

BGP (Border Gateway Protocol) - Réflexion optimale de la route

Contenu

[Introduction](#)

[Informations générales](#)

[Diagramme du réseau](#)

[Théorie](#)

[Implémentation IOS-XR](#)

[Configuration](#)

[Exemple de configuration](#)

[Ingénierie du trafic MPLS sur routeur racine](#)

[Dépannage](#)

[Conclusion](#)

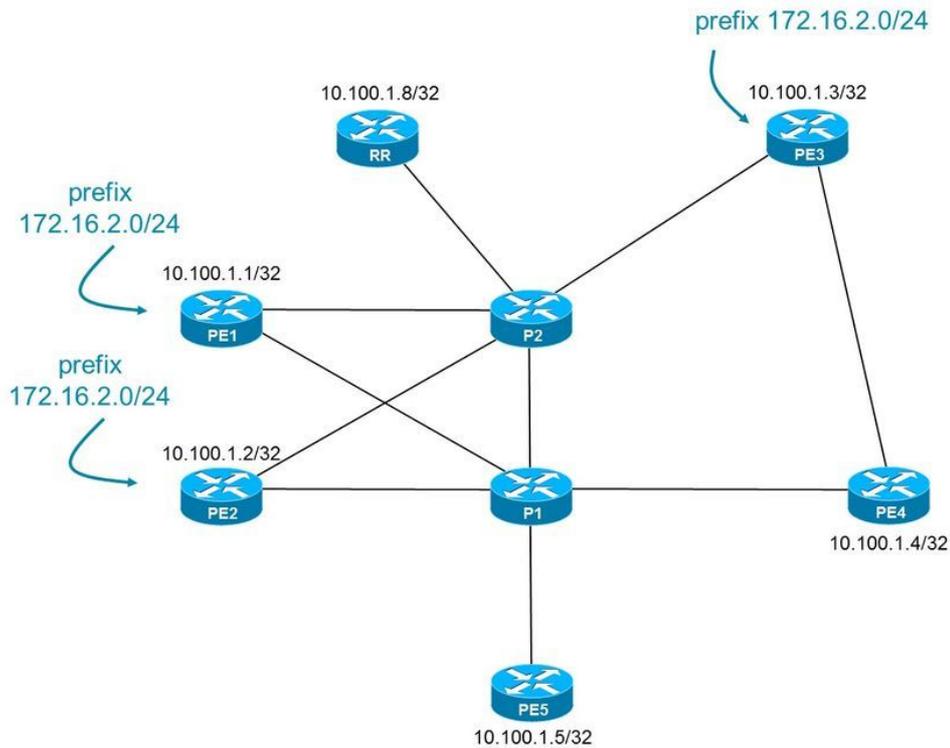
Introduction

Ce document décrit comment le routage peut être influencé lorsqu'il y a un ou plusieurs réflecteurs de route (RR) dans le réseau pour éviter un maillage complet entre les routeurs iBGP.

Informations générales

L'étape 8 de l'[algorithme de sélection du meilleur chemin BGP](#) consiste à préférer le chemin avec la métrique IGP la plus basse au prochain saut BGP. Ainsi, si toutes les étapes précédant l'étape 8 sont égales, l'étape 8 peut être le facteur déterminant sur le meilleur chemin du RR. Le coût IGP du RR au routeur iBGP publicitaire est ensuite déterminé par le placement du RR. Par défaut, le RR annonce uniquement le meilleur chemin à ses clients. Selon l'emplacement du RR, le coût IGP pour le routeur publicitaire peut être plus faible ou plus important. Si tous les coûts IGP des chemins sont les mêmes, alors il est probable qu'il finira par le séparateur du routeur publicitaire ayant l'ID de routeur BGP le plus faible.

