

BGP 边界网关协议

IGP=====内部网关协议=====RIP。EIGRP。OSPF。ISIS

EGP=====外部网关协议=====BGP

AS=====自治系统 (0-65535--2字节/4字节) =====自己管理的系统=====一个公司

为什么划分AS? ?

1、目前单独的某个IGP协议无法承载所有路由

RIP=====最多15跳

EIGRP=====特殊的公有协议 (支持的厂商不多)

OSPF/ISIS=====公有协议=====也有上限

2.公网要尽量避免动荡 (周期更新)

EIGRP/BGP只有触发更新

对于一个企业而言, 有必要运行BGP

AS 公有AS (类似于公网IP地址) ---- 0-64512

私有AS (类似于私网IP地址) ---- 64513之后

BGP什么情况下容易见到:

1.大型企业 (总部和分部)

2.数据中心

3.ISP骨干网

=====

路由器以什么为单位来划分AS?

本身为单位仅仅只属于一个AS

=====

BGP用单播建立邻居=====TCP179协议（类比成SSH TCP22）

1.直连

2.IGP（静态，动态）

跨设备建立BGP邻居-----建议用环回口建立=====相对更稳定

用环回口建立邻居的时候需要修改连接接口=====因为默认是物理口

用环回口建立EBGP邻居的时候还需要修改连接跳数=====因为默认是1跳（EBGP多跳）=====IBGP邻居不需要考虑==跳数设置为了255

BGP邻居表 =====最终状态 Established

EBGP=====建立在两个不同AS之间的邻居关系

IBGP=====建立在相同AS之间的邻居关系

BGP路由表

```
*> 11.1.1.1/32      12.1.1.1      0              0      1i
*> 111.1.1.0/24     12.1.1.1      0              0      1i
  i 11.1.1.1/32     12.1.1.1      0      100      0      1i
  i 111.1.1.0/24    12.1.1.1      0      100      0      1i
```

i代表是通过IBGP邻居学习到的，没有则代表通过EBGP邻居学习到的

BGP中的AS可以看出是一台路由器

EBGP邻居学习到的路由都是邻居的出接口地址

IBGP邻居学习到的路由的下一跳都是不变的

怎么办? ===修改下一跳

当我有EBGP邻居的时候, 那么我就对我所有的IBGP邻居去修改下一跳为我自己 [R2-bgp]peer 44.1.1.1 next-hop-local

*>的作用: 1.将该路由放进全局路由表中

2.将该路由继续传递给下一个BGP邻居

那么影响路由打上*>的原因???

1.同步====不需要考虑, 因为默认都是关闭的
undo synchronization

2.下一跳====可达就可以打上* , >代表最优

BGP配置:

1.AS号

2.让接口发送BGP报文=====IGP (network)

=====BGP (peer华为/neighbor思科)

3.宣告路由=====IGP (network) ===只需要
network的网段包括这个接口地址即可 (不在乎掩码) =宣告的是
接口

=====BGP (network*) ===宣告的是路
由===接口是多少位掩码就必须是多少位宣告===本地路由表中
怎么体现的就怎么宣告

BGP报文

Open=====建立邻居（携带很多的能力值）
Keepalive=====维持邻居（极小报文）
=====established
Update=====传递路由=====NLRI网络层可达信息（路由）
Route-refresh==路由刷新
Notification=====出错报文=====邻居建立失败

BGP状态：

IDLE/ACTIVE=====TCP三次握手
Opensent=====发送open
Openconfirm=====发送keepalive
Establish=====邻居建立，可以交互update

BGP防环机制：

EBGP防环====当我从一个BGP邻居收到路由之后，如果该路由所携带的AS-PATH（途径AS的集合）信息中有我的AS，则我会丢弃该路由=====有没有办法不让它防环

