

## BGP【边界网关协议】

### 一、动态路由协议的分类

- 1、内部网关路由协议【IGP】：RIP、OSPF、IS-IS
- 2、外部网关路由协议【EGP】：BGPv4

### 二、BGP的概念

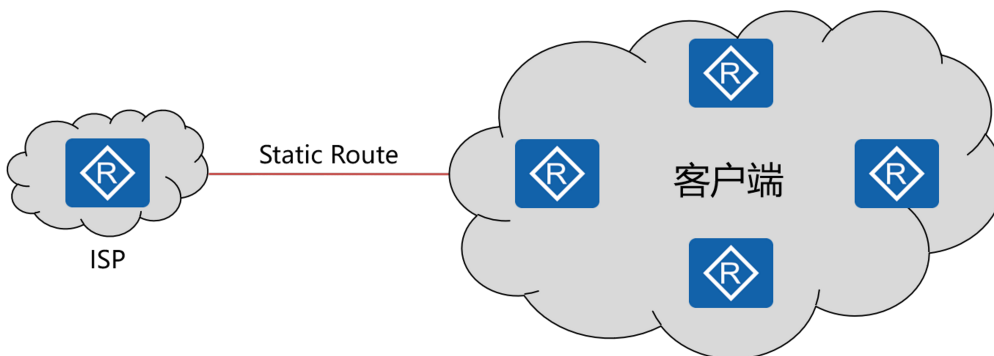
- 1、BGP属于外部网关路由协议，用来连接不同AS【自治系统】
- 2、主要用在ISP中，在各个ISP之间互联使用
- 3、亦可用在客户AS与ISP之间交换路由信息

### 三、自治系统【AS】

- 1、为解决组织与组织之间的冲突，网络划分为多个小的区域，称为自治系统【AS】
- 2、每个自治系统有一个标识号，由Internet注册机或服务提供商分配，范围1 – 65535
- 3、64512 – 65535为私有AS号码
- 4、所有的私有AS号码在外出时会被ISP进行替换

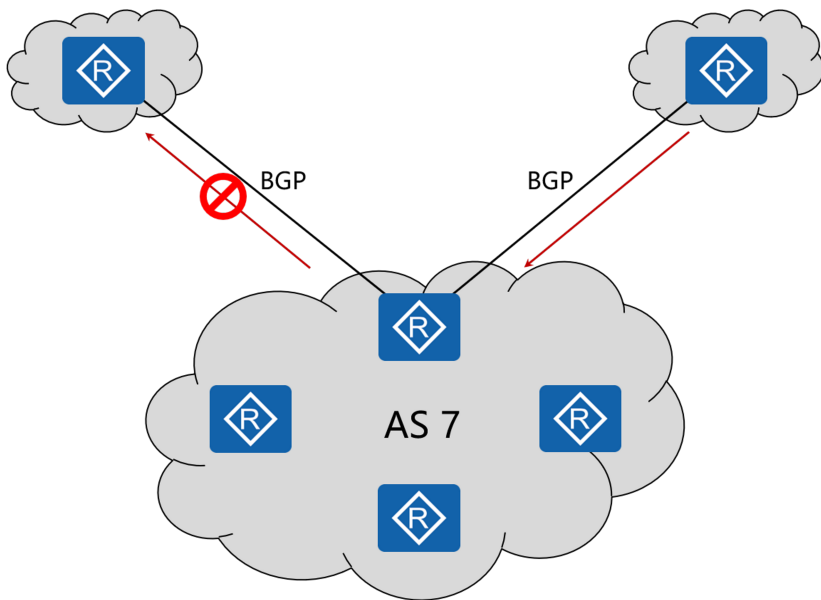
### 四、单宿主自治系统【Single Homed】

- 1、若一个AS到达外部网络有且只有一个出口时，它就被称为单宿主【Single-Homed】自治系统
- 2、该系统通常称为末节网络【Stub Network】
- 3、末节系统可以用一条缺省路由来到达外部网络
- 4、ISP向外界通告用户网络的3种方法：
  - 4.1、使用缺省配置
  - 4.2、使用IGP
  - 4.3、使用EGP



