

## 为了支持 IPv6 各协议做了哪些增加

### OSPFV3

新增 0x0008 Link-LSA

新增 0x2009 Intra-Area-Prefix-LSA

增加了 instance id

### ISISv6

新增两个 TLV

新增 IPv6 Reachability TLV ,Type 236

新增 IPv6 Interface Address TLV,Type 232

增加 NLPID 网络层协议 ID

### BGP4+

新增两个新的可选非过渡路径属性

新增 MP\_REACH\_NLRI 多协议可达 NLRI , type code=14

新增 MP\_UNREACH\_NLRI 多协议不可达 NLRI , type code=15

## BGP 邻居关系

1、BGP open 包影响邻居关系建立的参数？

( 1 ) 版本：一般都是 V4，不一致不能建立。

( 2 ) Router-id：router-id 不能一样，否则不可以建立邻居。

( 3 ) AS 号：建立邻居的时候会检测 open 报文中 as 号是否与本端配置的邻居 as 号一致，如果不一致，不能正常建立邻居

( 4 ) 能力属性：( 单播，组播，IPV4，IPV6，VPNv4，是否支持路由刷新以及 4 字节的 as 号 )

( 5 ) 认证：在 bgp 进程下指 peer 时配置认证，认证密钥要一致。  
注意：holdtime 可以不一致 ( 默认 180s，不一致采用小的 )，能力属性可以不一致，但是支持协议需有交集。

## BGP 邻居建立不成功的原因

1. AS 号或 peer 邻居地址出错；
2. BGP 的 router ID 是否有冲突；
3. BGP 对等体两端是否均采用环回口创建邻居；
4. 物理上非直连的 EBGP 邻居是否配置多跳；
5. 用于创建底层 TCP 的路由是否可达；
6. 创建 BGP 对等体两端认证配置是否一致；
7. BGP 对等体是否配置了 peer x.x.x.x ignore；
8. 是否配置了禁止 TCP 端口 179 的 ACL。

扩展问题 1: 为什么版本不一致无法建立邻居关系？

因为 BGP 各个版本之间不会向下兼容。

扩展问题 2：2 字节和 4 字节 AS 的区别，是否影响邻居关系建立？

2 字节 AS 的范围：1----65535，由于 2 字节 AS 号的范围较小，所以分为私网 AS 号（64512-65535）以及公网 AS 号。

4 字节 AS 的范围：1---- $2^{32}$ ，增大的 AS 号的范围，不分私网、公网 AS 号。

华为设备支持 4 字节 AS 号兼容 2 字节 AS 号。

2 字节和 4 字节 AS 邻居之间建立为旧邻居会话，4 字节和 4 字节 AS 邻居之间建立为新邻居会话。

AS 号 23456 用于衔接 2 字节 AS 和 4 字节 AS。

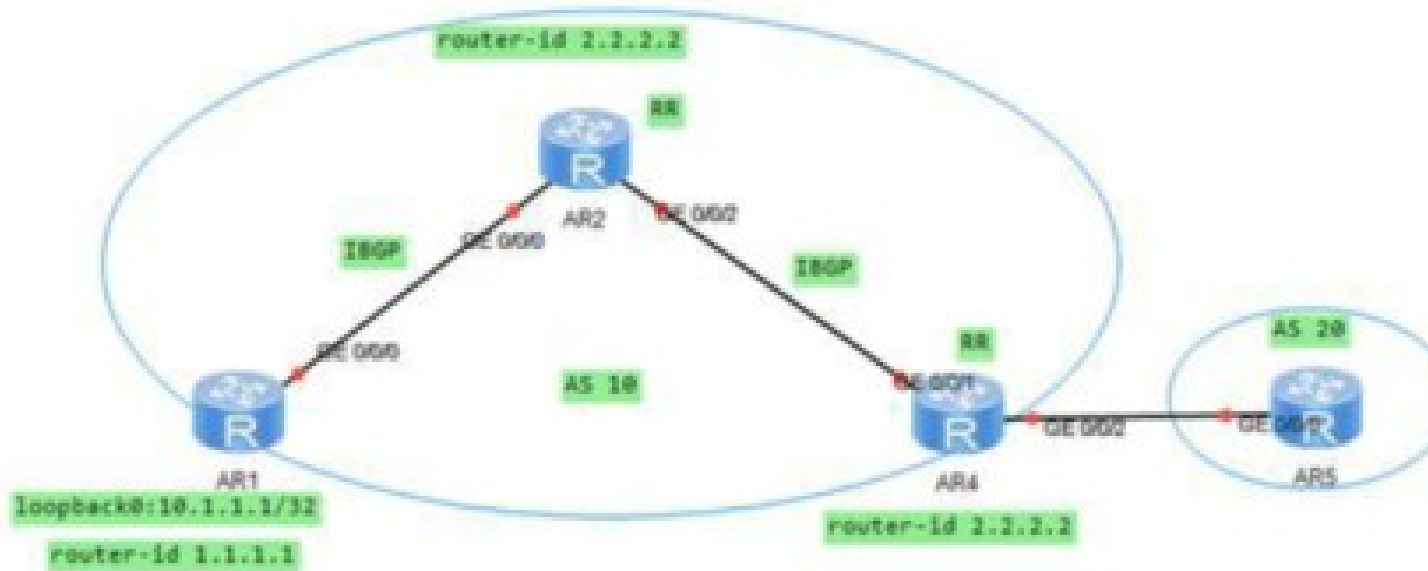
扩展问题 2：如何比较 AS 号？

对方发送过来的 open 报文中会携带 AS 号，然后在拿该 AS 号和 peer 的 AS 号比较，一致则通过；

扩展问题 3：为什么 router-id 一致无法建立邻居关系？

因为 router-id 的其中一个作用是防环。

假设 router-id 一致的话，会导致后期路由传递出现问题。如图所示：



扩展问题 4：那 BGP 中 router-id 的作用是什么？

- 1.标识一台 BGP 路由器
- 2.控制选路
- 3.防环(有 RR 的场景)

扩展问题 5：如果一台路由器的 holdtime 时间为 0S，一台路由器的 holdtime 时间为 180S，能否建立邻居关系？

可以，因为 holdtime 会向小的协商。即收到对方的 open 报文后，会发送 keepalive 报文进行确认。邻居建立后不会周期发送 keepalive 报文；（类似于 OSPF 的按需链路）