[R5-bgp]peer 44.1.1.1 allow-as-loop 1 允许传递过来的路由中所携带的和我本地相同的AS号出现几次（最大10次）

IBGP防环=====水平分割=====从一个IBGP邻居收到的路由将不会再传递给任何一个IBGP邻居=====IBGP路由只能一跳

怎么解决=====FULL-MESH（两两建立邻居---非常繁琐）

====RR反射器====三种角色：RR反射器（谁是镜子）

RR客户端

RR非客户端

最为广泛的解决方法

RR	客户端	非客户端	EBGP
客户端	可以	可以	可以
非客户端	可以	不可以	可以
EBGP	可以	可以	可以

唯一——一个不会反射的场景：从RR非客户端收到的路由不会反射给RR非客户端

[R3-bgp]peer 23.1.1.2 reflect-client

a。这条命令配置在哪台设备，那么这台设备就是RR反射器

b。这条命令中peer了谁。那么谁就是RR客户端

c。这条命令中没有peer谁。那么谁就是RR非客户端

====联邦（相对很少用）====将多个大AS分割成小AS

BGP有很多丰富的属性====不同的作用

官方分类：

公认属性====所有厂商设备都必须识别

公认必遵====必须携带在路由信息

Total Number of Routes: 3

	Network	NextHop	MED	LocPrf	PrefVal
Path/Ogn					
*>	11.1.1.1/32	12.1.1.1	0	0	1i
*>	111.1.1.0/24	12.1.1.1	0	0	1i
*>i	155.1.1.1/32	34.1.1.4	0	100	0 1i

Origin====起源属性====这个路由是怎么被宣告进BGP的
 的====i(network) ? (重分布) ====选路==i优于?

AS_Path====描述了这条路由所经过的AS路径

100 200 300 400

===最右端的AS号 (400) 代表始发AS

===最左端的AS号 (100) 代表邻居AS

===EBGP防环

===选路===AS-path中as数量越少越优

Next_hop ===下一跳

===全0==代表自己产生

===非0==代表从邻居学习

===前提：下一跳可达

===选路：自己产生的一定优于从邻居学习的

公认任意====可以不携带

Local_Pref ===本地优先级===选路，数值越大越优

===路由信息中所携带的LP值会在出AS的时候被清除，对端AS的设备在收到该路由的时候会为其赋予本地的LP值，默认是100，可以修改

Atomic_aggregate ==原子聚合===和路由汇总有关系
可选属性====不强制所有厂商能够识别

可选传递====虽然不能识别，但是可以传递给邻居

Aggregator====聚合者===和路由汇总有关系

Community 社团属性 *****====类似于一个标记=给路由
赋予一个标记从而产生一些用处

标准社团属性

AA:NN tag

众所周知==特殊作用

internet （默认，可以将路由传递给邻居）

no-advertise （不将该路由传递给任何BGP邻居）

no-export （不将该路由传递出本AS）

no-export-subconfed (不将该路由传递出本AS--联邦中的小AS)

扩展社团属性===MPLS VPN

可选不传递===不识别，也不传递

MED ===metric/cost 开销===选路

BGP选路

IGP===开销===等价/非等价

BGP===很多因素===13条选路原则===匹配：匹配即停止

===前提：下一跳一定要可达 *

	Network	NextHop	MED	LocPrf	PrefVal	Path/Ogn
*>	33.1.1.1/32	13.1.1.3	0		0	200i
*i		12.1.1.2	0	100	0	200i