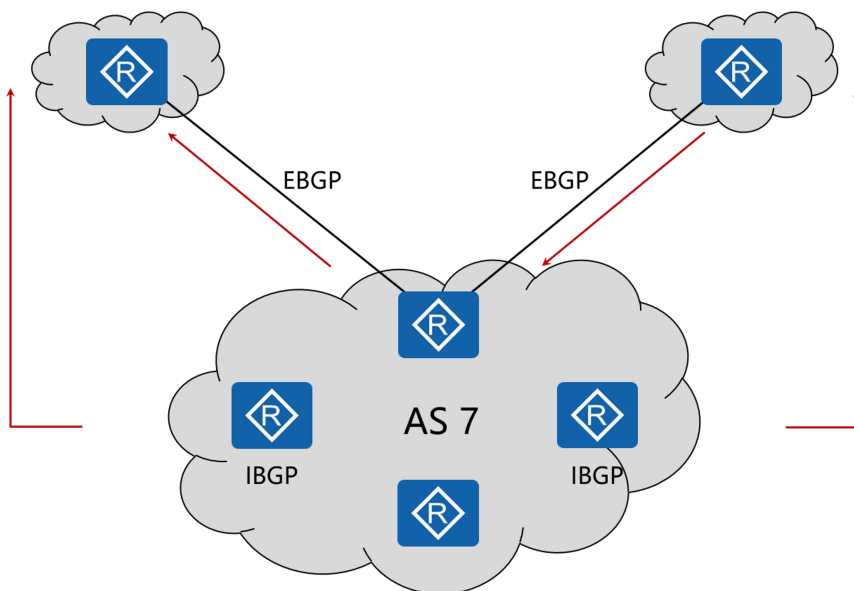
### 五、多宿主非渡越型自治系统

- 1、若一个AS到达外部网络有且不止一个出口，则称它为多宿主自治系统
- 2、连接到Internet的AS可以有到提供商的多个出口，非渡越型AS不允许渡越数据流传过它
- 3、渡越数据流是指源和目的地都在本AS之外的数据流



## 六、多宿主渡越型自治系统

- 1、若一个AS到达外部网络有且不止一个出口，则称它为多宿主自治系统
- 2、连接到Internet的AS可以有到提供商的多个出口，渡越型AS允许渡越数据流传过它
- 3、渡越数据流是指源在本AS，而目的地在本AS之外的数据流



## 七、何时不使用BGP

- 1、只有一条通往Internet或者其它AS的连接【通常使用缺省路由】
- 2、ISP的路由策略和路由选择与本AS无关
- 3、路由器的内存较低及CPU的处理能力较弱【硬件太差】
- 4、自治系统之间的链路带宽有限

## 八、BGP的路由更新

- 1、BGP用来在各个自治系统之间交换路由信息，同时确保生成无环路的拓扑结构
- 2、BGP更新由TCP协议承载，使用端口号179
- 3、RIP更新使用UDP端口号520，OSPF不使用第四层协议
- 4、BGP继承了TCP的可靠性和面向连接的特性

## 九、BGP的对等体

- 1、若两台路由器相互之间建立了一条基于TCP的BGP连接之后，就称它们为邻居或对等体
- 2、BGP对等体建立时，它们交换所有候选BGP【各自的路由条目】
- 3、交换之后，只有网络信息发生变化时才发送增量路由更新，而非整张路由表【触发更新、增量更新】
- 4、发送的增量路由更新含有路由前缀【网段及掩码】、AS路径、路径属性

- 5、BGP路由器保存着路由表版本号码，路由表发生变化，版本号增加
- 6、若没有路由变化要发送给对等体，则BGP周期性地发送存活保持【keepalive】信息来维持BGP连接
- 7、缺省情况下，19字节长的keepalive分组每隔60s发送一次

