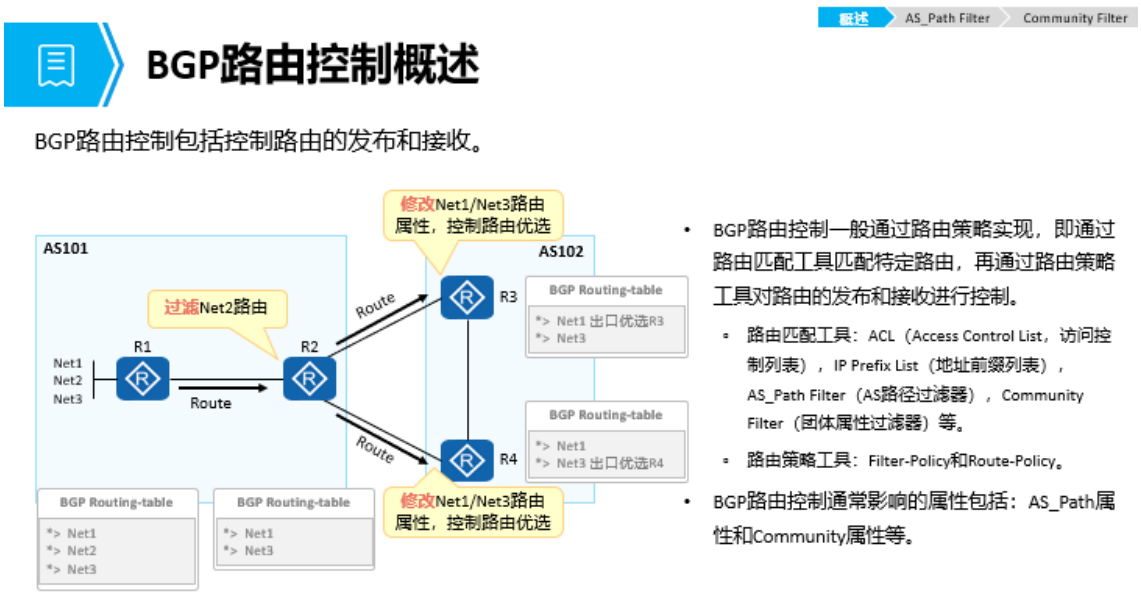
BGP 高级特性

- 在大型网络中通常会部署 BGP，相比于 IGP，BGP 拥有更加灵活的路由控制能力。每一条 BGP 路由都可以携带多个路径属性，针对其属性也有特有的路由匹配工具，包括：AS_Path Filter 和 Community Filter。根据实际组网需求，可以实施路由策略，控制路由的接收和发布。
- 同时，为了提升网络性能，BGP 提供了各种高级特性以及多种组网部署方案。
- 本课程将介绍 BGP 路由控制的原理与配置，介绍常用的 BGP 高级特性，包括：ORF、对等体组、安全特性，还会介绍 BGP 路由反射器的组网部署方式。



- 如图所示：
- R1 和 R2 在 AS101 中，建立 IBGP 邻居关系；R3 和 R4 在 AS102 中，分别与 R2 建立 EBGP 邻居关系。
- R1 有三个直连网段 Net1、Net2 和 Net3，并将三个网段通告进 BGP 路由中。

- 在 R2 可以通过 BGP 路由控制，过滤 Net2 的路由，则 R2 的 BGP 路由表中就没有 Net2 的路由条目。
- 在 R3 和 R4 可以通过 BGP 路由控制，分别修改 Net1 和 Net3 的路由属性，控制路由优选，使 AS102 中设备访问 Net1 优选 R3 作为出口设备，访问 Net3 优选 R4 作为出口设备。
- 说明：ACL、IP Prefix List、Filter-Policy、Route-Policy 和 BGP 路径属性等内容本课程不再赘述，详细内容请参考《HCIP-Datacom-Core Technology》课程。

正则表达式

- 正则表达式是按照一定的模板来匹配字符串的公式，由普通字符（例如字符 a 到 z）和特殊字符组成。
- 普通字符：匹配的对象是普通字符本身。
- 包括所有的大写和小写字母、数字、标点符号以及一些特殊符号。
- 例如：a 匹配 abc 中的 a，10 匹配 10.113.25.155 中的 10，@匹配 xxx@xxx.com 中的@。
- 特殊字符：配合普通字符匹配复杂或特殊的字符串组合。
- 位于普通字符之前或之后用来限制或扩充普通字符的独立控制字符或占位符。
- 用来描述它前面的字符的重复使用方式。
- 限定一个完整的范围。

特殊字符举例 (1)

- 类型1：

.	匹配任意单个的字符，包括空格	0.0 匹配0x0、020、...
^	匹配行首的位置，即一个字符串的开始	^10 匹配10.1.1.1，不匹配20.1.1.1
$	匹配行尾的位置，即一个字符串的结束	1$ 匹配10.1.1.1，不匹配10.1.1.2
_	下划线，匹配任意的一个分隔符 匹配一个逗号（,）、左花括号（{）、右花括号（}）、左圆括号、右圆括号 匹配输入字符串的开始位置（同^） 匹配输入字符串的结束位置（同$） 匹配一个空格	_10 匹配(10，{10，空格10等 10_ 匹配10)，10}，10空格等
|	管道字符，逻辑或。x|y，匹配x或y	100|200 匹配100或者200
\	转义字符，用来将下一个字符（特殊字符或普通字符）标记为普通字符	* 匹配*

- 类型2：

*	匹配前面的子正则表达式0次或多次	10* 匹配1、10、100、1000、...	(10)* 匹配空、10、1010、101010、...
+	匹配前面的子正则表达式1次或多次	10+ 匹配10、100、1000、...	(10)+ 匹配10、1010、101010、...
?	匹配前面的子正则表达式0次或1次	10? 匹配1或10	(10)? 匹配空或10

- 说明：圆括号()可以用来定义操作符的范围和优先度。例如，gr(a|e)y 等价于 gray|grey。

特殊字符举例 (2)

- 类型3：

[xyz]	匹配正则表达式中包含的任意一个字符	[123] 匹配255中的2
[^xyz]	匹配正则表达式中未包含的字符	[^123] 匹配123之外的任何字符
[a-z]	匹配正则表达式指定范围内的任意字符	[0-9] 匹配0到9之间的所有数字
[^a-z]	匹配正则表达式指定范围外的任意字符	[^0-9] 匹配所有非数字字符（即匹配0到9之外的任何字符）

- 请思考以下方框内的正则表达式可以匹配哪些字符串。

类型1

```
^a.$
^100_
^100$
100$|400$
^\(65000\)$
```

类型2

```
abc*d
abc+d
abc?d
a(bc)?d
```

类型3

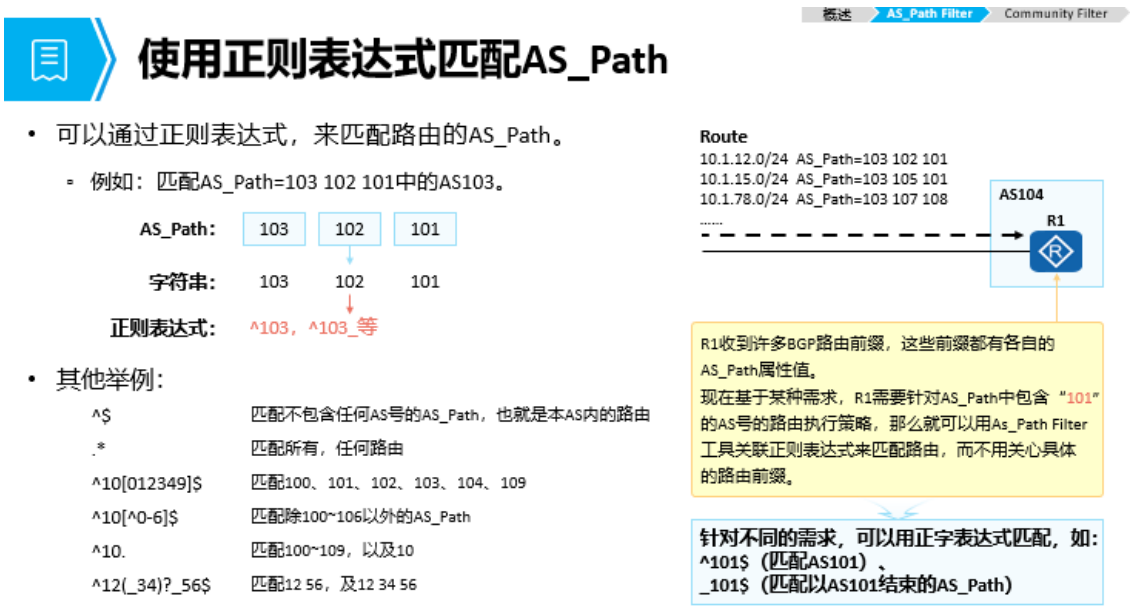
```
[abcd]
[a-c 1-2]$
[^act]$
[123].[7-9]
```

