BGP 路由协议详解

制作人：张选波

二〇〇九年六月二十二日

一、BGP 的概况

BGP 最新的版本是 BGP 第 4 版本（BGP4）,它是在 RFC4271 中定义的；一个路由器只能属于一个 AS。AS 的范围从 1-65535（64512-65535 是私有 AS 号），RFC1930 提供了 AS 号使用指南。

BGP 的主旨是提供一种域间路由选择系统，确保自主系统只能够无环地交换路由选择信息，BGP 路由器交换有关前往目标网络的路径信息。

BGP 是一种基于策略的路由选择协议，BGP 在确定最佳路径时考虑的不是速度，而是让 AS 能够根据多种 BGP 属性来控制数据流的传输。

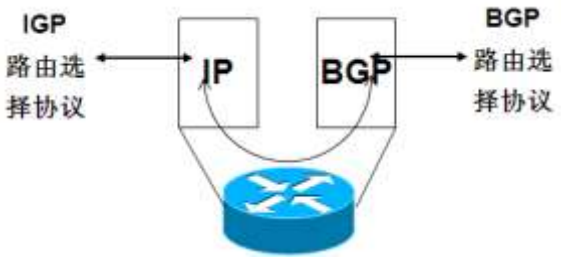
1、BGP 的特性

- BGP将传输控制协议（TCP）用作其传输协议。是可靠传输，运行在TCP的179端口上（目的端口）
- 由于传输是可靠的，所以BGP0使用增量更新，在可靠的链路上不需要使用定期更新，所以BGP使用触发更新。
- 类似于OSPF和ISIS路由协议的Hello报文，BGP使用keepalive周期性地发送存活消息（60s）（维持邻居关系）。
- BGP在接收更新分组的时候，TCP使用滑动窗口，接收方在发送方窗口达到一半的时候进行确定，不同于OSPF等路由协议使用1-to-1窗口。
- 丰富的属性值
- 可以组建可扩展的巨大的网络

2、BGP 的三张表

- 邻居关系表
  - 所有BGP邻居
- 转发数据库
  - 记录每个邻居的网络
  - 包含多条路径去往同一目的地，通过不同属性判断最好路径
  - 数据库包括BGP属性
- 路由表
  - 最佳路径放入路由表中
  - EBGP路由（从外部AS获悉的BGP路由）的管理距离为20
  - IBGP路由（从AS系统获悉的路由）管理距离为200

如下图所示。



- 邻居表，包含与之建立BGP连接的邻居
  - 使用命令show ip bgp summary可以查看到

```
Router#sh ip bgp summary
BGP router identifier 11.1.1.1, local AS number 100
BGP table version is 8, main routing table version 8
5 network entries using 585 bytes of memory
6 path entries using 312 bytes of memory
4/3 BGP path/bestpath attribute entries using 496 bytes of memory
1 BGP AS-PATH entries using 24 bytes of memory
0 BGP route-map cache entries using 0 bytes of memory
0 BGP filter-list cache entries using 0 bytes of memory
BGP using 1417 total bytes of memory
BGP activity 5/0 prefixes, 6/0 paths, scan interval 60 secs

Neighbor        V    AS MsgRcvd MsgSent   TblVer  InQ OutQ Up/Down  State/PfxRcd
10.1.1.1        4   100      14      18        8    0    0 00:09:32        2
11.1.1.2        4   200      12      16        8    0    0 00:07:03        1
```

- 转发表，从邻居那里获悉的的所有路由都被加入到BGP转发表中。
  - 使用命令show ip bgp可以查看

```
Router#sh ip bgp
BGP table version is 8, local router ID is 11.1.1.1
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
              r RIB-failure, S Stale
Origin codes: i - IGP, e - EGP, ? - incomplete

   Network          Next Hop            Metric LocPrf Weight Path
*> 10.1.1.0/24      0.0.0.0                  0         32768 i
* i                 10.1.1.1                 0    100      0 i
*> 11.1.1.0/24      0.0.0.0                  0         32768 i
*>i192.168.1.0      10.1.1.1                 0    100      0 i
*> 192.168.2.0      0.0.0.0                  0         32768 i
*> 192.168.3.0      11.1.1.2                 0             0 200 i
```

- 路由表，BGP路由选择进程从BGP转发表中选出前往每个网络的最佳路由，并加入到路由表中。
  - 使用命令show ip route bgp可以查看

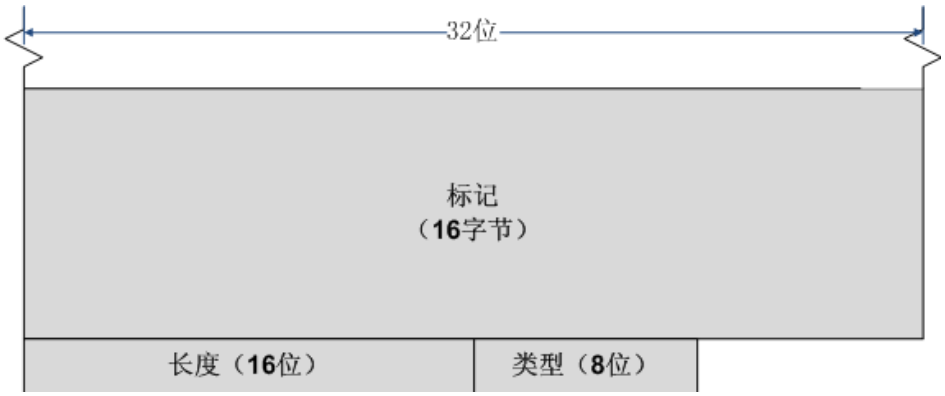
```
Router#sh ip route bgp
B       192.168.1.0/24 [200/0] via 10.1.1.1, 00:13:11
B       192.168.3.0/24 [20/0] via 11.1.1.2, 00:11:19
```

3、BGP 消息类型

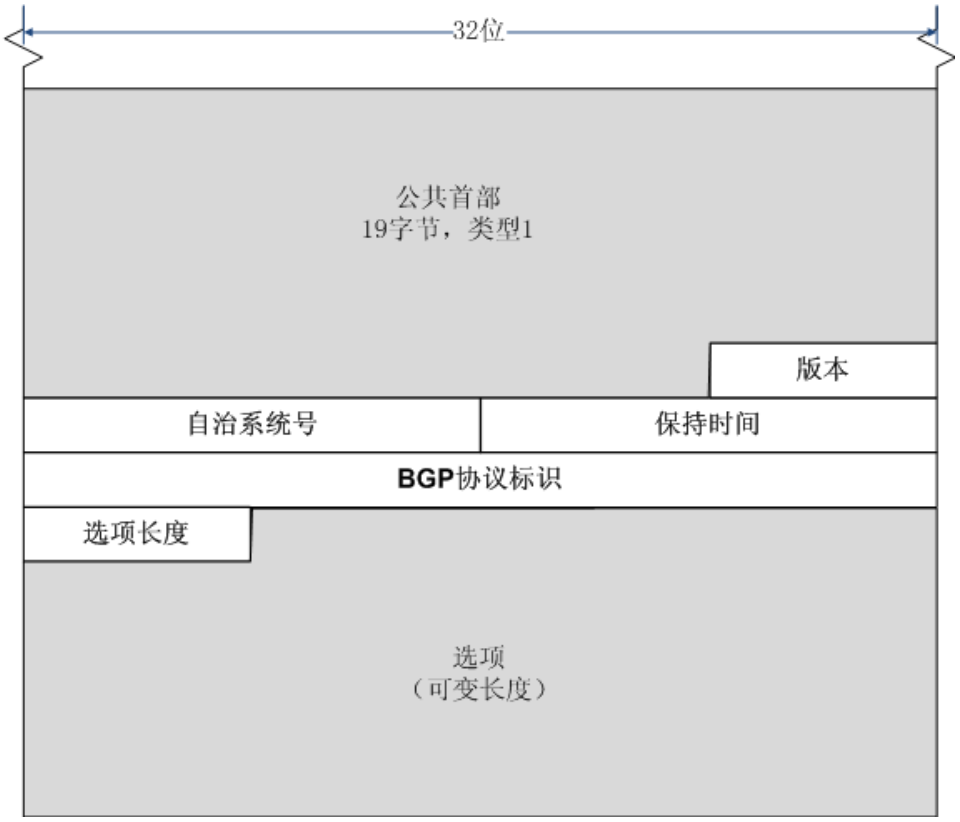
- open：用来建立最初的BGP连接。（包含hold-time,router-id）
- Keepalive：对等体之间周期性的交换这些消息以保持会话有效。（默认60秒）
- Update：对等体之间使用这些消息来交换网络层可达性信息。
- Notification：这些消息用来通知出错信息。

所有的BGP分组共享同样的公有首部，在学习不同类型的分组之前，先讨论公共首部，如下图所示，这个首部的字段如下。

- **标记**：这个16字节标记字段保留给鉴别用
- **长度**：这个2字节字段定义包括首部在内的报文总长度
- **类型**：这个1字节段定义分组的类型，用数值1至4定义BGP消息类型

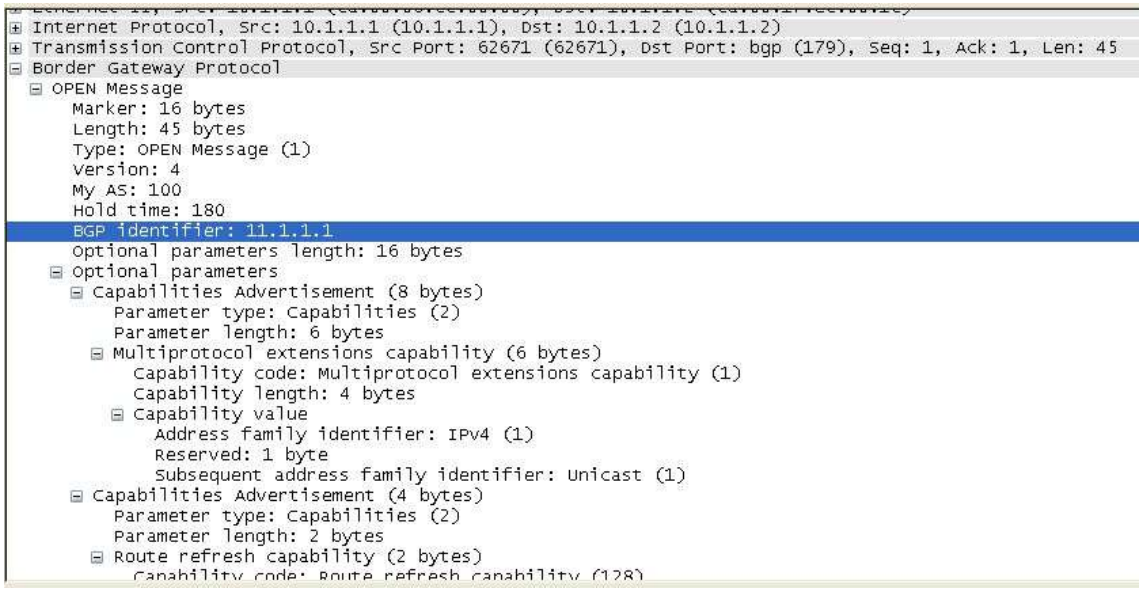


打开消息：主要是利用此报文建立邻居，运行BGP的路由器打开与邻居的TCP连接，并发送打开报文，如果邻居接受这种邻居关系，由响应保活报文。打开报文格式如下所示。

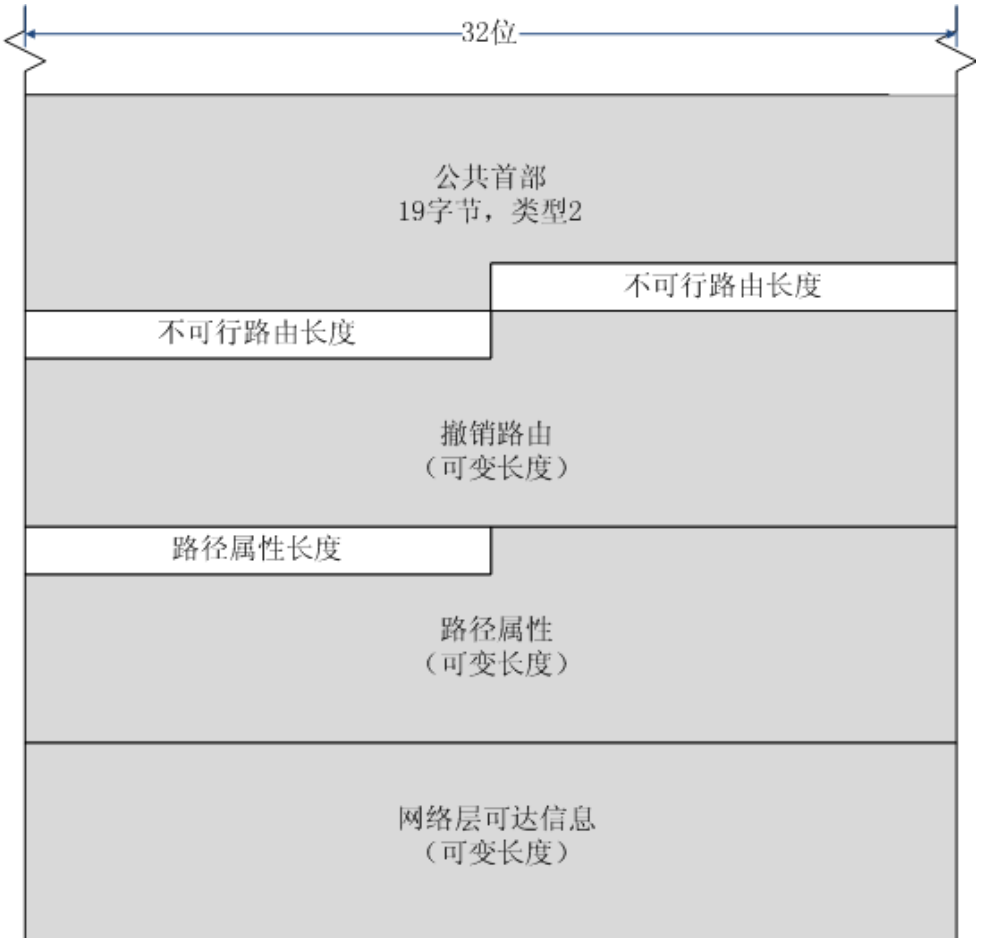


- **版本：**这个1字节字段定义BGP的版本，当前的版本是4
- **自治系统：**这个2字节字段定义自治系统号。
- **保持时间：**这个2字节字段定义一方从另一方收到保活报文或更新报文之前所经过的最大秒数，若路由器在保持时间的期间内没有收到这些报文中的一个，就认为对方是不工作的。
- **BGP协议标识：**这是2字节字段，这定义发送打开报文的路由器，为此，这个路由器通常使用它的IP地址中的一个作为BGP标识符。
- **选项长度：**打开报文还可以包含某些选项参数，若包含，则这个1字节字段定义选项参数总长度，若没有选项参数，则这个字段的值为0
- **选项参数：**若选项参数长度的值不是0,则表示有某些选项参数，每一个选项参数本身又有两个字段，参数长度和参数值，到现在已定义的唯一的选项参数是鉴别。

如下图是采用ethereal采集到的BGP的打开消息报文。



更新报文：更新报文是BGP协议的核心，路由器使用它来撤销以前已通知的终点和宣布到一个新终点的路由，或两者都有，应该注意：BGP可以撤销好几个在以前曾通知过的终点，但在单个更新报文中则只能通知一个新终点，如下所示。



- **不可行路由长度：**这个2字节字段定义下一字段的长度。
- **撤销路由：**这个字段列出必须从以前通知的清单中删除的所有路由
- **路径属性长度：**这个2字节字段定义下一个字段的长度
- **路径属性：**这个字段定义到这个报文宣布可达性的网络路径属性
- **网络层可达性信息：**这个字段定义这个报文真正通知的网络。它有一个长度字段和一个IP地址前缀，长度定义前缀中的位数。前缀定义这个网络地址的共同部分。例如，若这个网络是123.1.10.0/24，则网络前缀是24而前缀是123.1.10。

下图为，是采用ethereal采集到的BGP的更新消息报文。

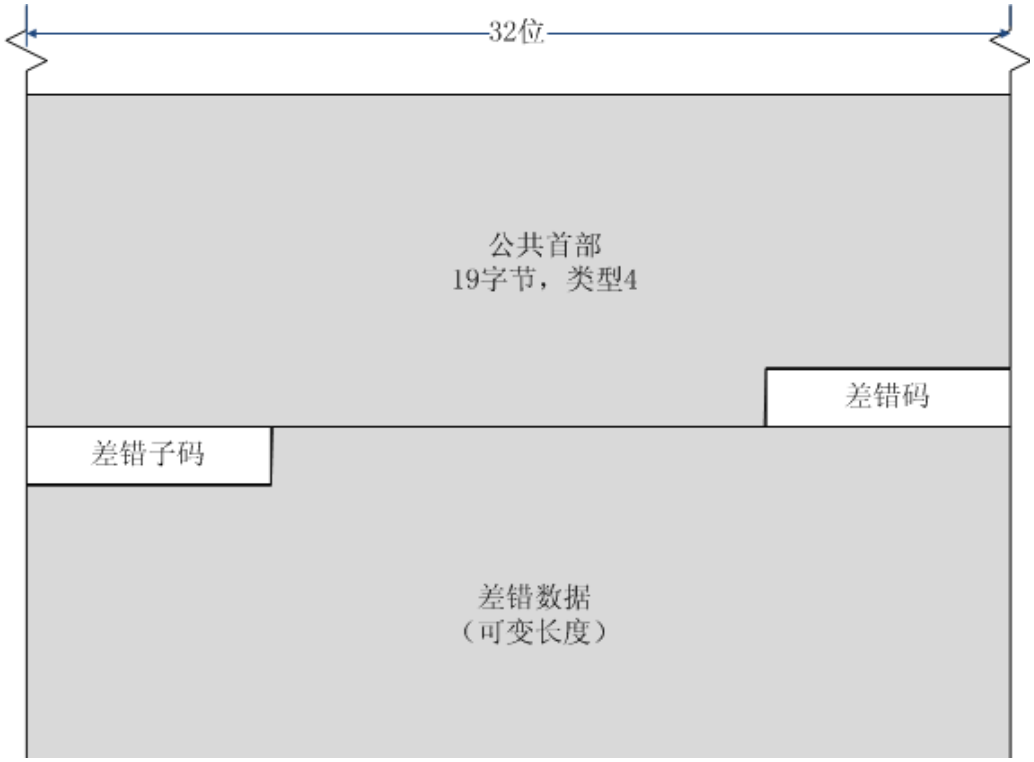
```
⊞ Transmission Control Protocol, Src Port: bgp (179), Dst Port: 62671 (62671), Seq: 65, Ack: 65, Len: 59
⊟ Border Gateway Protocol
  ⊟ UPDATE Message
      Marker: 16 bytes
      Length: 59 bytes
      Type: UPDATE Message (2)
      Unfeasible routes length: 0 bytes
      Total path attribute length: 32 bytes
    ⊟ Path attributes
      ⊟ ORIGIN: IGP (4 bytes)
        ⊞ Flags: 0x40 (Well-known, Transitive, Complete)
          Type code: ORIGIN (1)
          Length: 1 byte
          Origin: IGP (0)
      ⊟ AS_PATH: 200 (7 bytes)
        ⊟ Flags: 0x40 (Well-known, Transitive, Complete)
            0... .... = Well-known
            .1.. .... = Transitive
            ..0. .... = Complete
            ...0 .... = Regular length
          Type code: AS_PATH (2)
          Length: 4 bytes
        ⊟ AS path: 200
          ⊟ AS path segment: 200
              Path segment type: AS_SEQUENCE (2)
              Path segment length: 1 AS
              Path segment value: 200
      ⊟ NEXT_HOP: 10.1.2.2 (7 bytes)
        ⊟ Flags: 0x40 (Well-known, Transitive, Complete)
            0... .... = Well-known
            .1.. .... = Transitive
            ..0. .... = Complete
            ...0 .... = Regular length
          Type code: NEXT_HOP (3)
          Length: 4 bytes
          Next hop: 10.1.2.2 (10.1.2.2)
      ⊟ MULTI_EXIT_DISC: 0 (7 bytes)
        ⊟ Flags: 0x80 (Optional, Non-transitive, Complete)
            1... .... = Optional
            .0.. .... = Non-transitive
            ..0. .... = Complete
            ...0 .... = Regular length
          Type code: MULTI_EXIT_DISC (4)
          Length: 4 bytes
          Multiple exit discriminator: 0
      ⊟ LOCAL_PREF: 100 (7 bytes)
        ⊟ Flags: 0x40 (Well-known, Transitive, Complete)
            0... .... = Well-known
            .1.. .... = Transitive
            ..0. .... = Complete
            ...0 .... = Regular length
          Type code: LOCAL_PREF (5)
          Length: 4 bytes
          Local preference: 100
    ⊟ Network layer reachability information: 4 bytes
      ⊟ 12.1.1.0/24
          NLRI prefix length: 24
          NLRI prefix: 12.1.1.0 (12.1.1.0)
```

