

Reflexão de rota ideal do BGP (Border Gateway Protocol)

Contents

[Introduction](#)

[Informações de Apoio](#)

[Diagrama de Rede](#)

[Teoria](#)

[Implementação IOS-XR](#)

[Configurar](#)

[Exemplo de configuração](#)

[Engenharia de tráfego MPLS no roteador raiz](#)

[Troubleshoot](#)

[Conclusão](#)

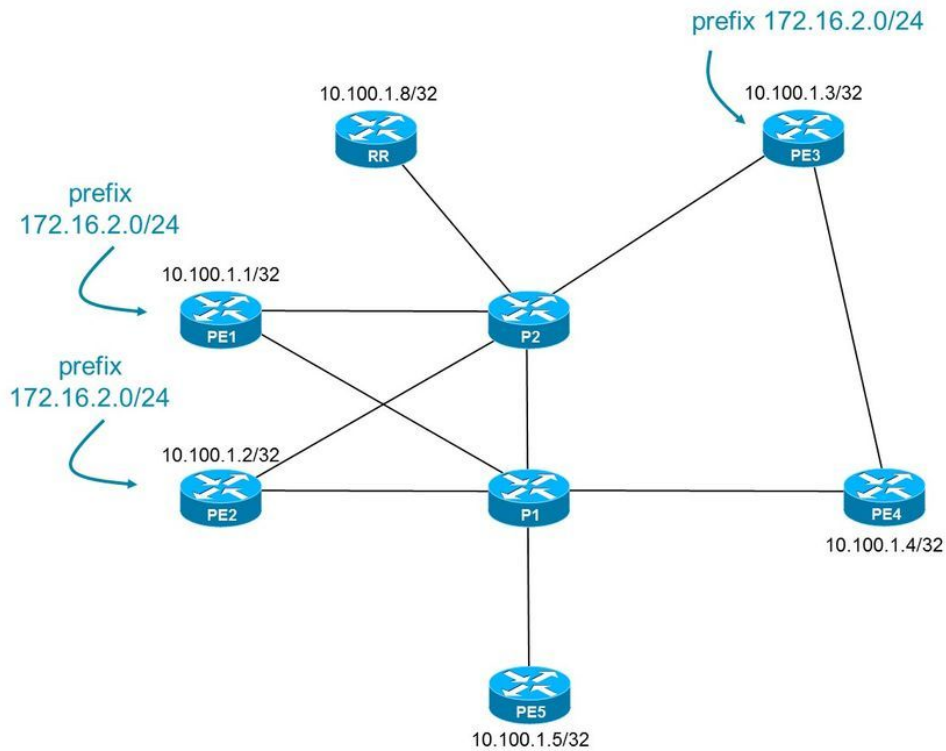
Introduction

Este documento descreve como o roteamento pode ser influenciado quando há um ou mais Refletores de Rota (RRs) na rede para evitar uma malha completa entre os roteadores iBGP.

Informações de Apoio

A etapa 8 no [algoritmo de seleção de melhor caminho BGP](#) é preferir o caminho com a menor métrica IGP ao salto seguinte BGP. Portanto, se todas as etapas antes da etapa 8 forem iguais, a etapa 8 pode ser o fator decisivo para qual é o melhor caminho no RR. O custo IGP do RR para o roteador iBGP de anúncio é então determinado pela colocação do RR. Por padrão, o RR apenas anuncia o melhor caminho para seus clientes. Dependendo de onde o RR é colocado, o custo IGP para o roteador de anúncio pode ser menor ou maior. Se todos os custos de IGP dos caminhos forem os mesmos, é provável que acabe até o ponto de quebra de tempo do roteador de anúncio com o menor ID de roteador BGP.