- 思考题：
- 类型 1：
- **^a.$**：匹配一个以字符 a 开始，以任意单一字符结束的字符串，如 a0，a!，ax 等。
- **^100_**：匹配以 100 为起始的字符串，如：100、100 200、100 300 400 等。
- **^100$**：只匹配 100。

- **100$|400$**：匹配以 100 或 400 结束的字符串，如：100、1400、300 400 等。
- **^\(65000\)$**：只匹配(65000)。
- 类型 2：
- **abc*d**：匹配 c 字符 0 次或多次，如：abd、abcd、abccd、abcccd、abccccdef 等。
- **abc+d**：匹配 c 字符 1 次或多次，如：abcd、abccd、abcccd、abccccdef 等。
- **abc?d**：匹配 c 字符 0 次或 1 次，如：abd、abcd、abcdef 等。
- **a(bc)?d**：匹配 bc 字符串 0 次或 1 次，如：ad、abcd、aaabcdef 等。
- 类型 3：
- **[abcd]**：匹配 abcd 中任意一个字符，即只要出现了 a、b、c、d 中的任意字符即可，如：ax、b！、abc、d0 等。
- **[a-c 1-2]$**：匹配以字符 a、b、c、1、2 结束的字符串，如：a、a1、62、xb、7ac 等。
- **[^act]$**：匹配不以字符 a、c、t 结束的字符串，如：ax、b！、d 等。
- **[123].[7-9]**：匹配如：1 7、2x9、348 等。
-

AS_Path 属性是 BGP 的公认必遵属性，所有的 BGP 路由都必须携带该属性。这个属性记录了 BGP 路由在传递过程中所经过的所有 AS 的号码。

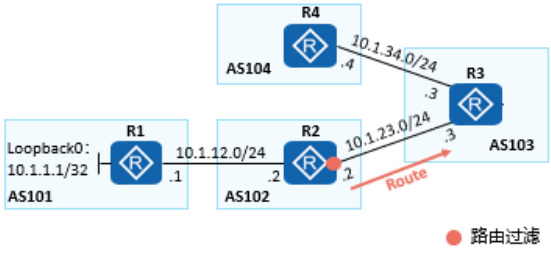
- AS_Path 属性值可以是 0 个、1 个或多个 AS 号码的集合。



- 在同一个过滤器编号下，可以定义多条过滤规则（permit 或 deny 模式）。在匹配过程中，这些规则之间是“或”的关系，即只要路由信息通过其中一项规则，就认为通过由该过滤器编号标识的这组 AS_Path Filter。
- 命令：[Huawei] **ip as-path-filter** { *as-path-filter-number* | *as-path-filter-name* } { **deny** | **permit** } *regular-expression*
- *as-path-filter-number*：指定的 AS 路径过滤器号。整数形式，取值范围 1~256。
- *as-path-filter-name*：指定的 AS 路径过滤器名称。字符串形式，区分大小写，不支持空格，长度范围是 1~51，且不能都是数字。当输入的字符串两端使用双引号时，可在字符串中输入空格。
- **deny**：指定 AS 路径过滤器的匹配模式为拒绝。
- **permit**：指定 AS 路径过滤器的匹配模式为允许。
- *regular-expression*：指定 AS 路径正则表达式。字符串形式，支持空格，取值范围是 1~255 个字符。
- AS 路径过滤器的默认行为是 deny，即路由如果没有在某一次过滤中被 permit 则最终不能通过该过滤器的过滤。如果一个过滤器中的所有过滤规则都是 deny，则没有路由能通

过该过滤器的过滤，这种情况下需要在多次（或一次）deny之后设置一次 permit，允许其余所有路由通过过滤器的过滤。

AS_Path Filter的配置举例 (1)



R2传递EBGP路由给R3，其中部分路由为R2本地始发，另一部分为AS101传递给R2再由R2更新给R3。

现在R2上部署路由策略，拒绝始发于AS101的路由。

1、创建AS_Path Filter。

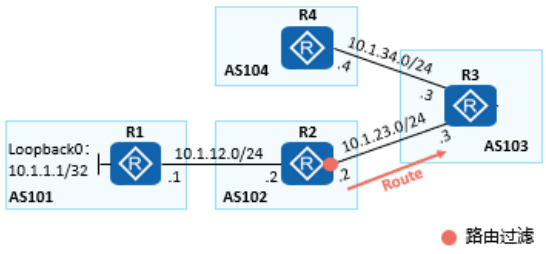
```
[R2] ip as-path-filter 1 deny _101$
[R2] ip as-path-filter 1 permit .*
```

拒绝始发于AS101的路由，但其他路由可以通过。

2、（直接调用方式）应用AS_Path Filter。

```
[R2] bgp 102
[R2-bgp] peer 10.1.23.3 as-number 103
[R2-bgp] ipv4-family unicast
[R2-bgp-af-ipv4] peer 10.1.23.3 as-path-filter 1 export
```


AS_Path Filter的配置举例 (2)



R2传递EBGP路由给R3，其中部分路由为R2本地始发，另一部分为AS101传递给R2再由R2更新给R3。

现在R2上部署路由策略，拒绝始发于AS101的路由。

1、创建AS_Path Filter。

```
[R2] ip as-path-filter 1 deny _101$
[R2] ip as-path-filter 1 permit .*
```

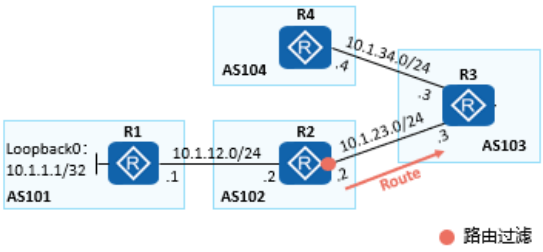
拒绝始发于AS101的路由，但其他路由可以通过。

2、（路由策略方式）应用AS_Path Filter。

```
[R2] route-policy AS_Path permit node 10
[R2-route-policy] if-match as-path-filter 1
[R2-route-policy] quit
[R2] bgp 102
[R2-bgp] peer 10.1.23.3 as-number 103
[R2-bgp] ipv4-family unicast
[R2-bgp-af-ipv4] peer 10.1.23.3 route-policy AS_Path export
```

概述 AS_Path Filter Community Filter

查看AS_Path Filter相关信息



1、查看AS_Path Filter。

```
[R2]display ip as-path-filter 1
As path filter number: 2
                deny       _101$
                permit     .*
```

2、显示BGP表中所有AS_Path被该正则表达式匹配的路由。

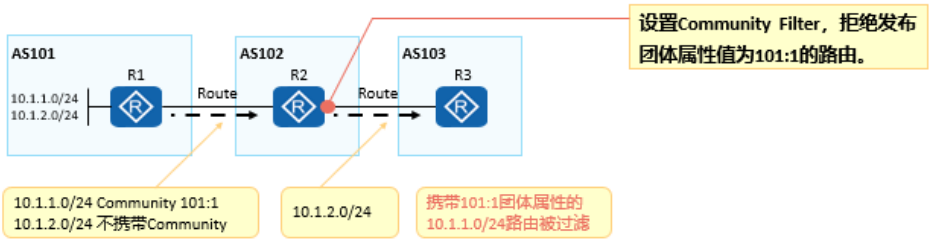
```
[R2]display bgp routing-table regular-expression _101$

Total Number of Routes: 1

BGP Local router ID is 10.1.12.2
Status codes: * - valid, > - best, d - damped,
              h - history,  i - internal, s - suppressed, S - Stale
              Origin : i - IGP, e - EGP, ? - incomplete
   Network       NextHop    MED  LocPrf  PrefVal  Path/Ogn
*> 10.1.1.1/32   10.1.12.1   0              0      101i
```

