

# Reflexión de ruta óptima del protocolo de gateway fronterizo (BGP)

## Contenido

[Introducción](#)

[Antecedentes](#)

[Diagrama de la red](#)

[Teoría](#)

[Implementación de IOS-XR](#)

[Configurar](#)

[Ejemplo de configuración](#)

[MPLS Traffic-Engineering en Router Raíz](#)

[Troubleshoot](#)

[Conclusión](#)

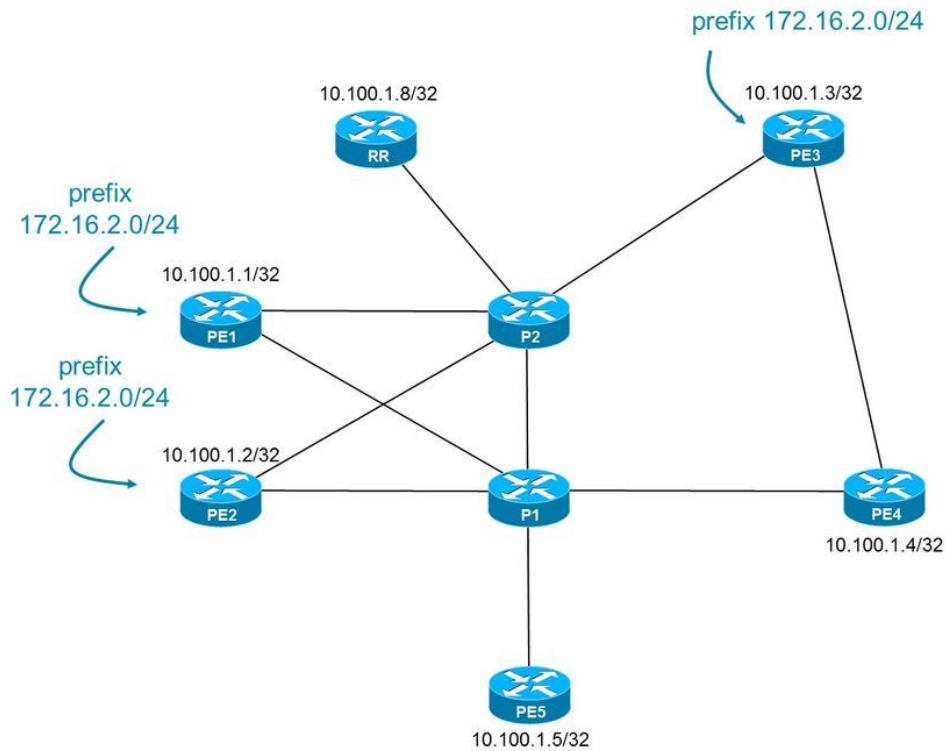
## Introducción

Este documento describe cómo puede influenciarse el ruteo cuando hay uno o más Reflectores de Ruta (RR) en la red para evitar una malla completa entre los routers iBGP.

## Antecedentes

El paso 8 en el [algoritmo de selección de la mejor trayectoria de BGP](#) es preferir la trayectoria con la métrica de IGP más baja al salto siguiente de BGP. Por lo tanto, si todos los pasos anteriores al paso 8 son iguales, entonces el paso 8 puede ser el factor decisivo en cuál es el mejor trayecto en el RR. El costo IGP del RR al router iBGP de anuncio se determina luego por la ubicación del RR. De forma predeterminada, el RR sólo anuncia la mejor trayectoria a sus clientes. Según dónde se ubique el RR, el costo IGP para el router de anuncio puede ser menor o mayor. Si todos los costos IGP de las trayectorias son iguales, es probable que termine hasta el desempate del router de anuncio con el ID de router BGP más bajo.

## Diagrama de la red



Los routers PE1, PE2 y PE3 anuncian el prefijo 172.16.2.0/24. Si todos los costos IGP de los links son iguales, entonces el RR verá la trayectoria de PE1, PE2 y PE3 con un costo IGP de 2. Al final, el RR elige el trayecto de PE1 como el mejor porque tiene el ID de router BGP más bajo. Este es el paso 11 en el algoritmo de selección de la mejor trayectoria de BGP. El resultado es que todos los routers PE, incluido PE4, elegirán PE1 como el router PE de salida para el prefijo 172.16.2.0/24. Desde el punto de vista de PE4, la trayectoria IGP más corta a cualquier router PE de salida es la trayectoria a PE3, con un costo IGP de 1. El costo IGP para cualquier otro router PE es 2. Para muchas redes, el hecho de transportar el tráfico a través de la red de tránsito de la manera más corta posible es importante. Esto se conoce como ruteo de papa caliente.