扩展问题 6：BGP 邻居建立时，TCP 主动端的如何确定？

先发起 TCP 连接的一端为主动端；

同时发起 TCP 连接建立的话，IP 地址大的一端为主动端；

指定了更新源的一端为主动端；

有明细路由的一端为主动端，  
只有默认路由的一端只能作为被动端；

扩展问题 7：BGP 有哪些认证方式，它们有什么不同？

BGP 只有 MD5 认证和 keychain 认证；

MD5 认证：只针对 TCP 进行认证，认证信息放在 TCP 报文的 option 字段中。

keychain 认证：针对 TCP 和 BGP 报文进行认证，认证信息放在 TCP 报文的 option 字段中，同时也放在 BGP 报文中 9。

扩展问题 8：华为默认就有的可选参数是什么？

4 字节 AS 号、IPV4 单播、路由刷新；

扩展问题 9：华为支持的可选参数有哪些？

1.多协议扩展能力

主地址族：IPV4 子地址族：单播（默认只能建立 IPV4 的单播邻居）、组播、VPNv4；

主地址族：IPV6 子地址族：单播、组播、

2.路由刷新能力（默认开启）

3.4 字节 AS 号支持能力（默认开启）

4.BGP 出方向路由过滤功能（需要手工配置）

## BGP 邻居状态机

BGP 对等体的交互过程中存在 6 种状态机：Idle、Connect、Active、OpenSent、OpenConfirm 和 Established。

（1）Idle：BGP 初始状态。在 Idle 状态下，BGP 拒绝邻居发送的连接请求。只有在收到本设备的 Start 事件后，BGP 才开始尝试和其它 BGP 对等体进行 TCP 连接，并转至 Connect 状态。

Start 事件是由一个操作者配置一个 BGP 过程，或者重置一个已经

存在的过程或者路由器软件重置 BGP 过程引起的。任何状态中收到 Notification 报文或 TCP 拆链通知等 Error 事件后，BGP 都会转至 Idle 状态。

( 2 ) Connect : 表示等待 TCP 连接完成，等待时间默认为 32 秒  
( 重传超时时间 )

如果完成则发送 open，然后进入 opensent 状态

如果 TCP 连接失败，进入 active 状态

//tcp 发起时，对方回复 tcp 连接失败 ( RST=1 )

如果 TCP 连接未完成 ( 32S )，停留在 connect 状态

//tcp 发起对方没回复

( 3 ) Active : 主动向对方发出 TCP 的重连接

如果连接成功则发送 open，然后进入 opensent 状态

如果 TCP 连接失败，进入 active 等待重传超时时间到期，如果 TCP 连接未完成，停留在 connect 状态

( 4 ) opensent : 已经发出 open 报文，协商来自邻居的 open 报文

如果同意邻居的 open 报文的参数，回复 keepalive，进入 openconfirm

如果不同意邻居的 open 报文的参数，回复 notification，退回 idle

( 5 ) openconfirm : 已经发出 keepalive 等待，等待来自邻居 keepalive 报文，等待时间 5s

如果收到对方的 keepalive 报文，则进入 established

如果 5s 内没有收到或者收到 notification 报文，则回到 idle

( 6 ) Established : 表示邻居关系已经完全建立，开始交换 update 报文

收到正确的 update、keepalive，不变

收到 route-refresh 报文，不变

收到错误的 update、keepalive 和收到 notification 或 TCP ( fin=1 ) 的断开连接报文，则转至 Idle

说明：在 BGP 对等体建立的过程中，通常可见的 3 个状态是：Idle、

Active 和 Established。

扩展问题 1：如果一直停在 idle 状态，则可能：

- 1、指 PEER 的地址在路由表中没有存在路由，无法发起 TCP 连接导致，
- 2、没有开启该地址族，
- 3、双方都没有明细路由，用的都是默认路由，
- 4、ebgp 邻居使用环回接口建立没有 ebgp 多跳

扩展问题 2：如果一直停在 connect 状态，则可能：

- 1、TCP 被过滤，对方收不到报文回复不了 TCP 确认
- 2、单通故障
- 3、一端配置认证一端没有认证，则会卡在 connect

扩展问题 3：如果一直停在 active 状态，则可能：

一般为配置错误导致：

- 1、邻居没有配置更新源
- 2、邻居 peer 中指的地址错误
- 3、邻居没有指 peer

扩展问题 4：没有配置 EBGp 邻居多跳，会停留在哪个状态？  
会停留在 idle；

扩展问题 5：认证不通过，会停留在哪个状态？  
会停在 connect，因为认证不通过所以直接丢弃 TCP 报文；

## BGP 报文

BGP 包含哪些报文，以及报文的作用分别是什么？

BGP 对等体间通过以下 5 种报文进行交互，其中 Keepalive 报文为周期性发送，其余报文为触发式发送：

( 1 ) Open 报文：用于建立 BGP 对等体连接。

( 2 ) Update 报文：用于在对等体之间交换路由信息。

每份 update 携带具有相同路径属性的路由，或者拥有对等体之间撤销不可达的路由信息；

( 3 ) Notification 报文：用于中断 BGP 连接。

当 BGP 检测到错误状态时，就向对等体发出 Notification 消息，之后 BGP 连接会立即中断。

( 4 ) Keepalive 报文：用于保持 BGP 连接。

BGP 会周期性的向对等体发出 Keepalive 消息，用来保持连接的有效性。

其消息格式中只包含报文头，没有附加其他任何字段。

( 5 ) Route-refresh 报文：用于在改变路由策略后请求对等体重新发送路由信息。只有支持路由刷新 ( Route-refresh ) 能力的 BGP 设备会发送和响应此报文。

扩展问题 1：BGP 路由器如何撤销路由？

通过发送 update 报文，在 update 报文中会携带撤销路由 ( Withdrawn Routes ) 的消息，此时的 update 报文中只携带需要撤销的路由的路由前缀和掩码；

```
Type: UPDATE Message (2)
Withdrawn Routes Length: 5
> Withdrawn Routes
Total Path Attribute Length: 34
> Path attributes
> Network Layer Reachability Information (NLRI)
```

