

Border Gateway Protocol (BGP) - Optimale Routen-Reflektion

Inhalt

[Einführung](#)

[Hintergrundinformationen](#)

[Netzwerkdiagramm](#)

[Theorie](#)

[IOS-XR-Implementierung](#)

[Konfigurieren](#)

[Konfigurationsbeispiel](#)

[MPLS Traffic Engineering auf Root-Router](#)

[Fehlerbehebung](#)

[Schlussfolgerung](#)

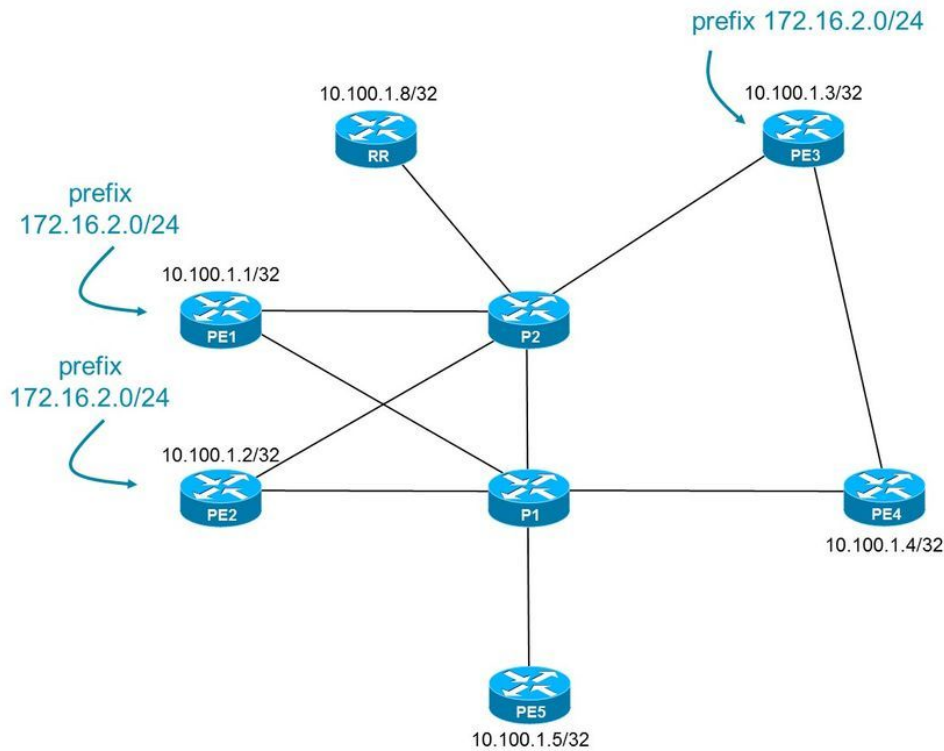
Einführung

In diesem Dokument wird beschrieben, wie das Routing beeinflusst werden kann, wenn ein oder mehrere Routen-Reflektoren (Route Reflectors, RRs) im Netzwerk vorhanden sind, um ein Full-Mesh zwischen iBGP-Routern zu vermeiden.

Hintergrundinformationen

Schritt 8 im [BGP-Algorithmus zur Pfadauswahl](#) besteht darin, den Pfad mit der niedrigsten IGP-Metrik dem BGP Next Hop vorzuziehen. Wenn also alle Schritte vor Schritt 8 gleich sind, kann Schritt 8 der entscheidende Faktor für den besten Pfad auf dem RR sein. Die IGP-Kosten vom RR zum iBGP-Werberouter werden durch die Platzierung des RR bestimmt. Standardmäßig kündigt der RR seinen Clients nur den besten Pfad an. Je nachdem, wo der RR platziert wird, können die IGP-Kosten für den Werberouter kleiner oder größer sein. Wenn alle IGP-Kosten für die Pfade identisch sind, wird dies wahrscheinlich zum Grenzwert des Werberouters führen, der über die niedrigste BGP-Router-ID verfügt.