Diagrama de Rede



Os roteadores PE1, PE2 e PE3 anunciam o prefixo 172.16.2.0/24. Se todo o custo IGP dos links for o mesmo, o RR verá o caminho de PE1, PE2 e PE3 com um custo IGP de 2. No final, o RR escolhe o caminho de PE1 como o melhor porque tem o ID de roteador BGP mais baixo. Esta é a etapa 11 do algoritmo de seleção de melhor caminho BGP. O resultado é que todos os roteadores PE, incluindo PE4, escolherão PE1 como o roteador PE de saída para o prefixo 172.16.2.0/24. Do ponto de vista do PE4, o caminho IGP mais curto para qualquer roteador PE de saída é o caminho para PE3, com um custo IGP de 1. O custo IGP para qualquer outro roteador PE é 2. Para muitas redes, o fato de transportar o tráfego pela rede de trânsito da forma mais rápida possível é importante. Isso é conhecido como roteamento de batata quente.

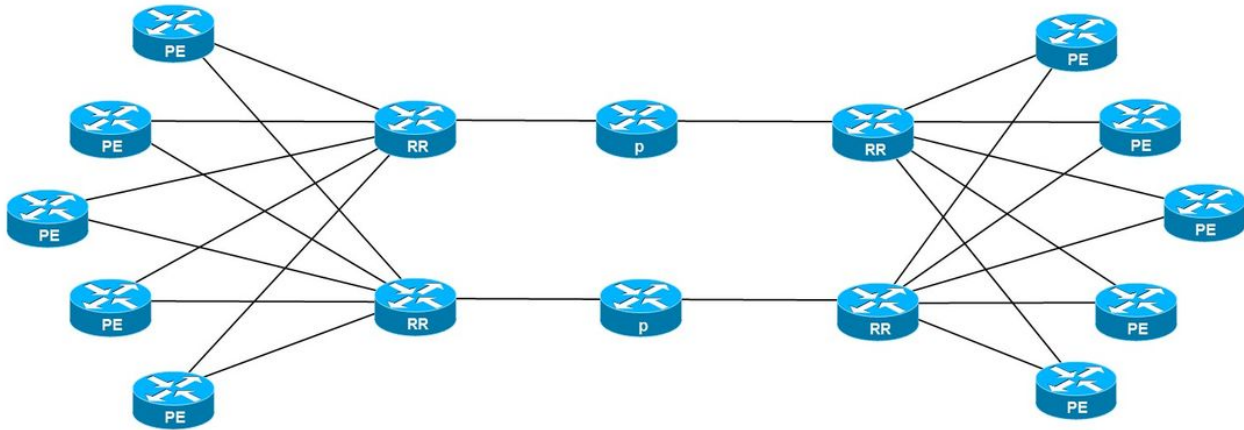
Há outro motivo possível para que o RR tenha escolhido o melhor caminho de PE1. Se na imagem, o custo Interior Gateway Protocol (IGP) do link P2-PE3 é 10 e todos os outros links ainda têm um custo IGP de 1, então o RR não escolheria o caminho de PE3 como o melhor, mesmo que PE3 tivesse o menor ID de roteador BGP.

Se o administrador dessa rede quiser ter roteamento de batata quente, um mecanismo deve estar em vigor para que, quando houver RRs na rede, o roteador de ingresso ainda possa aprender o caminho para o roteador de saída mais próximo na rede iBGP. O recurso BGP Add Path pode conseguir isso. No entanto, com esse recurso, os RR e os roteadores de borda devem ter um código mais recente que entenda o recurso. Com o recurso de reflexão de rota ideal de BGP, este não é um requisito. Esse recurso permitirá que o RR envie o melhor caminho para o roteador de borda de BGP de entrada, com base no que o RR acha ser o melhor caminho da perspectiva desse roteador BGP de entrada.

Outra solução que permitiria o roteamento de batata quente quando os RR são implantados é o posicionamento em linha dos RR. Esses RRs não são RRs dedicados, que executam somente o BGP e o IGP. Esses RR em linha estão no caminho de encaminhamento e são colocados na rede para que tenham seu próprio conjunto de clientes RR, para que possam refletir o melhor caminho para cada cliente RR, que também é o melhor caminho da perspectiva desse cliente RR.