扩展问题 2：撤销路由时，update 报文会携带路径属性吗？

不会；

扩展问题 3：update 报文的发送条件？

- 1.初次建立邻居关系时，会通过 update 报文交换路由信息；
- 2.有新增或者撤销路由信息时，会触发更新，发送 update 报文；
- 3.收到 refresh-route 报文时，会发送 update 报文；
- 4.在路由器上使用命令 refresh bgp all export 时，也会发送 update 报文；

扩展问题 4：路由刷新报文的作用？

因为 BGP 是不会周期性的更新路由信息的，所以我们需要使用 refresh-route 报文，来实现当接口入方向上策略发生改变时，可以在不中断 BGP 邻居关系的前提下，应用新的策略；

扩展问题 5：BGP 为什么需要有路由刷新才能刷新路由？

- 1.BGP 路由没有周期性更新路由的机制，都是触发更新的，需要通过路由刷新通知的邻居路由器才会发送 update，让 update 报文重新经过入方向的策略器；
- 2.BGP 路由为了稳定，所有策略配置完成之后，手工执行策略，让所有的策略生效

IBGP 和 EBGP 邻居关系区别，配置时要注意什么问题？

( 1 ) EBGP：运行于不同 AS 之间的 BGP 称为 EBGP

1 一般使用物理接口建立邻居

2 存在冗余链路时，一般使用环回接口建立

此时要注意：

环回接口路由可达

指定更新源地址

设置 EBGP 邻居多跳

( 2 ) IBGP：运行于同一 AS 内部的 BGP 称为 IBGP。



- 1 一般使用环回接口，跨跳建立邻居；
- 2 要保证 loopback 接口路由可达，并且是明细路由可达，默认路由无法主动发起 TCP 连接的建立，除非作为被动端；
- 3 要指定更新源地址。

扩展问题 1:BGP 的邻居关系和 IGP 的邻居关系有什么区别？

- 1.手工指定邻居，自动发现邻居
- 2.BGP 单播建立，可以非直连；IGP 一般组播建立，一般要直连
- 3.BGP 存在多种邻居关系 EBGp 和 IBGP；IGP 一般只有一种邻居关系

扩展问题 2:EBGP 邻居使用 loopback 接口建立邻居的时候需要注意什么问题？

- 1.因为 EBGp 邻居之间是没有运行 IGP 协议，所以要保证 loopback 接口可达；
- 2.注意指定更新源地址；
- 3.配置 EBGp 邻居多跳；

扩展问题 3:什么情况下 EBGp 邻居要用环回接口建立邻居关系？  
EBGP 邻居之间存在多条冗余链路时，建议使用环回接口建立，如果都用物理链路间会存在多个 BGP 会话。

扩展问题 4:用环回口建立邻居关系有什么好处？

- 1.比较稳定，因为 loopback 接口为逻辑上的接口，一般不会 down 掉，除非手工删除，或者设备故障；
- 2.当去往对等体存在多条路径时，用环回接口建立邻居关系，具有冗余性；

扩展问题 5:从 EBGp 邻居收到的路由，传给 EBGp 邻居时一定会改变下一跳吗？

默认情况下是会改变的，但是可以通过配置命令使下一跳不改变，或者在一些特殊场景下，发出路由给 EBGP 邻居时下一跳是不改变的，主要是为了防止出现次优路径的问题。

```
bgp 100
```

```
peer 192.168.23.3 next-hop-invariable
```

扩展问题 6: BGP 需要传递和承载的路由数量比 IGP 多吗？为什么？

BGP 协议承载是要比 IGP 协议多；

因为 IGP 是在 AS 内使用，用于发现、计算和维护路由信息的，对设备的性能开销较大，所以不能用于承载过多的路由信息；

BGP 是在 AS 之间使用，就是用于传递和控制路由的，所以可以承载大量的路由信息；

## BGP 通告原则

### ( 1 ) 只将 BGP 的最优路由发布给对等体

当存在多条到达同一目的地址的有效路由时，BGP 设备只将最优路由发布给对等体。当一条路由的前缀标示“>”，表示最优路由。

扩展问题 1: 成为最优路由条件是什么？

1. 成为有效路由（路由的前缀标示\*），也就是要先满足以下条件：

a ) 下一跳地址可达

b ) 如果开启同步，必须满足同步条件

c ) 路由的前缀和路由的下一跳地址不能一致

2. 成为有效路由之后，根据选路规则选出的最优路由

### ( 2 ) 从 EBGP 对等体收到的最优路由，发布给所有 EBGP 和 IBGP 对等体

1 . 向 EBGP 对等体发送路由信息时，会改变 BGP 路由的下一跳地址，下一跳地址为发送路由器的更新源地址；

2 . 向 IBGP 对等体发送路由信息时，不会改变 BGP 路由的下一跳

地址，可能会导致 IBGP 邻居收到此跳路由时下一跳地址不可达，无法使用该路由；

解决：

- a ) 针对 IBGP 邻居使用 next-hop-local 命令，使收到的 EBGp 路由传递指定 IBGP 邻居时将下一跳地址改变为本路由器更新源地址。
- b ) 将下一跳地址宣告进 IGP，使下一跳地址 IGP 可达。

扩展问题 2:为什么向 IBGP 对等体发送路由时，不会改变 BGP 路由的下一跳地址？

- 1.BGP 协议的特点，它认为一个 AS 是一个整体，类似一台路由器，只有将 BGP 路由从 AS 传出的时候，才会改变下一跳；
- 2.下一跳不改变，可以引导 AS 内的路由器访问目的网络时有统一的出口，方便做流量控制；
- 3.因为在一个 AS 内，下一跳不修改的话，本 AS 内的路由器可以根据这个下一跳地址找到离开本 AS 的一条最优的路径。（场景：MA 网络中）

（3）从 IBGP 对等体获得的路由，只发布给 EBGp 对等体，不发送给 IBGP 邻居。

IBGP 水平分割原则，主要用于 AS 内环路防止。IBGP 水平分割会导致路由传递问题，会使 AS 内的 IBGP 无法学习到路由；

解决：

- 1 全互联，让 AS 内使用 BGP 路由器都建立 IBGP 的邻居关系（会导致 AS 内 BGP 的会话数过多， $N(N-1)/2$ ）
- 2 反射器（RR），能打破水平分割的原则进行路由传递
- 3. 联盟