Netzwerkdiagramm



Die Router PE1, PE2 und PE3 geben das Präfix 172.16.2.0/24 bekannt. Wenn alle IGP-Kosten für die Verbindungen identisch sind, wird dem RR der Pfad von PE1, PE2 und PE3 mit einem IGP-Preis von 2 angezeigt. Am Ende wählt der RR den Pfad von PE1 aus, da er über die niedrigere BGP-Router-ID verfügt. Dies ist Schritt 11 des BGP-Algorithmus zur Pfadauswahl. Das Ergebnis ist, dass alle PE-Router, einschließlich PE4, PE1 als Egress-PE-Router für das Präfix 172.16.2.0/24 auswählen. Vom PE4-Standpunkt aus ist der kürzere IGP-Pfad zu einem beliebigen Egress-PE-Router der Pfad zu PE3 mit einem IGP-Preis von 1. Die IGP-Kosten für einen anderen PE-Router sind 2. Für viele Netzwerke ist es wichtig, den Verkehr so schnell wie möglich durch das Transit-Netzwerk zu transportieren. Dies wird als Hot-Potato-Routing bezeichnet.

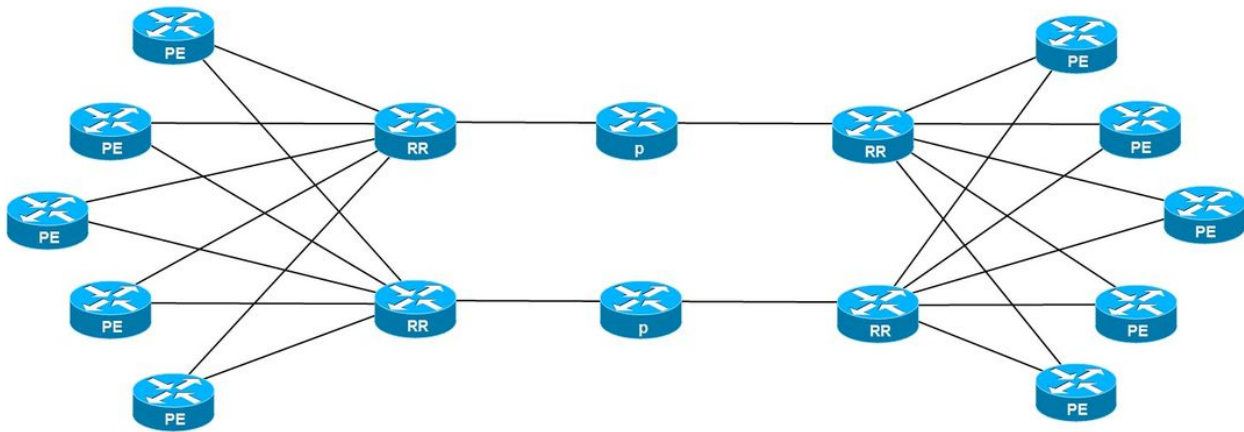
Es gibt einen weiteren Grund dafür, dass RR den besten Pfad von PE1 ausgewählt hat. Wenn im Bild die IGP-Kosten (Interior Gateway Protocol) für die Verbindung P2-PE3 10 und für alle anderen Verbindungen immer noch die IGP-Kosten 1 betragen, wählt der RR den Pfad von PE3 nicht als den besten aus, selbst wenn PE3 die niedrigste BGP-Router-ID aufweist.

Wenn der Administrator dieses Netzwerks Hot-Potato-Routing nutzen möchte, muss ein Mechanismus vorhanden sein, damit der Eingangs-Router den Pfad zum nächsten Ausgangs-Router im iBGP-Netzwerk trotzdem erkennen kann, wenn RRs im Netzwerk vorhanden sind. Dies kann durch die BGP-Funktion Add Path erreicht werden. Bei dieser Funktion müssen die RRs und die Grenzurter jedoch über neueren Code verfügen, der die Funktion versteht. Mit der Funktion "BGP Optimal Route Reflection" ist dies nicht erforderlich. Diese Funktion ermöglicht dem RR, den besten Pfad zum BGP-Grenzurter des Eingangs zu senden, je nachdem, was der RR für den besten Pfad aus Sicht dieses Eingangs-BGP-Routers hält.

