

BGP ルート リフレクションと複数のクラスタ ID

内容

[概要](#)

[BGP ルート リフレクションの説明](#)

[ルート リフレクション構成例](#)

[デフォルト設定のシングル クラスタ](#)

[クライアント間リフレクションを無効にしたシングル クラスタ](#)

[2つのクラスタ \(内部サイトやサイト間のルート リフレクション\)](#)

[2つのクラスタ \(クライアント間リフレクションなし\)](#)

[クラスタ リストとループ防止](#)

[クライアントと非クライアント間のリフレクション](#)

[クラスタ内部のリフレクション](#)

[クラスタ間のリフレクション](#)

[MCID とループ防止](#)

[参考資料](#)

[関連するシスコ サポート コミュニティ ディスカッション](#)

概要

この記事では、Border Gateway Protocol (BGP) ルート リフレクションの各種シナリオと、複数のクラスタ ID の使用法について説明します。BGP 概念、特にクラスタとルート リフレクションの専門知識が前提になります。

BGP ルート リフレクションの説明

BGP スピーカーは BGP が有効化されたルータです。デフォルトでは、BGP スピーカーは iBGP に登録されたプレフィックスを iBGP ピアにアドバタイズしません。これはループ防止を維持するために行われます。RFC4456 では、iBGP スピーカー間のフル メッシュの必要性をなくすルート リフレクション機能が導入されています。ルート リフレクタがプレフィックスを反映すると、CLUSTER_LIST と呼ばれるオプションの非遷移属性を、独自のクラスタ ID を追加することにより作成または変更します。この属性は、ループ防止に使用されます。ルータがルータ自身のクラスタ ID を含む CLUSTER_LIST の更新を受信すると、この更新は廃棄されます。

デフォルトでは、クラスタ ID は BGP ルータ ID 値に設定されますが、任意の 32 ビット値に設定できます。複数のクラスタ ID (MCID) 機能により、隣接クラスタ ID を割り当てることができません。したがって、3 つのタイプのルート リフレクション シナリオがあります。

1. クライアントと非クライアント間
2. 同じクラスタ内のクライアント間 (クラスタ内部)
3. 異なるクラスタのクライアント間 (クラスタ間)