扩展问题 3:IBGP 和 EBGp 路由器之间是如何传递路由的？会出现什么问题？有 RR 的情况呢？

从 EBGp 对等体收到的最优路由，发布给所有 EBGp 和 IBGP 对等

体，传给 IBGP 邻居可能会出现不可达的问题；  
从 IBGP 对等体获得的路由，只发布给 EBGP 对等体，不发送给 IBGP 邻居，带来路由传递的问题，可以使用 RR 解决。

扩展问题 4:反射器 (RR) 和联盟的区别？

反射器 (RR) 是打破水平分割的原则进行路由传递，配置简单；  
联盟是将原有的 AS 划分为多个子 AS，是原本的 IBGP 邻居关系变为 EBGP 邻居关系，这样就可以实现路由的传递，配置复杂。

#### ( 4 ) IBGP 和 IGP 同步。

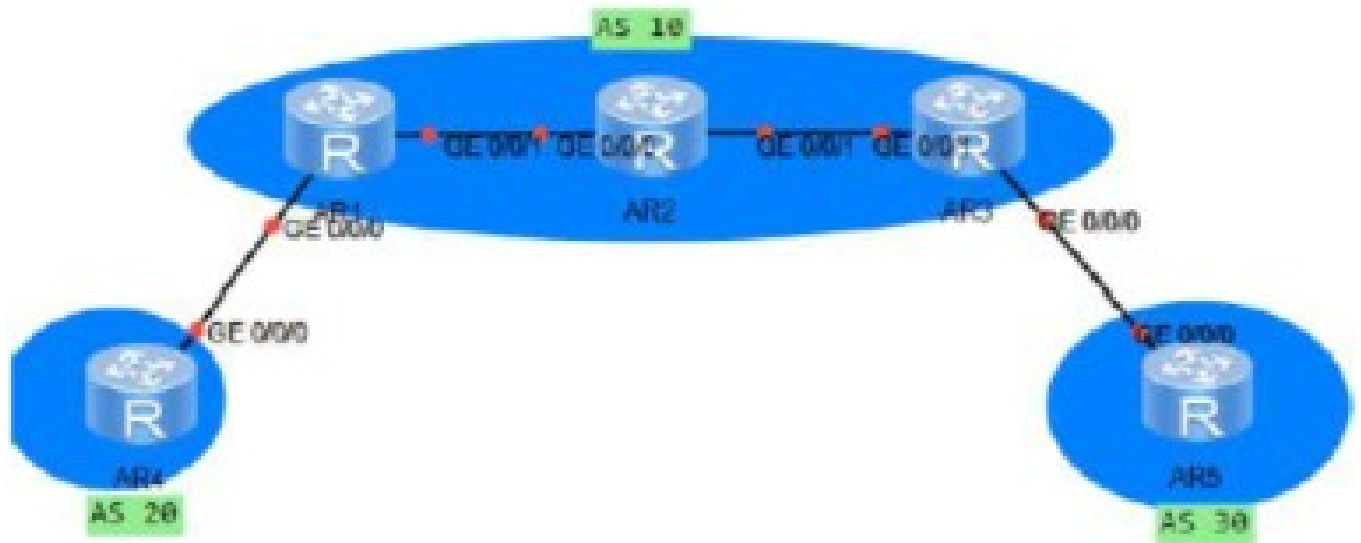
从 IBGP 学习到的 BGP 路由，必须也从 IGP 也学习到此路由，否则此路由为无效路由，不会传递给 BGP 邻居，目的是为了路由黑洞的问题；

说明：BGP 同步虽然可以避免路由黑洞问题，但是会带来路由信息传递的问题。

不建议开启 BGP 同步功能，因为 BGP 主要功能之一为传递路由信息，出现路由黑洞问题采用其他方式解决。

现在 VRP 平台无法开启同步机制，采用其他机制解决路由黑洞问题：

- 1 在数据转发所需经过的路径上也运行 BGP
  - 2 将所需的 BGP 路由引入到 IGP 中
  - 3 MPLS ( R1 和 R3 上使用命令：route recursive-lookup tunnel )
- 场景：



扩展问题 5:如何使用 MPLS 解决 BGP 路由黑洞的问题？

- 1、在一个 AS 内，为 IBGP 邻居的更新源地址事先通过 MPLS LDP 或者静态手工隧道建立起 LSP 隧道；
- 2、配置命令使去往目的网段时，发送给下一跳迭代进隧道中；

扩展问题 6:成为 BGP 路由的方式？

1.network：

将路由表中存在的路由通过 network 的方式成为 BGP 路由

注意：network 的路由前缀和掩码要与路由表中一致

2.import：

将特定协议的路由通过 import 的方式成为 BGP 路由

3.路由聚合：

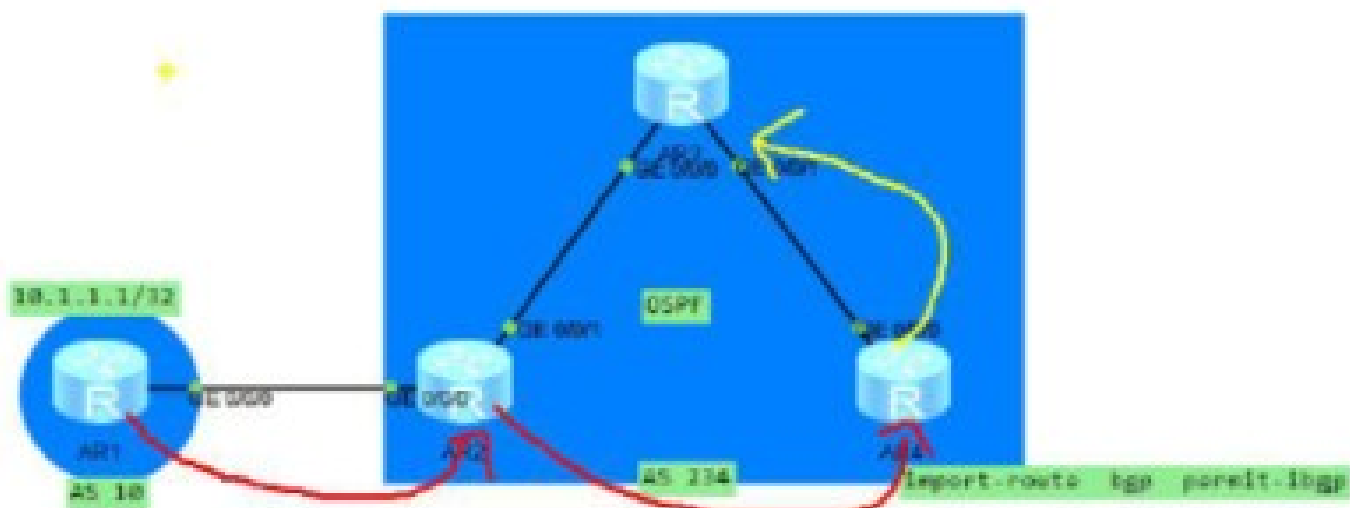
通过手工汇总或者自动汇总产生 BGP 路由

扩展问题 7:将 BGP 路由引入到 IGP 中需要注意哪些问题？

- 1、利用策略控制路由，避免将大量 BGP 路由引入到 IGP 中导致性能较差的设备工作不正常，无法承载。
- 2、默认情况下，只会将 EBGP 邻居学习到的路由引入到 IGP 中，I