Eine weitere Lösung, die Hot-Potato-Routing bei der Bereitstellung von RRs ermöglicht, ist die In-Line-Platzierung von RRs. Bei diesen RRs handelt es sich nicht um dedizierte RRs, auf denen nur BGP und das IGP ausgeführt werden. Diese Inline-RRs befinden sich im Weiterleitungspfad und werden im Netzwerk platziert, sodass sie über eigene RR-Clients verfügen, sodass sie den besten Pfad zu jedem RR-Client wiedergeben können, der auch aus der Sicht dieses RR-Clients der beste Pfad ist.

Wie in diesem Bild gezeigt, werden die RRs im Netzwerk platziert, sodass sie eine kleine Gruppe

nahegelegener RR-Clients bedienen können. Aufgrund des Netzwerkdesigns erhalten die RR-Clients die besten Pfade, die aus ihrer Sicht die besten Pfade sind, von den RRs, sodass im Netzwerk Hot-Potato-Routing möglich ist.



Theorie

Die optimale Routen-Reflexion für das BGP wird in IETF Draft-ietf-idr-bgp-optimales-route-reflexion beschrieben.

Mit der BGP-Lösung für optimale Routen-Reflexion kann der RR einen bestimmten besten Pfad an einen bestimmten BGP-Border-Router senden. Der RR kann einen anderen besten Pfad an verschiedene BGP-Border-Router oder einen Satz von Border-Routern senden. Bei den Grenzroutern muss es sich um RR-Clients des RRs handeln. Das Ziel besteht darin, dass jeder BGP-Grenz-Router für eingehende Anrufe über einen anderen BGP-Router für Ausgang oder Ausgang für dasselbe Präfix verfügen kann. Wenn der Eingangs-Border-Router den Datenverkehr immer an den geschlossenen AS-Exit-Router weiterleiten kann, ermöglicht dies Hot-Potato-Routing.

Das Problem besteht darin, dass der RR normalerweise nur den gleichen besten Pfad zu jedem BGP-Border-Router sendet, wodurch Hot-Potato-Routing verhindert wird. Um dieses Problem zu lösen, muss der RR in der Lage sein, je nach dem BGP-Grenz-Eingangs-Router unterschiedliche Pfade für dasselbe Präfix zu berechnen. Die Berechnung des besten Pfads im RR erfolgt auf Basis der Position des BGP-Eingangs-BGP-BGP-Border-Routers. Daher berechnet der RR den besten BGP-Pfad aus Sicht des Eingangs-Border-Routers. Der RR, der dies nur tun kann, ist der RR, der aus IGP-Sicht, wo sich der RR und der/die Eingangs-Border Router befinden, ein vollständiges Bild der Netzwerktopologie hat. Damit diese Anforderung erfüllt werden kann, muss das IGP ein Link-State-Routing-Protokoll sein.

In diesem Fall kann der RR eine SPF-Berechnung (Shortest Path First) mit dem Eingangs-Border-Router als Root der Struktur ausführen und die Kosten für jeden anderen Router berechnen. Auf diese Weise sind die Kosten vom Eingangs-Border-Router bis zu allen anderen Ausgangs-Border-

Routern bekannt. Diese spezielle SPF-Berechnung mit einem anderen Router als Root wird als Reverse SPF (rSPF) bezeichnet. Dies ist nur möglich, wenn der RR alle BGP-Pfade von allen BGP-Grenzroutern erfasst. Es können so viele rSPFs ausgeführt werden, wie es RR-Clients gibt. Dadurch wird die CPU-Last beim RR etwas erhöht.

Die Lösung ermöglicht, dass die Berechnung des besten Pfads auf dem BGP-Algorithmus zur Auswahl des besten Pfads basiert. Dies führt dazu, dass der RR den besten Pfad aus Sicht des Eingangs-Border-Routers auswählt, an den der RR den Pfad sendet. Dies bedeutet, dass der beste Pfad basierend auf den kürzesten IGP-Kosten für den nächsten BGP-Hop ausgewählt wird. Die Lösung ermöglicht außerdem die Auswahl des besten Pfades basierend auf konfigurierten Richtlinien. Die Router an den Eingangsgrenzen können ihre besten Pfade basierend auf einer konfigurierten Richtlinie auswählen, nicht auf den niedrigsten IGP-Kosten. Die Lösung ermöglicht dem RR, die optimale Routen-Reflektion entweder über die IGP-Kosten (Standort im Netzwerk) oder über eine konfigurierte Richtlinie oder beides zu implementieren. Wenn beide bereitgestellt werden, wird zuerst die Richtlinie angewendet, und dann erfolgt die optimale IGP-basierte Routen-Reflektion auf den verbleibenden Pfaden.

