

ボーダーゲートウェイプロトコル(BGP)の最適なルートリフレクション

内容

[概要](#)

[背景説明](#)

[ネットワーク図](#)

[理論](#)

[IOS-XRの実装](#)

[設定](#)

[設定例](#)

[ルートルータでのMPLSトラフィックエンジニアリング](#)

[トラブルシューティング](#)

[結論](#)

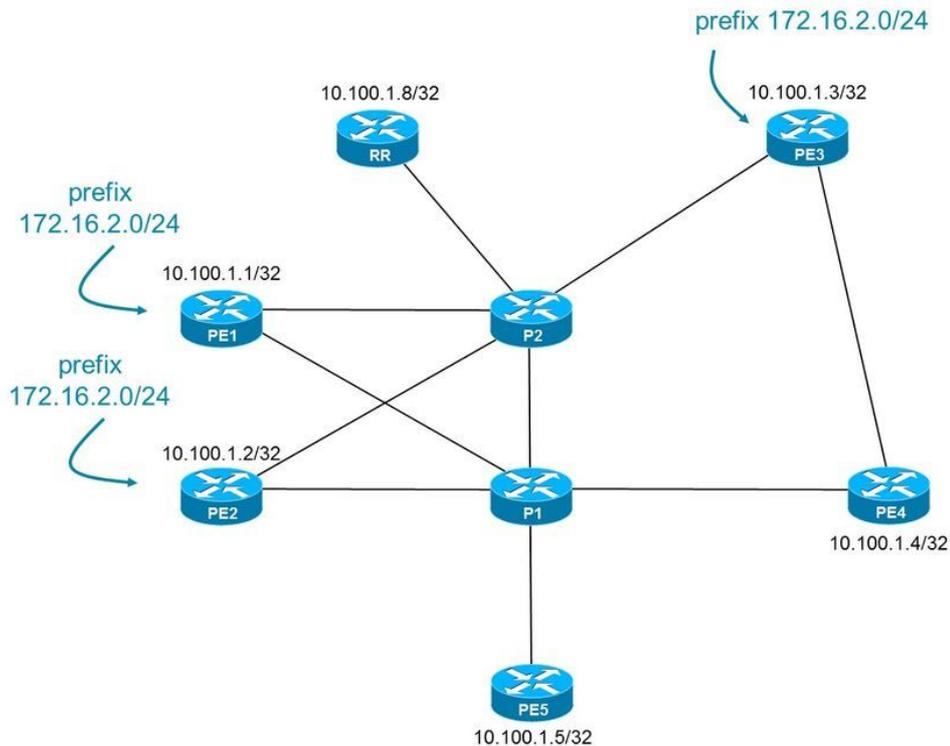
概要

このドキュメントでは、iBGPルータ間のフルメッシュを回避するために、ネットワーク内に1つ以上のルートリフレクタ(RR)が存在する場合に、ルーティングがどのように影響を受けるかについて説明します。

背景説明

[BGPベストパス選択アルゴリズムのステップ8](#)は、BGPネクストホップに最も低いIGPメトリックを持つパスを優先することです。したがって、ステップ8の前のすべてのステップが等しい場合、ステップ8がRR上のベストパスの決定要因になる可能性があります。次に、RRからアドバタイジングiBGPルータまでのIGPコストは、RRの配置によって決定されます。デフォルトでは、RRはクライアントへのベストパスのみをアドバタイズします。RRの配置場所によっては、アドバタイジングルータへのIGPコストが小さいか、大きくなることがあります。パスのすべてのIGPコストが同じ場合は、最も低いBGPルータIDを持つアドバタイジングルータのタイブレーカーが発生する可能性があります。

ネットワーク図



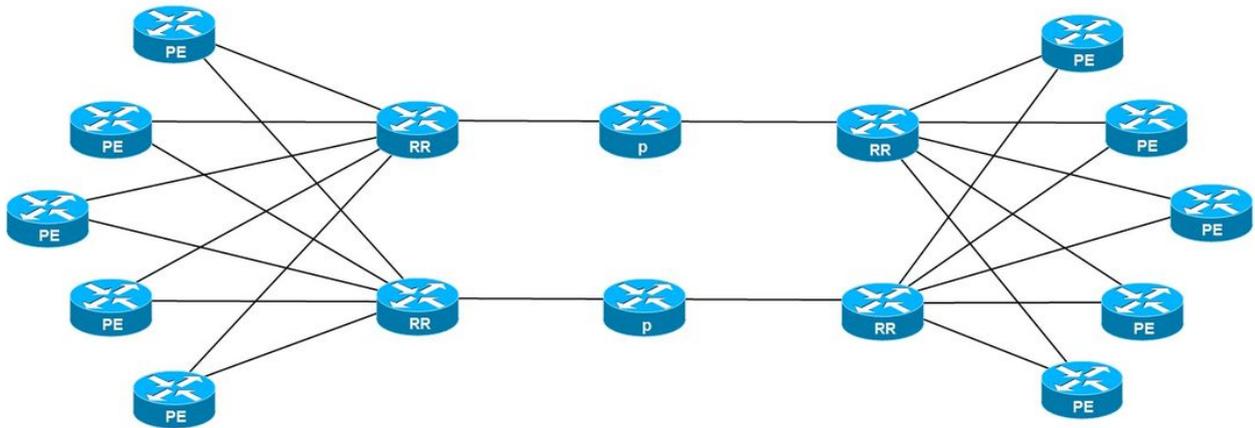
ルータPE1、PE2、およびPE3はプレフィックス172.16.2.0/24をアドバタイズします。リンクのすべてのIGPコストが同じ場合、RRはIGPコスト2のPE1、PE2、およびPE3からのパスを参照しますBGPルータIDが低い。これは、BGPベストパス選択アルゴリズムのステップ11です。その結果、PE4を含むすべてのPEルータは、プレフィックス172.16.2.0/24の出力PEルータとしてPE1を選択します。PE4の観点から、出力PEルータへの短いIGPパスは、IGPコスト1のPE3へのパスです多くのネットワークでは、可能な限り最短の方法でトラフィックを中継ネットワーク経由で転送するという事実が重要です。これはホットポテトルーティングと呼ばれます。