BGP 路由引入到 IGP 中需要额外配置命令激活。

扩展问题 8:为什么默认不会将 IBGP 邻居的路由引入到 IGP 中？  
将 IBGP 邻居的路由引入到 IGP 中时，如果引入不当会出现环路场景：



扩展问题 9: bgp 协议中同步的作用是什么？为什么要使用同步？举例说明如何使用同步技术？

(1) BGP 同步的概念

在 IBGP 路由加入路由表并发布给 EBGP 邻居之前，会检查 IGP 路由表，只有在 IGP 也知道这条 IBGP 路由时，它才会被加入到路由表，并发布给 BGP 邻居

(2) 同步的作用（防止路由黑洞出现）

避免出现误导外部 AS 路由器的现象发生，防止一个 AS（不是所有的路由器都运行 BGP）内部出现路由黑洞，即向外部通告了一个本 AS 不可达的虚假的路由

## BGP 属性

路由属性是对路由的特定描述，所有的 BGP 路由属性都可以分为以下 4 类：

( 1 ) 公认必遵属性：所有 BGP 设备都可以识别此类属性，且必须存在于 Update 报文中。如果缺少这类属性，路由信息就会出错。

### 1 Origin 属性：

用来定义路径信息的来源，标记一条路由是怎么成为 BGP 路由的。它有以下 3 种类型：

IGP：具有最高的优先级。通过 network 命令注入到 BGP 路由表的路由，其 Origin 属性为 IGP。

EGP：优先级次之。通过 EGP 得到的路由信息，其 Origin 属性为 EGP。

Incomplete：优先级最低。通过其他方式学习到的路由信息。比如 BGP 通过 import-route 命令引入的路由，其 Origin 属性为 Incomplete。

### 2 AS\_Path 属性：

a ) 按矢量顺序记录了某条路由从本地到目的地址所要经过的所有 AS 编号。

当 BGP Speaker 将这条路由通告到 EBGP 对等体时，便会在 Update 报文中创建一个携带本地 AS 号的 AS\_Path 列表。

当 BGP Speaker 将这条路由通告给 IBGP 对等体时，便会在 Update 报文中创建一个空的 AS\_Path 列表。

当 BGP Speaker 将这条路由通告给 EBGP 对等体时，便会把本地 AS 编号添加在 AS\_Path 列表的最前面（最左面）。收到此路由的 BGP 设备根据 AS\_Path 属性就可以知道去目的地址所要经过的 AS。离本地 AS 最近的相邻 AS 号排在前面，其他 AS 号按顺序依次排列。当 BGP Speaker 将这条路由通告给 IBGP 对等体时，不会改变这条路由相关的 AS\_Path 属性。

b ) AS 间防环：在接收路由时，设备如果发现 AS\_Path 列表中有本 AS 号，则不接收该路由，从而避免了 AS 间的路由环路。

c ) 控制选路：当收到相同前缀的 BGP 路由，优选 AS\_Path 列表 AS 号个数少的。

### 3 Next\_Hop 属性

作用：记录了路由的下一跳信息。

BGP 的下一跳属性和 IGP 的有所不同，不一定就是邻居设备的 IP 地址。

通常情况下，Next\_Hop 属性遵循下面的规则：

1 BGP Speaker 在向 EBGP 对等体发布某条路由时，会把该路由信息的下一跳属性设置为本地与对端建立 BGP 邻居关系的接口地址。

2 BGP Speaker 将本地始发路由发布给 IBGP 对等体时，会把该路由信息的下一跳属性设置为本地与对端建立 BGP 邻居关系的接口地址。

3 BGP Speaker 在向 IBGP 对等体发布从 EBGP 对等体学来的路由时，并不改变该路由信息的下一跳属性。

**(2) 公认任意属性：**所有 BGP 设备都可以识别此类属性，但不要求必须存在于 Update 报文中，即就算缺少这类属性，路由信息也不会出错。

### **1 Local\_Pref 属性**

作用：表明路由的 BGP 优先级，用于判断流量离开 AS 时的最佳路由。

使用场景：当 BGP 的设备通过不同的 IBGP 对等体得到目的地址相同但下一跳不同的多条路由时，将优先选择 Local\_Pref 属性值较高的路由。

a)Local\_Pref 属性仅在 IBGP 对等体之间有效，不通告给其他 AS。

b)Local\_Pref 属性可以手动配置，如果路由没有配置 Local\_Pref 属性，BGP 选路时将该路由的 Local\_Pref 值按缺省值 100 来处理。

### **2 ATOMIC-AGGREGATE：原子聚合属性**

作用：用于提醒管理员哪些路由被抑制了明细；

产生条件：手工聚合，并抑制明细；

**(3) 可选过渡属性：**BGP 设备可以不识别此类属性，如果 BGP



设备不识别此类属性，但它仍然会接收这类属性，并通告给其他对等体。

**1 AGGREGATOR**：聚合者，用于标识是哪个 AS 中的哪一台路由器上进行聚合的；携带 ROUTER-ID, 以及 AS 号

**2 COMMUNITY**：相当于 IGP 的 TAG，有基本的团体属性和扩展的团体属性

**(4) 可选非过渡属性**：BGP 设备可以不识别此类属性，如果 BGP 设备不识别此类属性，则会被忽略该属性，且不会通告给其他对等体。

**1 MED ( MULTI\_EXIT\_DISC )**：

作用：用于判断流量进入 AS 时的最佳路由。

应用场景：当一个运行 BGP 的设备通过不同的 EBGP 对等体得到目的地址相同但下一跳不同的多条路由时，在其它条件相同的情况下，将优先选择 MED 值较小者作为最佳路由。