Como mostrado nesta imagem, os RR são colocados na rede para que eles possam atender a um pequeno conjunto de clientes RR próximos. Devido ao projeto de rede, os clientes RR

recebem os melhores caminhos que são os melhores caminhos do ponto de vista deles, dos RR para que possa haver roteamento de batata quente na rede.



Teoria

A Reflexão de Rota Otimizada de BGP é descrita em IETF draft-ietf-idr-bgp-Optimal-Route-Reflection.

A solução BGP Optimal Route Reflection permite que o RR envie um melhor caminho específico para um roteador de borda BGP específico. O RR pode optar por enviar um melhor caminho diferente para roteadores de borda BGP diferentes ou para um conjunto de roteadores de borda. Os roteadores de borda devem ser clientes RR do RR. O objetivo é que cada roteador de borda BGP de entrada possa ter um roteador BGP de saída ou saída diferente para o mesmo prefixo. Se o roteador de borda de ingresso puder sempre encaminhar o tráfego para o roteador de saída AS fechado, isso permite o roteamento de batata quente.

O problema é que o RR normalmente envia o mesmo melhor caminho para cada roteador de borda BGP, o que impede o roteamento de batata quente. Para resolver isso, você precisa que o RR seja capaz de calcular diferentes melhores caminhos para o mesmo prefixo, dependendo do roteador de borda de BGP de entrada. O cálculo do melhor caminho no RR é feito com base na posição do roteador de borda de BGP de entrada. Portanto, o RR executará o cálculo do melhor caminho de BGP da perspectiva do roteador de borda de entrada. O RR que só pode fazer isso é o RR que tem a imagem completa da topologia da rede da perspectiva IGP onde estão os roteadores RR e de borda de entrada. Para que esse requisito seja atendido, o IGP deve ser um protocolo de roteamento link-state.

Nesse caso, o RR pode executar um cálculo SPF (Shortest Path First) com o roteador de borda de entrada como raiz da árvore e calcular o custo para todos os outros roteadores. Dessa forma, o custo do roteador de borda de entrada para todos os outros roteadores de borda de saída será conhecido. Esse cálculo SPF especial com outro roteador como raiz é conhecido como SPF inverso (rSPF). Isso só pode ser feito se o RR aprender todos os caminhos de BGP de todos os

roteadores de borda de BGP. Pode haver tantos rSPFs executados quanto clientes RR. Isso aumentará um pouco a carga da CPU no RR.

A solução permite que o cálculo do melhor caminho seja baseado no algoritmo de seleção do melhor caminho BGP, o que levará ao RR escolher o melhor caminho da perspectiva do roteador de borda de entrada para o qual o RR envia o caminho. Isso significa que o melhor caminho será escolhido com base no menor custo IGP para o salto seguinte do BGP. A solução também permite que o melhor caminho seja escolhido com base em alguma política configurada. Os roteadores de borda de ingresso podem escolher seus melhores caminhos com base em alguma política configurada e não no menor custo IGP. A solução permite que o RR implemente a reflexão de rota ideal no custo IGP (local na rede) ou em alguma política configurada, ou ambos. Se ambos forem implantados, a política será aplicada primeiro e, em seguida, a reflexão de rota otimizada baseada em IGP ocorrerá nos caminhos restantes.

Implementação IOS-XR

A implementação IOS-XR permite até três nós raiz para o cálculo do rSPF. Se você tiver muitos clientes RR em um grupo de atualização, não haverá necessidade de um cálculo rSPF por cliente RR se esses clientes RR tiverem a mesma política e/ou os mesmos custos IGP para os diferentes roteadores de borda de BGP de saída. Este último geralmente significa que os clientes RR estão co-localizados (provavelmente no mesmo POP). Se for esse o caso, não há necessidade de configurar cada cliente RR como raiz. A implementação IOS-XR permite configurar três, a primária, a secundária e a raiz terciária, por conjunto de clientes RR, para fins de redundância. Para que o recurso BGP ORR se aplique a qualquer cliente RR, esse cliente RR deve ser configurado para fazer parte de um grupo de política ORR.