RRがPE1から最適なパスを選択する別の理由が考えられます。図では、リンクP2-PE3のInterior Gateway Protocol(IGP)コストが10であり、他のすべてのリンクのIGPコストが1のみである場合、RRはPEから3からのパスを選択します最も低いBGPルータID。

このネットワークの管理者がホットポテトルーティングを使用する場合は、ネットワークにRRがある場合でも、入力ルータがiBGPネットワーク内で最も近い出力ルータへのパスを学習できるように、メカニズムを設定する必要があります。BGP機能Add Pathはこれを実現できます。ただし、この機能では、RRと境界ルータに、その機能を理解する最新のコードが必要です。BGP Optimal Route Reflectionの機能では、これは必須ではありません。この機能を使用すると、RRが入力BGPルータから見たベストパスを基に、入力BGPポータルルータにベストパスを送信できます。

RRの導入時にホットポテトルーティングを可能にするもう1つのソリューションは、RRのインライン配置です。これらのRRは専用RRではなく、BGPとIGPのみを実行します。これらのインラインRRは転送パスに配置され、独自のRRクライアントセットを持つようにネットワークに配置されます。これにより、各RRクライアントへのベストパスも反映されます。これは、RRクライアントの観点からのベストパスです。

次の図に示すように、RRはネットワークに配置され、近くのRRクライアントの小さなセットを提供できます。ネットワーク設計により、RRクライアントは、ネットワーク内でホットポテトルーティングが可能になるように、RRからRRの観点からベストパスであるベストパスを受信します。



理論

BGPの最適ルートの反映については、IETFドラフトdraft-ietf-idr-bgp-optimal-route-reflectionを参照してください。

BGP Optimal Route Reflection (FRR ; 最適ルート反射) ソリューションを使用すると、RRは特定のBGPボーダルータに特定のベストパスを送信できます。RRは、異なるBGP境界ルータまたは境界ルータのセットに異なるベストパスを送信することを選択できます。境界ルータは、RRのRRクライアントである必要があります。目標は、各入力BGP境界ルータが、同じプレフィクスに対して異なる出力BGPルータまたは出力BGPルータを持つことです。入力ボーダルータが常にトラフィックを閉じたAS-exitルータに転送できる場合、これはホットポテトルーティングを可能にします。

問題は、RRが通常、各BGPボーダルータに同じベストパスを送信するだけで、ホットポテトルーティングが防止されることです。これを解決するには、RRが入力BGPボーダルータに応じて同じプレフィクスに対して異なるベストパスを計算できる必要があります。RRでのベストパスの計算は、入力BGPボーダルータの位置に基づいて行われます。したがって、RRは、入力ボーダルータの観点からBGPベストパスの計算を実行します。これを実行できるRRは、RRと入力ボーダルータが配置されているIGPの観点から、ネットワークのトポロジの全体像を把握したRRです。この要件を満たすには、IGPがリンクステートルーティングプロトコルである必要があります。

この場合、RRは入力境界ルータをツリーのルートとして使用してShortest Path First(SPF)計算を実行し、他のすべてのルータへのコストを計算できます。これにより、入力ボーダルータから他のすべての出力ボーダルータまでのコストが分かります。別のルータをルートとするこの特別なSPF計算は、リバースSPF(rSPF)と呼ばれます。これは、RRがすべてのBGPボーダルータからすべてのBGPパスを学習した場合にのみ実行できます。RRクライアントと同じ数のrSPFが実行されている可能性があります。これにより、RRのCPU負荷が多少増加します。

このソリューションでは、BGPベストパス選択アルゴリズムに基づいてベストパスを計算できます。これにより、RRがパスを送信する入力境界ルータの観点からベストパスを選択することになります。これは、BGPネクストホップへの最短のIGPコストに基づいてベストパスが選択される

ことを意味します。このソリューションでは、設定されたポリシーに基づいてベストパスを選択することもできます。入力ポータルータは、最も低いIGPコストではなく、設定されたポリシーに基づいてベストパスを選択できます。このソリューションにより、RRはIGPコスト（ネットワーク上の場所）または設定されたポリシーのいずれかまたは両方に最適なルートリフレクションを実装できます。両方が展開されている場合、最初にポリシーが適用され、その後、残りのパスでIGPベースの最適なルートリフレクションが発生します。

IOS-XRの実装

IOS-XRの実装では、rSPFの計算に最大3つのルートノードを使用できます。1つのアップデートグループに多数のRRクライアントがある場合、それらのRRクライアントが異なる出力BGPポータルータに対して同じポリシーまたは同じIGPコストを持っている場合、RRクライアントごとに1つのrSPF計算は不要です。後者の場合、通常はRRクライアントが同じPOPに配置されている（同じPOP内にある可能性が高い）ことを意味します。その場合、各RRクライアントをルートとして設定する必要はありません。IOS-XRの実装では、冗長性を確保するために、RRクライアントのセットごとに、プライマリ、セカンダリ、およびターシャリルートの3つを設定できます。BGP OR機能を任意のRRクライアントに適用するには、そのRRクライアントをORポリシーグループの一部として設定する必要があります。