MED 属性仅在相邻两个 AS 之间传递，收到此属性的 AS 一方不会再将其通告给任何其他第三方 AS。

MED 属性可以手动配置，如果路由没有配置 MED 属性，BGP 选路时将该路由的 MED 值按缺省值 0 来处理。

**2 Cluster-list 属性**：

路由反射器和它的客户机组成一个集群 ( Cluster )，使用 AS 内唯一的 Cluster ID 作为标识。

作用：为了防止集群间产生路由环路，路由反射器使用 Cluster\_List 属性，记录路由经过的所有集群的 Cluster ID。

当一条路由第一次被 RR 反射的时候，RR 会把本地 Cluster ID 添加到 Cluster List 的前面。如果没有 Cluster\_List 属性，RR 就创建一个。当 RR 接收到一条更新路由时，RR 会检查 Cluster List。

如果 Cluster List 中已经有本地 Cluster ID，丢弃该路由；如果没有本地 Cluster ID，将其加入 Cluster List，然后反射该更新路由。

**3 Originator-ID 属性**：

Originator ID 由 RR 产生，使用的 Router ID 的值标识路由的始发者。

作用：防止集群内产生路由环路。

当一条路由第一次被 RR 反射的时候，RR 将 Originator\_ID 属性加入这条路由，标识这条路由的发起设备。如果一条路由中已经存在了 Originator\_ID 属性，则 RR 不会创建新的 Originator\_ID 属性。当设备接收到这条路由的时候，将比较收到的 Originator ID 和本地的 Router ID，如果两个 ID 相同，则不接收此路由。

扩展问题 1:为什么这么分类？

BGP 协议是用于路由的传递和控制选路的，同时现网中有各种各样的场景和需求，所以要对 BGP 的各个属性进行分类，便于了解各种属性的特征及应用场景，每种分类对应不同的应用场景；厂家也可以根据该属性属于哪种大类，来决定是否设计时支持该属性。

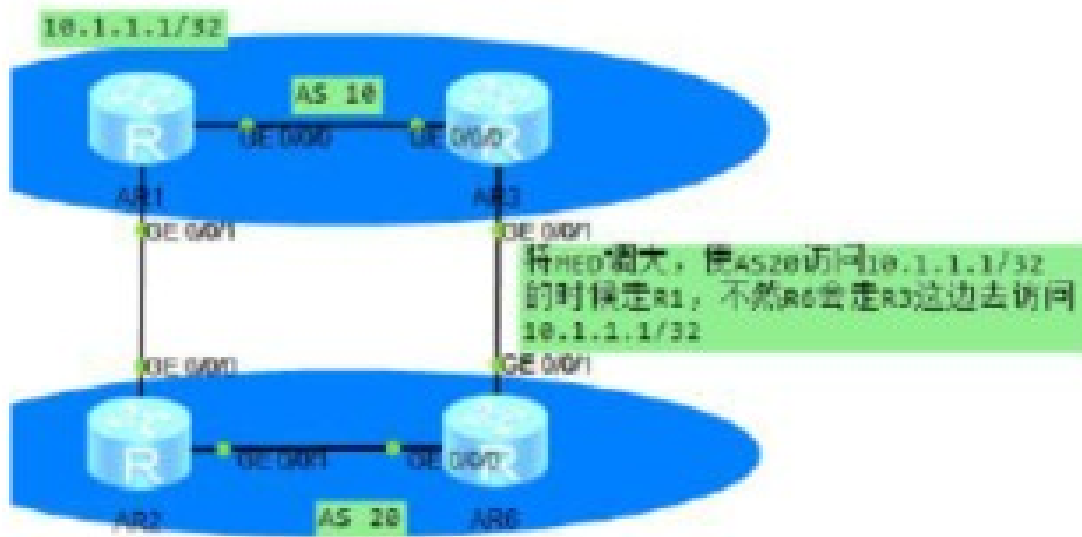
扩展问题 2:BGP 路由器如何得知相应的属性属于哪一类？

每个路径属性中有一个 flags 位标识

```
~ Path attributes
  ~ Path Attribute - ORIGIN: IGP
    ~ Flags: 0x40, Transitive, Well-known, Complete
      0... .... = Optional: Not set
      .1... .... = Transitive: Set
      ..0. .... = Partial: Not set
      ...0 .... = Extended-Length: Not set
      .... 0000 = Unused: 0x0
      Type Code: ORIGIN (1)
      Length: 1
      Origin: IGP (0)
```

扩展问题 3:介绍 med，并举场景？

Med 是多出口鉴别器，主要用于控制其他的 AS 如何来访问本 AS。



扩展问题 4: 可选非传递属性可传递的范围？

MED 传递给 EBGp 邻居后，只能在 EBGp 邻居所在的 AS 内部传递。

CLUSTER\_LIST 只能在本 AS 内传递。

ORIGINATOR\_ID 只能在 AS 内传递。

扩展问题 5: Community 在 MPLS 里的扩展属性有哪些？

Community 在 MPLS 里的扩展属性有 Route-target, SoO, Cost-community；

扩展问题 6: MP-BGP 新增加了哪一类的属性，分别哪些？

MP-BGP 中新增了可选非过渡属性：

MP\_REACH\_NLRI：Multiprotocol Reachable NLRI，多协议可达 NLRI。

用于发布可达路由及下一跳信息。

MP\_UNREACH\_NLRI：Multiprotocol Unreachable NLRI，多协议不可达 NLRI。

用于撤销不可达路由。

```

v Path Attribute - MP_REACH_NLRI
  > Flags: 0x80, Optional, Non-transitive, Complete
    Type Code: MP_REACH_NLRI (14)
    Length: 64
    Address family identifier (AFI): IPv6 (2)
    Subsequent address family identifier (SAFI): Unicast (1)
  v Next hop network address (32 bytes)
    Next Hop: 2001:db8::2
    Next Hop: fe80::c002:bff:fe7e:0
    Number of Subnetwork points of attachment (SNPA): 0
  v Network layer reachability information (27 bytes)
    > 2001:db8:2:2::/64
    > 2001:db8:2:1::/64
    > 2001:db8:2::/64

