策略路由与DHCP共同使用的典型配置

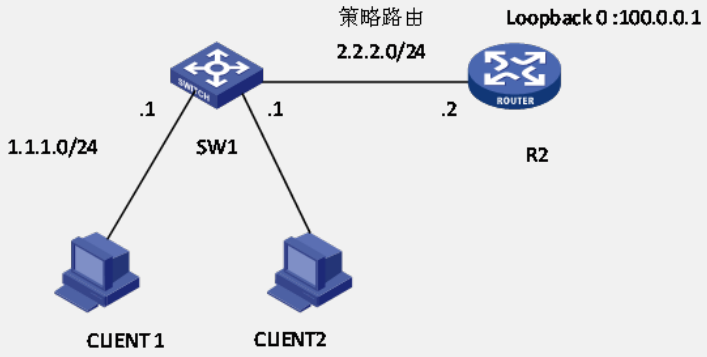
策略路由 DHCP 郭威 2014-07-21 发表

策略路由与DHCP共同使用的典型配置

一、 组网需求：

在中端交换机V5平台下（S7500E在R67XX版本之后，S10500在R12XX版本之后）的同一个三层虚接口下，既配置DHCP SERVER或者DHCP RELAY，同时又通过PBR或QOS的方式做策略路由。希望下连PC能够通过DHCP获取到地址，之后再通过PBR或者QOS做策略路由。

二、 组网图：



保证SW1与R2的连通性，其中SW1与CLIENT 之间是1.1.1.0/24网段，SW1与R2之间是2.2.2.0/24网段。R2上有一个LOOPBACK 0地址，为100.0.0.1，同时配置默认路由指回SW1。

三、 配置步骤：

下文均以DHCP SERVER为例进行分析，实际情况中DHCP RELAY也是相同支持的。

1、在SW1配置DHCP SERVER

SW1为CLIENT1和CLIENT2的网关，CLIENT 1和CLIENT 2通过DHCP获取地址，SW1做DHCP SERVER。

```
#
dhcp server ip-pool test
 network 1.1.1.0 mask 255.255.255.0
 gateway-list 1.1.1.1
#
interface Vlan-interface1
 ip address 1.1.1.1 255.255.255.0
#
```

此时CLIENT 1正常获取到了地址，在SW1上查看

```
[H3C]dis dhcp se ip al
Pool utilization: 0.39%
 IP address       Client-identifier/    Lease expiration          Type
                  Hardware address
 1.1.1.2          2c59-e501-a055        Jan  4 2015 12:41:40      Auto:COMMITTED

 --- total 1 entry ---
```

2、在SW1上配置PBR策略

```
#
[H3C]dis acl 3000
Advanced ACL  3000, named -none-, 1 rule,
ACL's step is 5
 rule 0 permit ip
#
```

```
[H3C]dis policy-based-route test
 policy-based-route : test
   Node 10 permit :
     if-match acl 3000
     apply ip-address next-hop 2.2.2.2        //通过PBR下一跳指向R2
#
```

此时从CLIENT 1 ping 100.0.0.1。由于在SW1上的路由表中没有到达100.0.0.1的路由，因此互PING能通证明流量匹配了PBR策略。

```
C:\Users\g10026>ping 100.0.0.1 -t
正在 Ping 100.0.0.1 具有 32 字节的数据:
来自 100.0.0.1 的回复: 字节=32 时间=14ms TTL=254
来自 100.0.0.1 的回复: 字节=32 时间=1ms TTL=254
来自 100.0.0.1 的回复: 字节=32 时间=1ms TTL=254
100.0.0.1 的 Ping 统计信息:
   数据包: 已发送 = 3，已接收 = 3，丢失 = 0 (0% 丢失)，
往返行程的估计时间(以毫秒为单位):
最短 = 1ms，最长 = 14ms，平均 = 5ms
```

此时发现CLIENT 1能够获取DHCP地址，且流量在SW1上匹配了PBR策略。但是这是在先获取地址后，再配置PBR策略的情况，此时不涉及DHCP地址的获取。让CLIENT重新获取地址，会发现此时无法正常获取地址。通过在CLIENT 上抓包查看，只有DHCP DISCOVER报文，而无任何的回应。

