第4章 路由协议故障诊断与排除

ISSUE 3.0

日期：

课程目标

学习完本课程，您应该能够：

- 掌握RIP协议的故障诊断和排除
- 掌握OSPF协议的故障诊断和排除
- 掌握BGP协议的故障诊断和排除

目录

- RIP故障诊断和排除
- OSPF故障诊断和排除
- BGP故障诊断和排除

RIP故障诊断和排除

- RIP协议简介
- RIP故障排查基础知识
- RIP故障排查基本方法
- RIP典型案例

RIP协议概述

- **RIP是Routing Information Protocol（路由信息协议）的简称**
- **RIP路由协议是距离矢量路由协议的一个具体实现**
- **RIP协议适用于中小型网络，有RIPv1和RIPv2**
- **RIPv2使用组播（224.0.0.9）发送，支持验证和VLSM**

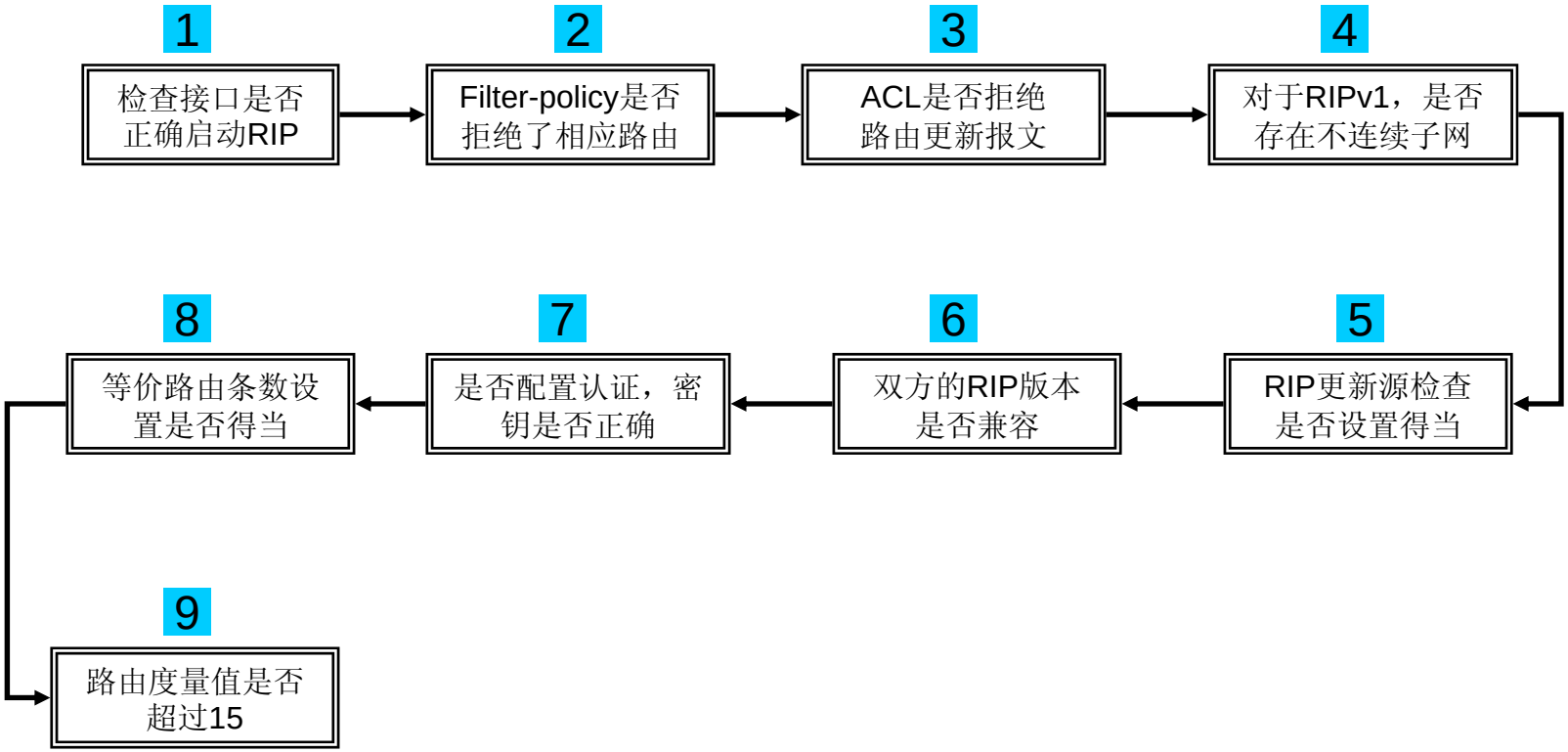
RIP排障基础知识（1）

- RIP度量值
- 重要计时器
  - → 更新计时器
  - → 失效计时器
  - → 保持计时器
  - → 垃圾收集计时器
- 水平分割
- 毒性逆转
- 发送路由更新原则
- 接受路由更新原则

RIP排障基础知识（2）

- **RIP版本2的改进**
  - → 组播方式发送路由更新
  - → 支持可变长子网掩码
  - → 下一跳地址
  - → 支持认证

RIP路由无法加入路由表故障排查



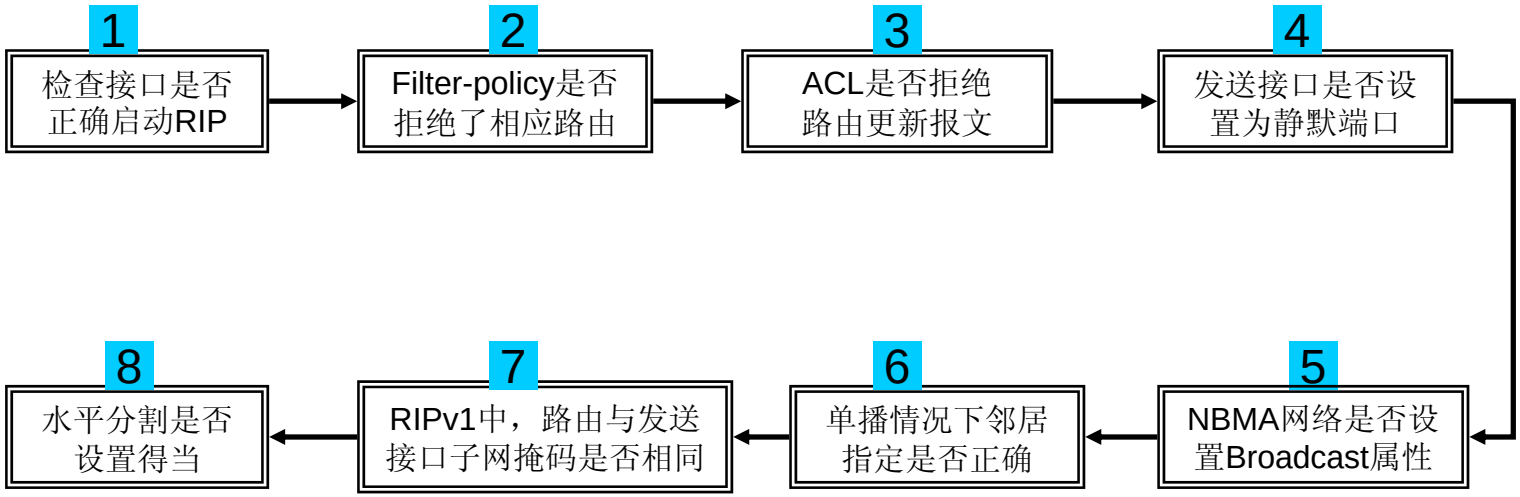
故障排除方法（1）

- **接口是否正确启动RIP协议**
  - → Network命令包含两层含义
    - 在接口地址包含在network主网络内的三层接口上使能RIP
    - 在RIP更新中发布相应的路由
- **Filter-policy是否设置正确**
  - → Filter-policy命令设置不当，拒绝了相应路由加入路由表
- **ACL配置是否拒绝了路由更新报文**
- **RIP版本1不连续子网问题**
  - → RIPv1是有类路由协议，在不连续子网的情况下，会导致子网路由缺失。

故障排除方法（2）

- **RIP更新的源合法性检查**
  - → 更新报文源地址是否与本端同一网段
  - → 更新报文源地址是否对端接口地址
- **RIP版本不兼容**
- **RIP认证配置是否正确**
- **引入或接收的RIP路由度量值过大**
  - → 确保路由在网络中传播时不会由于度量值达到16而被丢弃

RIP路由更新无法发送故障排查



故障排除方法（1）

- **Network命令配置是否正确**
  - → 错误的配置或没有配置network命令，RIP就不会在那个接口上运行,路由无法发送
- **Filter-policy、ACL是否设置正确**
- **发送接口是否设置为静默端口**
  - → 输出接口设置为静默端口，任何RIP更新都不会从该接口发送出去
- **非广播网络是否支持广播流量**
  - → NBMA（帧中继）网络，设置Broadcast属性

故障排除方法（2）

- 单播情况下RIP邻居是否正确
- 路由更新是否与接口子网掩码相同
- 水平分割引起的问题
  - → 点到多点帧中继网络，开启或关闭水平分割都可能产生问题，可能的情况下最好采用划分多个子接口的方式解决问题

RIP相关排障命令

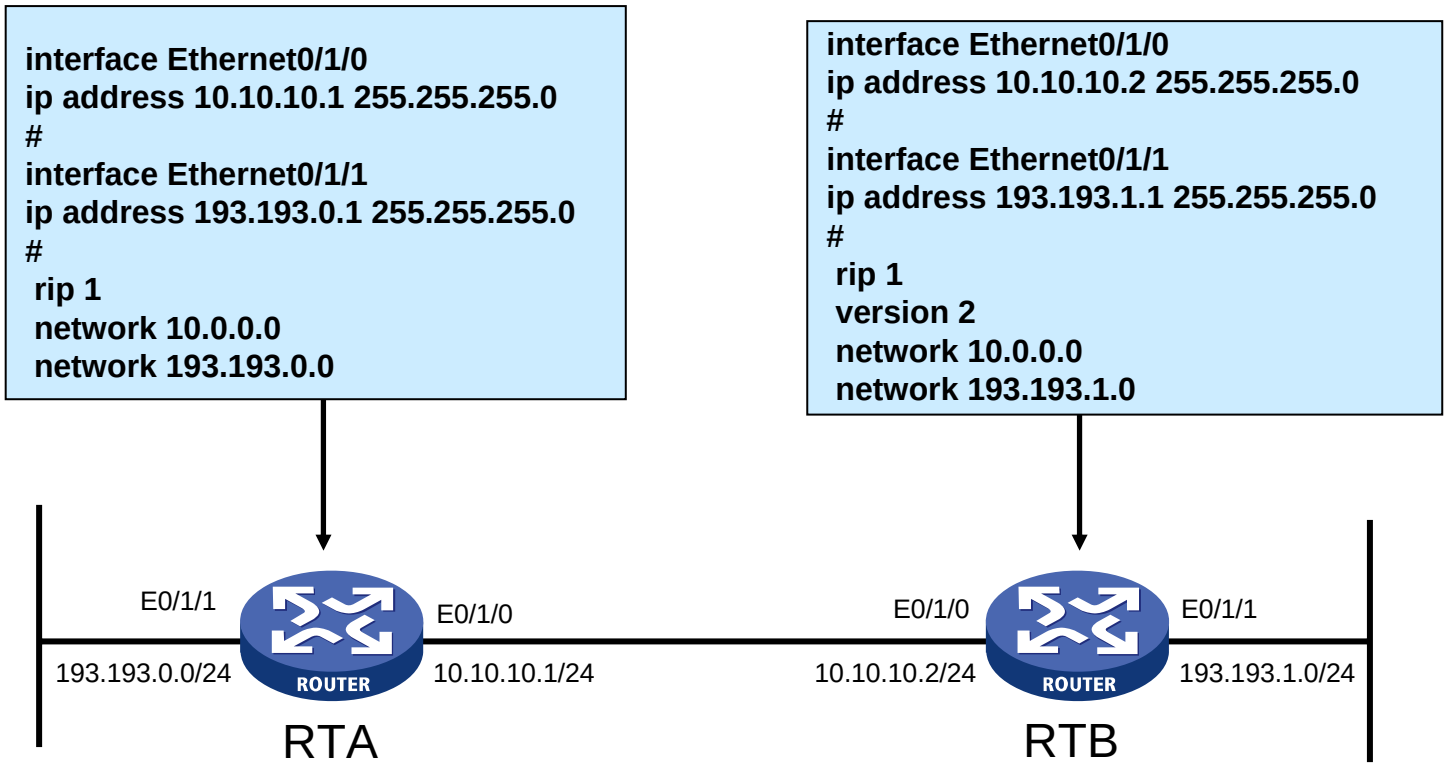
- **display rip**
  - → 显示RIP当前运行状态及配置信息

```
[H3C] display rip
Public VPN-instance name :
   RIP process : 1
RIP process : 1
      RIP version : 1
      Preference : 100
      Checkzero : Enabled
      Default-cost : 0
      Summary : Enabled
      Hostroutes : Enabled
      Maximum number of balanced paths : 6
      Update time   :  30 sec(s)  Timeout
time       :  180 sec(s)
      Suppress time :  120 sec(s)  Garbage-
collect time :  120 sec(s)
```