保活报文：是用来告诉对方自己是工作的，保活报文只包括公共首部，如下图所示。

下图为，是采用ethereal采集到的BGP的保活报文。

```
⊞ Frame 85 (73 bytes on wire, 73 bytes captured)
⊞ Ethernet II, Src: ca:00:17:ec:00:1d (ca:00:17:ec:00:1d), Dst: ca:00:17:14:00:1c (ca:00:17:14:00:1c)
⊞ Internet Protocol, Src: 10.1.2.1 (10.1.2.1), Dst: 10.1.2.2 (10.1.2.2)
⊟ Transmission Control Protocol, Src Port: 23502 (23502), Dst Port: bgp (179), Seq: 0, Ack: 0, Len: 19
    Source port: 23502 (23502)
    Destination port: bgp (179)
    Sequence number: 0    (relative sequence number)
    [Next sequence number: 19    (relative sequence number)]
    Acknowledgement number: 0    (relative ack number)
    Header length: 20 bytes
  ⊞ Flags: 0x0018 (PSH, ACK)
    Window size: 15679
    Checksum: 0xbb28 [correct]
⊟ Border Gateway Protocol
  ⊟ KEEPALIVE Message
      Marker: 16 bytes
      Length: 19 bytes
      Type: KEEPALIVE Message (4)
```

通知报文：当检测出差错状态或路由器打算关闭连接时，路由器就发送通知报文，如下图所示。



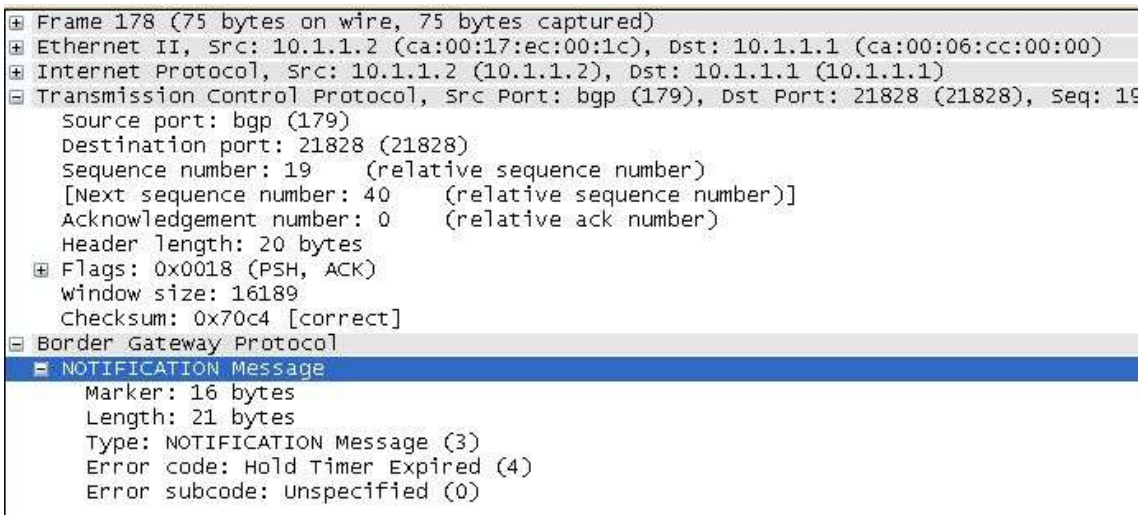
- **差错码：**这个1字节字段定义差错的种类
- **差错子码：**这个1字节字段进一步定义每一种差错的类型
- **差错数据：**这个字段可用来给出关于该差错的更多的诊断信息

具体的差错码，如下表所示。

差错码	差错码说明	差错子码说明
1	报文首部差错	3种不同的子码：同步问题（1），坏的报文长度（2），坏的报文类型（3）

2	打开报文差错	6种不同的子码：不支持的版本（1），坏的对等AS（2），坏的BGP标识符（3），不支持的可选参数（4），鉴别失败（5），不可接受的保持时间（6）
3	更新报文差错	11种不同的子码：错误形成的属性表（1），不能识别的熟知属性（2），丢失熟知属性（3），属性标志差错（4），属性长度差错（5），非法起点属性（6），AS路由选择环路（7），无效的下一路属性（8），可选属性差错（9），无效的网络字段（10），错误形成的AS_PATH（11）
4	保持计时器截止期到	未定义子码
5	有限状态机差错	定义过程的差错，未定义子码
6	关闭	未定义子码

下图为，是采用ethereal采集到的BGP的通知报文。



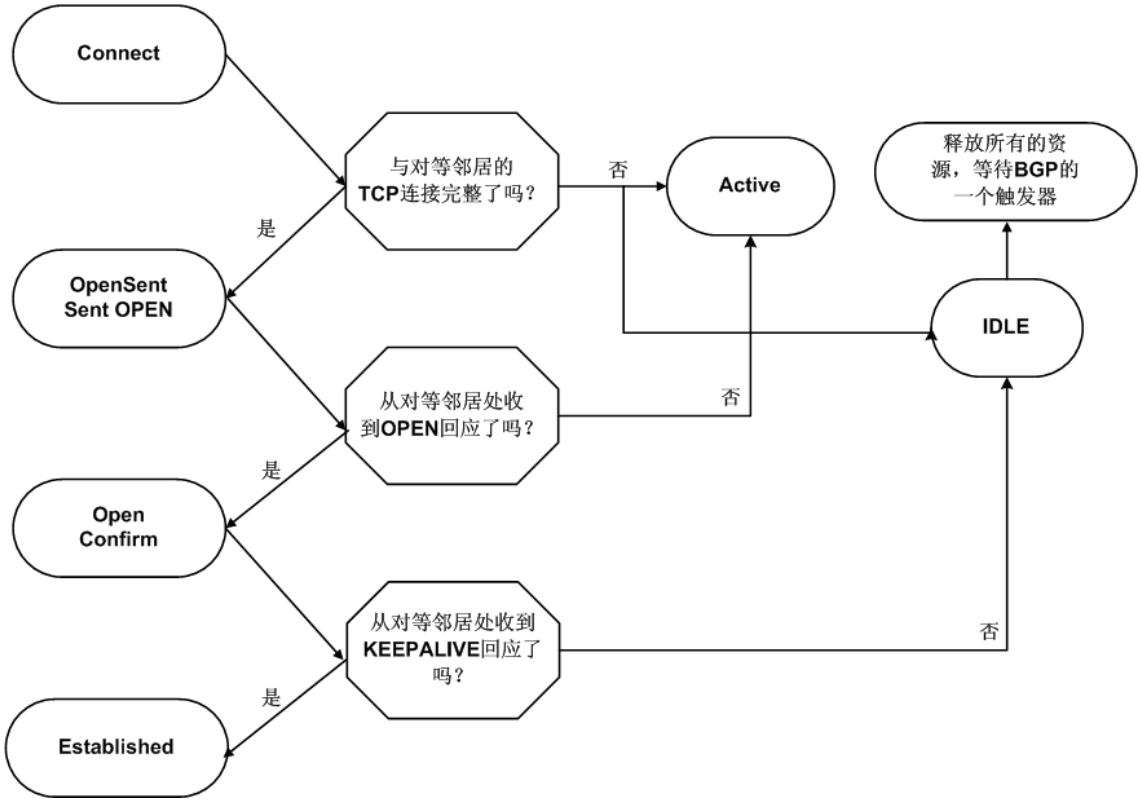
4、建立邻居的过程

在两个BGP发言人交换信息之前，BGP都要求建立邻居关系，BGP不是动态地发现所感兴趣的运行BGP的路由器，相反，BGP使用一个特殊的邻居IP地址来配置的。

BGP使用周期性的Keepalive分组来确认BGP邻居的可访问性。

Keepalive计时器是保持时间（Hold Time）的三分之一，如果发给某一特定BGP邻居三个连续的Keepalive分组都丢失的话，保持时间计时器超时，那个邻居被视为不可达，RFC1771对保持时间的建议是90秒，Keepalive计时器的建议值是30秒。

按照RFC1771，BGP建立邻居关系要经历以下几个阶段，如下图所示。



- **Idle**：在此状态下不分配网络资源，不允许传入的BGP连接。当在持续性差错条件下，经常性的重启会导致波动。因此，在第一次进入到空闲状态后，路由器会设置连接重试定时器，在定时器到期时才会重新启动BGP，思科的初始连接重试时间为60秒，以后每次连接重试时间都是之前的两倍，也就是说，连接等待时间呈指数关系递增。
- **Connect:**(已经建立完成了TCP三次握手)，BGP等待TCP连接完成，如果连接成功，BGP在发送了OPEN分组给对方之后，状态机变为OpenSent状态，如果连接失败，根据失败的原因，状态机可能演变到Active，或是保持Connect，或是返回Idle。
- **Active:**在这个状态下，初始化一个TCP连接来建立BGP间的邻居关系。如果连接成功，BGP在发送了OPEN分组给对方之后，状态机变为OpenSent状态，如果连接失败，可能仍处在Active状态或返回Idle状态。
- **OpenSent**：BGP发送OPEN分组给对方之后，BGP在这一状态下等待OPEN的回应分组，如果回应分组成功收到，BGP状态变为OpenConfirm，并给对方发送一条Keepalive分组，如果没有接到回应分组，BGP状态重新变为Idle或是Active。
- **OpenConfirm**：这时，距离最后的Established状态只差一步，BGP在这个状态下等待对方的Keepalive分组，如果成功接收，状态变为Established，否则，因为出现错误，BGP状态将重新变为Idle。
- **Established**：这是BGP对等体之间 可以交换信息的状态，可交换的信息包括UPDATE分组、KeepAlive分组和Notification分组。

connect和active都是TCP连接阶段，ACTIVE是发起方，connect是应答方。可以使用命令show ip bgp summary、debug ip bgp events、debug ip bgp来查看。

```
Router#debug ip bgp
BGP debugging is on for address family: IPv4 Unicast
*Jun 23 22:00:05.619: BGP: 11.1.1.2 went from Idle to Active
```

```
*Jun 23 22:00:05.627: BGP: 11.1.1.2 open active delayed 30128ms (35000ms max, 28% jitter)
*Jun 23 22:00:06.215: BGP: 11.1.1.2 passive open to 11.1.1.1
*Jun 23 22:00:06.219: BGP: 11.1.1.2 went from Active to Idle
*Jun 23 22:00:06.219: BGP: 11.1.1.2 went from Idle to Connect
*Jun 23 22:00:06.227: BGP: 11.1.1.2 rcv message type 1, length (excl. header) 26
*Jun 23 22:00:06.227: BGP: 11.1.1.2 rcv OPEN, version 4, holdtime 180 seconds
*Jun 23 22:00:06.231: BGP: 11.1.1.2 went from Connect to OpenSent
*Jun 23 22:00:06.231: BGP: 11.1.1.2 sending OPEN, version 4, my as: 100, holdtime 180 seconds
*Jun 23 22:00:06.231: BGP: 11.1.1.2 rcv OPEN w/ OPTION parameter len: 16
*Jun 23 22:00:06.231: BGP: 11.1.1.2 rcvd OPEN w/ optional parameter type 2 (Capability) len 6
*Jun 23 22:00:06.235: BGP: 11.1.1.2 OPEN has CAPABILITY code: 1, length 4
*Jun 23 22:00:06.235: BGP: 11.1.1.2 OPEN has MP_EXT CAP for afi/safi: 1/1
*Jun 23 22:00:06.235: BGP: 11.1.1.2 rcvd OPEN w/ optional parameter type 2 (Capability) len 2
*Jun 23 22:00:06.235: BGP: 11.1.1.2 OPEN has CAPABILITY code: 128, length 0
*Jun 23 22:00:06.239: BGP: 11.1.1.2 OPEN has ROUTE-REFRESH capability(old) for all address-families
*Jun 23 22:00:06.239: BGP: 11.1.1.2 rcvd OPEN w/ optional parameter type 2 (Capability) len 2
*Jun 23 22:00:06.239: BGP: 11.1.1.2 OPEN has CAPABILITY code: 2, length 0
*Jun 23 22:00:06.239: BGP: 11.1.1.2 OPEN has ROUTE-REFRESH capability(new) for all address-families
BGP: 11.1.1.2 rcvd OPEN w/ remote AS 200
*Jun 23 22:00:06.243: BGP: 11.1.1.2 went from OpenSent to OpenConfirm
*Jun 23 22:00:06.243: BGP: 11.1.1.2 send message type 1, length (incl. header) 45
*Jun 23 22:00:06.359: BGP: 11.1.1.2 went from OpenConfirm to Established
*Jun 23 22:00:06.363: %BGP-5-ADJCHANGE: neighbor 11.1.1.2 Up
```

5、建立 IBGP 邻居

IBGP 运行在 AS 内部，不需要直连。IBGP 有水平分割，建议使用 Full Mesh，由于 Full Mesh 不具有扩展性，为了解决 IBGP 的 Full Mesh 问题，使用路由反射器（RR）和联邦两种方法来解决。主要减少了 backbone IGP 中的路由。

Neighbor 后所指的地址可达。发起方不能是缺省路由，应答方不能是缺省路由。

可以使用下面两种方法来建立 IBGP 邻居：

- 邻居之间可以通过各自的一个物理接口建立对等关系，该对等关系是通过属于它们共享的子网的 IP 地址来建立的。
- 邻居之间也可以通过使用环回接口建立对等关系。

在 IBGP 中，由于假定了 IBGP 邻居在物理上直接相连的可能性不大，所以将 IP 分组头中的 TTL 域设置为 255。

6、建立 EBGP 邻居

EBGP 运行在 AS 与 AS 之间的边界路由器上，默认情况下，需要直连或使用静态路由，如果不是直连，必须指 EBGP 多跳，**Neighbor x.x.x.x ebgp-multihop [1-255]** 不选择为最大值，255 跳。

可以使用下面两种方法来建立 EBGP 邻居：

- 邻居之间可以通过各自的一个物理接口建立对等关系。
- 邻居之间也可以通过使用环回接口建立对等关系。

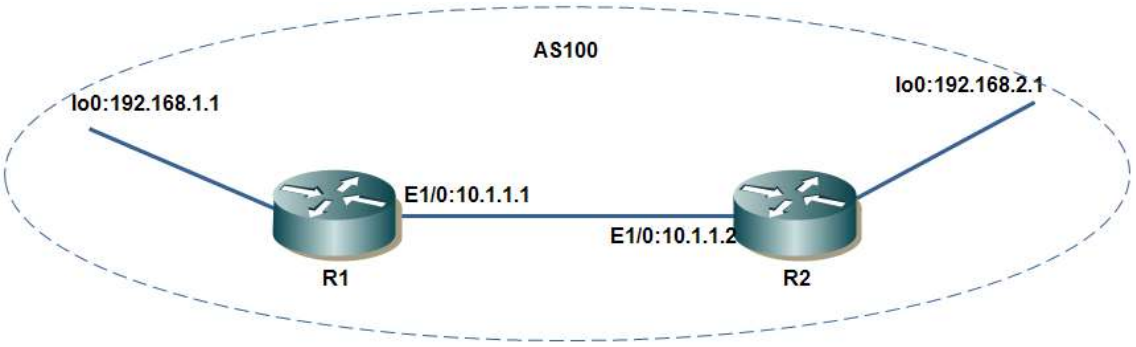
7、neighbor *ip-address* remote-as *number* 命令

例：**neighbor 10.1.1.1 remote-as 100**

指定对方属于哪一个AS。所指的10.1.1.1地址，必须在IGP中可达。

- 允许邻居用这个地址来访问我的 179 端口，但没有指明访问本路由器的哪个地址，只检查源地址。
- 本路由器以更新源地址去访问 neighbor 后面这个地址的 179 端口，是否可以建立 TCP 链接要看对方是否允许我的更新源来访问它。

示例：



R1/R2 两台路由器运行 RIPv2，都将环回口宣告进 RIP。这时假如在两台路由器之间运行 IBGP 邻居关系：

R1：neighbor 192.168.2.1 remote-as 1

R2：neighbor 10.1.1.1 remote-as 1

双方都没有写更新源。(neighbor x.x.x.x update-source lo0 代表本路由器的更新源为 lo0 口，BGP 的包以这个接口的地址为源地址发送出去。)

一边指环回口，一边指直连接口。可以建立邻居。这里有 2 个 TCP 的 session，其中只有 R1 去访问 R2 的环回口的 179 端口的 TCP session 可以建立。可以用 **show tcp brief** 查看。

```
Router#sh tcp brief
TCB        Local Address         Foreign Address       (state)
65693960   10.1.1.1.51124        192.168.2.1.179       ESTAB
```

这时在 R2 上写上确定更新源命令：**neighbor 10.1.1.1 update-source lo1**，这时即可建立 2 条 TCP session。可以使用命令 Show tcp brief 查看到 2 条 TCP session 在建立，当一条 establish 完成后，另一条过会即消失。

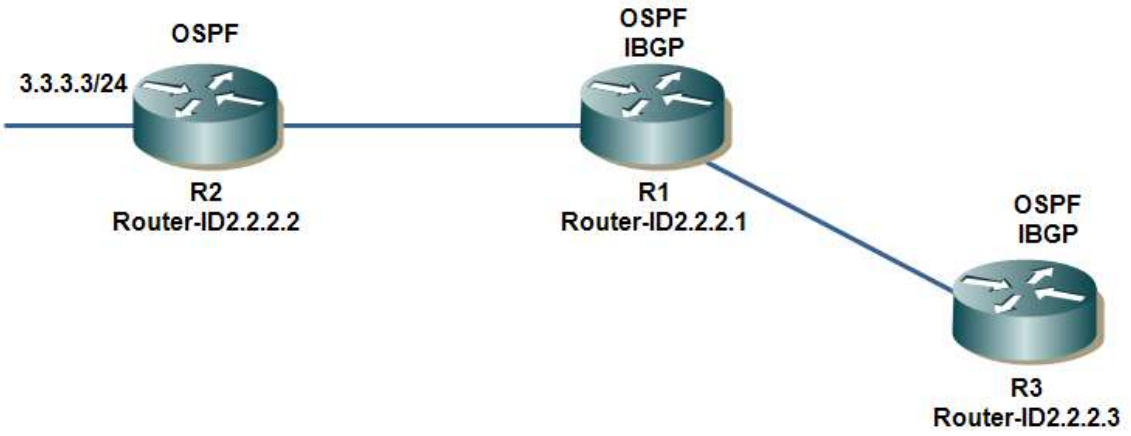
```
Router#sh tcp brief
TCB        Local Address         Foreign Address       (state)
65693960   10.1.1.1.51124        192.168.2.1.179       CLOSED
65693E14   10.1.1.1.37992        192.168.2.1.179       ESTAB
Router#sh tcp brief
TCB        Local Address         Foreign Address       (state)
65693E14   10.1.1.1.37992        192.168.2.1.179       ESTAB
```

注：路由器建立 BGP 邻居写两条正确的 neighbor 命令，是为了冗余。

8、IBGP 的同步

- BGP 同步规则指出，BGP 路由器不应用使用通过 IBGP 获悉的路由或将其通告给外部邻居，除非该路由是本地的或是通过 IGP 获悉的。
- 同步开启意味着，从一个 IBGP 邻居学来的路由，除非从 IGP 中也同样学习到，否则不可能被选为最优。
- 如果 IGP 为 OSPF，那么在 IGP 中，这些前缀的 router-id 也必须与通告这些前缀的 bgp 的 router-id 相匹配。才有可能被选为最优。

实例说明：如下图所示



R1、R2、R3 同为 OSPF area 0 中路由器（每台路由器的 router-id 如上图所示），R2 上一条路由 3.3.3.0/24 宣告进 OSPF。

R1、R3 运行 IBGP，R1 将 3.3.3.0/24 的前缀引入 BGP，传给 R3。这时 R3 既从 OSPF area0 中的 R2 学习到该前缀，又从 IBGP 对等体 R1，学习到该前缀，如果 R3 的 synchronizaion 是开启的，检查同步，在 R3 的 BGP 转发表里：

R1

```
router ospf 10
 router-id 2.2.2.1
 log-adjacency-changes
 network 10.1.1.0 0.0.0.255 area 0
 network 11.1.1.0 0.0.0.255 area 0
!
router bgp 100
 synchronization
 bgp log-neighbor-changes
 redistribute ospf 10
 neighbor 11.1.1.2 remote-as 100
 no auto-summary
```

R3

```
router ospf 10
 router-id 2.2.2.3
```

```
 log-adjacency-changes
 network 11.1.1.0 0.0.0.255 area 0
!
router bgp 100
 synchronization
 bgp log-neighbor-changes
 neighbor 11.1.1.1 remote-as 100
 no auto-summary
```

R3#sh ip bgp

```
BGP table version is 30, local router ID is 11.1.1.2
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
              r RIB-failure, S Stale
Origin codes: i - IGP, e - EGP, ? - incomplete

   Network          Next Hop            Metric LocPrf Weight Path
* i3.3.3.3/32      10.1.1.1                 11    100      0 ?
* i10.1.1.0/24     11.1.1.1                  0    100      0 ?
r>i11.1.1.0/24     11.1.1.1                  0    100      0 ?
```

R3#sh ip bgp 3.3.3.3

```
BGP routing table entry for 3.3.3.3/32, version 26
Paths: (1 available, no best path)
  Not advertised to any peer
  Local
    10.1.1.1 (metric 20) from 11.1.1.1 (11.1.1.1)
      Origin incomplete, metric 11, localpref 100, valid, internal, not synchronized
```

说明同步检查没有通过，当把 R1 的 bgp 的 router-id 改为 2.2.2.2 时，R3 这时检查同步就可以通过了。

R1

```
router ospf 10
 router-id 2.2.2.1
 log-adjacency-changes
 network 10.1.1.0 0.0.0.255 area 0
 network 11.1.1.0 0.0.0.255 area 0
!
router bgp 100
 synchronization
 bgp router-id 2.2.2.2
 bgp log-neighbor-changes
 redistribute ospf 10
 neighbor 11.1.1.2 remote-as 100
```

no auto-summary

R3#sh ip bgp

BGP table version is 37, local router ID is 11.1.1.2

Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,

r RIB-failure, S Stale

Origin codes: i - IGP, e - EGP, ? - incomplete

Network	Next Hop	Metric	LocPrf	Weight	Path
r>i3.3.3.3/32	**10.1.1.1**	**11**	**100**	**0**	**?**
r>i10.1.1.0/24	11.1.1.1	0	100	0	?
r>i11.1.1.0/24	11.1.1.1	0	100	0	?

