

Reflexão de Rota BGP e Várias IDs de Cluster

Contents

[Introduction](#)

[Descrição da Reflexão de Rota BGP](#)

[Exemplos de Configuração de Reflexão de Rota](#)

[Cluster único com configurações padrão](#)

[Cluster único com reflexão cliente a cliente desativada](#)

[Dois clusters, reflexão de rota entre sites e dentro do local](#)

[Dois clusters, sem reflexão cliente a cliente](#)

[Lista de clusters e prevenção de loop](#)

[Reflexão entre cliente e não cliente](#)

[Reflexão entre clusters](#)

[Reflexão entre clusters](#)

[MCIDs e prevenção de loop](#)

[Referências](#)

[Discussões relacionadas da comunidade de suporte da Cisco](#)

Introduction

Este artigo descreve diferentes cenários de reflexão de rota BGP (Border Gateway Protocol) e uso de várias IDs de cluster. Pressupõe-se o conhecimento prévio dos conceitos de BGP, especialmente clusters e reflexão de rota.

Descrição da Reflexão de Rota BGP

Um alto-falante BGP é um roteador ativado para BGP. Por padrão, os alto-falantes de BGP não anunciam os prefixos aprendidos de iBGP aos peers de iBGP - isso é feito para manter a prevenção de loop. O RFC4456 apresenta o recurso de reflexão de rota que elimina a necessidade de malha completa entre os alto-falantes iBGP. Quando o refletor de rota reflete um prefixo, ele cria/modifica um atributo não transitório opcional chamado CLUSTER_LIST adicionando seu próprio ID de cluster a ele. Este atributo é usado para prevenção de loop: quando o roteador recebe uma atualização que CLUSTER_LIST contém o próprio ID de cluster do roteador, essa atualização é descartada.

Por padrão, o ID do cluster é definido com o valor de ID do roteador BGP, mas pode ser definido como um valor arbitrário de 32 bits. O recurso MCID (Multiple Cluster IDs, várias IDs de cluster) permite atribuir IDs de cluster por vizinho. Portanto, há 3 tipos de cenários de reflexão de rota.

1. Entre cliente e não cliente
2. Entre clientes no mesmo cluster (intra-cluster)
3. Entre clientes em clusters diferentes (entre clusters)

Exemplos de Configuração de Reflexão de Rota

A seguir estão alguns cenários de reflexão do roteador e respectivos exemplos de configuração.

