Preferred BGP Route Selection

Foreword

- The border gateway protocol (BGP) is a widely deployed routing protocol in the globe. BGP defines multiple path attributes and has various routing policy tools, providing flexible route control and path selection.
- Operations on BGP route attributes may affect route selection and therefore affect network traffic. Therefore, it is important to master Rules for Selecting a Preferred BGP Route.
- This course illustrates Rules for Selecting a Preferred BGP Route.

Objectives

- Upon completion of this course, you will be able to:
  - Describe rules for selecting a preferred BGP route.
  - Learn BGP route control.

Contents

Rules for Selecting a Preferred BGP Route

When multiple routes to the same destination network segment exist, BGP selects routes in the following sequence:

Discards the route whose next hop is unreachable.

1. Prefers the route with the largest Preferred-Value attribute value.
2. Prefers the route with the largest Local_Preference value.

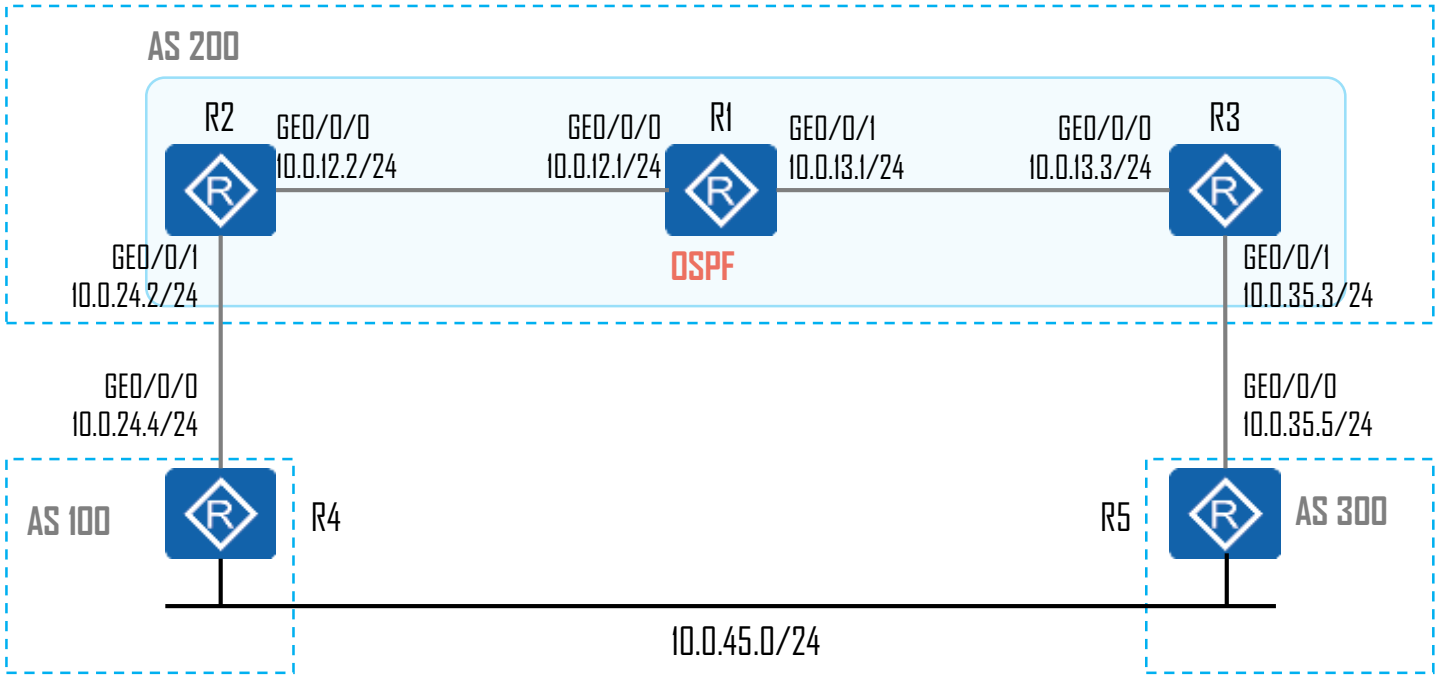
↑ A larger value indicates a better route.

3. Prefers the locally originated BGP route, which takes precedence over the route learned from a peer. The locally summarized route, automatically summarized route, route learned by using the **network** command, route learned by using the **import-route** command, and route learned from a peer are in descending order of priority.
4. Prefers the route with the shortest AS_Path.
5. Prefers the route with the shortest AS_Path. The routes with Origin attributes of IGP, EGP, and Incomplete are in descending order of priority.
6. Prefers the route with the lowest MED.
7. Prefers routes learned from EBGP peers to routes learned from IBGP peers.
8. Prefers the route with the smallest IGP metric to the next hop.
9. Prefers the route with the shortest Cluster_List length.
10. Prefers the route advertised by the device with the smallest router ID (Originator_ID).
11. Prefers the route learned from the peer with the smallest IP address.

↓ A smaller value indicates a better route.

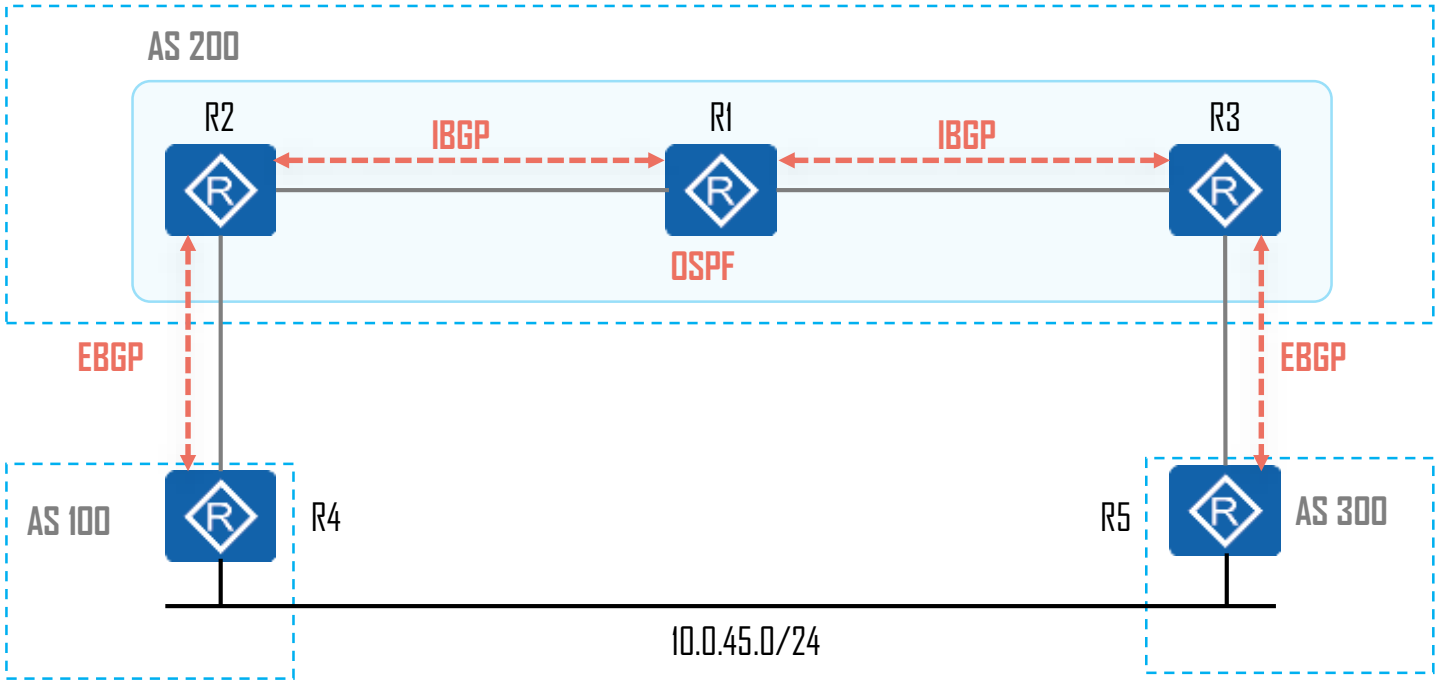
If the preceding eight attributes are the same, routes work in load balancing mode.

Topology (1)



- The figure shows ASs and interconnection addresses. Loopback0 interfaces are created on all devices, and the IP address is 10.0.x.x (x indicates the device ID). All devices use addresses of Loopback0 interfaces as router IDs.
- OSPF runs in AS 200 and OSPF is enabled on internal interconnection interfaces (excluding the interfaces connected to external AS) and loopback interfaces.

Topology (2)



- Loopback0 interfaces are used for establishing IBGP peer relationships in ASs, and directly connected interfaces are used for establishing EBGP peer relationships between ASs.
- R4 and R5 have the same network segment 10.0.45.0/24. The import-route command can be used to import the direct routes of this network segment to BGP so as to verify BGP route selection.

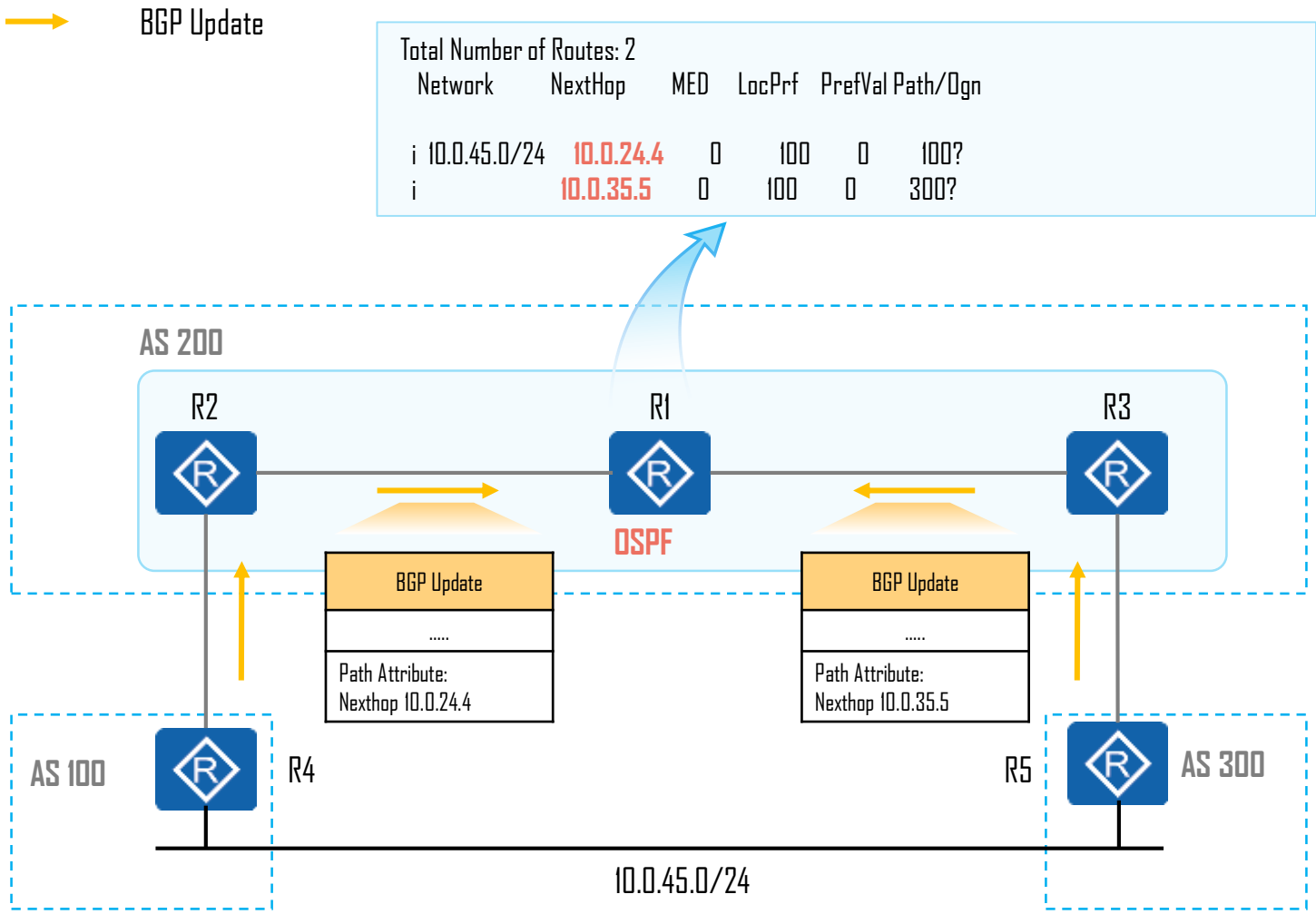
Rules for Selecting a Preferred BGP Route

When multiple routes to the same destination network segment exist, BGP selects routes in the following sequence:

Discards the route whose next hop is unreachable.

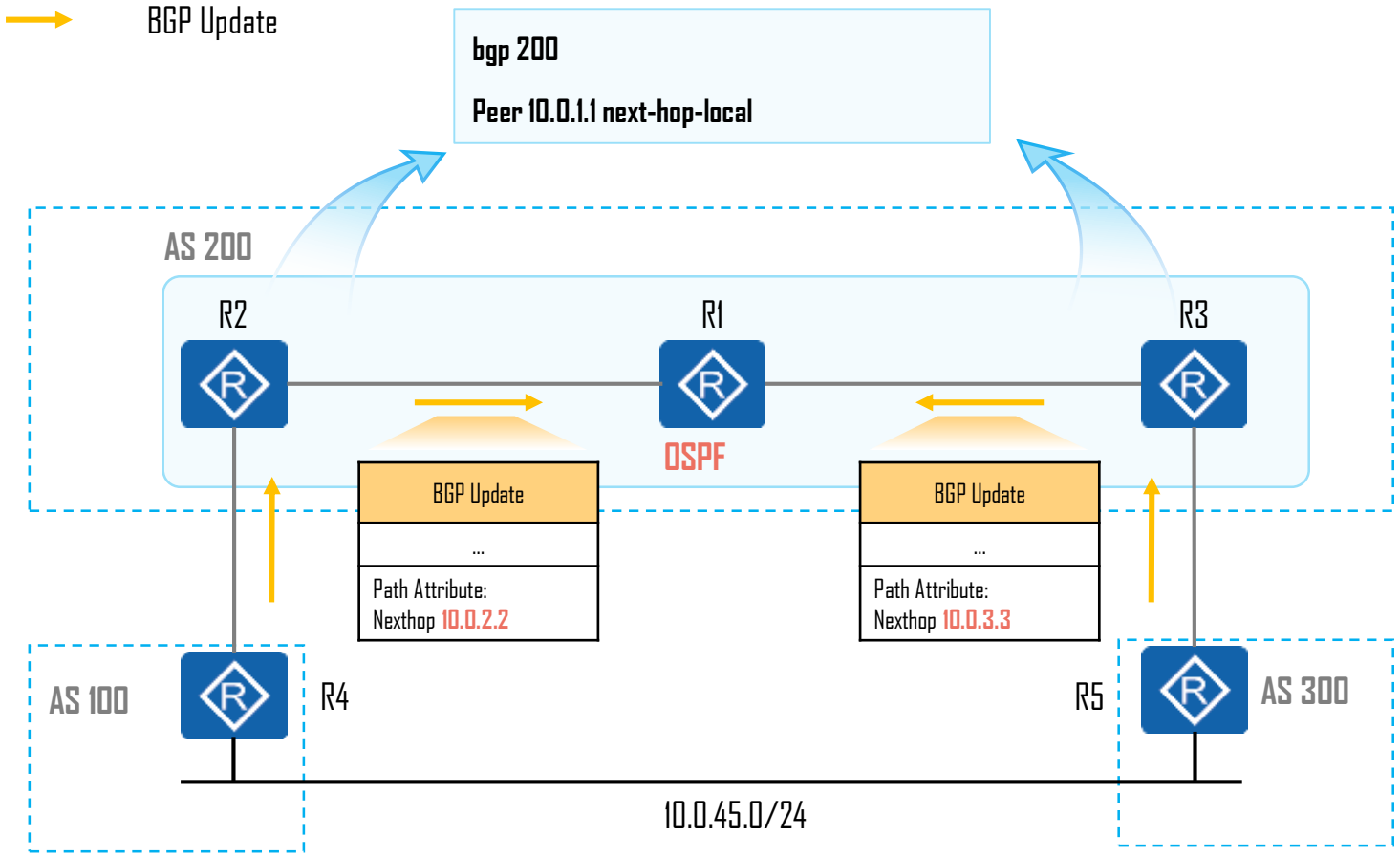
1. Prefers the route with the largest Preferred-Value attribute value.
2. Prefers the route with the largest Local_Preference value.
3. Prefers the locally originated BGP route, which takes precedence over the route learned from a peer. The locally summarized route, automatically summarized route, route learned by using the **network** command, route learned by using the **import-route** command, and route learned from a peer are in descending order of priority.
4. Prefers the route with the shortest AS_Path.
5. Prefers the route with the shortest AS_Path. The routes with Origin attributes of IGP, EGP, and Incomplete are in descending order of priority.
6. Prefers the route with the lowest MED.
7. Prefers routes learned from EBGP peers to routes learned from IBGP peers.
8. Prefers the route with the smallest IGP metric to the next hop.
9. Prefers the route with the shortest Cluster_List length.
10. Prefers the route advertised by the device with the smallest router ID (Originator_ID).
11. Prefers the route learned from the peer with the smallest IP address.

