# 一、BGP 公开数据解析

## 1.参考资料

http://www.routeviews.org/
http://data.caida.org/datasets/routing/routeviews-prefix2as/

## 2.routeviews 数据处理

在 routeviews 上面下载的 bgp 路由信息是二进制的，MRT 格式，需要工具来解析。

**bgpdump 解析：**

1）安装 bgpdump: 在 ubuntu 18.04，用“sudo apt install bgpdump”

bgpdump 工具获得：（网上教程，实际使用时，我采用了直接 apt install）

a. 从 http://www.ris.ripe.net/source/bgpdump/下载获得压缩文件；

b. 用 tar zxvf xxx 解压缩文件；

c. 用 cd 命令进入解压缩后的文件夹；

d. 运行./configure；

e. 输入 make；

f. 将生成的可执行文件 bgpdump 文件拷贝到需要解析的文件的目录下；）

2）输入 bgpdump -m xxx.bz2(routeviews 上下载下来的文件) > xxx.txt(解析结果输出文件名)；

3）输出内容解析，bgpdump -m outputs data in the following column order:

示例：

*TABLE_DUMP|1004140086|B|209.244.2.115|3356|3.0.0.0/8|3356 701 80|IGP|*
*209.244.2.115|0|0|3356:3 3356:86 3356:575 3356:666 3356:668 3356:680 3356:2008|*
*NAG||*

各字段的含义，依次是：

- BGP Protocol
- timestamp (in epoch format)
- W/A/B (withdrawal/announcement/routing table)
  (withdrawal:BGP 退出表示先前宣布的前缀不可用。)
- Peer IP 当前 collector 所在 AS 的一个 BGP 对等体的 IP（应该是当前 collector 的一个邻居，它向这个 collecor 传播了此条路由）
- Peer ASN
- Prefix（起始 IP 段，ASPath 中最右侧 AS 号所包含的 IP 段）
- ASPath
- Origin Protocol (typically always IGP)
  路由的 Origin 属性代码。显示在每条路由的最后面。

IGP：路由在起始 AS 的内部，使用 network 命令通过 BGP 通告路由时，其 Origin 属性为 IGP，用 i 表示。

EGP：通过 EGP 得到的路由信息，其 Origin 属性为 EGP，用 e 表示。

Incomplete：表示路由的来源无法确定。BGP 通过 import-route (BGP)命令引入的路由，其 Origin 属性为 Incomplete。

- Next Hop 当前 collector 所在 AS 要去往 prefix 的下一跳，应该就是 Peer IP
- LocalPref
  本地优先级
- MED
  （BGP 路由的 MED 度量值，作用类似于 IGP 路由的 Cost，也称为 Metric）
- Community strings
- Atomic Aggregator
  NAG 没自动聚合，AG 自动聚合
- Aggregator

**zebra 解析：**

https://blog.csdn.net/chrissata/article/details/10005071

1）进入 zebra 所在目录

2）解压：bzip2 -d xxx.bz2

3）cat xxx | ./zebra-dump-parser.pl > rib.20160628.2200.txt

**公开可访问 BGP 路由器：**

有一些公开可访问的路由器，可以通过 telnet 登陆来观察一些 BGP 表信息。

查看 bgp RIB 表：show ip bgp

BGP RIB 会列出到达某一网络的所有可用路由，其中带了“>”标识的是当前路由器使用的路由.

https://blog.csdn.net/achejq/article/details/12424537?utm_medium=distribute.pc_relevant.none-task-blog-BlogCommendFromMachineLearnPai2-3.channel_param&depth_1-utm_source=distribute.pc_relevant.none-task-blog-BlogCommendFromMachineLearnPai2-3.channel_param

# 二、BGP 协议基础知识

## 1. BGP 消息类型

1）Open

建立一个 TCP 连接后，双方互相发送一个 open 消息，用于标识自己的一些参数。

2）Keepalive

3）Update

4）Notification（用于差错通知）

5）Route-refresh

## 2.BGP 报文交互中分为 Speaker 和 Peer 两种角色

**Speaker：**发送 BGP 报文的设备称为 BGP 发言者（Speaker），它接收或产生新的报文信息，并发布（Advertise）给其它 BGP Speaker。

**Peer：**相互交换报文的 Speaker 之间互称对等体（Peer）。若干相关的对等体可以构成对等体组（Peer Group）。

## 3. BGP 的三张表

1）**Neighbor table** 邻居表
2）**BGP table** 保存从邻居学到的路由信息
(我们在 routeviews 中可以看到的 mrt 信息，应该是 BGP table 的信息)
3）**BGP route table** 路由表，从 BGP table 中挑选出的到达各目标网络的最优路由。

## 4. MRT 格式

Multi-Threaded Routing Toolkit，用于导出路由协议消息、状态变化和路由信息库内容。

https://tools.ietf.org/html/rfc6396

## 5.BGP 属性

### 1）Origin 属性

IGP、EGP、Incomplete

### 2）AS_Path 属性

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。AS_Path 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 A 结束。

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。P18

当 BGP Speaker 传播**自身引入的路由**时：

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 传播**从其他 BGP Speaker 的 Update 报文中学习到的路由**时：