- **debugging rip packet**
  - → 打开RIP报文调试信息开关

```
[H3C] debugging rip 1 packet
*Jan 13 03:04:44:313 2009 RTA
RM/6/RMDEBUG: RIP 1 : Receive response
from 10.1.1.2 on Ethernet0/1/0.1
*Jan 13 03:04:44:328 2009 RTA
RM/6/RMDEBUG:   Packet : vers 1, cmd
response, length 24
*Jan 13 03:04:44:328 2009 RTA
RM/6/RMDEBUG:     AFI 2, dest 10.3.1.0, cost 1
*Jan 13 03:04:44:344 2009 RTA
RM/6/RMDEBUG: RIP 1 : Can not find interface
for source address.
*Jan 13 03:04:46:313 2009 RTA
RM/6/RMDEBUG: RIP 1 : Sending response on
interface Ethernet0/1/0.1 from 10.1.1.1 to
255.255.255.255
```

RIP典型案例一（1）



- 故障现象：

  → 两台运行RIP协议的路由器在物理连接正常的情况下一台可以学习到路由，另一台无法学习到路由。

RIP典型案例一（2）

- 排障过程

→在路由器上执行display ip routing-table命令查看路由表

-RTA可以学习到RTB直连的193.193.1.0/24的路由，RTB无法学习到RTA直连的193.193.0.0/24的路由

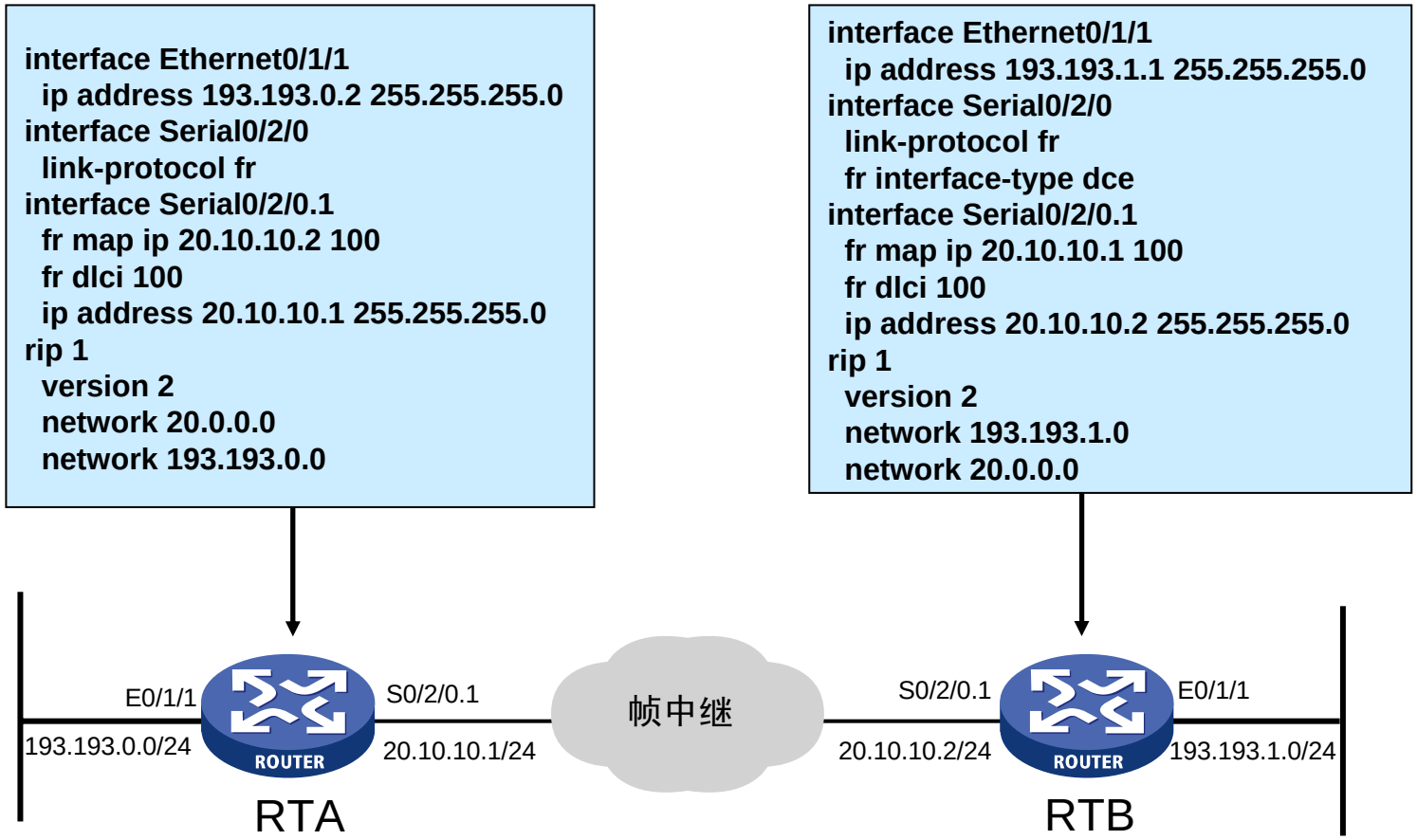
→在RTB上用命令debugging rip 调试开关，发现原因是版本不匹配

***Dec 4 20:03:55:484 2008 RTB RM/6/RMDEBUG: RIP 1 : Receive response from 10.10.10.1 on Ethernet0/1/0**
***Dec 4 20:03:55:500 2008 RTB RM/3/RMDEBUG: RIP 1 : Ignoring this packet. Version is not configured.**

RIP典型案例一（3）

- 解决方案
  - → 将RTA的RIP版本调整为版本2后，RTA和RTB都可以互相学习到正确的路由
- 原因分析
  - → 缺省情况下，启用RIP的接口发送RIP版本1报文，同时可以识别RIP版本2的报文。
  - → 如果显式指定接口使用RIP版本2，则启用RIP的接口只能识别RIP版本2的报文，忽略RIP版本1 报文。

RIP典型案例二（1）



- 故障现象

  → RTA与RTB采用点到多点子接口通过帧中继网络互联，RTA与RTB都无法学习对方路由。

RIP典型案例二（2）

- 排障过程

→打开RTB和RTB debug rip 调试开关

-RTA和RTB的接口都在定期发送RIP更新，但是双方都没有收到对方发过来的RIP更新

→用Ping来检查RTA与RTB间直连链路的连通性，正常

→检查RTA和RTB上帧中继的静态MAP

```
<RTB>display fr map-info
Map Statistics for interface Serial0/2/0 (DCE)
  DLCI = 100, IP 20.10.10.1, Serial0/2/0.1
 create time = 2008/12/05 02:21:25, status = ACTIVE
encapsulation = ietf, vlink = 10
```

→因输出中没有参数broadcast，说明链路不允许发送广播报文或组播报文。

RIP典型案例二（3）

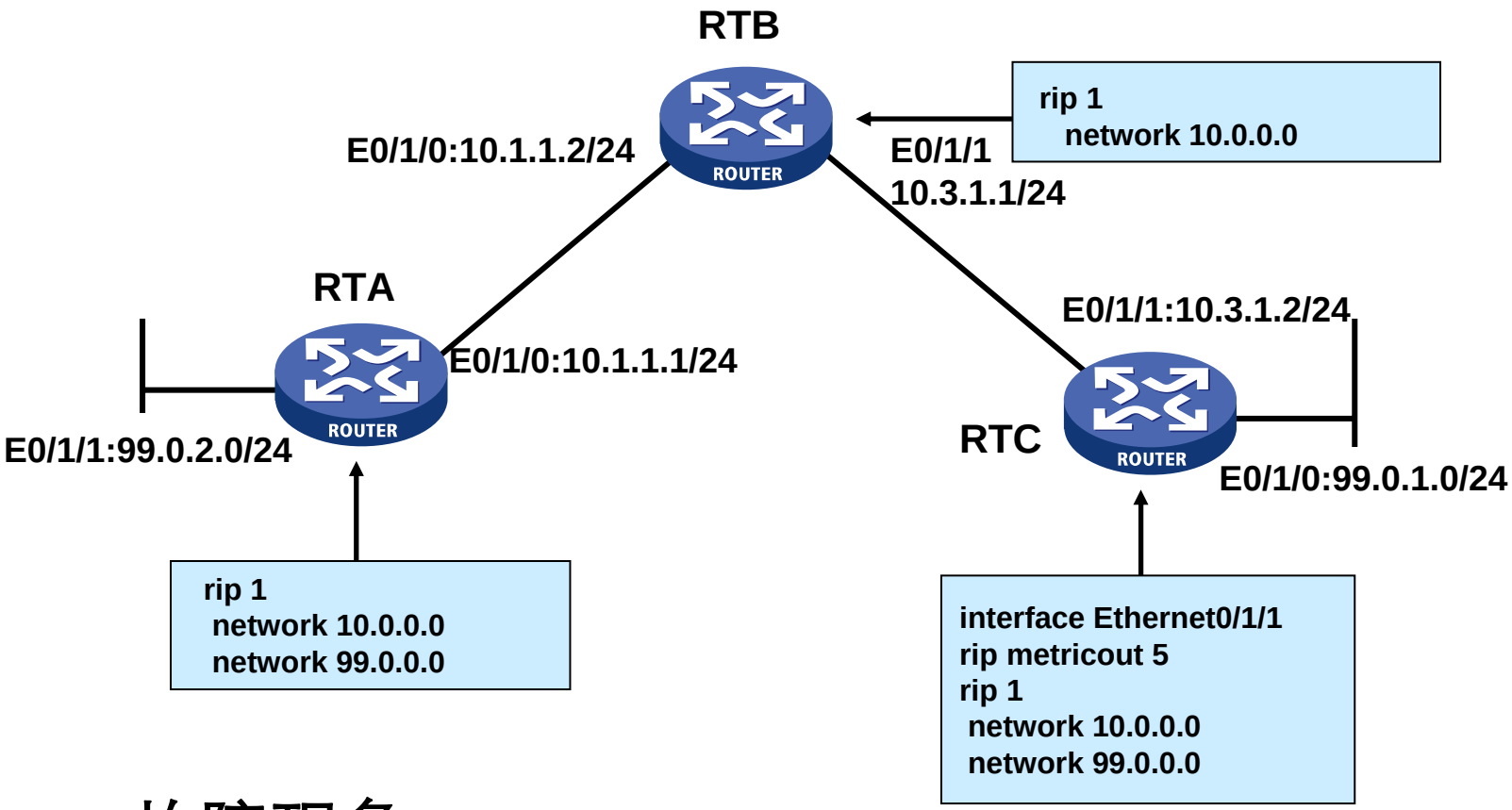
- 解决方案
  - → 调整RTA和RTB的配置，增加静态MAP的广播属性，允许广播或组播报文在链路上传送

RTA
interface Serial0/2/0.1
fr map ip 20.10.10.2 100
broadcast

RTB
interface Serial0/2/0.1
fr map ip 20.10.10.1 100
broadcast

- 建议与总结
  - → 确保下层（物理层、数据链路层、网络层）协议工作正常，然后再对路由协议本身进行排障。

RIP典型案例三（1）



- 故障现象

  → RTC无法访问RTA的子网99.0.2.0/24,RTA无法访问RTC的子网99.0.1.0/24。

RIP典型案例三（2）

- 排障过程
  - → 在RTA及RTC上使用display ip routing-table命令查看路由表，未发现相应子网的路由
  - → 在RTB和RTC打开调试开关，分别观察RTB发出的更新报文及RTC接受的更新报文

RTB
***Jan 13 06:14:38:391 2009 RTB RM/6/RMDEBUG: RIP 1 : Sending response on interface Ethernet0/1/1 from 10.3.1.1 to 255.255.255.255**
***Jan 13 06:14:38:438 2009 RTB RM/6/RMDEBUG: AFI 2, dest 99.0.0.0, cost 2**

RTC
***Jan 13 06:16:09:31 2009 RTC RM/6/RMDEBUG: RIP 1 : Receive response from 10.3.1.1 on Ethernet0/1/1**
***Jan 13 06:16:09:47 2009 RTC RM/6/RMDEBUG: AFI 2, dest 99.0.0.0, cost 2**
***Jan 13 06:16:09:62 2009 RTC RM/3/RMDEBUG: RIP 1 : Ignoring route 99.0.0.0. Its major net addr is same as the local interface's.**