BGP OR機能は、アドレスファミリごとに有効になります。

リンクステートプロトコルが必要です。OSPFまたはIS-ISの可能性もあります。

IOS XRは、BGPネクストホップへのIGPコストに基づいてBGP OR機能を実装するだけであり、一部の設定済みポリシーに基づいたものではありません。

同じアウトバウンドポリシーを持つBGPピアは、同じアップデートグループに配置されます。通常、これはRRでのiBGPの場合です。機能BGP ORが有効な場合、異なるORグループからのピアは異なるアップデートグループに含まれます。これは論理的です。RRから異なるBGP ORグループのRRクライアントに送信されるアップデートは、BGPのベストパスが異なるので、異なります。

rSPFの計算結果はデータベースに保存されます。

ORSPFは、BGP OR機能に必要なIOS-XRの新しいコンポーネントです。ORSPFは次の点を考慮します。

1. リンクステート情報の収集とリンクステートデータベースの維持
2. rSPFを実行し、ポリシーグループごとにSPTを維持
3. メトリックを使用してSPTからRIBにプレフィックスをダウンロードする

データベースは、リンクステート情報をリンクステートIGPまたはBGP-LSから直接取得します。

rSPFの計算では、RRクライアントからエリア/レベルの他のルータへの最短パスを示すトポロジが作成されます。

トポロジ内のすべてのルータでハングするルートは、ORRグループポリシーの特別なRIBテーブルとAFI/SAFIごとに保存されます。このテーブルはRSIによって作成されます。このテーブルには、プライマリルートをルートとしてrSPFが計算したルートが入力されます。プライマリルートが使用不能になると、セカンダリルートがルートになり、OR RIBテーブルにルートが入力されます。同じことが三次ルートにも当てはまります。

設定

必要な最小限の設定：

1. BGPネイバーの特定のグループに対して、BGPのアドレスファミリに対してORを有効にする必要があります
2. BGPネイバーのグループごとに、少なくとも1つのルートを設定する必要があります。オプションで、セカンダリルートとターシャリルートを設定できます。
3. IGPからBGPへのORルートの再配布を有効にする必要があります。

設定例

最初の図に示すように、RRはBGP OR機能を備えたIOS-XRルータです。

他のすべてのルータはIOSを実行しています。これらのルータにはBGP OR機能はありません。

PE1、PE2、およびPE3は、AFI/SAFI 1/1 (IPv4ユニキャスト) のプレフィクス172.16.2.0/24をアドバタイズします。RRはPE1とPE2にPE3と同じように近づいています。すべてのリンクのIGPコストは1です。このプレフィクスのベストパスは、R1をネクストホップとして持つパスで、BGPルータIDが最小です。

```
RP/0/0/CPU0:RR#show bgp ipv4 unicast 172.16.2.0/24 bestpath-compare
BGP routing table entry for 172.16.2.0/24
Versions:
  Process          bRIB/RIB   SendTblVer
  Speaker          34         34
Last Modified: Mar  7 20:29:48.156 for 11:36:44
Paths: (3 available, best #1)
  Advertised to update-groups (with more than one peer):
    0.3
  Path #1: Received by speaker 0
  Advertised to update-groups (with more than one peer):
    0.3
  Local, (Received from a RR-client)
    10.100.1.1 (metric 3) from 10.100.1.1 (10.100.1.1)
      Origin IGP, metric 0, localpref 100, valid, internal, best, group-best
      Received Path ID 0, Local Path ID 1, version 34
      best of local AS, Overall best
  Path #2: Received by speaker 0
  Not advertised to any peer
  Local, (Received from a RR-client)
    10.100.1.2 (metric 3) from 10.100.1.2 (10.100.1.2)
      Origin IGP, metric 0, localpref 100, valid, internal, add-path
      Received Path ID 0, Local Path ID 6, version 33
      Higher router ID than best path (path #1)
  Path #3: Received by speaker 0
  ORR bestpath for update-groups (with more than one peer):
    0.1
  Local, (Received from a RR-client)
    10.100.1.3 (metric 5) from 10.100.1.3 (10.100.1.3)
      Origin IGP, metric 0, localpref 100, valid, internal, add-path
      Received Path ID 0, Local Path ID 7, version 34
      Higher IGP metric than best path (path #1)
```

PE4は、ネクストホップとしてPE1を持つパスを受信します。そのため、PE4のホットポテトルーティングはありません。

PE4にホットポテトルーティングを設定する場合は、PE1、PE2、およびPE3によってアドバタイズされたプレフィックス(プレフィックス172.16.2.0/24など)に対して、PE1は出力点としてPE3を持つ必要があります。つまり、PE4のパスは、ネクストホップとしてPE3を持つパスになります。このOR設定を使用して、RRからネクストホップPE3のルートをPE4に送信できます。

```
router ospf 1
distribute bgp-1s
  area 0
  interface Loopback0
  !
  interface GigabitEthernet0/0/0/0
    network point-to-point
  !
  !
  !
!

router bgp 1
  address-family ipv4 unicast
  optimal-route-reflection ipv4-orr-group 10.100.1.4
  !
  address-family vpnv4 unicast
  !
  neighbor 10.100.1.1
  remote-as 1
  update-source Loopback0
  address-family ipv4 unicast
  route-reflector-client
  !
  !
  neighbor 10.100.1.2
  remote-as 1
  update-source Loopback0
  address-family ipv4 unicast
  route-reflector-client
  !
  !
  neighbor 10.100.1.3
  remote-as 1
  update-source Loopback0
  address-family ipv4 unicast
  route-reflector-client
  !
  !
  neighbor 10.100.1.4
  remote-as 1
  update-source Loopback0
  address-family ipv4 unicast
  optimal-route-reflection ipv4-orr-group
  route-reflector-client
  !
  !
  neighbor 10.100.1.5
  remote-as 1
  update-source Loopback0
  address-family ipv4 unicast
  route-reflector-client
  !
  !
!
```

IGPがIS-ISの場合：

```

router isis 1
net 49.0001.0000.0000.0008.00
  distribute bgp-ls
address-family ipv4 unicast
metric-style wide
!
interface Loopback0
address-family ipv4 unicast
!
!
interface GigabitEthernet0/0/0/0
address-family ipv4 unicast
!
!
!

