

## M-LAG典型案例及FAQ(第3期)

### H3C交换机技术专题篇

网，  
交换那些事儿M-LAG技术专题新系列  
M-LAG典型案例及FAQ  
每期推送2个案例+4个FAQ  
让你在遇到类似问题时不迷茫不发愁  
解决方案手到擒来~  
往期内容可点击上方话题复习

### 本期内容

#### 典型案例：

- Case1 M-LAG+EVPN组网leaf单挂终端无法访问外网
- Case2 M-LAG组网下挂终端ping设备上地址不通

#### FAQ:

- Q1 M-LAG+VRRP场景下流量负载分担?
- Q2 M-LAG单挂设备有没有特殊配置要求?
- Q3 M-LAG成员接口物理down是否会影响M-LAG group的状态?
- Q4 M-LAG+EVPN组网中, border设备配置group ip后为何不以虚地址发布路由?

### DRNI更名为M-LAG

原DR口>>M-LAG接口  
原DR group>>M-LAG group  
原IP链路>>peer-link链路  
原IP口>>peer-link接口  
本文涉及命令行暂不做修改

#### Case1

#### M-LAG+EVPN组网leaf单挂终端无法访问外网

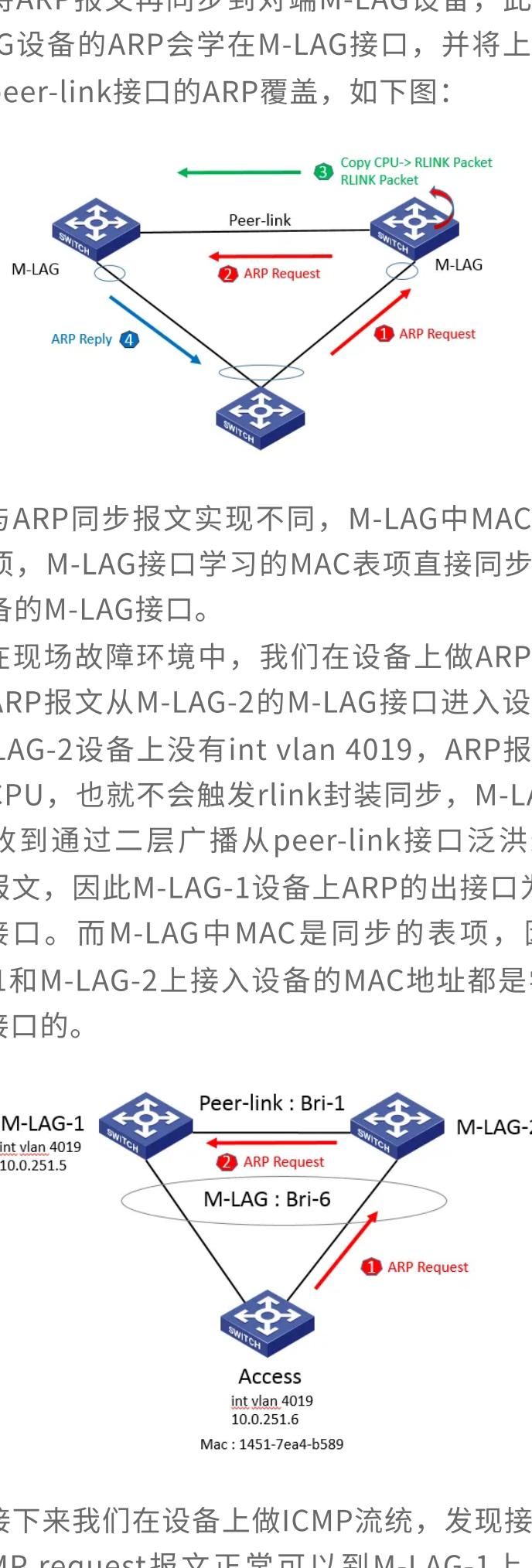
### 组/网/说/明

设备及版本：

S12508X-AF R2820

S6860-54HF F2707

组网：两台S6860起M-LAG作为一组leaf设备，上行两台S12508X-AF同样起M-LAG作为spine/border设备，leaf与spine/border之间建立隧道，leaf设备下的终端全单挂。下图为局部组网图。