## 十、BGP的消息类型

### 1、16字节的标记字段【Flag】

用来认证进入的BGP消息或检测两个BGP对等体之间同步的丢失

### 2、2字节的长度字段

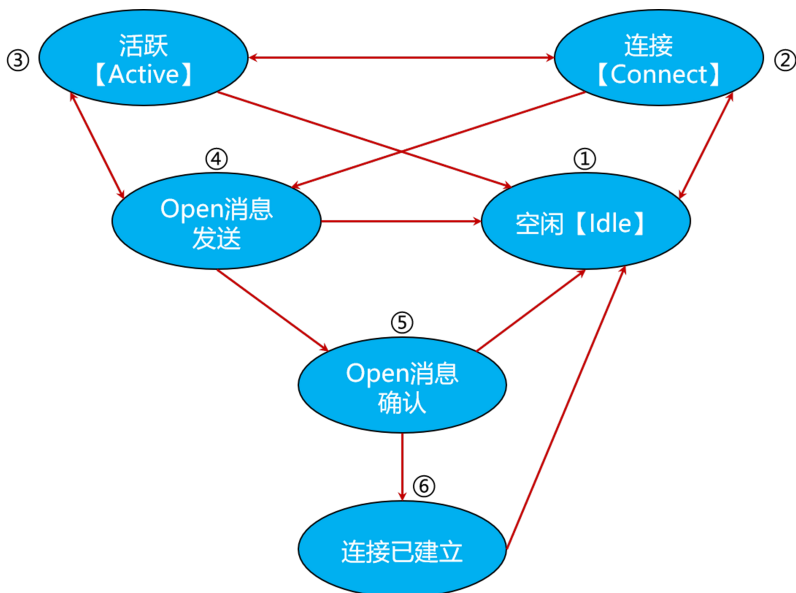
BGP消息的总长度，最小为19字节，最大为4096字节

### 3、1字节的类型字段【4种可能的状态】

- 3.1、发起连接【Open】消息：在BGP建立对等体时使用，其中包含：BGP版本号、AS号码、路由器ID等信息
- 3.2、存活保持【Keepalive】消息：在BGP稳定后，保活时使用，保持时间为3倍的Keepalive，即180s
- 3.3、通知消息：在BGP路由器发生错误时使用
- 3.4、更新消息：在BGP需要发送更新时使用

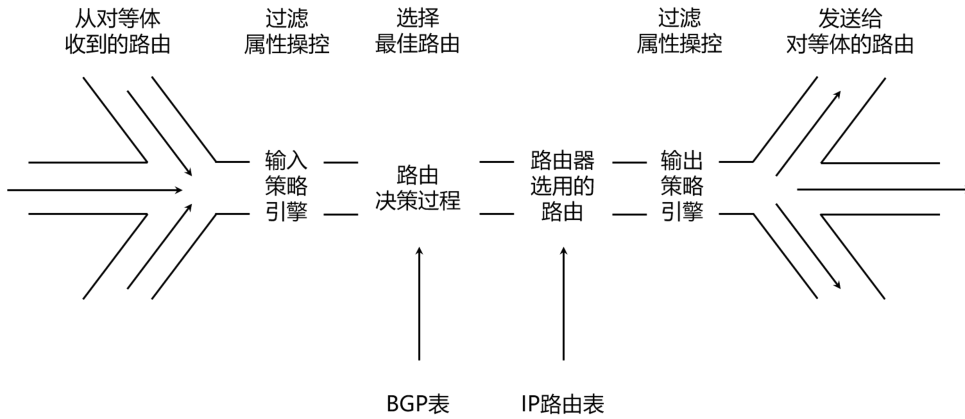
## 十一、BGP的对等体关系建立过程

- 1、空闲【Idle】：发起自己到其他BGP对等体的传输连接并侦听其他对等体所发起的连接
- 2、连接【Connect】：等待TCP连接完成，如果成功则进入Open消息发送状态，若失败则进入活跃状态，尝试再次连接，若人为手动停止，则返回空闲状态
- 3、活跃【Active】：发起TCP连接，若成功则进入Open消息发送状态，处于活跃状态的BGP仍可侦听其他对等体发起的连接，若人为手动停止，则返回空闲状态
- 4、Open消息发送【Open Sent】：BGP等待对等体发送Open消息，若检测成功，则发送keepalive消息，并采用数值最小的协商，其通过对比双方的AS号码来识别是否是同一AS，若相同，则是IBGP，若不同，则是EBGP。若失败则退回空闲状态
- 5、Open消息确认【Open Ack】：BGP等待keepalive通知，若收到则进入连接已建立阶段，若在3倍的keepalive内没有收到，则退回到空闲状态
- 6、连接已建立【Established】：此时，邻居关系建立完成，开始每60s发送keepalive消息，若人为手动停止，则返回空闲状态