Discarding the Route Whose Next Hop Is Unreachable (1)



- When R4 and R5 advertise the BGP routes 10.0.45.0/24 to AS 200, the Next_Hop attribute values of the routes are 10.0.24.4 and 10.0.34.5.
- R2 and R3 do not modify the Next_Hop attribute when advertising routes to R1. The next hops of the two BGP routes 10.0.45.0/24 learned by R1 are 10.0.24.4 and 10.0.34.5.
- When R1 performs recursive query for next hops of BGP routes, route recursion fails because OSPF is not activated on the interfaces connecting R2 and R3 to external ASs. As a result, the next hop of the BGP route 10.0.45.0/24 on R1 is unreachable.
- Run the **display bgp routing** command on R1 to check the BGP routing table. The command output shows that the BGP route 10.0.45.0/24 is invalid.

Discarding the Route Whose Next Hop Is Unreachable (2)



- Run the **next-hop-local** command on R2 and R3 to change the Next_Hop attribute value to the local source address.
- When R2 and R3 advertise BGP routes to R1, the Next_Hop attribute values of the routes are changed to 10.0.2.2 and 10.0.3.3.
- The two next-hop addresses can be successfully recursed on R1, and the next-hop address of the BGP route becomes reachable.

Unless otherwise specified, devices in all subsequent cases use basic configuration. In addition to the basic configuration, R2 and R3 are configured with **peer next-hop-local**.

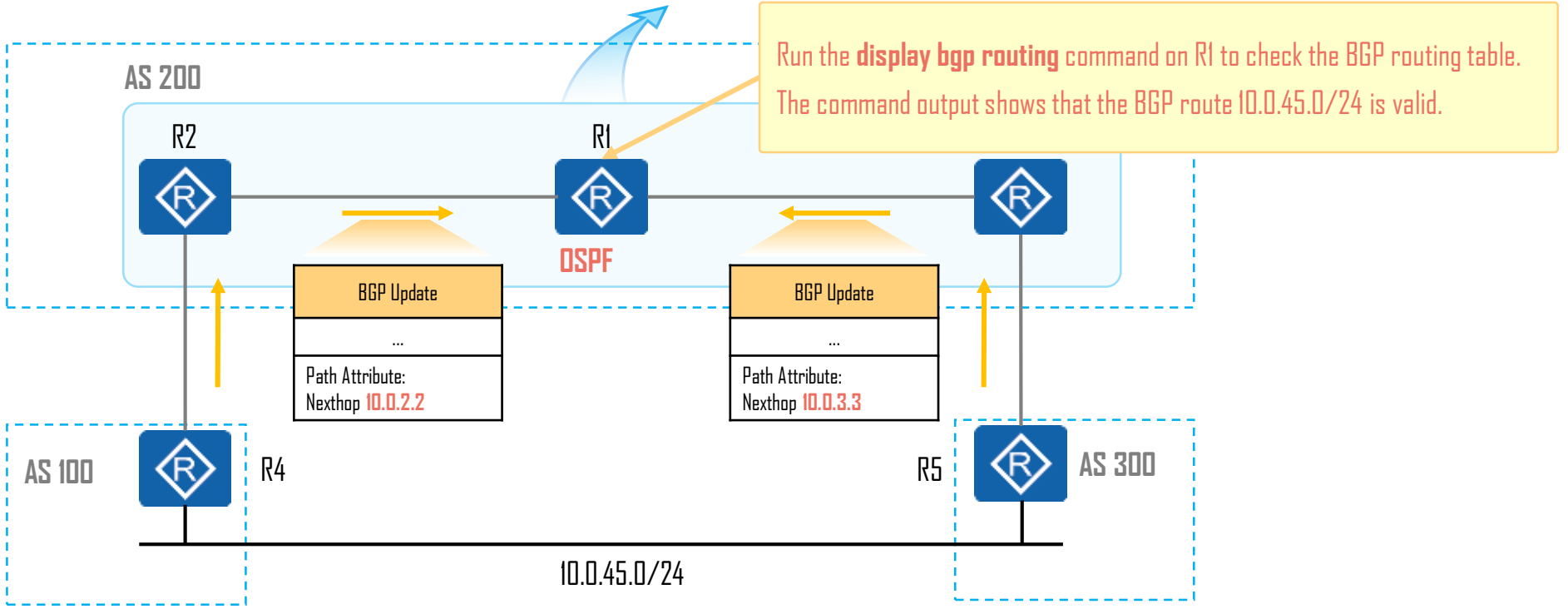
Discarding the Route Whose Next Hop Is Unreachable (3)

BGP Update

```
Total Number of Routes: 2
     Network         NextHop    MED    LocPrf    PrefVal Path/Ogn

 *>i 10.0.45.0/24    10.0.2.2     0       100       0     100?
 * i                 10.0.3.3     0       100       0     300?
```

Run the **display bgp routing** command on R1 to check the BGP routing table. The command output shows that the BGP route 10.0.45.0/24 is valid.



Why is the BGP route with the next hop of 10.0.2.2 the optimal route when next hops of two BGP routes are reachable?

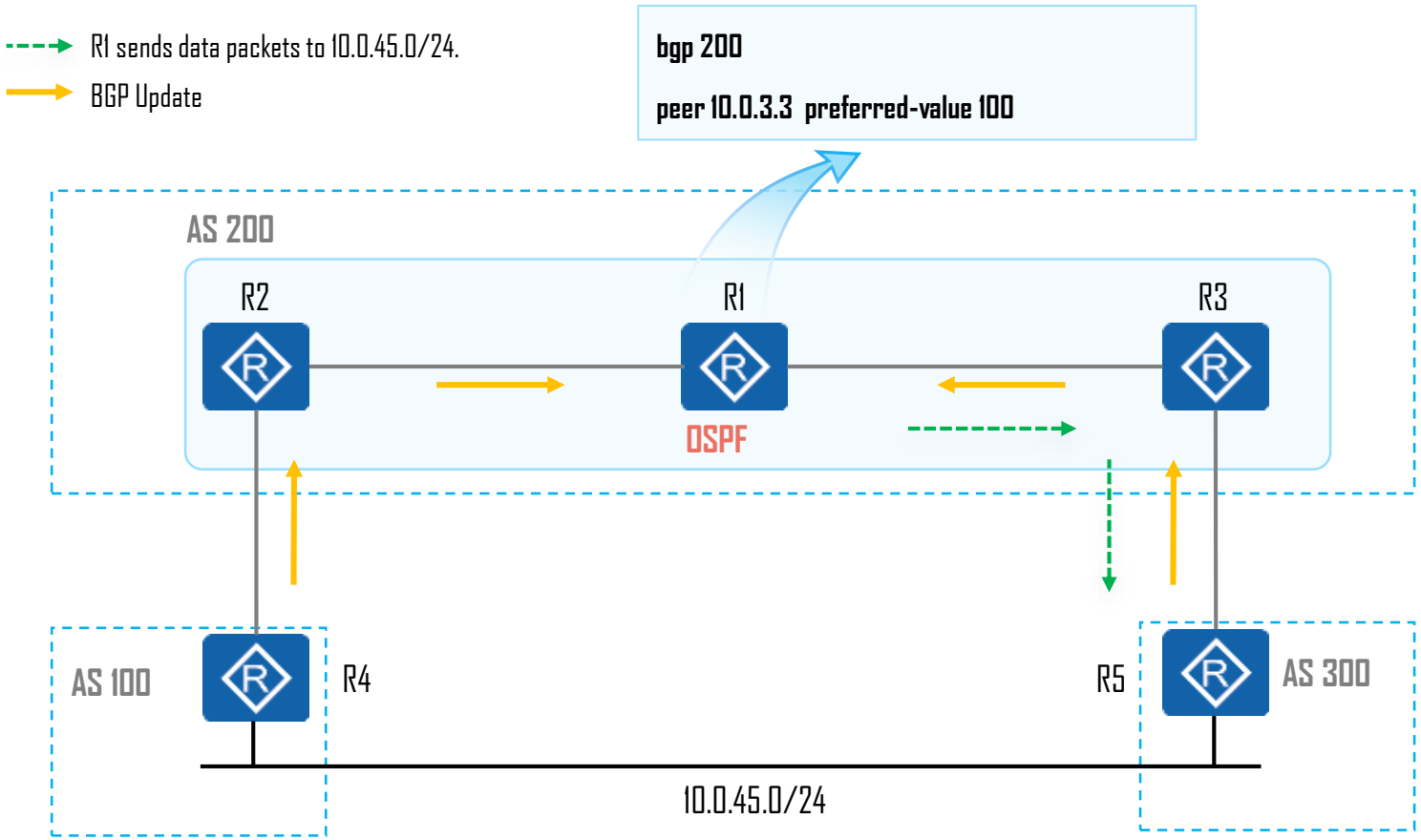
Rules for Selecting a Preferred BGP Route

When multiple routes to the same destination network segment exist, BGP selects routes in the following sequence:

Discards the route whose next hop is unreachable.

1. **Prefers the route with the largest Preferred-Value attribute value.**
2. Prefers the route with the largest Local_Preference value.
3. Prefers the locally originated BGP route, which takes precedence over the route learned from a peer. The locally summarized route, automatically summarized route, route learned by using the **network** command, route learned by using the **import-route** command, and route learned from a peer are in descending order of priority.
4. Prefers the route with the shortest AS_Path.
5. Prefers the route with the shortest AS_Path. The routes with Origin attributes of IGP, EGP, and Incomplete are in descending order of priority.
6. Prefers the route with the lowest MED.
7. Prefers routes learned from EBGP peers to routes learned from IBGP peers.
8. Prefers the route with the smallest IGP metric to the next hop.
9. Prefers the route with the shortest Cluster_List length.
10. Prefers the route advertised by the device with the smallest router ID (Originator_ID).
11. Prefers the route learned from the peer with the smallest IP address.

Changing the Preferred-Value Attribute



Run the **preferred-value** command to change the Preferred-Value attribute value of the BGP route advertised by R3 to 100, which takes precedence over the BGP route with the default Preferred-Value attribute value advertised by R2. Then R1 preferentially selects the BGP route 10.0.45.0/24 advertised by R3.

Checking the BGP Routing Table of R1

```
[R1] display bgp routing-table
BGP Local router ID is 10.0.1.1
 Status codes: * - valid, > - best, d - damped,
               h - history,  i - internal, s - suppressed, S - Stale
               Origin : i - IGP, e - EGP, ? - incomplete

 Total Number of Routes: 4
     Network      NextHop        MED        LocPrf      PrefVal      Path/Ogn
*>i  10.0.45.0/24               10.0.3.3   0           100          300 i
* i                             10.0.2.2   0           0            100 i
```

The BGP route advertised by R3 at 10.0.3.3 has a higher Preferred-Value attribute value (100), so R1 prefers the BGP route 10.0.45.0/24 advertised by R3.

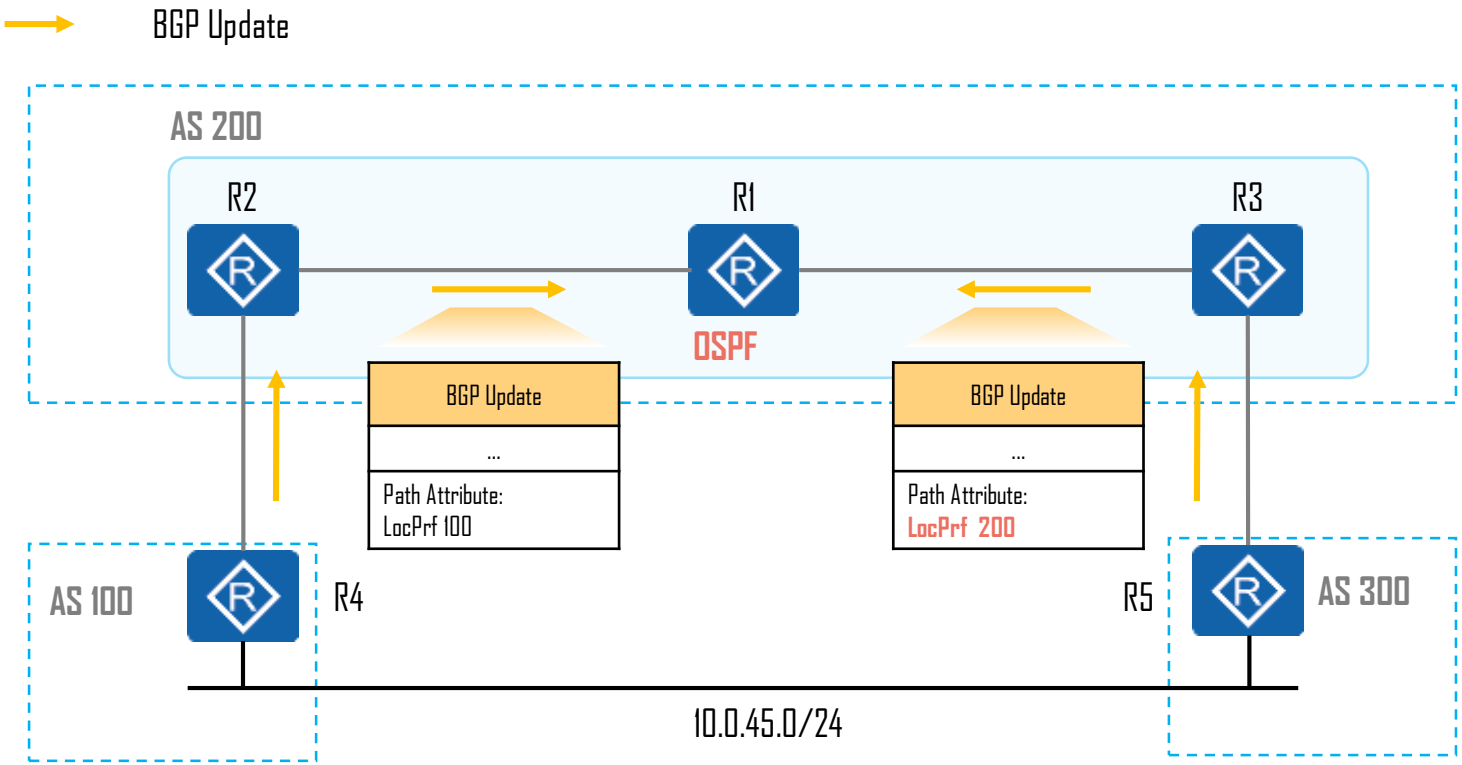
Rules for Selecting a Preferred BGP Route

When multiple routes to the same destination network segment exist, BGP selects routes in the following sequence:

Discards the route whose next hop is unreachable.

1. Prefers the route with the largest Preferred-Value attribute value.
2. **Prefers the route with the largest Local_Preference value.**
3. Prefers the locally originated BGP route, which takes precedence over the route learned from a peer. The locally summarized route, automatically summarized route, route learned by using the **network** command, route learned by using the **import-route** command, and route learned from a peer are in descending order of priority.
4. Prefers the route with the shortest AS_Path.
5. Prefers the route with the shortest AS_Path. The routes with Origin attributes of IGP, EGP, and Incomplete are in descending order of priority.
6. Prefers the route with the lowest MED.
7. Prefers routes learned from EBGP peers to routes learned from IBGP peers.
8. Prefers the route with the smallest IGP metric to the next hop.
9. Prefers the route with the shortest Cluster_List length.
10. Prefers the route advertised by the device with the smallest router ID (Originator_ID).
11. Prefers the route learned from the peer with the smallest IP address.

Changing the Local_Preference Attribute (1)



Run the following commands on R3.