```

扩展问题 7:如何理解簇的概念？

路由反射器及其客户机的集合，一台客户机不一定属于一个簇(即一台客户机可以属于多个簇)

扩展问题 8:联盟在传递路由时，as\_path 表现形式是怎样的？

1.当 BGP Speaker 将这条路由通告给 EBGP 对等体时，便会把本地 AS 编号添加在 AS\_Path 列表的最前面（最左面）。收到此路由的 BGP 设备根据 AS\_Path 属性就可以知道去目的地址所要经过的 AS。离本地 AS 最近的相邻 AS 号排在前面，其他 AS 号按顺序依次排列

2.当 BGP Speaker 将这条路由通告给 IBGP 对等体时，不会改变这条路由相关的 AS\_Path 属性

扩展问题 9:AS-Path 属性有 4 个值，你知道吗？

( 1 ) AS\_SEQUENCE

( 2 ) AS\_SET

( 3 ) AS\_CONFED\_SEQUENCE

#### ( 4 ) AS\_CONFED\_SET

SET 和 SEQUENCE 的不同之处在于：

SET 选项下的 AS 列表通常用于路由聚合，将来自不同 AS 的 AS 号无序排列在 AS 列表里；

而 SEQUENCE 选项下的 AS 列表是有序的，每经过一个 AS 都会将其 AS 号排列在列表的前端。

AS\_CONFED 仅仅只能应用于 BGP 联盟的情况下，一旦路由信息向外部 AS 更新时，AS\_CONFED 将会被删去。

扩展问题 10:联盟在路由传递给外部 as 时，怎么知道要剥离成员 a s 的 as 号？

因为在配置 confederation 的时候，会配置一条命令 confederation peer-as 64513，让联盟内的 BGP 路由器知道同一个联盟内的子 A S 有哪些成员，而没有被指定的则为其他的联盟，所以当从联盟发出的时候会将子 AS 的 AS 号剥离（即将被圆括号括起来的 AS 号剥离）；

扩展问题 11:ATOMIC-AGGREGATE 属性的作用以及产生的条件？

Atomic-aggregate 属性用于通知管理员哪些丢失的 BGP 属性；

产生条件：手工聚合并抑制明细；

扩展问题 12:aggregate 属性的作用？

用于标识是哪一个 AS 中的，哪一台路由器对路由进行了汇总（AS 号以及 router-id）；

**BGP 选路规则：**

**1．优选协议首选值（Preferred\_Value）最高的路由。**

协议首选值是华为特有属性，该属性仅在本地有效，默认为 0，越大越优。

**2．优选本地优先级（Local\_Preference）最高的路由。**

如果路由没有本地优先级，BGP 选路时将该路由按缺省的本地优先级 100 来处理。Local-Pref 用于选择流量离开 AS 时的最佳路由，也就是控制流量从哪个出口离开 AS，只在 IBGP 中传递，默认为 100，越大越优。

### 3. 优选本地生成的路由 ( 本地生成的路由优先级高于从邻居学来的路由 )

本地生成的路由包括通过 network 命令或 import-route 命令引入的路由、手动聚合路由和自动聚合路由。

手动聚合>自动聚合>network>import>从对等体学到的

1. 优选聚合路由 ( 聚合路由优先级高于非聚合路由 )。
2. 通过 aggregate 命令生成的手动聚合路由的优先级高于通过 summary automatic 命令生成的自动聚合路由。
3. 通过 network 命令引入的路由的优先级高于通过 import-route 命令引入的路由。

### 4. 优选 AS\_Path 短的路由

执行 bestroute as-path-ignore 命令后，BGP 选路时，忽略 AS\_Path 的比较。

### 5. 比较 Origin 属性，起源类型 IGP>EGP>Incomplete

### 6. 优选 MED ( Multi Exit Discriminator ) 值最低的路由。

BGP 只比较来自同一个 AS ( 不包括联盟的子 AS ) 的路由的 MED 值。即，只有两条路由的 AS\_SEQUENCE 属性的第一个 AS 号相同时，BGP 才会比较二者的 MED 值。MED 属性主要作用是用来控制来自邻居 AS 的流量从哪个入口进入到本 AS 中。默认 MED 值为 0，越小越优。

路由的 MED 值为 0，如果接收到的路由条目经过一个 AS 进行中转，那么 MED 值将会丢失，设置为空。MED 值为空实际上等同于值为 0

执行 compare-different-as-med 命令后，BGP 将强制比较来自不同自治系统中的邻居的路由的 MED 值。

### 7. 优选从 EBGp 学来的路由 ( EBGp>IBGP )。

## 8. 优选下一跳 IGP 的 Metric 最小的路由。

如果配置了负载分担，当上述所有规则相同，且存在多条 As\_Path 完全相同的外部路由，则根据配置的路由条数选择多条路由进行负载分担。

## 9. 优选 Cluster\_List 最短的路由。

## 10. 优选 Router\_ID 最小的路由器发布的路由。

## 11. 优选具有较小 IP 地址的邻居学来的路由。

扩展问题 2:第 3 条选路原则和第 5 条选路原则有什么区别？

第 3 条选路原则，是针对自己产生的路由，当设备通过 network 和 import 产生相同的路由时，只会将 network 产生的路由发给 BGP 对等体。

第 5 条选路原则，是针对邻居产生的路由，路由器从两个不同的邻居处收到，通过 network 和 import 产生相同的路由，那么路由器会优选 network 产生的路由。

扩展问题 3:来自不同 AS 的相同 BGP 路由，会比较 MED 值吗？

不会比较 MED。当然也可以通过配置一条命令：compare-different-as-med 使设备即使从不同 AS 收到相同路由时，也会比较 MED 值；

扩展问题 4:BGP 比较到哪一条就可以实现负载分担？

如果在前八条都一致，配置 maximum load-balancing >=2 命令后，满足如下所有条件的多条 BGP 路由会成为等价路由，形成负载分担：

原始下一跳不相同。

都是聚合路由，或者都不是聚合路由；

AS\_Path 属性完全相同；

扩展问题 5:为什么设计者要让 EBGp 优于 IBGP ？



此时属于选路规则第 7 条，证明上面的选路规则是一样的。  
IBGP 过来的路由是在本 AS 内传递了之后才传递给本台路由器的，  
(经过域内其他设备)  
EBGP 邻居传来的路由是离目的网络更近的路径，  
这样设计，可以在一定程度上防止次优路径