```

注：アドレスファミリリンクステートは、グローバルまたはBGPネイバーの下で設定する必要はありません。

ルートルータでのMPLSトラフィックエンジニアリング

rSPFを実行するには、RRがIGPデータベースで設定済みのルートアドレスを見つける必要があります。ISISでは、ルータIDがISISデータベースに存在します。OSPFの場合、OSPF LSAにはルータIDがありません。解決策は、ルートルータにRR上で設定されたルートアドレスに一致するマルチプロトコルラベルスイッチング(MPLS)TEルータIDをアドバタイズさせることです。

OSPFの場合、BGP ORを機能させるには、ルートルータに追加の設定が必要です。このMPLS TEルータIDをアドバタイズするには、すべてのルートルータで最小限のMPLS TE設定が必要です。コマンドの正確な最小セットは、ルートルータのオペレーティングシステムによって異なります。ルートルータのMPLS TE設定では、OSPFが不透明エリアLSA (タイプ10) でMPLS TEルータIDをアドバタイズするように、MPLS TEの最小限の設定を有効にする必要があります。

RRに、設定されたルートルータアドレスに一致するMPLS TEルータIDを持つ不透明エリアLSAが存在すると、rSPFが実行され、RR上のBGPが最適なルートをアドバタイズできます。

IOSルータの場合、ルートルータのOSPFに必要な最小限の設定は次のとおりです。

```

!
interface GigabitEthernet0/2
 ip address 10.1.34.4 255.255.255.0
 ip ospf network point-to-point
mpls traffic-eng tunnels
!

router ospf 1
mpls traffic-eng router-id Loopback0
  mpls traffic-eng area 0
  router-id 10.200.1.155
  network 10.0.0.0 0.255.255.255 area 0
!

```

次の点に注意してください。

- 特定のOSPFエリアでMPLS TEが有効になっている
- MPLS TEルータIDは、RR上の設定済みルートアドレスと一致するように設定されています
- MPLS TEが少なくとも1つのインターフェイスに設定されている

- RSVP-TEを設定する必要はありません
- エリア内の他のルータにMPLS TEを設定する必要はありません

IOS-XRルータの場合、ルートルータのOSPFに必要な最小限の設定は次のとおりです。

```
!
router ospf 1
router-id 5.6.7.8
area 0
mpls traffic-eng
interface Loopback0
!
interface GigabitEthernet0/0/0/0
network point-to-point
!
!
mpls traffic-eng router-id 10.100.1.11
!
mpls traffic-eng
!
```

上記の設定がルートルータ上にある場合、RRはOSPFデータベース内にMPLS TEルータIDを持つ必要があります。

```
RP/0/0/CPU0:RR#show ospf 1 database
```

```
OSPF Router with ID (10.100.1.99) (Process ID 1)
```

```
Router Link States (Area 0)
```

Link ID	ADV Router	Age	Seq#	Checksum	Link count
10.1.12.1	10.1.12.1	1297	0x8000002b	0x006145	3
10.100.1.2	10.100.1.2	646	0x80000025	0x00fb6f	7
10.100.1.3	10.100.1.3	1693	0x80000031	0x003ba9	5
10.100.1.99	10.100.1.99	623	0x8000001e	0x00ade1	3
10.200.1.155	10.200.1.155	28	0x80000002	0x009b2e	5

```
Type-10 Opaque Link Area Link States (Area 0)
```

Link ID	ADV Router	Age	Seq#	Checksum	Opaque ID
1.0.0.0	10.200.1.155	34	0x80000001	0x00a1ad	0
1.0.0.3	10.200.1.155	34	0x80000001	0x0057ff	3

```
RP/0/0/CPU0:RR#show ospf 1 database opaque-area adv-router 10.200.1.155
```

```
OSPF Router with ID (10.100.1.99) (Process ID 1)
```

```
Type-10 Opaque Link Area Link States (Area 0)
```

```
LS age: 184
Options: (No TOS-capability, DC)
LS Type: Opaque Area Link
Link State ID: 1.0.0.0
Opaque Type: 1
Opaque ID: 0
Advertising Router: 10.200.1.155
LS Seq Number: 80000001
Checksum: 0xa1ad
```

Length: 28

MPLS TE router ID : 10.100.1.4

Number of Links : 0

LS age: 184

Options: (No TOS-capability, DC)

LS Type: Opaque Area Link

Link State ID: 1.0.0.3

Opaque Type: 1

Opaque ID: 3

Advertising Router: 10.200.1.155

LS Seq Number: 80000001

Checksum: 0x57ff

Length: 132

Link connected to Point-to-Point network

Link ID : 10.100.1.3 (all bandwidths in bytes/sec)

Interface Address : 10.1.34.4

Neighbor Address : 10.1.34.3

Admin Metric : 1

Maximum bandwidth : 125000000

Maximum reservable bandwidth global: 0

Number of Priority : 8

Priority 0 : 0 Priority 1 : 0

Priority 2 : 0 Priority 3 : 0

Priority 4 : 0 Priority 5 : 0

Priority 6 : 0 Priority 7 : 0

Affinity Bit : 0

IGP Metric : 1

Number of Links : 1

MPLS TEルーターID(10.100.1.4)とOSPFルーターIDが異なることに注意してください。

PE4には、プレフィックスのネクストホップとしてPE3があります (ネクストホップへの正しいIGPメトリック)。

PE4#**show bgp ipv4 unicast 172.16.2.0**

BGP routing table entry for 172.16.2.0/24, version 37

Paths: (1 available, best #1, table default)