RIP典型案例三（3）

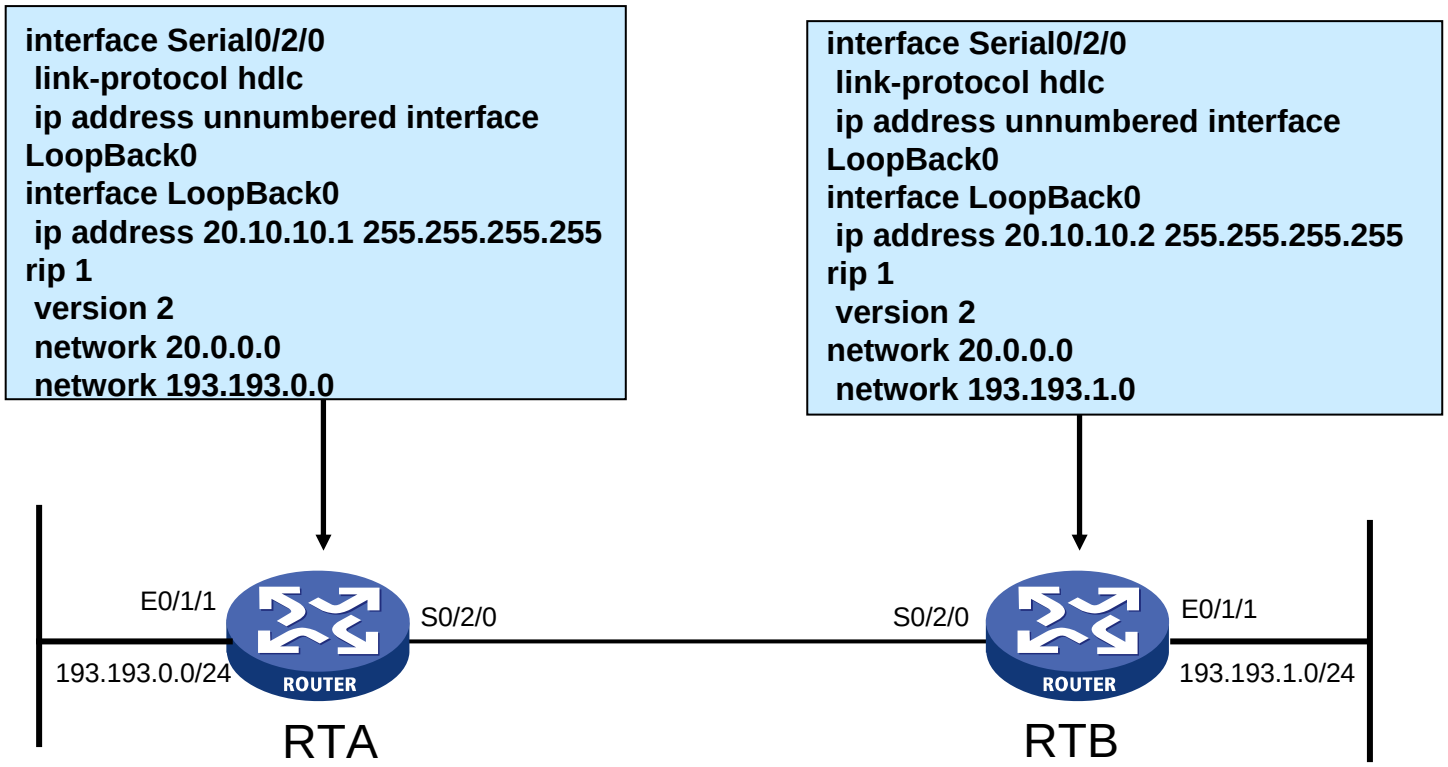
- 排障过程

  → 从RTB和RTC的调试信息来看，RTC接收到RTB发送的99.0.0.0/8的路由更新,但是由于RTC本身有一个接口属于99.0.0.0/8的子网，根据RIP路由的接收原则，该更新被忽略

- 解决方案

  → 使用RIPv2，并在RTA和RTC上取消自动聚合

RIP典型案例四（1）



- 故障现象

→RTA和RTB都无法学习到对端设备的路由。

RIP典型案例四（2）

- 排障过程

→在RTA上ping RTB 20.10.10.2不通，增加静态路由后可以实现互通

→在RTA和RTB上用命令debugging rip查看RIP调试信息，发现问题原因

-RTB在收到路由更新后会检查更新报文中的源地址。如果源地址与端口的IP地址不在同一个网段，则会导致源检查失败，相应的路由无法加入路由表。

```
*Dec  5 06:43:33:109 2008 RTB RM/6/RMDEBUG: RIP 1 : Can not find interface for source address.
```

- 解决方案

→ 在RTA、RTB上关闭更新源检查

```
[H3C-rip-1]undo validate-source-address
```

目录

- RIP故障诊断和排除
- OSPF故障诊断和排除
- BGP故障诊断和排除

OSPF故障排除

- OSPF协议简介
- OSPF故障排查基础知识
- OSPF故障排查基本方法
- OSPF典型案例分析

OSPF协议简介

- 无路由自环
- 可适应大规模网络
- 路由变化收敛速度快
- 支持区域划分
- 支持等值路由
- 支持验证
- 支持路由分级管理
- 支持以组播地址发送协议报文

OSPF排障基础知识（1）

- **OSPF协议号（IP 89）**
- **Router ID的选择**
  - → 优选最大的loopback，其次为最大的接口地址
- **OSPF Hello报文**
  - → 类型1，用于在路由器之间形成OSPF邻居关系
  - → 可以组播方式或单播方式（NBMA网络）发送
  - → 如子网掩码不同，则邻居建立失败
- **OSPF DD报文**
  - → 类型2，用于对OSPF 链路状态数据库进行描述
  - → 用于在LSDB数据交换期间的主从确认。RouterID大者为主设备

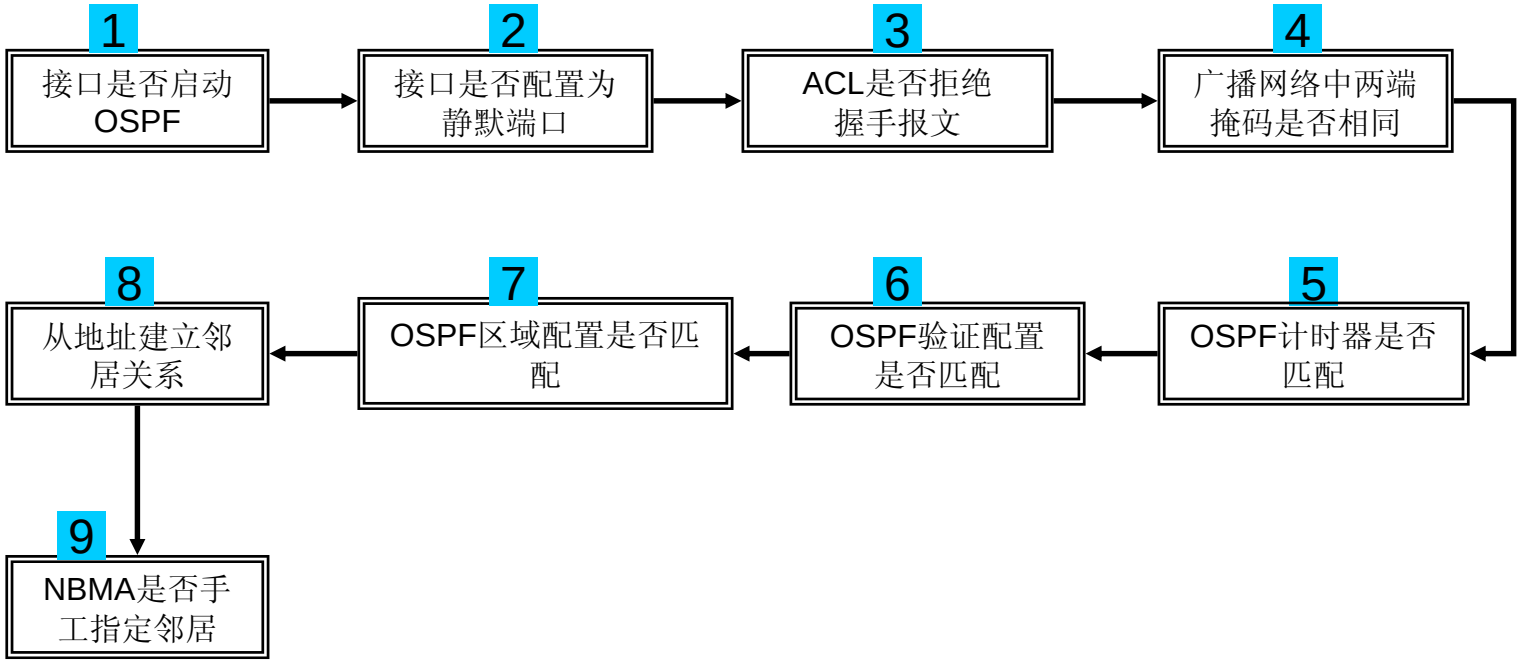
OSPF排障基础知识（2）

- **OSPF LSA**
  - → 6种LSA，LSA老化时间3600秒，每隔1800秒重新泛洪。
  - → Router LSA描述链路类型，仅在本区域内泛洪
  - → 汇总LSA负责在区域间传递路由信息，由ABR生成。包括类型3（Summary LSA）和类型4（Summary ASBR ）。
- **OSPF特殊区域**
  - → 存根区域（Stub Area）
  - → 完全存根区域（Totally Stubby Area）
- **OSPF 转发地址（FA）**
  - → 通常在5类LSA中，FA填写为0.0.0.0。满足以下条件时，FA设置为非0
    - OSPF在下一跳接口启动
    - 下一跳接口非静默端口
    - 下一条接口非P2P或P2MP接口
  - → FA设置为非0的目的是路径优选

OSPF常见故障

- OSPF无法形成邻居关系
- OSPF邻接关系停滞在异常状态
- OSPF路由无法通告
- OSPF路由无法加入路由表
- SPF重复计算

OSPF无法形成邻居关系故障排查（1）



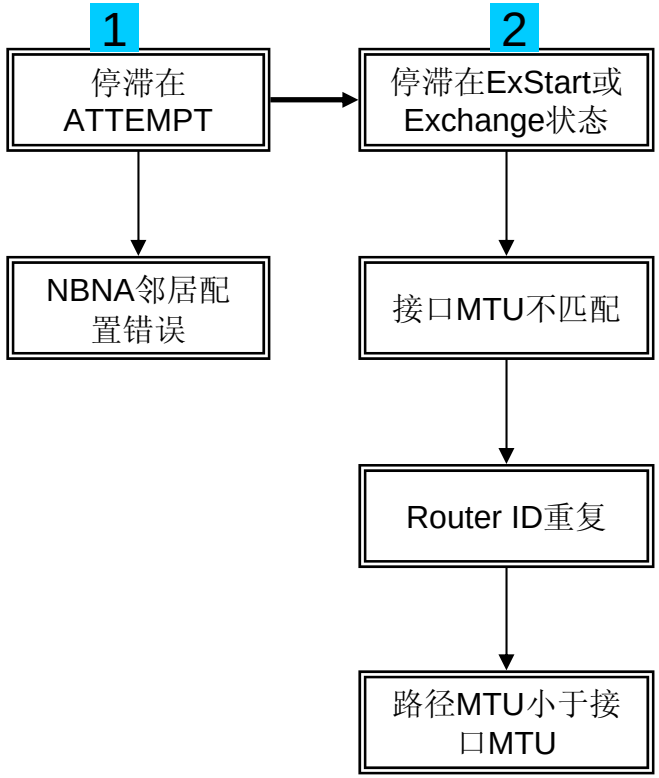
OSPF无法形成邻居关系故障排查（2）

- **接口是否启动OSPF**
  - → OSPF的运行是基于设备接口的，如果OSPF没有在接口启动，那么邻居关系肯定无法形成
- **接口是否配置为静默端口**
  - → 设置为静默端口时，不能发送OSPF Hello报文
- **ACL是否拒绝了Hello报文**
  - → OSPF组播地址为224.0.0.5
- **广播网络中两端接口子网掩码是否相同**
  - → 如果两端接口属于不同的IP子网，那么邻居关系无法形成
- **两端OSPF计时器设定值是否匹配**

OSPF无法形成邻居关系故障排查（3）

- **OSPF验证配置是否匹配**
- **OSPF区域配置是否匹配**
  - → 区域类型或区域ID不匹配则不会形成邻居关系
- **OSPF邻居是否使用从地址建立**
  - → OSPF邻居关系只能使用接口的主地址进行建立，从地址无法建立邻居关系
- **NBMA网络是否指定邻居**
  - → OSPF网络类型为NBMA时，必须手工指定邻居的IP地址，否则端口无法发送Hello报文，无法形成邻居关系。