Cluster único com configurações padrão

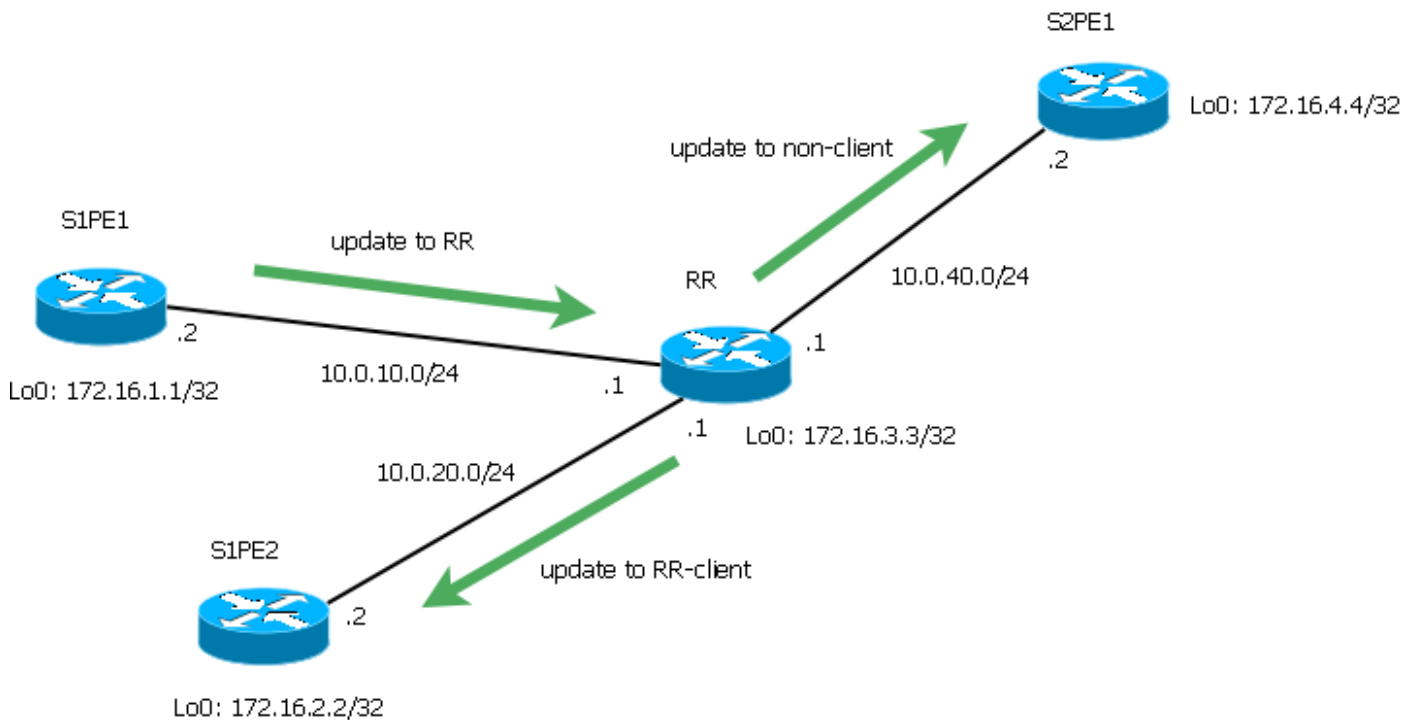


Figure 1

A seguinte configuração foi feita no roteador RR atuando como refletor de rota.

```
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
```

Nesse caso, S1PE1 e S1PE2 são clientes de RR enquanto S2PE1 não é cliente. Em designs convencionais, o roteador não cliente será refletor de rota para roteadores no próximo nível de hierarquia, mas neste exemplo, apenas outro PE é usado para simplicidade.

```
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
```

Essas saídas mostram que o RR recebe o prefixo 172.16.1.1/32 do S1PE1 e o reflete para o S1PE2 do cliente e para o S2PE1 não cliente. Nesse caso específico, a atualização também é enviada de volta para S1PE1, mas acontece porque S1PE1 e S1PE2 têm a mesma política de roteamento e, portanto, formam o mesmo grupo de atualização.