R3#sh ip bgp 3.3.3.3

BGP routing table entry for 3.3.3.3/32, version 35

Paths: (1 available, best #1, table Default-IP-Routing-Table, RIB-failure(17))

Flag: 0x820

Not advertised to any peer

Local

10.1.1.1 (metric 20) from 11.1.1.1 (2.2.2.2)

Origin incomplete, metric 11, localpref 100, valid, internal, **synchronized, best**

- 关闭同步的条件
  - 将 EBGP 的路由重分布进 IGP
  - 本 AS 不为其他 AS 提供穿越服务（末节的 AS）
  - 穿越路径上所有路由器都运行 BGP

二、BGP 属性

路由器发送关于目标网络的 BGP 更新消息，更新的度量值被称为路径属性。属性可以是公认的或可选的、强制的或自由决定的、传递的或非传递的。属性也可以是部分的。并非组织的和有组合都是合法的，路径属性分为 4 类：

- ——公认强制的
- ——公认自由决定的
- ——可选传递的
- ——可选非传递的
- 只有可选传递属性可被标记为部分的

公认属性

- 是公认所有 BGP 实现都必须能够识别的属性。这早些属性被传递给 BGP 邻居。
- 公认强制属性必须出现在路由描述中，公认自由决定属性可以不出现在路由描述中

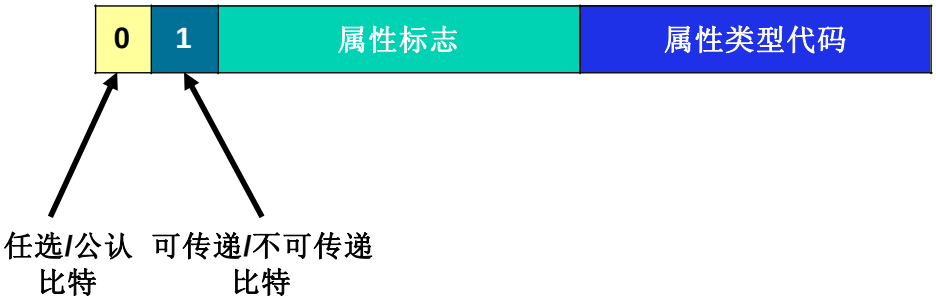
可选属性

- 非公认属性被称为可选的，可选属性可以是传递的或非传递的
- 可选属性不要求所有的 BGP 实现都支持
- 对于不支持的可选传递属性，路由器将其原封不动地传递给其他 BGP 路由器，在这种情况下，属性被标记为部分的。
- 对于可选非传递属性，路由器必须将其删除，而不将其传递给其他 BGP 路由器

BGP 定义属性

- 公认强制属性
- 公认自由决定
- 可选传递属性
- 可选非传递属性

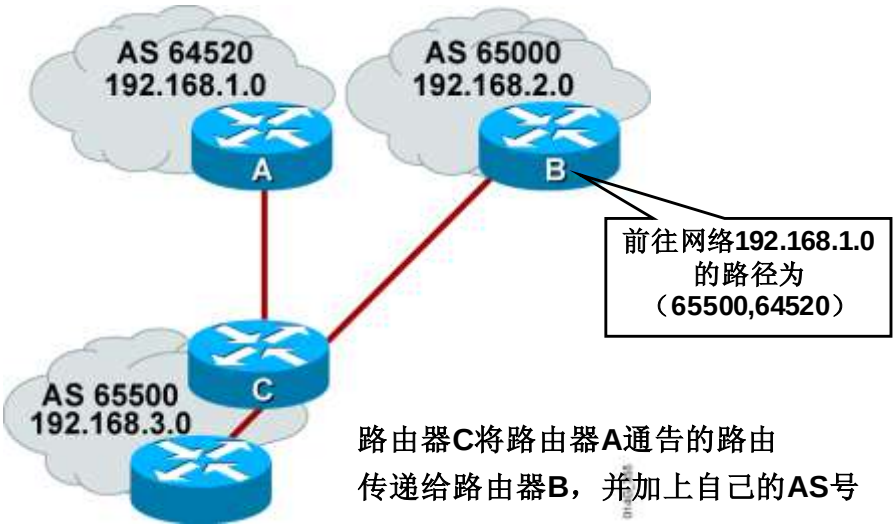
BGP 每条更新消息都有一个长度可变的路径属性序列<属性类型，属性长度，属性值>，如果第 1 比特是 0,则属于是公认属性，如果它是 1,则该属性是任选属性，如果第 2 比特是 0,则该属性是不可传递的，如果它是 1,则属性是可传递的，公认属性总是可传递的，属性标志域中的第 3 个比特指示任选可传递属性中的信息是部分的（值为 1）还是完整的（值为 0），第 4 个比特确定该属性长度是 1 字还是 2 字节，标志域其他 4 个比特总为 0.属性类型代码字节含有属性代码。如下图所示。



1、AS 路径属性（AS-path）

AS_PATH 是一个公认必选的属性，它用 AS 号的顺序来描述 AS 间的路径或到 NLRI 所明确的目的地的路由。

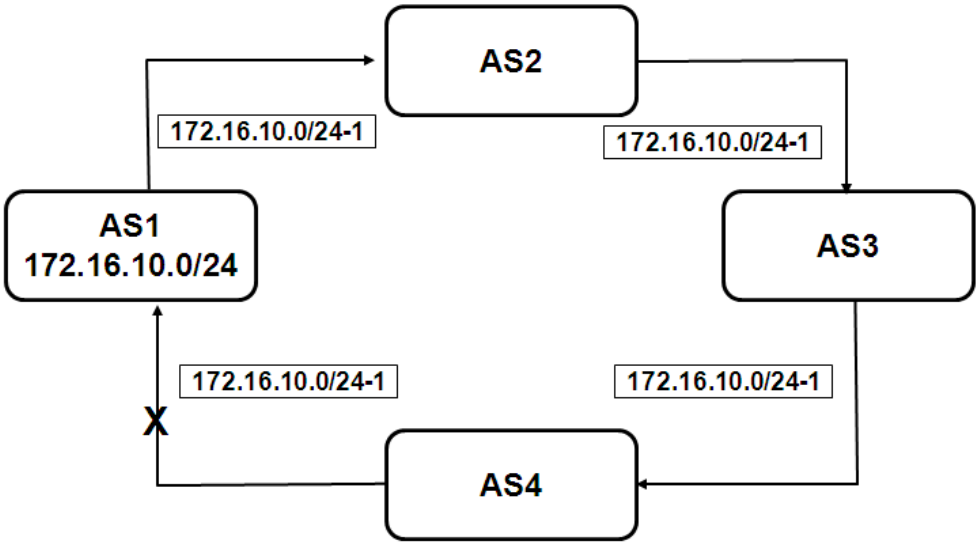
当每个运行运行 BGP 的路由器发起一条路由——当它在自己的 AS 域内公布一个有关目的地 NLRI——它将自己的 AS 号附加到 AS_PATH 中。当后续的运行 BGP 的路由器向外部的对端公布路由，它将自己的 AS 号附加到 AS_PATH 中。AS 可以描述所有它经过的自治系统，以最近的 AS 开始，以发起者的 AS 结束。如下图所示。



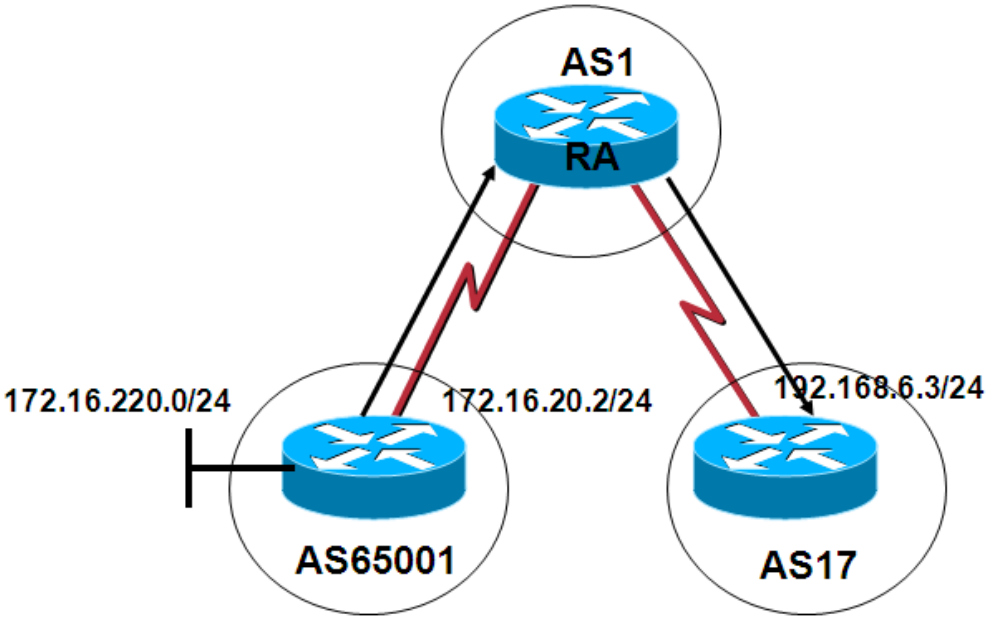
只有将更新消息发送给在另一个 AS 域内的邻居时，BGP 路由器才将它的 AS 号加到 AS_PATH 中，也就是说只有在两个 EBGP 对等体之间公布路由时，AS 号才被附加到 AS_PATH 中。

可以通过使用 AS 附加改变其公布路由的 AS_PATH 来影响数据流的流向。

AS_PATH 属性的另一个功能就是避免环路，如果 BGP 路由器从它的外部邻居收到一条路由，而该路由 AS_PATH 包含这个 BGP 路由器自己的 AS 号。于是该路由器就知道是条环路路由。如下图所示。



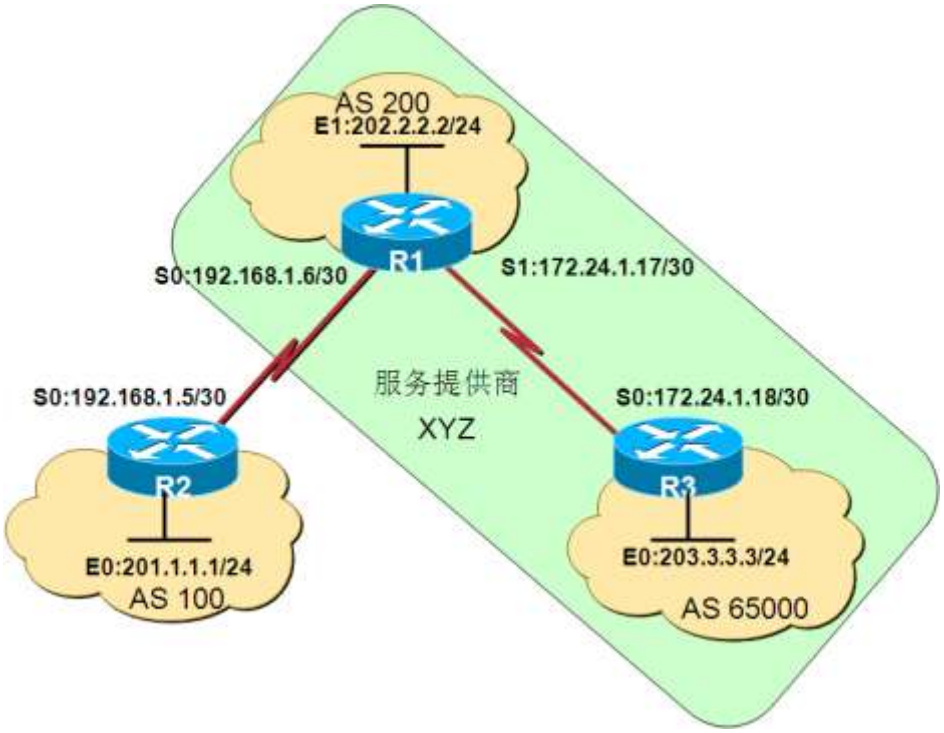
实例说明：如图所示。



可以将私有的 AS 号进行隐藏，如下配置所示。

```
RA:
RA(config)#router bgp 1
RA(config-router)#neighbor 172.16.20.2 remote-as 65001
RA(config-router)#neighbor 192.168.6.3 remote-as 17
RA(config-router)#neighbor 192.168.6.3 remove-private-as
```

下面是 AS 属性的另一个实例，如下图所示。



R1 在发送更新的时候，剥除私有 AS 号；并且不将 AS100 的路由传播给其客户路由器 R3，配置如下所示。

```
R2:
R2(config)#router bgp 100
R2(config-router)#no synchronization
R2(config-router)#neighbor 192.168.1.6 remote-as 200
R2(config-router)#network 201.1.1.0
R1:
R1(config)#router bgp 200
R1(config-router)#no synchronization
R1(config-router)#neighbor 192.168.1.5 remote-as 100
R1(config-router)#neighbor 172.24.1.18 remote-as 65000
R1(config-router)#network 202.2.2.0
R1(config-router)#neighbor 192.168.1.5 remove-private-as
R1(config)#ip as-path access-list 1 deny ^100$
R1(config)#ip as-path access-list 1 permit .*
R1(config-router)#neighbor 172.24.1.18 filter-list 1 out
R3:
R3(config)#router bgp 65000
R3(config-router)#no synchronization
R3(config-router)#neighbor 172.24.1.17 remote-as 200
R3(config-router)#network 203.3.3.0
```

聚合后继承明细路由的属性，在大括号里面的 as-path 在计算长度时，只算一个。在联盟内小括号里面的 AS 号，在选路时，不计算到 as－path 长度里面。

增加 as－path 的长度，可以用 route-map 里面的 set as-path prepend 来做，如：

```
neighbor 1.1.1.1 route-map AS {in|out}
route-map AS
set as-path prepend 10 10
```

在 neighbor 的入向做 as-path prepend。是在 as-path 靠近我的地方加长度，如：**10 10 2i** 。10 10 是新加的。

而在 neighbor 的出向做 as-path prepend。是在 AS 起源的方向加 path 长度，如：**2 10 10**i 。10 10 是新加的。

在 as-path prepend 的后面还有一个参数，last－as，如：

```
route-map AS
set as-path prepend last-as ?
  <1-10>  number of last-AS prepends
```

意思是将离我最近的 AS，将它的 AS 号在 as－path 里面再重复出现几次。这个 10 看起来可以和 allowas-in 里面的 10 对应起来。

假如 as-path prepend 与 as-path prepend last-as 合用的时候，last-as 先生效，然后 prepend 再生效。

减小 as-path 的长度，如用联盟和 remove-private-AS 等可以实现。

注意：Remove－private－as，如果在 as-path 里交替出现私有和公有的 AS 号，这样将无法将私有 AS 号去掉。在起源的时候，连续的时候才有效。

bgp bestpath as-path ignore(隐藏命令), 这条命令可以使我们在选路时，跳过 as－path 的选路，直接往下继续选择最优路径。

2、源头属性（Origin）

源头是公认强制属性，它定义了路径信息的源头。

IGP：路由在起始 AS 的内部，使用 network 命令通过 BGP 通告路由时，通常属于这种情况，在 BGP 表中，IGP 源头用 i 表示

EGP：路由是通过 EGP 获悉的，在 BGP 表中用 e 表示。

不完全：路由的源头未知或是通过其他方法获悉的，在 BGP 表中，不完整源头用？表示

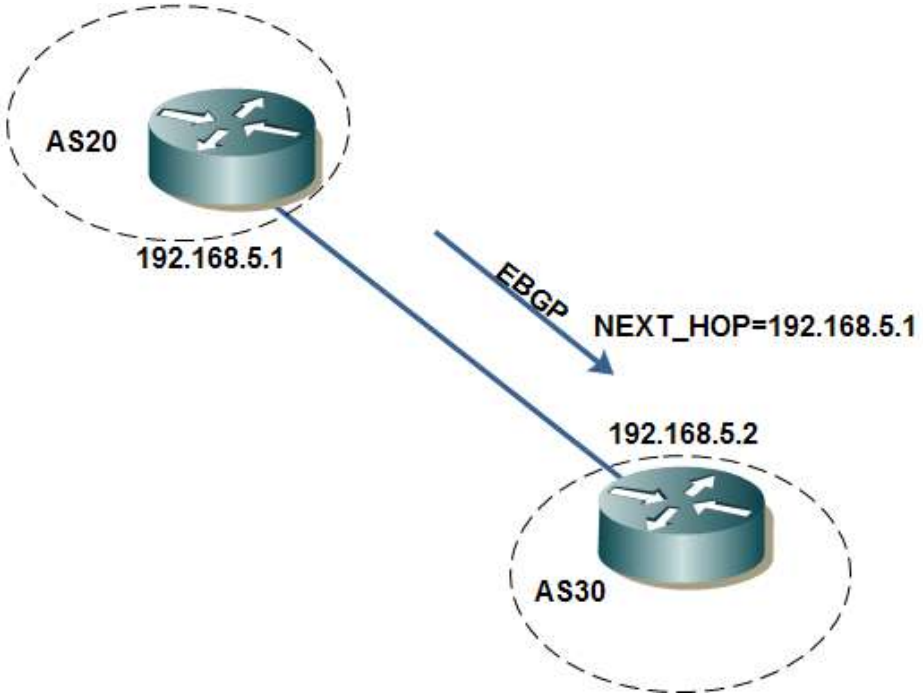
如下示例所示。

```
RouterA# show ip bgp

BGP table version is 23, local router ID is 192.168.1.49
Status codes: s suppressed, d damped, h history, * valid, > best, i -
internal
Origin codes: i - IGP, e - EGP, ? - incomplete

   Network          Next Hop            Metric LocPrf Weight Path
*> 10.0.0.0         10.1.1.100               0             0 65200 i
*> 172.16.10.0/24   10.1.1.100               0             0 65200 i
*> 172.16.11.0/24   10.1.1.100               0             0 65200 i
*>i172.26.1.16/28   192.168.1.50             0    100      0 i
*>i172.26.1.32/28   192.168.1.50             0    100      0 i
*>i172.26.1.48/28   192.168.1.50             0    100      0 i
*> 192.168.1.0      0.0.0.0                  0         32768 i
*> 192.168.2.0      10.1.1.100                             0 65200 65102 i
*> 192.168.2.64/28  10.1.1.100                             0 65200 65102 i
* i192.168.101.0    192.168.1.34             0    100      0 i
*>i                 192.168.1.18             0    100      0 i
```

2、下一跳属性（NEXT_HOP）

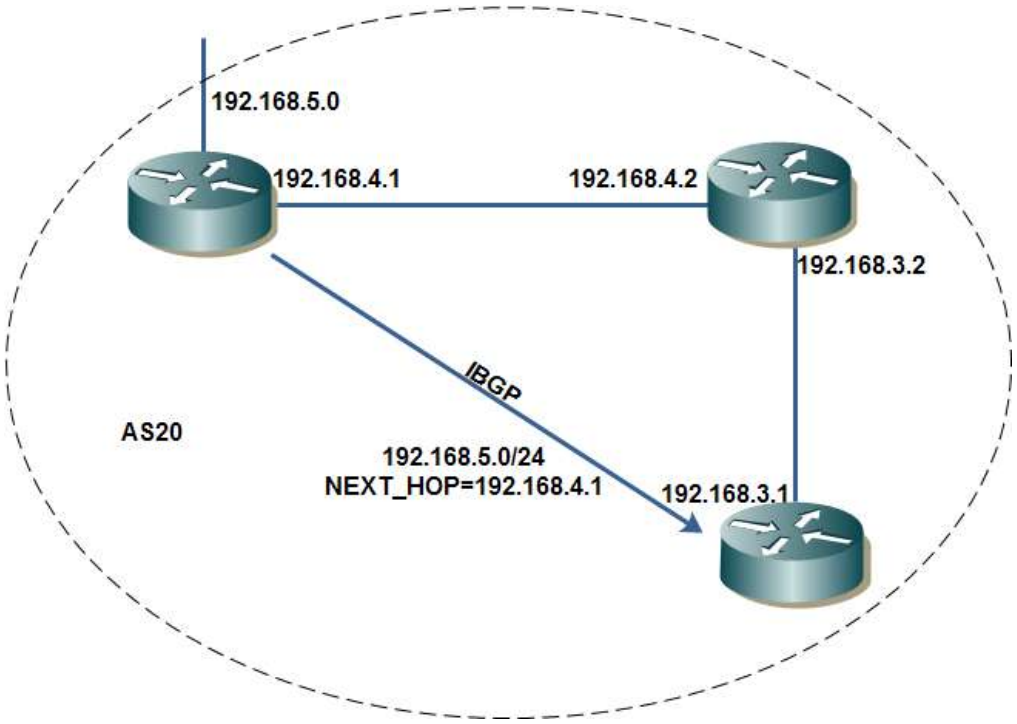
该为公认必选属性，描述了到公布目的地的路径下一跳路由器的 IP 地址。由 BGP NEXT_HOP 属性所描述的 IP 地址不经常是邻居路由器的 IP 地址，要遵循下面的规则：