O recurso BGP ORR é ativado por família de endereços.

É necessário um protocolo link-state. Pode ser OSPF ou IS-IS.

O IOS XR implementa somente o recurso BGP ORR com base no custo IGP para o salto seguinte BGP, e não com base em alguma política configurada.

Os peers BGP com a mesma política de saída são colocados no mesmo grupo de atualização. Geralmente, esse é o caso do iBGP no RR. Quando o recurso BGP ORR está ativado, os peers de diferentes grupos ORR estarão em diferentes grupos de atualização. Isso é lógico, porque as atualizações enviadas do RR para os clientes RR em diferentes grupos de BGP ORR serão diferentes, porque o melhor caminho de BGP é diferente.

O resultado dos cálculos do rSPF é armazenado em um banco de dados.

O ORRSPF é o novo componente no IOS-XR necessário para o recurso BGP ORR. O ORRSPF cuida de:

1. Coleta de informações de link-state e manutenção do banco de dados de link-state
2. Executando rSPFs e mantendo os SPTs, por grupo de política
3. Download dos prefixos do SPT para o RIB com as métricas

O banco de dados obtém suas informações de link-state diretamente do IGP de link-state ou do BGP-LS.

Os cálculos do rSPF resultam em uma topologia que mostra o caminho mais curto do cliente RR para qualquer outro roteador na área/nível.

As rotas suspensas em cada roteador na topologia são armazenadas em uma tabela RIB especial para a política de grupo ORR e por AFI/SAFI. Esta tabela é criada pelo RSI. A tabela é preenchida pelas rotas calculadas pelos rSPFs com a raiz primária como raiz. Se a raiz primária ficar indisponível, a raiz secundária será a raiz e preencherá as rotas na tabela ORR RIB. O mesmo se aplica à raiz terciária.

Configurar

A configuração mínima necessária:

1. O ORR precisa ser ativado para a família de endereços do BGP, para grupos específicos de vizinhos do BGP
2. Para cada grupo de vizinhos BGP, pelo menos uma raiz precisa ser configurada. Opcionalmente, uma raiz secundária e terciária pode ser configurada.
3. A redistribuição das rotas ORR do IGP para o BGP precisa ser habilitada.

Exemplo de configuração

Como mostrado na primeira imagem, o RR é um roteador IOS-XR com o recurso BGP ORR.

Todos os outros roteadores estão executando o IOS. Esses roteadores não têm o recurso BGP ORR.

PE1, PE2 e PE3 anunciam o prefixo 172.16.2.0/24 em AFI/SAFI 1/1 (unicast IPv4). O RR é igualmente próximo a PE1 e PE2 do que a PE3. O custo IGP de todos os links é 1. O melhor caminho para esse prefixo é aquele com R1 como próximo salto devido ao menor ID de roteador BGP.

```
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)

O PE4 receberá o caminho com PE1 como próximo salto. Portanto, não há roteamento de batata quente para PE4.

Se você quiser ter roteamento de batata quente no PE4, então para os prefixos anunciados por PE1, PE2 e PE3 (por exemplo, o prefixo 172.16.2.0/24), PE1 deve ter PE3 como ponto de saída. Isso significa que o caminho no PE4 deve ser aquele com PE3 como próximo salto. Podemos fazer com que o RR envie a rota com o próximo salto PE3 para PE4 com essa configuração ORR.