ルート リフレクション構成例

以下は、ルート リフレクション シナリオとそれぞれの設定例です。

デフォルト設定のシングル クラスタ

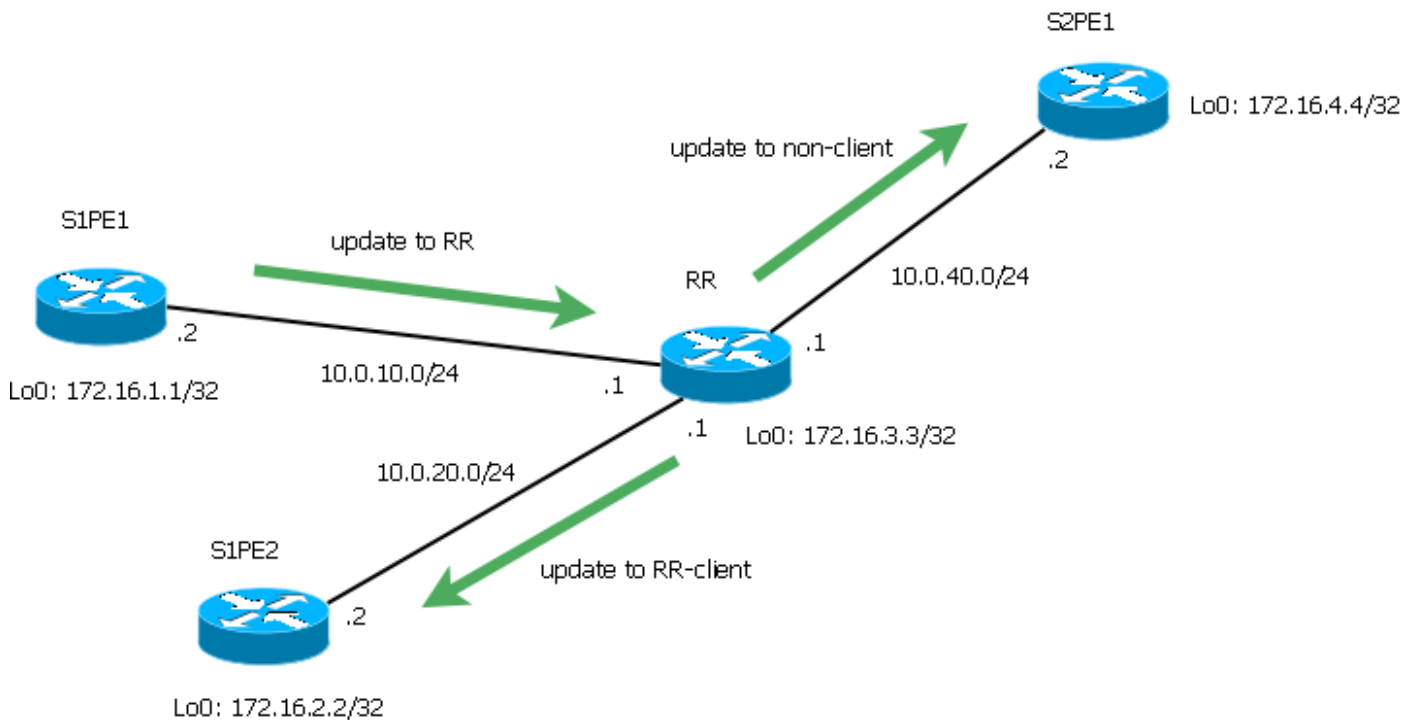


図 1 :

ルート リフレクタとして機能するルータ RR で次の設定が行われました。

```
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
```

この場合、S1PE1 と S1PE2 は RR のクライアントであり、S2PE1 は非クライアントです。従来の設計では、非クライアントのルータは次の階層レベルのルータのルート リフレクタになりますが、この例では簡単にするために別の PE を使用します。

```
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 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
```

これらの出力は RR が S1PE1 から 172.16.1.1/32 のプレフィックスを受け取り、クライアント S1PE2 および非クライアント S2PE1 に反映していることを示します。この特定のケースでは、更新も S1PE1 に戻されますが、これは、S1PE1 と S1PE2 が同じルーティングポリシーを持つために同じ更新グループを形成するゆえに発生します。