Not advertised to any peer

Refresh Epoch 1

Local

10.100.1.3 (metric 2) from 10.100.1.8 (10.100.1.8)

Origin IGP, metric 0, localpref 100, valid, internal, best

Originator: 10.100.1.3, Cluster list: 10.100.1.8

rx pathid: 0, tx pathid: 0x0

PE5には、プレフィックスのネクストホップとしてPE1が引き続き存在します (ネクストホップへの正しいIGPメトリックを使用)。

PE5#**show bgp ipv4 unicast 172.16.2.0/24**

BGP routing table entry for 172.16.2.0/24, version 13

Paths: (1 available, best #1, table default)

Not advertised to any peer

Refresh Epoch 1

Local

10.100.1.1 (metric 3) from 10.100.1.8 (10.100.1.8)

Origin IGP, metric 0, localpref 100, valid, internal, best

Originator: 10.100.1.1, Cluster list: 10.100.1.8
rx pathid: 0, tx pathid: 0x0

トラブルシュート

RRのプレフィクスを確認します。

```
RP/0/0/CPU0:RR#show bgp ipv4 unicast 172.16.2.0
BGP routing table entry for 172.16.2.0/24
Versions:
  Process          bRIB/RIB   SendTblVer
  Speaker          19         19
Last Modified: Mar  7 16:41:20.156 for 03:07:40
Paths: (3 available, best #1)
  Advertised to update-groups (with more than one peer):
    0.3
  Path #1: Received by speaker 0
  Advertised to update-groups (with more than one peer):
    0.3
  Local, (Received from a RR-client)
    10.100.1.1 (metric 3) from 10.100.1.1 (10.100.1.1)
      Origin IGP, metric 0, localpref 100, valid, internal, best, group-best
      Received Path ID 0, Local Path ID 1, version 14
  Path #2: Received by speaker 0
  Not advertised to any peer
  Local, (Received from a RR-client)
    10.100.1.2 (metric 3) from 10.100.1.2 (10.100.1.2)
      Origin IGP, metric 0, localpref 100, valid, internal, add-path
      Received Path ID 0, Local Path ID 4, version 14
  Path #3: Received by speaker 0
  ORR bestpath for update-groups (with more than one peer):
    0.1
  Local, (Received from a RR-client)
    10.100.1.3 (metric 5) from 10.100.1.3 (10.100.1.3)
      Origin IGP, metric 0, localpref 100, valid, internal, add-path
      Received Path ID 0, Local Path ID 5, version 19
```

他のベストパスにadd-pathが追加され、ベストパス以外にもアドバタイズできることに注意してください。RRとそのクライアントの間では、add path機能は使用されません。パスはパス識別子でアドバタイズされません。

ルートが特定のBGPネイバーに (まだ) アドバタイズされていることを確認します。

ネイバーPE4へのネクストホップは、プレフィックス172.16.2.0/24のPE3です。

```
RP/0/0/CPU0:RR#show bgp ipv4 unicast neighbors 10.100.1.4 advertised-routes
Network          Next Hop          From              AS Path
172.16.1.0/24    10.100.1.5       10.100.1.5       i
172.16.2.0/24    10.100.1.3      10.100.1.3       i
```

Processed 2 prefixes, 2 paths

ネイバーPE5へのネクストホップは、プレフィックス172.16.2.0/24のPE1です。

```
RP/0/0/CPU0:RR#show bgp ipv4 unicast neighbors 10.100.1.5 advertised-routes
Network          Next Hop          From              AS Path
```

```
172.16.1.0/24      10.100.1.8      10.100.1.5      i
172.16.2.0/24      10.100.1.1      10.100.1.1      i
```

ネイバー10.100.1.4は、ORポリシーが設定されているため、自身のアップデートグループにあります。

```
RP/0/0/CPU0:RR#show bgp ipv4 unicast update-group
```

```
Update group for IPv4 Unicast, index 0.1:
```

```
Attributes:
```

```
Neighbor sessions are IPv4
Internal
Common admin
First neighbor AS: 1
Send communities
Send GSHUT community if originated
Send extended communities
Route Reflector Client
ORR root (configured): ipv4-orr-group; Index: 0
4-byte AS capable
Non-labeled address-family capable
Send AIGP
Send multicast attributes
Minimum advertisement interval: 0 secs
```

```
Update group desynchronized: 0
```

```
Sub-groups merged: 0
```

```
Number of refresh subgroups: 0
```

```
Messages formatted: 8, replicated: 8
```

```
All neighbors are assigned to sub-group(s)
```

```
Neighbors in sub-group: 0.1, Filter-Groups num:1
```

```
Neighbors in filter-group: 0.3(RT num: 0)
```

```
10.100.1.4
```

```
Update group for IPv4 Unicast, index 0.3:
```

```
Attributes:
```

```
Neighbor sessions are IPv4
Internal
Common admin
First neighbor AS: 1
Send communities
Send GSHUT community if originated
Send extended communities
Route Reflector Client
4-byte AS capable
Non-labeled address-family capable
Send AIGP
Send multicast attributes
Minimum advertisement interval: 0 secs
```

```
Update group desynchronized: 0
```

```
Sub-groups merged: 1
```

```
Number of refresh subgroups: 0
```

```
Messages formatted: 12, replicated: 42
```

```
All neighbors are assigned to sub-group(s)
```

```
Neighbors in sub-group: 0.3, Filter-Groups num:1
```

```
Neighbors in filter-group: 0.1(RT num: 0)
```

```
10.100.1.1      10.100.1.2      10.100.1.3
```

```
10.100.1.5
```

show orspf databaseコマンドは、ORグループとそのルートを表示します。

```
RP/0/0/CPU0:RR#show orrspf database
```

```
ORR policy: ipv4-orr-group, IPv4, RIB tableid: 0xe0000012
```