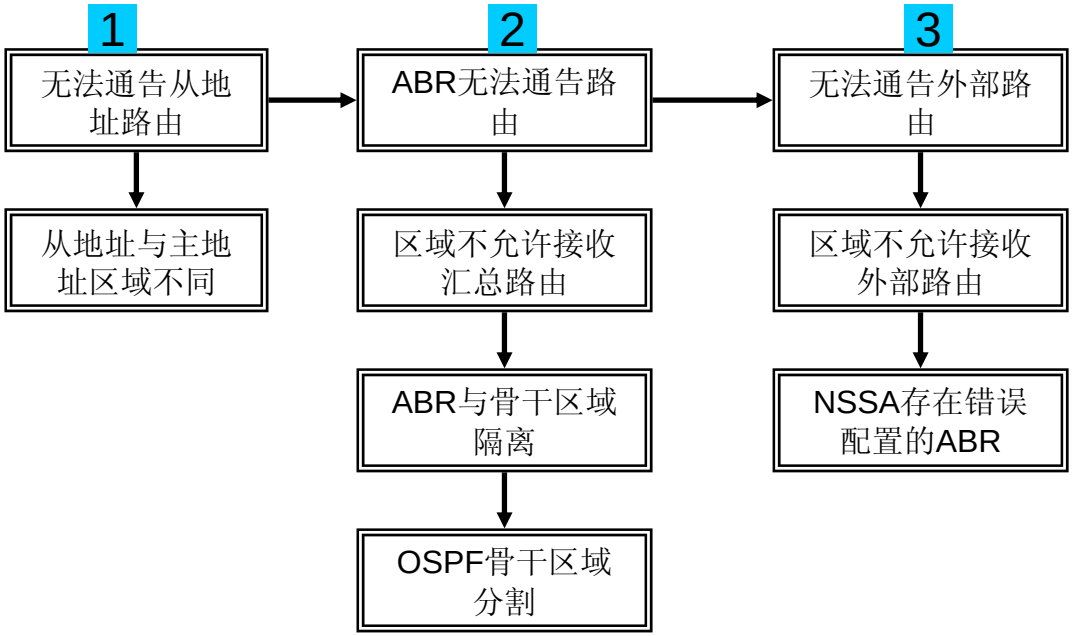
OSPF邻居关系停滞于异常状态故障排查（1）



OSPF邻居关系停滞于异常状态故障排查（2）

- **邻居关系停滞与ATEMPT**
  - → 仅仅在网络类型是NBMA的情况下
  - → Hello发出未收到回应，最常见原因是NBMA邻居配置错误
- **邻居关系停滞于Exstart或Exchange状态**
  - → 接口MTU设置不匹配
    - DD报文中携带了接口的MTU信息
  - → 邻居Router ID重复
    - 通过Router ID的信息确定邻居的主从关系
  - → 路径MTU小于接口MTU
    - 大的OSPF报文将在传输路径上被丢弃，导致邻居双方无法完成完整的数据库信息交互

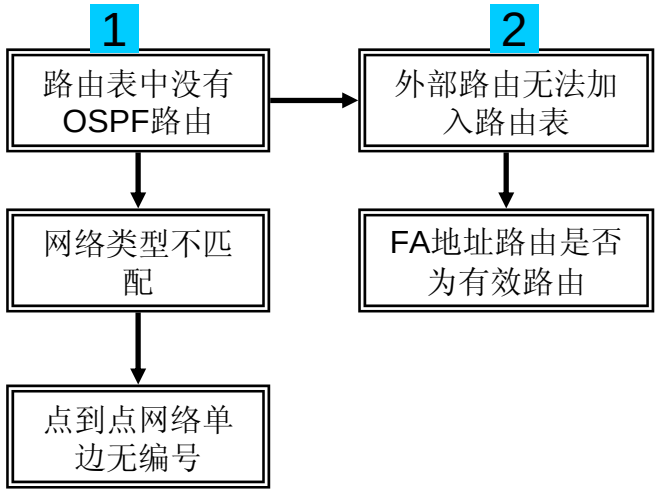
OSPF路由无法通告故障排查（1）



OSPF路由无法通告故障排查（2）

- **OSPF无法通告从地址的路由**
  - → 主从地址必须属于相同区域
- **ABR无法通告路由**
  - → 区域不允许接收汇总路由
    - OSPF的区域为完全存根区域或完全NSSA区域
  - → ABR与骨干区域隔离
    - ABR相连的区域必须有一个是骨干区域
  - → OSPF骨干区域分割
    - 如果OSPF的骨干区域分割，ABR可能无法生成全部的区域间路由
- **无法通告外部路由**
  - → 区域不允许接收外部路由
  - → NSSA区域存在设置错误的ABR
    - NSSA区域存在配置错误的ABR而且其Router ID较大

OSPF路由无法加入路由表故障排查（1）



OSPF路由无法加入路由表故障排查（2）

- 路由表没有OSPF路由
  - → OSPF网络类型不匹配
    - 如果OSPF邻居两边的网络类型设置不匹配，则数据库中网络类型不匹配，OSPF不会在路由表中添加路由
  - → 点到点网络单边无编号
    - 有编号和无编号接口的链路数据值不匹配，导致了OSPF数据库中的不一致，因此不会在OSPF路由表中添加路由。
- OSPF外部路由无法加入路由表
  - → 转发地址不能通过OSPF内部路由达到
    - OSPF外部路由中会携带转发地址信息，如果该转发地址非零，那么OSPF必须能够通过区域内或区域间路由到达该转发地址，否则该外部路由不会加入OSPF路由表。

SPF重复计算故障排查

- **链路抖动引起SPF重复计算**
  - → 链路抖动，导致区域内的路由器重新运行SPF算法
- **Router ID重复引起SPF重复计算**
  - → Router ID重复，将会导致OSPF拓扑数据库处于混乱状态，SPF频繁计算

OSPF排障相关命令介绍（1）

- **display ospf brief**

  → OSPF路由选择进程的概要信息

```
[RTA]display ospf brief

                OSPF Process 1 with Router ID 150.1.1.1
                    OSPF Protocol Information

 RouterID: 150.1.1.1
 Spf-schedule-interval: 5
 Routing preference: Inter/Intra: 10 External: 150
 Default ASE parameters: Metric: 1 Tag: 1 Type: 2
 SPF computation count: 0
 Area Count: 0    Nssa Area Count: 0
```

OSPF排障相关命令介绍（2）

- **display ospf interface**

  → OSPF相关的接口信息

```
[H3C]display       ospf interface

            OSPF Process 1 with Router ID 3.3.3.3
                      Interfaces
Area: 0.0.0.0
   IP Address     Type   State    Cost   Pri    DR          BDR
   1.1.1.2        NBMA  DR       1562   1      1.1.1.2      1.1.1.1


[H3C ] display ospf interface serial 1/0

 OSPF Process 1 with Router ID 150.1.1.1
                      Interfaces
    Interface: 150.1.1.1 (Serial0/0) --> 150.1.1.2
      Cost: 1562 State: PtoP    Type: PointToPoint
      Priority: 1
      Timers: Hello 10, Dead 40, Poll 40, Retransmit 5, Transmit Delay 1
```

OSPF排障相关命令介绍（3）

- **display ospf peer**

  → 显示OSPF邻居信息

```
<RTD>display ospf peer

            OSPF Process 1 with Router ID 3.3.3.3
                    Neighbors

 Area 0.0.0.0 interface 1.1.1.2(Serial1/0)'s neighbor(s)
 RouterID: 2.1.1.1        Address: 1.1.1.1
     State: Full  Mode: Nbr is Slave  Priority: 1
     DR: 1.1.1.2  BDR: 1.1.1.1
     Dead timer expires in 103s
     Neighbor has been up for 04:41:32
```

OSPF排障相关命令介绍（4）

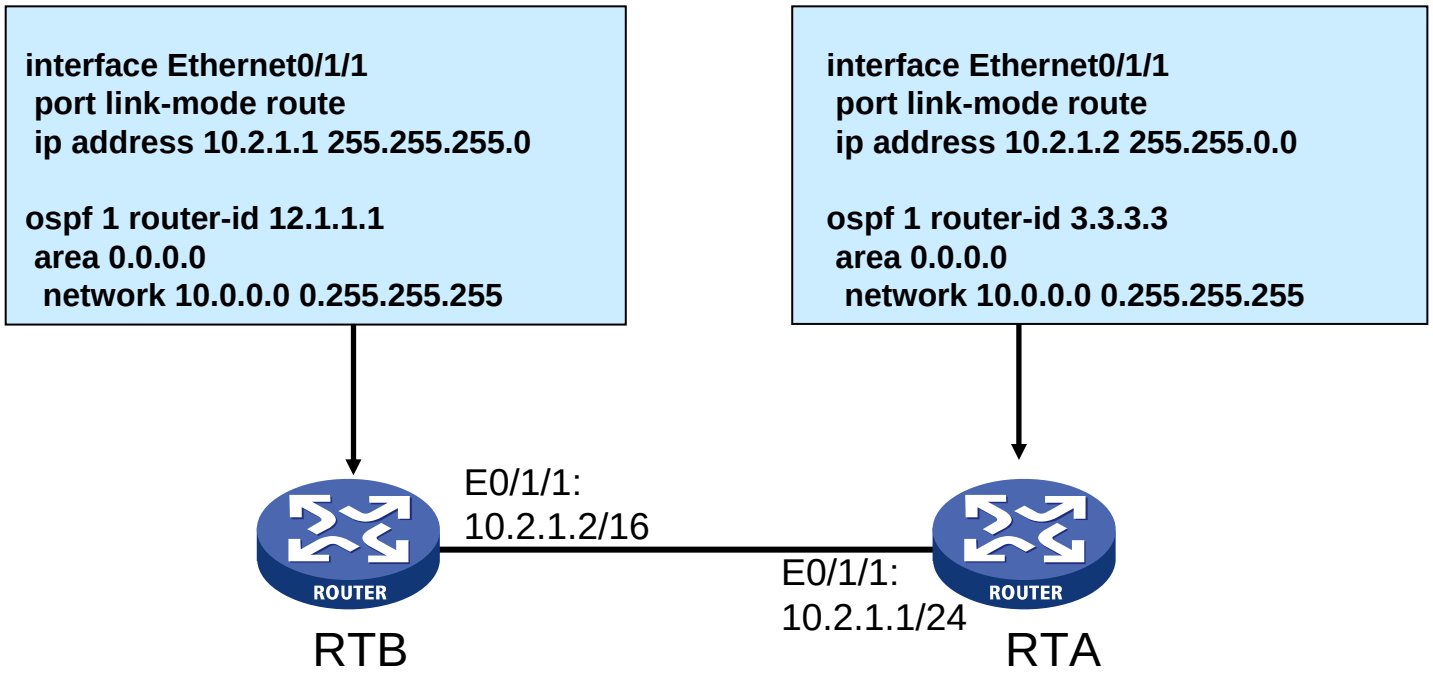
- **display ospf error**

  → 显示OSPF错误信息

```
[H3C] display ospf error

  OSPF Process 1 with Router ID 192.168.80.100
OSPF Packet Error Statistics
0 : OSPF Router ID confusion                0 : OSPF bad packet
0 : OSPF bad version                        0 : OSPF bad checksum
0 : OSPF bad area ID                        0 : OSPF drop on unnumber interface
0 : OSPF bad virtual link                   0 : OSPF bad authentication type
0 : OSPF bad authentication key             0 : OSPF packet too small
0 : OSPF Neighbor state low                 0 : OSPF transmit error
0 : OSPF interface down                     0 : OSPF unknown neighbor
0 : HELLO: Netmask mismatch                 0 : HELLO: Hello timer mismatch
0 : HELLO: Dead timer mismatch              0 : HELLO: Extern option mismatch
0 : HELLO: NBMA neighbor unknown            0 : DD: MTU option mismatch
0 : DD: Unknown LSA type                    0 : DD: Extern option mismatch
0 : LS ACK: Bad ack                         0 : LS ACK: Unknown LSA type
0 : LS REQ: Empty request                   0 : LS REQ: Bad request
0 : LS UPD: LSA checksum bad                0 : LS UPD: Received less recent LSA
0 : LS UPD: Unknown LSA type
```