始发于AS101的10.1.1.1/32路由被AS_Path Filter过滤

概述 AS_Path Filter Community Filter

路由匹配工具：Community Filter

- Community Filter与Community属性配合使用，可以在不便使用IP Prefix List和AS_Path Filter时，降低路由管理难度。
- 团体属性过滤器有两种类型：
  - 基本Community Filter。匹配团体号或公认Community属性。
  - 高级Community Filter。使用正则表达式匹配团体号。



- Community 属性为可选过渡属性，可以标识具有相同特征的路由，而不用考虑零散路由前缀和繁多的 AS 号。即可以将某些路由分配特定的 Community 属性值，之后就可以基于 Community 值而不是网络号/掩码来匹配路由并执行相应的路由策略。

概述 AS_Path Filter Community Filter

Community属性

- 公认Community属性

团体属性名称	团体属性号	说明
Internet	0（0x00000000）	设备在收到具有此属性的路由后，可以向任何BGP对等体发送该路由。缺省情况下，所有的路由都属于Internet团体。
No_Advertise	4294967042（0xFFFFFF02）	设备收到具有此属性的路由后，将不向任何BGP对等体发送该路由。
No_Export	4294967041（0xFFFFFF01）	设备收到具有此属性的路由后，将不向AS外发送该路由。
No_Export_Subconfed	4294967043（0xFFFFFF03）	设备收到具有此属性的路由后，将不向AS外发送该路由。如果使用了联盟，也不向联盟内其他子AS发布此路由。

- Community属性格式：
  - 一个Community属性值的长度为32bit，可使用两种形式呈现：
    - 十进制整数格式。
    - AA：NN格式，其中AA表示AS号，NN是自定义的编号。指定不向自治系统外部通告路由。

概述 AS_Path Filter Community Filter

设置Community的基础配置命令

1. 在路由策略中，设置路由的Community属性值

[Huawei-route-policy] **apply community** { *community-number* | *aa:nn* | **internet** | **no-advertise** | **no-export** | **no-export-subconfed** } [**additive**]

2. 将团体属性发布给对等体（组）

[Huawei-bgp-af-ipv4] **peer** { *group-name* | *ipv4-address* | *ipv6-address* } **advertise-community**

缺省情况下，BGP不将团体属性发布给任何对等体（组）。

- 命令：[Huawei-route-policy] **apply community** { *community-number* | *aa:nn* | **internet** | **no-advertise** | **no-export** | **no-export-subconfed** } [**additive**]
- *community-number* | *aa:nn*：指定团体属性中的团体号。一条命令中最多可以配置 32 个团体号。整数形式，community-number 的取值范围是 0~4294967295，aa 和 nn 的取值范围都是 0~65535。
- **internet**：表示可以向任何对等体发送匹配的路由。缺省情况下，所有的路由都属于 Internet 团体。
- **no-advertise**：表示不向任何对等体发送匹配的路由。即收到具有此属性的路由后，不能发布给任何其他的 BGP 对等

体。

- **no-export**：表示不向 AS 外发送匹配的路由，但发布给其它子自治系统。即收到具有此属性的路由后，不能发布到本地 AS 之外。
- **no-export-subconfed**：表示不向 AS 外发送匹配的路由，也不发布给其它子自治系统，即收到具有此属性的路由后，不能发布给任何其他的子自治系统。
- **additive**：表示追加路由的团体属性。

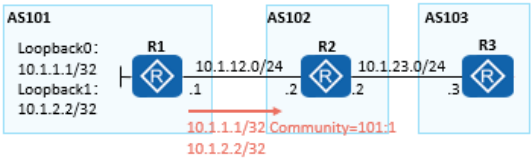
- 命令：[Huawei] **ip community-filter** { **basic** *comm-filter-name* | *basic-comm-filter-num* } { **permit** | **deny** } [*community-number* | *aa:nn* | **internet** | **no-export-subconfed** | **no-advertise** | **no-export**]
- **basic** *comm-filter-name*：指定基本团体属性过滤器名称。字符串形式，区分大小写，取值范围是 1~51 个字符，且不能都是数字。
- *basic-comm-filter-num*：指定基本团体属性过滤器号。整数形式，取值范围 1~99。
- **deny**：指定团体属性过滤器的匹配模式为拒绝。
- **permit**：指定团体属性过滤器的匹配模式为允许。
- *community-number*：指定团体号。整数形式，取值范围 0~4294967295。

- *aa:nn*：指定团体号。一条命令最多可以指定 20 个团体属性号。aa 和 nn 都是整数形式，取值范围都是 0~65535。
- **internet**：表示可以向任何对等体发送匹配的路由。
- **no-export-subconfed**：指定不向自治系统外部通告路由。如果使用了联盟，则不会向联盟中的其他子自治系统通告路由。
- **no-advertise**：指定不通告给其他对等体。
- **no-export**：指定不向自治系统外部通告路由。如果使用了联盟，则不向联盟外部通告路由，但会通告给联盟中的其他子自治系统。
- 命令：[Huawei] **ip community-filter** { **advanced** *comm-filter-name* | *adv-comm-filter-num* } { **permit** | **deny** } *regular-expression*
- **advanced** *comm-filter-name*：指定高级团体属性过滤器名称。字符串形式，区分大小写，取值范围是 1～51 个字符，且不能都是数字。
- *adv-comm-filter-num*：指定高级团体属性过滤器号。整数形式，取值范围 100~199。
- *regular-expression*：指定团体属性正则表达式。字符串形式，支持空格，区分大小写，取值范围是 1~255。

- 命令：[Huawei-route-policy] **if-match community-filter** { *basic-comm-filter-num* [**whole-match**] | *adv-comm-filter-num* }

- 命令：[Huawei-route-policy] **if-match community-filter** *comm-filter-name* [**whole-match**]
- *basic-comm-filter-num*：指定基本团体属性过滤器号。整数形式，取值范围是 1~99。
- *adv-comm-filter-num*：指定高级团体属性过滤器号。整数形式，取值范围是 100~199。
- *comm-filter-name*：指定团体属性过滤器名称。字符串形式，区分大小写，不支持空格，长度范围是 1~51，且不能都是数字。当输入的字符串两端使用双引号时，可在字符串中输入空格。
- **whole-match**：表示完全匹配，即所有的团体都必须出现。仅对基本团体属性过滤器生效。

配置Community属性 (1)



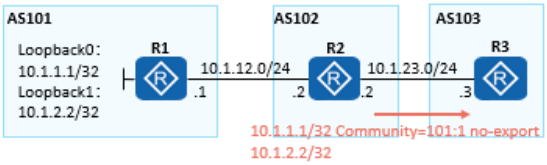
在R1上部署路由策略，使其在通告BGP路由10.1.1.1/32时携带Community属性值101:1，其他路由不携带。

1、R1上部署路由策略为路由添加Community属性，并允许将Community属性传给EBGP对等体R2。

```
[R1] ip ip-prefix 1 permit 10.1.1.1 32
[R1] route-policy Community permit node 10
[R1-route-policy] if-match ip-prefix 1
[R1-route-policy] apply community 101:1
[R1-route-policy] quit
[R1] route-policy Community permit node 20
[R1-route-policy] quit
[R1] bgp 101
[R1-bgp] peer 10.1.12.2 as-number 102
[R1-bgp] peer 10.1.12.2 route-policy Community export
[R1-bgp] peer 10.1.12.2 advertise-community
[R1-bgp] network 10.1.1.1 32
[R1-bgp] network 10.1.2.2 32
```