## 十二、BGP的路由选择

- 1、从对等体处收到的路由
- 2、输入策略引擎【可配置进站策略】
- 3、路由决策过程【可接收路由形成BGP表】
- 4、路由器使用的路由【选择最佳路径，生成路由表】
- 5、输出策略引擎【可配置出站策略】
- 6、通告给对等体的路由



### 十三、策略路由【Policy-Based-Route | PBR】

- 1、使得网络设备不仅能够基于报文的目的IP地址进行数据转发，更能基于其他元素进行数据转发，例如源IP地址、源MAC地址、目的MAC地址、源端口号、目的端口号、VLAN-ID等等
- 2、用户还可以使用ACL匹配特定的报文，然后针对该ACL进行PBR部署
- 3、若设备部署了PBR，则被匹配的报文优先根据PBR的策略进行转发，即【PBR的优先级高于传统路由表】
- 4、策略路由 —— 策略凌驾于路由

### 十四、BGP的路由策略

- 1、通过配置策略，可影响BGP默认的选路规则
- 2、Route-Policy是一个策略工具，用于过滤路由信息，以及为过滤后的路由信息设置路由属性
- 3、配置好的策略，应用在BGP建立对等体的【import | export】方向上

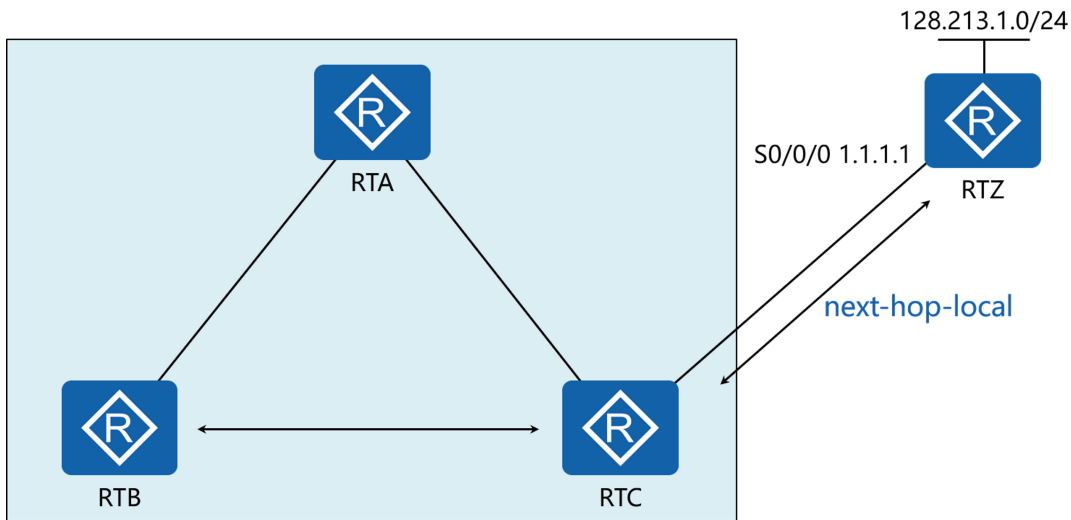
```
[Huawei]acl 2001
[Huawei-acl-basic-2001]rule permit source 192.168.1.0 0.0.0.255
[Huawei-acl-basic-2001]rule deny source any
[Huawei]route-policy atnet permit node 10
[Huawei-route-policy]if-match acl 2001
[Huawei-route-policy]apply preference 200
[Huawei-route-policy]quit
[Huawei-bgp]peer 2.2.2.2 route-policy atnet import
```

