

路由器在多个OSPF进程下的路由学习问题

目录

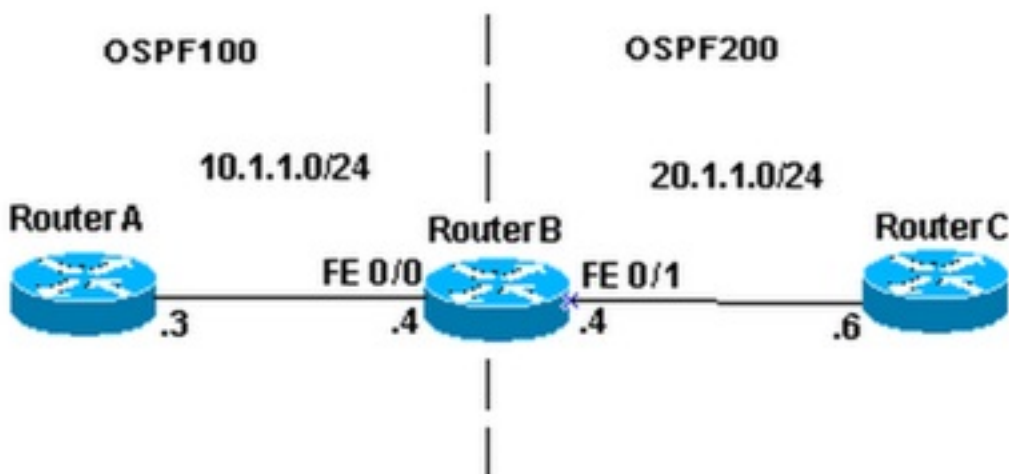
- [软硬件平台](#)
- [问题描述](#)
- [故障诊断步骤](#)
- [经验总结](#)
- [相关命令](#)
- [相关错误信息](#)
- [其他相关文档](#)

软硬件平台

路由器及多层交换机设备，所有软件平台

问题描述

OSPF是我们维护工作普遍应用的一种路由协议，在使用过程中我们经常会发现客户会遇到这样的问题。当路由器B配置了两个OSPF进程（如下图A），为了实现路由的备份，用户在两个进程内发布了两条相同的路由条目，例如在如下图中分别在路由器A和路由器C上发布了路由条目：172.168.0.1/32



在客户原先的网络设计中，将路由器A发布的路由cost值为1，而路由器C发布的metric值被人为增加cost为200。客户认为根据OSPF协议的选路比较原则，在路由器B的路由表中应该优先选择metric最小的路由，因而会选择路由器A的接口作为该路由的下一跳，而路由器C发布的路由作为本条路由的备份。

```
RouterB#show ip route
Codes: L - local, C - connected, S - static, R - RIP, M - mobile, B - BGP
       D - EIGRP, EX - EIGRP external, O - OSPF, IA - OSPF inter area
       N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
       E1 - OSPF external type 1, E2 - OSPF external type 2
       i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
       ia - IS-IS inter area, * - candidate default, U - per-user static route
       o - ODR, P - periodic downloaded static route, + - replicated route
```

```
Gateway of last resort is not set

    1.0.0.0/8 is variably subnetted, 3 subnets, 2 masks
L       20.1.1.0/24 is directly connected, Ethernet0/1
L       10.1.1.0/24 is directly connected, Ethernet0/0
O 172.168.0.1 [110/11] via 10.1.1.3, 00:53:32, Ethernet0/0
```

按照客户的网络设计，当FE0/0端口故障中断的时候，172.168.0.1/32路由的下一跳将切换到RouterC上，直到FE0/0端口恢复，OSPF路由收敛完毕之后，OSPF路由表应该比较metric值而重新选举RouterA作为该路由的下一跳。事实上，客户却在实际网络中发现，故障恢复后，下一跳依然选择RouterC！这是为什么呢？

```
RouterB#show ip route
Codes: L - local, C - connected, S - static, R - RIP, M - mobile, B - BGP
       D - EIGRP, EX - EIGRP external, O - OSPF, IA - OSPF inter area
       N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
       E1 - OSPF external type 1, E2 - OSPF external type 2
       i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
       ia - IS-IS inter area, * - candidate default, U - per-user static route
       o - ODR, P - periodic downloaded static route, + - replicated route
```

```
Gateway of last resort is not set

    1.0.0.0/8 is variably subnetted, 3 subnets, 2 masks
L       20.1.1.0/24 is directly connected, Ethernet0/1
L       10.1.1.0/24 is directly connected, Ethernet0/0
O 172.168.0.1 [110/210] via 20.1.1.6, 01:00:32, Ethernet0/1
```