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。

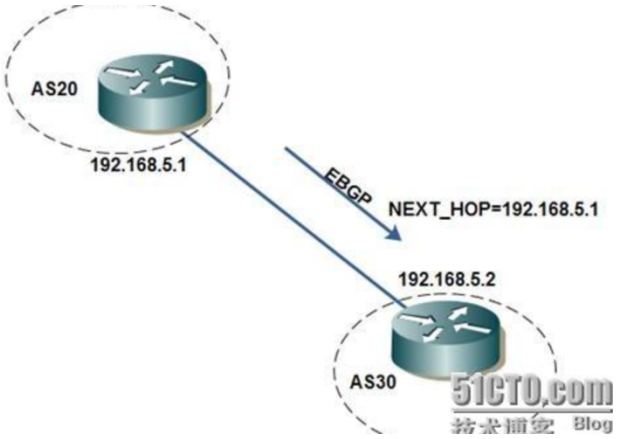
当 BGP Speaker 将这条路由通告给 IBGP 对等体时，不会改变这条路由相关的 AS_Path 属性。

**3）Next_hop 属性**

记录了路由的下一跳信息，不一定就是邻居设备的 IP 地址。

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。

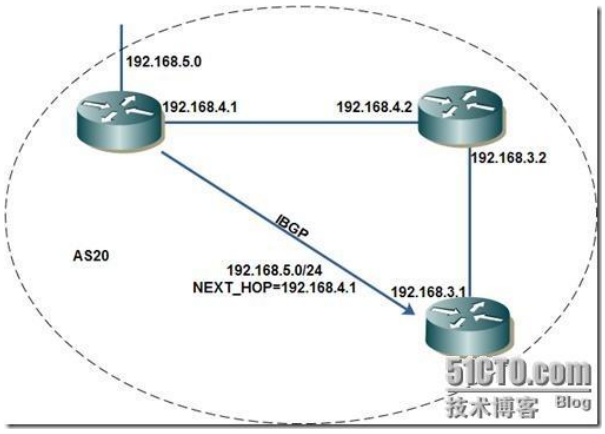
如果正在进行路由宣告的路由器和接收的路由器在不同的自治系统中，NEXT_HOP 是正在宣告路由器接口的 IP 地址：




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。

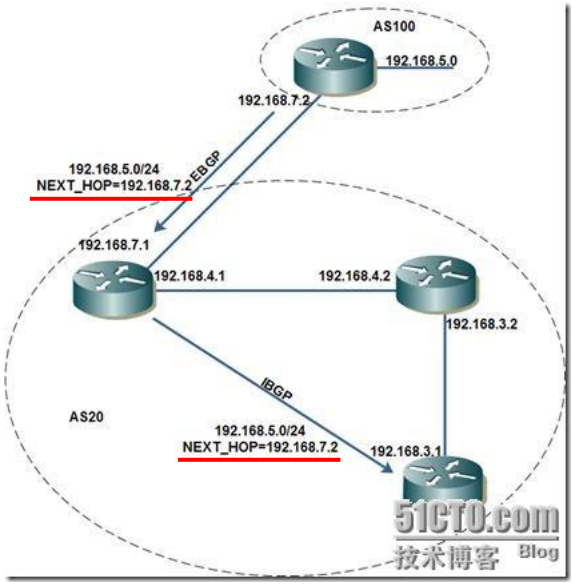
正在进行路由宣告的路由器和接收的路由器在同一个 AS 内，并且更新消息的 NLRI 指明的目的地也在同一个 AS 内，那么 NEXT_HOP 就是宣告路由的邻居的 IP 地址：

（192.168.5.0 和图中 IP 都在同一个 AS20，这时候 4.1 向 3.1 宣告路由，就会告诉他，下一跳是我 4.1；3.1 向 3.2 宣告，则会把下一跳设置为 3.1）




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。

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。




**4）Local_pref 属性**

表明路由器的 BGP 优先级，**用于判断流量离开 AS 时的最佳路由**。

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。Local_Pref 属性仅在 IBGP 对等体之间有效，不通告给其他 AS

**5）MED 属性**

Multi-Exit Discriminator 属性用于**判断流量进入 AS 时的最佳路由**

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

**6）团体属性**

## 6.路由选择策略

**当到达同一目的地存在多条路由时，BGP 选择路由的策略，依次如下：**

1）优选协议首选值（PrefVal）最高的路由。

协议首选值（PrefVal）是华为设备的特有属性，该属性仅在本地有效。

2）优选本地优先级（Local_Pref）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。

3）依次优选手动聚合路由、自动聚合路由、network 命令引入的路由、import-route 命令引入的路由、从对等体学习的路由。

4）优选 AS 路径（AS_Path）最短的路由。

5）依次优选 Origin 类型为 IGP、EGP、Incomplete 的路由。

6）对于来自同一 AS 的路由，优选 MED 值最低的路由。

7）依次优选 EBGP 路由、IBGP 路由、LocalCross 路由、RemoteCross 路由。PE 上某个 VPN 实例的 VPNv4 路由的 ERT 匹配其他 VPN 实例的 IRT 后复制到该 VPN 实例，称为 LocalCross；从远端 PE 学习到的 VPNv4 路由的 ERT 匹配某个 VPN 实例的 IRT 后复制到该 VPN 实例，称为 RemoteCross。

8）优选到 BGP 下一跳 IGP 度量值（metric）最小的路由。

9）优选 Cluster_List 最短的路由。

10）优选 Router ID 最小的设备发布的路由。

11）优选从具有最小 IP Address 的对等体学来的路由。