クライアント間リフレクションを無効にしたシングル クラスタ

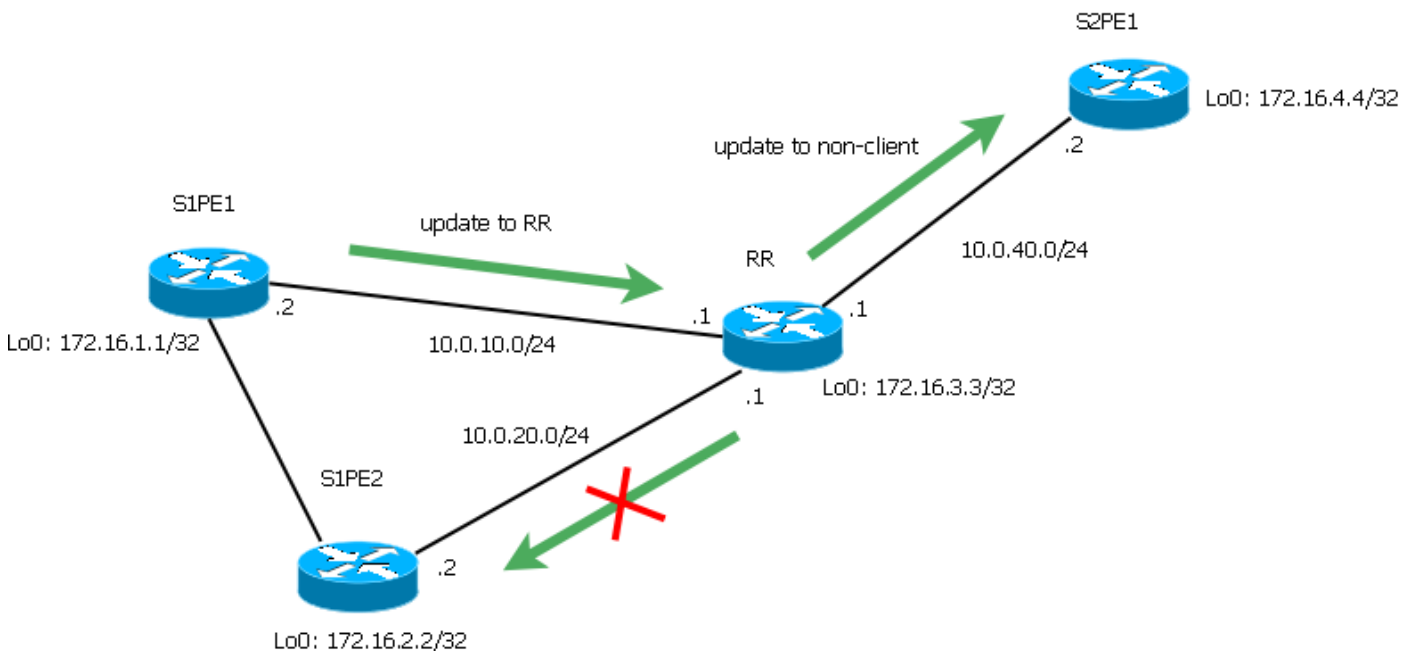


図 2

ルート リフレクタとして機能するルータ RR で次の設定が行われました。

```
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
```

AS1 が部分的にメッシュになっているとします。S1PE1 と S1PE2 が iBGP ネイバーシップを形成します (たとえば、それらが同じサイトにあり、ネットワークが更新を処理する方法を最適化したい)。この場合、RR はクライアント間リフレクションを無効にしており、S1PE1 から来る 172.16.1.1/32 を非クライアント S2PE1 にだけ反映します。

```
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
```

2 つのクラスタ (内部サイトやサイト間のルート リフレクション)

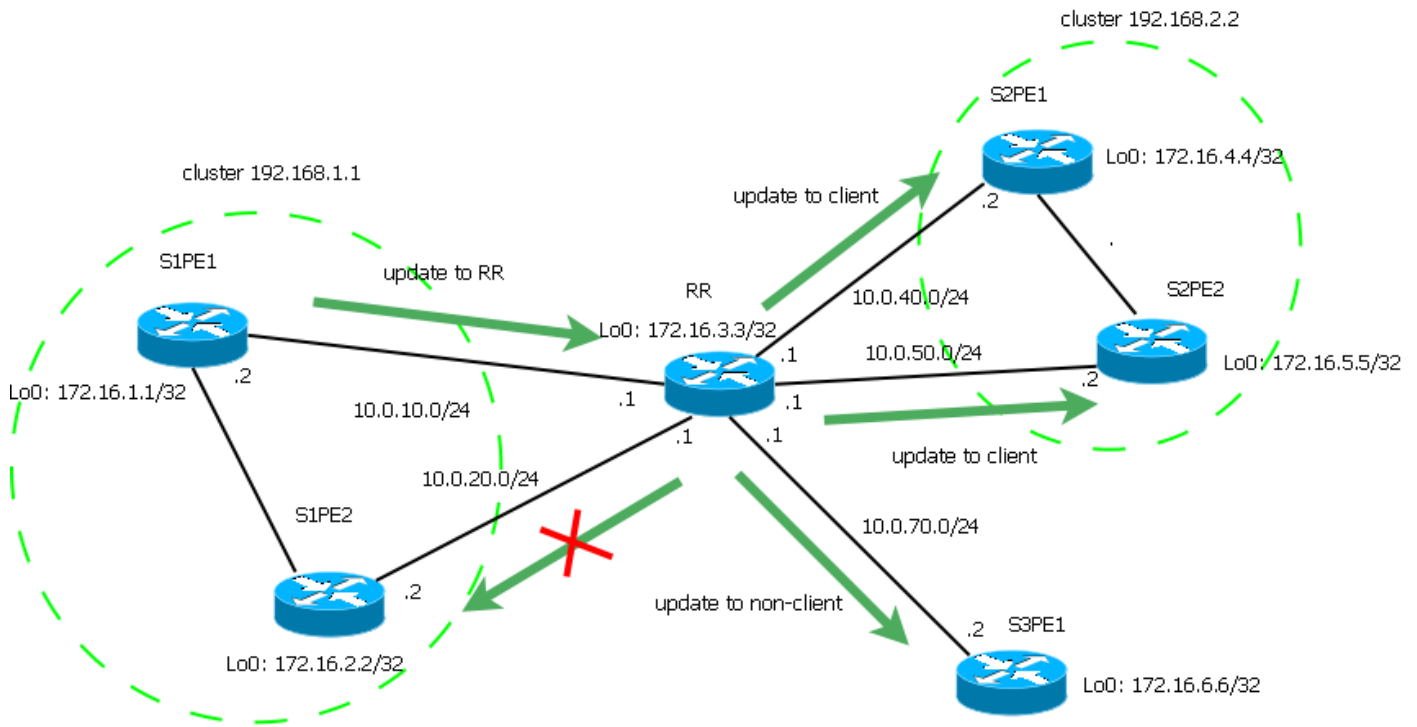


図 3

ルート リフレクタとして機能するルータ RR で次の設定が行われました。

```
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
```