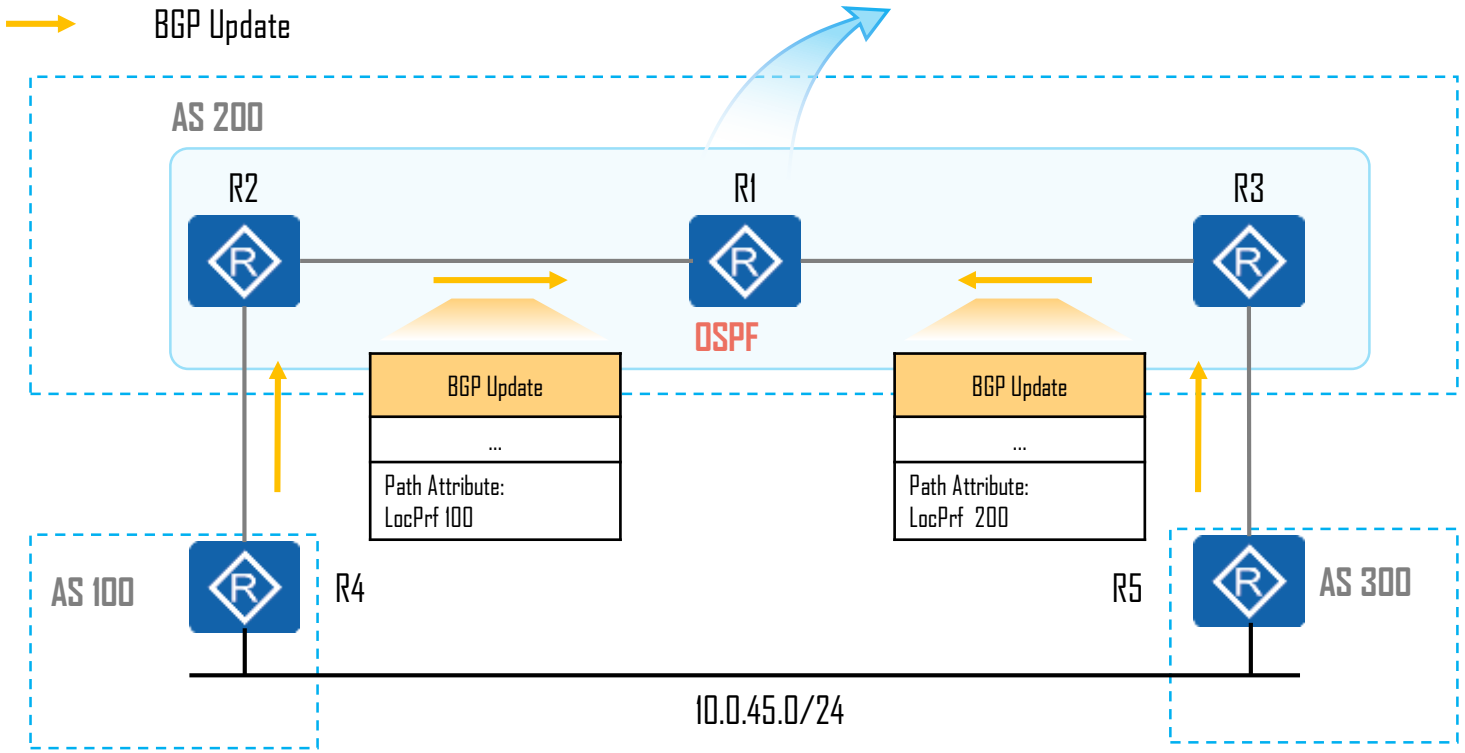
```
ip ip-prefix local_pref index 10 permit 10.0.45.0 24
#
route-policy local_pref permit node 10
 if-match ip-prefix local_pref
 apply local-preference 200
route-policy local_pref permit node 20
#
bgp 200
peer 10.0.1.1 route-policy local_pref export
```

Configure a routing policy on R3 to change the Local_Preference value of the BGP route 10.0.45.0/24 advertised to R1.

Changing the Local_Preference Attribute (2)

```
Total Number of Routes: 2
     Network         NextHop   MED    LocPrf    PrefVal Path/Ogn

*>i  10.0.45.0/24    10.0.3.3    0       200       0     300?
* i                   10.0.2.2    0       100       0     100?
```



If the next hops of the routes are reachable and the Preferred-Value attribute values of the routes are the same, R1 compares the Local_Preference attribute values of the routes. The Local_Preference attribute value of the BGP route advertised by R3 is 200, which is greater than that of the BGP route advertised by R2. Therefore, R1 prefers the BGP route advertised by R3.

Rules for Selecting a Preferred BGP Route

When multiple routes to the same destination network segment exist, BGP selects routes in the following sequence:

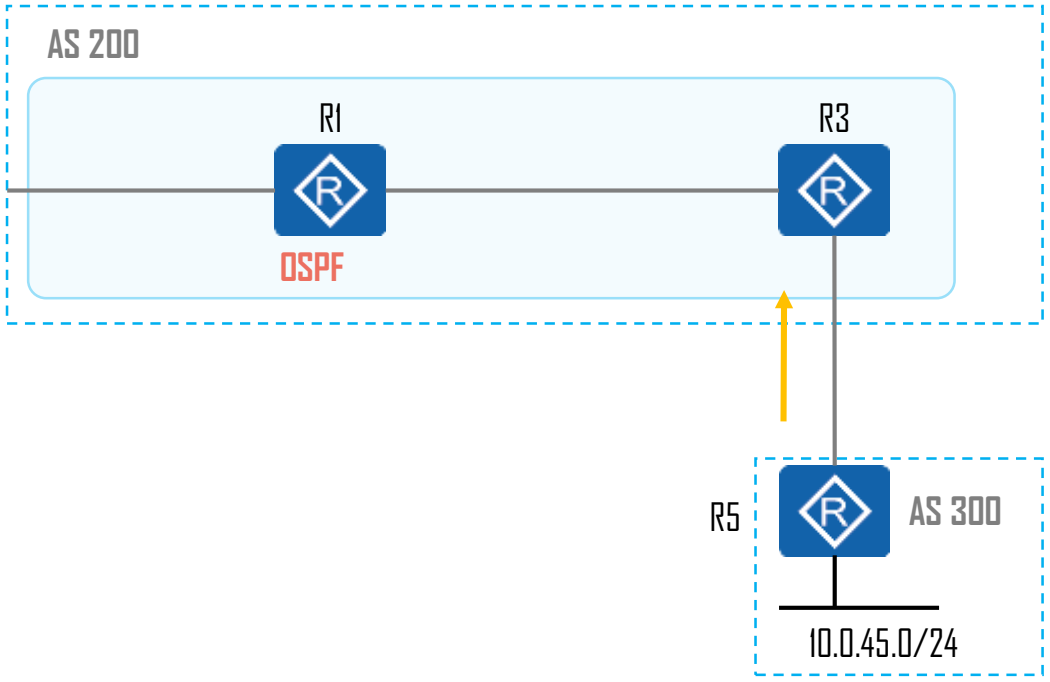
Discards the route whose next hop is unreachable.

1. Prefers the route with the largest Preferred-Value attribute value.
2. Prefers the route with the largest Local_Preference value.
3. **Prefers the locally originated BGP route, which takes precedence over the route learned from a peer. The locally summarized route, automatically summarized route, route learned by using the network command, route learned by using the import-route command, and route learned from a peer are in descending order of priority.**
4. Prefers the route with the shortest AS_Path.
5. Prefers the route with the shortest AS_Path. The routes with Origin attributes of IGP, EGP, and Incomplete are in descending order of priority.
6. Prefers the route with the lowest MED.
7. Prefers routes learned from EBGP peers to routes learned from IBGP peers.
8. Prefers the route with the smallest IGP metric to the next hop.
9. Prefers the route with the shortest Cluster_List length.
10. Prefers the route advertised by the device with the smallest router ID (Originator_ID).
11. Prefers the route learned from the peer with the smallest IP address.

Preferring the Locally Generated Route

- In the case of identical conditions, a locally generated route is preferred, and the route learned from a peer has the secondary priority.
- In addition, locally generated routes may be learned in multiple ways. When the same route is learned in multiple ways, the following routes are in descending order of priority:
  - Summarized route by manually running the **aggregate** command in the BGP view
  - Automatically summarized route by running the **summary automatic** command
  - Route imported using the **network** command
  - Route imported using the **import-route** command

Manual Route Summarization (1)



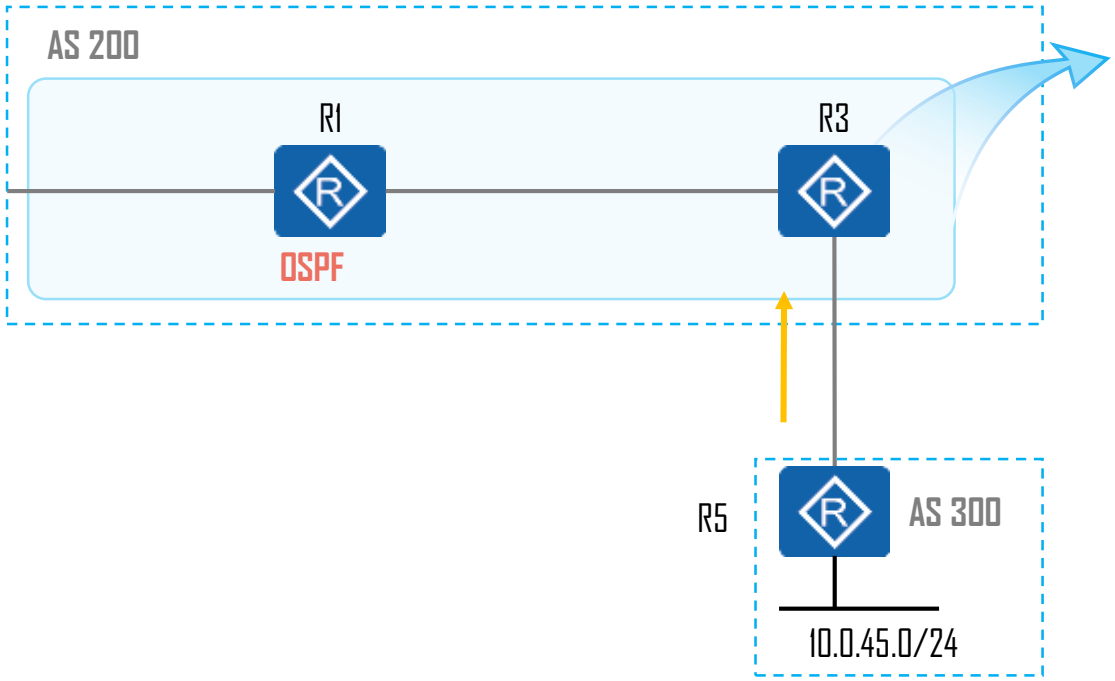
To manually summarize routes on R3, configure two static routes pointing to null0 on R3 and import them to BGP.

- Run the following commands on R3.

```
ip route-static 10.0.45.0 255.255.255.128 null0
ip route-static 10.0.45.128 255.255.255.128 null0
bgp 200
 aggregate 10.0.45.0 255.255.255.0 detail-suppressed
 import-route static
```

- Configure two static routes on R3, import the static routes to BGP using the **import-route** command, run the **aggregate** command to manually summarize the routes, and specify **detail-suppressed** suppress advertisement of specific routes.

Manual Route Summarization (2)



	Network	NextHop	MED	LocPrf	PrefVal	Path/Ogn
*>	10.0.45.0/24	127.0.0.1			0	?
*		10.0.35.5	0		0	300?
s>	10.0.45.0/25	0.0.0.0	0		0	?
s>	10.0.45.128/25	0.0.0.0	0		0	?

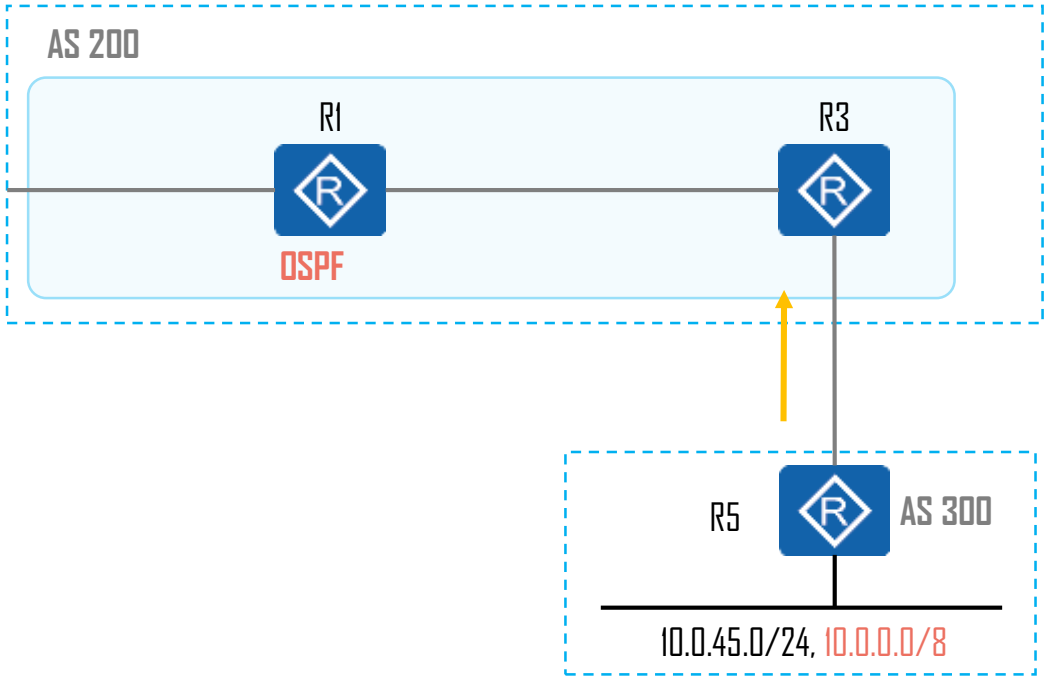
- The BGP routing table of R3 contains two BGP routes 10.0.45.0/24.
  - Locally generated route: Static routes are imported to BGP and are manually summarized.
  - Route advertised by R5 at 10.0.35.5
- The two routes do not have Local_Preference or Preferred-Value attribute values on R3. R3 then compares the source of the two routes and prefers the route that is manually summarized.

Manual Route Summarization (3)

```
[R3]display bgp routing-table 10.0.45.0 24
 BGP local router ID : 10.0.3.3
 Local AS number : 200
 Paths:   2 available, 1 best, 1 select
 BGP routing table entry information of 10.0.45.0/24:
 Aggregated route.
 Route Duration: 00h00m14s
 Direct Out-interface: NULL0
 Original nexthop: 127.0.0.1
 Qos information : 0x0
 AS-path Nil, origin incomplete, pref-val 0, valid, local, best, select, active,
pre 255
 Aggregator: AS 200, Aggregator ID 10.0.3.3, Atomic-aggregate
 Advertised to such 2 peers:
   10.0.35.5
   10.0.1.1
```

- Run the **display bgp routing-table 10.0.45.0 24** command on R3 to check detailed information about BGP route 10.0.45.0/24. The command output shows that there are two valid routes, and the manually summarized route is better.
- This example verifies that the locally generated BGP route is better than the BGP route learned from a peer.

Automatic Route Summarization (1)



In this case, the configurations on R1, R3, and R5 are irrelevant to the configurations that have been performed in the example of manual summarization.

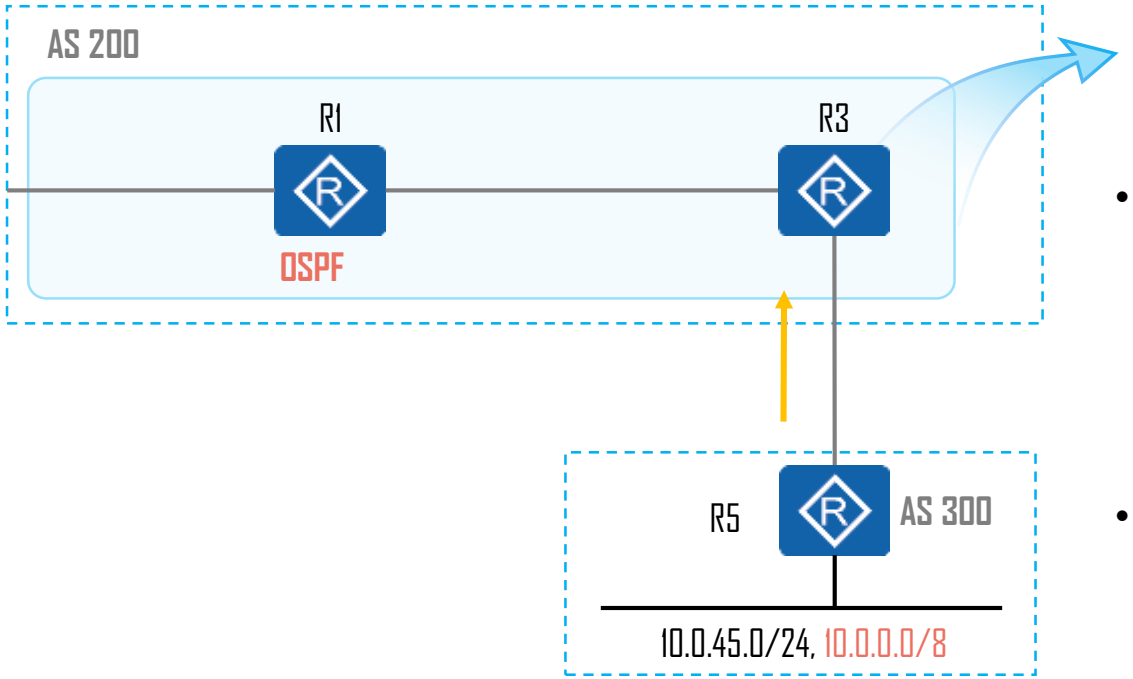
- Run the following commands on R3.

```
ip route-static 10.0.45.0 255.255.255.128 null0
ip route-static 10.0.45.128 255.255.255.128 null0

bgp 200
summary automatic
import-route static
```

- Configure two static routes on R3, import the static routes to BGP using the **import-route** command, and enable automatic summarization. BGP summarizes routes by natural network segment. For example, class A addresses 10.1.1.1/24 and 10.2.1.1/24 on the non-natural network segment are summarized into class A address 10.0.0.0/8 on the natural network segment. In addition, BGP advertises only the summarized route to peers.
- On R3, you can view that the route is summarized to 10.0.0.0/8.
- R5 imports the route 10.0.0.0/8 and advertises it to R3.