OSPF典型案例一（1）



- 故障现象

→ 两台路由器之间的OSPF邻居关系无法建立

OSPF典型案例一（2）

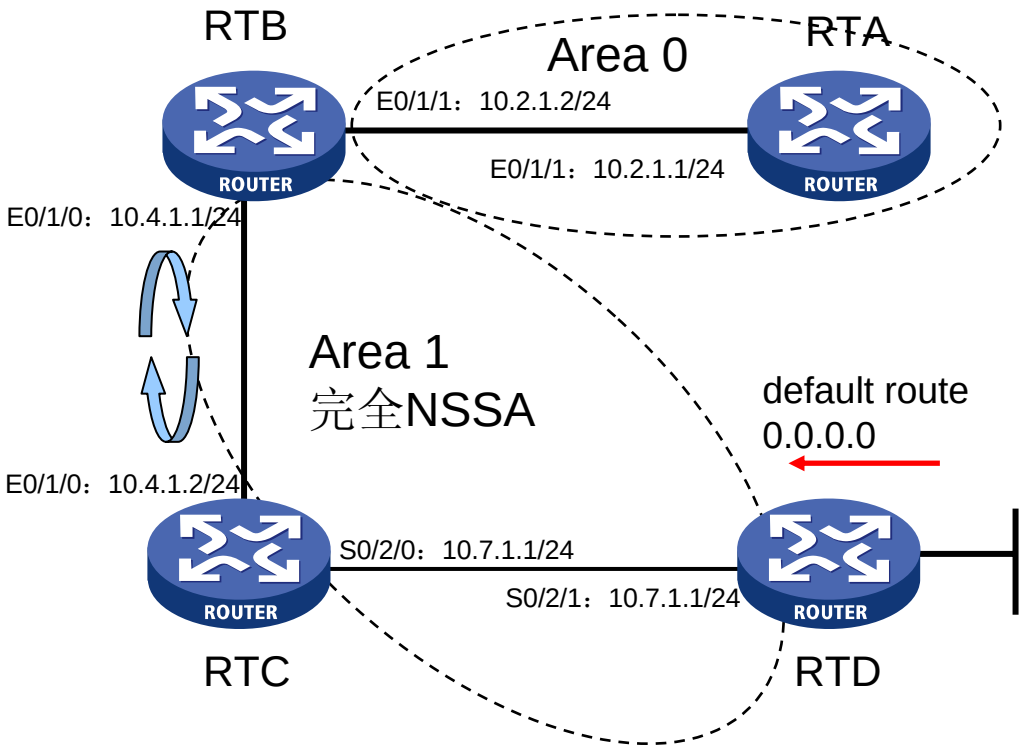
- 排障过程
  - → 在RTA上Ping RTB，可达
  - → 在RTA上用debugging ospf packet查看调试信息

```
*Dec 20 09:56:17:31 2008 RTA RM/6/RMDEBUG:OSPF 1: RECV Packet.
*Dec 20 09:56:17:46 2008 RTA RM/6/RMDEBUG:Source Address: 10.2.1.2
*Dec 20 09:56:17:46 2008 RTA RM/6/RMDEBUG:Destination Address: 224.0.0.5
*Dec 20 09:56:17:46 2008 RTA RM/6/RMDEBUG:Ver# 2, Type: 1, Length: 44.
*Dec 20 09:56:17:62 2008 RTA RM/6/RMDEBUG:Router: 3.3.3.3, Area: 0.0.0.0, Checksum:
60053.
*Dec 20 09:56:17:62 2008 RTA RM/6/RMDEBUG:AuType: 00, Key(ascii): 0 0 0 0 0 0 0 0.
*Dec 20 09:56:17:62 2008 RTA RM/6/RMDEBUG:Hello: netmask mismatch.
```

以上信息表明，RTA收到了RTB发出的OSPF Hello报文，但是由于RTB的接口掩码与本地接口的掩码不匹配，导致OSPF无法完成邻居建立过程

- 解决方案
  - → 修改RTB的掩码与RTA的相同

OSPF典型案例二（1）



- 故障现象
  - → OSPF多区域组网。配置完成后发现RTB和RTC均无法访问外部区域；
  - → 通过Trace命令进行路径检查，发现在RTB和RTC之间形成环路。

OSPF典型案例二（2）

- RTA上的配置：

```
ospf 1 router-id 12.1.1.1
 area 0.0.0.0
  network 10.0.0.0 0.255.255.255
```

- RTB上的配置：

```
ospf 1 router-id 3.3.3.3
 area 0.0.0.0
  network 10.2.1.2 0.0.0.0
 area 0.0.0.1
  network 10.4.1.1 0.0.0.0
  nssa no-summary
```

- RTC上的配置：

```
ospf 1 router-id 11.1.1.1
 area 0.0.0.1
  network 10.4.1.2 0.0.0.0
  network 10.7.1.1 0.0.0.0
  nssa no-summary
```

- RTD上的配置：

```
ospf 1 router-id 13.1.1.1
 area 0.0.0.1
  network 10.7.1.2 0.0.0.0
  nssa default-route-advertise
#
 ip route-static 0.0.0.0 0.0.0.0 NULL0
preference 5
```

OSPF典型案例二（3）

- 排障过程

→查看RTC和RTB的路由表，发现RTC的缺省路由指向RTB而RTB的缺省路由指向RTC，报文在RTB和RTC之间发生环路。

- 原因分析

→RTB上缺省路由的产生

-由于RTD是ASBR，所以向NSSA区域发布第7类缺省路由，RTB将此缺省路由加入路由表，下一跳指向RTC

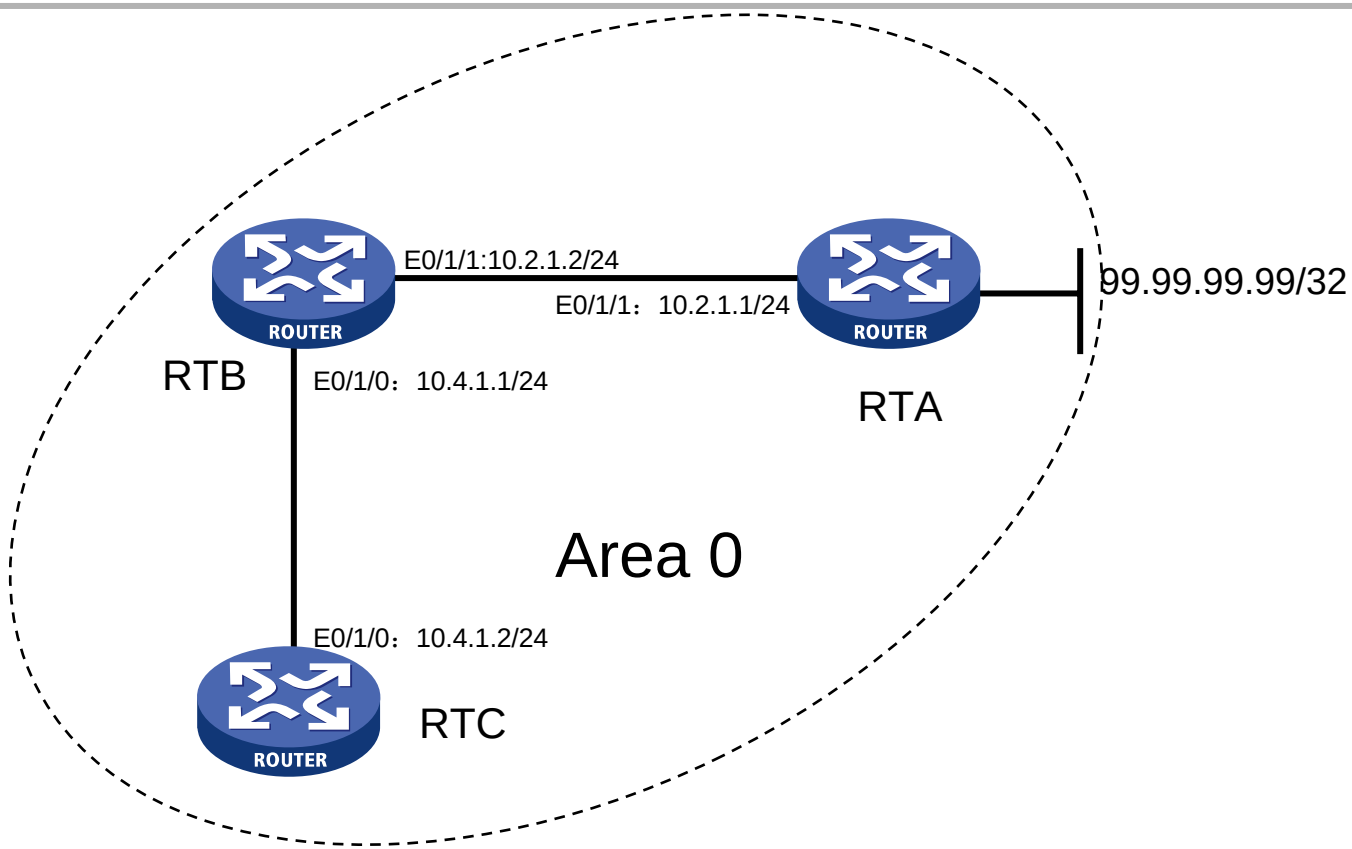
→RTC上缺省路由的产生

-RTB是ABR，因此RTB向Area1发布第3类缺省路由，此路由优先级高于第7类缺省路由，所以RTC将此缺省路由加入路由表，下一跳指向RTB

- 解决方案

→调整RTD的NSSA配置，去掉区域下发布NSSA缺省路由的命令nssa default-route-advertise后，在RTB和RTC之间的路由环路消失，网络恢复正常

OSPF典型案例三（1）



- 故障现象

→ 在RTB上访问外部区域的99.99.99.99时，丢包严重

OSPF典型案例三（2）

- RTA上的配置：

```
ospf 1 router-id 11.1.1.1
 import-route static
 area 0.0.0.0
  network 10.0.0.0 0.255.255.255
#
 ip route-static 10.9.1.0 255.255.255.0
NULL0
 ip route-static 99.99.99.99
255.255.255.255 NULL0
```

- RTB上的配置：

```
ospf 1 router-id 3.3.3.3
 area 0.0.0.0
  network 10.0.0.0 0.255.255.255
```

- RTC上的配置：

```
interface loop0
ip address 11.1.1.1 255.255.255.255
ospf 1
 area 0.0.0.0
  network 10.0.0.0 0.255.255.255
```

OSPF典型案例三（3）

- 排障过程
  - → 多次查看RTB的路由表，发现99.99.99.99这条路由不断震荡，时而出现在路由表中时而又从路由表中消失。
  - → 在RTB上通过display ospf brief 查看OSPF统计信息发现OSPF SPF计算次数正在快速增加。

```
<RTB>display ospf brief
。。。。。。
SPF Computation Count: 234
```

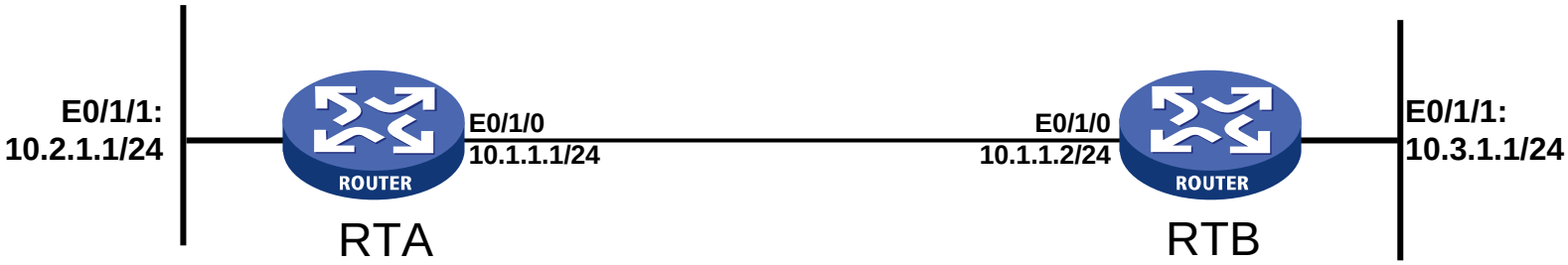
```
<RTB>display ospf brief
。。。。。。
SPF Computation Count: 237
```

  - → 在RTB上查看LSDB，发现Router 11.1.1.1 的LSA Age很小，而且LSA 99.99.99.99正在反复被老化。

OSPF典型案例三（4）

- 排障过程
  - → 通过上面的检查与分析，可以基本判断网络中存在Router ID冲突，导致路由异常以及大量的SPF重复计算发生。
  - → 检查三台路由器的配置发现RTC上配置了一个IP地址为11.1.1.1的loopback接口。而由于RTC没有指定Router ID，因此设备自动选取了loopback接口的地址作为自己的Router ID。
- 解决方案
  - → 调整RTC的Router ID后，网络恢复正常。

OSPF典型案例四（1）



- 故障现象

  → 路由器无法学习到对方的OSPF路由

OSPF典型案例四（2）

- RTA上的配置：

```
interface Ethernet0/1/0
 port link-mode route
 ip address 10.1.1.1 255.255.255.0
#
interface Ethernet0/1/1
 port link-mode route
 ip address 10.2.1.1 255.255.255.0
#
ospf 1
 area 0.0.0.0
  network 10.1.1.1 0.0.0.0
  network 10.2.1.1 0.0.0.0
```