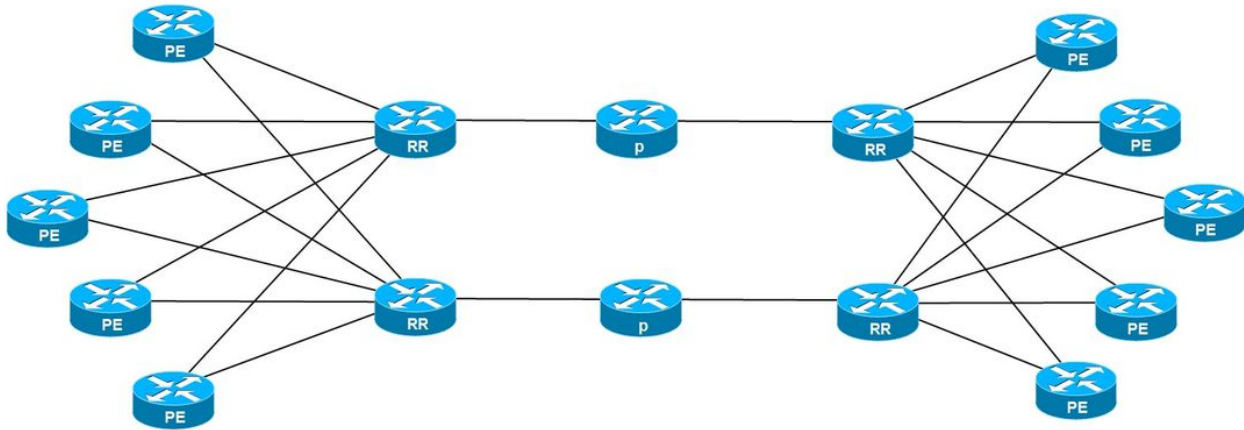
Hay otra razón posible para que RR haya elegido la mejor trayectoria de PE1. Si en la imagen, el costo del protocolo de gateway interior (IGP) del enlace P2-PE3 es 10 y todos los demás links todavía tienen un costo IGP de 1, entonces el RR no elegiría la ruta de PE3 como la mejor, incluso si PE3 tuviera el ID de router BGP más bajo.

Si el administrador de esta red desea tener un ruteo de papa caliente, entonces debe haber un mecanismo para que cuando haya RR en la red, el router de ingreso pueda aprender la trayectoria al router de salida más cercano en la red iBGP. La función BGP Add Path puede lograr esto. Sin embargo, con esa función, los RR y los routers de borde deben tener código más reciente que entienda la función. Con la función BGP Optimal Route Reflection, esto no es un requisito. Esta función permitirá que el RR envíe la mejor trayectoria al router de borde BGP de ingreso, en base a lo que el RR piensa que es la mejor trayectoria desde la perspectiva de ese router BGP de ingreso.

Otra solución que permitiría el ruteo de papa caliente cuando se implementan RR, es la ubicación en línea de RR. Estos RR no son RR dedicados, que sólo ejecutan BGP y el IGP. Estos RR en línea se encuentran en la trayectoria de reenvío y se colocan en la red para que tengan su propio conjunto de clientes RR, de modo que puedan reflejar la mejor trayectoria a cada cliente RR, que también es la mejor trayectoria desde la perspectiva de ese cliente RR.

Como se muestra en esta imagen, los RR se colocan en la red para que sean un pequeño conjunto de clientes RR cercanos a los que puedan servir. Debido al diseño de la red, los clientes

RR reciben las mejores trayectorias que son las mejores trayectorias desde su punto de vista, desde los RR para que pueda haber ruteo de papa caliente en la red.



## Teoría

La Reflexión de Ruta Óptima BGP se describe en el borrador IETF draft-ietf-idr-bgp-óptimo-route-revision.

La solución BGP Optimal Route Reflection permite al RR enviar una mejor trayectoria específica a un router de borde BGP específico. El RR puede elegir enviar una mejor trayectoria diferente a diferentes routers de borde BGP o conjunto de routers de borde. Los routers de borde deben ser clientes RR del RR. El objetivo es que cada router de borde BGP de ingreso puede tener un router BGP de salida o salida diferente para el mismo prefijo. Si el router de borde de ingreso siempre puede reenviar el tráfico al router AS-exit de cierre, esto permite el ruteo de papa caliente.