Diagramme du réseau



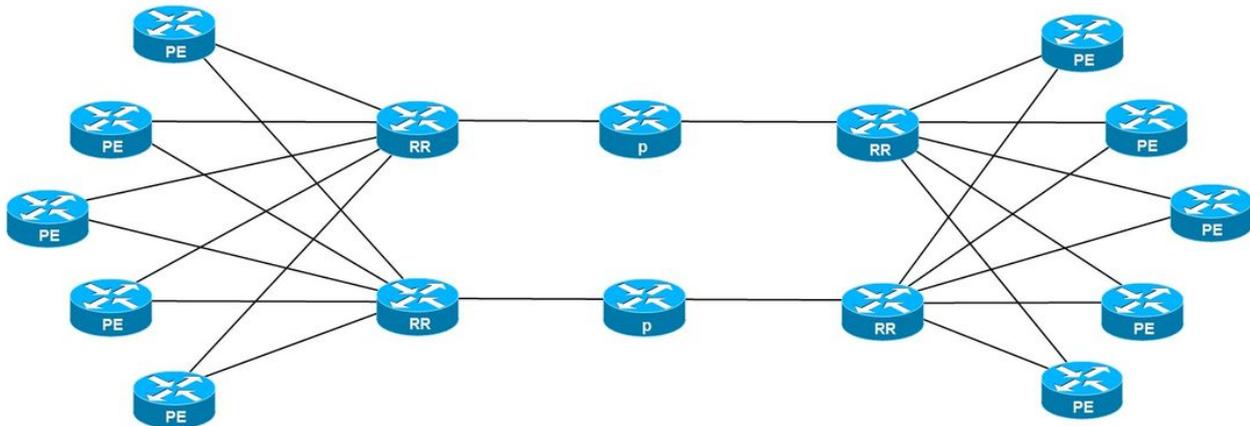
Les routeurs PE1, PE2 et PE3 annoncent le préfixe 172.16.2.0/24. Si tous les coûts IGP des liaisons sont identiques, le RR verra le chemin depuis PE1, PE2 et PE3 avec un coût IGP de 2. En fin de compte, le RR choisit le chemin de PE1 comme étant le meilleur car il possède l'ID de routeur BGP le plus faible. Il s'agit de l'étape 11 de l'algorithme de sélection du meilleur chemin BGP. Résultat : tous les routeurs PE, y compris PE4, choisiront PE1 comme routeur PE de sortie pour le préfixe 172.16.2.0/24. Du point de vue de PE4, le chemin IGP le plus court vers n'importe quel routeur PE de sortie est le chemin vers PE3, avec un coût IGP de 1. Le coût IGP pour tout autre routeur PE est de 2. Pour de nombreux réseaux, le fait de transporter le trafic via le réseau de transit de la manière la plus rapide possible est important. C'est ce que l'on appelle le routage à patate chaude.

Il y a une autre raison possible pour que RR ait choisi le meilleur chemin à partir de PE1. Si dans l'image, le coût IGP (Interior Gateway Protocol) de la liaison P2-PE3 est de 10 et que tous les autres liens ont un coût IGP de 1, alors le RR ne choisirait pas le meilleur chemin de PE3, même si PE3 avait l'ID de routeur BGP le plus faible.

Si l'administrateur de ce réseau souhaite disposer d'un routage à chaud, alors un mécanisme doit être en place pour que, lorsqu'il y a des RR dans le réseau, le routeur d'entrée puisse toujours apprendre le chemin vers le routeur de sortie le plus proche dans le réseau iBGP. La fonction BGP Add Path peut atteindre cet objectif. Cependant, avec cette fonctionnalité, les routeurs RR et les routeurs périphériques doivent avoir un code plus récent qui comprend la fonctionnalité. Avec la fonctionnalité de réflexion de route optimale BGP, ce n'est pas obligatoire. Cette fonctionnalité permettra au RR d'envoyer le meilleur chemin vers le routeur BGP d'entrée, en fonction de ce que le RR pense être le meilleur chemin du point de vue de ce routeur BGP d'entrée.