### 问/题/描/述

现场M-LAG+EVPN组网，两台S6860起M-LAG作为leaf，远端两台12508X-AF起M-LAG作为Spine\border设备，测试发现leaf下单挂的终端无法访问外网，leaf设备上终端的arp、mac表项学习正常。在终端ping外网时抓包，ICMP报文从leaf设备发出，spine/border设备也收到了ICMP报文，但并未上行继续向外发出，可以确认丢包发生在spine/border上。

### 过/程/分/析

spine/border未将报文继续转发出去，我们首先怀疑从leaf发送过来的报文是否封装正确，从抓包来看，报文是正常的，VXLAN外层源自IP为隧道源目IP，内层为原始报文源目IP：

既然报文没有问题，那为什么spine/border没有正常解封装报文并转发呢？在“[交换那些事儿 | 技术专题篇 - DRNI+EVPN组网进阶介绍](#)”中我们讲解过，设备收到带封装的报文后，如果其封装的源自IP在设备上可以找到对称隧道，就可以解封装，反之则无法正常解封装并转发。我们在leaf上查看到spine/border的隧道只有如下一条，抓到的报文也是封装的这个隧道的源自IP：

**source 60.1.1.5, destination 60.1.1.251**  
抓到的报文也是封装的这个隧道的源自IP；  
在spine/border上查看到leaf的隧道有两条：  
**source 60.1.1.251, destination 60.1.1.1**  
**source 60.1.1.251, destination 60.1.1.2**

很明显spine/border上这两条隧道都不是报文封装源自IP的对称隧道，因此spine/border就是无法正常解封装并转发这个报文的。

那为什么leaf和spine/border上的隧道会不对称呢？我们进一步查看配置信息发现，60.1.1.1和60.1.1.2为两台leaf设备的loopback 0地址，60.1.1.5和60.1.1.251为leaf和spine/border设备的group ip。且由于leaf下存在单挂终端，所以leaf上配置了`evpn drni local local-ip remote remote-ip`命令；而spine/border设备上引入了5类路由，也配置了`nexthop evpn-drni group-address`命令：

**Leaf:**  
evpn drni group 60.1.1.5  
evpn drni local 60.1.1.2 remote 60.1.1.1  
spine\border:  
evpn drni group 60.1.1.251  
nexthop evpn-drni group-address

在这种配置下，由于leaf全单挂，leaf设备会以group ip发布3类路由，以local ip发布2类路由，spine/border只会通过group ip发布引入的5类路由。leaf收到spine/border发布的5类路由后，根据路由下一跳建立虚地址隧道 source 60.1.1.5, destination 60.1.1.251；spine/border上会收到leaf发布的3类和2类路由，但由于spine/border上没有VSI配置，无法接收3类路由，因此只能通过2类路由的下一跳建立到leaf实地址的隧道。

### 问/题/描/述

在spine/border上配置`vxlans default-decapsulation source interface`命令强制解封装即可解决问题。配置该命令后，spine/border收到带VXLAN封装的报文后会直接强制解封装，再走转发流程正常转发。

每多处理一个问题  
对M-LAG的理解也更多一层  
经验案例多看看  
典型配置多学习  
大家都可以说大佬~  
有问题欢迎留言哦~

按照规范配置，在两台M-LAG设备上都正常配置VLAN4019的三层虚接口即可。

**M-LAG三层网关配置及表项同步机制可参考：**

[交换那些事儿 | 技术专题篇 - DRNI+VRRP及MAC表项同步机制介绍](#)

[交换那些事儿 | 技术专题篇 - DRNI+EVPN组网ARP及MAC表项同步机制介绍](#)

**Case2**

#### M-LAG组网下挂终端ping设备上地址不通

### 组/网/说/明

设备及版本：S10508 R7596P09

组网：两台S10508起M-LAG，没有配置VLAN双活或者VRRP网关，单台M-LAG设备上有终端对应的VLAN，M-LAG接口对接下行设备。下图为局部组网图。



### 问/题/描/述

现场M-LAG组网，M-LAG设备上未起VLAN双活或者VRRP网关，设备下行通过M-LAG接口BAGG6对接一台接入设备，BAGG6放通VLAN4019。其中只有M-LAG-1设备上有interface vlan-interface4019地址，地址为10.0.251.5，M-LAG-2设备上未起VLAN4019的三层虚接口，接入设备上int vlan 4019的地址为10.0.251.6，现场测试从接入设备上ping 10.0.251.5不通。