IOS-XR-Implementierung

Die IOS-XR-Implementierung ermöglicht die rSPF-Berechnung auf bis zu drei Root-Knoten. Wenn Sie viele RR-Clients in einer Aktualisierungsgruppe haben, ist keine rSPF-Berechnung pro RR-Client erforderlich, wenn diese RR-Clients dieselbe Richtlinie und/oder dieselben IGP-Kosten für die verschiedenen BGP-Grenz-Router für Ausgangs-BGP haben. Letzteres bedeutet in der Regel, dass die RR-Clients am gleichen Standort arbeiten (wahrscheinlich im selben POP). In diesem Fall muss nicht jeder RR-Client als Root konfiguriert werden. Die IOS-XR-Implementierung ermöglicht die Konfiguration von drei, den primären, den sekundären und den tertiären Root-Clients pro RR-Client, um Redundanzzwecke zu ermöglichen. Damit die BGP ORR-Funktion auf jeden RR-Client angewendet werden kann, muss dieser RR-Client als Teil einer ORR-Richtliniengruppe konfiguriert werden.

Die BGP ORR-Funktion ist pro Adressfamilie aktiviert.

Ein Link-State-Protokoll ist erforderlich. Dabei kann es sich um OSPF oder IS-IS handeln.

IOS XR implementiert die BGP ORR-Funktion nur basierend auf den IGP-Kosten für den BGP Next Hop und nicht basierend auf einer konfigurierten Richtlinie.

Die BGP-Peers mit derselben Richtlinie für ausgehenden Datenverkehr werden in derselben Aktualisierungsgruppe angeordnet. Dies ist in der Regel bei iBGP auf dem RR der Fall. Wenn die Funktion BGP ORR aktiviert ist, befinden sich die Peers verschiedener ORR-Gruppen in unterschiedlichen Aktualisierungsgruppen. Dies ist logisch, da sich die Aktualisierungen, die vom RR an die RR-Clients in verschiedenen BGP ORR-Gruppen gesendet werden, unterscheiden, da der beste BGP-Pfad unterschiedlich ist.

Das Ergebnis der rSPF-Berechnungen wird in einer Datenbank gespeichert.

ORRSPF ist die neue Komponente in IOS-XR, die für die BGP ORR-Funktion benötigt wird. ORRSPF übernimmt die Pflege:

1. Erfassen der Link-State-Informationen und Pflegen der Link-State-Datenbank
2. Ausführen von rSPFs und Pflegen der SPTs pro Richtliniengruppe
3. Herunterladen der Präfixe von der SPT zur RIB mit den Metriken

Die Datenbank erhält ihre Link-State-Informationen entweder direkt vom Link-State IGP oder vom BGP-LS.

Die rSPF-Berechnungen führen zu einer Topologie, die den kürzesten Pfad vom RR-Client zu einem anderen Router im Bereich/auf der Ebene anzeigt.

Die Routen, die von jedem Router in der Topologie hängen, werden in einer speziellen RIB-Tabelle für die ORR-Gruppenrichtlinie und pro AFI/SAFI gespeichert. Diese Tabelle wird von RSI erstellt. Die Tabelle wird durch die von den rSPFs berechneten Routen aufgefüllt, wobei der primäre Root als Root festgelegt wird. Wenn der primäre Root nicht verfügbar ist, ist der sekundäre Root der Root und füllt die Routen in die ORR RIB-Tabelle auf. Dasselbe gilt für die tertiäre Wurzel.

Konfigurieren

Minimale erforderliche Konfiguration:

1. ORR muss für die Adressfamilie von BGP und für bestimmte Gruppen von BGP-Nachbarn aktiviert werden.
2. Für jede Gruppe von BGP-Nachbarn muss mindestens ein Root konfiguriert werden. Optional kann ein sekundärer und tertiärer Root konfiguriert werden.
3. Die Neuverteilung der ORR-Routen vom IGP zum BGP muss aktiviert werden.

Konfigurationsbeispiel

Wie im ersten Bild gezeigt, ist der RR ein IOS-XR-Router mit der BGP ORR-Funktion.