- 命令：[R1] **route-policy** Community **permit node** 20
- 通过配置该命令允许 10.1.2.2/32 路由被正常通告。

配置Community属性 (2)



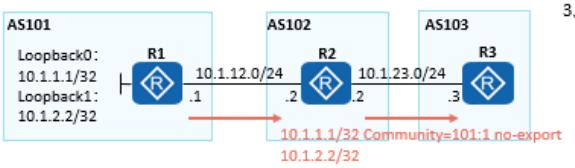
在R1上部署路由策略，使其在通告BGP路由10.1.1.1/32时携带Community属性值101:1，其他路由不携带。

在R2上部署路由策略，使其在通告BGP路由10.1.1.1/32时追加Community属性no-expert。

2、R2配置允许将团体属性传给EBGP对等体R3。

```
[R2] ip ip-prefix 1 permit 10.1.1.1 32
[R2] route-policy Community permit node 10
[R2-route-policy] if-match ip-prefix 1
[R2-route-policy] apply community no-expert additive
[R2-route-policy] quit
[R2] route-policy Community permit node 20
[R2-route-policy] quit
[R2] bgp 102
[R2-bgp] peer 10.1.12.1 as-number 101
[R2-bgp] peer 10.1.23.3 as-number 102
[R2-bgp] peer 10.1.23.3 advertise-community
[R2-bgp] peer 10.1.23.3 route-policy Community export
```

配置Community属性 (3)



在R1上部署路由策略，使其在通告BGP路由10.1.1.1/32时携带Community属性值101:1，其他路由不携带。

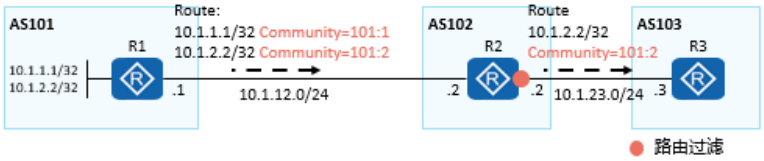
在R2上部署路由策略，使其在通告BGP路由10.1.1.1/32时追加Community属性no-expert。

3、在R3上查看BGP路由信息。

```
[R3] bgp 103
[R3-bgp] peer 10.1.23.2 as-number 102
[R3-bgp] quit

[R3] display bgp routing-table 10.1.1.1
 BGP local router ID : 10.1.23.3
 Local AS number : 103
 Paths:   1 available, 1 best, 1 select
 BGP routing table entry information of 10.1.1.1/32:
 From: 10.1.23.2 (10.1.12.2)
 Route Duration: 00h00m21s
 Direct Out-interface: GigabitEthernet0/0/2
 Original nexthop: 10.1.23.2
 Qos information : 0x0
 Community:<101:1>, no-export
 AS-path 102 101, origin igp, pref-val 0, valid, external, best,
select, active, pre 255
 Not advertised to any peer yet
```

在R3上查看10.1.1.1/32路由，发现携带101:1 no-export两个团体属性

配置Community Filter (1)



R2传递路由给EBGP对等体R3，在R2上部署路由策略，过滤掉携带101:1的Community属性值的路由。

1、配置Community Filter，匹配Community中包含101:1的路由。

```
[R2] ip community-filter 1 permit 101:1
```

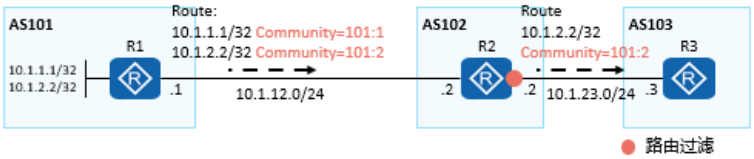
2、调用Community Filter。

```
[R2] route-policy Community deny node 10
[R2-route-policy] if-match community-filter 1
[R2-route-policy] quit
[R2] route-policy Community permit node 20
[R2-route-policy] quit
[R2] bgp 102
[R2-bgp] peer 10.1.23.3 route-policy Community export
```

3、查看R2的Community Filter信息。

```
[R2]display ip community-filter 1
Community filter Number: 1
    permit 101:1
```

配置Community Filter (2)



R2传递路由给EBGP对等体R3，在R2上部署路由策略，过滤掉携带101:1的Community属性值的路由。

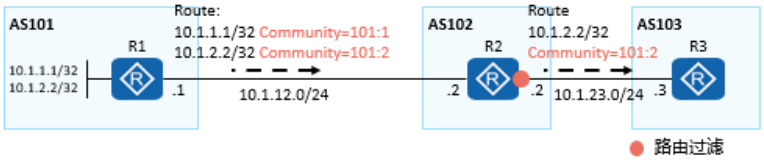
4、查看R2的BGP路由信息。

```
[R2]dis bgp routing-table 10.1.1.1
BGP local router ID : 10.1.12.2
Local AS number : 102
Paths:   1 available, 1 best, 1 select
BGP routing table entry information of 10.1.1.1/32:
From: 10.1.12.1 (10.1.1.1)
Route Duration: 00h13m39s
Direct Out-interface: GigabitEthernet0/0/1
Original nexthop: 10.1.12.1
Qos information : 0x0
Community:<101:1>
AS-path 101, origin igp, MED 0, pref-val 0, valid, external, best, select, active, pre 255
Not advertised to any peer yet
```

在R2上查看10.1.1.1/32路由，发现携带101:1团体属性。

通过Community Filter对其进行过滤。

配置Community Filter (3)



R2传递路由给EBGP对等体R3，在R2上部署路由策略，过滤掉携带101:1的Community属性值的路由。

5、查看R3的路由表项。

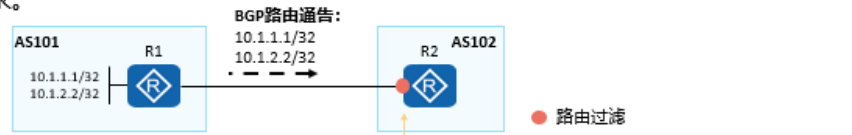
```
[R3]display bgp routing-table
 BGP Local router ID is 10.1.23.3
 Status codes: * - valid, > - best, d - damped,
               h - history,  i - internal, s - suppressed, S - Stale
               Origin : i - IGP, e - EGP, ? - incomplete
 Total Number of Routes: 1
      Network      NextHop      MED    LocPrf    PrefVal  Path/Ogn
 *>   10.1.2.2/32  10.1.23.2                     0        102 101i
```

R3接收不到10.1.1.1/32的BGP路由信息。

ORF 对等体组 安全特性

邻居按需发布路由

- 如果设备希望只接收自己需要的路由，但对端设备又无法针对每个与它连接的设备维护不同的出口策略。此时，可以通过配置BGP基于前缀的ORF（Outbound Route Filters，出口路由过滤器）来满足两端设备的需求。



R2在接收口进行路由过滤，仅接收路由10.1.1.1/32。

对于被过滤掉的路由（如：10.1.2.2/32）而言，实际上R1没有必要通告给R2。

- BGP基于前缀的ORF能力，能将本端设备配置的基于前缀的入口策略通过路由刷新报文发送给BGP邻居。BGP邻居根据这些策略（刷新报文中）构造出口策略，在路由发送时对路由进行过滤。
- 这样不仅避免了本端设备接收大量无用的路由，降低了本端设备的CPU使用率，还有效减少了BGP邻居的配置工作，降低了链路带宽的占用率。

ORF 对等体组 安全特性

ORF的基础配置命令

1. 配置对等体（组）基于IP地址前缀列表的路由过滤策略。

```
[Huawei-bgp-af-ipv4] peer { group-name | ipv4-address } ip-prefix ip-prefix-name { import | export }
```

2. 使能BGP对等体（组）基于地址前缀的ORF功能。

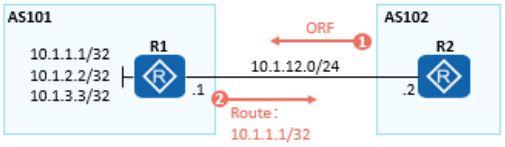
```
[Huawei-bgp] peer { group-name | ipv4-address } capability-advertise orf [ non-standard-compatible ] ip-prefix { both | receive | send } [ standard-match ]
```