```
router ospf 1
distribute bgp-ls
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 vpv4 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
!
!
!

```

Se o IGP for IS-IS:

```

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

```

Note: O link-state da família de endereços não precisa ser configurado globalmente ou sob os vizinhos BGP.

Engenharia de tráfego MPLS no roteador raiz

O RR precisa encontrar o endereço raiz configurado no banco de dados IGP para executar o rSPF. No ISIS, o ID do roteador está presente no banco de dados do ISIS. Para OSPF, não há ID de roteador presente nos LSAs do OSPF. A solução é fazer com que os roteadores raiz anunciem o ID do roteador MPLS (Multi Protocol Label Switching) TE correspondente ao endereço raiz configurado no RR.

Para o OSPF, os roteadores raiz precisam de configuração adicional para fazer o BGP ORR funcionar. É necessária uma configuração mínima de TE MPLS em qualquer roteador raiz para anunciar esse ID de roteador TE MPLS. O conjunto de comandos mínimo exato depende do sistema operacional do roteador raiz. A configuração do TE MPLS no roteador raiz precisa ter a configuração mínima para o TE MPLS habilitada para que o OSPF anuncie o ID do roteador TE MPLS em um LSA de área opaca (tipo 10).

Quando o RR tiver um LSA de área opaca com o ID de roteador TE MPLS correspondente ao endereço do roteador raiz configurado, o rSPF poderá ser executado e o BGP no RR poderá anunciar a rota ideal.

A configuração mínima necessária para o OSPF no roteador raiz se for um roteador IOS é:

```

!
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
!

```

Observe que:

- O TE de MPLS está ativado na área específica do OSPF
- o ID do roteador TE MPLS está configurado correspondendo ao endereço raiz configurado no RR
- O TE MPLS está configurado em pelo menos uma interface
- não há necessidade de configurar o RSVP-TE
- não há necessidade de ter o TE MPLS configurado em nenhum outro roteador na área

A configuração mínima necessária para o OSPF no roteador raiz se for um roteador IOS-XR é:

```

!
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
!

```

Se a configuração acima estiver no roteador raiz, o RR deverá ter o ID do roteador TE MPLS no banco de dados OSPF.

```
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: 0xalad
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

Observe que o ID do roteador MPLS TE (10.100.1.4) e o ID do roteador OSPF são diferentes.

PE4 tem PE3 como próximo salto para o prefixo (com a métrica IGP correta para o próximo salto):

```
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 ainda tem PE1 como o próximo salto para o prefixo (com a métrica IGP correta para o próximo salto):

```
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
```

Troubleshoot

Verifique o prefixo no RR:

```
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
```

Observe que o caminho adicional foi adicionado aos outros caminhos não melhores, para que também possam ser anunciados, além do melhor caminho. O recurso adicionar caminho não é usado entre o RR e seus clientes: os caminhos não são anunciados com um identificador de caminho.

Verifique se as rotas estão (ainda) anunciadas aos vizinhos BGP específicos.

Para o vizinho PE4, o próximo salto é PE3 para o prefixo 172.16.2.0/24:

```
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

Para o vizinho PE5, o próximo salto é PE1 para o prefixo 172.16.2.0/24:

```
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

O vizinho 10.100.1.4 está em seu próprio grupo de atualização devido à política ORR em vigor:

```
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

O comando show orrspf database mostra o grupo ORR e suas raiz,

```
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

O mesmo comando com a palavra-chave detail fornece o custo da raiz do rSPF para um outro roteador/prefixo na mesma área OSPF:

```
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

O ID da tabela foi atribuído pelo RSI para o grupo ORR e para o AFI/SAFI:

```
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

O custo da raiz (R4/10.100.1.4) do rSPF para o outro roteador é o mesmo que o custo visto com show ip route ospf em PE4:

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
[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

O RIB para o grupo BGP ORR:

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

O comando `show bgp neighbor` mostra se o peer é uma raiz ORR:

```
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)
```

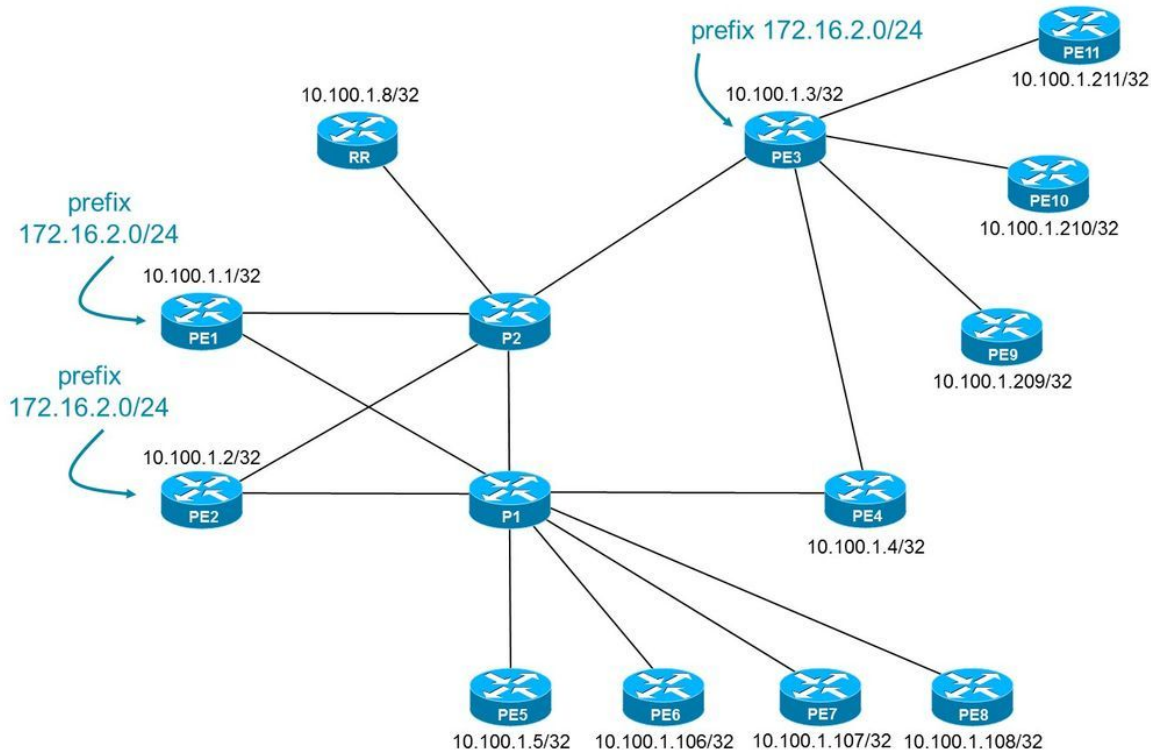
Time since last notification sent to neighbor: 01:17:42

Error Code: administrative reset

Notification data sent:

None

Como mostrado nesta imagem, vários conjuntos de clientes RR configurados



Há um conjunto de clientes RR conectados ao PE3 e outro conjunto conectado ao P1. Cada cliente RR em cada conjunto está em uma distância igual a qualquer roteador de borda de BGP de saída.

```
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
!
!
!

```

O banco de dados do orrspf para ambos os grupos:

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

Se, para um grupo, a raiz primária estiver inativa ou inacessível, a raiz secundária será a raiz real usada. Neste exemplo, a raiz primária do grupo ipv4-orr-group-1 não pode ser alcançada. A raiz secundária tornou-se a raiz real:

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

```
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: 10.100.1.209
```

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

Conclusão

A Reflexão de Rota Otimizada (ORR - Optimal Route Reflektion) de BGP é um recurso que permite o roteamento de batata quente em uma rede iBGP quando há refletores de rota, sem a necessidade de software de sistema operacional mais recente nos roteadores de borda. O pré-requisito é que o IGP seja um protocolo de roteamento link-state.