Cluster único com reflexão cliente a cliente desativada

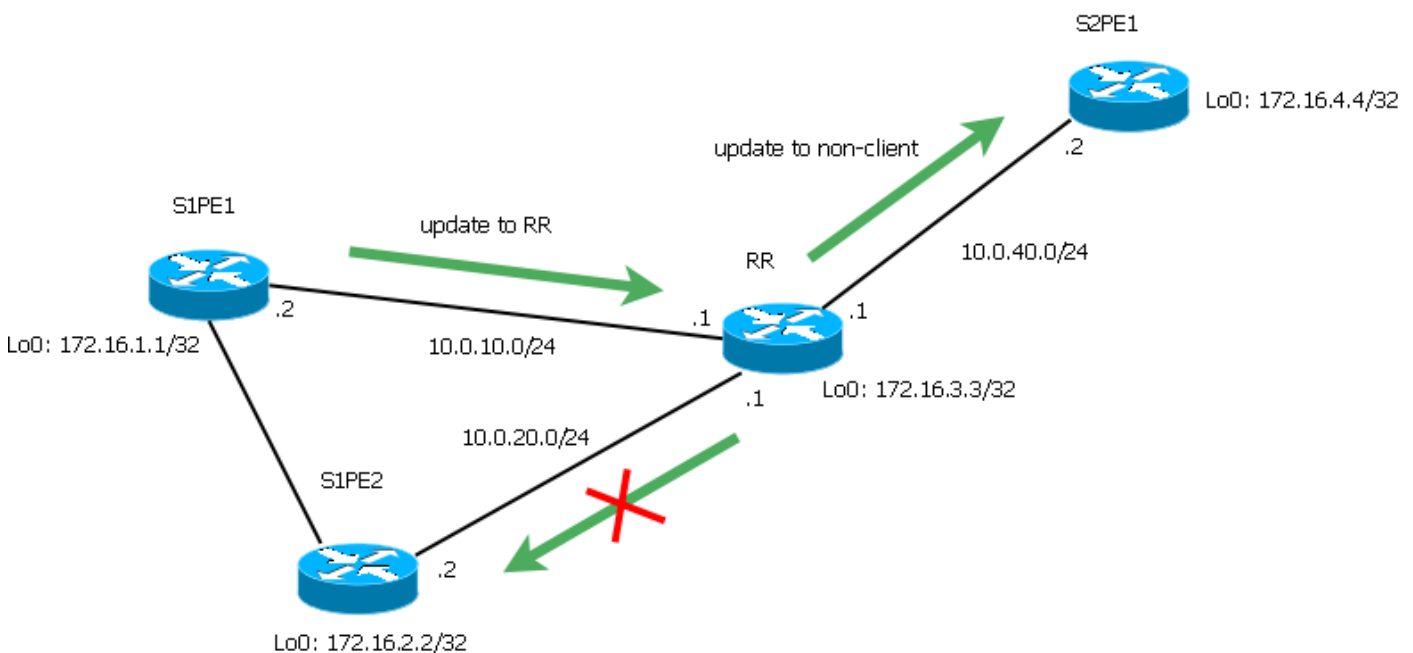


Figure 2

A seguinte configuração foi feita no roteador RR atuando como refletor de rota.

```
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
```

Vamos supor que AS1 esteja parcialmente em mesh: O S1PE1 e o S1PE2 formam a vizinhança do iBGP (por exemplo, eles estão localizados no mesmo site e queremos otimizar a forma como os processos de rede são atualizados). Nesse caso, o RR tem a reflexão cliente-cliente desativada e reflete 172.16.1.1/32 vindo do S1PE1 somente para o S2PE1 não-cliente.

```
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
```

Dois clusters, reflexão de rota entre sites e dentro do local

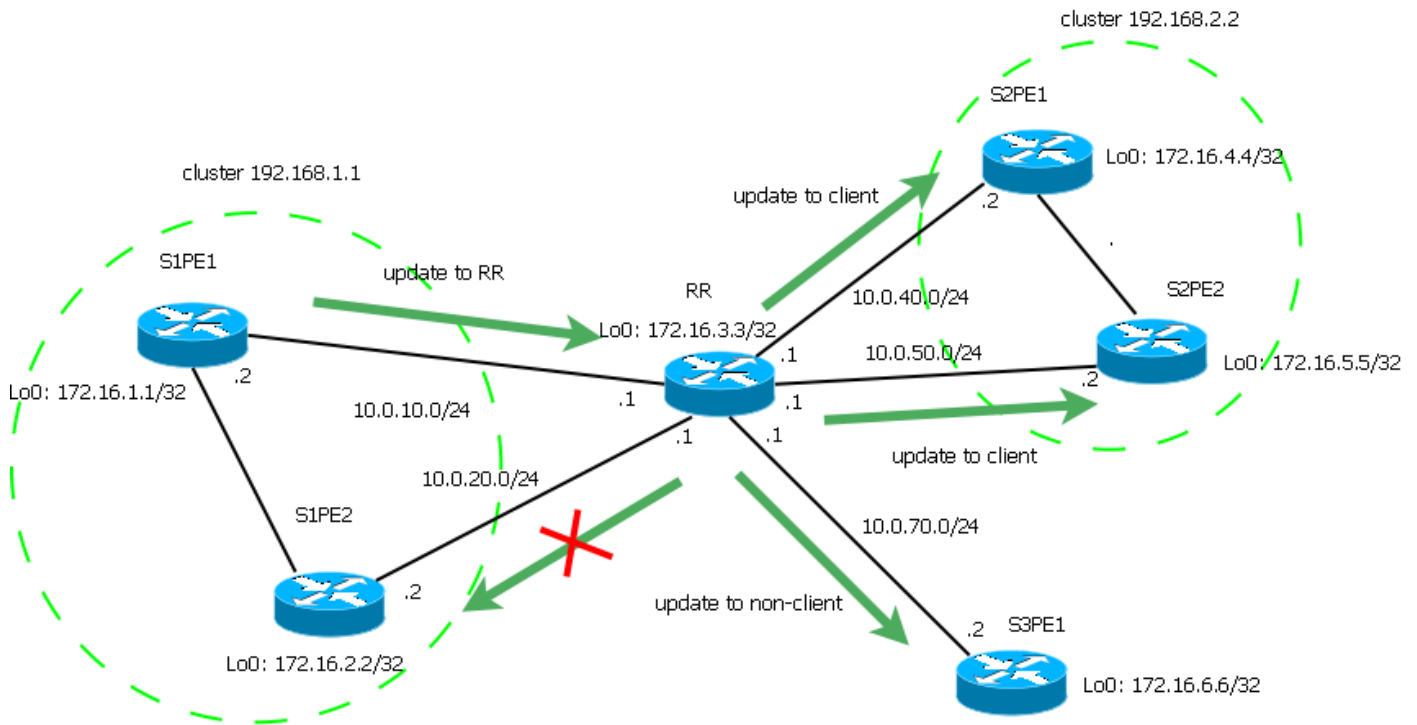


Figure 3

A seguinte configuração foi feita no roteador RR atuando como refletor de rota.

```
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
```