注意：ORF功能需要对等体两端同时开启。

- 命令：[Huawei-bgp-af-ipv4] **peer** { *group-name* | *ipv4-address* } **ip-prefix** *ip-prefix-name* { **import** | **export** }
- **import**：对由指定对等体（组）接收的路由应用过滤策略。
- **export**：对向指定对等体（组）发送的路由应用过滤策略。
- 命令：[Huawei-bgp] **peer** { *group-name* | *ipv4-address* } **capability-advertise orf** [**non-standard-compatible**] **ip-prefix** { **both** | **receive** | **send** } [**standard-match**]
- **non-standard-compatible**：指定与非标准设备兼容。
- **both**：表示允许发送和接收 ORF 报文。

- **receive**：表示只允许接收 ORF 报文。
- **send**：表示只允许发送 ORF 报文。
- **standard-match**：指定按照 RFC 标准规定的前缀匹配规则来匹配路由。

ORF配置举例



1、R2上配置基于10.1.1.1/32的路由过滤策略，并使能ORF功能，发送ORF报文。

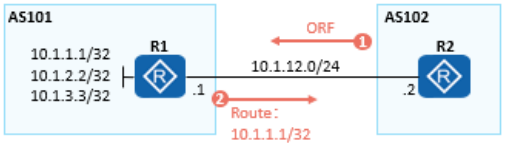
```
[R2] ip ip-prefix 1 permit 10.1.1.1 32
[R2] bgp 102
[R2-bgp] peer 10.1.12.1 as-number 101
[R2-bgp] peer 10.1.12.1 ip-prefix 1 import
[R2-bgp] peer 10.1.12.1 capability-advertise orf ip-prefix send
```

R2只期望R1通告10.1.1.1/32路由，R2通过向R1推送ORF报文来达到这个目的。

2、R1上使能ORF功能，接收ORF报文。

```
[R1] bgp 101
[R1-bgp] peer 10.1.12.2 as-number 102
[R1-bgp] peer 10.1.12.2 capability-advertise orf ip-prefix receive
[R1-bgp] network 10.1.1.1 32
[R1-bgp] network 10.1.2.2 32
[R1-bgp] network 10.1.3.3 32
```


查看ORF配置



1、在R1上查看R2发出的基于地址前缀的ORF信息。

```
[R1]display bgp peer 10.1.12.2 orf ip-prefix
Total number of ip-prefix received: 1
Index       Action      Prefix      MaskLen  MinLen  MaxLen
10          Permit      10.1.1.1    32
```

2、在R2上查看BGP路由信息。

```
[R2]display bgp routing-table peer 10.1.12.1 received-routes

BGP Local router ID is 10.1.12.2
Status codes: * - valid, > - best, d - damped,
              h - history,  i - internal, s - suppressed, S - Stale
              Origin : i - IGP, e - EGP, ? - incomplete

Total Number of Routes: 1
   Network     NextHop     MED     LocPrf    PrefVal  Path/Ogn
*> 10.1.1.1/32 10.1.12.1   0                 0        101i
```

在R1上查看ORF信息，发现需要提供10.1.1.1/32路由。

在R2上查看BGP路由信息，发现只接收到了10.1.1.1/32路由。

BGP对等体组

- 对等体组（Peer Group）是一些具有某些相同策略的对等体的集合。当一个对等体加入对等体组中时，该对等体将获得与所在对等体组相同的配置。当对等体组的配置改变时，组内成员的配置也相应改变。
- 在大型BGP网络中，对等体的数量会很多，其中很多对等体具有相同的策略，在配置时会重复使用一些命令，利用对等体组可以简化配置。

- 对等体组中的单个对等体也可以配置自己的发布路由与接收路由的策略。

BGP对等体组的基础配置命令

1. 创建对等体组。

   [Huawei-bgp] **group** *group-name* [**external** | **internal**]

   在BGP视图、BGP-VPN实例IPv4地址族视图、BGP-VPN实例IPv6地址族视图下创建对等体组。

2. （可选）为指定的对等体组配置AS号。

   [Huawei-bgp] **peer** *group-name* **as-number** { *as-number-plain* | *as-number-dot* }

   配置EBGP对等体组的AS号。对于IBGP对等体组，AS号为IBGP的本地AS号。

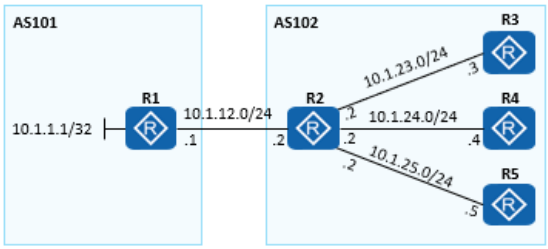
3. 将对等体加入对等体组。

   [Huawei-bgp] **peer** { *ipv4-address* | *ipv6-address* } **group** *group-name*

4. 指定发送BGP报文的源接口，并指定发起连接时使用的源地址。

   [Huawei-bgp] **peer** *group-name* **connect-interface** *interface-type interface-number* [*ipv4-source-address*]

BGP对等体组配置举例



R2作为ASBR从EBGP邻居R1收到路由后，会向所有IBGP邻居（R3、R4、R5）发送。如果R2支持BGP对等体组功能，那么它的BGP转发性能将得到较大提升。

设备	接口	接口地址	设备	接口	接口地址
R1	Loopback0	10.1.1.1/32	R4	Loopback0	10.1.4.4/32
R2	Loopback0	10.1.2.2/32	R5	Loopback0	10.1.5.5/32
R3	Loopback0	10.1.3.3/32	/		

1、完成R1的BGP基本配置。

```
[R1] bgp 101
[R1-bgp] peer 10.1.12.2 as-number 102
[R1-bgp] network 10.1.1.1 32
```

2、完成R2的EBGP基本配置和IBGP对等体组基本配置。

```
[R2] bgp 102
[R2-bgp] peer 10.1.12.1 as-number 101
[R2-bgp] group in internal
[R2-bgp] peer 10.1.3.3 group in
[R2-bgp] peer 10.1.4.4 group in
[R2-bgp] peer 10.1.5.5 group in
[R2-bgp] peer in connect-interface Loopback 0
```

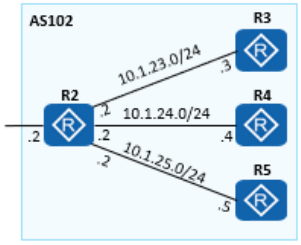
3、完成R3的IBGP基本配置。

```
[R3] bgp 102
[R3-bgp] peer 10.1.2.2 as-number 102
[R3-bgp] peer 10.1.2.2 connect-interface Loopback 0
```

R4与R5的配置与R3相似，不再赘述。

- 如图所示：假设 AS102 内，通过静态路由或 OSPF 实现 AS102 内部网络可达，配置省略。

查看BGP对等体组配置



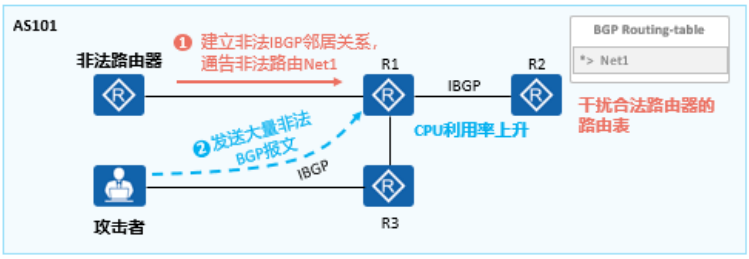
1、查看R2的BGP对等体组信息。

```
[R2]display bgp group in
 BGP peer-group: in
 Remote AS: 102
 Authentication type configured: None
 Type : internal
 Configured hold timer value: 180
 Keepalive timer value: 60
 Connect-retry timer value: 32
 Minimum route advertisement interval is 15 seconds
 Connect-interface has been configured
 PeerSession Members:
  10.1.3.3        10.1.4.4        10.1.5.5

 Peer Preferred Value: 0
 No routing policy is configured
 Peer Members:
 Peer      V  AS  MsgRcvd  MsgSent  OutQ  Up/Down  State       PrefRcv
 10.1.3.3  4  102     5        7       0    00:03:33  Established   0
 10.1.4.4  4  102     5        6       0    00:03:11  Established   0
 10.1.5.5  4  102     4        6       0    00:02:52  Established   0
```