Configured root: primary: 10.100.1.4, secondary: NULL, tertiary: NULL
Actual Root: 10.100.1.4

Number of mapping entries: 1

同じコマンドにdetailキーワードを指定すると、同じOSPFエリア内の他のルータ/プレフィクスにrSPFのルートのコストが提供されます。

RP/0/0/CPU0:RR#show orrspf database detail

ORR policy: ipv4-orr-group, IPv4, RIB tableid: 0xe0000012
Configured root: primary: 10.100.1.4, secondary: NULL, tertiary: NULL
Actual Root: 10.100.1.4

Prefix	Cost
10.100.1.6	2
10.100.1.1	3
10.100.1.2	3
10.100.1.3	2
10.100.1.4	0
10.100.1.5	3
10.100.1.7	3
10.100.1.8	4

Number of mapping entries: 9

テーブルIDは、RSIによってORグループとAFI/SAFIに割り当てられました。

RP/0/0/CPU0:RR#show rsi table-id 0xe0000012

TBL_NAME=ipv4-orr-group, AFI=IPv4, SAFI=Ucast TBL_ID=0xe0000012 in VRF=default/0x60000000 in VR=default/0x20000000
Refcnt=1
VRF Index=4 TCM Index=1
Flags=0x0 LST Flags=(0x0) NULL

RP/0/0/CPU0:RR#show rib tables

Codes: N - Prefix Limit Notified, F - Forward Referenced
D - Table Deleted, C - Table Reached Convergence

VRF/Table	SAFI	Table ID	PrfxLmt	PrfxCnt	TblVersion	N	F	D	C
default/default	uni	0xe0000000	5000000	22	128	N	N	N	Y
**nVSatellite/default	uni	0xe0000010	5000000	2	4	N	N	N	Y
default/ipv4-orr-grou	uni	0xe0000012	5000000	9	27	N	N	N	Y
default/default	multi	0xe0100000	5000000	0	0	N	N	N	Y

rSPFのルート(R4/10.100.1.4)から他のルータへのコストは、PE4でshow ip route ospfで表示されるコストと同じです。

PE4#show ip route ospf

Codes: L - local, C - connected, S - static, R - RIP, M - mobile, B - BGP
D - EIGRP, EX - EIGRP external, O - OSPF, IA - OSPF inter area
N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
E1 - OSPF external type 1, E2 - OSPF external type 2
i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
ia - IS-IS inter area, * - candidate default, U - per-user static route
o - ODR, P - periodic downloaded static route, H - NHRP, l - LISP
a - application route
+ - replicated route, % - next hop override, p - overrides from PfR

Gateway of last resort is not set

```
10.0.0.0/8 is variably subnetted, 20 subnets, 2 masks
O    10.100.1.1/32 [110/3] via 10.1.7.6, 4d05h, GigabitEthernet0/1
O    10.100.1.2/32 [110/3] via 10.1.7.6, 4d05h, GigabitEthernet0/1
O    10.100.1.3/32 [110/2] via 10.1.8.3, 4d06h, GigabitEthernet0/2
O    10.100.1.5/32 [110/3] via 10.1.7.6, 4d05h, GigabitEthernet0/1
O    10.100.1.6/32 [110/2] via 10.1.7.6, 4d05h, GigabitEthernet0/1
O    10.100.1.7/32 [110/3] via 10.1.8.3, 4d06h, GigabitEthernet0/2
O    10.100.1.8/32 [110/4] via 10.1.8.3, 4d05h, GigabitEthernet0/2
    [110/3] via 10.1.7.6, 4d05h, GigabitEthernet0/1
O    10.100.1.8/32 [110/4] via 10.1.8.3, 4d05h, GigabitEthernet0/2
    [110/4] via 10.1.7.6, 4d05h, GigabitEthernet0/1
```

BGP ORグループのRIB:

```
RP/0/0/CPU0:RR#show route afi-all safi-all topology ipv4-orr-group
```

IPv4 Unicast Topology ipv4-orr-group:

```
-----
Codes: C - connected, S - static, R - RIP, B - BGP, (>) - Diversion path
D - EIGRP, EX - EIGRP external, O - OSPF, IA - OSPF inter area
N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
E1 - OSPF external type 1, E2 - OSPF external type 2, E - EGP
i - ISIS, L1 - IS-IS level-1, L2 - IS-IS level-2
ia - IS-IS inter area, su - IS-IS summary null, * - candidate default
U - per-user static route, o - ODR, L - local, G - DAGR, l - LISP
A - access/subscriber, a - Application route
M - mobile route, r - RPL, (!) - FRR Backup path
```

Gateway of last resort is not set

```
o    10.100.1.1/32 [255/3] via 0.0.0.0, 14:14:52, Unknown
o    10.100.1.2/32 [255/3] via 0.0.0.0, 14:14:52, Unknown
o    10.100.1.3/32 [255/2] via 0.0.0.0, 00:04:53, Unknown
o    10.100.1.4/32 [255/0] via 0.0.0.0, 14:14:52, Unknown
o    10.100.1.5/32 [255/3] via 0.0.0.0, 14:14:52, Unknown
o    10.100.1.6/32 [255/2] via 0.0.0.0, 14:14:52, Unknown
o    10.100.1.7/32 [255/3] via 0.0.0.0, 14:14:52, Unknown
o    10.100.1.8/32 [255/4] via 0.0.0.0, 14:14:52, Unknown
```

```
RP/0/0/CPU0:RR#show rsi table name ipv4-orr-group
```