- RTB上的配置：

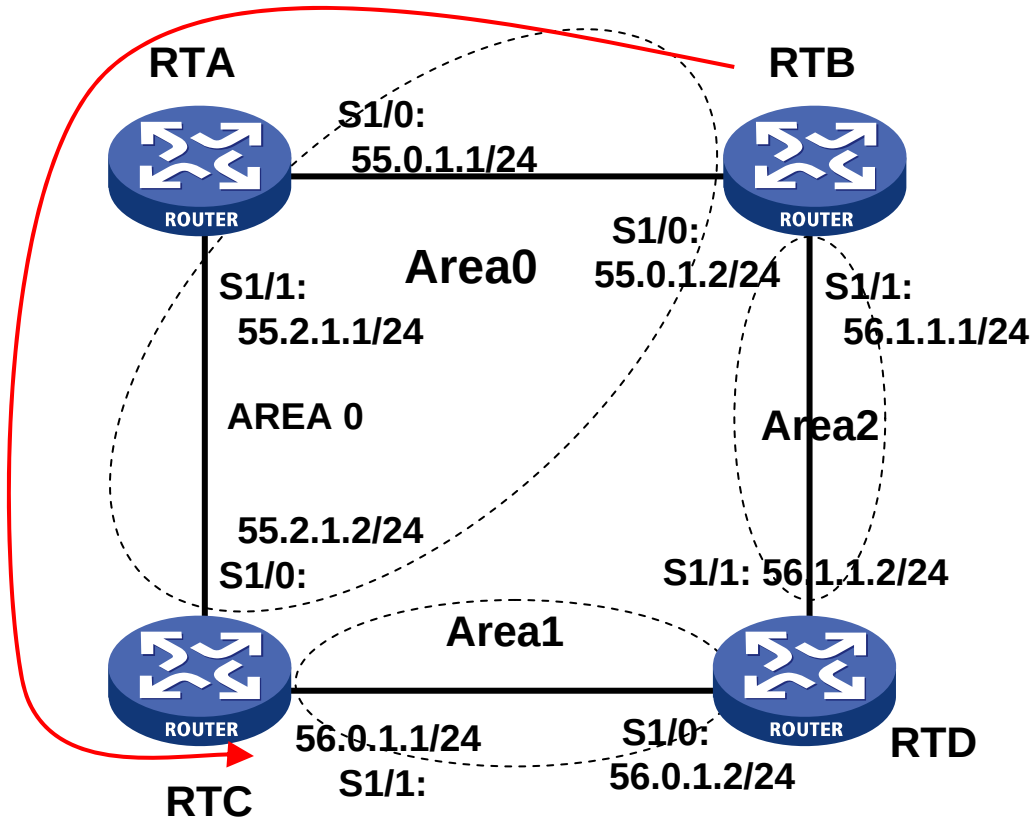
```
interface Ethernet0/1/0
 port link-mode route
 ip address 10.1.1.2 255.255.255.0
 ospf network-type p2p
#
interface Ethernet0/1/1
 port link-mode route
 ip address 10.3.1.1 255.255.255.0
#
ospf 1
 area 0.0.0.0
  network 10.1.1.2 0.0.0.0
  network 10.3.1.1 0.0.0.0
```

OSPF典型案例四（3）

- 排障过程
  - → 使用display ospf peer命令发现邻居关系正常建立
  - → 使用display ospf interface 命令发现接口网络类型不匹配。
- 解决方案
  - → 将RTA的E0/1/0接口的OSPF网络类型调整为P2P，与RTB保持一致，路由学习正常，问题解决。

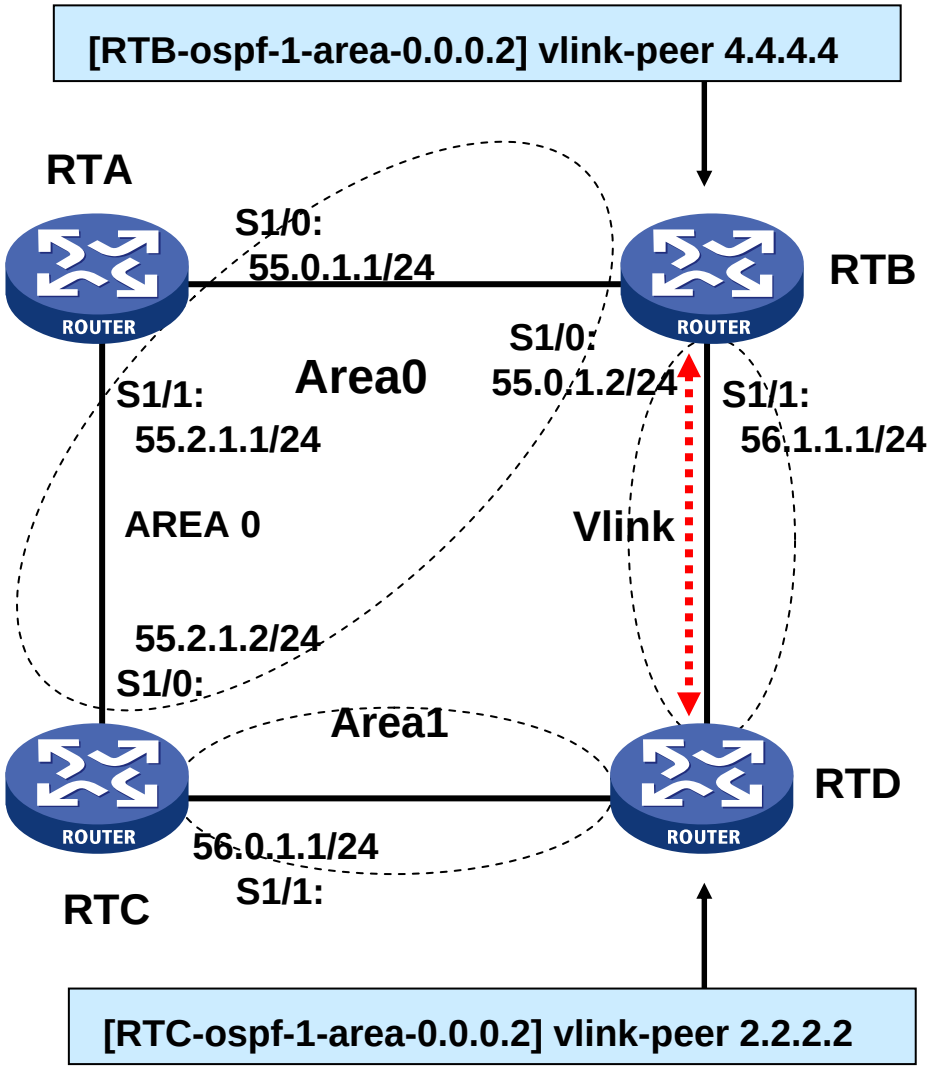
OSPF典型案例五（1）

- 故障现象
  - → 观察RTB路由表，发现从RTB到56.0.1.0/24的路径为RTB->RTA->RTC-> 56.0.1.0
  - → 观察RTC路由表，发现RTC到56.1.1.0/24的路径为RTC->RTA->RTB-> 56.1.1.0



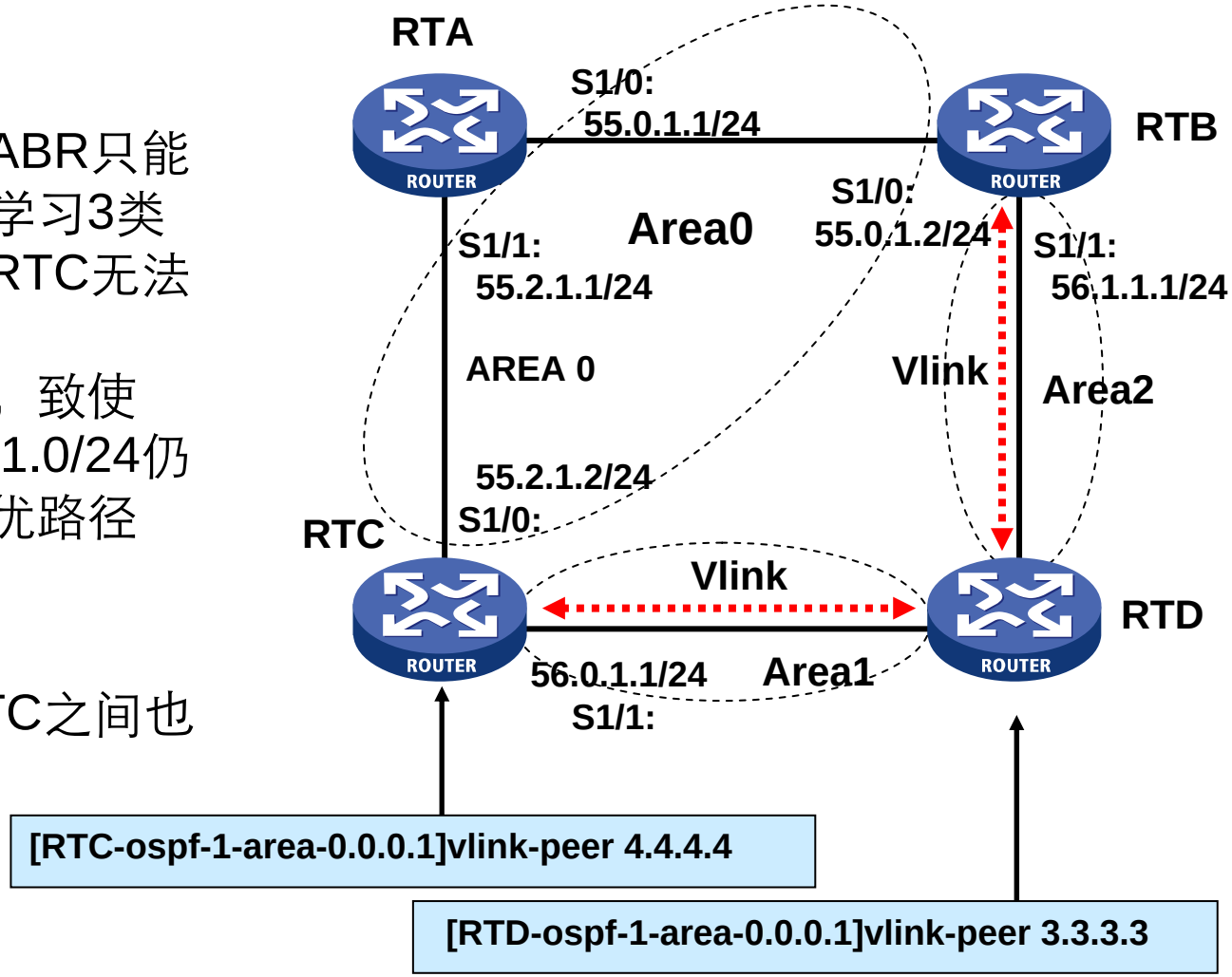
OSPF典型案例五（2）

- 排障过程
  - → 故障原因是RTD没有连接到Area0，不是ABR，不能在区域间传递路由
  - → 在RTB和RTD之间配置虚连接，使RTD成为ABR，从而使RTD与RTB之间能够交换域间路由信息
  - → 查看路由表，发现RTB到56.0.1.0/24采用了最优路由。即路径为RTB->RTD->56.0.1.0



OSPF典型案例五（3）

- 排障过程
  - → 但是，因为ABR只能从骨干区域学习3类路由，所以RTC无法从RTD学习56.1.1.0/24，致使RTC到56.1.1.0/24仍然选择了次优路径
- 解决方案
  - → 在RTD与RTC之间也配置虚连接