## BGP 防环

### 1、BGP 有哪些防环机制？

#### ( 1 ) AS 内防环：

- 1 IBGP 邻居之间通过 IBGP 的水平分割防止环路。
- 2 配置路由反射器之后，路由反射器通过 CLUSTER\_LIST 防止环路；路由反射器的客户端通过 ORIGINATOR\_ID 防止环路。
- 3 在 confederation 内部，confederation 成员 AS 间使用成员的 AS 号来防环。

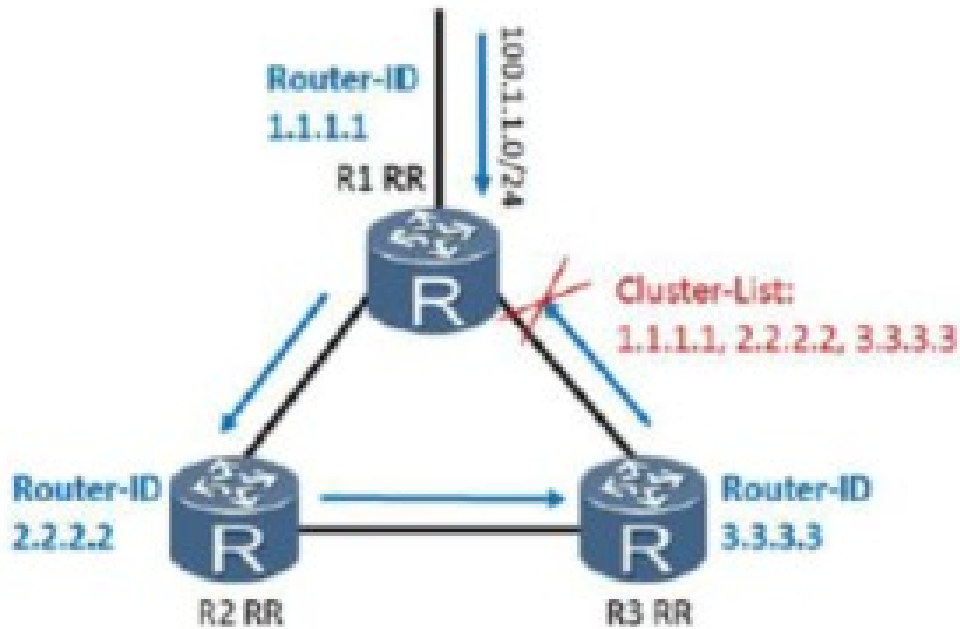
#### ( 2 ) AS 间防环：EBGP 邻居通过 AS\_PATH 防止环路。

### 2、BGP 路由经过 RR，会添加几种属性？如何避免环路的？

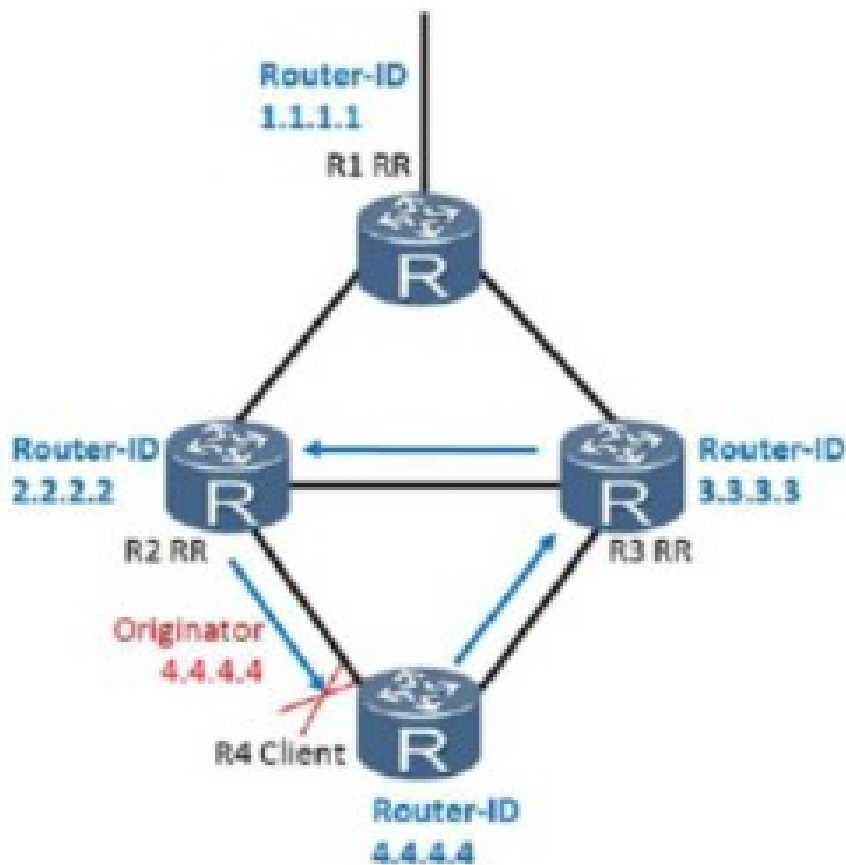
#### ( 1 ) cluster-list ( 路由反射器 )

cluster-list 就是路由沿途经反射器反射，RR 默认就会将自己的 cluster-id(默认为 router-id)写进 cluster-list，接收路由器，一旦发现 cluster-list 里面有跟自己 cluster-id 一样的，就会拒收路由，从而起到防环的作用。





( 2 ) originator-id ( 路由反射器的客户端 )  
路由器发现 originator-id 跟自己 router-id 一致就不收  
一台客户机不一定属于一个簇(即一台客户机可以属于多个簇 )



扩展问题 4:一定要接收带有本 AS 的路由时可以怎么做？为什么要设计这种机制？

- ( 1 ) 可以使用 peer allow-as-loop
- ( 2 ) 可以使用 route-policy ( apply as-path overwrite , 覆盖 as 号 )
- ( 3 ) 做聚合 ( 不加 as-set )
- ( 4 ) 可以使用 peer fake-as

一个公司如果有两个站点，但是只申请了一个 as 号的时候。  
此时就需要做相应的操作，不然会因为 AS-Path 防环机制而导致一边学习不到路由