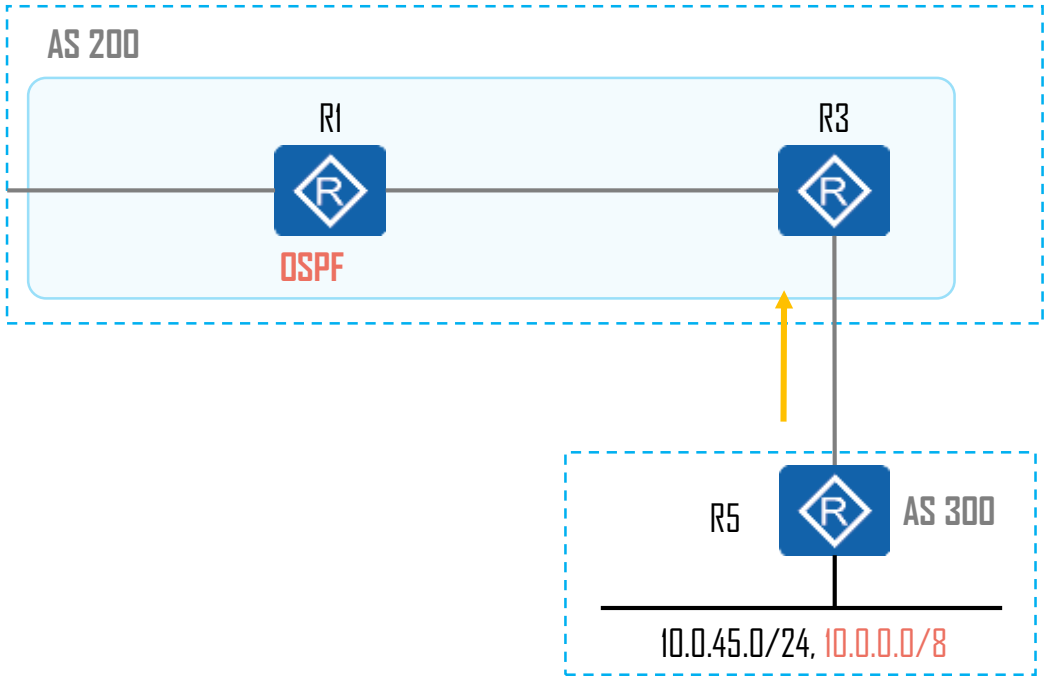
Automatic Route Summarization (2)



	Network	NextHop	MED	LocPrf	PrefVal	Path/Ogn
*>	10.0.0.0	127.0.0.1			0	?
*		10.0.35.5	0		0	300?

- The BGP routing table on R3 contains two BGP routes 10.0.0.0.
  - Locally generated route: Static routes are imported to BGP and automatically aggregated.
  - Route advertised by R5 at 10.0.35.5
- The two routes do not have Local_Preference or Preferred-Value attribute values on R3. R3 then compares the source of the two routes and prefers the route that is automatically summarized.

Automatic Route Summarization (3)



- Perform manual summarization on R3.

```
bgp 200
 aggregate 10.0.0.0 255.0.0.0 detail-suppressed
```

- Check the routing table of R3.

	Network	NextHop	MED	LocPrf	PrefVal	Path/Ogn
*>	10.0.0.0	127.0.0.1			0	?
*		127.0.0.1			0	?
*		10.0.35.5	0		0	300?

- The locally generated BGP route is preferred. However, there are two locally generated BGP routes, and the routing entry cannot help determine whether the preferred route is manually or automatically summarized.

Automatic Route Summarization (4)

```
BGP local router ID : 10.0.3.3
Local AS number : 200
Paths:   3 available, 1 best, 1 select
BGP routing table entry information of 10.0.0.0/8:
Aggregated route.
Route Duration: 00h08m17s
Direct Out-interface: NULL0
Original nexthop: 127.0.0.1
Qos information : 0x0
AS-path Nil, origin incomplete, pref-val 0, valid, local, best, select, active,
pre 255
Aggregator: AS 200, Aggregator ID 10.0.3.3, Atomic-aggregate
Advertised to such 2 peers:
  10.0.35.5
  10.0.1.1
```

- Run the **display bgp routing-table 10.0.0.0** command on R3 to check detailed information about the BGP route 10.0.0.0/8. The command output shows that there are three valid routes, among which the optimal route is generated by route summarization and has the Atomic-aggregate attribute. The command output shows that the route is manually summarized.
- On R3, the manually summarized route is better than the automatically summarized route.
- This example verifies that the manually summarized route is better than the automatically summarized route.

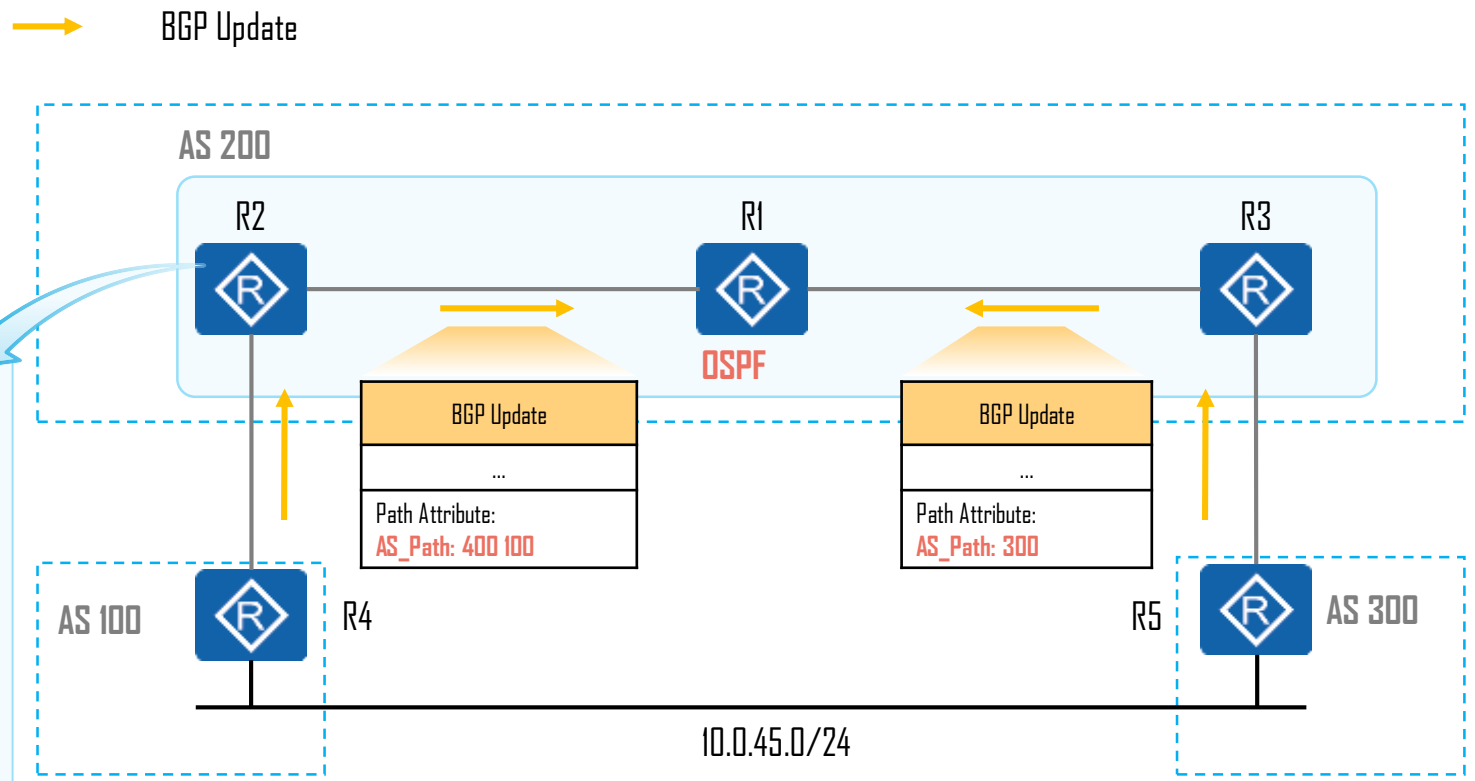
Rules for Selecting a Preferred BGP Route

When multiple routes to the same destination network segment exist, BGP selects routes in the following sequence:

Discards the route whose next hop is unreachable.

1. Prefers the route with the largest Preferred-Value attribute value.
2. Prefers the route with the largest Local_Preference value.
3. Prefers the locally originated BGP route, which takes precedence over the route learned from a peer. The locally summarized route, automatically summarized route, route learned by using the network command, route learned by using the import-route command, and route learned from a peer are in descending order of priority.
4. **Prefers the route with the shortest AS_Path.**
5. Prefers the route with the shortest AS_Path. The routes with Origin attributes of IGP, EGP, and Incomplete are in descending order of priority.
6. Prefers the route with the lowest MED.
7. Prefers routes learned from EBGP peers to routes learned from IBGP peers.
8. Prefers the route with the smallest IGP metric to the next hop.
9. Prefers the route with the shortest Cluster_List length.
10. Prefers the route advertised by the device with the smallest router ID (Originator_ID).
11. Prefers the route learned from the peer with the smallest IP address.

Preferring the Route with the Shortest AS_Path (1)




```
ip ip-prefix as_path index 10 permit 10.0.45.0 24
#
route-policy as_path permit node 10
 if-match ip-prefix as_path
 apply as-path 400 additive
route-policy as_path permit node 20
#
bgp 200
peer 10.0.1.1 route-policy local_pref export
```

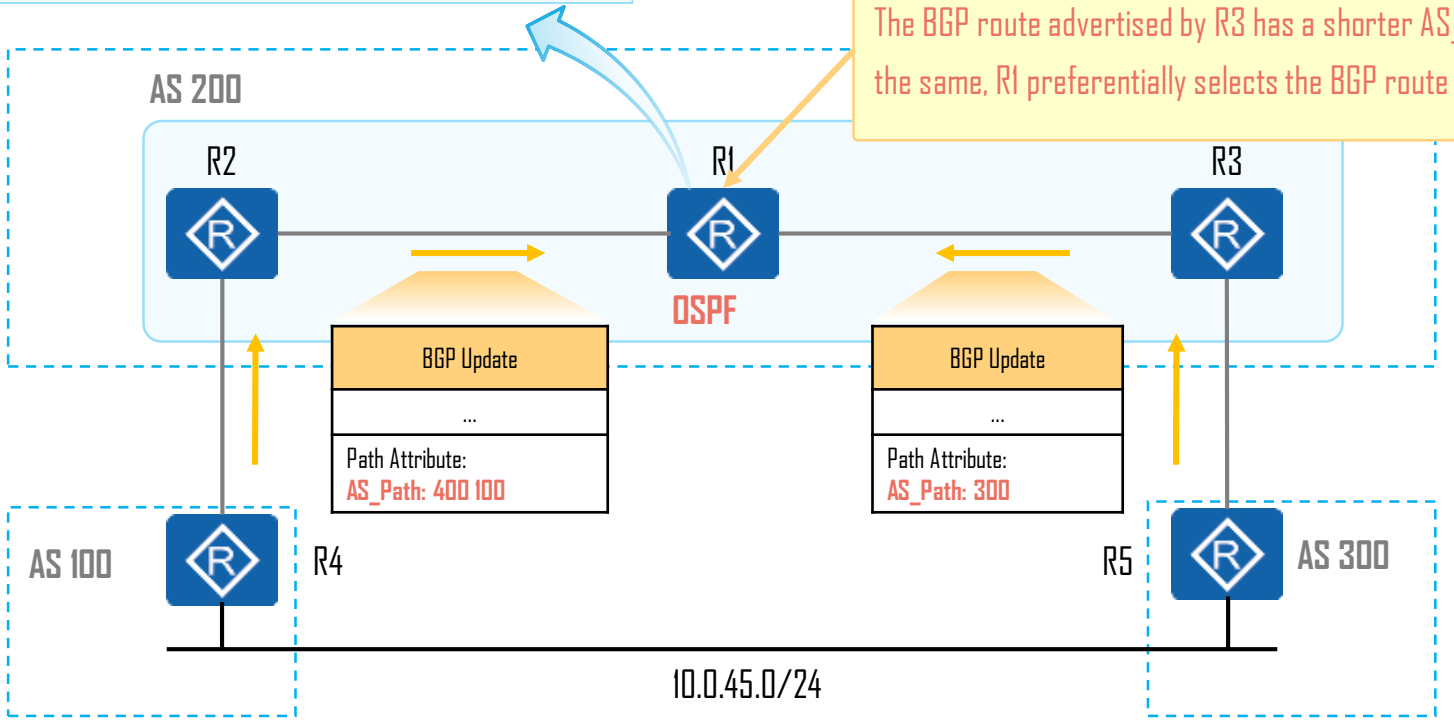
Configure a routing policy on R2 to change the AS_Path attribute value of the BGP route advertised to R1.

Preferring the Route with the Shortest AS_Path (2)

```
Total Number of Routes: 2
     Network         NextHop   MED    LocPrf    PrefVal Path/Ogn

*>i  10.0.45.0/24    10.0.3.3        0         100       0      300?
* i                   10.0.2.2        0         100       0      400 100?
```

The BGP route advertised by R3 has a shorter AS_Path. If the preceding rules are the same, R1 preferentially selects the BGP route advertised by R3.



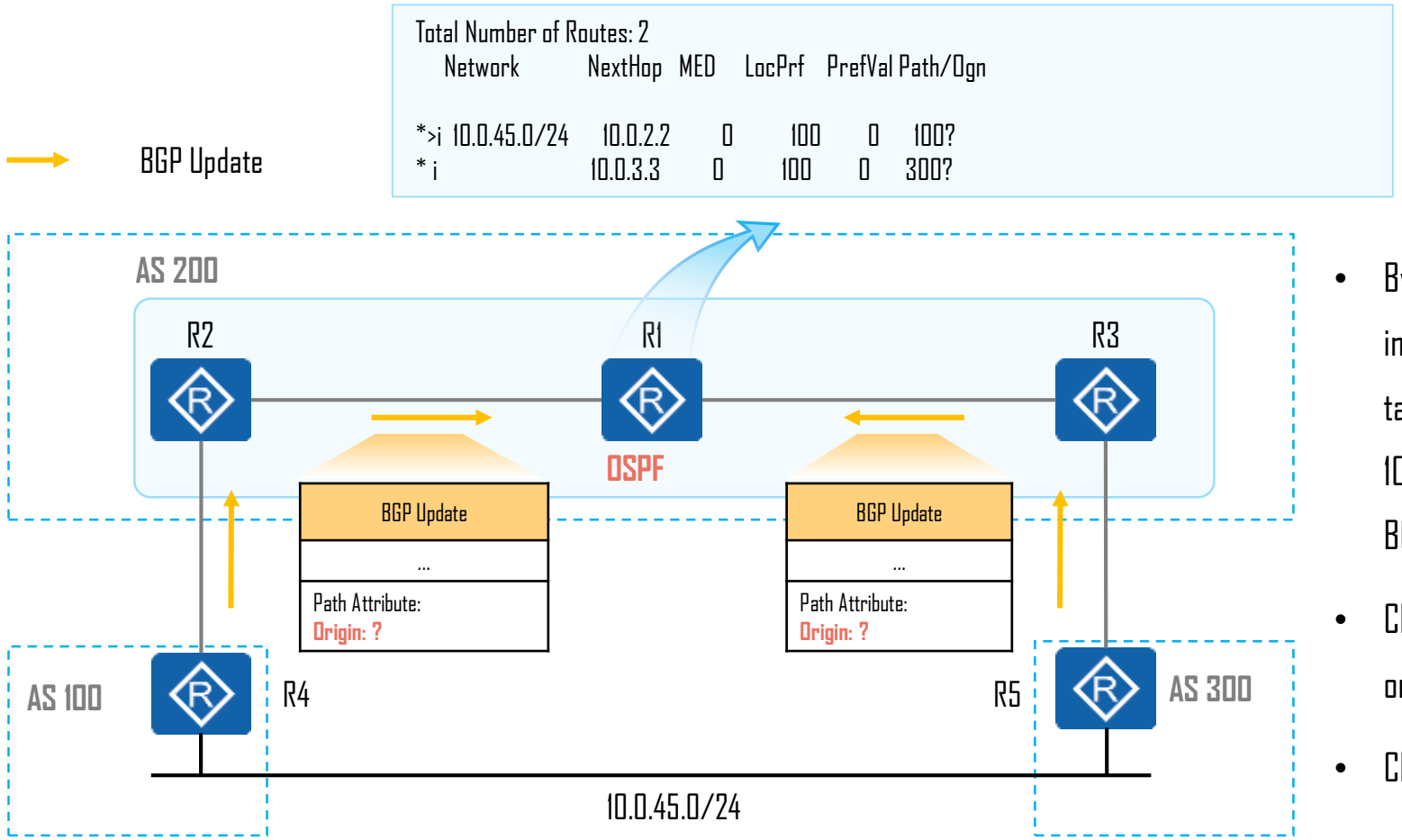
Rules for Selecting a Preferred BGP Route

When multiple routes to the same destination network segment exist, BGP selects routes in the following sequence:

Discards the route whose next hop is unreachable.