可以看出，R2的对等体组名为in，对等体组中有三个IBGP成员，分别为：10.1.3.3、10.1.4.4、10.1.5.5，并且都成功建立了邻居关系。

BGP安全性

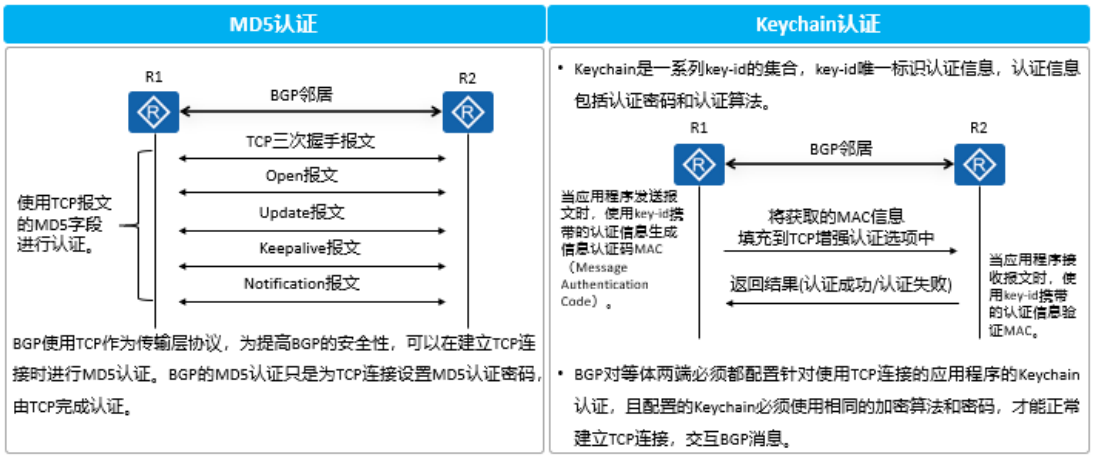
- 常见BGP攻击主要有两种：
  - 建立非法BGP邻居关系，通告非法路由条目，干扰正常路由表。
  - 发送大量非法BGP报文，路由器收到后上送CPU，导致CPU利用率升高



- BGP使用认证和GTSM（Generalized TTL Security Mechanism，通用TTL安全保护机制）两个方法保证BGP对等体间的交互安全。

BGP认证

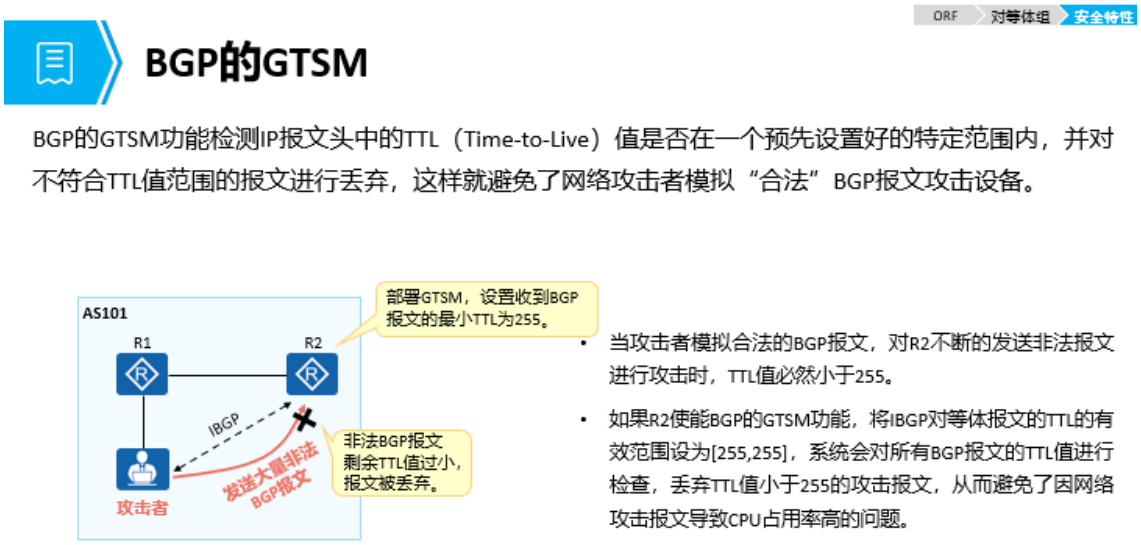
BGP认证分为MD5认证和Keychain认证，对BGP对等体关系进行认证可以预防非法BGP邻居建立。



- BGP 使用 TCP 作为传输协议，只要 TCP 数据包的源地址、目的地址、源端口、目的端口和 TCP 序号是正确的，BGP 就会认为这个数据包有效，但数据包的大部分参数对于攻击者来说是不难获得的。为了保证 BGP 免受攻击，可以在 BGP 邻居之间使用 MD5 认证或者 Keychain 认证来降低被攻击的可能性。
- MD5 算法配置简单，配置后生成单一密码，需要人为干预才可以更换密码。
- Keychain 具有一组密码，可以根据配置自动切换，但是

配置过程较为复杂，适用于对安全性能要求比较高的网络。

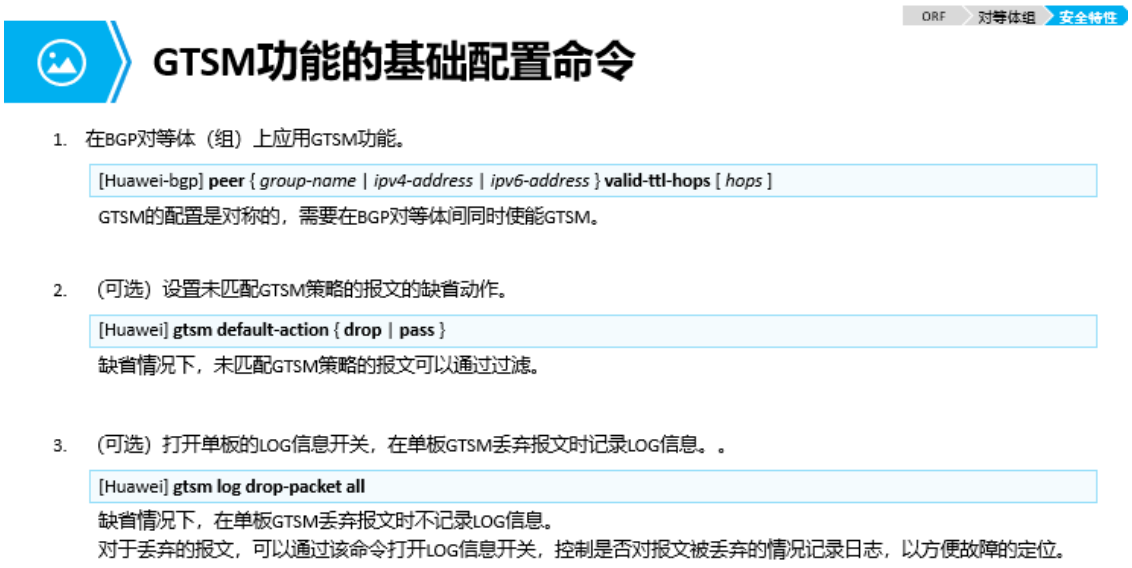
- 注意：BGP 的 MD5 认证与 BGP 的 Keychain 认证互斥。



- 如图所示：如果没有使能 BGP 的 GTSM 功能，设备收到大量非法 BGP 报文后，发现是发送给本机的报文，会直接上送控制层面处理。这时将会因为控制层面处理大量攻击报文，导致设备 CPU 占用率高，系统异常繁忙。