目录

- RIP故障诊断和排除
- OSPF故障诊断和排除
- BGP故障诊断和排除

BGP故障排除

- BGP路由协议简介
- BGP故障排查基础知识
- BGP故障排查基本方法
- BGP典型案例分析

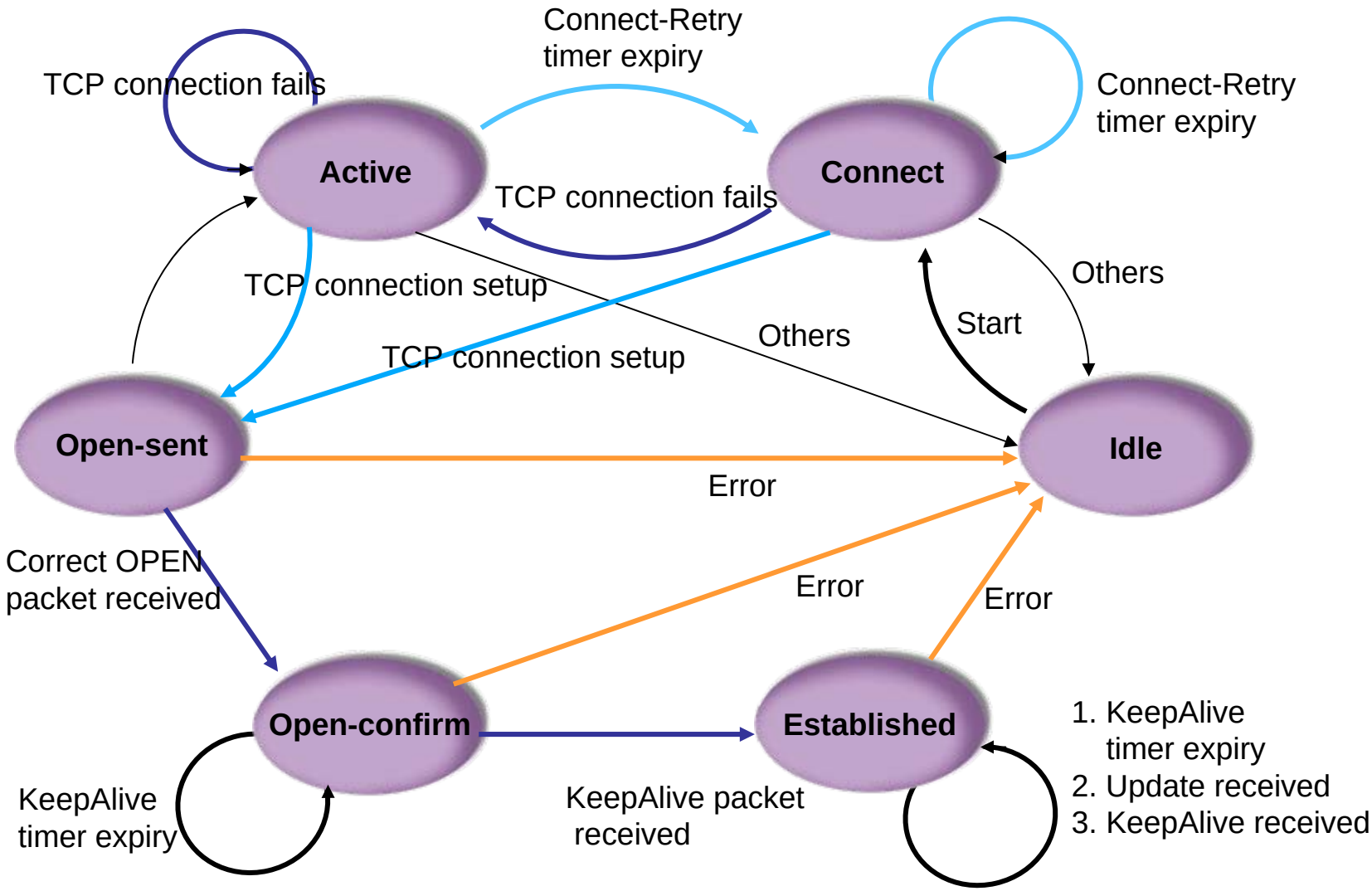
BGP协议简介

- BGP是外部路由协议，用来在AS之间传递路由信息
- 是一种距离矢量的路由协议，从设计上避免了环路的发生
- 为路由附带属性信息
- 传送协议：TCP，端口号179
- 支持CIDR（无类别域间选路）
- 路由更新：只发送增量路由
- 丰富的路由过滤和路由策略

BGP报文

- **Open报文**
  - → 协商运行参数，包括版本、AS号、保持计时器、ID、验证等
- **Update报文**
  - → 传送路由增量信息，如更新或撤销路由
- **Keepalive报文**
  - → 防止保持定时器超时
- **Notification报文**
  - → 检测到错误后发送。接收者会中断BGP连接

BGP状态机


Connect-Retry
timer expiry
TCP connection fails
Connect-Retry
timer expiry
Active
Connect
TCP connection fails
Others
TCP connection setup
Start
Others
TCP connection setup
Open-sent
Idle
Error
Correct OPEN
packet received
Error
Error
Open-confirm
Established
1. KeepAlive
timer expiry
2. Update received
3. KeepAlive received
KeepAlive
timer expiry
KeepAlive packet
received

BGP通告路由的策略

- BGP Speaker只选最优的给自己使用
- BGP Speaker只把自己使用的路由通告给相邻体
- BGP Speaker从EBGP获得的路由会向它所有BGP相邻体通告（包括EBGP和IBGP）
- BGP Speaker从IBGP获得的路由不向它的IBGP相邻体通告
- BGP Speaker 从IBGP获得的路由是否通告给它的EBGP相邻体要依IGP和BGP同步的情况来决定
- 连接一建立，BGP Speaker将把自己所有BGP路由通告给新相邻体

BGP选择路由的顺序

- 首先丢弃下一跳（NEXT_HOP）不可达的路由；
- 优选Preferred-value值最大的路由；
- 优选本地优先级（LOCAL_PREF）最高的路由；
- 优选本路由器始发的路由；
- 优选AS路径（AS_PATH）最短的路由；
- 依次选择ORIGIN类型为IGP、EGP、Incomplete的路由；
- 优选MED值最低的路由；
- 优选选择从EBGP学到的路由；
- 优选下一跳IGP Cost值最低的路由；
- 优选CLUSTER_LIST长度最短的路由；
- 优选ORIGINATOR_ID 最小的路由；
- 优选Router ID最小的路由器发布的路由。

BGP路由环路避免

- 在路由的AS-Path属性中记录着所有途经的AS，BGP路由器将丢弃收到的任何一条带有本地AS的路由，这就避免了AS之间的环路；
- 从本AS内部得到的路由不再在本AS内部转发，从而避免了AS内部的环路。

大规模BGP网络解决办法

- **BGP路由反射器**
  - → 一个或一组路由器作为IBGP会话的中心点。其它路由器都与中心点建立IBGP邻居关系，依靠中心点的反射进行路由交换
  - → 反射原则
    - 如果路由更新是从非客户机收到的，仅反射给客户机。
    - 如果路由更新是从客户机收到的，反射给所有非客户机以及客户机
    - 如果路由更新是从EBGP相邻体收到的，反射给所有的客户机和非客户机
- **BGP联盟**
  - → 一个AS可以被分为多个子AS，子AS内使用IBGP全连接，子AS之间以及联盟本身与外部AS之间使用特殊的EBGP连接

BGP路由协议常见故障

- BGP邻居无法建立
- 已经建立好的邻居关系又失败了
- BGP路由无法发布
- BGP路由无法接收
- 路由选择不一致

BGP排障基本步骤（1）

- **BGP邻居无法建立问题**
  - → 查看是否配置了正确的邻居、AS号
  - → 检查邻居能否Ping通，由于一台路由器可能有多个接口能够到达对端，应使用扩展Ping命令检查，指定Ping包的源IP地址为建立邻居关系的地址。如果不能Ping通检查IGP路由表中是否有邻居的路由。
  - → 检查是否配置了禁止TCP端口179的ACL，如果有，取消对179端口的禁止。
  - → 如果使用loopback接口建邻居，查看是否配置了peer connect-interface命令
  - → 如果是EBGP邻居，检查和对端建邻居的接口是否up；
  - → 如果是EBGP邻居，且EBGP连接在物理上不是直连的，检查是否配置了peer ebgp-max-hop。默认情况下，EBGP邻居的TTL被置为1，如果不是直连，必须配置peer group-name ebgp-max-hop。
  - → 通过debugging bgp open报文查看是否和对端的能力协商不通过

BGP排障基本步骤（2）

- 已经建立好的邻居关系又失败了
  - → MTU问题
    - 使用扩展的Ping命令检查是否存在MTU问题，ping –s可指定Ping包的包长。
  - → QoS问题
    - 检查是否在接口上设置了流量整形或物理接口限速。
  - → MTU和QoS设置不当可能导致大的Update报文被丢弃，由于TCP的重传机制，当发送多个大的Update报文时，可能产生大量等待重传的Update报文，从而抑制了keepalive报文的正常发送，当连续收不到keepalive报文时，BGP认为邻居已经Down。
  - → 网络拥塞问题
    - 网络拥塞可能导致Keepalive报文收发失常，邻居状态不断改变；另外，如果到达邻居的路由是通过IGP（如OSPF）发现的，网络拥塞可能导致路由丢失，从而使邻居间的连接中断。

BGP排障基本步骤（3）

- **BGP路由无法发布**
  - → 使用display bgp peer命令查看BGP邻居是否已经建立；
  - → 查看路由表中是否存在所需的IGP路由。
    - BGP自己无法生成路由，只能由IGP学习路由，然后BGP再引入。使用network命令引入路由时，在路由表中一定要存在该路由才能够引入。而且network发布的路由必须与路由表中的路由精确匹配才能发布，即路由的掩码长度要匹配。
  - → 查看BGP是否配置了路由引入，将IGP路由引入到BGP中；
  - → 查看BGP是否配置了路由策略将路由过滤掉；
  - → IBGP对等体没有全连接造成路由无法发布。BGP规定从IBGP对等体收到的路由信息不能向另一个IBGP对等体发送。

BGP排障基本步骤（4）

- **BGP路由接收问题**
  - → 检查路由信息中的AS_PATH是否包含本路由器的AS，如果路由信息中AS_PATH中包含本路由器的AS，则该路由被丢弃
  - → 检查路由信息中的cluster-list是否包含本路由器的cluster-id，如果是的话，该路由被丢弃。
  - → 检查路由信息中的originator-id是否包含本路由器的originator-id，如果是的话，该路由被丢弃。
  - → 查看是否是由于路由迭代的原因造成的。迭代后下一跳不可达的路由不能加入路由表。
  - → 查看路由表中是否存在其他路由和BGP路由相同，在路由的优先级中，BGP的优先级最低，如果有其他路由存在，BGP路由不会生效。
  - → 查看BGP的配置，看是否配置了入口路由策略将BGP路由信息过滤掉。

BGP排障基本步骤（5）

- 路由选择不一致问题
  - → 查看选路策略中是否需要比较MED。缺省情况下，只比较来自同一AS的路由的MED值，如果要比较来自不同AS的路由的MED值，应使用命令compare-different-as-med。
  - → 查看是否启动了同步。IGP路由表中不存在的IBGP路由不能作为最佳路由，即使它具有高的本地优先级，也就是说，未经同步的路由不能被选为最佳路由。

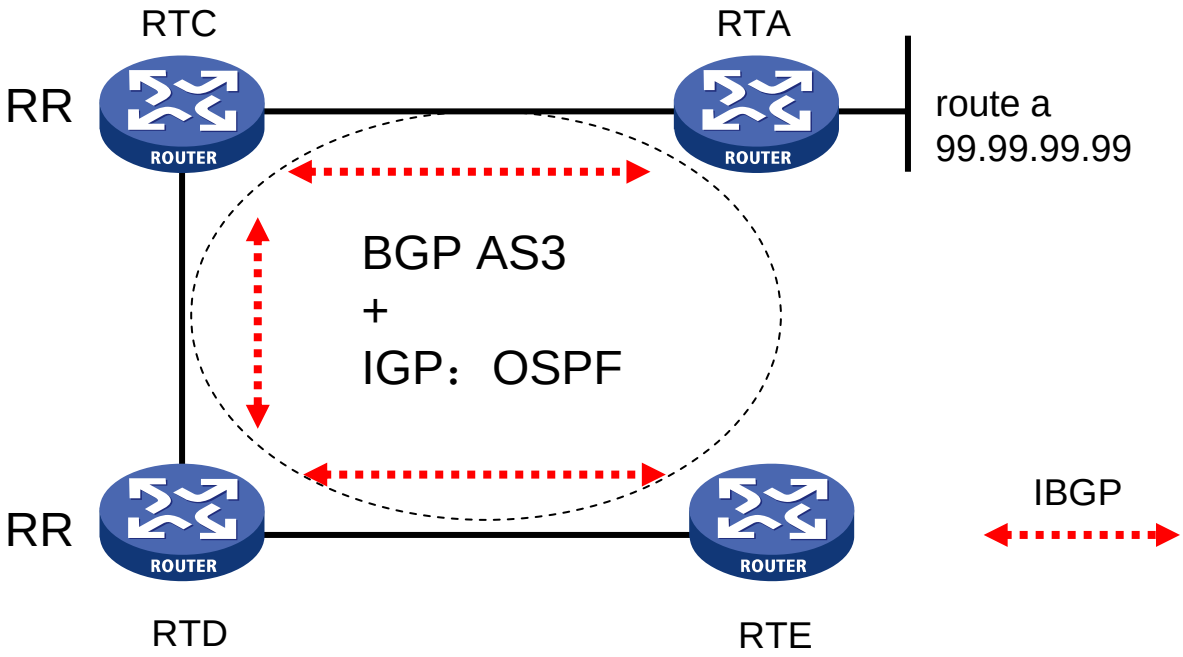
BGP排障相关命令介绍 （1）

- **display bgp peer**
  - → 显示对等体信息
- **display bgp routing-table**
  - → 查看bgp路由表信息
- **display bgp routing-table as-path-acl**
  - → 查看匹配过滤列表的路由信息
- **display ip as-path**
  - → 查看配置的AS_PATH列表规则