Nesse caso, ambos os PEs no Site 1 formam o cluster 192.168.1.1, enquanto ambos os PEs no Site 2 formam o cluster 192.168.2.2. O S3PE1 não é um cliente. Os PEs no Site 1 têm sessão iBGP direta, a reflexão intragrupo é desabilitada para o cluster 192.168.1.1, mas ainda está habilitada para o cluster 192.168.2.2. A reflexão entre clusters está habilitada.

```
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
```

O prefixo 172.16.1.1/32 recebido do S1PE1 é refletido para os clientes no cluster 192.168.2.2 e para os não-clientes. Ao mesmo tempo, o prefixo 172.16.4.4/32 recebido de S2PE1 é refletido para todos os clientes e não clientes.

```
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

Você também pode desativar a reflexão de rota entre sites para o cluster 192.168.2.2, mas, nesse caso, os clientes nesse cluster devem ter uma malha completa de sessões iBGP:

```
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
-------------	---	----------	-----------------

A reflexão dentro do local também pode ser desativada para todos os clusters:

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

Dois clusters, sem reflexão cliente a cliente

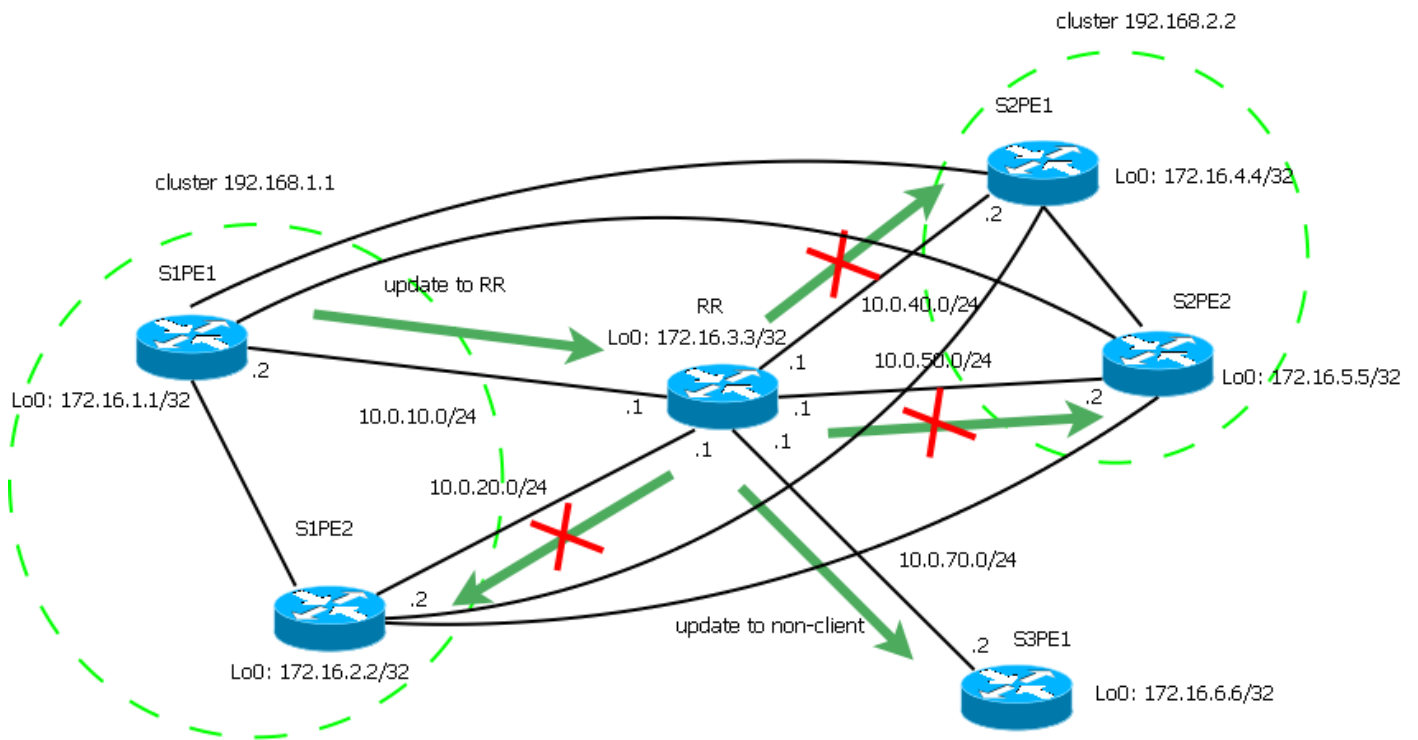


Figure 4

A seguinte configuração foi feita no roteador RR atuando como refletor de rota.

```
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
```

É possível desativar a reflexão entre clusters e entre clusters. Nesse caso, somente a reflexão entre clientes e não clientes será realizada.

```
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
```

Os prefixos 172.16.1.1/32 e 172.16.4.4/32 são originados pelos clusters 192.168.1.1 e 192.168.2.2, respectivamente. Ambos os prefixos são refletidos somente no S3PE1 não cliente. Nesse caso, todos os clientes devem estar totalmente em malha. Geralmente, nesse cenário específico, os MCIDs não fazem realmente sentido (o mesmo comportamento pode ser alcançado com um único cluster), mas ainda podem ser usados se você quiser ter listas de cluster diferentes para rotas de vizinhos diferentes.

Note: Não é possível ativar a reflexão entre clusters (para cluster específico ou para todos os clusters) enquanto a reflexão entre clusters está desativada.

Lista de clusters e prevenção de loop