VR=default:

```
TBL_NAME=ipv4-orr-group, AFI=IPv4, SAFI=Ucast TBL_ID=0xe0000012 in VRF=default/0x60000000 in
VR=default/0x20000000
Refcnt=1
VRF Index=4 TCM Index=1
Flags=0x0 LST Flags=(0x0) NULL
```

show bgp neighbor コマンドは、ピアがORルートであるかどうかを表示します。

```
RP/0/0/CPU0:RR#show bgp neighbor 10.100.1.4
```

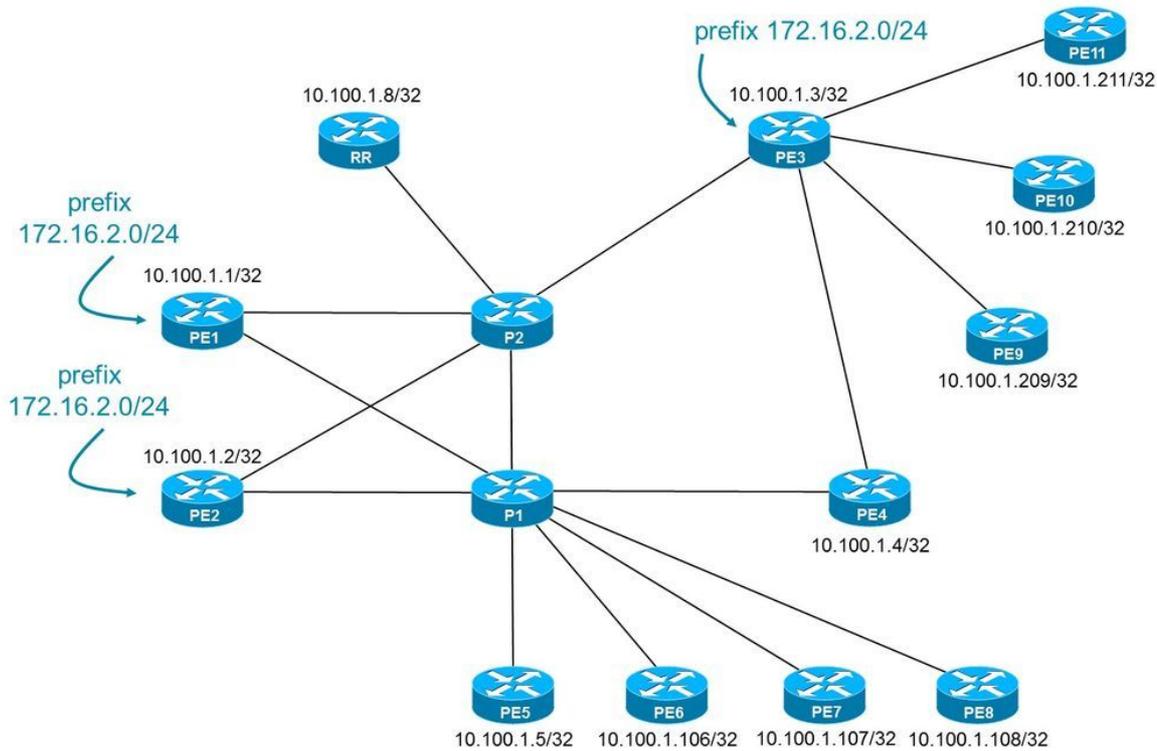
```
BGP neighbor is 10.100.1.4
Remote AS 1, local AS 1, internal link
Remote router ID 10.100.1.4
Cluster ID 10.100.1.8
BGP state = Established, up for 01:17:41
NSR State: None
```

```
Last read 00:00:52, Last read before reset 01:18:30
Hold time is 180, keepalive interval is 60 seconds
Configured hold time: 180, keepalive: 60, min acceptable hold time: 3
Last write 00:00:34, attempted 19, written 19
Second last write 00:01:34, attempted 19, written 19
Last write before reset 01:17:43, attempted 19, written 19
Second last write before reset 01:18:43, attempted 19, written 19
Last write pulse rcvd Mar  8 10:20:13.779 last full not set pulse count 12091
Last write pulse rcvd before reset 01:17:42
Socket not armed for io, armed for read, armed for write
Last write thread event before reset 01:17:42, second last 01:17:42
Last KA expiry before reset 01:17:43, second last 01:18:43
Last KA error before reset 00:00:00, KA not sent 00:00:00
Last KA start before reset 01:17:43, second last 01:18:43
Precedence: internet
Non-stop routing is enabled
Multi-protocol capability received
Neighbor capabilities:
  Route refresh: advertised (old + new) and received (old + new)
  4-byte AS: advertised and received
  Address family IPv4 Unicast: advertised and received
Received 6322 messages, 0 notifications, 0 in queue
Sent 5782 messages, 4 notifications, 0 in queue
Minimum time between advertisement runs is 0 secs
Inbound message logging enabled, 3 messages buffered
Outbound message logging enabled, 3 messages buffered

For Address Family: IPv4 Unicast
BGP neighbor version 41
Update group: 0.1 Filter-group: 0.1 No Refresh request being processed
Route-Reflector Client
ORR root (configured): ipv4-orr-group; Index: 0
  Route refresh request: received 0, sent 0
  0 accepted prefixes, 0 are bestpaths
  Cumulative no. of prefixes denied: 0.
  Prefix advertised 2, suppressed 0, withdrawn 0
  Maximum prefixes allowed 1048576
  Threshold for warning message 75%, restart interval 0 min
  AIGP is enabled
  An EoR was received during read-only mode
  Last ack version 41, Last synced ack version 0
  Outstanding version objects: current 0, max 2
  Additional-paths operation: None
  Send Multicast Attributes
  Advertise VPNv4 routes enabled with option
  Advertise VPNv6 routes is enabled with Local with stitching-RT option

Connections established 6; dropped 5
Local host: 10.100.1.8, Local port: 25176, IF Handle: 0x00000000
Foreign host: 10.100.1.4, Foreign port: 179
Last reset 01:17:42, due to User clear requested (CEASE notification sent - administrative
reset)
Time since last notification sent to neighbor: 01:17:42
Error Code: administrative reset
Notification data sent:
  None
```