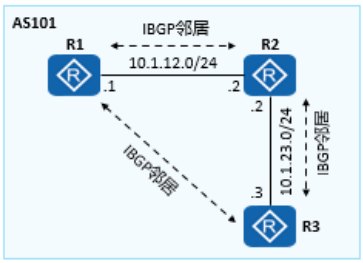
- 命令：[Huawei-bgp] **peer** { *group-name* | *ipv4-address* | *ipv6-address* } **keychain** *keychain-name*
- *keychain-name*：指定 Keychain 名称。字符串形式，长度范围是 1~47，不区分大小写。字符不包括问号和空格，但

是当输入的字符串两端使用双引号时，可在字符串中输入空格。



- 命令：[Huawei-bgp] **peer** { *group-name* | *ipv4-address* | *ipv6-address* } **valid-ttl-hops** [*hops*]
- *hops*：指定需要检测的 TTL 跳数值。整数形式，取值范围是 1~255，缺省值是 255。如果配置为 hops，则被检测的报文的 TTL 值有效范围为[255–hops+1, 255]
- 命令：[Huawei] **gtsm default-action** { **drop** | **pass** }
- **drop**：未匹配 GTSM 策略的报文不能通过过滤，报文被丢弃。
- **pass**：未匹配 GTSM 策略的报文可以通过过滤。
- 命令：[Huawei] **gtsm log drop-packet all**
- **all**：所有单板。

GTSM配置举例 (1)



R1、R2、R3均处于AS101中，且使用Loopback0接口建立IBGP全互联。同时需要在三台路由器上开启GTSM功能，防止CPU类型的攻击。

设备	接口	接口地址
R1	Loopback0	10.1.1.1/32
R2	Loopback0	10.1.2.2/32
R3	Loopback0	10.1.3.3/32

1、建立IBGP全互联。

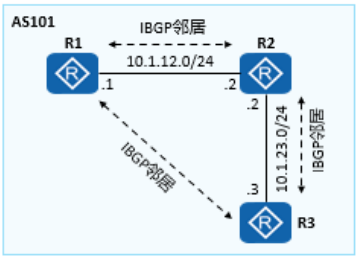
```
[R1] bgp 101
[R1-bgp] peer 10.1.2.2 as-number 101
[R1-bgp] peer 10.1.2.2 connect-interface Loopback 0
[R1-bgp] peer 10.1.3.3 as-number 101
[R1-bgp] peer 10.1.3.3 connect-interface Loopback 0
[R1-bgp] network 10.1.1.1 32
```

```
[R2] bgp 101
[R2-bgp] peer 10.1.1.1 as-number 101
[R2-bgp] peer 10.1.1.1 connect-interface Loopback 0
[R2-bgp] peer 10.1.3.3 as-number 101
[R2-bgp] peer 10.1.3.3 connect-interface Loopback 0
```

```
[R3] bgp 101
[R3-bgp] peer 10.1.1.1 as-number 101
[R3-bgp] peer 10.1.1.1 connect-interface Loopback 0
[R3-bgp] peer 10.1.2.2 as-number 101
[R3-bgp] peer 10.1.2.2 connect-interface Loopback 0
```

- 如图所示：
- 假设 AS101 内，通过静态路由或 OSPF 实现 AS101 内部网络可达，配置省略。
- R1 通告 Loopback0 接口地址到 BGP 中。

GTSM配置举例 (2)



R1、R2、R3均处于AS101中，且使用Loopback0接口建立IBGP全互联。同时需要在三台路由器上开启GTSM功能，防止CPU类型的攻击。

设备	接口	接口地址
R1	Loopback0	10.1.1.1/32
R2	Loopback0	10.1.2.2/32
R3	Loopback0	10.1.3.3/32

2、在R1与R2间开启GTSM。由于两台路由器直连，因此TTL到达对方的有效范围是[255, 255]，所以此处的valid-ttl-hops值取1。

```
[R1-bgp] peer 10.1.2.2 valid-ttl-hops 1
```

```
[R2-bgp] peer 10.1.1.1 valid-ttl-hops 1
```

3、在R2与R3间开启GTSM。由于两台路由器直连，因此TTL到达对方的有效范围是[255, 255]，所以此处的valid-ttl-hops值取1。

```
[R2-bgp] peer 10.1.3.3 valid-ttl-hops 1
```

```
[R3-bgp] peer 10.1.2.2 valid-ttl-hops 1
```

4、在R1与R3间开启GTSM。由于两台路由器经过R2连接，经过一跳后，TTL到达对方的有效范围是[254, 255]，所以此处的valid-ttl-hops值取2。

```
[R1-bgp] peer 10.1.3.3 valid-ttl-hops 2
```

```
[R3-bgp] peer 10.1.1.1 valid-ttl-hops 2
```

查看GTSM配置

```
[R1]display bgp peer 10.1.3.3 verbose
  BGP Peer is 10.1.3.3,  remote AS 101
  Type: IBGP link
  BGP version 4, Remote router ID 10.1.3.3
  Update-group ID: 1
  BGP current state: Established, Up for 00h02m17s
  BGP current event: KATimerExpired
  BGP last state: OpenConfirm
  BGP Peer Up count: 1
  Received total routes: 0
  Received active routes total: 0
  Advertised total routes: 1
  Port:  Local - 179              Remote - 51077
  Configured: Connect-retry Time: 32 sec
  Configured: Active Hold Time: 180 sec       Keepalive Time:60 sec
  Received  : Active Hold Time: 180 sec
  Negotiated: Active Hold Time: 180 sec       Keepalive Time:60 sec
  Peer optional capabilities:
  Peer supports bgp multi-protocol extension
  Peer supports bgp route refresh capability
  Peer supports bgp 4-byte-as capability
  Received: Total 5 messages
                          Update messages      0
                          Open messages        1
                          KeepAlive messages   3
                          Notification messages 0
                          Refresh messages     1
  Sent: Total 8 messages
                          Update messages      2
                          Open messages        2
                          KeepAlive messages   3
                          Notification messages 0
                          Refresh messages     1
  Authentication type configured: None
  Last keepalive received: 2020/06/22 17:34:13 UTC-08:00
  Last keepalive sent   : 2020/06/22 17:34:13 UTC-08:00
  Last update   sent    : 2020/06/22 17:34:02 UTC-08:00
  Minimum route advertisement interval is 15 seconds
  Optional capabilities:
  Route refresh capability has been enabled
  4-byte-as capability has been enabled
  Connect-interface has been configured
  GTSM has been enabled, valid-ttl-hops: 2
  Peer Preferred Value: 0
  Routing policy configured:
  No routing policy is configured
```

接右边

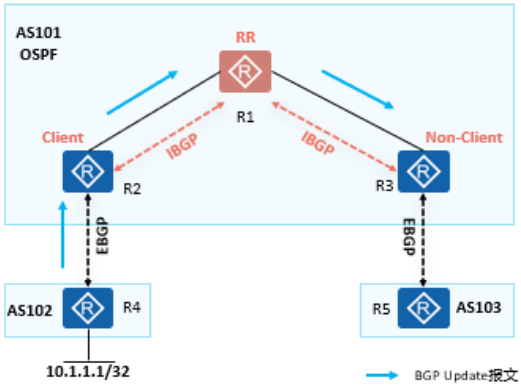
可以通过**display bgp peer**命令用来查看BGP对等体信息。

通过指定**verbose**参数，可以查看如：BGP定时器信息、收发路由数量、邻居支持的能力以及已经使能的配置等等。

R1使能了GTSM功能，且与IBGP邻居10.1.3.3（R3）间的有效跳数为2。

路由反射器

引入路由反射器，可以简化IBGP全互联的需求，也可以减轻网络和CPU的负担。



- 引入路由反射器之后存在3种角色：
  - RR（Route Reflector）：路由反射器
  - Client：客户机
  - Non-Client：非客户机
- RR会将学习的路由反射出去，从而使得IBGP路由在AS内传播无需建立IBGP全互联。
- 当RR收到对等体发来的路由，首先使用BGP选路策略来选择最佳路由。在向IBGP邻居发布学习到的路由信息时，RR会按照一定规则来发布路由。