如果正在进行路由宣告的路由器和接收的路由器在不同的自治系统中，NEXT_HOP 是正在宣告路由器接口的 IP 地址，如下图所示。

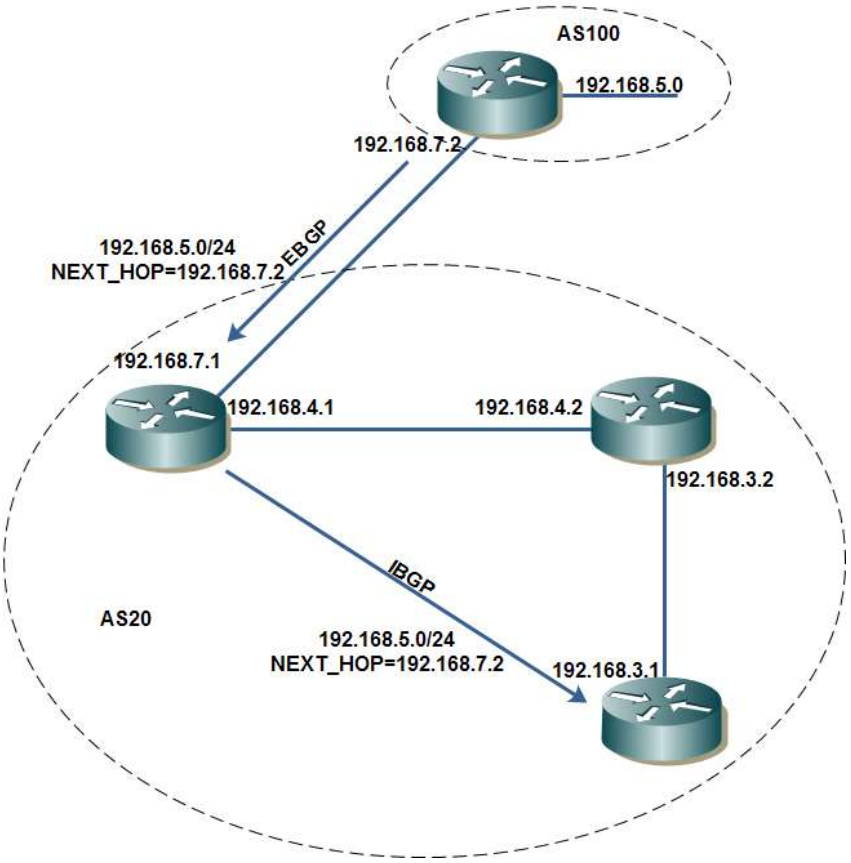


如果正在进行路由宣告的路由器和接收的路由器在同一个 AS 内，并且更新消息的 NLRI

指明的目的地也在同一个 AS 内，那么 NEST_HOP 就是宣告路由的邻居的 IP 地址。如下图所示。



如果正在宣告的路由器和接收的路由器是内部对等体，并且更新消息的 NLRI 指明目的地在不同的 AS，则 NEXT_HOP 就是学习到路由的外部对等实体的 IP 地址。如下图所示。



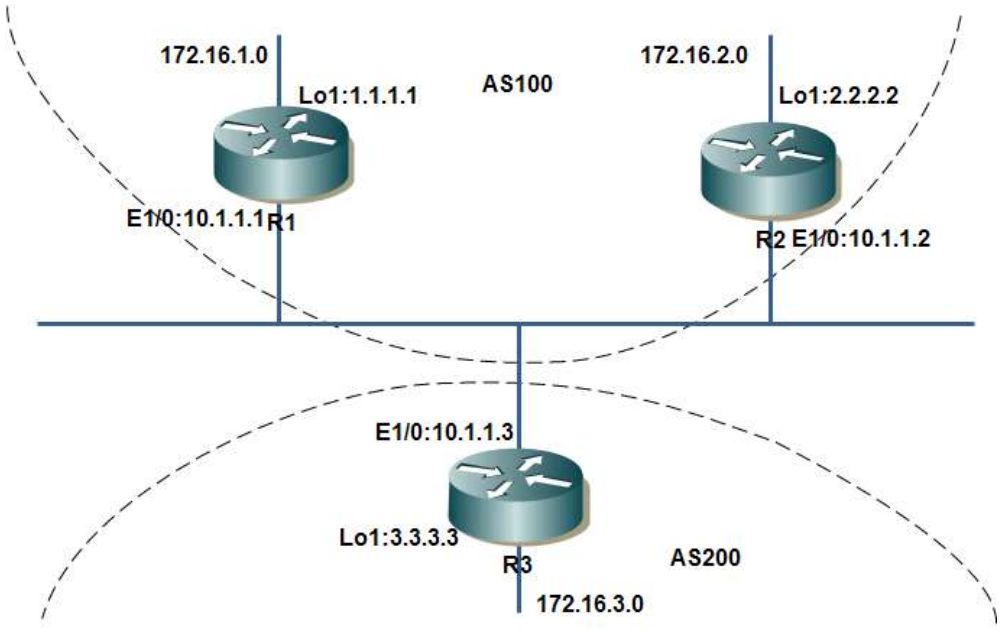
从上面图可以知道，在去往 192.168.5.0 的网段中会出现路径不可达的情况，解决这个问题的方法是保证内部路由器知道与两处自治系统相连的外部网络，可以使用静态路由的办法，但实际的做法是在外部端口上以被动模式运行 IGP。但在某种情况下，该方法并不理想。

第二种方法是采用配置选项来做，这个配置选项被称做 next-hop-self。

下面具体详述了下一跳的不可达的解决方法：

- 解决下一跳不可达的方法：
  - 静态路由
  - 在 IBGP 邻居所处的 IGP 中宣告
  - 将与 EBGP 直连的网络重分布进 IGP
  - neighbor x.x.x.x next-hop-self（将指向 EBGP 邻居更新源的地址变为自己的更新源地址）（RR 有的版本会将下一跳改变）
- 一般情况下，在本路由器上将直连的网络引入 BGP，下一跳为 0.0.0.0，本路由器聚合的路由的下一跳也为 0.0.0.0。
- 在本路由器上将从 IGP 学来的路由引入 BGP 时，在本路由器上看 BGP 的转发表，下一跳为 IGP 路由的下一跳。在多访问网络环境中，用直连接口建立邻居关系，会产生第三方下一跳。

实例说明：如下图所示，



R2 与 R1 是 IBGP 邻居，R1 与 R3 是 EBGP 邻居，当用直连接口建邻居时，R2 引入 BGP 的前缀 172.16.2.0/24，在 R3 的 bgp 转发表里，将显示为 R2 的多访问网络接口地址（如：10.1.1.2）。产生第三方下一跳的现象。

- 如果 R1、R2、R3 全部用直连接口建邻居时会产生第三方下一跳。
- 如果 R1、 R2 用环回口而 R1、R3 用直连建立邻居时，会产生第三方下一跳。
- 如果 R1、R2 用直连而 R1、R3 用回环口时，不会产生第三方下一跳，如下所示配置。
- 如果 R1、R2、R3 都用环回口建立邻居，则不会产生第三方下一跳，如下配置所示。

R3#show ip bgp

BGP table version is 4, local router ID is 172.16.3.1

Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,

r RIB-failure, S Stale

Origin codes: i - IGP, e - EGP, ? - incomplete

Network	Next Hop	Metric	LocPrf	Weight	Path
*> 172.16.1.0/24	10.1.1.1	0		0	100 i
***> 172.16.2.0/24**	**10.1.1.2**			**0**	**100 i**
*> 172.16.3.0/24	0.0.0.0	0		32768	i

R1

```
router bgp 100
 no synchronization
 bgp log-neighbor-changes
 network 172.16.1.0 mask 255.255.255.0
 neighbor 3.3.3.3 remote-as 200
 neighbor 3.3.3.3 ebgp-multihop 2
 neighbor 3.3.3.3 update-source Loopback1
 neighbor 10.1.1.2 remote-as 100
 no auto-summary
```

R3

```
router bgp 200
 no synchronization
 bgp log-neighbor-changes
 network 172.16.3.0 mask 255.255.255.0
 neighbor 1.1.1.1 remote-as 100
 neighbor 1.1.1.1 ebgp-multihop 2
 neighbor 1.1.1.1 update-source Loopback1
 no auto-summary
```

R3#sh ip bgp

BGP table version is 8, local router ID is 172.16.3.1

Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,

r RIB-failure, S Stale

Origin codes: i - IGP, e - EGP, ? - incomplete

Network	Next Hop	Metric	LocPrf	Weight	Path
*> 172.16.1.0/24	1.1.1.1	0		0	100 i
***> 172.16.2.0/24**	**1.1.1.1**			**0**	**100 i**
*> 172.16.3.0/24	0.0.0.0	0		32768	i

R1

```
router bgp 100
 no synchronization
 bgp log-neighbor-changes
 network 172.16.1.0 mask 255.255.255.0
 neighbor 2.2.2.2 remote-as 100
 neighbor 2.2.2.2 update-source Loopback1
 neighbor 3.3.3.3 remote-as 200
 neighbor 3.3.3.3 ebgp-multihop 2
 neighbor 3.3.3.3 update-source Loopback1
 no auto-summary
```

R2

```
router bgp 100
 no synchronization
 bgp log-neighbor-changes
 network 172.16.2.0 mask 255.255.255.0
 neighbor 1.1.1.1 remote-as 100
 neighbor 1.1.1.1 update-source Loopback1
 no auto-summary
```

R3

```
router bgp 200
 no synchronization
 bgp log-neighbor-changes
 network 172.16.3.0 mask 255.255.255.0
 neighbor 1.1.1.1 remote-as 100
 neighbor 1.1.1.1 ebgp-multihop 2
 neighbor 1.1.1.1 update-source Loopback1
 no auto-summary
```

R3#sh ip bgp

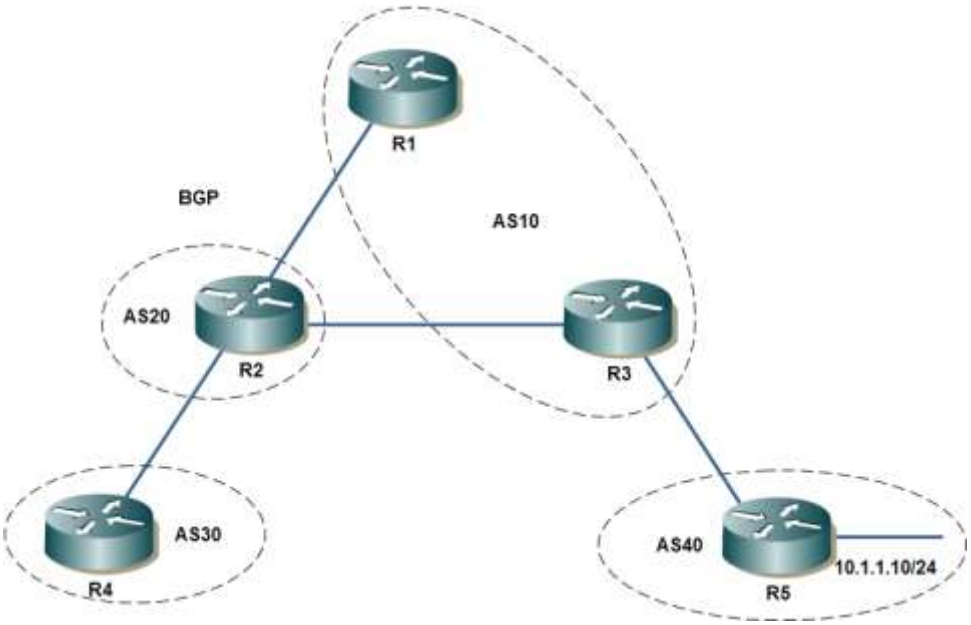
```
BGP table version is 10, local router ID is 172.16.3.1
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
              r RIB-failure, S Stale
Origin codes: i - IGP, e - EGP, ? - incomplete
```

Network	Next Hop	Metric	LocPrf	Weight	Path
*> 172.16.1.0/24	1.1.1.1	0		0	100 i
***> 172.16.2.0/24**	**1.1.1.1**			**0**	**100 i**
*> 172.16.3.0/24	0.0.0.0	0		32768	i

第三方下一跳：收到路由更新的源地址与将要发出去的接口地址在同一网段的时候，路由的下一跳不改变，为原来路由更新的源地址。

- 有时虽然路由的下一跳可达，但会出现访问网络出现环路的现象。

实例说明：



R5、R3，R1、R2 为 EBGP 邻居关系，R1、R3 为 IBGP 邻居关系。那么 R5 通过 BGP 传给 R3 的路由（如 10.1.1.0/24），R3 通过 IBGP 传给 R1，R1 通过 EBGP 传给 R2，这时 R2 访问 10.1.1.0/24 这个网络的下一跳就在 R1 上。这时 R2 去访问 R5 的时候，就会产生环路。

则 R2（走下一跳）——R1（走物理链路）——R2，这样环路产生了。

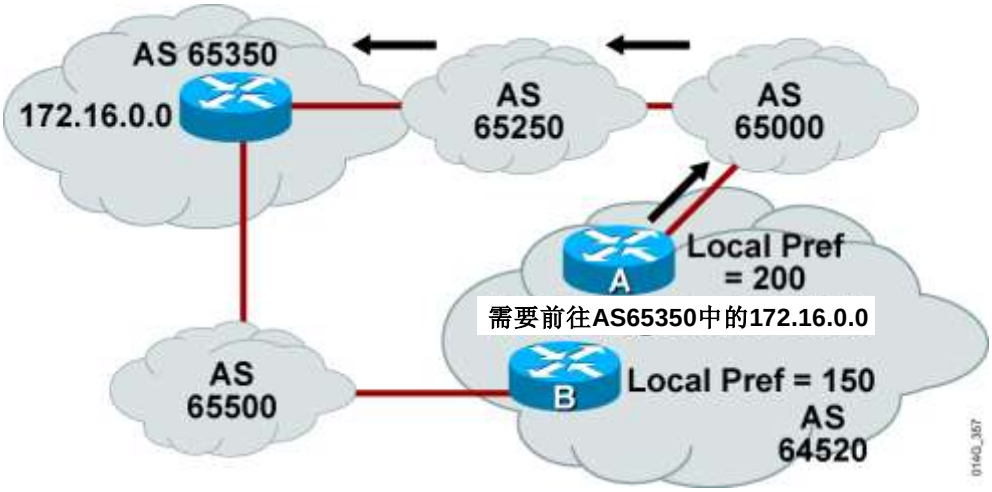
解决方法：

- **neighbor x.x.x.x next-hop-unchanged** (此命令只能用在 EBGP 多跳的环境下，将路由的下一跳，从自己的更新源地址改变为从 IBGP 学来的下一跳地址)（这时路由的下一跳在路由表里将改变。）
- **neighbor x.x.x.x route-map XX {in|out}**然后在 route-map 里面 set ip next-hop 来改变前缀的下一跳。(在路由表里下一跳会改变。)
- 策略路由 PBR，强制命令 R2 到 10.1.1.0/24 的时候走 R3。（路由表里下一跳不会改变）

3、本地优先级属性（Local_preference）

本地优先级是公认自由决定的属性，它告诉 AS 中的路由器，那条路径离开 AS 的首选路径。本地优先级越高，路径被选中的可能性越大。本地优先级这种属性只能在同一个 AS 中的路由器之间交换，本是优先级只适用于内部邻居，用于内部对等体之间的 Update 消息。

本地优先级，可以在本 AS 和大联盟内传递。越大越优先。影响路由器的出站流量。默认情况下，local-preference 为 100。

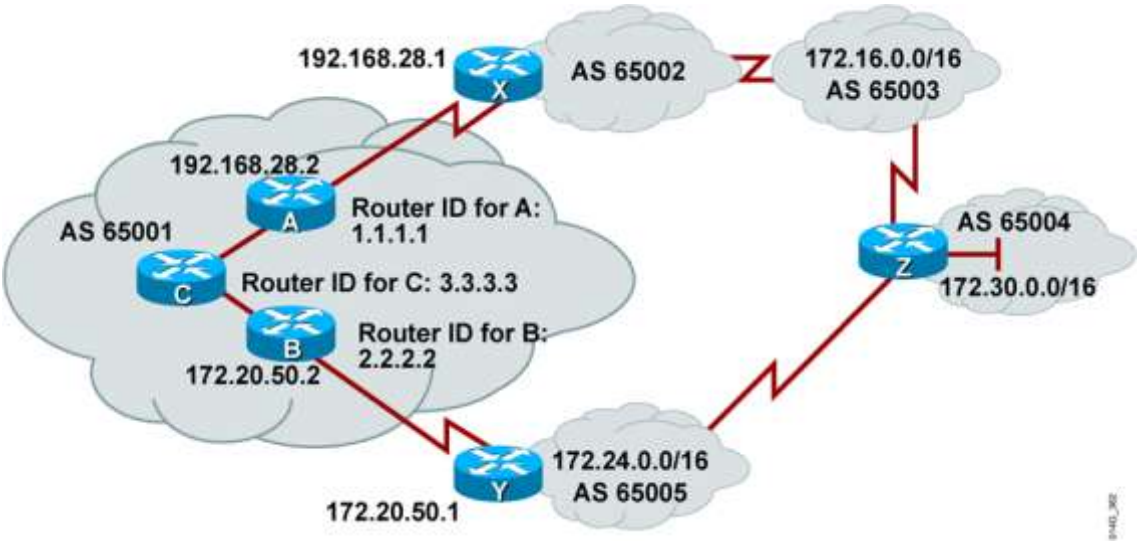


使用下面的命令，如下图所示

```
Router(config-router)#
bgp default local-preference value
```

是将路由器收到的所有外部 BGP 路由的默认本地优先级修改为指定值。对 IBGP 邻居路由器传过来的路由，不会改变它们的 local-preference。如果将一个 IBGP 邻居传来的路由传给另外一个 IBGP 邻居，那我必须是 RR。

实例说明：如下图所示。



未使用本地优先级操作路径，如下所示路由器 C 的 BGP 表。

```
RouterC# show ip bgp

BGP table version is 7, local router ID is 3.3.3.3
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal
Origin codes: i - IGP, e - EGP, ? - incomplete
   Network          Next Hop       Metric LocPrf Weight Path
* i172.16.0.0       172.20.50.1             100       0 65005 65004 65003 i
*>i                 192.168.28.1            100       0 65002 65003 i
*>i172.24.0.0       172.20.50.1             100       0 65005 i
* i                 192.168.28.1            100       0 65002 65003 65004 65005 i
*>i172.30.0.0       172.20.50.1             100       0 65005 65004 i
* i                 192.168.28.1            100       0 65002 65003 65004i
```

在路由器 A 上修改本地优先级，如下所示。

```
router bgp 65001
neighbor 2.2.2.2 remote-as 65001
neighbor 3.3.3.3 remote-as 65001
neighbor 2.2.2.2 remote-as 65001 update-source loopback0
neighbor 3.3.3.3 remote-as 65001 update-source loopback0
neighbor 192.168.28.1 remote-as 65002
neighbor 192.168.28.1 route-map local_pref in
!
route-map local_pref permit 10
match ip address 65
set local-preference 400
!
route-map local_pref permit 20
!
access-list 65 permit 172.30.0.0 0.0.255.255
```

在使用本地优先操纵后的路径，查看路由器 C 的 BGP 表。

```
RouterC# show ip bgp

BGP table version is 7, local router ID is 3.3.3.3
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal
Origin codes: i - IGP, e - EGP, ? - incomplete
   Network          Next Hop       Metric LocPrf Weight Path
* i172.16.0.0       172.20.50.1             100       0 65005 65004 65003 i
*>i                 192.168.28.1            100       0 65002 65003 i
*>i172.24.0.0       172.20.50.1             100       0 65005 i
* i                 192.168.28.1            100       0 65002 65003 65004 65005 i
* i172.30.0.0       172.20.50.1             100       0 65005 65004 i
*>i                 192.168.28.1            400       0 65002 65003 65004i
```

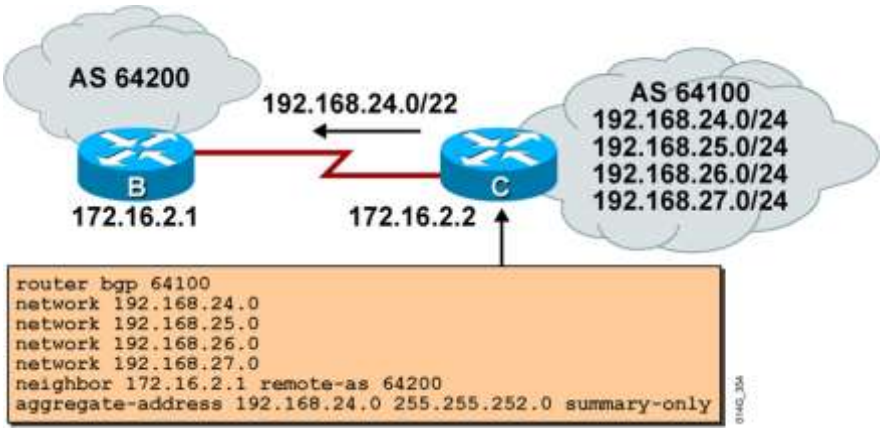
4、原子聚合属性

原子聚合是一个公认自决的属性。类型代码为 6，它告诉邻接 AS，始发路由器对路由进行了聚合。可以使用下面的命令进行配置，

```
Router(config-router)#
aggregate-address ip-address mask [summary-only]
[as-set]
```