3、修改PBR策略路由

由于PBR策略的下发，会导致所有流量均重定向到R2上，从而SW1无法处理DHCP报文，因此通过修改PBR的策略，将DHCP 报文单独匹配出来。

```
#
[H3C]dis acl 3001
Advanced ACL  3001, named -none-, 1 rule,
ACL's step is 5
 rule 0 permit udp destination-port eq bootps (22 times matched)
#
[H3C]dis policy-based-route test
 policy-based-route : test
   Node 0 permit :
     if-match acl 3001
   Node 10 permit :
     if-match acl 3000
     apply ip-address next-hop 2.2.2.2
#
```

实现的重点是写一条ACL 3001，以匹配目的端口为BOOTPS的DHCP协议报文。

可以看到此时CLIENT 成功获取了地址，抓包查看报文交互完整。

7	11:06:17.445323000	DHCP	0.0.0.0	255.255.255.255	342	DHCP Discover - Transaction ID 0x74e1e5cf
8	11:06:17.982457000	DHCP	1.1.1.1	255.255.255.255	350	DHCP Offer - Transaction ID 0x74e1e5cf
9	11:06:17.983683000	DHCP	0.0.0.0	255.255.255.255	370	DHCP Request - Transaction ID 0x74e1e5cf
10	11:06:17.996310000	DHCP	1.1.1.1	255.255.255.255	350	DHCP ACK - Transaction ID 0x74e1e5cf
11	11:06:21.293217000	DHCP	1.1.1.2	255.255.255.255	342	DHCP Inform - Transaction ID 0xee6da89a
12	11:06:21.295396000	DHCP	1.1.1.1	1.1.1.2	350	DHCP ACK - Transaction ID 0xee6da89a
13	11:07:44.578848000	DHCP	1.1.1.2	255.255.255.255	342	DHCP Inform - Transaction ID 0x120a162a
14	11:07:44.760851000	DHCP	1.1.1.1	1.1.1.2	350	DHCP ACK - Transaction ID 0x120a162a

4、通过QOS策略实现需求

通过在物理接口下配置QOS策略，也可以实现需求，具体配置如下

先取消有关PBR的配置，

```
#
traffic classifier test1 operator and
 if-match acl 3001
traffic classifier test operator and
 if-match acl 3000
#
traffic behavior test1            //此处不配置动作
traffic behavior test
 redirect next-hop 2.2.2.2 fail-action forward
#
interface GigabitEthernet2/0/32
 port link-mode bridge
 qos apply policy test inbound
```

```
#
[H3C-GigabitEthernet2/0/32]dis qos po int

  Interface: GigabitEthernet2/0/32

  Direction: Inbound

  Policy: test
   Classifier: test1 (Failed)
     Operator: AND
     Rule(s) : If-match acl 3001
     Behavior: test1
      -none-
   Classifier: test
     Operator: AND
     Rule(s) : If-match acl 3000
     Behavior: test
      Redirect enable:
        Redirect type: next-hop
        Redirect destination:
          2.2.2.2
        Redirect fail-action: forward
#
```

可以看到此时CLIENT 能够正常获取地址，且匹配QOS 的策略路由。

```
C:\Users\g10026>ping 100.0.0.1

正在 Ping 100.0.0.1 具有 32 字节的数据:
来自 100.0.0.1 的回复: 字节=32 时间=8ms TTL=254
来自 100.0.0.1 的回复: 字节=32 时间=1ms TTL=254
来自 100.0.0.1 的回复: 字节=32 时间=1ms TTL=254
来自 100.0.0.1 的回复: 字节=32 时间=1ms TTL=254

100.0.0.1 的 Ping 统计信息:
    数据包: 已发送 = 4，已接收 = 4，丢失 = 0 (0% 丢失)，
往返行程的估计时间(以毫秒为单位):
最短 = 1ms，最长 = 8ms，平均 = 2ms
```

四、 配置关键点：

1、ACL中的DHCP协议端口号，如果没有bootps，则需要指定UDP端口号为67；
2、由于PBR节点的匹配顺序是从小到大依次匹配，因此需要将匹配DHCP报文的ACL所在节点放在前面，以避免该报文被其他节点所匹配。
3、QOS策略中有不同CB对时，是按照配置顺序，先下发先生效的，匹配一个就不再往下匹配，因此需要将匹配DHCP报文的ACL的CB对放在最前面。
4、DHCP RELAY与DHCP SERVER的配置相同。