この場合、Site 1 の両方の PE がクラスタ 192.168.1.1 を形成し、Site 2 の両方の PE がクラスタ 192.168.2.2 を形成します。S3PE1 は、非クライアントです。Site 1 の PE には直接 iBGP セッションがあり、クラスタ内部リフレクションはクラスタ 192.168.1.1 では無効になっていますが、クラスタ 192.168.2.2 ではまだ有効になっています。クラスタ間リフレクションが有効です。

```
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
```

S1PE1 から受信したプレフィックス 172.16.1.1/32 は、クラスター 192.168.2.2 のクライアントおよび非クライアントに反映されません。同時に、S2PE1 から受信したプレフィックス 172.16.4.4/32 は、すべてのクライアントおよび非クライアントに反映されます。

```
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

クラスタ 192.168.2.2 のサイト内部のルートリフレクションも無効にできますが、この場合、そのクラスタのクライアントには iBGP セッションのフル メッシュが必要です。

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**

サイト内部のリフレクションは、すべてのクラスタで無効にできます。

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

2 つのクラスタ (クライアント間リフレクションなし)

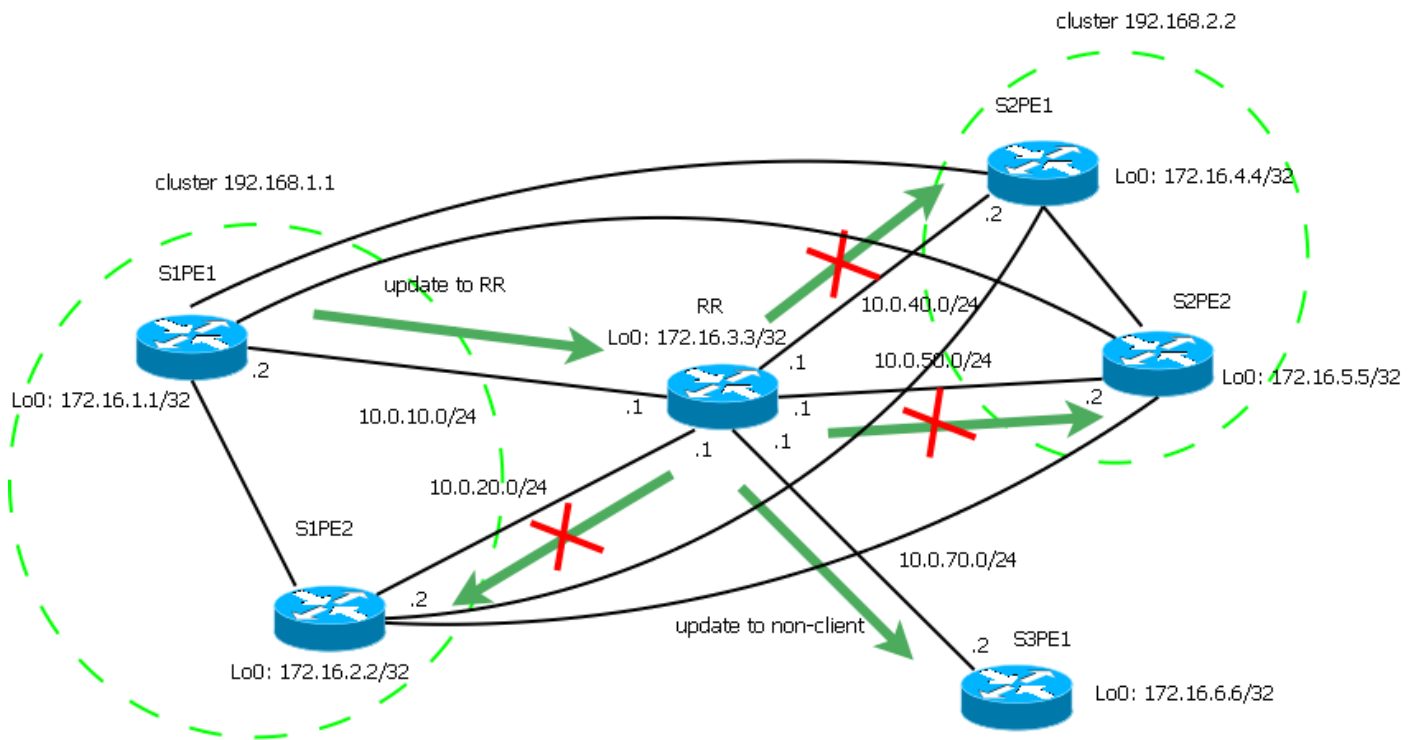


図 4

ルート リフレクタとして機能するルータ RR で次の設定が行われました。