命令只聚合已经包含在 BGP 表中的网络，这与使用 network 来通告汇总路由要求不同，后者要求网络必须出 现在 IP 路由选择表中

配置命令 aggregate-address 后，一条与汇总路由对应的指向 null0 的 BGP 路由将自动被加入到 IP 路由表中。如下示例所示。



可以使用 show ip bgp 命令来查看

```
routerC# show ip bgp

BGP table version is 28, local router ID is 172.16.2.1
Status codes: s = suppressed, * = valid, > = best, and i = internal
Origin codes : i = IGP, e = EGP, and ? = incomplete

   Network          Next Hop        Metric  LocPrf    Weight     Path
*> 192.168.24.0/22  0.0.0.0              0              32768     i
s> 192.168.24.0     0.0.0.0              0              32768     i
s> 192.168.25.0     0.0.0.0              0              32768     i
s> 192.168.26.0     0.0.0.0              0              32768     i
s> 192.168.27.0     0.0.0.0              0              32768     i
```

关于原子聚合的详细内容在以后的章节中详细说明。

5、权重属性

cisco 私有的参数。本地有效。缺省条件下，本地始发的路径具有相同的 WEIGHT 值（即 32768），所有其他的路径的 weight 值为 0。越大越优选。影响路由器的出站流量。

权重只影响当前路由器，指定邻居的权重。使用下面命令来修改权重。

```
neighbor {ip-address | peer-group-name} wiight weight
```

可以在 neighbor 的入向设置。范围 0－65535。Neighbor 1.1.1.1 weight 10，从对等体 1.1.1.1 接收过来的所有路由的 weight 值都设置为 10。

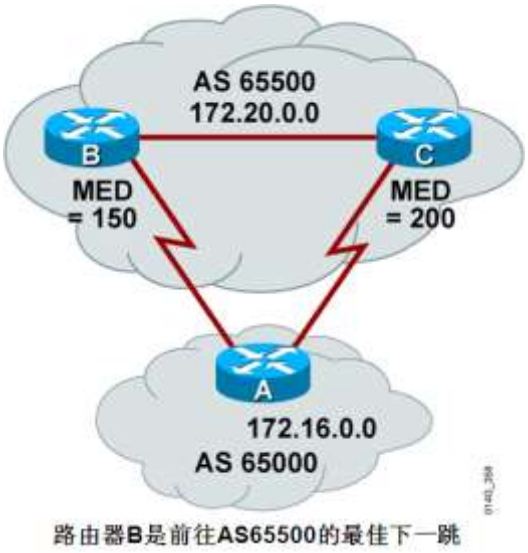
还可以用 route-map 来设定，可以将特定路由的 weight 值改变。如下所示：

```
Neighbor 1.1.1.1 route-map AA in

Route-map AA permit 10
```

```
Match ip address prefix AA
Set weight 10

Route-map AA permit 20
```

6、MED 属性

MED 属性也被称为度量值，是一种可选非传递属性。承载于 EBGP 的 Update 消息中。MED 用于向外部邻居指出进入 AS 的首选路径，当入口有多个时，AS 可以使用 MED 来动态地影响其他 AS 如何选择进入路径，在 BGP 中，MED 是唯一一个可影响数据如何进入 AS 的属性。度量值越小，路径被选中的可能性越大。与本地优先级不同，MED 是在自主系统之间交换的。MED 影响进入 AS 的数据流，而本地优先级影响离开 AS 的数据流。如下图所示。



Metric 和 med：BGP 的 metric 对 IBGP 同样有效。特指 med：从 EBGP 收到的 metric 比较的时候才叫 MED，MED 是借用了 BGP 的 metric 在 EBGP 的时候进行比较。MED（多出口区分）比较 EBGP 的 metric 找到最优的出口。

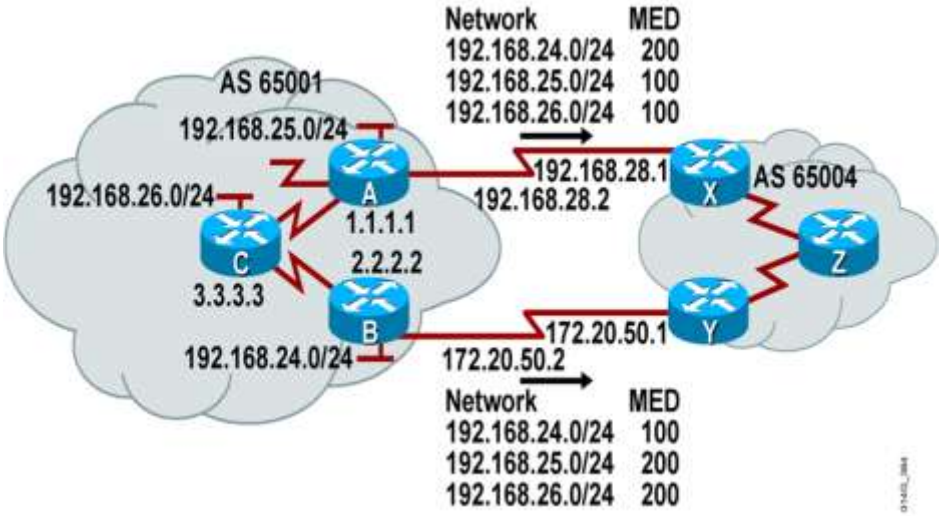
MED 相当于 IGP 路由的 metric 值，越小越优先。在新的 IOS 中，将 IGP 中的路由重分布进 BGP，BGP 将自动继承 IGP 路由的 metric 值。在老的 IOS 里，如果需要继承需要在重分布时加 route-map，如：

```
Redistribute rip route-map RE
Route-map RE
set metric-type internal
```

默认情况下，只有在两条路径的第一个（邻近的）AS 相同的情况下才会进行比较：任何联盟内的子自治系统都被忽略。任何多跳路径，只有在 AS_SEQUENCE 中的第一个 AS 相同的情况下，才会比较 MED；任何打头的 AS_CONFED_SEQUENCE 都将被忽略。如果激活了 **bgp always-compare-med**，那么对于所有路径都比较 MED，而不考虑是否来自同一个 AS。如果使用了这个选项，就应该在整个 AS 中都这样做，以避免路由选择环路。

实例说明：如下拓扑图

下面是一个使用策略路由来实现修改 MED 值的案例。

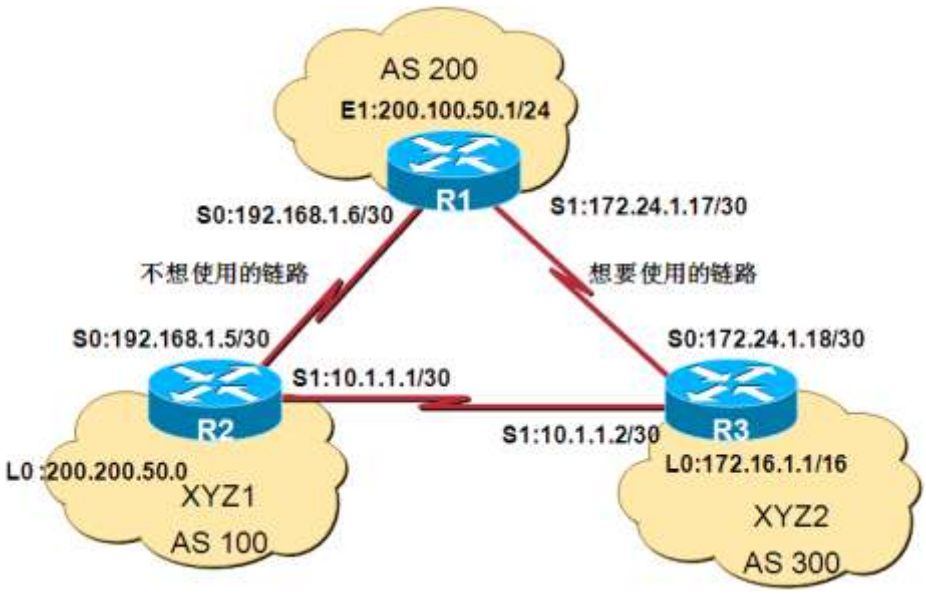



```
Router A's Configuration:
router bgp 65001
neighbor 2.2.2.2 remote-as 65001
neighbor 3.3.3.3 remote-as 65001
neighbor 2.2.2.2 update-source loopback0
neighbor 3.3.3.3 update-source loopback0
neighbor 192.168.28.1 remote-as 65004
neighbor 192.168.28.1 route-map med_65004 out
!
route-map med_65004 permit 10
match ip address 66
set metric 100
route-map med_65004 permit 100
set metric 200
!
access-list 66 permit 192.168.25.0 0.0.0.255
access-list 66 permit 192.168.26.0 0.0.0.255
```

```
Router B's Configuration:
router bgp 65001
neighbor 1.1.1.1 remote-as 65001
neighbor 3.3.3.3 remote-as 65001
neighbor 1.1.1.1 update-source loopback0
neighbor 3.3.3.3 update-source loopback0
neighbor 172.20.50.1 remote-as 65004
neighbor 172.20.50.1 route-map med_65004 out
!
route-map med_65004 permit 10
match ip address 66
set metric 100
route-map med_65004 permit 100
set metric 200
!
access-list 66 permit 192.168.24.0 0.0.0.255
```

```
RouterZ# show ip bgp

BGP table version is 7, local router ID is 122.30.1.1
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal
Origin codes: i - IGP, e - EGP, ? - incomplete
   Network          Next Hop        Metric LocPrf Weight Path
*>i192.168.24.0     172.20.50.2        100    100        0 65001 i
* i                 192.168.28.2       200    100        0 65001 i
* i192.168.25.0     172.20.50.2        200    100        0 65001 i
*>i                 192.168.28.2       100    100        0 65001 i
* i192.168.26.0     172.20.50.2        200    100        0 65001 i
*>i                 192.168.28.2       100    100        0 65001 i
```

实例说明：如下图所示，此实例采用了本地优先级与 MED 属性



```
R2:
R2(config)#router bgp 100
R2(config-router)#neighbor 10.1.1.2 remote-as 300
R2(config-router)#neighbor 192.168.1.6 remote-as 200
R2(config-router)#network 200.200.50.0
R1:
R1(config)#router bgp 200
R1(config-router)#neighbor 192.168.1.5 remote-as 100
R1(config-router)#neighbor 172.24.1.18 remote-as 300
R1(config-router)#network 200.100.50.0
R3:
R3(config)#router bgp 300
R3(config-router)#neighbor 10.1.1.1 remote-as 100
R3(config-router)#neighbor 172.24.1.17 remote-as 200
R3(config-router)#network 172.16.0.0
R2:
R2(config)#route-map viaas300
R2(config-route-map)#set local-preference 150
R2(config)#router bgp 100
R2(config-router)#neighbor 10.1.1.2 route-map viaas300 in
```

```
R2:
R2(config)#router bgp 100
R2(config-router)#neighbor 10.1.1.2 remote-as 300
R2(config-router)#neighbor 192.168.1.6 remote-as 200
R2(config-router)#network 200.200.50.0
R1:
R1(config)#router bgp 200
R1(config-router)#neighbor 192.168.1.5 remote-as 100
R1(config-router)#neighbor 172.24.1.18 remote-as 300
R1(config-router)#network 200.100.50.0
R3:
R3(config)#router bgp 300
R3(config-router)#neighbor 10.1.1.1 remote-as 100
R3(config-router)#neighbor 172.24.1.17 remote-as 200
R3(config-router)#network 172.16.0.0
R2:
R2(config)#route-map viaas300
R2(config-route-map)#set local-preference 150
R2(config)#router bgp 100
R2(config-router)#neighbor 10.1.1.2 route-map viaas300 in
```

7、共同体属性

BGP 团体是一组共享某些共同特性的目的地，用于简化路由策略的执行，一个团体并不被限制在一个网络或一个 AS 之中。是另一种过滤入站或出站 BGP 路由的方法。

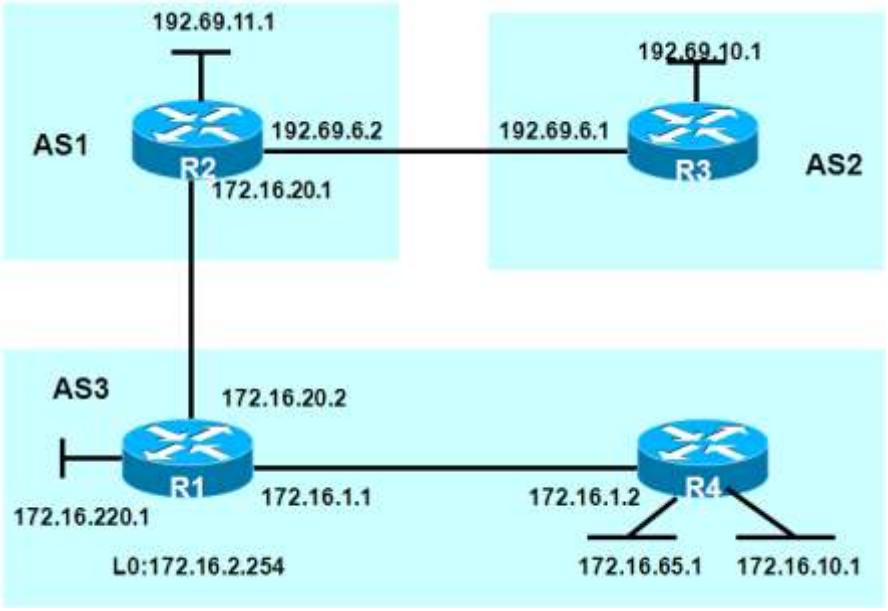
COMMUNITY 属性是一组 4 个 8 位组的数值，RFC1997 规定，前 2 个 8 位组表示自治系统，后 2 个 8 位组表示出于管理目的而定义的标识符，格式为 AA:NN，而思科的默认格式为 NN:AA，可以使用命令 ip bgpcommunity new-format 将思科默认格式改为 RFC1997 的标准格式。

团体属性是一个可传递属性，类型代码为 8。

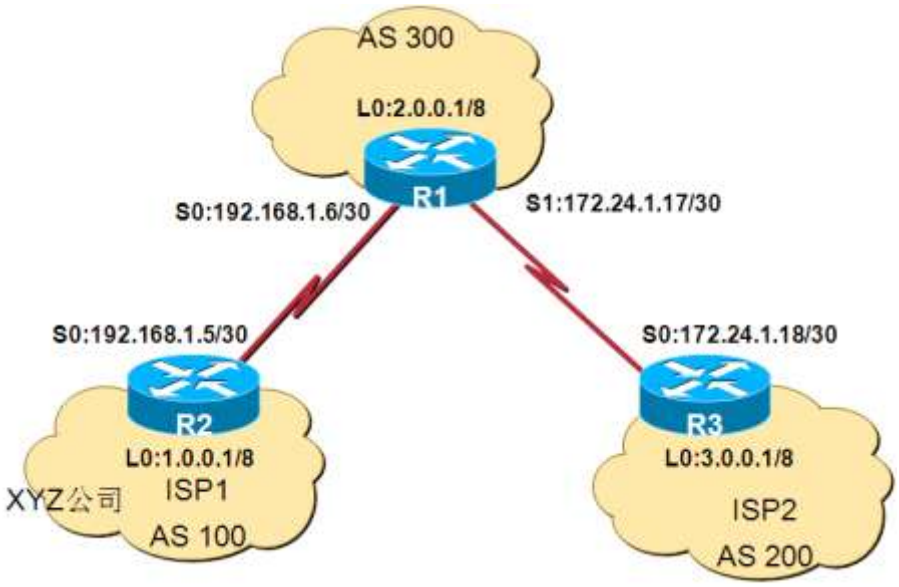
- no_export——如果接收到的路由携带该数值，不通告到 EBGP 对等体。如果配置了联盟，则不能将此路由宣告到联盟之外。
- no_advertise——如果接收到的路由携带该数值，不通告给任何对等体，包括 EBGP 和 IBGP。

- internet——无任何值，所有路由器默认情况下都属于该团体，带此属性的路由在被收到后，应该被通告给所有的其他路由器
- local_as——带有此属性的路由在被收到后，应该被通告给本地 AS 域内的对等体，但不应该被通告给外部系统中的对等体，包括同一个联盟内其它自治系统中的对等体。

实例说明：如下图所示。




```
R1(config)#router bgp 3
R1(config-router)#network 172.16.1.0 mask 255.255.255.0
R1(config-router)#network 172.16.10.0 mask 255.255.255.0
R1(config-router)#network 172.16.65.0 mask 255.255.255.192
R1(config-router)#network 172.16.220.0 mask 255.255.255.0
R1(config-router)#neighbor 172.16.1.2 remote-as 3
R1(config-router)#neighbor 172.16.1.2 update-source l0
R1(config-router)#neighbor 172.16.20.1 remote-as 1
R1(config-router)#neighbor 172.16.20.1 send-community
R1(config-router)#neighbor 172.16.20.1 route-map mymap out
R1(config-router)#exit
R1(config)#route-map mymap permit 10
R1(config-route-map)#match ip address 1
R1(config-route-map)#set community no-export
R1(config-route-map)#exit
R1(config-route-map)#route-map mymap permit 20
R1(config-router)#exit
R1(config)#access-list 1 permit 172.16.65.0 0.0.0.255
```




```
R2(config)#route bgp 100
R2(config-router)#neighbor 192.168.1.6 remote-as 200
R2(config-router)#network 1.0.0.0
R2(config-router)#no auto-summary
R2(config-router)#no synchronization
R2(config-router)#neighbor 192.168.1.6 route-map mymap out
R2(config-router)#neighbor 192.168.1.6 send-community
R2(config)#router-map mymap permit 10
R2(config-route-map)#match ip add 1
R2(config-route-map)#set community no-export
R2(config-route-map)#route-map mymap permit 20
R2(config-route-map)#exit
R2(config)#access-list 1 permit 1.0.0.0 0.255.255.255
```

三、BGP 路由汇总

BGP 的汇总有 2 种：

A．汇总：summary

静态路由手工汇总指向null 0，再network引入BGP。

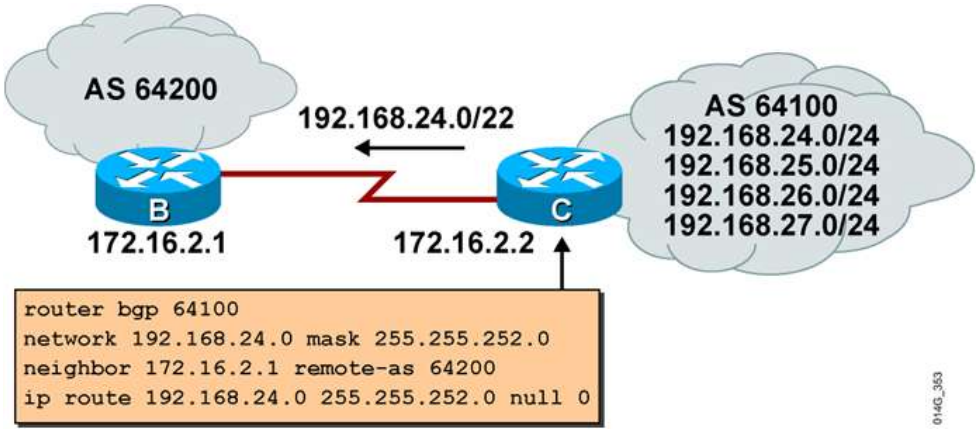
如果明细路由断了，汇总仍然会被引入，且缺乏灵活性。

```
Router(config-router)#
 network network-number [mask network-mask]
```

```
Router(config)#
ip route prefix mask null0
```

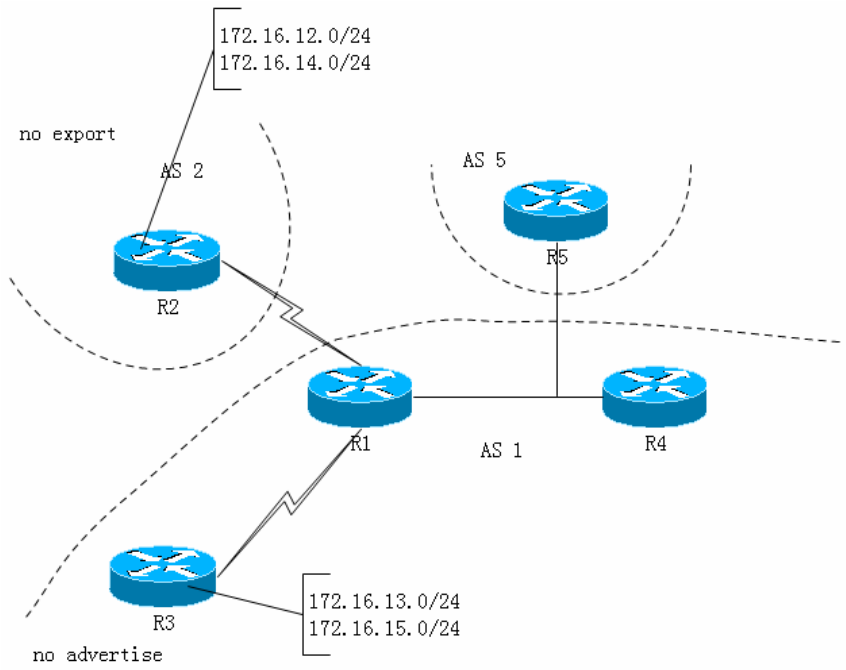
命令network要求路由选择表中有与指定的前缀或掩码完全匹配的条目，为满足这种要求，可配置一条指向接口null0的静态路由，如果IGP执行汇总，则路由选择中可能已以有这样的静态路由。