根据往期内容我们知道，现场这种配置是不规范的，那为什么这种不规范的配置会导致不通呢？

对于直连不通的情况，我们首先要确认ARP/MAC的学习情况，查看表项发现，在M-LAG-1设备上，接入设备的ARP学习在了peer-link接口，MAC学习在M-LAG接口：

[M-LAG-1]display arp 10.0.251.6  
Type: S-Static D-Dynamic O-Openflow R-Rule M-Multiport I-Invalid

IP address MAC Address VLAN ID State Port/Nickname Aging Type  
10.0.251.6 1451-7ea4-b589 4019 BAGG1 1200 D

[M-LAG-1]display mac 1451-7ea4-b589  
MAC Address VLAN ID State Port/Nickname Aging  
1451-7ea4-b589 4019 Learned BAGG6 Y

正常情况下，ARP和MAC都应该学习在M-LAG接口，才能实现M-LAG的正常转发。我们先来复习一下M-LAG组网中，两台M-LAG设备间ARP和MAC的同步过程：

M-LAG设备的M-LAG接口在收到ARP请求报文后，会正常二层泛洪广播一份到对端M-LAG设备，对端M-LAG设备上的arp学习在Peer-link接口；同时本端将收到的ARP请求报文上送平台后，还会通过rlink报文将ARP同步到对端M-LAG设备，此时对端M-LAG设备的ARP会学到M-LAG接口，并将上一步学习在peer-link接口的ARP覆盖，如下图：



与ARP同步报文实现不同，M-LAG中MAC同步的是表项，M-LAG接口学习的MAC表项直接同步到另一台设备的M-LAG接口。

在现场故障环境中，我们在设备上做ARP流统计，发现ARP报文从M-LAG-2的M-LAG接口进入设备，由于M-LAG-2设备上没有int vlan 4019，ARP报文不会上送CPU，也就不会触发rlink封装同步，M-LAG-1上只会收到通过二层广播从M-LAG-2设备上来的ARP报文，因此M-LAG-1设备上arp表项都是学在M-LAG接口的。

既然报文没有问题，那为什么spine/border没有正常解封装报文并转发呢？在“[交换那些事儿 | 技术专题篇 - DRNI+EVPN组网进阶介绍](#)”中我们讲解过，设备收到带封装的报文后，如果其封装的源自IP在设备上可以找到对称隧道，就可以解封装，反之则无法正常解封装并转发。我们在leaf上查看到spine/border的隧道只有如下一条，抓到的报文也是封装的这个隧道的源自IP：

**source 60.1.1.5, destination 60.1.1.251**  
抓到的报文也是封装的这个隧道的源自IP；  
在spine/border上查看到leaf的隧道有两条：  
**source 60.1.1.251, destination 60.1.1.1**  
**source 60.1.1.251, destination 60.1.1.2**

很明显spine/border上这两条隧道都不是报文封装源自IP的对称隧道，因此spine/border就是无法正常解封装并转发这个报文的。

那为什么leaf和spine/border上的隧道会不对称呢？我们进一步查看配置信息发现，60.1.1.1和60.1.1.2为两台leaf设备的loopback 0地址，60.1.1.5和60.1.1.251为leaf和spine/border设备的group ip。且由于leaf下存在单挂终端，所以leaf上配置了`evpn drni local local-ip remote remote-ip`命令；而spine/border设备上引入了5类路由，也配置了`nexthop evpn-drni group-address`命令：

**Leaf:**  
evpn drni group 60.1.1.5  
evpn drni local 60.1.1.2 remote 60.1.1.1  
spine\border:  
evpn drni group 60.1.1.251  
nexthop evpn-drni group-address

在这种配置下，由于leaf全单挂，leaf设备会以group ip发布3类路由，以local ip发布2类路由，spine/border只会通过group ip发布引入的5类路由。leaf收到spine/border发布的5类路由后，根据路由下一跳建立虚地址隧道 source 60.1.1.5, destination 60.1.1.251；spine/border上会收到leaf发布的3类和2类路由，但由于spine/border上没有VSI配置，无法接收3类路由，因此只能通过2类路由的下一跳建立到leaf实地址的隧道。