El problema es que el RR normalmente sólo envía la misma mejor trayectoria a cada router de borde BGP, lo que evita el ruteo de papas calientes. Para resolver esto, necesita que el RR sea capaz de calcular diferentes mejores trayectorias para el mismo prefijo dependiendo del router de borde BGP de ingreso. El mejor cálculo de trayectoria en el RR se realiza en base a la posición del router de borde BGP de ingreso. Por lo tanto, el RR realizará el cálculo del mejor trayecto BGP desde la perspectiva del router de borde de ingreso. El RR que sólo puede hacer esto es el RR que tiene la imagen completa de la topología de la red desde la perspectiva IGP donde se encuentran los routers de borde de ingreso y RR. Para cumplir este requisito, el IGP debe ser un protocolo de ruteo de estado de link.

En ese caso, el RR puede ejecutar un cálculo de trayecto más corto primero (SPF) con el router de borde de ingreso como la raíz del árbol y calcular el costo para cada otro router. De esta manera, se conocerá el costo del router de borde de ingreso a todos los demás routers de borde de salida. Este cálculo SPF especial con otro router como la raíz, se denomina SPF Inversa (rSPF). Esto sólo puede hacerse si el RR aprende todas las trayectorias BGP de todos los routers

de borde BGP. Podría haber tantos rSPF ejecutados como clientes RR. Esto incrementará la carga de la CPU de alguna manera en el RR.

La solución permite que el mejor cálculo del trayecto se base en el algoritmo de selección del mejor trayecto BGP, que llevará al RR eligiendo el mejor trayecto desde la perspectiva del router de borde de ingreso al que el RR envía el trayecto. Esto significa que la mejor trayectoria se elegirá en función del costo IGP más corto para el siguiente salto BGP. La solución también permite seleccionar la mejor ruta en función de algunas políticas configuradas. Los routers de borde de ingreso podrían elegir sus mejores trayectorias en función de alguna política configurada, y no en el costo IGP más bajo. La solución permite al RR implementar la reflexión de ruta óptima en el costo IGP (ubicación en la red) o en alguna política configurada, o en ambas. Si se implementan ambas, la política se aplica primero y luego la reflexión de ruta óptima basada en IGP ocurrirá en las trayectorias restantes.

## Implementación de IOS-XR

La implementación de IOS-XR permite hasta tres nodos raíz para el cálculo de rSPF. Si tiene muchos clientes RR en un grupo de actualización, entonces no hay necesidad de un cálculo rSPF por cliente RR si esos clientes RR tendrán la misma política y/o los mismos costos IGP para los diferentes routers de borde BGP de salida. Esto último significa generalmente que los clientes RR están ubicados de manera conjunta (probablemente en el mismo POP). Si ese es el caso, no hay necesidad de configurar cada cliente RR como raíz. La implementación de IOS-XR permite configurar tres, la raíz primaria, la secundaria y la raíz terciaria, por conjunto de clientes RR, con fines de redundancia. Para que la función BGP ORR se aplique a cualquier cliente RR, ese cliente RR se debe configurar para formar parte de un grupo de políticas ORR.

La función BGP ORR se habilita por familia de direcciones.

Se requiere un protocolo de estado de link. Puede ser OSPF o IS-IS.

IOS XR solamente implementa la función BGP ORR basándose en el costo IGP para el siguiente salto BGP, y no en base a alguna política configurada.

Los peers BGP con la misma política saliente se colocan en el mismo grupo de actualización. Este suele ser el caso de iBGP en el RR. Cuando la función BGP ORR está habilitada, los peers de diferentes grupos ORR estarán en diferentes grupos de actualización. Esto es lógico, porque las actualizaciones enviadas desde el RR a los clientes RR en diferentes grupos ORR BGP serán diferentes, porque la mejor trayectoria de BGP es diferente.

El resultado de los cálculos de rSPF se almacena en una base de datos.

ORRSPF es el nuevo componente en IOS-XR que se necesita para la función BGP ORR. ORRSPF se encarga de:

1. Recopilación de la información de estado de link y mantenimiento de la base de datos de estado de link
2. Ejecución de rSPF y mantenimiento de los SPT, por grupo de políticas
3. Descarga de los prefijos del SPT al RIB con las métricas

La base de datos obtiene su información de estado de link directamente desde el IGP de estado de link o desde BGP-LS.

Los cálculos de rSPF resultan en una topología que muestra la trayectoria más corta del cliente

RR a cualquier otro router en el área/nivel.

Las rutas que cuelgan de cada router en la topología se almacenan en una tabla RIB especial para la política de grupo ORR y por AFI/SAFI. RSI crea esta tabla. La tabla se completa con las rutas calculadas por los rSPF con la raíz primaria como la raíz. Si la raíz primaria deja de estar disponible, entonces la raíz secundaria es la raíz y rellena las rutas en la tabla ORR RIB. Lo mismo se aplica a la raíz terciaria.