命令network告诉BGP通告哪些网络，而不如何通告，仅当描写的网络号出现在IP路由选择中后，BGP才会通告它，如下图所示。



B．聚合：aggregate

聚合路由在本路由器上生成一条聚合路由，下一跳为0.0.0.0。



aggregate-address 172.16.12.0 255.255.252.0 ?

advertise-map	**Set condition to advertise attribute**
as-set	**Generate AS set path information**
attribute-map	**Set attributes of aggregate**
route-map	**Set parameters of aggregate**
summary-only	**Filter more specific routes from updates**
suppress-map	**Conditionally filter more specific routes from updates**
nlri	
<cr>	

- Advertise-map
  - 只对 advertise-map 里面匹配的路由进行聚合。当 advertise-map 里面匹配的明细路由全部消失后，即使聚合路由范围内还有其他明细路由，聚合路由也将消

失。当与 as－set 合用时，只继承 advertise-map 里面匹配的明细路由的属性。如果用 summary-only，会将所有的明细包括没有在 advertise-map 里面匹配的路由一起抑制。

- As-set
  - 聚合路由继承明细路由的属性，包括：as-path、local_preference、community、origin－code。与 advertise-map 合用，只继承 advertise-map 里面匹配的明细路由的属性。如果继承了 as-path 属性，继承的 as-path 如果没有在大括号{ }中显示，则有几个算几个 AS；如果继承 AS 是在大括号中排列的，那么只算一个 AS 号。只关心 AS 的号码，不关心顺序。
- As－path、as－seq（as－path）原子聚合不带任何 AS。AS-SET 首先是区别于 atomic-aggregate，产生了 AS 的序列，序列中无分先后顺序，这一点也不同于有明确顺序的 AS-SEQEUENCE
- Attribute－map 和 route－map
- 这两个参数一样，可以将聚合路由的属性清除掉（除了 as-path 属性），添加自己需要添加的属性。Attribute-map 与 as－set 的合用时，能否将聚合的路由的属性重置。（OK 可以改）
- Summary-only 将聚合路由所包括的所有路由都抑制掉，被抑制的路由在 bgp 的转发表里，显示为 s，代表 suppress 的意思。发送更新时，只发送聚合路由。可以与
- neighbor 1.1.1.1 unsuppress-map XX 合用，对特定邻居漏过特定的明细路由。
- Suppress－map，将 suppress－map 里面匹配的路由抑制掉，被抑制的路由在 bgp 的转发表里，显示为 s，代表 suppress 的意思。发送更新时，只发送聚合路由和没有被抑制的明细路由。可以与
- neighbor 1.1.1.1 unsuppress-map XX 合用，对特定邻居漏过特定的明细路由。

四、BGP 路由决策

BGP 的 RIB 包括三部分：

- Adj-RIBs-In：存储了从对等体学习到的路由理新中未经处理的路由信息，这些包含在 Adj-RIBs-In 中的路由被认为是可行路由。
- Loc-RIB：包含了 BGP 发言者对 Adj-RIBs-In 中的路由应用本地策略之后选定的路由
- Adj-RIBs-Out：包含了 BGP 发言者向对等体宣告路由。

BGP 有三个部分既可以是 3 个不同的数据库，也可以是利用指针来区分不同部分的单一数据库。BGP 路由决策通过对 Adj-RIBs-In 中的路由应用本地路由策略，且向 Loc-RIB 和 Adj-RIBs-Out 中输入选定或修改的路由进行路由选择。其有三个阶段。

第一阶段：计算每条可行路由的优先级

第二阶段：从所有可用路由中为特定目的地选出最佳路由，并将其安装到 Loc-RIB 中。

第三阶段：将相应的路由加入到 Adj-RIBs-Out 中，以便向对等体进行宣告。

以下为 BGP 选路原则的 13 条：

（1）weight

cisco 私有的参数。本地有效。缺省条件下，本地始发的路径具有相同的 WEIGHT 值（即 32768），所有其他的路径的 weight 值为 0。越大越优选。影响路由器的出站流量。

（2）local-preference

本地优先级，可以在本 AS 和大联盟内传递。越大越优先。影响路由器的出站流量。默认

情况下，local-preference 为 100。

（3）本地起源

路由器本地始发的路径优先。在 BGP 的转发表里显示为 0.0.0.0。依次降低的优先级顺序是：default-originate(针对每个邻居配置)、default-informaiton-originate（针对每种地址簇配置）、network、redistribute、aggregate－address。

（4）as－path

评估 as-path 的长度，as-path 列表最短的路径优先。

聚合后继承明细路由的属性，在大括号里面的 as-path 在计算长度时，只算一个。在联盟内小括号里面的 AS 号，在选路时，不计算到 as－path 长度里面。

（5）起源代码

评估路由的 origin code 属性，有 3 个 i<e<?。i 代表用 network 将 IGP 引入 BGP 的，或者是聚合等路由，e 代表 EGP，？代表重分布进 BGP 的路由。i 为 0，e 为 1，？为 3。越小越优。

(6)MED

metric 传递不能传出 AS。例外：始发路由器可以 metric 传给邻居，可以是 IBGP/EBGP，但是 EBGP 再传不出去。

MED 相当于 IGP 路由的 metric 值，越小越优先。

（7）EBGP 优于 IBGP

这里 EBGP>联盟内的 EBGP>IBGP。

（8）最近的 IGP 邻居

这里是指 peer 的更新源在我的路由表里显示，哪个最近哪个最优。

OSPF 是否考虑 O、OIA、OE1、OE2？只看 cost 不看 O/OIA/OE。

（9）如果配置了 maximum-path[ibgp]n，如果存在多条等价的路径，会插入多条路径。BGP 默认 maximum-path=1，只能有一条最优路径，但可以通过命令来改变，如果没有 IBGP 参数，默认只能做 EBGP 的负载均衡。做负载均衡还有一个条件，就是上面的 8 条都比不出哪条最优的情况下，才有可能出现负载均衡。

做了 BGP 的负载均衡后，在 BGP 的转发表里还是一个最优，但在路由表里可以出现 2 个下一跳。

（10）最老的

与本端最早建立邻居关系的 peer，被优选。因为它最稳定。但一般不考虑，会跳过这个继续往下选。

如果以下任一条件为真，这一步将会被忽略：

启用了 bgp bestpath compare-routerid，多条路径具有相同的 router-id，因为这些路由都是从同一台路由器接收过来的；当前没有最佳路径。缺乏当前最佳路径的例子发生在正在通告最佳路径的邻居失效的时候。

（11）最低的 ROUTER-ID

BGP 优选来自具有最低的路由器 ID 的 BGP 路由器的路由。Router-id 是路由器上最高的 IP 地址，并且优选环回口。也可以通过 bgp router-id 命令静态的设定路由器 ID。如果路径包含 RR 属性，那么在路径选择过程中，就用 originator－id 来替代路由器 ID。

（12）多跳路径的始发路由器 ID 相同，那么选择 CLUSTER_LIST 长度短的，因为每经过一个 RR，cluster-list 会加上这个 RR 的 router-id

如果多条路径的始发 router-id 相同，那么 BGP 将优选 cluster-list 长度最短的路径。这种情况仅仅出现在 BGP RR 的环境下。

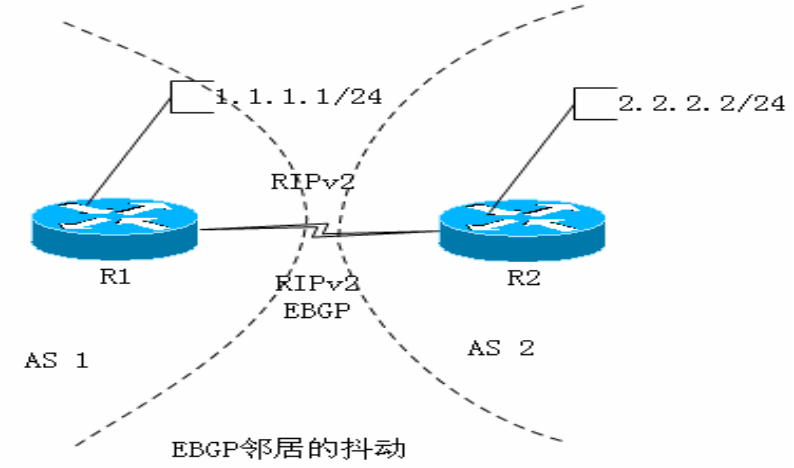
（13）BGP 优选来自于最低的邻居地址的路径。是 BGP 的 neighbor 配置中的那个地址，如果是环回口，则看环回口地址的高低。

BGP 优选来自于最低的邻居地址的路径。这是 BGP 的 neighbor 配置中所使用的 IP 地址，并且它对应于与本地路由器建立 TCP 连接的远端对等体。

五、路由翻动（route flaps）和路由惩罚（route dampening）

路由翻动产生的原因有很多种比如：链路不稳定、路由器接口故障、ISP 工程施工、管理员错误配置和错误故障检查等等都能造成路由翻动，由于路由翻动会造成每台路由器重新计算路由，从而消耗了大量的网络带宽和路由器的 CPU 资源。

BGP 邻居的 flaping



当 R1 与 R2 两台路由器运行 IGP 协议，并且建立 EBGP 的邻居关系，用环回口建立邻居关系。这时假如 R1、R2 将他们的更新源通告进了 BGP，然后通过 BGP 传递给对方，这时由于从 EBGP 学到的路由的 AD 为 20，大于 IGP 的默认 AD，这时会产生邻居的 flaping 现象。

这时 show ip bgp summary 可以看到每经过 60 秒 BGP table version is 1, main routing table version 1 会改变一次。BGP 转发表里变化了多少次。

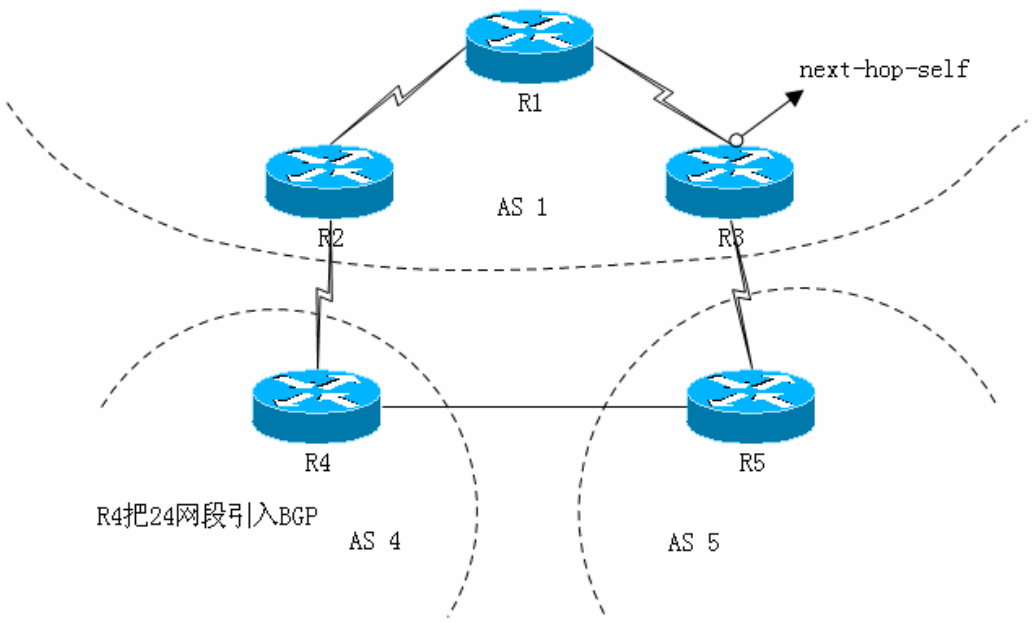
用 debug ip bgp、debug ip bgp update 来查看 BGP 的 flaping。

解决方法：

（1）EBGP 建邻居时不要将环回口引入 BGP。

（2）Network ＋ backdoor

BGP 路由下一跳的 flaping



R1、R2、R3 因为属于同一个 AS，所以运行一个 IGP，R2-R4，R3-R5 之间的链路并没有通告进 IGP 中。R1、R2、R3 IBGP 对等体关系，R3 在指 R1 时，打了 neighbor 1.1.1.1 next-hop-self；R4、R2 ，R5、R3 ，R4、R5 为 EBGP 对等体关系，它们都拿直连接口建立邻居关系。

这时 R4 将它的环回口 4.4.4.0/24 和 R2-R4 的直连网络 24.0.0.0/24 引入 BGP，这时在 R1 上就会产生路由下一跳 flaping 的现象。这时 show ip bgp summary 可以看到每经过 60 秒 BGP table version is 1, main routing table version 1 会改变一次。

解决方法：

（1）静态路由（R1 上静态路由）

（2）在 IBGP 邻居所处的 IGP 中宣告

（3）将与 EBGP 直连的网络重分布进 IGP

（4）neighbor x.x.x.x next-hop-self（R2 指 R1 时输入）

路由惩罚（route dampening）由 RFC2439 描述，它主要由以下三个目的：

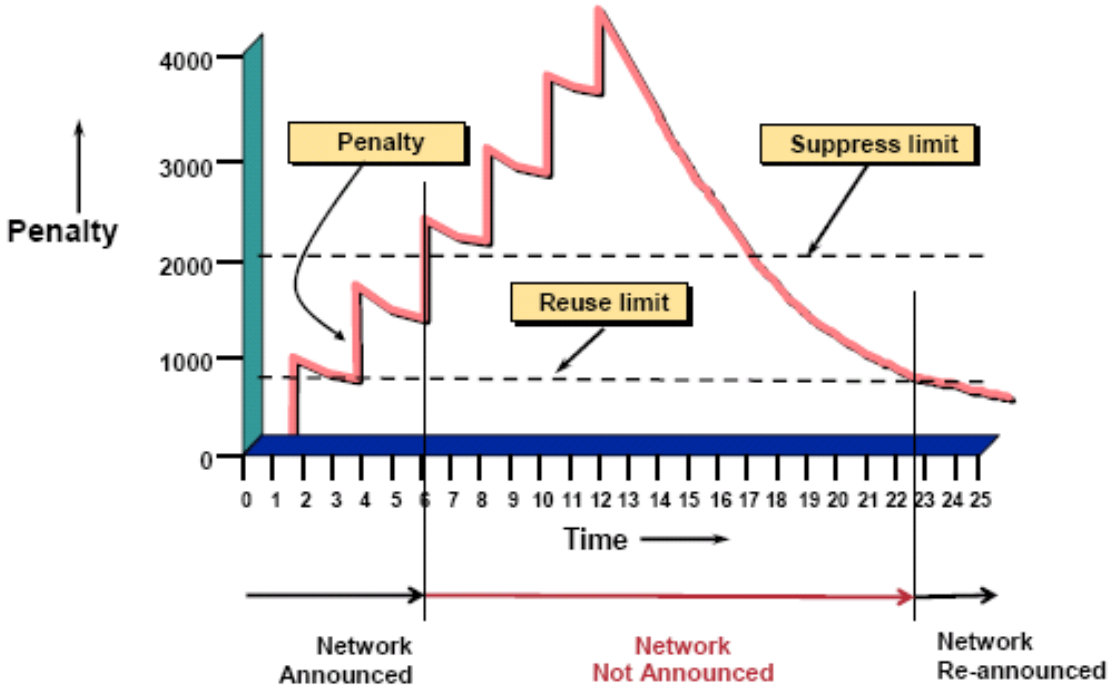
- 提供了一种机制，以减少由于不稳定路由引起的路由器处理负载
- 防止持续的路由抖动
- 增强了路由的稳定性，但不牺牲表现良好的（well－behaved）路由的收敛时间。

ROUTER BG 1
BG DAMP 15 750 2000 60 ---- 针对所有的路由。
BG DAMP ROUTE-MAP XXX
ROUTE-MAP XXX
MATIP ADD PREFIS XX
SET DAM 15 750 2000 60 ---DEFAULT
IP PREFIX XX PERMIT 1.1.1.0/24
SH IP BG 1.1.1.0

SH IP BG DAM PARA

Dampening 为每一条前缀维护了一个路由抖动的历史记录。Dampening 算法包含以下几个参数：

- 历史记录－－－－当一条路由 flaping 后，改路由就会被分配一个惩罚值，并且它的惩罚状态被设置为 history。
- 惩罚值（penalty）－－－－路由每 flaping 一次，这个惩罚值就会增加。默认的路由 flaping 惩罚值为 1000。如果只有路由属性发生了变化，那么惩罚值为 500。这个值是硬件编码的。
- 抑制门限（suppress limit）－－－－如果惩罚值超过了抑制门限，改路由将被惩罚或 dampen。路由状态将由 history 转变为 damp 状态。默认值的抑制门限是 2000，它可以被设置。
- 惩罚状态（damp state）－－－－当路由处于惩罚状态时，路由器在最佳路径选择中将不考虑这条路径，因此也不会把这条前缀通告给它的对等体。
- 半衰期（half life）－－－－在一半的生命周期的时间内，路由的惩罚值将被减少，半衰期的缺省值是 15 分钟。路由的惩罚值每 5 秒钟减少一次。半衰期的值可以被设置。
- 重用门限（reuse limit）－－－－路由的惩罚值不断的递减。当惩罚值降到重用门限以下时，改路由将不再被抑制。缺省的重用门限为 750。路由器每 10 秒钟检查一次那些不需要被抑制的前缀。重用门限时可以被配置的。当惩罚值达到了重用门限的一半时，这条前缀的历史记录（history）将被清除，以便更有效率的使用内存。
- 最大抑制门限/最大抑制时间－－－－如果路由在短时间内表现出极端的不稳定性，然后又稳定下来，那么累计的惩罚值可能会导致这条路由在过长的时间里一直处于惩罚状态。这就是设置最大抑制门限的基本目的。如果路由表现出连续的不稳定性，那么惩罚值就停留在它的上限上，使得路由保持在惩罚状态。最大抑制门限是用公式计算出来的。最大抑制时间为一条路由停留在惩罚状态的最长时间。默认为 60 分钟（半衰期的 4 倍）可以配置。
  - 最大抑制门限＝重用门限×2（最大抑制时间÷半衰期）
  - 由于最大抑制门限为公式算出来的，所以有可能最大抑制门限≤抑制门限，当这种情况发生时，dampening 的设置是没有效果的。如重用门限＝750，抑制门限＝3000，半衰期＝30 分钟，最大抑制时间＝60 分钟。按照这样的配置，算出来的最大抑制门限为 3000，
  - 与抑制门限一样，因为必须超过抑制门限，才能对路由进行 dampening，所以这时 dampening 的设置没有效果。



BGP 的 dampening 仅仅影响 EBGP 的路由。Dampening 是基于每条路径的路由而操作的。如果一条前缀具有两条路径，并且其中一条被惩罚了，那么另一条前缀仍然是可用的，可以通告给 BGP 对等体。

命令：

bgp dampening [route-map XX] [{Half-life reuse-limit suppress-limit Maximum-time }]

如果挂了 route-map，那么就在 route-map 里面匹配特定 EBGP 路由，来设置 dampening 值。

检查命令：

show ip protocol

sh ip bgp dampening ?

dampened-paths 只显示（清除）被抑制的路由。

flap-statistics 显示（清除）所有出现摆动的路由以及该路由出现摆动的次数。

parameters Display details of configured dampening parameters

show ip bgp neighbors 1.1.1.1 dampened-routes

show ip bgp neighbors 1.1.1.1 flap-statistics

六、路由反射器

由于 IBGP 的水平分割问题，所以 IBGP 需要 Full Mesh。由于整个 IBGP full mesh 的话，需要建的 session 数为 n*(n-1)/2。不具有扩展性。所以产生两种解决方法，路由反射器是其中一种，而另一种则是联邦。

路由反射器是被配置为允许它把通过 IBGP 所获悉的路由通告到其他 IBGP 对等体的路由器，路由器反射器与其他路由器有部分 IBGP 对等关系，这些路由器被称为客户。客户间的对等是不需要的，因为路由反射器将在客户间传递通告。 如下图所示。

其优点：减少 AS 内 BGP 邻居关系的数量，从而减少了 TCP 连接数；在 AS 内可以有多个路由反射器，即是为了冗余也是为了分成组，以进一步减少所需 IBGP 会话的数量。路由反射器的路由器可以与非路由反射器的路由器共存，所以配置更简单。

RFC1966 中定义了 3 条 RR 用来决定要宣告哪条路由的规则，具体使用时取决于路由是如何学习到的。

- 如果路由学习自非客户 IBGP 对等体，则仅反射给客户路由器。
- 如果路由学习自某客户，则反射给所有非客户和客户路由器（发起该路由的客户除外）。
- 如果路由学习自 EBGP 对等体，则反射给所有非客户和客户路由器