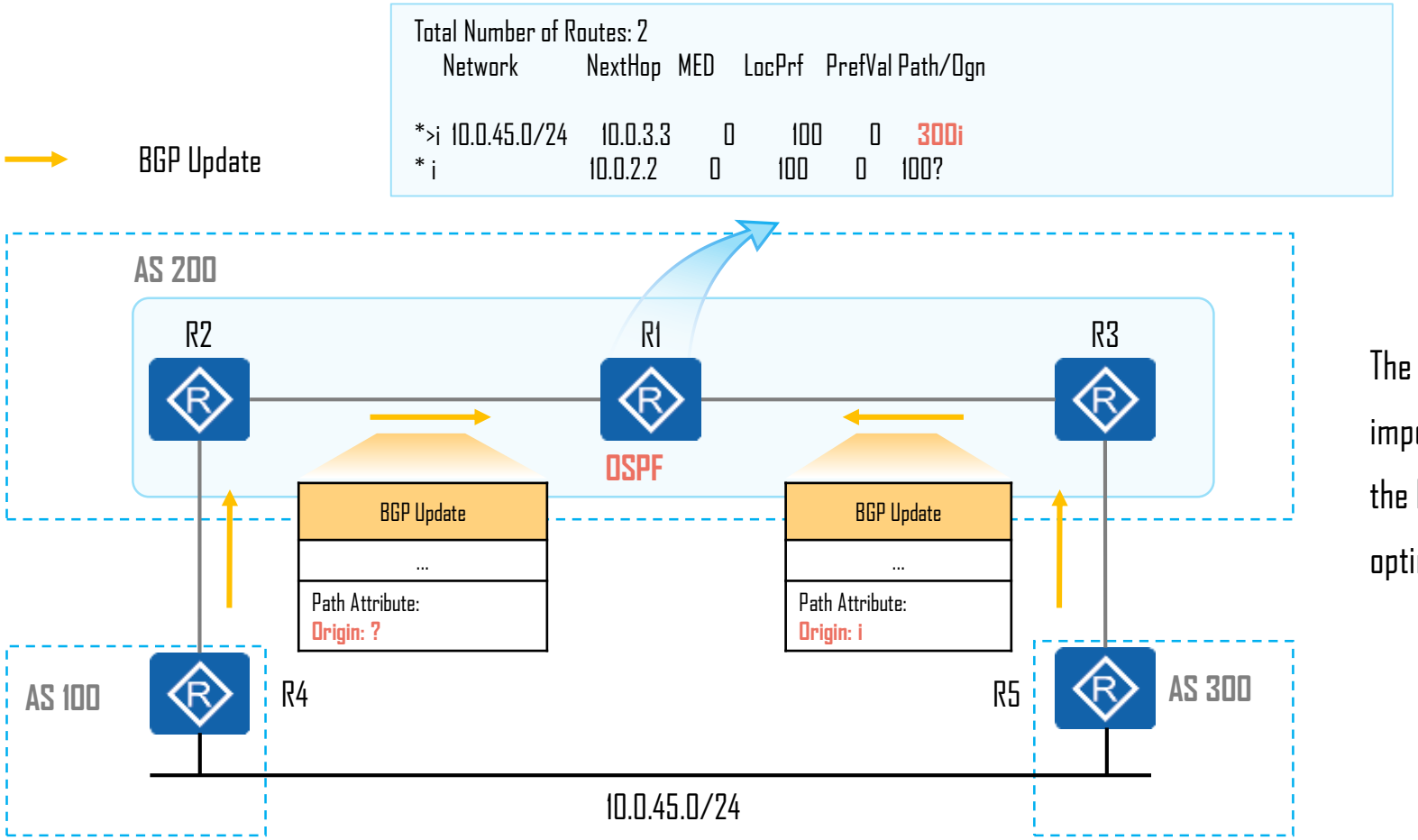
1. Prefers the route with the largest Preferred-Value attribute value.
2. Prefers the route with the largest Local_Preference value.
3. Prefers the locally originated BGP route, which takes precedence over the route learned from a peer. The locally summarized route, automatically summarized route, route learned by using the network command, route learned by using the import-route command, and route learned from a peer are in descending order of priority.
4. Prefers the route with the shortest AS_Path.
5. **Prefers the route with the shortest AS_Path. The routes with Origin attributes of IGP, EGP, and Incomplete are in descending order of priority.**
6. Prefers the route with the lowest MED.
7. Prefers routes learned from EBGP peers to routes learned from IBGP peers.
8. Prefers the route with the smallest IGP metric to the next hop.
9. Prefers the route with the shortest Cluster_List length.
10. Prefers the route advertised by the device with the smallest router ID (Originator_ID).
11. Prefers the route learned from the peer with the smallest IP address.

Origin Attribute Verification (1)



- By default, R4 and R5 use the **import-route** command to import the routes 10.0.45.0/24 to BGP. In the BGP routing table of R1, the Origin attributes of the two BGP routes 10.0.45.0/24 are both "?". R1 preferentially selects the BGP route imported by R4.
- Change the command used to import routes to **network** on R5.
- Check the BGP routing table on R1.

Origin Attribute Verification (2)



The Origin attribute of the BGP route 10.0.45.0/24 imported by R5 is "i". If the preceding rules are the same, the BGP route with the origin type being "i" becomes the optimal route.

Rules for Selecting a Preferred BGP Route

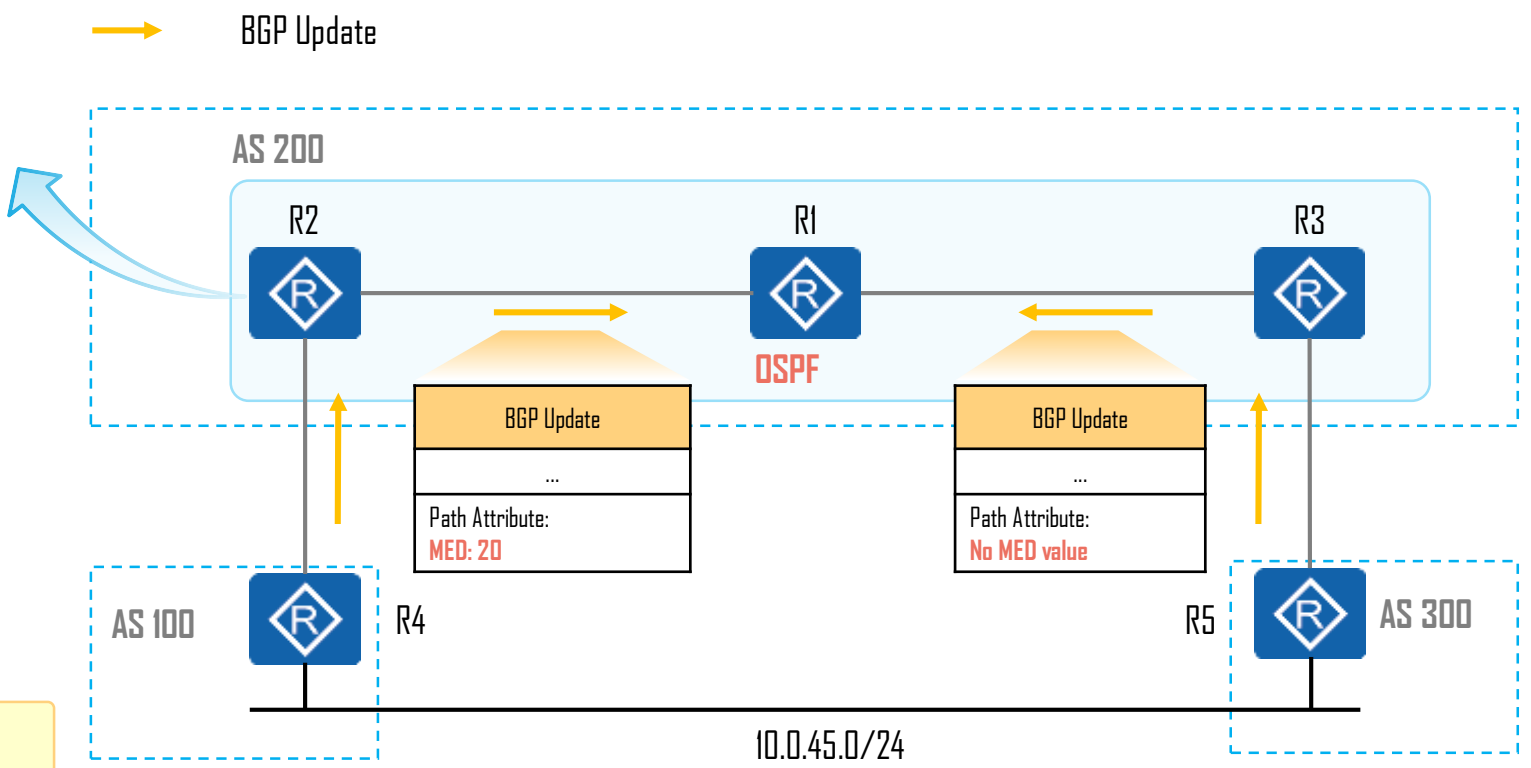
When multiple routes to the same destination network segment exist, BGP selects routes in the following sequence:

Discards the route whose next hop is unreachable.

1. Prefers the route with the largest Preferred-Value attribute value.
2. Prefers the route with the largest Local_Preference value.
3. Prefers the locally originated BGP route, which takes precedence over the route learned from a peer. The locally summarized route, automatically summarized route, route learned by using the network command, route learned by using the import-route command, and route learned from a peer are in descending order of priority.
4. Prefers the route with the shortest AS_Path.
5. Prefers the route with the shortest AS_Path. The routes with Origin attributes of IGP, EGP, and Incomplete are in descending order of priority.
6. **Prefers the route with the lowest MED.**
7. Prefers routes learned from EBGP peers to routes learned from IBGP peers.
8. Prefers the route with the smallest IGP metric to the next hop.
9. Prefers the route with the shortest Cluster_List length.
10. Prefers the route advertised by the device with the smallest router ID (Originator_ID).
11. Prefers the route learned from the peer with the smallest IP address.

Preferring the Route with the Smallest MED (1)

```
ip ip-prefix med index 10 permit 10.0.45.0 24
#
route-policy med permit node 10
 if-match ip-prefix med
 apply cost 20
route-policy med permit node 20
#
bgp 200
peer 10.0.1.1 route-policy med export
compare-different-as-med
```



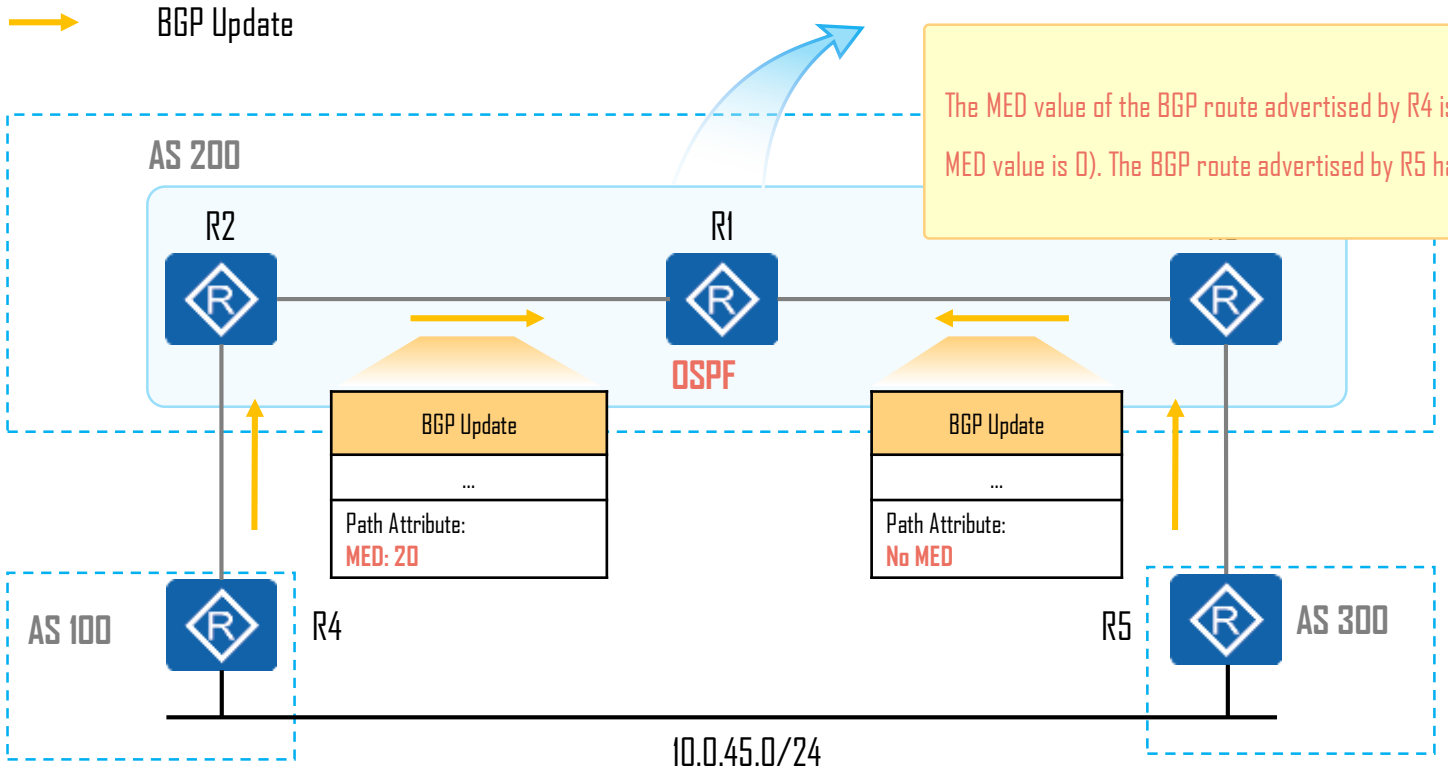
By default, BGP compares the MED values of routes from the same AS and to the same network segment. You can use a command to enable BGP to compare the MED attribute values of the same routes from different ASs.

Configure a routing policy on R2 to change the MED attribute values of the BGP routes advertised to R1.

Preferring the Route with the Smallest MED (2)

```
Total Number of Routes: 2
     Network         NextHop   MED    LocPrf    PrefVal Path/Ogn

*>i  10.0.45.0/24    10.0.3.3    0       100       0    300?
* i                   10.0.2.2   20       100       0    100?
```



The MED value of the BGP route advertised by R4 is 20, and the BGP route advertised by R5 does not carry the MED value (the default MED value is 0). The BGP route advertised by R5 has a smaller MED value, and R1 preferentially selects the BGP route advertised by R5.