Une autre solution qui permettrait le routage à chaud lorsque des RR sont déployés est le placement en ligne des RR. Ces RR ne sont pas des RR dédiés, qui exécutent uniquement BGP et IGP. Ces RR en ligne se trouvent sur le chemin de transfert et sont placés dans le réseau de sorte qu'ils aient leur propre ensemble de clients RR, de sorte qu'ils puissent refléter le meilleur chemin vers chaque client RR, qui est également le meilleur chemin du point de vue de ce client RR.

Comme l'illustre cette image, les RR sont placés dans le réseau de sorte qu'ils puissent servir un petit ensemble de clients RR proches. En raison de la conception du réseau, les clients RR reçoivent les meilleurs chemins qui sont les meilleurs de leur point de vue, des RR, de sorte qu'il peut y avoir un routage de patates chaudes dans le réseau.



Théorie

BGP Optimal Route Reflection est décrit dans IETF draft-ietf-idr-bgp-optimal-route-reflection.

La solution BGP Optimal Route Reflection permet au RR d'envoyer un meilleur chemin spécifique vers un routeur de frontière BGP spécifique. Le RR peut choisir d'envoyer un autre meilleur chemin vers différents routeurs de frontière BGP ou un ensemble de routeurs de frontière. Les routeurs périphériques doivent être des clients RR du RR. L'objectif est que chaque routeur BGP d'entrée puisse avoir un routeur BGP de sortie ou de sortie différent pour le même préfixe. Si le routeur de périphérie d'entrée peut toujours transférer le trafic vers le routeur de sortie de système autonome le plus fermé, cela permet le routage de patate chaude.

Le problème est que le RR envoie normalement le même meilleur chemin vers chaque routeur de frontière BGP, ce qui empêche le routage de patate chaude. Afin de résoudre ce problème, vous avez besoin du RR pour pouvoir calculer différents meilleurs chemins pour le même préfixe en fonction du routeur de frontière BGP d'entrée. Le meilleur calcul de chemin sur le RR est effectué en fonction de la position du routeur de frontière BGP d'entrée. Par conséquent, le RR effectuera le meilleur calcul de chemin BGP du point de vue du routeur de frontière d'entrée. Le RR qui peut seulement faire ceci est le RR qui a l'image complète de la topologie du réseau du point de vue de l'IGP où le RR et les routeurs de frontière d'entrée sont situés. Pour que cette condition soit remplie, le protocole IGP doit être un protocole de routage à état de liens.

Dans ce cas, le RR peut exécuter un calcul SPF (Shortest Path First) avec le routeur de frontière d'entrée comme racine de l'arborescence et calculer le coût pour chaque autre routeur. De cette manière, le coût du routeur de frontière d'entrée à tous les autres routeurs de frontière de sortie sera connu. Ce calcul SPF spécial avec un autre routeur en tant que racine est appelé SPF

inverse (rSPF). Cela ne peut être fait que si le RR apprend tous les chemins BGP de tous les routeurs de frontière BGP. Il peut y avoir autant de rSPF exécutés qu'il y a de clients RR. Cela augmentera quelque peu la charge du CPU sur le RR.

La solution permet que le calcul du meilleur chemin soit basé sur l'algorithme de sélection du meilleur chemin BGP, ce qui conduira le RR à choisir le meilleur chemin du point de vue du routeur de frontière d'entrée auquel le RR envoie le chemin. Cela signifie que le meilleur chemin sera choisi en fonction du coût IGP le plus court vers le prochain saut BGP. La solution permet également de choisir le meilleur chemin en fonction d'une stratégie configurée. Les routeurs périphériques d'entrée peuvent choisir leurs meilleurs chemins en fonction d'une politique configurée, et non en fonction du coût IGP le plus bas. La solution permet au RR de mettre en oeuvre la réflexion de route optimale soit sur le coût IGP (emplacement sur le réseau), soit sur une politique configurée, ou les deux. Si les deux sont déployés, la stratégie est appliquée en premier, puis la réflexion de route optimale basée sur IGP se produit sur les chemins restants.