```
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
```

クラスタ内部とクラスタ間のリフレクションの両方を無効にできます。この例では、クライアントと非クライアント間のリフレクションのみが実行されます。

```
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
```

プレフィックス 172.16.1.1/32 および 172.16.4.4/32 は、それぞれクラスター 192.168.1.1 および 192.168.2.2 によって生成されます。これらのプレフィックスはいずれも、非クライアント S3PE1 にのみ反映されます。この場合、すべてのクライアントはフル メッシュでなければなりません。通常、この特定のシナリオでは、MCID は実際には意味がありません（同じ動作はシングル クラスターで実現できます）。ただし、異なるネイバーからのルートに異なるクラスター リストを使用する場合にはまだ使用できます。

注： クラスター間リフレクションが無効になっている間は、クラスター内部リフレクションを（特定のクラスターに対してもすべてのクラスターに対しても）有効にすることはできません。

クラスター リストとループ防止

RR はプレフィックスを反映すると、オプションの非遷移属性 CLUSTER_LIST にクラスター ID を追加します。また、オプションの非遷移属性 ORIGINATOR_ID をピアのルータ ID に設定します。これは、RR にプレフィックスをアドバタイズします。

MCID が使用され RR がプレフィックスを反映する場合、RR はそのプレフィックスを自分にアドバタイズしたピアに構成されているクラスター ID を使用します。そのピアに特定のクラスター ID が設定されていない場合は、グローバル クラスター ID が使用されます。

いくつかの例を見てみましょう。RR では、すべての形式のルート リフレクションが有効になっています。グローバルクラスター ID は 172.16.3.3 であり、クラスター ID 192.168.1.1 および

192.168.2.2 がそれぞれサイト 1 およびサイト 2 の PE に設定されています (上記のトポロジ図参照)。

```
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
```

クライアントと非クライアント間のリフレクション

```
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
```

非クライアント S2PE3 は、クラスタ 192.168.1.1 からのプレフィックス 172.16.1.1/32 を受信します。クラスタ ID 192.168.1.1 がクラスタ リストに追加されます。また、クラスタ 192.168.2.2 からのプレフィックス 172.16.4.4/32 も受信します。クラスタ ID 192.168.2.2 がクラスタ リストに追加されます。

```
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
```

クライアント S1PE1 は、非クライアントからのプレフィックス 172.16.6.6/32 を受信します。グローバル クラスタ ID 172.16.3.3 がクラスタリストに追加されます。

クラスタ内部のリフレクション

```
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 はクラスタ 192.168.1.1 に属し、同じクラスタ 192.168.1.1 に属する S1PE1 からのプレフィックス 172.16.1.1/32 を受信します。クラスタ ID 192.168.1.1 がクラスタのリストに追加されます。

クラスタ間のリフレクション

```
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 はクラスタ 192.168.2.2 に属し、クラスタ 192.168.1.1 からのプレフィックス 172.16.1.1/32 を受信します。クラスタ ID は 192.168.1.1 に設定されます。

S1PE1 はクラスタ 192.168.1.1 に属し、クラスタ 192.168.2.2 からのプレフィックス 172.16.4.4/32 を受信します。クラスタ ID は 192.168.2.2 に設定されます。

MCID とループ防止

ルータがルータ自身のクラスタ ID を含むクラスタ リストがあるプレフィックスの更新を受信した場合、その更新は破棄されます。MCID を使用する場合、設定済みのクラスタ ID (グローバルまたはネイバーごと) のいずれかを含む更新は破棄されます。

参考資料

[RFC 4456](#)

[BGP-複数のクラスタ ID](#)