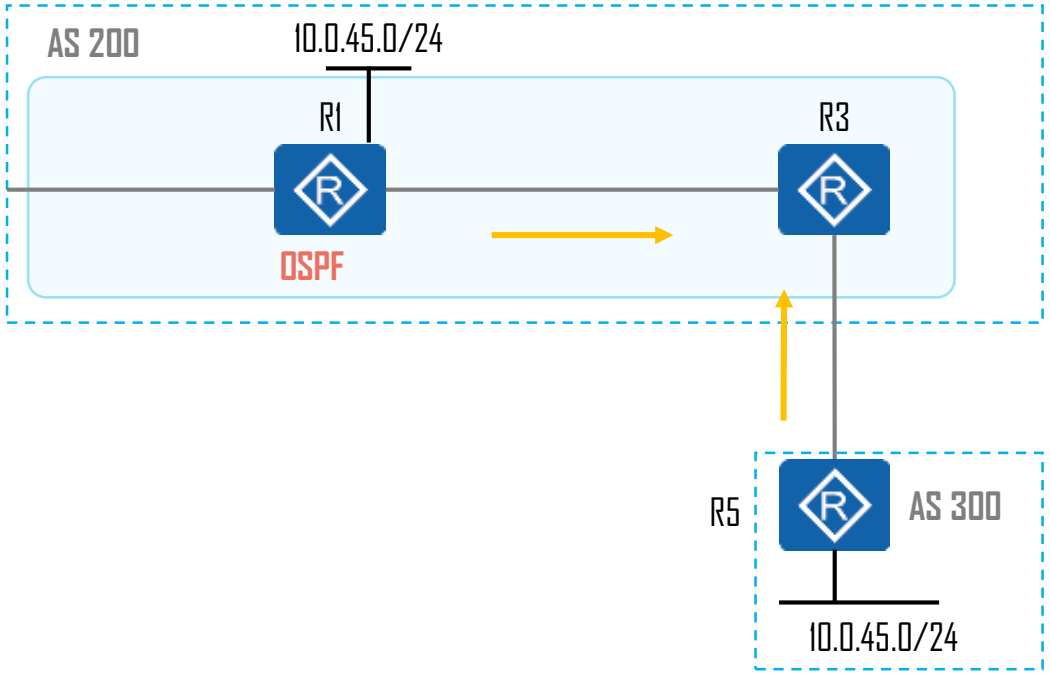
Rules for Selecting a Preferred BGP Route

When multiple routes to the same destination network segment exist, BGP selects routes in the following sequence:

Discards the route whose next hop is unreachable.

1. Prefers the route with the largest Preferred-Value attribute value.
2. Prefers the route with the largest Local_Preference value.
3. Prefers the locally originated BGP route, which takes precedence over the route learned from a peer. The locally summarized route, automatically summarized route, route learned by using the network command, route learned by using the import-route command, and route learned from a peer are in descending order of priority.
4. Prefers the route with the shortest AS_Path.
5. Prefers the route with the shortest AS_Path. The routes with Origin attributes of IGP, EGP, and Incomplete are in descending order of priority.
6. Prefers the route with the lowest MED.
7. **Prefers routes learned from EBGP peers to routes learned from IBGP peers.**
8. Prefers the route with the smallest IGP metric to the next hop.
9. Prefers the route with the shortest Cluster_List length.
10. Prefers the route advertised by the device with the smallest router ID (Originator_ID).
11. Prefers the route learned from the peer with the smallest IP address.

Preferring a Route Learned from an EBGP Peer (1)



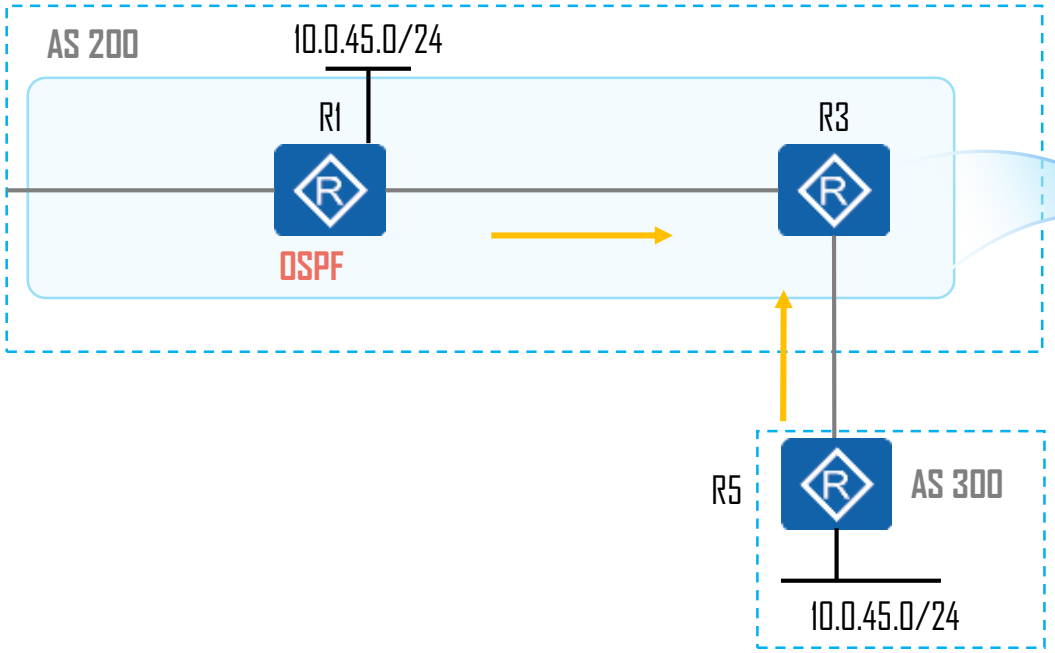
Create a static route 10.0.45.0/24 pointing to null0 on R1 and advertise the route to BGP. Ensure that the AS_Path attribute values of the BGP routes advertised by R1 and R5 to R3 are the same, configure a routing policy to add the AS_Path attribute to the route advertised by R1 to R3. The AS_Path attribute value is 500.

- Run the following commands on R1.

```
ip route-static 10.0.45.0 255.255.255.0 null0
ip ip-prefix ebgp index 10 permit 10.0.45.0 24
#
route-policy ebgp permit node 10
 if-match ip-prefix ebgp
 apply as-path 500 additive
route-policy ebgp permit node 20
#
bgp 200
  import-route static
  peer 10.0.3.3 route-policy ebgp export
```

- R3 will receive the BGP route 10.0.45.0/24 advertised by R1 and R5, and the preceding route selection rules cannot determine the optimal route.

Preferring a Route Learned from an EBGP Peer (2)



	Network	NextHop	MED	LocPrf	PrefVal	Path/Ogn
*>	10.0.45.0/24	10.0.35.5	0		0	300?
* i		10.0.1.1	0	100	0	500?

R5 and R1 are the EBGP peer and IBGP peer, respectively. The BGP routes advertised by EBGP peers take precedence over the BGP routes advertised by IBGP peers. R3 preferentially selects the BGP routes advertised by R5.

Preferring a Route Learned from an EBGP Peer (3)

BGP routing table entry information of 10.0.45.0/24:

From: 10.0.1.1 (10.0.1.1)

Route Duration: 00h06m43s

Relay IP Nexthop: 10.0.13.1

Relay IP Out-Interface: GigabitEthernet0/0/0

Original nexthop: 10.0.1.1

Qos information : 0x0

AS-path 500, origin incomplete, MED 0, localpref 100, pref-val 0, valid, internal, pre 255, IGP cost 1, **not preferred for peer type**

Not advertised to any peer yet

Run the **display bgp routing-table 10.0.45.0 24** command on R3 to check detailed information about BGP routes. The command output is as follows:

not preferred for peer type

The route is not selected because the peer type is not preferred.

Rules for Selecting a Preferred BGP Route

When multiple routes to the same destination network segment exist, BGP selects routes in the following sequence:

Discards the route whose next hop is unreachable.

1. Prefers the route with the largest Preferred-Value attribute value.
2. Prefers the route with the largest Local_Preference value.
3. Prefers the locally originated BGP route, which takes precedence over the route learned from a peer. The locally summarized route, automatically summarized route, route learned by using the network command, route learned by using the import-route command, and route learned from a peer are in descending order of priority.
4. Prefers the route with the shortest AS_Path.
5. Prefers the route with the shortest AS_Path. The routes with Origin attributes of IGP, EGP, and Incomplete are in descending order of priority.
6. Prefers the route with the lowest MED.
7. Prefers routes learned from EBGP peers to routes learned from IBGP peers.
8. **Prefers the route with the smallest IGP metric to the next hop.**
9. Prefers the route with the shortest Cluster_List length.
10. Prefers the route advertised by the device with the smallest router ID (Originator_ID).
11. Prefers the route learned from the peer with the smallest IP address.

IGP Cost

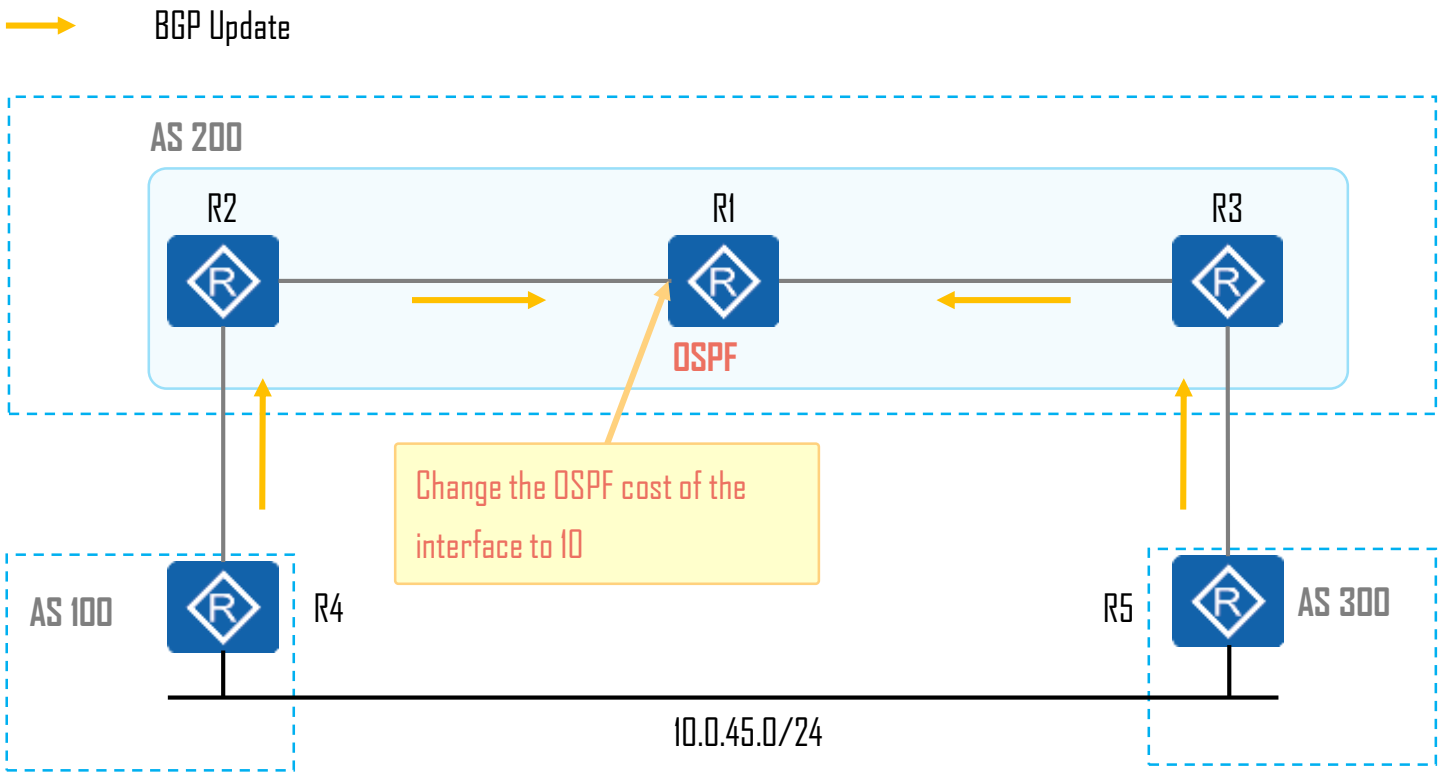
BGP local router ID : 10.0.1.1
Local AS number : 200
Paths: 2 available, 1 best, 1 select
BGP routing table entry information of 10.0.45.0/24:
From: 10.0.3.3 (10.0.3.3)
Route Duration: 00h22m35s
Relay IP Nexthop: 10.0.13.3
Relay IP Out-Interface: GigabitEthernet0/0/1
Original nexthop: 10.0.3.3
Qos information : 0x0
AS-path 300, origin incomplete, MED 0, localpref 100, pref-val 0, valid, internal, best, select, active, pre 255, **IGP cost 1**
Not advertised to any peer yet

- The IGP cost is displayed in the detailed BGP route information. The IGP cost is the cost of the route to the original next hop in the local IP routing table.

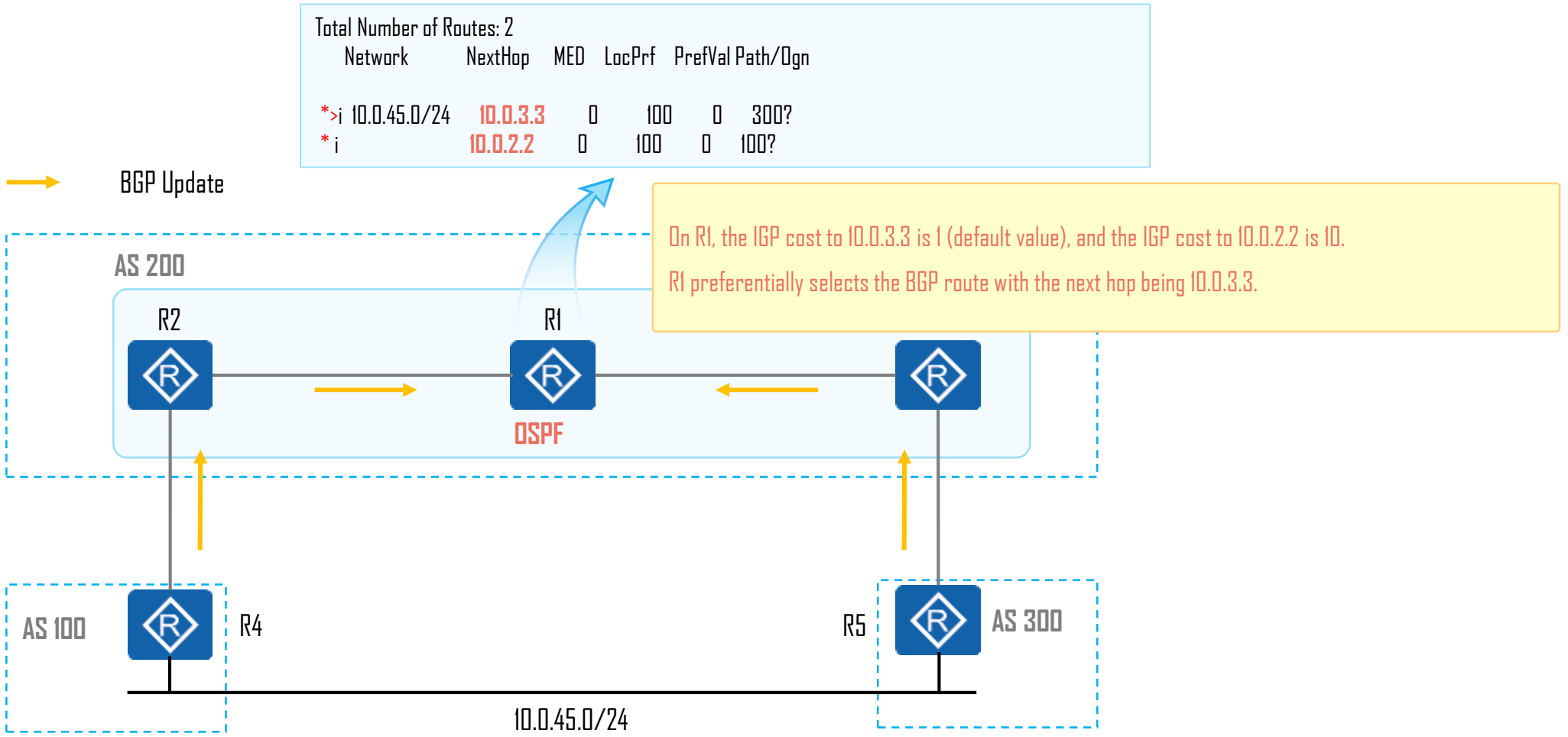
Destination/Mask	Proto	Pre	Cost	NextHop	Interface
10.0.3.3/32	OSPF	10	1	10.0.13.3	GigabitEthernet0/0/1

- If the preceding seven rules cannot determine the optimal BGP route, the IGP cost of the next hop is compared.

Preferring the Route with the Smallest IGP Cost (1)



Preferring the Route with the Smallest IGP Cost (2)



Preferring the Route with the Smallest IGP Cost (3)

BGP routing table entry information of 10.0.45.0/24:

From: 10.0.2.2 (10.0.2.2)

Route Duration: 00h24m07s

Relay IP Nexthop: 10.0.12.2

Relay IP Out-Interface: GigabitEthernet0/0/0

Original nexthop: 10.0.2.2

Qos information : 0x0

AS-path 100, origin incomplete, MED 0, localpref 100, pref-val 0, valid, internal, pre 255, **IGP cost 10, not preferred for IGP cost**

Not advertised to any peer yet

- Run the **display bgp routing-table 10.0.45.0 24** command on R1 to check detailed information about BGP routes. The command output shows that the IGP cost of the BGP route with the next hop being 10.0.2.2 changes to 10 and the IGP cost of the BGP route with the next hop being 10.0.3.3 is 1 (default value). Therefore, R1 preferentially selects the BGP route with the next hop being 10.0.3.3.
- The following information is displayed in the detailed routing information of R1:

  not preferred for IGP cost

  The route is not selected because of the IGP cost.