Auf allen anderen Routern wird IOS ausgeführt. Diese Router verfügen nicht über die BGP ORR-Funktion.

PE1, PE2 und PE3 geben das Präfix 172.16.2.0/24 in AFI/SAFI 1/1 (IPv4-Unicast) an. Der RR ist ebenso nah an PE1 und PE2 wie an PE3. Die IGP-Kosten für alle Verbindungen sind 1. Der beste Pfad für dieses Präfix ist derjenige mit R1 als Next-Hop aufgrund der niedrigsten BGP-Router-ID.

```
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 empfängt den Pfad mit PE1 als Next-Hop. Es gibt also kein Hot-Potato-Routing für PE4.

Wenn für PE4 Hot-Potato-Routing erforderlich ist, dann sollte für die von PE1, PE2 und PE3 angekündigten Präfixe (z. B. das Präfix 172.16.2.0/24) PE1 als Ausgangspunkt PE3 haben. Dies bedeutet, dass der Pfad auf PE4 derjenige mit PE3 als Next-Hop sein muss. Mit dieser ORR-Konfiguration kann der RR die Route mit Next-Hop-PE3 an PE4 senden.

```

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
!
!
!

```

Wenn das IGP IS-IS ist:

```

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

```

Hinweis: Der Verbindungsstatus der Adressfamilie muss nicht global oder unter den BGP-Nachbarn konfiguriert werden.

MPLS Traffic Engineering auf Root-Router

Der RR muss die konfigurierte Root-Adresse in der IGP-Datenbank finden, um den rSPF auszuführen. In ISIS ist die Router-ID in der ISIS-Datenbank vorhanden. Für OSPF ist in den OSPF-LSAs keine Router-ID vorhanden. Die Lösung besteht darin, dass die Root-Router die MPLS-TE-Router-ID (Multi Protocol Label Switching) ankündigen, die mit der konfigurierten Root-Adresse des RR übereinstimmt.

Für OSPF benötigen die Root-Router eine zusätzliche Konfiguration, damit BGP ORR funktioniert. Für jeden Root-Router ist eine minimale MPLS-TE-Konfiguration erforderlich, um diese MPLS-TE-Router-ID anzukündigen. Der genaue minimale Befehlssatz hängt vom Betriebssystem des Root-Routers ab. Bei der MPLS-TE-Konfiguration auf dem Root-Router muss die Mindestkonfiguration für MPLS TE aktiviert sein, sodass OSPF die MPLS-TE-Router-ID in einem LSA mit deckendem Bereich (Typ 10) ankündigt.

Sobald der RR über ein opak-Area LSA mit einer MPLS-TE-Router-ID verfügt, die mit der konfigurierten Root-Router-Adresse übereinstimmt, kann rSPF ausgeführt werden, und das BGP auf dem RR kann die optimale Route ankündigen.

Die erforderliche Mindestkonfiguration für OSPF auf dem Root-Router, wenn es sich um einen IOS-Router handelt, lautet wie folgt:

```

!
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
!

```

Beachten Sie, dass

- MPLS TE ist im spezifischen OSPF-Bereich aktiviert.
- Die MPLS-TE-Router-ID wird mit der konfigurierten Root-Adresse des RR konfiguriert.
- MPLS-TE wird auf mindestens einer Schnittstelle konfiguriert.
- RSVP-TE muss nicht konfiguriert werden.
- MPLS TE muss auf keinem anderen Router im Bereich konfiguriert werden.

Bei einem IOS-XR-Router ist die minimale Konfiguration für OSPF auf dem Root-Router erforderlich:

```

!
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
!

```

Wenn die obige Konfiguration auf dem Root-Router implementiert ist, muss die RR-Einheit die MPLS-TE-Router-ID in der OSPF-Datenbank haben.

```
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

Beachten Sie, dass sich die MPLS-TE-Router-ID (10.100.1.4) und die OSPF-Router-ID unterscheiden.