Quando o RR reflete um prefixo, ele adiciona o ID do cluster ao atributo opcional não transitivo CLUSTER_LIST. Também define o atributo não transitivo opcional ORIGINATOR_ID para o ID do roteador do peer, que anunciou o prefixo para o RR.

Quando os MCIDs são usados e o RR reflete o prefixo, ele usa o ID de cluster configurado para o peer que anunciou esse prefixo ao RR. Se esse peer não tiver um ID de cluster específico configurado, será usado o ID de cluster global.

Vamos ver alguns exemplos. RR tem todas as formas de reflexão de rota ativadas. A ID global do cluster é 172.16.3.3, as IDs de cluster 192.168.1.1 e 192.168.2.2 estão definidas como PEs no

local 1 e no local 2, respectivamente (consulte o diagrama de topologia acima).

```
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
```

Reflexão entre cliente e não cliente

```
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
```

O S2PE3 não cliente recebe o prefixo 172.16.1.1/32 originado pelo cluster 192.168.1.1 - a ID do cluster 192.168.1.1 é adicionada à lista de clusters. Ele também recebe o prefixo 172.16.4.4/32 originado pelo cluster 192.168.2.2 - a ID do cluster 192.168.2.2 é adicionada à lista de clusters.

```
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
```

O cliente S1PE1 recebe o prefixo 172.16.6.6/32 originado por um não cliente - a ID de cluster global 172.16.3.3 é adicionada à lista de clusters.

Reflexão entre clusters

```
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
```

O S1PE2 pertence ao cluster 192.168.1.1 e recebe o prefixo 172.16.1.1/32 originado pelo S1PE1 que também pertence ao cluster 192.168.1.1. A ID do cluster 192.168.1.1 é adicionada à lista de clusters.

Reflexão entre clusters

```
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 pertence ao cluster 192.168.2.2 e recebe o prefixo 172.16.1.1/32 originado pelo cluster 192.168.1.1 - o ID do cluster é definido como 192.168.1.1.

O S1PE1 pertence ao cluster 192.168.1.1 e recebe o prefixo 172.16.4.4/32 originado pelo cluster 192.168.2.2 - o ID do cluster é definido como 192.168.2.2.

MCIDs e prevenção de loop

Se o roteador receber a atualização do prefixo em que lista de clusters contém o ID do cluster do próprio roteador, a atualização será descartada. Se os MCIDs forem usados, a atualização que contém qualquer um dos IDs de cluster configurados (global ou por vizinho) será descartada.

Referências

[RFC 4456](#)

[BGP-Várias IDs de Cluster](#)