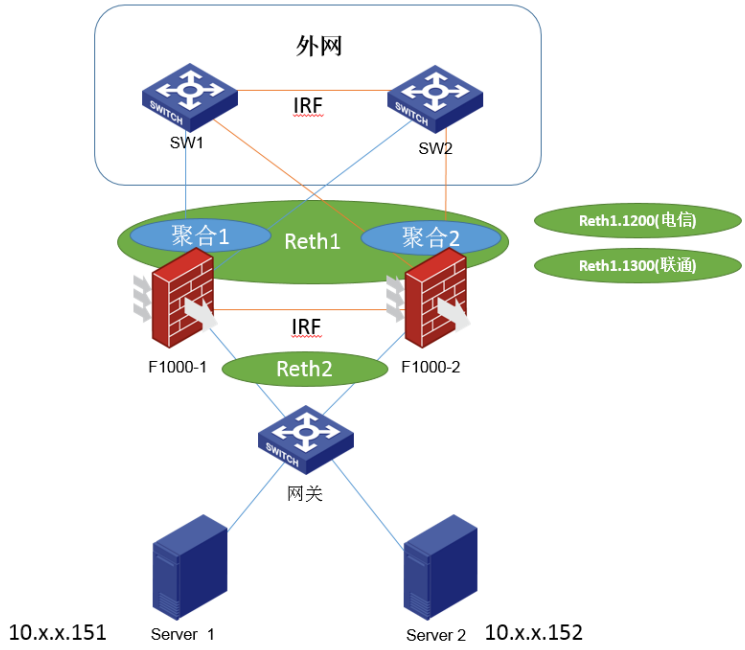
某局点内部服务器访问外网，有的能ping通，有的ping不通

NAT IRF zhiliao_HABbRf 2022-04-28 发表

组网及说明

1. 防火墙IRF堆叠，访问外网时一共有两条链路，分别是电信和联通
2. 外网设备也使用IRF，与防火墙使用二层动态聚合口相连，使用vlan1200和vlan1300为业务vlan
3. 防火墙使用冗余子接口对接，使用reth1.1200（电信）和reth1.1300（联通）
4. 防火墙配置有nat server+反向，保证内网服务器使用nat server的global地址访问，且server 1使用电信出去，server 2使用联通出去



问题描述

Server 1和Server 2 ping www.baidu.com 没有问题，ping106.15.251.68不通

过程分析

1. 查看防火墙上的nat配置，客户分别在reth1.1200和reth1.1300配置了两台服务器的nat server，没有问题

interface Reth1.1200
ip address 116.x.x.194 255.255.255.192
nat server protocol icmp global 116.x.x.197 inside 10.x.x.151 rule 1 icmp counting
#
interface Reth1.1300
ip address 220.x.x.130 255.255.255.224
nat server protocol icmp global 220.x.x.133 inside 10.x.x.152 rule 1 icmp counting

2. 经确认外网交换机的接口为二层聚合口，不是三层聚合口，故不会存在回包跨框转发的情况
3. 查看防火墙会话信息，发现两个问题：

l 防火墙有发出去的包，但没有回来的包
l 接口配置了nat，但没有转换成公网地址

display session table ipv4 source-ip 10.x.x.151 destination-ip 106.15.251.68 verbose

Initiator:
  Source   IP/port: 10.x.x.151/9714
  Destination IP/port: 106.15.251.68/2048
  DS-Lite tunnel peer: -
  VPN instance/VLAN ID/Inline ID: -/-/-
  Protocol: ICMP(1)
  Inbound interface: Reth2.1021
  Source security zone: APP
Responder:
  Source   IP/port: 106.15.251.68/9714
  Destination IP/port: 10.x.x.151/0
  DS-Lite tunnel peer: -
  VPN instance/VLAN ID/Inline ID: -/-/-
  Protocol: ICMP(1)
  Inbound interface: Reth1.1300
  Source security zone: Internet
State: ICMP_REQUEST
Application: ICMP
Rule ID: 11
Rule name:xx
Start time: 2022-04-25 10:10:56 TTL: 59s
Initiator->Responder: 114 packets   9576 bytes
Responder->Initiator: 0 packets   0 bytes

4. 分析为什么接口有配置nat，但服务器不转换成公网地址；使用debugging查看

acl ad 3000
rule 0 permit ip source ip 10.x.x.151 0 destination 106.15.251.68 0
rule 5 permit ip source ip 106.15.251.68 0 destination 10.x.x.151 0

IP packet debugging switch is on (ACL:3000)
IP info debugging switch is on (ACL:3000)
NAT packet debugging switch is on(acl:3000)

开启之后，只有debugging ip packet 的回显，其他两个没有，也证实了的确是没有进行nat转换

Receiving, interface = Reth2.1021
version = 4, headlen = 20, tos = 16
pktlen = 401, pktid = 24650, offset = 0, ttl = 63, protocol = 17
checksum = 17497, s = 10.x.x.151, d = 106.15.251.68
channelID = 0, vpn-InstanceIn = 0, vpn-InstanceOut = 0.
prompt: Receiving IP packet from interface Reth2.1021.
Payload: UDP
  source port = 7368, destination port = 5080
  checksum = 0x9b26, length = 381.

```
Sending, interface = Reth1.1300
version = 4, headlen = 20, tos = 16
pktlen = 401, pktid = 24650, offset = 0, ttl = 62, protocol = 17
checksum = 7253, s = 10.x.x.151, d = 106.x.x.68
channelID = 0, vpn-InstanceIn = 0, vpn-InstanceOut = 0.
prompt: Sending IP packet received from interface Reth2.1021 at interface Reth1.1300.
Payload: UDP
  source port = 7368, destination port = 5080
  checksum = 0x9b26, length = 381.
```

也是从reth2.1021接口接收的，这个没有问题。但这个包使用reth1.1300（联通）发出的，这个是有问题的，因为最初的需求就是10.x.x.151走电信，现在却发给了联通这条线，没有从预期接口发送，下一步就排查路由配置

```
===============display ip routing-table===============
Destination/Mask   Proto   Pre Cost       NextHop        Interface
0.0.0.0/0          Static  60  0          16.x.x.193     Reth1.1200
                                          220.x.x.129    Reth1.1300
```

可以看到这个是等价路由，这也就找到了问题的根本原因：

因为配置的等价路由，导致流量会随机分配给两个外网出口，分为两种情况：

当分配给reth1.1200后，会进行nat转换，可以成功访问外网

当分配给reth1.1300后，则直接进行转发，但源地址还是私网地址，会被对方服务器丢弃

解决方法

使用acl匹配源地址，然后使用策略路由的方式指定出接口

```
acl acl 2999
rule 0 permit ip source ip 10.x.x.151 0
acl acl 3888
rule 0 permit ip source ip 10.x.x.152 0
policy-based-route aaa permit node 10
if-match acl 2999
apply next-hop 16.x.x.193（电信下一条地址）
policy-based-route aaa permit node 20
if-match acl 3888
apply next-hop 220.x.x.129（联通下一条地址）
```