Implémentation IOS-XR

L'implémentation IOS-XR autorise jusqu'à trois noeuds racine pour le calcul rSPF. Si vous avez plusieurs clients RR dans un groupe de mise à jour, alors il n'est pas nécessaire d'effectuer un calcul rSPF par client RR si ces clients RR auront la même politique et/ou les mêmes coûts IGP pour les différents routeurs de périphérie BGP de sortie. Cela signifie généralement que les clients RR sont colocalisés (probablement dans le même POP). Si tel est le cas, il n'est pas nécessaire de configurer chaque client RR en tant que racine. L'implémentation IOS-XR permet de configurer trois, la racine principale, la racine secondaire et la racine tertiaire, par ensemble de clients RR, à des fins de redondance. Pour que la fonctionnalité ORR BGP s'applique à n'importe quel client RR, ce client RR doit être configuré pour faire partie d'un groupe de stratégies ORR.

La fonctionnalité BGP ORR est activée par famille d'adresses.

Un protocole à état de liens est requis. Il peut s'agir d'OSPF ou IS-IS.

IOS XR implémente uniquement la fonctionnalité BGP ORR en fonction du coût IGP au prochain saut BGP, et non en fonction d'une stratégie configurée.

Les homologues BGP ayant la même stratégie de sortie sont placés dans le même groupe de mise à jour. C'est généralement le cas pour iBGP sur le RR. Lorsque la fonctionnalité BGP ORR est activée, les homologues de différents groupes ORR se trouvent dans différents groupes de mise à jour. C'est logique, car les mises à jour envoyées du RR aux clients RR dans différents groupes BGP ORR seront différentes, car le meilleur chemin BGP est différent.

Le résultat des calculs rSPF est stocké dans une base de données.

ORRSPF est le nouveau composant dans IOS-XR qui est nécessaire pour la fonctionnalité ORR BGP. ORRSPF prend soin de :

1. Collecte des informations d'état des liaisons et gestion de la base de données d'état des liaisons
2. Exécution de rSPF et maintenance des SPT, par groupe de politiques
3. Téléchargement des préfixes du SPT vers le RIB à l'aide des métriques

La base de données obtient ses informations d'état de liens directement à partir de l'IGP d'état de liens ou de BGP-LS.

Les calculs rSPF aboutissent à une topologie indiquant le chemin le plus court entre le client RR et tout autre routeur de la zone/niveau.

Les routes qui raccrochent chaque routeur de la topologie sont stockées dans une table RIB spéciale pour la stratégie de groupe ORR et par AFI/SAFI. Cette table est créée par RSI. La table est remplie par les routes calculées par les rSPFs avec la racine principale comme racine. Si la racine principale devient indisponible, la racine secondaire est la racine et remplit les routes dans la table RIB ORR. Il en va de même pour la racine tertiaire.

Configuration

Configuration minimale requise :

1. La fonction ORR doit être activée pour la famille d'adresses de BGP, pour des groupes spécifiques de voisins BGP
2. Pour chaque groupe de voisins BGP, au moins une racine doit être configurée. Vous pouvez éventuellement configurer une racine secondaire et une racine tertiaire.
3. La redistribution des routes ORR à partir de l'IGP dans BGP doit être activée.

Exemple de configuration

Comme l'illustre la première image, le RR est un routeur IOS-XR avec la fonctionnalité BGP ORR.

Tous les autres routeurs exécutent IOS. Ces routeurs ne disposent pas de la fonctionnalité BGP ORR.

