



某局点通过我司4台S75E组建核心网络，其中4台S75E根据定位不同分为两组：中心机房和灾备机房，同一机房两台S75E通过聚合链路互联，不同机房S75E间通过2条三层线路互联，建立两个OSPF邻居，运行在OSPF Area 0中。各个支行通过两台MSR3016上行到分行，分别通过电信和联通线路接入核心网络，运行在Area 46中。近期客户进行网络优化，为减少网络中路由数目，客户把支行网络的OSPF Area设置为NSSA区域，发现原本负载均衡的网络流量都通过联通下行，路由学习出现问题。MSR3016-2上原本应该能学到两条去往业务网段的等价路由，现只有一条路由进入路由表；灾备中心S75E (S75C、S75D) 上原本应该只有一条去往业务网段地址，指向各自连接的下联MSR5040，先出现3条等价路由，指向中心机房S75E (S75A、S75B)。

现场经过信息了解，对相关路由信息进行了提取。某支行网络上的业务地址200.51.36.96/29，其网关在交换机3652EI-1和3652EI-2上(3652EI-1和3652EI-2配置了VRRP)，并通过引入直连路由方式引入了OSPF Area 46中。

逐跳查看路由的学习情况。

1、在支行上联路由器MSR3016-1查看路由由200.51.36.96/29：  
 200.51.36.96/29 O\_NSSA 150 1 200.51.40.194 Eth0/0  
 O\_NSSA 150 1 200.51.40.198 Eth0/1

MSR3016-1分别通过支行SW3652-1和SW3652-2学习到两条去往200.51.36.96/29的等价路由，路由学习正常。

2、在支行上联路由器MSR3016-2查看路由由200.51.36.96/29：  
 200.51.36.96/29 O\_NSSA 150 1 200.51.40.202 Eth0/0

MSR3016-2只学到一条去往200.51.36.96/29的路由，下一跳为S3652-2，路由学习异常，进一步查看MSR3016-2的OSPF Lsdb。

< MSR3016-2>display OSPF Lsdb NSSA 200.51.36.96

```

OSPF Process 1 with Router ID 199.51.46.129
  Area: 0.0.0.46
  Link State Database

  Type :NSSA
  LS ID :200.51.36.96
  Adv Rtr :199.51.46.2
  LS Age :1577
  Len :36
  Options : NP
  Seq# :800012c5
  Checksum :0x36eb
  Net Mask :255.255.255.248
  TOS 0 Metric: 1
  E Type :2
  Forwarding Address :199.51.46.2
  Tag :1

  Type :NSSA
  LS ID :200.51.36.96
    
```

Adv Rtr : 199.51.46.3  
LS Age : 29  
Len : 36  
Options : NP  
Seq# : 800012c5  
Checksum : 0xe775  
Net Mask : 255.255.255.248  
TOS 0 Metric: 1  
E Type : 2  
Forwarding Address : 200.51.40.202  
Tag : 1

通过查看MSR3016-2的OSPF Lsdb发现MSR3016-2收到两条去往业务网段200.51.36.96的LSA, 说明对端S3652路由发送没有问题。进一步查看为什么199.51.46.2 (S3652-1的Router-id) 发送的路由没有添加路由表。收到LSA, 却未优选, 考虑cost值的问题。二类外部路由cost值默认都为1。这是两条7类LSA, 且LSA的Forwarding Address地址不为0, 路由优选时会计算到Forwarding Address的cost值。进而查看MSR3016-2上到两个Forwarding Address的cost值。

< MSR3016-2> display ip routing-table

...

```
199.51.46.0/25   OSPF 10 2      200.51.40.206 Eth0/1
                 OSPF 10 2      200.51.40.202 Eth0/0
200.51.40.200/30 Direct 0 0     200.51.40.201 Eth0/0
```

MSR3016-2到199.51.46.0/25的cost为2, 到200.51.40.200为0。因此MSR3016-2会优选S3652-1发送的, Forwarding Address为200.51.40.200的路由, 无法形成等价。

### 3、查看S75A上查看业务路由的学习情况

< S7506-A>display ip routing-table | include 200.51.36.96

Routing Tables: Public

Destination/Mask	Proto	Pre	Cost	NextHop	Interface
200.51.36.96/29	O_ASE	150	1	200.51.100.58	GE3/0/9

S75A上学到一条去往200.51.36.96/29的路由, 下一跳为联通MSR5040, 路由学习正常。

### 4、查看S75C上查看业务路由的学习情况

dis ip routing-table 200.51.36.96

Routing Table : Public

Summary Count : 3

Destination/Mask	Proto	Pre	Cost	NextHop	InterForwarding Addressce
200.51.36.96/29	O_ASE	150	1	200.51.100.50	GE3/0/2
	O_ASE	150	1	200.51.100.1	GE2/0/15
	O_ASE	150	1	200.51.100.5	GE2/0/16

S75C上学到三条去往200.51.36.96/29的等价路由, 下一跳分别为电信MSR5040与之相连接口地址200.51.100.50, S75A与之互连接口地址200.51.100.1、200.51.100.5。路由学习“异常”。在网络变更之前4台75E上到200.51.36.96/29的路由都只有一条指向到各自连接的下联MSR5040上, 但现在灾备中心的这两台7503E各有3条等价路由。同时由于此处路由的问题导致连接到总行去的思科设备上到200.51.36.96/29的路由都指向到了S7506E-A和S7506E-B, 从而所有到支行的流量都只走联通线路。

我们分析下这三条路由是否真的异常。追溯下路由的形成:

