果路由器收到前缀的更新，则当集群列表包含路由器自己的集群ID时，该更新将被丢弃。如果使用MCID，包含任何已配置集群ID（全局或每个邻居）的更新将被丢弃。

相关信息

- RFC 4456
- BGP — 多个集群ID

关于此翻译
思科采用人工翻译与机器翻译相结合的方式将此文档翻译成不同语言，希望全球的用户都能通过各自的语言得到支持性的内容。

请注意：即使是最好的机器翻译，其准确度也不及专业翻译人员的水平。

Cisco Systems, Inc. 对于翻译的准确性不承担任何责任，并建议您总是参考英文原始文档（已提供链接）。