故障诊断步骤

在做故障诊断之前，首先让我们复习一下OSPF协议选路的原则，当单进程的一台OSPF路由器检查一个数据包的目的地址的时候，它应该通过下面的步骤选择最优路由：

1. 选择可以和目的地址最精确匹配的路由，例如路由表中存在路由条目:172.168.64.0/18, 172.16.64.0/24, 172.16.64.192.27，而目的地址是172.16.64.205,那么最后一条路由条目被选中。最精确的匹配总是应用最长匹配原则。
2. 区域内的路径总是优先于区域间的路径，E1的外部路径总是优先于E2外部路径
3. 不同的路由协议之间通过AD值进行比较路由，AD值相同的情况下会通过Metric进行比较

在我们的案例中，两台路由器发布的路由分别属于不同的OSPF进程，它们拥有相同的AD值和不同的metric值，那么是否真的按照客户的想法去实现的呢？让我们从ospf路由表看看详细信息：

```
RouterB#show ip ospf rib

      OSPF Router with ID (20.1.1.6) (Process ID 200)
      Base Topology (MTID 0)
OSPF local RIB
Codes: * - Best, > - Installed in global RIB
*   20.1.1.0/24, Intra, cost 10, area 0, Connected
    via 20.1.1.4, Ethernet0/1
*>  3.3.3.3/32, Intra, cost 11, area 0
    via 20.1.1.6, Ethernet0/1
*> 172.168.0.1/32, Intra, cost 210, area 0 via 20.1.1.6, Ethernet0/1 OSPF Router with
```

```
ID (10.1.1.3) (Process ID 100) Base Topology (MTID 0) OSPF local RIB Codes: * - Best,
> - Installed in global RIB * 10.1.1.0/24, Intra, cost 10, area 0, Connected via
10.1.1.4, Ethernet0/0 *> 1.1.1.1/32, Intra, cost 11, area 0 via 10.1.1.3, Ethernet0/0
* 172.168.0.1/32, Intra, cost 11, area 0 via 10.1.1.3, Ethernet0/0
```

我们发现OSPF确实选择cost值小的路由作为自己的下一跳，但是当我们手工clear OSPF 200的进程，当路由收敛之后，我们发现路由表又重新选择RouterA作为自己的下一跳了：

```
RouterB#show ip ospf rib

                OSPF Router with ID (20.1.1.6) (Process ID 200)
                Base Topology (MTID 0)
OSPF local RIB
Codes: * - Best, > - Installed in global RIB
*   20.1.1.0/24, Intra, cost 10, area 0, Connected
    via 20.1.1.4, Ethernet0/1
*>  3.3.3.3/32, Intra, cost 11, area 0
    via 20.1.1.6, Ethernet0/1
* 172.168.0.1/32, Intra, cost 210, area 0 via 20.1.1.6, Ethernet0/1 OSPF Router with
ID (10.1.1.3) (Process ID 100) Base Topology (MTID 0) OSPF local RIB Codes: * - Best,
> - Installed in global RIB * 10.1.1.0/24, Intra, cost 10, area 0, Connected via
10.1.1.4, Ethernet0/0 *> 1.1.1.1/32, Intra, cost 11, area 0 via 10.1.1.3, Ethernet0/0
*> 172.168.0.1/32, Intra, cost 11, area 0 via 10.1.1.3, Ethernet0/0
```

在我们反复初始化几次进程，我们会发现在RouterB上根据OSPF进程的启动收敛顺序，哪个进程率先把路由注入到ospf路由表，将优先选择它作为该路由的下一跳！这是因为METRIC值只会在同一进程内进行比较。在多进程OSPF的路由器上，不同进程的路由OSPF路由表只会注入OSPF路由表的顺序优先选择一条作为最佳路由放在全局路由表中，而不会针对metric属性进行选择，。

经验总结

由此我们可以看出，不同的OSPF进程在路由器上类似两个不同的路由协议，我们要进行路由比较的时候，可以根据AD值而不是metric值进行比较。因此我们再设计类似的网络的时候，可以通过修改不同进程的AD值达到用户期望的目的。

通过这个案例，我们知道对于路由协议相关命令的输出，接合自己的网络要知道正常预期输出结果，找到与预期行为不同的细节，就能分析出解决问题的线索。

相关命令

以下数据来自真实网络环境。为保护客户资料，隐去敏感信息，同时不影响故障排查示例。

一般说来，接收端对错误的探测更为敏感。我们常常从接收端查起。查看s1rx, s2rx, s3rx。在本例中，我们可以看到在s2rx的几条fabric link 探测到了错误。以下略去对s1rx, s3rx, 以及发送端的排查输出。

```
show ip route
show ip route x.x.x.x
show ip ospf
```

相关错误信息

无

其他相关文档

[OSPF 排错流程](#)

[OSPF多进程 排错指南](#)