PE1, PE2 et PE3 annoncent le préfixe 172.16.2.0/24 dans AFI/SAFI 1/1 (monodiffusion IPv4). Le RR est également proche de PE1 et PE2 que de PE3. Le coût IGP de toutes les liaisons est 1. Le meilleur chemin pour ce préfixe est celui avec R1 comme tronçon suivant en raison de l'ID de routeur BGP le plus faible.

```
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 recevra le chemin avec PE1 comme tronçon suivant. Il n'y a donc pas de routage de patates chaudes pour PE4.

Si vous souhaitez disposer d'un routage de patate chaude sur PE4, alors pour les préfixes annoncés par PE1, PE2 et PE3 (par exemple le préfixe 172.16.2.0/24), PE1 doit avoir PE3 comme point de sortie. Cela signifie que le chemin sur PE4 doit être celui avec PE3 comme tronçon suivant. Nous pouvons faire envoyer la route par RR avec le tronçon suivant PE3 vers PE4 avec cette configuration 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 IGP est 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: L'état de liaison de la famille d'adresses n'a pas besoin d'être configuré, globalement ou sous le ou les voisins BGP.

Ingénierie du trafic MPLS sur routeur racine

Le RR doit trouver l'adresse racine configurée dans la base de données IGP, afin d'exécuter rSPF. Dans ISIS, l'ID de routeur est présent dans la base de données ISIS. Pour OSPF, aucun ID de routeur n'est présent dans les LSA OSPF. La solution consiste à demander aux routeurs racine d'annoncer l'ID de routeur TE MPLS (Multi Protocol Label Switching) correspondant à l'adresse racine configurée sur le RR.

Pour OSPF, les routeurs racine ont besoin d'une configuration supplémentaire pour faire fonctionner BGP ORR. Une configuration MPLS TE minimale est nécessaire sur n'importe quel routeur racine afin d'annoncer cet ID de routeur MPLS TE. L'ensemble de commandes minimal exact dépend du système d'exploitation du routeur racine. La configuration MPLS TE sur le routeur racine doit avoir la configuration minimale pour MPLS TE activée afin qu'OSPF annonce l'ID de routeur MPLS TE dans une LSA de zone opaque (type 10).

Une fois que le RR a une LSA de zone opaque avec l'ID de routeur MPLS TE correspondant à l'adresse de routeur racine configurée, rSPF peut s'exécuter et BGP sur le RR peut annoncer la route optimale.

La configuration minimale requise pour OSPF sur le routeur racine s'il s'agit d'un routeur IOS est la suivante :

```
!
```

```

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
!

```

Notez que :

- MPLS TE est activé dans la zone OSPF spécifique
- L'ID de routeur MPLS TE est configuré pour correspondre à l'adresse racine configurée sur le RR
- L'équipement terminal MPLS est configuré sur au moins une interface
- RSVP-TE n'est pas nécessaire
- Il n'est pas nécessaire de configurer MPLS TE sur un autre routeur de la zone

La configuration minimale requise pour OSPF sur le routeur racine s'il s'agit d'un routeur IOS-XR est la suivante :

```

!
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 configuration ci-dessus est en place sur le routeur racine, RR doit avoir l'ID de routeur TE MPLS dans la base de données 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

Notez que l'ID de routeur MPLS TE (10.100.1.4) et l'ID de routeur OSPF sont différents.

PE4 a PE3 comme tronçon suivant pour le préfixe (avec la métrique IGP correcte au tronçon suivant) :

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 a toujours PE1 comme tronçon suivant pour le préfixe (avec la métrique IGP correcte au tronçon suivant) :

```
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
```

Dépannage

Vérifiez le préfixe du 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
```

Notez que add-path a été ajouté aux autres chemins non optimaux, afin qu'ils puissent également

être annoncés, en plus du meilleur chemin. La fonction d'ajout de chemin n'est pas utilisée entre le RR et ses clients : les chemins ne sont pas annoncés avec un identificateur de chemin.

Vérifiez que les routes sont (toujours) annoncées aux voisins BGP spécifiques.

Pour le voisin PE4, le tronçon suivant est PE3 pour le préfixe 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

Pour le voisin PE5, le tronçon suivant est PE1 pour le préfixe 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
```

Le voisin 10.100.1.4 se trouve dans son propre groupe de mises à jour en raison de la stratégie ORR en place :

```
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
```

La commande show orrspf database affiche le groupe ORR et ses racines,

```
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
```