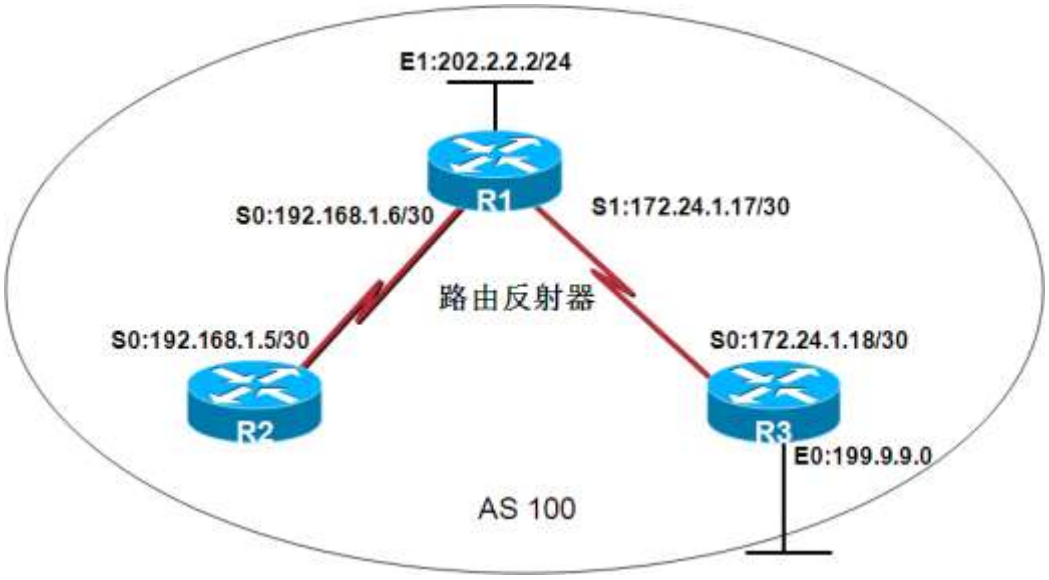
路由反射器的客户并不知道自己是客户。客户和非客户经过路由反射器反射的路由更新将会带上 cluster-list 和 originator，可用于 IBGP 防环。Cluster-id 默认为路由反射器自己的 router-id，可以通过命令 bgp cluster-id 1.1.1.1 来修改，cluster-id 为 32 位的值，可以写成点分十进制，也可以写成十进制数；originator 为 IBGP 内起源路由器的 router-id。路由反射器是 IBGP 的特性，出了 IBGP 后，路由反射器所有的特性消失（即路由携带的 cluster-list 和 originator 全部消失）。

neighbor 1.1.1.1 route-reflector-client

可以通过这条命令来将 IBGP 的 peer 1.1.1.1 变为自己的客户。建议对每个 IBGP 邻居都打上。

当路由反射器的客户 full mesh 时，可以用 no bgp client-to-client reflection 禁止客户到客户的路由反射。可以减少路由更新。

如下图为路由反射器的基本配置。

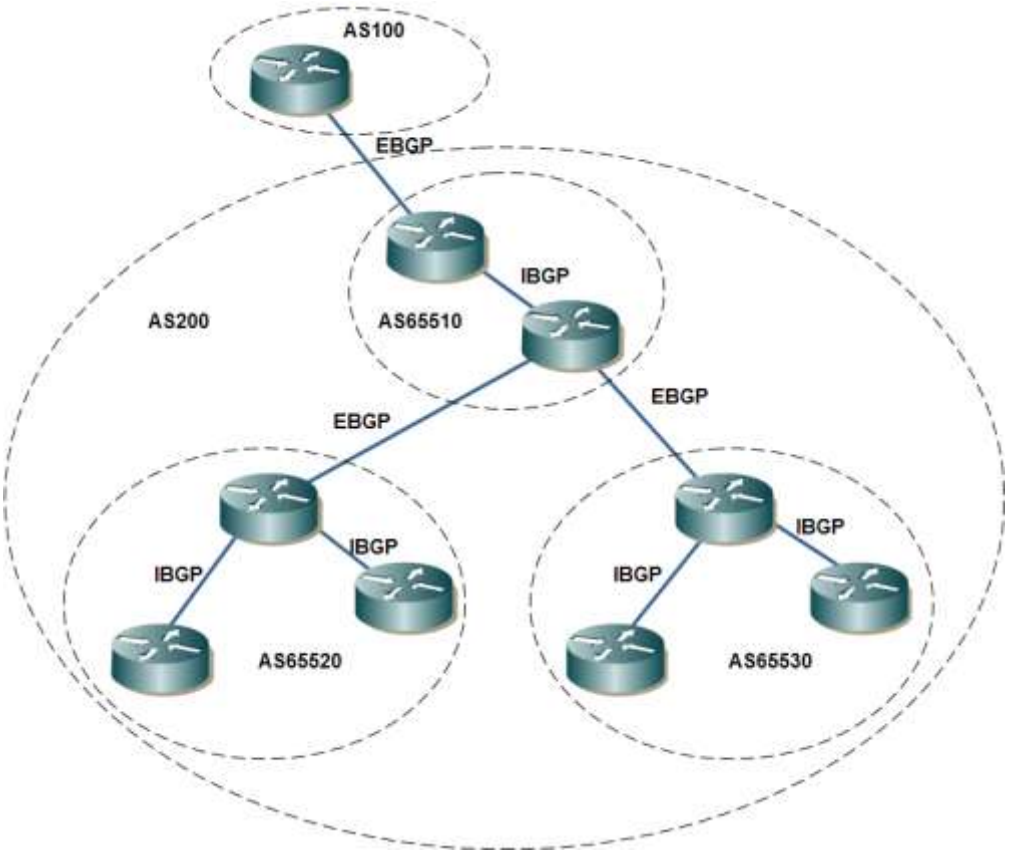



```
R1(config)#router bgp 100
R1(config-router)#neighbor 192.168.1.5 remote-as 100
R1(config-router)#neighbor 192.168.1.5 route-reflector-client
R1(config-router)#neighbor 172.24.1.18 remote-as 100
R1(config-router)#neighbor 172.24.1.18 route-reflector-clien
R1(config-router)#network 20.100.50.0
R1(config-router)#no auto-summary
R1(config-router)#no synchronization
R2(config-router)#neighbor 192.168.1.6 remote-as 100
R2(config-router)#no auto-summary
R2(config-router)#no synchronization
R3(config-router)#neighbor 172.24.1.17 remote-as 100
R3(config-router)#no auto-summary
R3(config-router)#no synchronization
R3(config-router)#network 199.9.9.0
R3(config-router)#aggregate-address 199.0.0.0 255.0.0.0
R3(config)#int l0
R3(config-if)#ip add 199.99.9.1 255.255.255.0
R1(config)#ip prefix-list supernetonly permit 199.0.0.0/8
R1(config-router)#neighbor 192.168.1.5 prefix-list supernetonly out
```

七、BGP 联邦

由于 IBGP 的水平分割问题，所以 IBGP 需要 full mesh。由于整个 IBGP full mesh 的话，需要建的 session 数为 n*(n-1)/2。不具有扩展性。所以产生两种解决方法，联邦是其中一种。联邦既有 EBGP 的特性，又有 IBGP 的特性。

联盟是另一种控制大量 IBGP 对等体的方法，它就是一个被细分为一组子自治系统（称为成员自治系统）的 AS。如下图所示。



联盟增加了两种类型的 AS_PATH 属性

AS_CONFED_SEQUENCE：一个去往特定目的地所经路径上的有序 AS 号列表，其用法与 AS_SEQUENCE 完全一样，区别在于该列表中的 AS 号属于本地联盟中的自治系统。

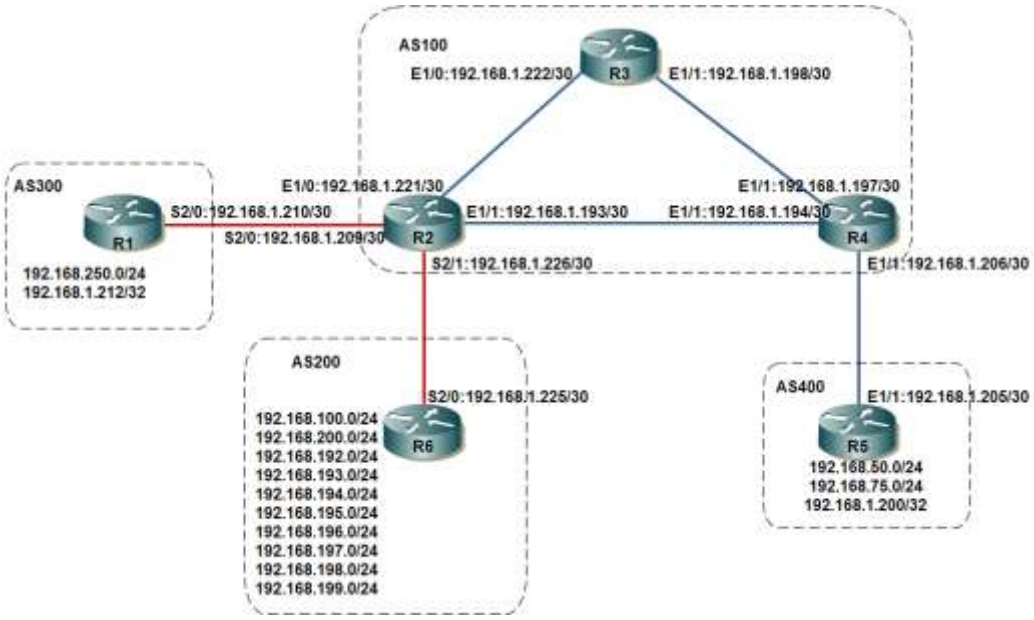
AS_CONFED_SET：一个去往特定目的地所经路径上的无序 AS 号列表，其用法与 AS_SET 完全一样，区别在于该列表中的 AS 号属于本地联盟中的自治系统。

由于 AS_PATH 发生被用于成员自治系统之间，因而保留了环路预防功能。将 Update 消息发送给联盟之外的对等体时，将从 AS_PATH 属性中剥离 AS_CONFED_SEQUENCE 和 AS_CONFED_SET 信息，而将联盟 ID 附加到 AS_PATH 中。

Local_preference 和 MED 可以在联邦内传递。联盟内的小 AS 号，在 as-path 里显示在小括号里，在 as-path 计算长度时，不被考虑。下一跳在联邦内传递不会改变。

八、配置样例 1

下面的示例中涉及到 BGP 的基本配置，涉及到一些基本的知识点，如 EBGP 多跳、更新源使用环回接口、路由映射发布团体属性等，如下图所示。



下面是其参考配置。

```
R1#sh running-config
!
interface Loopback0
 ip address 192.168.250.1 255.255.255.0
!
interface Loopback1
 ip address 192.168.1.213 255.255.255.252
!
interface Loopback6
 ip address 5.5.5.5 255.255.255.0
!
interface Serial2/0
 ip address 192.168.1.210 255.255.255.252
 serial restart-delay 0
!
router bgp 300
 no synchronization
 bgp log-neighbor-changes
 network 192.168.1.212 mask 255.255.255.252
 network 192.168.250.0
 neighbor 6.6.6.6 remote-as 100
 neighbor 6.6.6.6 ebgp-multihop 2
 neighbor 6.6.6.6 update-source Loopback6
 no auto-summary
!
ip route 6.6.6.0 255.255.255.0 192.168.1.209
```

```
!

R2#sh running-config
!
interface Loopback5
 ip address 1.1.1.1 255.255.255.255
!
interface Loopback6
 ip address 6.6.6.6 255.255.255.0
!
interface Ethernet1/0
 ip address 192.168.1.221 255.255.255.252
 duplex half
!
interface Ethernet1/1
 ip address 192.168.1.193 255.255.255.252
 duplex half
!
!
interface Serial2/0
 ip address 192.168.1.209 255.255.255.252
 serial restart-delay 0
!
interface Serial2/1
 ip address 192.168.1.226 255.255.255.252
 serial restart-delay 0
!
router ospf 10
 log-adjacency-changes
 passive-interface Serial2/0
 passive-interface Serial2/1
 network 1.1.1.1 0.0.0.0 area 0
 network 192.168.1.192 0.0.0.3 area 0
 network 192.168.1.220 0.0.0.3 area 0
!
router bgp 100
 no synchronization
 bgp log-neighbor-changes
 neighbor 2.2.2.2 remote-as 100
 neighbor 2.2.2.2 update-source Loopback5
 neighbor 2.2.2.2 next-hop-self
 neighbor 3.3.3.3 remote-as 100
 neighbor 3.3.3.3 update-source Loopback5
```

```
 neighbor 3.3.3.3 next-hop-self
 neighbor 5.5.5.5 remote-as 300
 neighbor 5.5.5.5 ebgp-multihop 2
 neighbor 5.5.5.5 update-source Loopback6
 neighbor 192.168.1.225 remote-as 200
 no auto-summary
!
ip route 5.5.5.0 255.255.255.0 192.168.1.210
!

R3#sh running-config
!
interface Loopback5
 ip address 2.2.2.2 255.255.255.255
!
interface Ethernet1/0
 ip address 192.168.1.222 255.255.255.252
 duplex half
!
interface Ethernet1/1
 ip address 192.168.1.198 255.255.255.252
 duplex half
!
router ospf 10
 log-adjacency-changes
 network 2.2.2.2 0.0.0.0 area 0
 network 192.168.1.196 0.0.0.3 area 0
 network 192.168.1.220 0.0.0.3 area 0
!
router bgp 100
 no synchronization
 bgp log-neighbor-changes
 neighbor 1.1.1.1 remote-as 100
 neighbor 1.1.1.1 update-source Loopback5
 neighbor 3.3.3.3 remote-as 100
 neighbor 3.3.3.3 update-source Loopback5
 no auto-summary
!

R4#sh running-config
!
interface Loopback5
 ip address 3.3.3.3 255.255.255.255
```

```
!
interface Ethernet1/0
 ip address 192.168.1.194 255.255.255.252
 duplex half
!
interface Ethernet1/1
 ip address 192.168.1.197 255.255.255.252
 duplex half
!
interface Ethernet1/2
 ip address 192.168.1.206 255.255.255.252
 duplex half
!
!
router ospf 10
 log-adjacency-changes
 passive-interface Ethernet1/2
 network 3.3.3.3 0.0.0.0 area 0
 network 192.168.1.192 0.0.0.3 area 0
!
router bgp 100
 no synchronization
 bgp log-neighbor-changes
 neighbor 1.1.1.1 remote-as 100
 neighbor 1.1.1.1 update-source Loopback5
 neighbor 1.1.1.1 next-hop-self
 neighbor 2.2.2.2 remote-as 100
 neighbor 2.2.2.2 update-source Loopback5
 neighbor 2.2.2.2 next-hop-self
 neighbor 192.168.1.205 remote-as 400
 no auto-summary
!
R5#sh running-config
!
interface Loopback0
 ip address 192.168.50.1 255.255.255.0
!
interface Loopback1
 ip address 192.168.75.1 255.255.255.0
!
interface Loopback3
 ip address 192.168.1.201 255.255.255.252
!
```

```
interface Ethernet1/0
 ip address 192.168.1.205 255.255.255.252
 duplex half
!
router bgp 400
 no synchronization
 bgp log-neighbor-changes
 network 192.168.1.200 mask 255.255.255.252
 network 192.168.50.0
 network 192.168.75.0
 neighbor 192.168.1.206 remote-as 100
 no auto-summary
!
R6#sh running-config
!
interface Loopback0
 ip address 192.168.100.1 255.255.255.0
!
interface Loopback1
 ip address 192.168.200.1 255.255.255.0
!
interface Loopback3
 ip address 192.168.1.217 255.255.255.252
!
interface Loopback10
 ip address 192.168.192.1 255.255.255.0
!
interface Loopback11
 ip address 192.168.193.1 255.255.255.0
!
interface Loopback12
 ip address 192.168.194.1 255.255.255.0
!
interface Loopback13
 ip address 192.168.195.1 255.255.255.0
!
interface Loopback14
 ip address 192.168.196.1 255.255.255.0
!
interface Loopback15
 ip address 192.168.197.1 255.255.255.0
!
interface Loopback16
```

```
 ip address 192.168.198.1 255.255.255.0
!
interface Loopback17
 ip address 192.168.199.1 255.255.255.0
!
interface Serial2/0
 ip address 192.168.1.225 255.255.255.252
 serial restart-delay 0
!
router bgp 200
 no synchronization
 bgp log-neighbor-changes
 network 192.168.1.216 mask 255.255.255.252
 network 192.168.100.0
 network 192.168.192.0
 network 192.168.193.0
 network 192.168.194.0
 network 192.168.195.0
 network 192.168.196.0
 network 192.168.197.0
 network 192.168.198.0
 network 192.168.199.0
 network 192.168.200.0
 aggregate-address 192.168.192.0 255.255.248.0
 neighbor 192.168.1.226 remote-as 100
 neighbor 192.168.1.226 send-community
 neighbor 192.168.1.226 route-map community out
 no auto-summary
!
!
no ip http server
no ip http secure-server
!
!
access-list 110 permit ip host 192.168.192.0 host 255.255.248.0
!
route-map community permit 10
 match ip address 110
 set community none
!
route-map community permit 20
 set community no-export
!
```

使用一些常用命令来查看其状态信息，如下所示：

```
R2#sh ip bgp
BGP table version is 18, local router ID is 6.6.6.6
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
              r RIB-failure, S Stale
Origin codes: i - IGP, e - EGP, ? - incomplete

   Network          Next Hop            Metric LocPrf Weight Path
*>i192.168.1.200/30 3.3.3.3                  0    100      0 400 i
*> 192.168.1.212/30 5.5.5.5                  0             0 300 i
*> 192.168.1.216/30 192.168.1.225            0             0 200 i
*>i192.168.50.0     3.3.3.3                  0    100      0 400 i
*>i192.168.75.0     3.3.3.3                  0    100      0 400 i
*> 192.168.100.0    192.168.1.225            0             0 200 i
*> 192.168.192.0    192.168.1.225            0             0 200 i
*> 192.168.192.0/21 192.168.1.225            0             0 200 i
*> 192.168.193.0    192.168.1.225            0             0 200 i
*> 192.168.194.0    192.168.1.225            0             0 200 i
*> 192.168.195.0    192.168.1.225            0             0 200 i
*> 192.168.196.0    192.168.1.225            0             0 200 i
*> 192.168.197.0    192.168.1.225            0             0 200 i
*> 192.168.198.0    192.168.1.225            0             0 200 i
*> 192.168.199.0    192.168.1.225            0             0 200 i
*> 192.168.200.0    192.168.1.225            0             0 200 i
*> 192.168.250.0    5.5.5.5                  0             0 300 i

R2#sh ip bgp summary
BGP router identifier 6.6.6.6, local AS number 100
BGP table version is 18, main routing table version 18
17 network entries using 1989 bytes of memory
17 path entries using 884 bytes of memory
5/4 BGP path/bestpath attribute entries using 620 bytes of memory
3 BGP AS-PATH entries using 72 bytes of memory
1 BGP community entries using 24 bytes of memory
0 BGP route-map cache entries using 0 bytes of memory
0 BGP filter-list cache entries using 0 bytes of memory
BGP using 3589 total bytes of memory
BGP activity 17/0 prefixes, 17/0 paths, scan interval 60 secs

Neighbor        V    AS  MsgRcvd  MsgSent   TblVer  InQ  OutQ  Up/Down
State/PfxRcd
2.2.2.2         4   100      15       18       18    0    0 00:12:09
        0
```

```
3.3.3.3          4   100       16        18        18      0     0 00:12:17        3
5.5.5.5          4   300       16        18        18      0     0 00:12:49        2
192.168.1.225    4   200       17        18        18      0     0 00:12:03        12

R4#sh ip bgp
BGP table version is 18, local router ID is 3.3.3.3
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
              r RIB-failure, S Stale
Origin codes: i - IGP, e - EGP, ? - incomplete

   Network          Next Hop            Metric LocPrf Weight Path
*> 192.168.1.200/30 192.168.1.205            0             0 400 i
*>i192.168.1.212/30 1.1.1.1                  0    100      0 300 i
*>i192.168.1.216/30 1.1.1.1                  0    100      0 200 i
*> 192.168.50.0     192.168.1.205            0             0 400 i
*> 192.168.75.0     192.168.1.205            0             0 400 i
*>i192.168.100.0    1.1.1.1                  0    100      0 200 i
*>i192.168.192.0    1.1.1.1                  0    100      0 200 i
*>i192.168.192.0/21 1.1.1.1                  0    100      0 200 i
*>i192.168.193.0    1.1.1.1                  0    100      0 200 i
*>i192.168.194.0    1.1.1.1                  0    100      0 200 i
*>i192.168.195.0    1.1.1.1                  0    100      0 200 i
*>i192.168.196.0    1.1.1.1                  0    100      0 200 i
*>i192.168.197.0    1.1.1.1                  0    100      0 200 i
*>i192.168.198.0    1.1.1.1                  0    100      0 200 i
*>i192.168.199.0    1.1.1.1                  0    100      0 200 i
*>i192.168.200.0    1.1.1.1                  0    100      0 200 i
*>i192.168.250.0    1.1.1.1                  0    100      0 300 i

R4#sh ip route
Codes: 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

Gateway of last resort is not set

B     192.168.192.0/24 [200/0] via 1.1.1.1, 00:12:47
      1.0.0.0/32 is subnetted, 1 subnets
O        1.1.1.1 [110/11] via 192.168.1.193, 00:13:28, Ethernet1/0
```

```
B       192.168.193.0/24 [200/0] via 1.1.1.1, 00:12:47
        2.0.0.0/32 is subnetted, 1 subnets
O          2.2.2.2 [110/21] via 192.168.1.193, 00:13:28, Ethernet1/0
B       192.168.194.0/24 [200/0] via 1.1.1.1, 00:12:47
B       192.168.75.0/24 [20/0] via 192.168.1.205, 00:12:49
        3.0.0.0/32 is subnetted, 1 subnets
C          3.3.3.3 is directly connected, Loopback5
B       192.168.195.0/24 [200/0] via 1.1.1.1, 00:12:47
B       192.168.196.0/24 [200/0] via 1.1.1.1, 00:12:47
B       192.168.197.0/24 [200/0] via 1.1.1.1, 00:12:47
B       192.168.198.0/24 [200/0] via 1.1.1.1, 00:12:48
B       192.168.199.0/24 [200/0] via 1.1.1.1, 00:12:48
B       192.168.200.0/24 [200/0] via 1.1.1.1, 00:12:48
B       192.168.250.0/24 [200/0] via 1.1.1.1, 00:12:49
B       192.168.50.0/24 [20/0] via 192.168.1.205, 00:12:51
        192.168.1.0/30 is subnetted, 7 subnets
B          192.168.1.200 [20/0] via 192.168.1.205, 00:12:51
C          192.168.1.204 is directly connected, Ethernet1/2
C          192.168.1.192 is directly connected, Ethernet1/0
C          192.168.1.196 is directly connected, Ethernet1/1
B          192.168.1.216 [200/0] via 1.1.1.1, 00:12:48
O          192.168.1.220 [110/20] via 192.168.1.193, 00:13:30, Ethernet1/0
B          192.168.1.212 [200/0] via 1.1.1.1, 00:12:49
B       192.168.100.0/24 [200/0] via 1.1.1.1, 00:12:48
B       192.168.192.0/21 [200/0] via 1.1.1.1, 00:12:48

R4#sh ip bgp summary
BGP router identifier 3.3.3.3, local AS number 100
BGP table version is 18, main routing table version 18
17 network entries using 1989 bytes of memory
17 path entries using 884 bytes of memory
5/4 BGP path/bestpath attribute entries using 620 bytes of memory
3 BGP AS-PATH entries using 72 bytes of memory
0 BGP route-map cache entries using 0 bytes of memory
0 BGP filter-list cache entries using 0 bytes of memory
BGP using 3565 total bytes of memory
BGP activity 17/0 prefixes, 17/0 paths, scan interval 60 secs

Neighbor        V    AS  MsgRcvd  MsgSent   TblVer  InQ  OutQ  Up/Down
State/PfxRcd
1.1.1.1         4   100       19       17       18    0     0 00:13:37
14
2.2.2.2         4   100       16       17       18    0     0 00:13:25
0
192.168.1.205   4   400       17       19       18    0     0 00:13:30
3
```