PE4 hat PE3 als Next-Hop für das Präfix (mit der richtigen IGP-Metrik für den nächsten Hop):

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 hat weiterhin PE1 als Next-Hop für das Präfix (mit der richtigen IGP-Metrik für den nächsten Hop):

```
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
```

Fehlerbehebung

Überprüfen Sie das Präfix auf dem 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
```

Beachten Sie, dass Add-Path zu den anderen nicht besten Pfaden hinzugefügt wurde, sodass sie

neben dem besten Pfad auch angekündigt werden können. Die Funktion zum Hinzufügen von Pfaden wird zwischen dem RR und seinen Clients nicht verwendet: Die Pfade werden nicht mit einer Pfadkennung angekündigt.

Stellen Sie sicher, dass die Route(n) (noch) den spezifischen BGP-Nachbarn angekündigt wird(en).

Für den Nachbarn PE4 ist der Next-Hop PE3 für das Präfix 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

Für den Nachbarn PE5 ist der Next-Hop PE1 für das Präfix 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
```

Der Nachbar 10.100.1.4 befindet sich aufgrund der geltenden ORR-Richtlinie in seiner eigenen Aktualisierungsgruppe:

```
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

```

Der Befehl `show orrspf database` zeigt die ORR-Gruppe und deren Root(s),

```
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
```

Der gleiche Befehl mit dem detail-Schlüsselwort stellt die Kosten des Roots des rSPF für den anderen Router/das Präfix im gleichen OSPF-Bereich bereit:

```
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
```