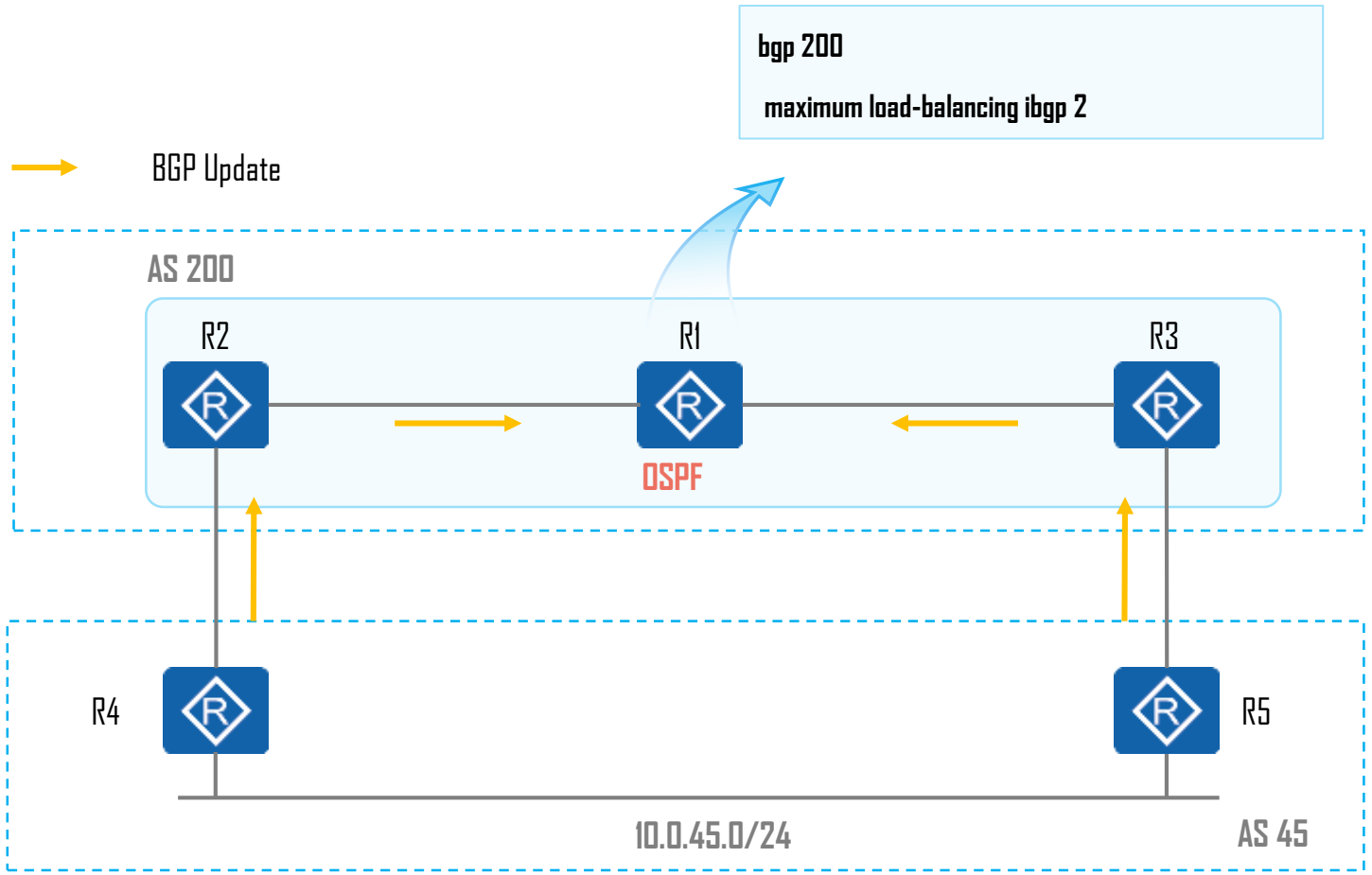
Load Balancing Among BGP Routes

- On a large network, there may be multiple valid BGP routes to the same destination. The device will select and add the optimal BGP route to its routing table for traffic forwarding.
- This, however, will result in uneven load balancing of much traffic. Configuring BGP load balancing can enable the device to add these multiple equal-cost BGP routes to its routing table, implementing traffic load balancing and reducing network congestion.
- After BGP load balancing is configured, the device will still select the optimal route among the multiple routes and advertise only this route to its peers.
- After BGP load balancing is enabled on a device, only the BGP routes that meet specified conditions can be used as equal-cost routes for load balancing.

Conditions for Load Balancing Among Equal-Cost BGP Routes

- The Preferred-Value attribute values are the same.
- The Local_Preference attribute values are the same.
- All the routes are summarized or non-summarized routes.
- AS_Path attribute values are the same.
- Origin types (IGP, EGP, or incomplete) are the same.
- The MED attribute values are the same.
- All the routes are EBGP or IBGP routes.
- The IGP metric values within an AS are the same.
- AS_Path attribute values are the same.

Configuring BGP Load Balancing



In the figure, if no routing policy or configuration is performed for the two BGP routes on R1, the first eight rules cannot determine the optimal route. Therefore, you can configure load balancing among IBGP routes.

Verifying the Configuration of Load Balancing Among BGP Routes

```
[R1]display ip routing-table 10.0.45.0 24
Route Flags: R - relay, D - download to fib
------------------------------------------------------------------------------
Routing Table : Public
Summary Count : 2
Destination/Mask    Proto   Pre  Cost      Flags NextHop         Interface

10.0.45.0/24        IBGP    255  0           RD  10.0.2.2        GigabitEthernet0/0/0
                    IBGP    255  0           RD  10.0.3.3        GigabitEthernet0/0/1
```

The equal-cost routes to 10.0.45.0/24 exist in the IP routing table.

```
[R1]display bgp routing-table

 BGP Local router ID is 10.0.1.1
 Status codes: * - valid, > - best, d - damped,
               h - history,  i - internal, s - suppressed, S - Stale
               Origin : i - IGP, e - EGP, ? - incomplete
 Total Number of Routes: 2
      Network            NextHop        MED        LocPrf    PrefVal Path/Ogn

 *>i  10.0.45.0/24       10.0.2.2             0          100          0      45?
 * i                     10.0.3.3             0          100          0      45?
```

There is only one optimal route in the BGP routing table.

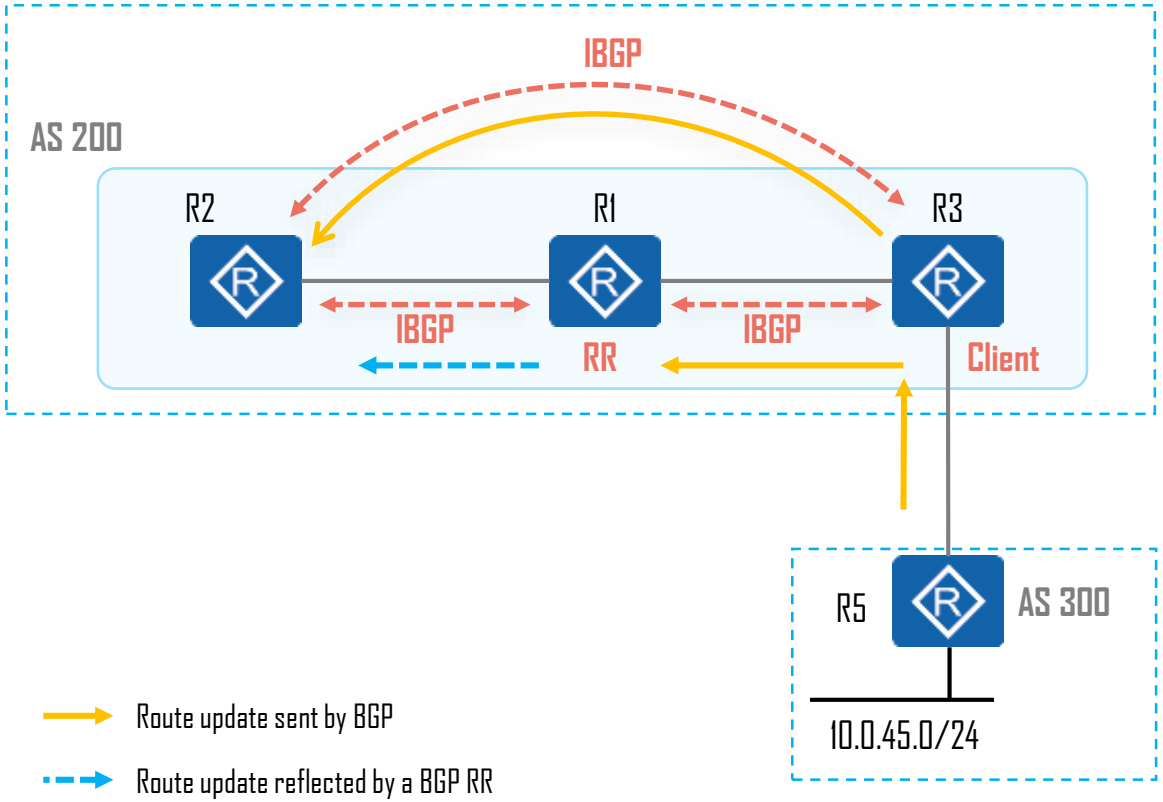
Rules for Selecting a Preferred BGP Route

When multiple routes to the same destination network segment exist, BGP selects routes in the following sequence:

Discards the route whose next hop is unreachable.

1. Prefers the route with the largest Preferred-Value attribute value.
2. Prefers the route with the largest Local_Preference value.
3. Prefers the locally originated BGP route, which takes precedence over the route learned from a peer. The locally summarized route, automatically summarized route, route learned by using the network command, route learned by using the import-route command, and route learned from a peer are in descending order of priority.
4. Prefers the route with the shortest AS_Path.
5. Prefers the route with the shortest AS_Path. The routes with Origin attributes of IGP, EGP, and Incomplete are in descending order of priority.
6. Prefers the route with the lowest MED.
7. Prefers routes learned from EBGP peers to routes learned from IBGP peers.
8. Prefers the route with the smallest IGP metric to the next hop.
9. **Prefers the route with the shortest Cluster_List length.**
10. Prefers the route advertised by the device with the smallest router ID (Originator_ID).
11. Prefers the route learned from the peer with the smallest IP address.

Preferring the Route with the Shortest Cluster_List Length (1)

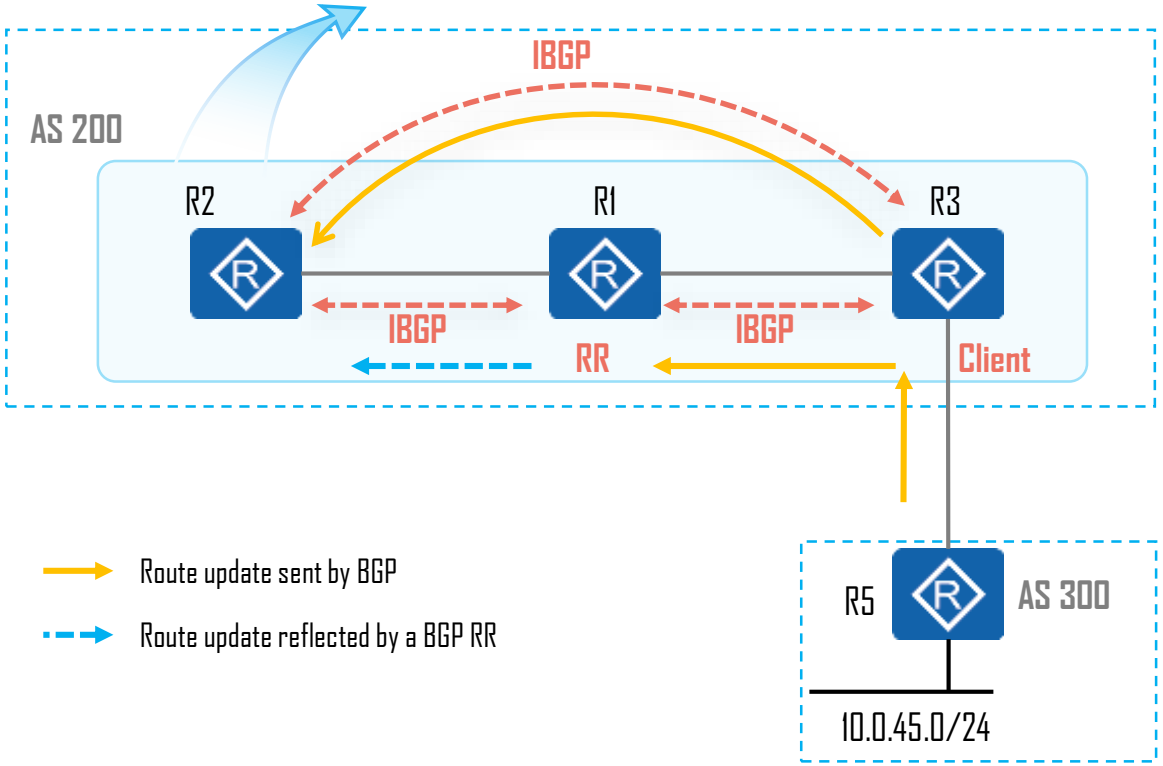


- The following configuration is performed:
  - Configure only R5 to advertise the route 10.0.45.0/24 to BGP.
  - Configure R1 as the RR and R3 as the client of R1.
  - Establish an IBGP peer relationship between R2 and R3 based on loopback interfaces.
- R2 receives the BGP route 10.0.45.0/24 advertised by R3 and the BGP route 10.0.45.0/24 reflected by R1.
- By default, the preceding rules cannot determine the optimal route. In this case, the route is selected based on Cluster_List.

Preferring the Route with the Shortest Cluster_List Length (2)

```
Total Number of Routes: 2
     Network         NextHop   MED    LocPrf    PrefVal Path/Ogn

*>i  10.0.45.0/24    10.0.3.3    0       100       0     300?
* i                  10.0.3.3    0       100       0     300?
```



Based on the BGP routing table, it cannot be determined whether the BGP route reflected by R1 or advertised by R3 is preferred. To check detailed information about BGP routes, run the **display bgp routing 10.0.45.0 24** command.

Preferring the Route with the Shortest Cluster_List Length (3)

BGP routing table entry information of 10.0.45.0/24:

From: 10.0.1.1 (10.0.1.1)

Route Duration: 00h03m10s

Relay IP Nexthop: 10.0.12.1

Relay IP Out-Interface: GigabitEthernet0/0/0

Original nexthop: 10.0.3.3

Qos information : 0x0

AS-path 300, origin incomplete, MED 0, localpref 100, pref-val 0, valid, internal, pre 255, IGP cost 2, **not preferred for Cluster List**

Originator: 10.0.3.3

Cluster list: 10.0.1.1

Not advertised to any peer yet

- The route reflected by R1 is not the optimal route due to the following reason: not preferred for Cluster List
- The BGP route that R3 directly advertises to R2 does not pass through the RR and therefore does not have the Cluster_List attribute. That is, the Cluster_List attribute value of the BGP route advertised by R3 is considered as 0, which is smaller than the Cluster_List attribute value (1) of the BGP route reflected by R1. Therefore, the BGP route advertised by R3 is preferred.