次の図に示すように、複数のRRクライアントセットが設定されています



PE3に接続された1セットのRRクライアントと、P1に接続された別のセットがあります。各セットのすべてのRRクライアントは、出力BGPポータルと等しい距離にあります。

```

router bgp 1
  address-family ipv4 unicast
    optimal-route-reflection ipv4-orr-group-1 10.100.1.4 10.100.1.209 10.100.1.210
    optimal-route-reflection ipv4-orr-group-2 10.100.1.5 10.100.1.106 10.100.1.107
  !
  ...
  neighbor 10.100.1.4
    remote-as 1
    update-source Loopback0
    address-family ipv4 unicast
      optimal-route-reflection ipv4-orr-group-1
      route-reflector-client
    !
    !
  neighbor 10.100.1.5
    remote-as 1
    update-source Loopback0
    address-family ipv4 unicast
      optimal-route-reflection ipv4-orr-group-2
      route-reflector-client
    !
    !
  neighbor 10.100.1.106
    remote-as 1
    update-source Loopback0
    address-family ipv4 unicast
      optimal-route-reflection ipv4-orr-group-2
      route-reflector-client
    !
    !
  neighbor 10.100.1.107
    remote-as 1
    update-source Loopback0
  
```

```

address-family ipv4 unicast
  optimal-route-reflection ipv4-orr-group-2
  route-reflector-client
!
!
neighbor 10.100.1.108
remote-as 1
update-source Loopback0
address-family ipv4 unicast
  optimal-route-reflection ipv4-orr-group-2
  route-reflector-client
!
!
neighbor 10.100.1.209
remote-as 1
update-source Loopback0
address-family ipv4 unicast
  optimal-route-reflection ipv4-orr-group-1
  route-reflector-client
!
!
neighbor 10.100.1.210
remote-as 1
update-source Loopback0
address-family ipv4 unicast
  optimal-route-reflection ipv4-orr-group-1  route-reflector-client
!
!
neighbor 10.100.1.211
remote-as 1
update-source Loopback0
address-family ipv4 unicast
  optimal-route-reflection ipv4-orr-group-1
  route-reflector-client
!
!
!

```

両方のグループのorospfデータベース :

RP/0/0/CPU0:RR#**show orrspf database detail**

ORR policy: ipv4-orr-group-1, IPv4, RIB tableid: 0xe0000012
 Configured root: primary: 10.100.1.4, secondary: 10.100.1.209, tertiary: 10.100.1.210
 Actual Root: 10.100.1.4

Prefix	Cost
10.100.1.1	3
10.100.1.2	3
10.100.1.3	2
10.100.1.4	0
10.100.1.5	3
10.100.1.6	2
10.100.1.7	3
10.100.1.8	4
10.100.1.106	3
10.100.1.107	3
10.100.1.108	3
10.100.1.209	3
10.100.1.210	3
10.100.1.211	3

ORR policy: ipv4-orr-group-2, IPv4, RIB tableid: 0xe0000013
 Configured root: primary: 10.100.1.5, secondary: 10.100.1.106, tertiary: 10.100.1.107

Actual Root: 10.100.1.5

Prefix	Cost
10.100.1.1	3
10.100.1.2	3
10.100.1.3	4
10.100.1.4	3
10.100.1.5	0
10.100.1.6	2
10.100.1.7	3
10.100.1.8	4
10.100.1.106	3
10.100.1.107	3
10.100.1.108	3
10.100.1.209	5
10.100.1.210	5
10.100.1.211	5

Number of mapping entries: 30

グループでプライマリルートがダウンしているか到達不能な場合、セカンダリルートが実際に使用されるルートになります。この例では、グループipv4-or-group-1のプライマリルートに到達できません。セカンダリルートが実際のルートになります。

```
RP/0/0/CPU0:RR#show orrspf database ipv4-or-group-1
```

ORR policy: ipv4-or-group-1, IPv4, RIB tableid: 0xe0000012

Configured root: primary: 10.100.1.4, secondary: 10.100.1.209, tertiary: 10.100.1.210

Actual Root: 10.100.1.209

Prefix	Cost
10.100.1.1	4
10.100.1.2	5
10.100.1.3	2
10.100.1.5	5
10.100.1.6	4
10.100.1.7	3
10.100.1.8	4
10.100.1.106	5
10.100.1.107	5
10.100.1.108	5
10.100.1.209	0
10.100.1.210	3
10.100.1.211	3

Number of mapping entries: 14

結論

BGP Optimal Route Reflection(ORR)は、ルータリフレクタが存在する場合にiBGPネットワークでホットポットルーティングを可能にする機能です。境界ルータに新しいオペレーティングシステムソフトウェアは必要ありません。前提条件は、IGPがリンクステートルーティングプロトコルであることです。