## Configurar

La configuración mínima necesaria:

1. ORR debe estar habilitado para la familia de direcciones de BGP, para grupos específicos de vecinos BGP
2. Para cada grupo de vecinos BGP, se debe configurar al menos una raíz. Opcionalmente, se puede configurar una raíz secundaria y terciaria.
3. La redistribución de las rutas ORR del IGP al BGP debe estar habilitada.

## Ejemplo de configuración

Como se muestra en la primera imagen, el RR es un router IOS-XR con la función BGP ORR.

Todos los demás routers ejecutan IOS. Estos routers no tienen la función BGP ORR.

PE1, PE2 y PE3 anuncian el prefijo 172.16.2.0/24 en AFI/SAFI 1/1 (unidifusión IPv4). El RR está igual de cerca de PE1 y PE2 que de PE3. El costo IGP de todos los links es 1. La mejor trayectoria para este prefijo es la con R1 como salto siguiente debido al ID de router BGP más bajo.

```
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 recibirá la trayectoria con PE1 como salto siguiente. Por lo tanto, no hay ruteo de papa caliente para PE4.

Si desea tener un ruteo de papa caliente en PE4, entonces para los prefijos anunciados por PE1, PE2 y PE3 (por ejemplo el prefijo 172.16.2.0/24), PE1 debe tener PE3 como punto de salida. Esto significa que la trayectoria en PE4 debe ser la con PE3 como salto siguiente. Podemos hacer que el RR envíe la ruta con el siguiente salto PE3 a PE4 con esta configuración ORR.

```
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
!
!
!

```

Si el IGP es 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
!
!
!

```

**Nota:** El estado de link de la familia de direcciones no necesita ser configurado globalmente o bajo los vecinos BGP.

## MPLS Traffic-Engineering en Router Raíz

El RR necesita encontrar la dirección raíz configurada en la base de datos IGP, para ejecutar el rSPF. En ISIS, el ID de router está presente en la base de datos de ISIS. Para OSPF, no hay ningún router-ID presente en los LSA OSPF. La solución es hacer que los routers raíz anuncien el ID de router TE de conmutación de etiquetas multiprotocolo (MPLS) que coincida con la dirección raíz configurada en el RR.

Para OSPF, los routers raíz necesitan configuración adicional para hacer que el BGP ORR funcione. Se necesita una configuración MPLS TE mínima en cualquier router raíz para anunciar este ID de router MPLS TE. El conjunto de comandos mínimo exacto depende del sistema operativo del router raíz. La configuración MPLS TE en el router raíz necesita tener habilitada la configuración mínima para MPLS TE de modo que OSPF anuncie el ID del router MPLS TE en un LSA de área opaca (tipo 10).

Una vez que el RR tiene un LSA de área opaca con el router-ID MPLS TE que coincide con la dirección del router raíz configurada, el rSPF puede ejecutarse y el BGP en el RR puede anunciar la ruta óptima.

La configuración mínima necesaria para OSPF en el router raíz si es un router IOS es:

```

!
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:

- MPLS TE se habilita en el área OSPF específica
- el router-ID MPLS TE se configura coincidiendo con la dirección raíz configurada en el RR
- MPLS TE se configura en al menos una interfaz
- no es necesario tener RSVP-TE configurado
- no es necesario que MPLS TE se configure en ningún otro router del área

La configuración mínima necesaria para OSPF en el router raíz si es un router IOS-XR es:

```

!
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
!

```

Si la configuración anterior está en vigor en el router raíz, entonces RR debe tener el router-ID MPLS TE en la base de datos 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 el router-ID MPLS TE (10.100.1.4) y el router-ID OSPF son diferentes.

PE4 tiene PE3 como salto siguiente para el prefijo (con la métrica IGP correcta para el salto siguiente):

```
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 todavía tiene PE1 como salto siguiente para el prefijo (con la métrica IGP correcta para el salto siguiente):

```
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 el prefijo en el 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 add-path se agregó a las otras trayectorias no mejores, de modo que también se puedan anunciar, además de la mejor trayectoria. La función add path no se utiliza entre el RR y sus clientes: las trayectorias no se anuncian con un identificador de trayectoria.

Verifique que las rutas se anuncien (todavía) a los vecinos BGP específicos.

Para el PE4 vecino, el salto siguiente es PE3 para el prefijo 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	<b>10.100.1.3</b>	10.100.1.3	i

Processed 2 prefixes, 2 paths

Para el PE5 vecino, el salto siguiente es PE1 para el prefijo 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	<b>10.100.1.1</b>	10.100.1.1	i

El vecino 10.100.1.4 está en su propio grupo de actualización debido a la política ORR implementada:

```
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
```

El comando `show orrspf database` muestra el grupo ORR y sus raíces,

```
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
```

El mismo comando con la palabra clave `detail` proporciona el costo de la raíz del rSPF entre el otro router/prefijo en el mismo 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
```

El ID de tabla fue asignado por el RSI para el grupo ORR y para el 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

El costo de la raíz (R4/10.100.1.4) del rSPF para cada router es el mismo que el costo que se ve con `show ip route ospf` en 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

**EI RIB para el grupo ORR BGP:**

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

El comando **show bgp neighbor** muestra si el peer es una raíz 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)
```

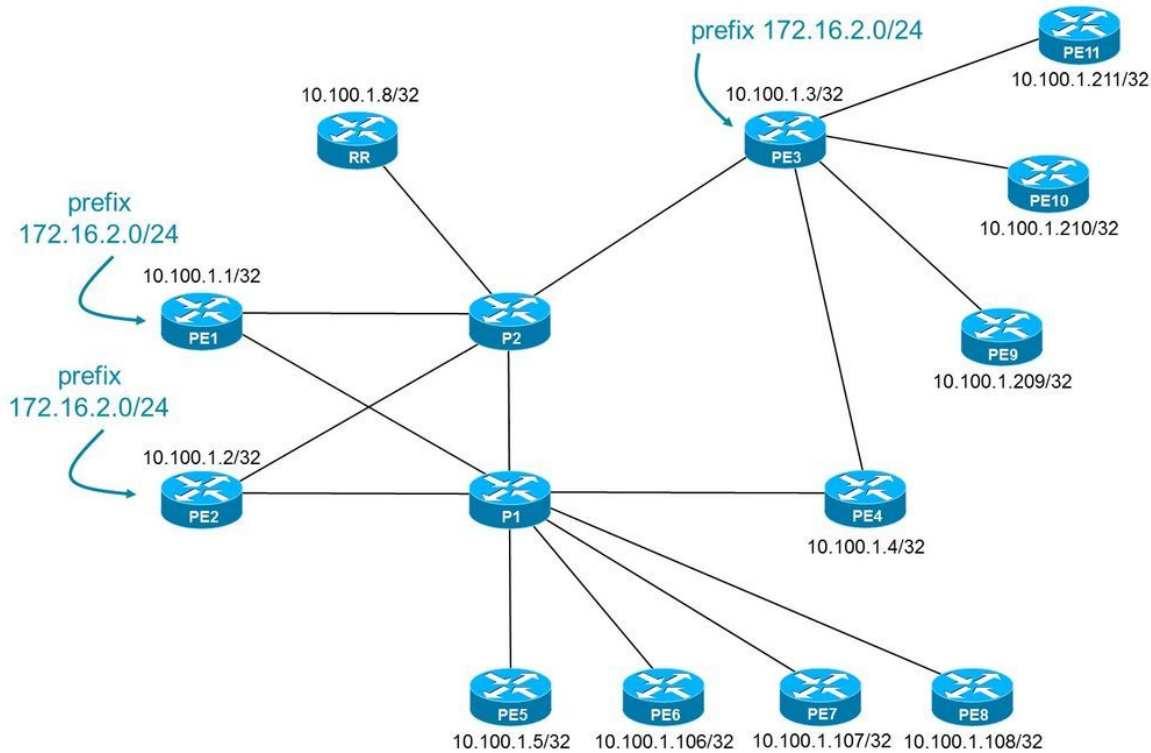
Time since last notification sent to neighbor: 01:17:42

Error Code: administrative reset

Notification data sent:

None

Como se muestra en esta imagen, varios conjuntos de clientes RR configurados



Hay un conjunto de clientes RR conectados a PE3 y otro conjunto conectado a P1. Cada cliente RR en cada conjunto se encuentra a la misma distancia que cualquier router de borde BGP de salida.

```
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
  !
  !
!
```

La base de datos orspf para ambos 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

Si para un grupo la raíz primaria está inactiva o inalcanzable, entonces la raíz secundaria será la raíz real utilizada. En este ejemplo, la raíz primaria del grupo ipv4-o-group-1 es inalcanzable. La raíz secundaria se convirtió en la raíz 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

## Conclusión

La función BGP Optimal Route Reflection (ORR) es una función que permite el ruteo de papa caliente en una red iBGP cuando hay reflectores de ruta presentes, sin la necesidad de un software de sistema operativo más nuevo en los routers de borde. El prerrequisito es que el IGP es un protocolo de ruteo de estado de link.