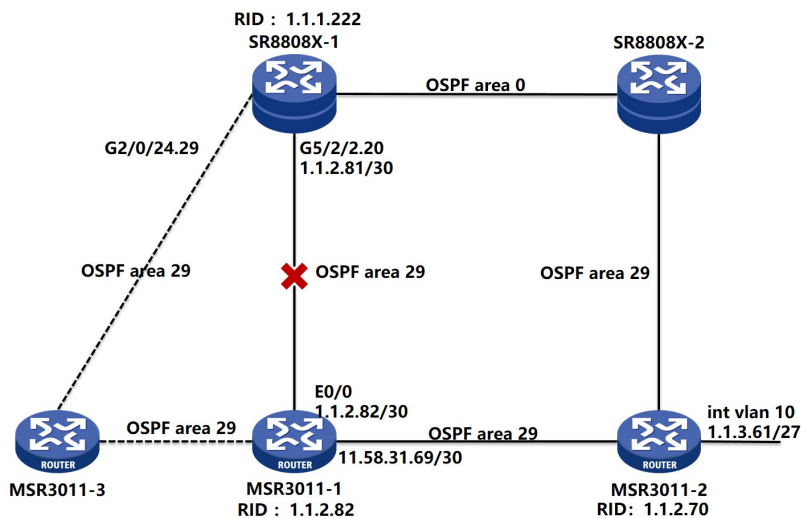


某局点SR8808X运营商线路故障OSPF邻居down但路由未消失经验案例

OSPF 郭昊 2018-03-18 发表



客户反馈现场有四台设备，分别为两台SR8808X及MSR3011-1和MSR3011-2。两台SR8808X之间网段属于OSPF area 0，SR8808X和MSR3011以及两台MSR3011之间网段属于OSPF area 29。MSR3011-2将interface vlan 10地址1.1.3.61引入到OSPF中，正常情况下SR8808X-1到1.1.3.61的路由由下一跳指向MSR3011-1。当SR8808X-1和MSR3011-1之间的运营商线路故障时，两台设备间的OSPF邻居断开了，但是SR8808X-1上面到达1.1.3.61的路由仍指向MSR3011-1，出接口仍为G5/2/2.20，没有切换到其他线路上去，导致转发不通。将SR8808X-1的G5/2/2.20接口shutdown后，路由才能正常切换到其他线路，与1.1.3.61能够互通。

正常情况下，SR8808X-1上查看到达1.1.3.61的路由信息为，

```
[pbcnmhhr01]dis ip routing-table 1.1.3.61
```

```
Summary Count : 3
```

Destination/Mask	Proto	Pre	Cost	NextHop	Interface
0.0.0.0/0	BGP	100	0	11.127.251.165	GE2/0/17
1.1.0.0/15	Static	60	0	0.0.0.0	NULL0
1.1.3.32/27	O_NSSA2	150	1	1.1.2.82	GE5/2/2.20

//路由由下一跳指向MSR3011-1，出接口为G5/2/2.20，该路由是OSPF NSSA区域引入的外部路由。

对应的LSDB信息为，

```
[pbcnmhhr01]dis ospf lsdb nssa 1.1.3.32
```

```
OSPF Process 100 with Router ID 1.1.1.222
```

```
Area: 0.0.0.0
```

```
Link State Database
```

```
Area: 0.0.0.29
```

```
Link State Database
```

```
Type : NSSA
```

```
LS ID : 1.1.3.32
```

```
Adv Rtr : 1.1.2.82
```

```
LS Age : 187
```

```
Len : 36
```

```
Options : NP
```

```
Seq# : 800056ff
```

```
Checksum : 0x4d55
```

```
Net Mask : 255.255.255.224
```

```
TOS 0 Metric: 1
```

```
E Type : 2
```

```
Forwarding Address : 1.1.2.82
```

```
Tag : 1
```