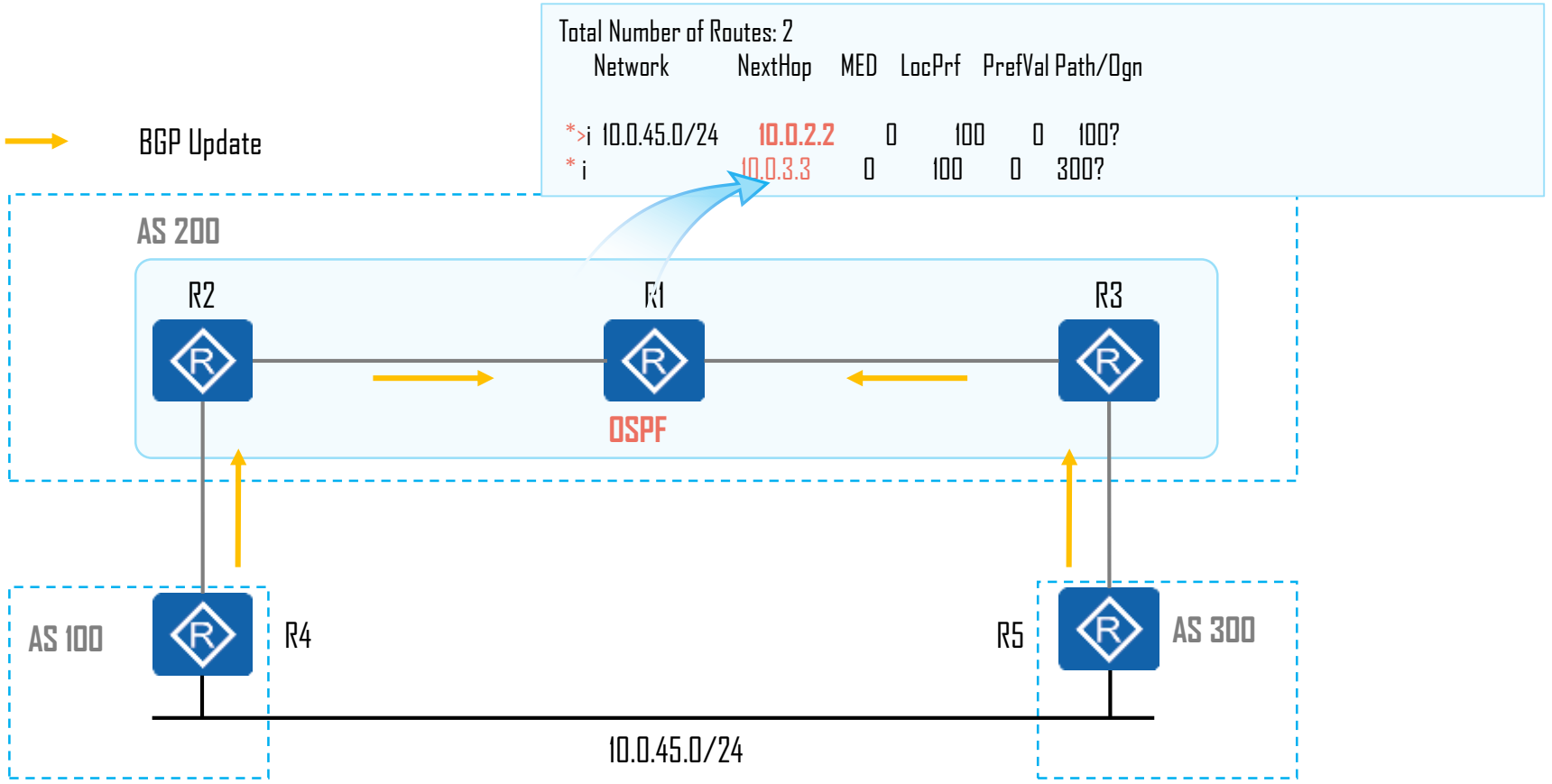
Rules for Selecting a Preferred BGP Route

When multiple routes to the same destination network segment exist, BGP selects routes in the following sequence:

Discards the route whose next hop is unreachable.

1. Prefers the route with the largest Preferred-Value attribute value.
2. Prefers the route with the largest Local_Preference value.
3. Prefers the locally originated BGP route, which takes precedence over the route learned from a peer. The locally summarized route, automatically summarized route, route learned by using the network command, route learned by using the import-route command, and route learned from a peer are in descending order of priority.
4. Prefers the route with the shortest AS_Path.
5. Prefers the route with the shortest AS_Path. The routes with Origin attributes of IGP, EGP, and Incomplete are in descending order of priority.
6. Prefers the route with the lowest MED.
7. Prefers routes learned from EBGP peers to routes learned from IBGP peers.
8. Prefers the route with the smallest IGP metric to the next hop.
9. Prefers the route with the shortest Cluster_List length.
10. **Prefers the route advertised by the device with the smallest router ID (Originator_ID).**
11. Prefers the route learned from the peer with the smallest IP address.

Preferring the Route with the Smallest Router ID (1)



In the preceding topology, R1 receives BGP route 10.0.45.0/24 from both R2 and R3 by default, and the preceding route selection rules cannot determine the optimal route. Therefore, R1 selects the BGP route advertised by the peer with the smallest router ID based on the preceding route selection rules. In this example, the BGP route advertised by R2 is preferred.

Preferring the Route with the Smallest Router ID (2)

BGP routing table entry information of 10.0.45.0/24:

From: 10.0.3.3 (10.0.3.3)

Route Duration: 00h40m15s

Relay IP Nexthop: 10.0.13.3

Relay IP Out-Interface: GigabitEthernet0/0/1

Original nexthop: 10.0.3.3

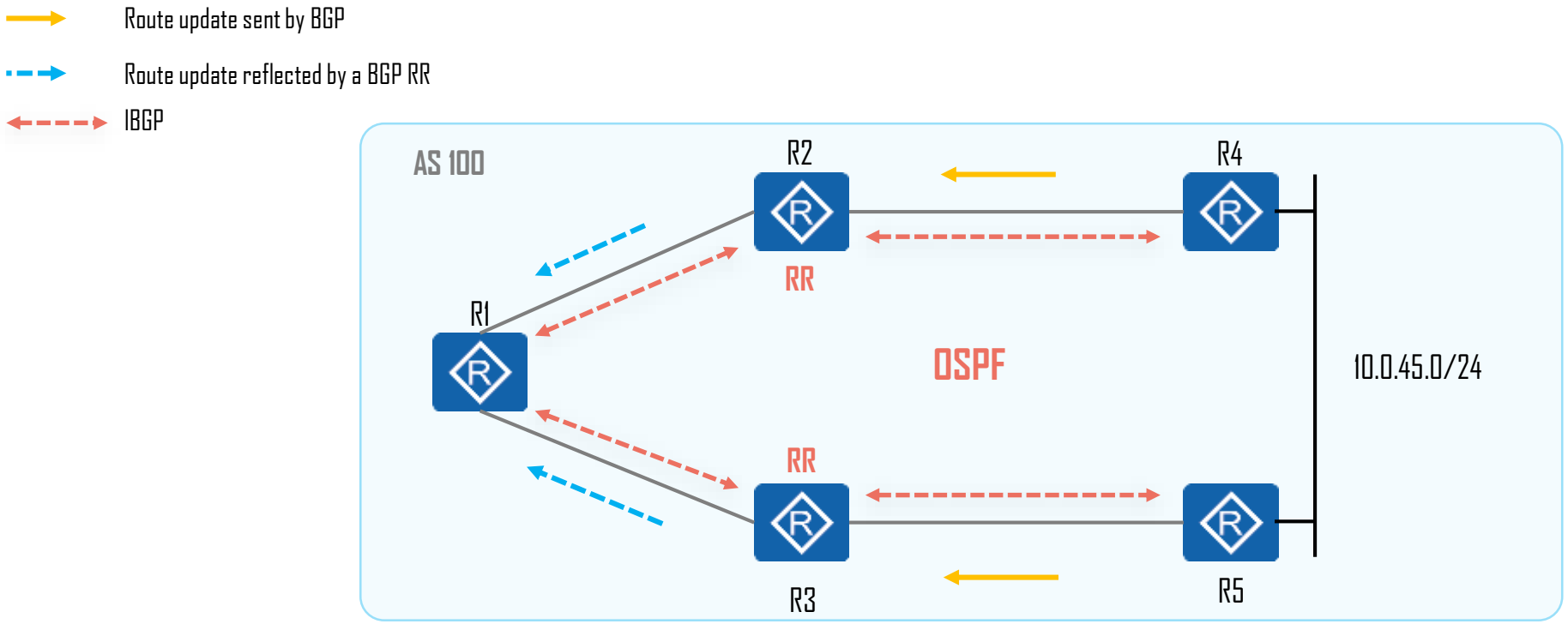
Qos information : 0x0

AS-path 300, origin incomplete, MED 0, localpref 100, pref-val 0, valid, internal, pre 255, IGP cost 1, **not preferred for router ID**

Not advertised to any peer yet

Check detailed information about the BGP routing table of R1. The command output shows that the BGP route from 10.0.3.3 is not preferred because of the router ID.

Preferring the Route with the Smallest Originator_ID (1)

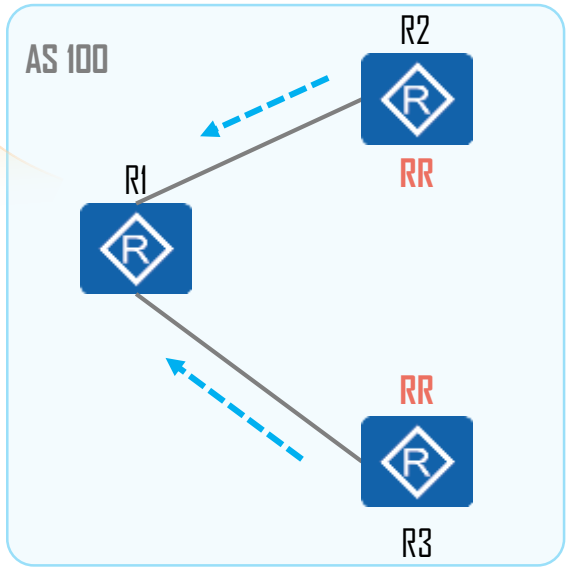


If a BGP route carries the Originator_ID attribute, the router compares the Originator_ID values of the routes and selects the BGP route with the smallest Originator_ID value.

Preferring the Route with the Smallest Originator_ID (2)

```
Total Number of Routes: 2
     Network         NextHop     MED    LocPrf    PrefVal Path/Ogn

 *>i 10.0.45.0/24    10.0.4.4      0       100       0       ?
 * i                 10.0.5.5      0       100       0       ?
```



Route update sent by BGP

Route update reflected by a BGP RR

```
BGP routing table entry information of 10.0.45.0/24:
From: 10.0.3.3 (10.0.3.3)
Route Duration: 00h33m15s
Relay IP Nexthop: 10.0.13.3
Relay IP Out-Interface: GigabitEthernet0/0/1
Original nexthop: 10.0.5.5
Qos information : 0x0
AS-path Nil, origin incomplete, MED 0, localpref 100, pref-val 0, valid, internal, pre 255, IGP cost 2, not
preferred for router ID
Originator: 10.0.5.5
Cluster list: 10.0.3.3
Not advertised to any peer yet
```

The BGP route reflected by R3 is not selected because of the router ID. The router ID here refers to the Originator ID (router ID of the original route advertiser).

Rules for Selecting a Preferred BGP Route

When multiple routes to the same destination network segment exist, BGP selects routes in the following sequence:

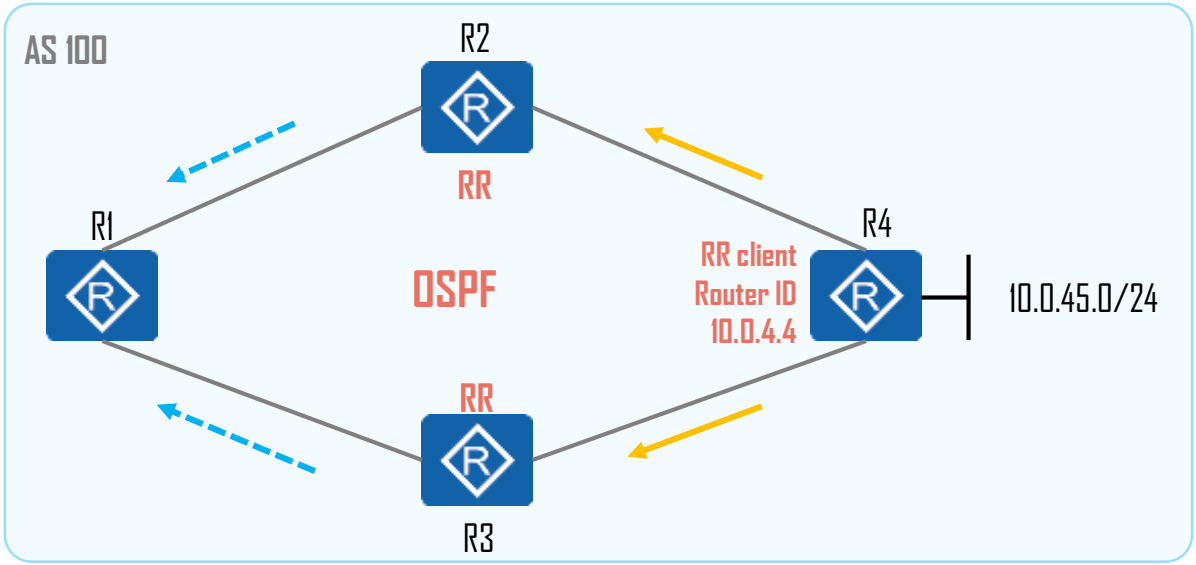
Discards the route whose next hop is unreachable.

1. Prefers the route with the largest Preferred-Value attribute value.
2. Prefers the route with the largest Local_Preference value.
3. Prefers the locally originated BGP route, which takes precedence over the route learned from a peer. The locally summarized route, automatically summarized route, route learned by using the network command, route learned by using the import-route command, and route learned from a peer are in descending order of priority.
4. Prefers the route with the shortest AS_Path.
5. Prefers the route with the shortest AS_Path. The routes with Origin attributes of IGP, EGP, and Incomplete are in descending order of priority.
6. Prefers the route with the lowest MED.
7. Prefers routes learned from EBGP peers to routes learned from IBGP peers.
8. Prefers the route with the smallest IGP metric to the next hop.
9. Prefers the route with the shortest Cluster_List length.
10. Prefers the route advertised by the device with the smallest router ID (Originator_ID).
11. **Prefers the route learned from the peer with the smallest IP address.**

Preferring the Route from the Device with the Smallest IP Address (1)

Route update sent by BGP

Route update reflected by a BGP RR

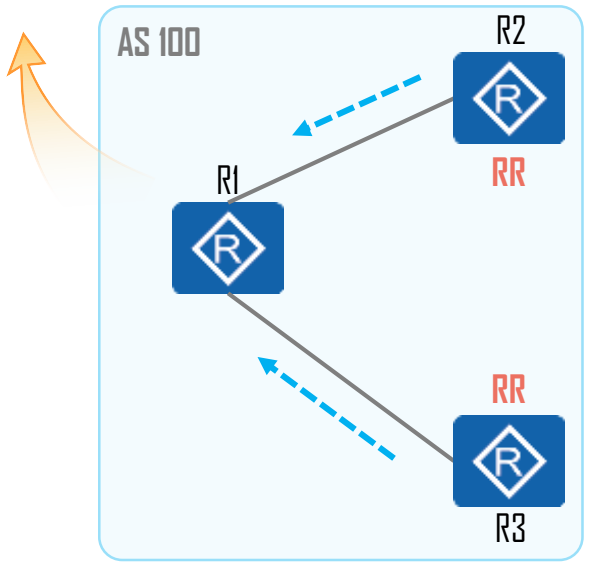


- If the preceding rules cannot determine the optimal route, the route from the device with the smallest IP address is preferred.
- In the preceding topology, R2 and R3 are connected to R4. R4 functions as the RR client and advertises routes to BGP only on R4. In this case, the BGP routes reflected by R2 and R3 have the same Originator ID 10.0.4.4.

Preferring the Route from the Device with the Smallest IP Address (2)

```
Total Number of Routes: 2
     Network         NextHop     MED    LocPrf    PrefVal Path/Ogn

 *>i 10.0.45.0/24    10.0.4.4      0        100       0        ?
 * i                  10.0.4.4      0        100       0        ?
```



Route update sent by BGP

Route update reflected by a BGP RR

```
BGP routing table entry information of 10.0.45.0/24:
From: 10.0.3.3 (10.0.3.3)
Route Duration: 00h01m07s
Relay IP Nexthop: 10.0.12.2
Relay IP Out-Interface: GigabitEthernet0/0/0
Original nexthop: 10.0.4.4
Qos information : 0x0
AS-path Nil, origin incomplete, MED 0, localpref 100, pref-val 0, valid, intern
al, pre 255, IGP cost 2, not preferred for peer address
Originator:  10.0.4.4
Cluster list: 10.0.3.3
Not advertised to any peer yet
```

The BGP route reflected by R3 is not selected because the peer address is larger. The peer address of the route reflected by R2 is 10.0.2.2, and the peer address of the route reflected by R3 is 10.0.3.3. Therefore, the BGP route reflected by R3 is not selected.

Quiz

1. (Essay) When a BGP route received from an EBGP peer is advertised to an IBGP peer, how does the EBGP peer change the value of the Next_Hop attribute to its own source address?
2. (TorF) If the preceding three rules are the same, BGP compares the AS_Path attribute values. If the AS_Path attribute values are the same, BGP compares the AS numbers.

Summary

- BGP selects optimal routes based on path attributes. This allows BGP to select optimal routes based on path attributes in different scenarios.
- BGP defines a set of detailed optimal path selection algorithms, which enable routers to select the optimal path in any complex and highly redundant network environment.
- BGP route selection rules are frequently used in practice and need to be mastered.