R5#sh ip bgp
BGP table version is 18, local router ID is 192.168.75.1
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
r RIB-failure, S Stale
Origin codes: i - IGP, e - EGP, ? - incomplete

Network	Next Hop	Metric	LocPrf	Weight	Path
*> 192.168.1.200/30	0.0.0.0	0		32768	i
*> 192.168.1.212/30	192.168.1.206			0	100 300 i
*> 192.168.1.216/30	192.168.1.206			0	100 200 i
*> 192.168.50.0	0.0.0.0	0		32768	i
*> 192.168.75.0	0.0.0.0	0		32768	i
*> 192.168.100.0	192.168.1.206			0	100 200 i
*> 192.168.192.0	192.168.1.206			0	100 200 i
*> 192.168.192.0/21	192.168.1.206			0	100 200 i
*> 192.168.193.0	192.168.1.206			0	100 200 i
*> 192.168.194.0	192.168.1.206			0	100 200 i
*> 192.168.195.0	192.168.1.206			0	100 200 i
*> 192.168.196.0	192.168.1.206			0	100 200 i
*> 192.168.197.0	192.168.1.206			0	100 200 i
*> 192.168.198.0	192.168.1.206			0	100 200 i
*> 192.168.199.0	192.168.1.206			0	100 200 i
*> 192.168.200.0	192.168.1.206			0	100 200 i
*> 192.168.250.0	192.168.1.206			0	100 300 i

R5#sh ip route
Codes: 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

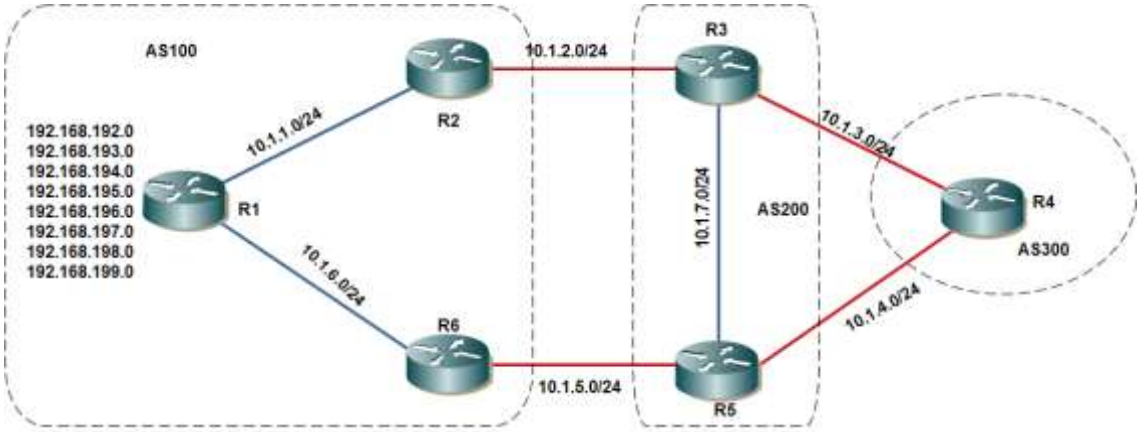
Gateway of last resort is not set

B 192.168.192.0/24 [20/0] via 192.168.1.206, 00:13:03
B 192.168.193.0/24 [20/0] via 192.168.1.206, 00:13:03
B 192.168.194.0/24 [20/0] via 192.168.1.206, 00:13:03
C 192.168.75.0/24 is directly connected, Loopback1
B 192.168.195.0/24 [20/0] via 192.168.1.206, 00:13:03
B 192.168.196.0/24 [20/0] via 192.168.1.206, 00:13:03

```
B       192.168.197.0/24 [20/0] via 192.168.1.206, 00:13:03
B       192.168.198.0/24 [20/0] via 192.168.1.206, 00:13:03
B       192.168.199.0/24 [20/0] via 192.168.1.206, 00:13:03
B       192.168.200.0/24 [20/0] via 192.168.1.206, 00:13:03
B       192.168.250.0/24 [20/0] via 192.168.1.206, 00:13:34
C       192.168.50.0/24 is directly connected, Loopback0
        192.168.1.0/30 is subnetted, 4 subnets
C          192.168.1.200 is directly connected, Loopback3
C          192.168.1.204 is directly connected, Ethernet1/0
B          192.168.1.216 [20/0] via 192.168.1.206, 00:13:04
B          192.168.1.212 [20/0] via 192.168.1.206, 00:13:35
B       192.168.100.0/24 [20/0] via 192.168.1.206, 00:13:04
B       192.168.192.0/21 [20/0] via 192.168.1.206, 00:13:04
```

九、配置样例 2

下面的示例中涉及到聚合路由内容，并将聚合路由使用 **community**、**router-map** 及 **prefix-list** 等功能实现过滤精细路由，拓扑图如下所示。



具体配置如下：

```
R1#sh running-config
!
hostname R1
!
interface Loopback0
 ip address 192.168.192.1 255.255.255.0
!
interface Loopback1
 ip address 192.168.193.1 255.255.255.0
!
interface Loopback2
```

```
 ip address 192.168.194.1 255.255.255.0
!
interface Loopback3
 ip address 192.168.195.1 255.255.255.0
!
interface Loopback4
 ip address 192.168.196.1 255.255.255.0
!
interface Loopback5
 ip address 192.168.197.1 255.255.255.0
!
interface Loopback6
 ip address 192.168.198.1 255.255.255.0
!
interface Loopback7
 ip address 192.168.199.1 255.255.255.0
!
interface Ethernet1/0
 ip address 10.1.1.1 255.255.255.0
 duplex half
!
interface Ethernet1/1
 ip address 10.1.6.1 255.255.255.0
 duplex half
!
router ospf 10
 log-adjacency-changes
 network 10.1.1.0 0.0.0.255 area 0
 network 10.1.6.0 0.0.0.255 area 0
 network 192.168.192.0 0.0.0.255 area 0
 network 192.168.193.0 0.0.0.255 area 0
 network 192.168.194.0 0.0.0.255 area 0
 network 192.168.195.0 0.0.0.255 area 0
 network 192.168.196.0 0.0.0.255 area 0
 network 192.168.197.0 0.0.0.255 area 0
 network 192.168.198.0 0.0.0.255 area 0
 network 192.168.199.0 0.0.0.255 area 0
-----------------------------------------------------------------------------------------------------------------
R2#sh running-config
interface Ethernet1/0
 ip address 10.1.1.2 255.255.255.0
 duplex half
!
```

```
interface Serial2/0
 ip address 10.1.2.1 255.255.255.0
 serial restart-delay 0
!
router ospf 10
 log-adjacency-changes
 network 10.1.1.0 0.0.0.255 area 0
!
router bgp 100
 no synchronization
 bgp log-neighbor-changes
 aggregate-address 192.168.192.0 255.255.248.0
 redistribute ospf 10 metric 50
 neighbor 10.1.2.2 remote-as 200
 neighbor 10.1.2.2 send-community
 neighbor 10.1.2.2 route-map community out
 neighbor 10.1.6.2 remote-as 100
 no auto-summary
!
access-list 110 permit ip host 192.168.192.0 host 255.255.248.0
!
route-map community permit 10
 match ip address 110
 set community none
!
route-map community permit 20
 set community no-export
!
-----------------------------------------------------------------------------------------------------------------------
R3#sh running-config
interface Ethernet1/0
 ip address 10.1.7.1 255.255.255.0
 duplex half
!
interface Serial2/0
 ip address 10.1.2.2 255.255.255.0
 serial restart-delay 0
!
interface Serial2/1
 ip address 10.1.3.1 255.255.255.0
 serial restart-delay 0
!
router bgp 200
```

```
 no synchronization
 bgp log-neighbor-changes
 neighbor 10.1.2.1 remote-as 100
 neighbor 10.1.3.2 remote-as 300
 neighbor 10.1.7.2 remote-as 200
 no auto-summary
!
```

```
R4#sh running-config
interface Serial2/0
 ip address 10.1.3.2 255.255.255.0
 serial restart-delay 0
!
interface Serial2/1
 ip address 10.1.4.1 255.255.255.0
 serial restart-delay 0
!
router bgp 300
 no synchronization
 bgp log-neighbor-changes
 neighbor 10.1.3.1 remote-as 200
 neighbor 10.1.4.2 remote-as 200
 no auto-summary
!
```

```
R5#sh running-config
interface Ethernet1/0
 ip address 10.1.7.2 255.255.255.0
 duplex half
!
interface Serial2/0
 ip address 10.1.4.2 255.255.255.0
 serial restart-delay 0
!
interface Serial2/1
 ip address 10.1.5.1 255.255.255.0
 serial restart-delay 0
!
router bgp 200
 no synchronization
 bgp log-neighbor-changes
 neighbor 10.1.4.1 remote-as 300
 neighbor 10.1.5.2 remote-as 100
```

```
 neighbor 10.1.7.1 remote-as 200
 no auto-summary
!
```

```
R6#sh running-config
interface Ethernet1/0
 ip address 10.1.6.2 255.255.255.0
 duplex half
!
interface Serial2/0
 ip address 10.1.5.2 255.255.255.0
 serial restart-delay 0
!
!
router ospf 10
 log-adjacency-changes
 network 10.1.6.0 0.0.0.255 area 0
!
router bgp 100
 no synchronization
 bgp log-neighbor-changes
 aggregate-address 192.168.192.0 255.255.248.0
 redistribute ospf 10 metric 50
 neighbor 10.1.1.2 remote-as 100
 neighbor 10.1.5.1 remote-as 200
 neighbor 10.1.5.1 send-community
 neighbor 10.1.5.1 route-map community out
 no auto-summary
!
!
ip prefix-list aggregate seq 5 permit 192.168.192.0/21
!
route-map community permit 10
 match ip address prefix-list aggregate
 set community none
!
route-map community permit 20
 set community no-export
!
```

配置完成以后，可以查看聚合路由表。

```
R4#sh ip bgp
BGP table version is 32, local router ID is 10.1.4.1
```

```
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
              r RIB-failure, S Stale
Origin codes: i - IGP, e - EGP, ? - incomplete

   Network          Next Hop            Metric LocPrf Weight Path
*  192.168.192.0/21 10.1.4.2                               0 200 100 i
*>                  10.1.3.1                                0 200 100 i

R3#sh ip bgp
BGP table version is 22, local router ID is 10.1.7.1
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
              r RIB-failure, S Stale
Origin codes: i - IGP, e - EGP, ? - incomplete

   Network          Next Hop            Metric LocPrf Weight Path
* i10.1.1.0/24     10.1.5.2                 50    100      0 100 ?
*>                 10.1.2.1                  0             0 100 ?
* i10.1.6.0/24     10.1.5.2                  0    100      0 100 ?
*>                 10.1.2.1                 50             0 100 ?
* i192.168.192.0/21 10.1.5.2                 0    100      0 100 i
*>                 10.1.2.1                  0             0 100 i
* i192.168.192.1/32 10.1.5.2                50    100      0 100 ?
*>                 10.1.2.1                 50             0 100 ?
* i192.168.193.1/32 10.1.5.2                50    100      0 100 ?
*>                 10.1.2.1                 50             0 100 ?
* i192.168.194.1/32 10.1.5.2                50    100      0 100 ?
*>                 10.1.2.1                 50             0 100 ?
* i192.168.195.1/32 10.1.5.2                50    100      0 100 ?
*>                 10.1.2.1                 50             0 100 ?
* i192.168.196.1/32 10.1.5.2                50    100      0 100 ?
*>                 10.1.2.1                 50             0 100 ?
* i192.168.197.1/32 10.1.5.2                50    100      0 100 ?
*>                 10.1.2.1                 50             0 100 ?
* i192.168.198.1/32 10.1.5.2                50    100      0 100 ?
*>                 10.1.2.1                 50             0 100 ?
* i192.168.199.1/32 10.1.5.2                50    100      0 100 ?
*>                 10.1.2.1                 50             0 100 ?

R5#sh ip bgp
BGP table version is 33, local router ID is 10.1.7.2
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
              r RIB-failure, S Stale
Origin codes: i - IGP, e - EGP, ? - incomplete
```

```
     Network          Next Hop            Metric LocPrf Weight Path
* i10.1.1.0/24       10.1.2.1                 0    100      0 100 ?
*>                   10.1.5.2                50             0 100 ?
* i10.1.6.0/24       10.1.2.1                50    100      0 100 ?
*>                   10.1.5.2                 0             0 100 ?
*> 192.168.192.0/21 10.1.5.2                  0             0 100 i
* i                  10.1.2.1                 0    100      0 100 i
* i192.168.192.1/32 10.1.2.1                50    100      0 100 ?
*>                   10.1.5.2                50             0 100 ?
* i192.168.193.1/32 10.1.2.1                50    100      0 100 ?
*>                   10.1.5.2                50             0 100 ?
* i192.168.194.1/32 10.1.2.1                50    100      0 100 ?
*>                   10.1.5.2                50             0 100 ?
* i192.168.195.1/32 10.1.2.1                50    100      0 100 ?
*>                   10.1.5.2                50             0 100 ?
* i192.168.196.1/32 10.1.2.1                50    100      0 100 ?
*>                   10.1.5.2                50             0 100 ?
* i192.168.197.1/32 10.1.2.1                50    100      0 100 ?
*>                   10.1.5.2                50             0 100 ?
* i192.168.198.1/32 10.1.2.1                50    100      0 100 ?
*>                   10.1.5.2                50             0 100 ?
* i192.168.199.1/32 10.1.2.1                50    100      0 100 ?
*>                   10.1.5.2                50             0 100 ?
```

使用下面的命令查看携带 NO-EPORT COMMUNITY 属性的路由情况

```
R3#sh ip bgp community no-export
BGP table version is 22, local router ID is 10.1.7.1
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
              r RIB-failure, S Stale
Origin codes: i - IGP, e - EGP, ? - incomplete

     Network          Next Hop            Metric LocPrf Weight Path
*> 10.1.1.0/24       10.1.2.1                 0             0 100 ?
*> 10.1.6.0/24       10.1.2.1                50             0 100 ?
*> 192.168.192.1/32 10.1.2.1                50             0 100 ?
*> 192.168.193.1/32 10.1.2.1                50             0 100 ?
*> 192.168.194.1/32 10.1.2.1                50             0 100 ?
*> 192.168.195.1/32 10.1.2.1                50             0 100 ?
*> 192.168.196.1/32 10.1.2.1                50             0 100 ?
*> 192.168.197.1/32 10.1.2.1                50             0 100 ?
*> 192.168.198.1/32 10.1.2.1                50             0 100 ?
*> 192.168.199.1/32 10.1.2.1                50             0 100 ?
```

R5#sh ip bgp community no-export
BGP table version is 33, local router ID is 10.1.7.2
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
r RIB-failure, S Stale
Origin codes: i - IGP, e - EGP, ? - incomplete

Network	Next Hop	Metric	LocPrf	Weight	Path
*> 10.1.1.0/24	10.1.5.2	50		0	100 ?
*> 10.1.6.0/24	10.1.5.2	0		0	100 ?
*> 192.168.192.1/32	10.1.5.2	50		0	100 ?
*> 192.168.193.1/32	10.1.5.2	50		0	100 ?
*> 192.168.194.1/32	10.1.5.2	50		0	100 ?
*> 192.168.195.1/32	10.1.5.2	50		0	100 ?
*> 192.168.196.1/32	10.1.5.2	50		0	100 ?
*> 192.168.197.1/32	10.1.5.2	50		0	100 ?
*> 192.168.198.1/32	10.1.5.2	50		0	100 ?
*> 192.168.199.1/32	10.1.5.2	50		0	100 ?

也可以在上面配置的基础上实现如下策略：

- 通过 R2-R3 链路来宣告 192.168.192.0/24、192.168.193.0/24、192.168.194.0/24
- 通过 R6-R5 链路来宣告 192.168.196.0/24、192.168.197.0/24、192.168.198.0/24
- 不宣告 192.168.195.0/24、192.168.199.0/24

具体配置如下：

R**2#sh running-config**
router bgp 100
no synchronization
bgp log-neighbor-changes
aggregate-address 192.168.192.0 255.255.248.0 suppress-map suppress
!
access-list 1 permit 192.168.195.0 0.0.0.255
access-list 1 permit 192.168.196.0 0.0.3.255
!
route-map suppress permit 10
match ip address 1

R6#sh running-config
router bgp 100
no synchronization
bgp log-neighbor-changes
aggregate-address 192.168.192.0 255.255.248.0 suppress-map suppress
!
ip prefix-list suppress seq 5 permit 192.168.192.0/22 le 24

```
ip prefix-list suppress seq 10 permit 192.168.199.0/24
!
route-map suppress permit 10
 match ip address prefix-list suppress
```

使用命令查看路由状态

```
R3#sh ip bgp
BGP table version is 39, local router ID is 10.1.7.1
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
              r RIB-failure, S Stale
Origin codes: i - IGP, e - EGP, ? - incomplete

   Network          Next Hop            Metric LocPrf Weight Path
* i10.1.1.0/24     10.1.5.2                 50    100      0 100 ?
*>                 10.1.2.1                  0             0 100 ?
* i10.1.6.0/24     10.1.5.2                  0    100      0 100 ?
*>                 10.1.2.1                 50             0 100 ?
* i192.168.192.0/21 10.1.5.2                 0    100      0 100 i
*>                 10.1.2.1                  0             0 100 i
* i192.168.192.1/32 10.1.5.2                50    100      0 100 ?
*>                 10.1.2.1                 50             0 100 ?
* i192.168.193.1/32 10.1.5.2                50    100      0 100 ?
*>                 10.1.2.1                 50             0 100 ?
* i192.168.194.1/32 10.1.5.2                50    100      0 100 ?
*>                 10.1.2.1                 50             0 100 ?
* i192.168.195.1/32 10.1.5.2                50    100      0 100 ?
* i192.168.196.1/32 10.1.5.2                50    100      0 100 ?
* i192.168.197.1/32 10.1.5.2                50    100      0 100 ?
* i192.168.198.1/32 10.1.5.2                50    100      0 100 ?
* i192.168.199.1/32 10.1.5.2                50    100      0 100 ?
```

```
R5#sh ip bgp
BGP table version is 62, local router ID is 10.1.7.2
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
              r RIB-failure, S Stale
Origin codes: i - IGP, e - EGP, ? - incomplete

   Network          Next Hop            Metric LocPrf Weight Path
*> 10.1.1.0/24     10.1.5.2                 50             0 100 ?
* i                 10.1.2.1                  0    100      0 100 ?
*> 10.1.6.0/24     10.1.5.2                  0             0 100 ?
* i                 10.1.2.1                 50    100      0 100 ?
```

*>	192.168.192.0/21	10.1.5.2	0		0 100 i
* i		10.1.2.1	0	100	0 100 i
*>	192.168.192.1/32	10.1.5.2	50		0 100 ?
* i		10.1.2.1	50	100	0 100 ?
*>	192.168.193.1/32	10.1.5.2	50		0 100 ?
* i		10.1.2.1	50	100	0 100 ?
*>	192.168.194.1/32	10.1.5.2	50		0 100 ?
* i		10.1.2.1	50	100	0 100 ?
*>	192.168.195.1/32	10.1.5.2	50		0 100 ?
*>	192.168.196.1/32	10.1.5.2	50		0 100 ?
*>	192.168.197.1/32	10.1.5.2	50		0 100 ?
*>	192.168.198.1/32	10.1.5.2	50		0 100 ?
*>	192.168.199.1/32	10.1.5.2	50		0 100 ?