//这一条LSA是MSR3011-1发布的，FA地址为MSR3011-1的E0/0地址，该接口过运营商线路与SR8808X-1相连

Type :NSSA
LS ID :1.1.3.32
Adv Rtr :1.1.2.70
LS Age :1366
Len :36
Options :NP
Seq# :80005702
Checksum :0xe5d1
Net Mask :255.255.255.224
TOS 0 Metric: 1
E Type :2
Forwarding Address :1.1.2.70
Tag :1

//这一条是MSR3011-2发布的1.1.3.61，不影响该问题

当SR8808X-1和MSR3011-1中间线路故障时，SR8808X感知不到中间线路故障，接口G5/2/2.20仍为up的，这个是正常的。但SR8808X-1经过4个OSPF hello间隔收不到MSR3011-1的OSPF hello报文、将OSPF邻居置down后，display ip routing-table 1.1.3.61以及display ospf lsdb nssa 1.1.3.32回显结果与上述相同，这是有问题的。

既然这条OSPF路由没有消失，说明设备认为对应的LSA还是有效，并且根据这条LSA正常计算出了路由。我们来看一下上面LSDB的信息，这条MSR3011-1发出的LSA，FA地址为1.1.2.82，为SR8808X-1的直连接口地址。设备收到5类LSA时，会根据FA地址进行迭代计算出到达LSA发布者的最佳路径。由于线路故障时SR8808X的接口仍是up的，所以SR8808X-1根据这条LSA计算路由时认为到达FA地址的最佳路径是从G5/2/2.20出去。因此是对应的LSA没有失效，导致路由没有切换到其他出口。

接下来就要看为何这条LSA在线路故障后仍未失效。线路故障时，SR8808X-1应该通过其他对应目的地址1.1.3.61的LSA重新计算路由。既然dis ospf lsdb显示LSA没变，说明SR8808X-1从其他线路收到了与线路故障前MSR3011-1发来的相同LSA，而最有可能造成这种情况的就是组网中还存在另一台属于OSPF area 29的设备，类似于我们上面组网图中的MSR3011-3。该设备将MSR3011-1发布的LSA传给了SR8808X-1，并且由于是区域内路由器，在传递LSA时没有更改FA地址，最终造成了SR8808X路由异常。后与客户确认，组网中确实存在这样一台类似于MSR3011-3的设备。

从上面这个过程来看，SR8808X、MSR3011对OSPF LSA的发布、转发以及路由的计算都是正常的，是组网导致了线路故障时路由不切换的现象。那么该如何解决呢？既然路由不切换直接原因是FA地址为SR88X直连网段地址，不难想到，可以更改MSR3011-1发出LSA的FA地址，使得线路故障时，SR8808X-1能够正常计算出到FA的最佳路径。

设备在向域内发布LSA、填写LSA中的FA时，对地址的选取有较为复杂的规则，此处简单列举两条：

1. 是否将环回口地址发布到对应area中，有则将环回口地址填到FA；
 2. 对于没有发布环回口的情况，设备会将先发布到对应area的接口地址填写到FA。
- 所以可通过在MSR3011-1上将环回口地址发布到OSPF area 29中来改变LSA的FA。

通过在MSR3011-1上将环回口地址发布到OSPF area 29的方式，改变MSR3011-1发出的LSA中的FA地址。这样，当SR8808X-1和MSR3011-1之间线路故障时，SR8808X-1根据从MSR3011-3学到的LSA计算路由时，会迭代查询MSR3011-1环回口的路由，重新选择正确的出接口。

对于该OSPF路由异常的问题，我们首先需要怀疑的不是设备本身，毕竟OSPF路由计算属于较为基本的功能，设备计算OSPF路由出错的可能性较小，原因更可能是组网不合理。在这个问题中，我们从查看路由出发，分析对应的LSDB信息以及路由的计算过程，这要求我们对OSPF的LSA发布及路由计算有一定的了解。

此外，现场反馈的组网信息可能并不完整，例如该问题中，现场人员可能并没有认为MSR3011-3会与问题现象相关，因此没有反馈。所以在分析问题时，不应局限于现场反馈的组网，还需在分析过程中考虑更多的可能性。