名称	操作对象	描述
路由策略 【route-policy】	路由信息	路由策略是一套用于对路由信息进行过滤、属性设置等操作的方法，通过对路由的操作或控制，来影响数据报文的转发路径
策略路由 【Policy-Based Routing   PBR】	数据报文	PBR直接对数据报文进行操作，通过多种手段匹配感兴趣的报文，然后执行丢弃或强制转发路径等操作

### 十五、BGP的属性

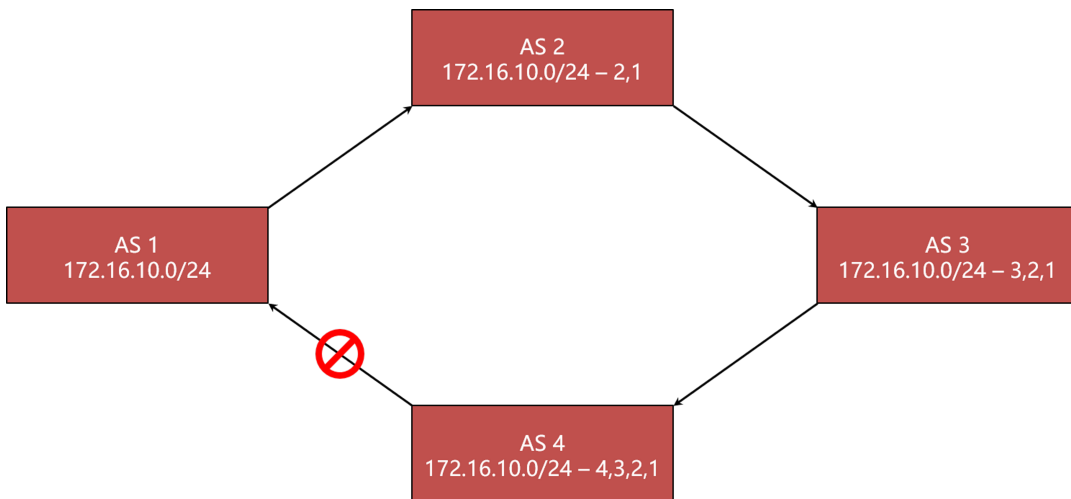
#### 1、下一跳【Next Hop】

AS的边界BGP路由器需要在与其内部路由器建立对等体关系时指定自身为它们的下一跳路由器，否则AS内部路由器将无法学习与到达AS外部

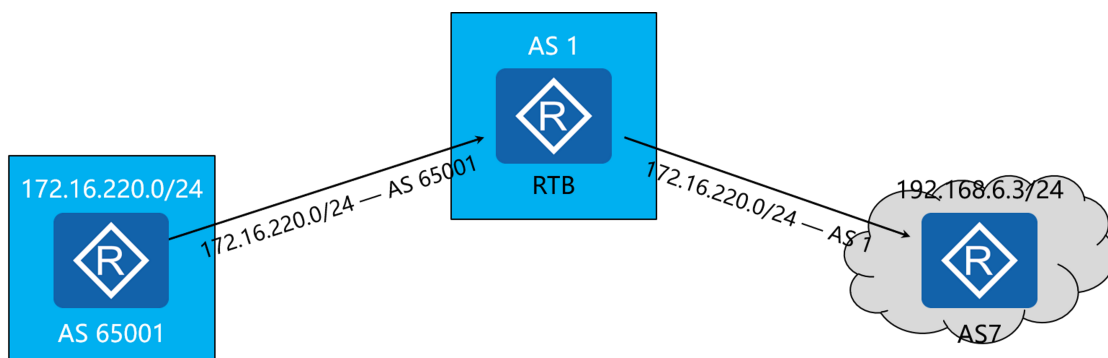


## 2、自治系统路径【AS Path】

2.1、在不同的AS之间传递路由信息，为防止环路的发生，每条路由信息都会添加其所属的AS号，或曾经经过的AS号码，通过检查AS号码，其最初的AS将不再接收本条路由



2.2、在ISP为客户网络分配私有AS号码后，客户在访问外部网络时，ISP将使用自身的公有AS号码替换掉客户的私有AS号码



```
[RTB-bgp]peer 192.168.6.3 public-as-only
```