1. 优选Preference_Value/Weight值最高的路由

PV华为私有属性===默认是0，取值范围0-65535，比大本路由器有效===只能在in方向修改

Weight 思科私有属性===默认是0/32768 取值范围0-65535，比大本路由器有效===只能在in方向修改

修改方式：

1.全局修改===影响所有从该邻居过来的路由

[R1-bgp]peer 12.1.1.2 preferred-value 11

2.精确修改===只影响特定的路由

===华为：ACL（匹配路由）+Route-policy（修改值）

acl number 2000

rule 5 permit source 33.1.1.1 0

route-policy PV permit node 10 默认大号拒绝

if-match acl 2000

apply preferred-value 111

route-policy PV permit node 9999 大号放行其他路由

bgp 100

peer 12.1.1.2 route-policy PV import

===思科：ACL（匹配路由）+Route-map（修改值）

2. 优选本地优先级（Local_Preference）最高的路由。

默认是100，本AS有效，除了AS出方向不能做之外其他都可以

LP值如果显示为红空，意味着这条路由是从EBGP邻居收到的，空代表本地LP值==默认是100，可以修改

bgp 100

default local-preference 200

3. 优选手动聚合>自动聚合>network>import>从对等体学到的。

本地产生的优于从邻居学习到的

聚合（AGG）：手动聚合

[R1-bgp]aggregate 33.0.0.0 8

a. 不会抑制明细路由（通告给邻居的明细，不是本地的明细）====建议在始发设备上配置

[R1-bgp]aggregate 33.0.0.0 8 detail-suppressed 抑制明细

b. 此汇总路由下一跳为null接口（黑洞路由--防环）

c, 会回传给邻居====一般建议在始发设备上做汇总

自动聚合====只能汇总重分布路由

[R3-bgp]summary automatic

Info: Automatic summarization is valid only for the routes imported through the import-route command.

思科设备的第三条其实没有啥作用：原因是因为思科本地产生的路由的weight值是32768，从邻居学到的路由的weight是0===第

一条就比掉了

4. 优选AS_Path短的路由。

经过的AS数量越少越优

要注意人为添加的AS尽量是AS-PATH已经存在的，防止EBGP防环

5. 起源类型IGP>EGP>Incomplete。

i优于e优于?

目前思科取消掉了e i优于?

6. 对于来自同一AS的路由，优选MED值小的。

MED：仅限于相邻的AS

默认是0，比小

7. 优选从EBGP学来的路由（EBGP>IBGP）。

协议优先级 EBGP 20 IBGP 200

=====

优选AS内部IGP的Metric最小的路由。

路由的下一跳的metric

优选Cluster_List最短的路由。====RR反射的次数

优选Originator_ID最小的路由。====RR-RID

优选Router_ID最小的路由器发布的路由。

优选具有较小IP地址的邻居学来的路由。

=====路由一定不是负载的

在其中需要插入一条：负载

=====

路由控制

抓路由

ACL====抓路由====调用在其他工具下====默认都是拒绝所有
====过来流量====调用在接口下====思科默认是拒绝所有
====华为默认是放行所有
====能不能对掩码进行精确控制====比较难

10.1.1.1/25	10.1.0000 0001 . 0000 0001
10.1.2.3/26	10.1.0000 0010 . 0000 0011
10.1.3.8/27	10.1.0000 0011 . 0000 1000
10.1.4.10/28	10.1.0000 0100 . 0000 1010
10.1.1.0/21	10.1.1.0/22

prefix-list 前缀列表====只能用来匹配路由，并且可以精确匹配

掩码====也是可以被Route-policy/route-map调用的

```
[R2]ip ip-prefix 10 permit 10.1.0.0 16 greater-equal 26 less-equal 27
```

16====代表固定多少为不变====定义了前缀路由 10.1

26====代表路由的掩码最少是26位

27====代表路由的掩码最大是27位

10.1.1.1/25

10.1.2.3/26

10.1.3.8/27

10.1.4.10/28

```
ip ip-prefix 10 permit A.A.A.A X ge Y le Z
```

```
ip ip-prefix 10 permit 10.1.0.0 16 ====全部?—
```

仅仅匹配10.1.0.0/16这一条路由

```
ip ip-prefix 10 permit 10.1.0.0 16 greater-equal 26
```

匹配==26-32的路由（不写le等同于le 32）

```
ip ip-prefix 10 permit 10.1.0.0 16 less-equal 27
```

匹配==16-27 (不写ge等用于ge X)

```
R1(config)#ip prefix-list 10 permit 192.168.1.1/12 ge 15 le 17
```

BGP匹配路由

匹配 AS-PATH

ip as-path-filter 正则表达式

* 任意。所有 (0-所有)

? 0次或者1次

^ 从什么什么开始

\$ 到什么什么结束

100 匹配经过AS100的任意路由

^100\$ 匹配只有AS100的路由

100* AS中有10(0重复0次或者多次)

匹配社团属性----AA: NN

community-filter

设置行为----route-policy/route-map