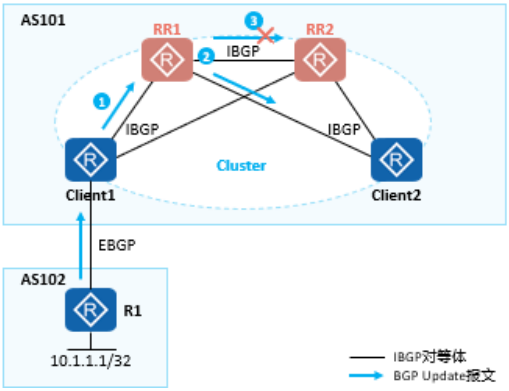
- 路由反射器相关角色：
- RR：允许把从 IBGP 对等体学到的路由反射到其他 IBGP 对等体的 BGP 设备，类似 OSPF 网络中的 DR。
- Client：与 RR 形成反射邻居关系的 IBGP 设备。在 AS 内部客户机只需要与 RR 直连。
- Non-Client：既不是 RR 也不是客户机的 IBGP 设备。在 AS 内部非客户机与 RR 之间，以及所有的非客户机之间仍然必须建立全互联关系。
- Originator（始发者）：在 AS 内部始发路由的设备。Ori

ginator_ID 属性用于防止集群内产生路由环路。

- Cluster（集群）：路由反射器及其客户机的集合。Cluster_List 属性用于防止集群间产生路由环路。
- 将一台 BGP 路由器指定为 RR 的同时，还需要指定其 Client。至于 Client 本身，无需做任何配置，它并不知晓网络中存在 RR。
- RR 发布路由规则：
- 从非客户机 IBGP 对等体学到的路由，发布给此 RR 的所有客户机。
- 从客户机学到的路由，发布给此 RR 的所有非客户机和客户机。
- 从 EBGP 对等体学到的路由，发布给所有的非客户机和客户机。

常见组网：备份RR组网

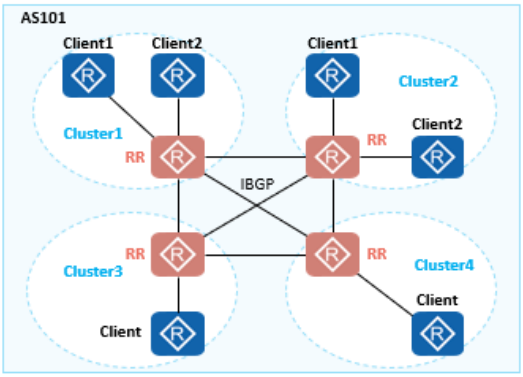
- 为增加网络的可靠性，防止单点故障对网络造成影响，有时需要在一个集群中配置一个以上的RR。
- 转发路径上的路由器与所有RR均建立IBGP关系，任意一个RR均有完整的BGP路由。



- RR1和RR2在同一个集群内，配置了相同的Cluster ID。
- 单级RR组网路由反射原理（图示以RR1的反射路径为例）：
  1. 当客户机Client1从EBGP对等体接收到一条更新路由，它将通过IBGP向RR1和RR2通告这条路由。
  2. RR1和RR2在接收到该更新路由后，将本地Cluster ID添加到Cluster List前面，然后向其他的客户机（Client2）反射，同时相互反射。
  3. RR1和RR2在接收到该反射路由后，检查Cluster List，发现自己的Cluster ID已经包含在Cluster List中。于是RR1和RR2丢弃该更新路由，从而避免了路由环路。

常见组网：多集群RR组网 (1)

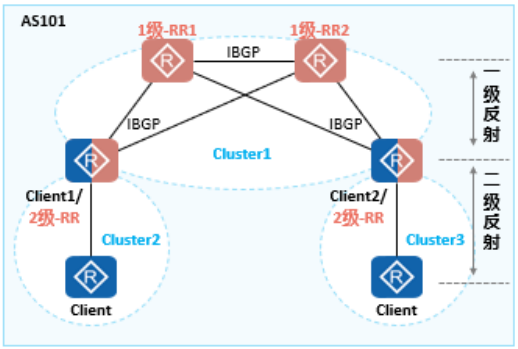
- 一个AS中可以存在多个集群，各个集群的RR之间建立IBGP对等体。
- 当RR所处的网络层相同时，可以将不同集群的RR全互联，形成同级RR。



- 一个骨干网AS可能被分成多个集群。各集群的RR互为非客户机关系，并建立IBGP全互联。
- 此时虽然每个客户机只与所在集群的RR建立IBGP连接，但所有RR和客户机都能收到全部路由信息。
- 如图所示：四个RR分别处于Cluster1、Cluster2、Cluster3、Cluster4中，它们之间互相建立IBGP连接，而每个客户机只与所在集群内的RR建立IBGP连接。

常见组网：多集群RR组网 (2)

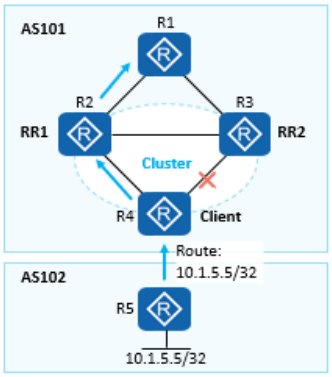
- 一个AS中可以存在多个集群，各个集群的RR之间建立IBGP对等体。
- 当RR所处的网络层不同时，可以将较低网络层次的RR配成客户机，形成分级RR。



- 在实际的RR部署中，常用的是分级RR的场景。
- 如图所示，AS101内部分为三个集群：
  - 其中Cluster1内的四台设备是核心路由器，采用备份RR的形式保证可靠性。Cluster1部署了两个1级-RR，其余两台路由器作为1级-RR的客户机。
  - Cluster2和Cluster3中分别部署了一个2级-RR，而2级-RR同时也是1级-RR的客户机。两个2级-RR之间无需建立IBGP连接关系。

单集群问题

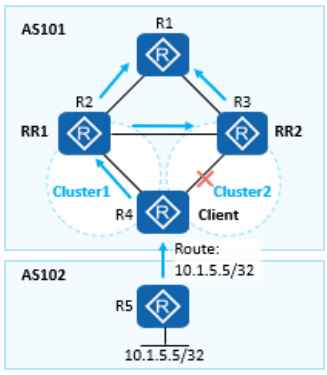
为了在基于RR的架构中提供所期望的冗余，正确的集群划分是非常重要的。



- 场景描述：
  - 如图AS101采用备份RR组网，RR1和RR2使用相同的Cluster ID，两台RR为R1访问10.1.5.5/32提供了冗余链路。
  - R4通告了10.1.5.5/32路由后，两台RR向R1通告，并相互之间通告。由于RR1和RR2有相同的Cluster ID，因此RR之间的更新消息会被丢弃。
- IBGP会话失效导致冗余失效：
  - 假设R3和R4间的IBGP会话失效（如：错误配置），由于R3忽略R2通告的10.1.5.5/32的路由，因此R1访问10.1.5.5/32就没有了冗余链路。

多集群设计

多集群设计不仅提供了针对链路失效的物理冗余，同时提供了针对客户机与RR之间的IBGP会话失效的逻辑冗余。



- 如图：将R2和R4划入Cluster1中，将R3和R4划入Cluster2中。
- 当R3和R4间的IBGP会话失效后，R3仍然可以继续转发流量，因为R3会学习R2通告的10.1.5.5/32路由。

思考题：

- （单选题）命令“ip as-path-filter 1 permit ^(100|200)$”中匹配到的 AS_Path 是什么？(　　)
- AS_Path 100
- AS_Path 200
- AS_Path 100 200
- AS_Path 100 或 AS_Path 200
- （判断题）BGP 的 GTSM 功能，可以预防非法 BGP 邻

居建立。(　　)

- 正确
- 错误
- （判断题）BGP 的备份 RR 组网中，主备 RR 收到对方反射的路由会丢弃，避免路由环路。(　　)
- 正确
- 错误

参考答案：

- D
- B
- A