1) 支行SW3引入直连路由200.51.36.96/29, 在OSPF Area46内泛洪type-7 LSA, SW1发布的type-7 LSA Forwarding Address填写199.51.46.2, SW2发布的type-7 LSA Forwarding Address填写200.51.40.202;

2) 比较到Forwarding Address地址的cost, 3016-2优选了3652-2通告的TYPE-7 LSA路由;

3) 分行电信、联通路由器是ABR, 且联通MSR50-2的Router-id大, 所以只有联通MSR50-2对收到type-7 LSA进行7转5, 并在Area 0内泛洪;

4) 联通MSR50-2实际收到两条200.51.36.96/29的type-7 LSA, 比较到Forwarding Address的cost, 到200.51.40.202的cost较小, 因此将这条LSA进行7转5, Forwarding Address地址继承type-7 LSA的Forwarding Address, 为200.51.40.202;

5) 75A/B/C/D均收到一条200.51.36.96/29的type-5 LSA, 其Forwarding Address是200.51.40.202。

5、查询75C/D路由表, 到200.51.40.200/30确存在三条等价路由, 这三条等价路由是经过正常路由计算得出的

子付甲					
200.51.40.200/30	4	Inter	200.51.100.50	199.51.0.13	
200.51.40.200/30	4	Inter	200.51.100.1	199.51.0.14	
200.51.40.200/30	4	Inter	200.51.100.5	199.51.0.14	

对于200.51.36.96/29的type-5 LSA, 加入路由表时, 直接将下一跳地址替换为到Forwarding Address路由的下一跳地址

200.51.36.96/29	1	Type2	1	200.51.100.50	199.51.0.14
200.51.36.96/29	1	Type2	1	200.51.100.1	199.51.0.14
200.51.36.96/29	1	Type2	1	200.51.100.5	199.51.0.14

因此S75C上学习到三条去往200.51.36.96/29的等价路由。

6、查询75A/B路由表, 到Forwarding Address地址200.51.40.200/30的路由由下一跳唯一, 所以到200.51.36.96/29的下一跳唯一。

综上, 该路由由计算结果是符合OSPF标准规定的正确计算结果。

一路分析下来我们也发现了问题产生的缘由都与Forwarding Address相关。包括5类7类路由由优选要计算到Forwarding Address的cost值, 7转5时Forwarding Address会被继承。而问题产生的根源在于分支SW引入直连路由200.51.36.96/29时, Forwarding Address填写不对称: SW1发布的type-7 LSA Forwarding Address填写199.51.46.2, 为设备上非互联虚接口地址; SW1发布的type-7 LSA Forwarding Address填写200.51.40.202, 互联接口地址。为什么这么填写? 这就涉及到Forwarding Address的填写规则。

先看下7类LSA Forwarding Address的填写规则:

7类LSA Forwarding Address的填写有三种可能: 下一跳地址、loopback地址、设备上使能OSPF的接口地址。

在满足以下所有条件的情况下Forwarding Address会填写下一跳地址:

- 1) 引入路由由接口已在OSPF中使能;
- 2) 引入路由由接口没有被设置为被动接口;
- 3) 引入路由由接口的OSPF网络类型不是P2P或者PM2P;
- 4) 引入路由由下一跳地址只有一个;
- 5) 引入路由由下一跳地址是落在出接口的OSPF使能网段之内

否则需要判断NSSA区域内是否使能Loopback接口, 有使能的话Forwarding Address填充Loopback地址, 否则将随机填写设备上使能OSPF的接口地址。

本案例中引入路由为直连路由, 接口未使能OSPF, 设备上也未配置Loopback地址。因此Forwarding Address随机填写的使能OSPF的接口地址, 导致了Forwarding Address填写不对称, 设备到达两个Forwarding Address cost值不一致, 从而无法形成等价的问题。

为什么Area 46未改成NSSA区域之前没有这个问题呢?

因为未改为NSSA区域时引入外部路由为5类外部路由, 5类路由填写Forwarding Address的规则与7类填写Forwarding Address的规则不一样。

5类LSA Forwarding Address的填写规则较为简单, 如下:

5类LSA Forwarding Address的填写有两种可能: 下一跳地址、0.0.0.0。

在满足以下所有条件的情况下Forwarding Address会填写下一跳地址:

- 1) 引入路由由接口已在OSPF中使能;
- 2) 引入路由由接口没有被设置为被动接口
- 3) 引入路由由接口的OSPF网络类型不是P2P或者PM2P
- 4) 引入路由由下一跳地址只有一个
- 5) 引入路由由下一跳地址是落在出接口的OSPF使能网段之内

其他情况全部填写为0.0.0.0, 因此网络未优化前没有选路的问题。

1、在设备上启用Loopback地址固化Forwarding Address, 需要在NSSA区域使能, 即时生效;

2、不使用NSSA区域, 5类LSA填写规则不一样, 不会有这个问题。

这个问题的发生有以下几个要点:

1. 7类LSA始发时, 对Forwarding Address地址的随机选择填充;
2. ABR做7转5时, 对多条7类LSA的选择, 主要是到FA地址的cost;
3. 7转5后, 5类LSA继承7类LSA的Forwarding Address, 计算路由时, 直接把到Forwarding Address路由由下一跳填充为到5类LSA网段的下一跳;
4. 在该特定组网下, 75C/D到Forwarding Address的路由恰好存在3条等价路由

这类纯网路规划问题比较少见, 一旦发现问题影响范围大且不易排查, 找到原因要更改网络规划也比较麻烦。因此在网络规划或变更时一定要考虑周全, 涉及OSPF 5类、7类外部路由时, 要重点关注Forwarding Address的填写, 使Forwarding Address的填写可控。