La même commande avec le mot clé detail fournit le coût de la racine du rSPF à l'autre routeur/préfixe dans la même zone 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
```

L'ID de table a été attribué par RSI pour le groupe ORR et pour l'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

Le coût de la racine (R4/10.100.1.4) du rSPF entre les routeurs est le même que le coût qui est affiché avec **show ip route ospf** sur 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

Le RIB du groupe 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

La commande **show bgp neighbor** indique si l'homologue est une racine 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

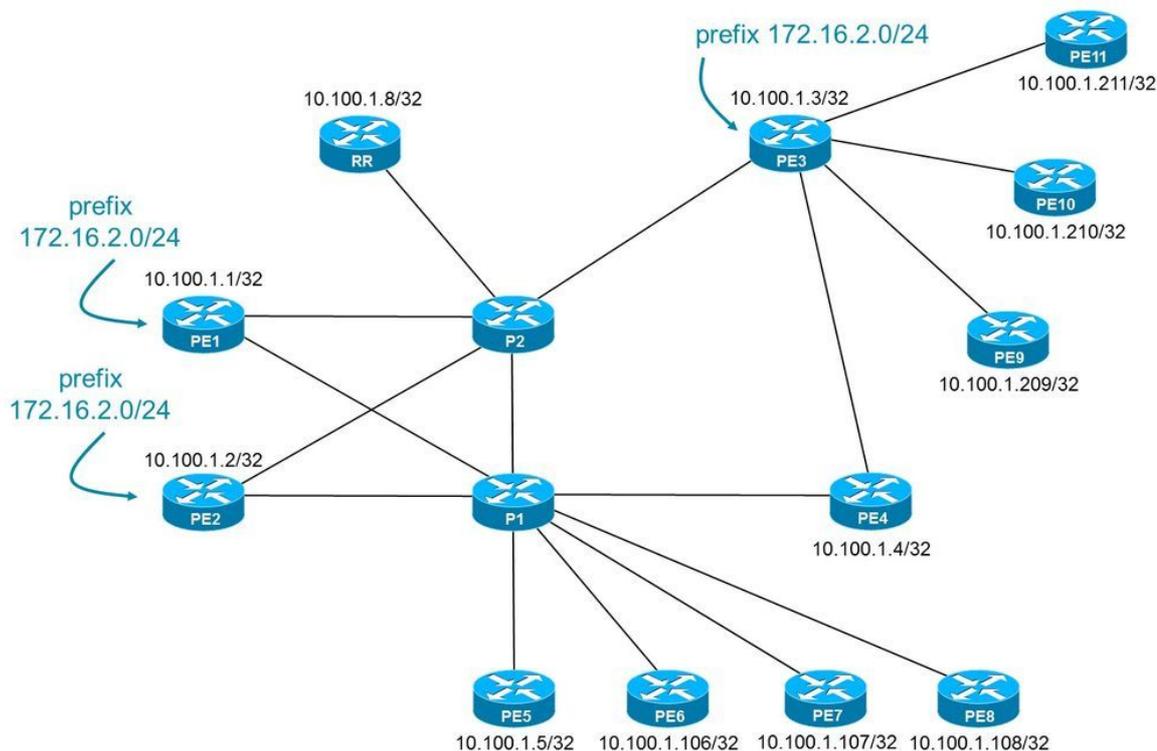
Last ack version 41, Last synced ack version 0

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

Comme le montre cette image, plusieurs ensembles de clients RR configurés



Il existe un ensemble de clients RR connectés à PE3 et un autre ensemble connecté à P1. Chaque client RR de chaque ensemble est à la même distance que n'importe quel routeur de frontière BGP de sortie.

```
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
!
!
!

```

Base de données orrspf pour les deux groupes :

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, pour un groupe, la racine principale est arrêtée ou inaccessible, la racine secondaire sera la racine réelle utilisée. Dans cet exemple, la racine principale du groupe ipv4-orr-group-1 est inaccessible. La racine secondaire est devenue la racine réelle :

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

Conclusion

La fonction de réflexion de route optimale BGP (ORR) permet le routage à chaud dans un réseau iBGP lorsque des réflecteurs de route sont présents, sans avoir besoin de logiciel de système d'exploitation plus récent sur les routeurs périphériques. La condition préalable est que le protocole IGP est un protocole de routage à état de liens.