### 问/题/描/述

在spine/border上配置`vxlans default-decapsulation source interface`命令强制解封装即可解决问题。配置该命令后，spine/border收到带VXLAN封装的报文后会直接强制解封装，再走转发流程正常转发。

既然报文没有问题，那为什么spine/border没有正常解封装报文并转发呢？在“[交换那些事儿 | 技术专题篇 - DRNI+EVPN组网进阶介绍](#)”中我们讲解过，设备收到带封装的报文后，如果其封装的源自IP在设备上可以找到对称隧道，就可以解封装，反之则无法正常解封装并转发。我们在leaf上查看到spine/border的隧道只有如下一条，抓到的报文也是封装的这个隧道的源自IP：

**source 60.1.1.5, destination 60.1.1.251**  
抓到的报文也是封装的这个隧道的源自IP；  
在spine/border上查看到leaf的隧道有两条：  
**source 60.1.1.251, destination 60.1.1.1**  
**source 60.1.1.251, destination 60.1.1.2**

很明显spine/border上这两条隧道都不是报文封装源自IP的对称隧道，因此spine/border就是无法正常解封装并转发这个报文的。

那为什么leaf和spine/border上的隧道会不对称呢？我们进一步查看配置信息发现，60.1.1.1和60.1.1.2为两台leaf设备的loopback 0地址，60.1.1.5和60.1.1.251为leaf和spine/border设备的group ip。且由于leaf下存在单挂终端，所以leaf上配置了`evpn drni local local-ip remote remote-ip`命令；而spine/border设备上引入了5类路由，也配置了`nexthop evpn-drni group-address`命令：

**Leaf:**  
evpn drni group 60.1.1.5  
evpn drni local 60.1.1.2 remote 60.1.1.1  
spine\border:  
evpn drni group 60.1.1.251  
nexthop evpn-drni group-address

在这种配置下，由于leaf全单挂，leaf设备会以group ip发布3类路由，以local ip发布2类路由，spine/border只会通过group ip发布引入的5类路由。leaf收到spine/border发布的5类路由后，根据路由下一跳建立虚地址隧道 source 60.1.1.5, destination 60.1.1.251；spine/border上会收到leaf发布的3类和2类路由，但由于spine/border上没有VSI配置，无法接收3类路由，因此只能通过2类路由的下一跳建立到leaf实地址的隧道。



### 问/题/描/述

在spine/border上配置`vxlans default-decapsulation source interface`命令强制解封装即可解决问题。配置该命令后，spine/border收到带VXLAN封装的报文后会直接强制解封装，再走转发流程正常转发。

既然报文没有问题，那为什么spine/border没有正常解封装报文并转发呢？在“[交换那些事儿 | 技术专题篇 - DRNI+EVPN组网进阶介绍](#)”中我们讲解过，设备收到带封装的报文后，如果其封装的源自IP在设备上可以找到对称隧道，就可以解封装，反之则无法正常解封装并转发。我们在leaf上查看到spine/border的隧道只有如下一条，抓到的报文也是封装的这个隧道的源自IP：

**source 60.1.1.5, destination 60.1.1.251**  
抓到的报文也是封装的这个隧道的源自IP；  
在spine/border上查看到leaf的隧道有两条：  
**source 60.1.1.251, destination 60.1.1.1**  
**source 60.1.1.251, destination 60.1.1.2**

很明显spine/border上这两条隧道都不是报文封装源自IP的对称隧道，因此spine/border就是无法正常解封装并转发这个报文的。

那为什么leaf和spine/border上的隧道会不对称呢？我们进一步查看配置信息发现，60.1.1.1和60.1.1.2为两台leaf设备的loopback 0地址，60.1.1.5和60.1.1.251为leaf和spine/border设备的group ip。且由于leaf下存在单挂终端，所以leaf上配置了`evpn drni local local-ip remote remote-ip`命令；而spine/border设备上引入了5类路由，也配置了`nexthop evpn-drni group-address`命令：

**Leaf:**  
evpn drni group 60.1.1.5  
evpn drni local 60.1.1.2 remote 60.1.1.1  
spine\border:  
evpn drni group 60.1.1.251<br