## 3、原子汇总【Atomic Aggregate】

该属性通知BGP路由器，将多个目的地汇总到一条路由更新中【即：手工汇总】，即使有到目的地更具体的路由，也不会将这些更具体的路由发送出去

```
[Huawei-bgp]network 192.168.1.0 24
[Huawei-bgp]aggregate 192.168.0.0 16 detail-suppressed
```

## 4、汇总子【Aggregator】

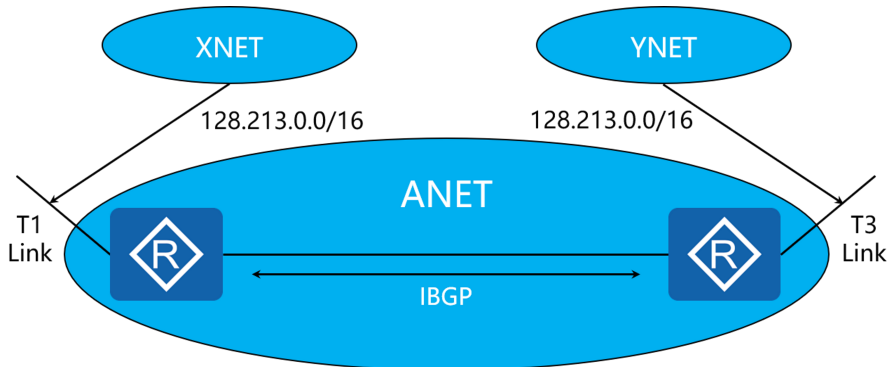
该属性使得在配置地址汇总时，也可以配置路由器在超网路由中包含它自己的路由器ID和本地AS号码。这个属性可以使得ISP管理员确定一条

特定的汇总路由是哪台路由器产生的，有助于排除故障【可与原子汇总属性一同使用】

[Huawei-bgp]aggregate 192.168.0.0 16 as-set

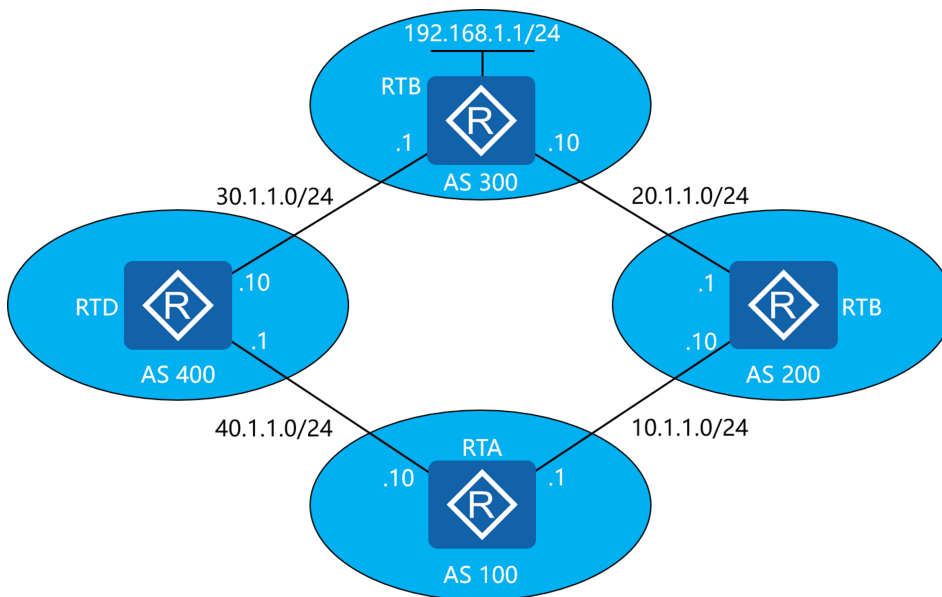
## 5、本地优先级【Local Preference】

- 5.1、本地优先级赋予一条路由的优选程度，用于比较它和其它到达同一目的地的路由
- 5.2、本地优先级只用于AS内部，只在IBGP对等体之间交换，不会被通告给EBGP对等体
- 5.3、本地优先级数值越高，路由的优选程度就越高【越高越优选】
- 5.4、需要与策略路由配合使用