Die Tabelle-ID wurde vom RSI für die ORR-Gruppe und für AFI/SAFI zugewiesen:

```
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

Die Kosten für den Root (R4/10.100.1.4) des rSPF zu den anderen Routern entsprechen den Kosten, die bei **show ip route ospf** auf PE4 festgestellt werden:

PE4#**show ip route ospf**

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

Gateway of last resort is not set

10.0.0.0/8 is variably subnetted, 20 subnets, 2 masks

- O 10.100.1.1/32 [110/3] via 10.1.7.6, 4d05h, GigabitEthernet0/1
- O 10.100.1.2/32 [110/3] via 10.1.7.6, 4d05h, GigabitEthernet0/1
- O 10.100.1.3/32 [110/2] via 10.1.8.3, 4d06h, GigabitEthernet0/2
- O 10.100.1.5/32 [110/3] via 10.1.7.6, 4d05h, GigabitEthernet0/1
- O 10.100.1.6/32 [110/2] via 10.1.7.6, 4d05h, GigabitEthernet0/1
- O 10.100.1.7/32 [110/3] via 10.1.8.3, 4d06h, GigabitEthernet0/2
- O 10.100.1.8/32 [110/4] via 10.1.8.3, 4d05h, GigabitEthernet0/2
- O 10.100.1.9/32 [110/3] via 10.1.7.6, 4d05h, GigabitEthernet0/1
- O 10.100.1.10/32 [110/2] via 10.1.7.6, 4d05h, GigabitEthernet0/1
- O 10.100.1.11/32 [110/3] via 10.1.8.3, 4d06h, GigabitEthernet0/2
- O 10.100.1.12/32 [110/3] via 10.1.7.6, 4d05h, GigabitEthernet0/1
- O 10.100.1.13/32 [110/2] via 10.1.7.6, 4d05h, GigabitEthernet0/1
- O 10.100.1.14/32 [110/3] via 10.1.8.3, 4d06h, GigabitEthernet0/2
- O 10.100.1.15/32 [110/3] via 10.1.7.6, 4d05h, GigabitEthernet0/1
- O 10.100.1.16/32 [110/2] via 10.1.7.6, 4d05h, GigabitEthernet0/1
- O 10.100.1.17/32 [110/3] via 10.1.8.3, 4d06h, GigabitEthernet0/2
- O 10.100.1.18/32 [110/3] via 10.1.7.6, 4d05h, GigabitEthernet0/1
- O 10.100.1.19/32 [110/2] via 10.1.7.6, 4d05h, GigabitEthernet0/1
- O 10.100.1.20/32 [110/3] via 10.1.8.3, 4d06h, GigabitEthernet0/2

Die RIB für die BGP ORR-Gruppe:

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

Der Befehl show bgp neighbor zeigt, ob der Peer ein ORR-Root ist:

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

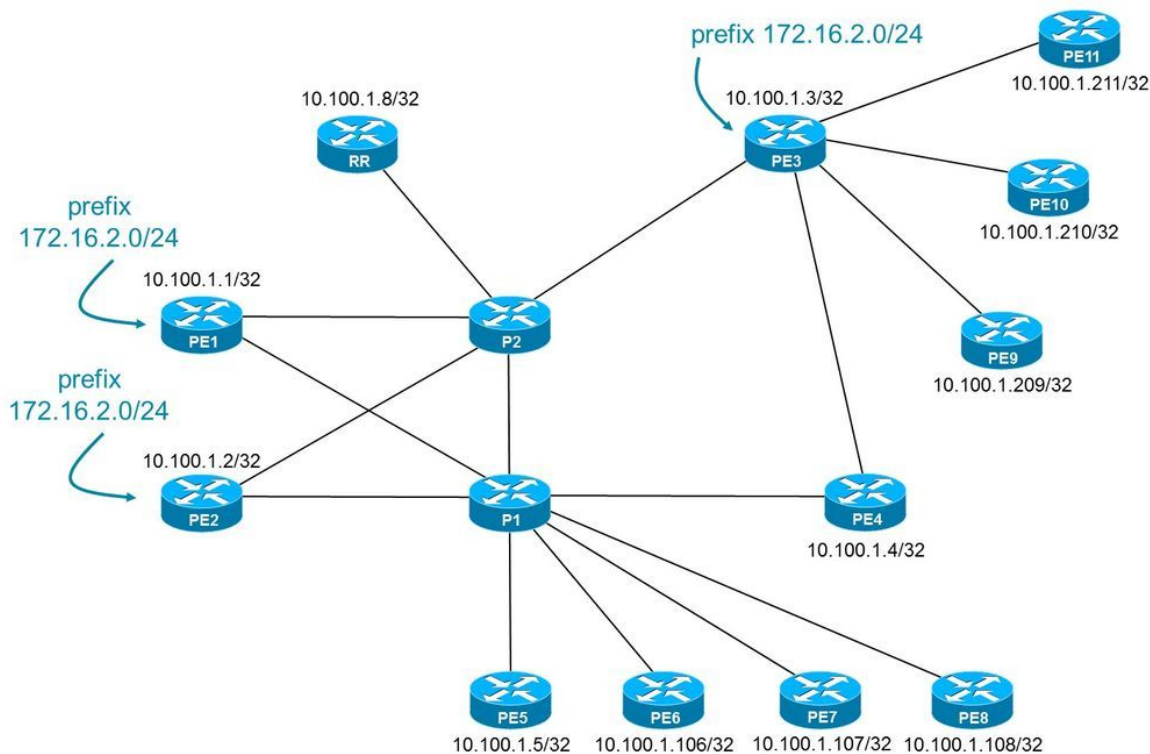
BGP neighbor is 10.100.1.4
Remote AS 1, local AS 1, internal link
Remote router ID 10.100.1.4
Cluster ID 10.100.1.8
BGP state = Established, up for 01:17:41
NSR State: None
Last read 00:00:52, Last read before reset 01:18:30
Hold time is 180, keepalive interval is 60 seconds
Configured hold time: 180, 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

Wie in diesem Bild gezeigt, wurden mehrere RR-Clients konfiguriert.



Ein RR-Client ist mit PE3 verbunden, ein weiterer mit P1 verbunden. Jeder RR-Client in jedem Set befindet sich in gleicher Entfernung zu jedem BGP-Grenz-Router am Ausgang.

```
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
!
!
!

```

Die orrspf-Datenbank für beide Gruppen:

```
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

Wenn für eine Gruppe der primäre Root ausgefallen oder nicht erreichbar ist, wird der sekundäre Root als tatsächlich verwendeter Root verwendet. In diesem Beispiel ist der primäre Root der Gruppe ipv4-orr-group-1 nicht erreichbar. Der sekundäre Root wurde zur eigentlichen Root:

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

Schlussfolgerung

BGP Optimal Route Reflection (ORR) ist eine Funktion, die Hot-Potato-Routing in einem iBGP-Netzwerk ermöglicht, wenn Routen-Reflektoren vorhanden sind, ohne dass neuere Betriebssystemsoftware auf den Grenzroutern benötigt wird. Voraussetzung ist, dass das IGP ein Link-State-Routing-Protokoll ist.