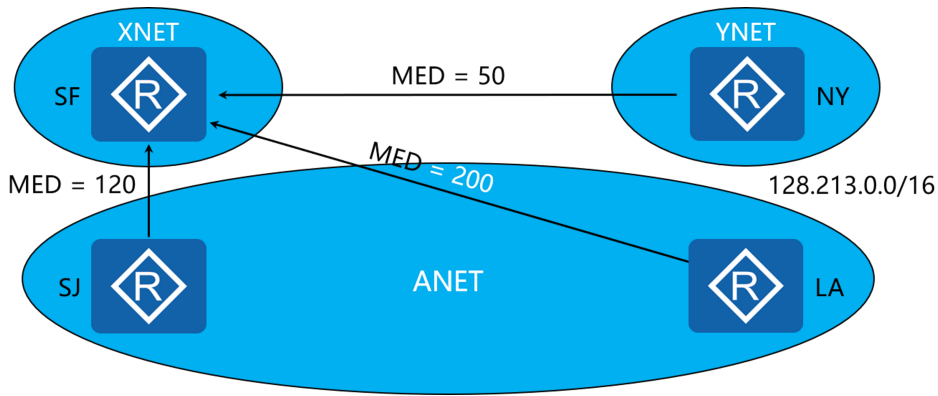
## 6、优先级值【Preference Value】

- 6.1、与本地优先级类似，不同之处在于优先级值是只对路由器本地有效，不在路由器之间交换
- 6.2、优先级值影响从不同ISP进入同一路由器的路由
- 6.3、优先级值的优先级比其它所有属性的优先级都高，是判定路由由优先级时最重要的属性
- 6.4、Huawei私有
- 6.5、需要与策略路由配合使用



## 7、多出口鉴别【Multiple Exit Discriminator, MED】

- 7.1、MED向外部邻居路由器指示进入一个有多个入口的AS的优先路径
- 7.2、路径的MED值越低，其优先级越高【越低越优选】
- 7.3、MED是在AS之间交换的
- 7.4、来自不同AS的MED不会做比较，除非使用compare-different-as-med单独设定
- 7.5、需要与策略路由配合使用



## 8、起源【Origin】

8.1、起源属性表明了BGP对路由的一种喜好程度

8.2、BGP允许下面3种类型的起源：

8.2.1、IGP —— 路由前缀在源AS内部

8.2.2、EGP —— 路由前缀是通过某种诸如BGP之类的外部网关协议【EGP】学到的

8.2.3、不完整起源 —— 路由前缀是通过其他一些方法学到的，诸如【重发布】

8.3、在BGP路由中，IBGP > EBGP > 不完整路由【重发布得到的路由】

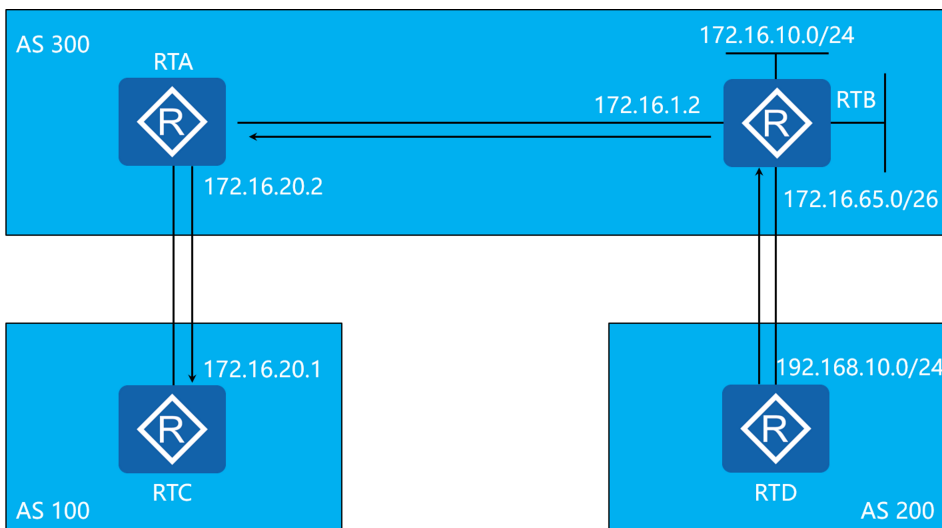
## 十六、BGP路由决策过程

- 1、如果下一跳不可达，则不考虑路由
- 2、BGP优先选用最大优先级值的路径
- 3、若优先级值相同，则选用最大本地优先级的路由
- 4、若具有相同的本地优先级，则优先选用起源于本路由器的路由
- 5、若本地优先级相同，优先选用最短AS路径的路由
- 6、若AS路径相同，则优先选用具有最低起源类型的路由
- 7、若起源相同，则选用最低MED值的路由
- 8、若以上所有值均相同，则选用Router-ID最小的路由器发布的路由的路径

## 十七、BGP的路由过滤

1、BGP可使用filter-policy过滤路由

2、eg：防止RTA将路由前缀192.168.10.0/24通告给C



```
[RTA]bgp 300
[RTA-bgp]peer 172.16.20.1 as-number 100
[RTA-bgp]peer 172.16.20.1 filter-policy 2001 export
[RTA-bgp]quit
[RTA]acl 2001
[RTA-acl-basic-2001]rule deny source 192.168.10.0 0.0.0.255
[RTA-acl-basic-2001]rule permit source 0.0.0.0 255.255.255.255
```

## 十八、使用前缀控制列表【ip ip-prefix】过滤BGP路由

1、实现高级访问控制功能

2、语法较前者简单易懂

3、关键字: less-equal、greater-equal

4、取值范围:

less-equal: 【length】至【less-equal-value】

greater-equal: 【greater-equal-value】至32

Length < greater-equal-value < less-equal-value <= 32

5、eg: 拒绝带有前缀192.0.0.0/8的、掩码长度可达24比特的路由, 并接受任何更具体的路由:

```
[Huawei]ip ip-prefix 1 deny 192.0.0.0 8 less-equal 24
```

```
[Huawei]ip ip-prefix 1 permit 192.0.0.0 8 greater-equal 25
```

```
[Huawei]bgp 1
```

```
[Huawei-bgp]peer 1.1.1.1 as-number 65001
```

```
[Huawei-bgp]peer 1.1.1.1 ip-prefix 1 export
```

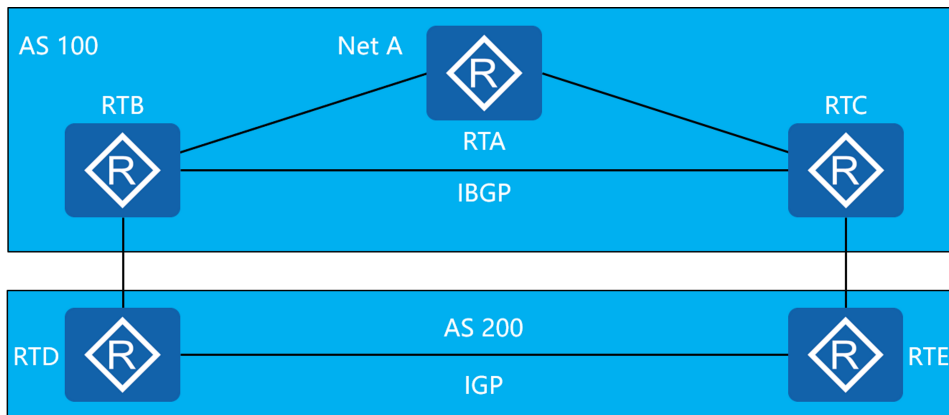
## 十九、BGP的重发布

1、路由信息可以通过重发布的形式注入到BGP中

2、重发布注入IGP路由信息可能会给其它AS带去不被采纳的、不正确的或不想要的路由信息

3、在IGP和BGP之间的相互重发布也可能导致错误路由信息的传播

4、因此Huawei不建议在IGP与BGP之间配置双向重发布



## 二十、BGP的配置

详细配置见实验手册