BGP排障相关命令介绍（2）

- **display acl**
  - → 查看配置的访问控制列表规则
- **display route-policy**
  - → 查看配置的Routing policy规则
- **debugging bgp all**
  - → 打开BGP所有报文调试信息开关
- **debugging bgp event**
  - → 打开BGP事件调试信息开关

BGP典型案例一（1）



- 网络描述
  - → RTC与RTD配置为路由反射器，RTA和RTE作为反射客户端。RTA与RTC建立IBGP邻居、RTC与RTD建立IBGP邻居、RTD与RTE建立IBGP邻居。
- 故障现象
  - → RTE无法学习到路由99.99.99.99

BGP典型案例一（2）

- RTA上的配置：

```
bgp 3
 network 99.99.99.99 255.255.255.255
 undo synchronization
 peer 3.3.3.3 as-number 3
 peer 3.3.3.3 connect-interface
LoopBack0
#
ospf 1
 area 0.0.0.0
  network 10.0.0.0 0.255.255.255
  network 12.1.1.1 0.0.0.0
```

- RTC上的配置：

```
bgp 3
 reflector cluster-id 3.3.3.3
 undo synchronization
 peer 12.1.1.1 as-number 3
 peer 11.1.1.1 as-number 3
 peer 12.1.1.1 reflect-client
 peer 12.1.1.1 connect-interface
LoopBack0
 peer 11.1.1.1 connect-interface
LoopBack0
#
ospf 1
 area 0.0.0.0
  network 10.0.0.0 0.255.255.255
  network 3.3.3.3 0.0.0.0
```

BGP典型案例一（3）

- RTD上的配置：

```
bgp 3
 reflector cluster-id 3.3.3.3
 undo synchronization
 peer 13.1.1.1 as-number 3
 peer 3.3.3.3 as-number 3
 peer 13.1.1.1 reflect-client
 peer 13.1.1.1 connect-interface
LoopBack0
 peer 3.3.3.3 connect-interface
LoopBack0
#
ospf 1
 area 0.0.0.0
  network 10.0.0.0 0.255.255.255
  network 11.1.1.1 0.0.0.0
```

- RTE上的配置：

```
bgp 3
 undo synchronization
 peer 11.1.1.1 as-number 3
 peer 11.1.1.1 connect-interface
LoopBack0
#
ospf 1
 area 0.0.0.0
  network 10.0.0.0 0.255.255.255
  network 13.1.1.1 0.0.0.0
ospf 1 router-id 11.1.1.1
 import-route static
 area 0.0.0.0
  network 10.0.0.0 0.255.255.255
#
 ip route-static 10.9.1.0 255.255.255.0
NULL0
 ip route-static 99.99.99.99
255.255.255.255 NULL0
```

BGP典型案例一（4）

●排障过程

→在RTE上查看路由表，发现没有收到99.99.99.99的BGP路由。

→在RTD上打开debugging bgp update开关并执行reset bgp all并查看。发现信息如下：

```
Dec 20 05:43:24:922 2008 RTD RM/6/RMDEBUG:

BGP. : Error identified while receiving UPDATE message
from the peer 3.3.3.3 and ignored

Reason: Received CLUSTERLIST Value equal to ClusterID of
the speaker.
```

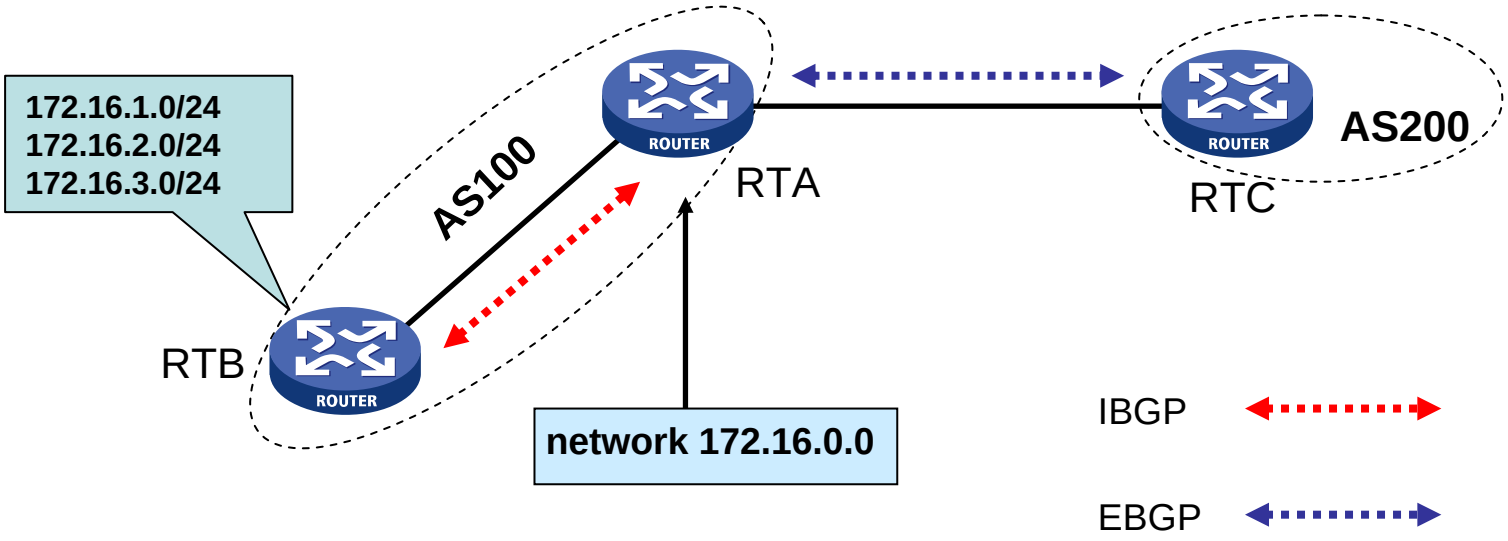
至此原因找到了，是由于RTD与RTC属于同一个反射组，其cluster id相同，所以RTD不会把从RTC收到的更新反射给RTE

●解决方案

→在使用路由反射器的时候，反射客户端必须与反射组的所有反射器建立IBGP邻居

→调整RTA的BGP配置，增加RTA与RTD的IBGP连接后，RTE可以正常学习RTA发出的路由。

BGP典型案例二（1）



- 故障描述
  - → RTA和RTB属于AS100，通过OSPF交换IGP路由，RTB通过OSPF发布本地直连地址路由，RTA与RTC建立EBGP邻居，并通过network 172.16.0.0命令将RTB上的直连路由发布出去。
  - → 但是RTC没有学习到相应的路由信息

BGP典型案例二（2）

- RTA上的配置：

```
ospf 1
  area 0.0.0.0
  net 150.1.1.0 0.0.0.255
bgp 100
    network 172.16.0.0
    group as200 external
    peer 133.1.1.2 group as200 as-number 100
```

- RTB上的配置：

```
interface ethernet 0/1/0
    ip address 172.16.1.1 255.255.255.0
  #
  interface ethernet 0/1/1
    ip address 172.16.2.1 255.255.255.0
  #
  interface ethernet 0/1/2
    ip address 172.16.3.1 255.255.255.0
ospf 1
  area 0
  net 150.1.1.0 0.0.0.255
  net 172.16.0.0 0.0.255.255
```

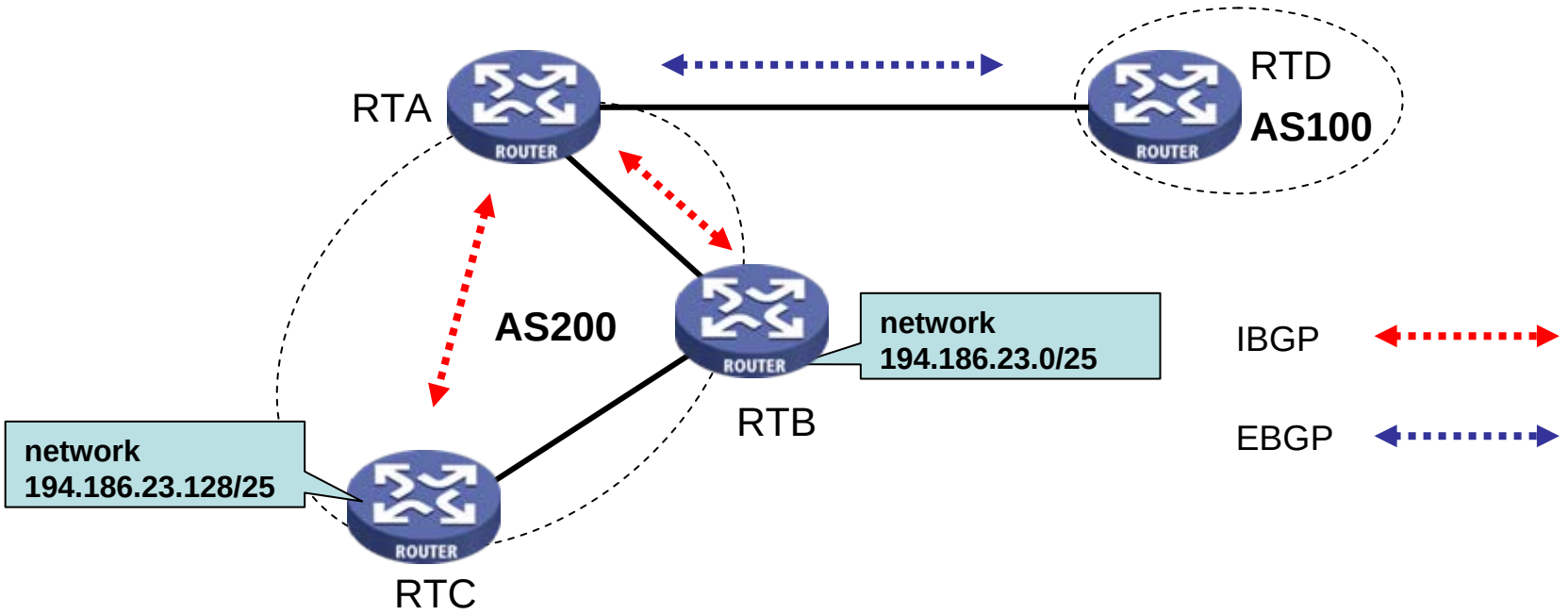
- RTC上的配置：

```
bgp 200
 undo synchronization
 group as100 external
 peer 133.1.1.1 group as100 as-number 100
```

BGP典型案例二（3）

- 排障过程：
  - → 在RTA上用display ip routing-table查看本地路由表，可以看到三条OSPF路由已经存在于本地路由表中。
  - → 用display bgp routing-table命令查看BGP路由表，没有发现任何BGP路由信息。
  - → 打开debugging，没有看到RTA向RTC发送update报文的调试信息。由此可以推断出应该是BGP路由发布的问题。
  - → 在RTA上输入network 172.16.1.0 mask 255.255.255.0命令，在BGP路由表中出现了路由172.16.1.0/24，RTC上也学习到了这条路由。
  - → 使用BGP的network命令发布路由时，必须保证本地路由表中待发布路由的前缀和掩码同network命令发布的路由完全匹配才能正常发布。
- 解决方案：
  - → 调整RTA上BGP的路由引入配置，用network命令发布对应的24位掩码网段，RTC上能够学习到相应的路由信息

BGP典型案例三（1）



- 故障现象
  - → RTA与RTB和RTC分别建立IBGP邻居关系，RTA与RTD建立EBGP邻居关系。RTC和RTB通过network发布25位掩码的明细路由，RTA将两条明细路由聚合为24位掩码后向IBGP和EBGP邻居发布。
  - → 网络194.186.23.128/25无法访问AS100，且RTA和RTB之间产生路由环路。

BGP典型案例三（2）

- RTA上的配置：

```
bgp 200
 aggregate 194.186.23.0 255.255.255.0
 undo synchronization
 group as100 external
 peer 133.1.1.2 group as100 as-number 100
 group as200 internal
 peer as200 next-hop-local
 peer 200.1.7.2 group as200
 peer 150.1.1.2 group as200
```

- RTB上的配置：

```
bgp 200
 network 194.186.23.0 255.255.255.128
 undo synchronization
 group as200 internal
 peer 200.1.7.1 group as200
```

- RTC上的配置：

```
bgp 200
 network 194.186.23.128 255.255.255.128
 undo synchronization
 group as200 internal
 peer 200.1.7.1 group as200
```

- RTD上的配置：

```
bgp 100
 import-route direct
 undo synchronization
 group as200 external
 peer 133.1.1.1 group as200 as-number 200
```

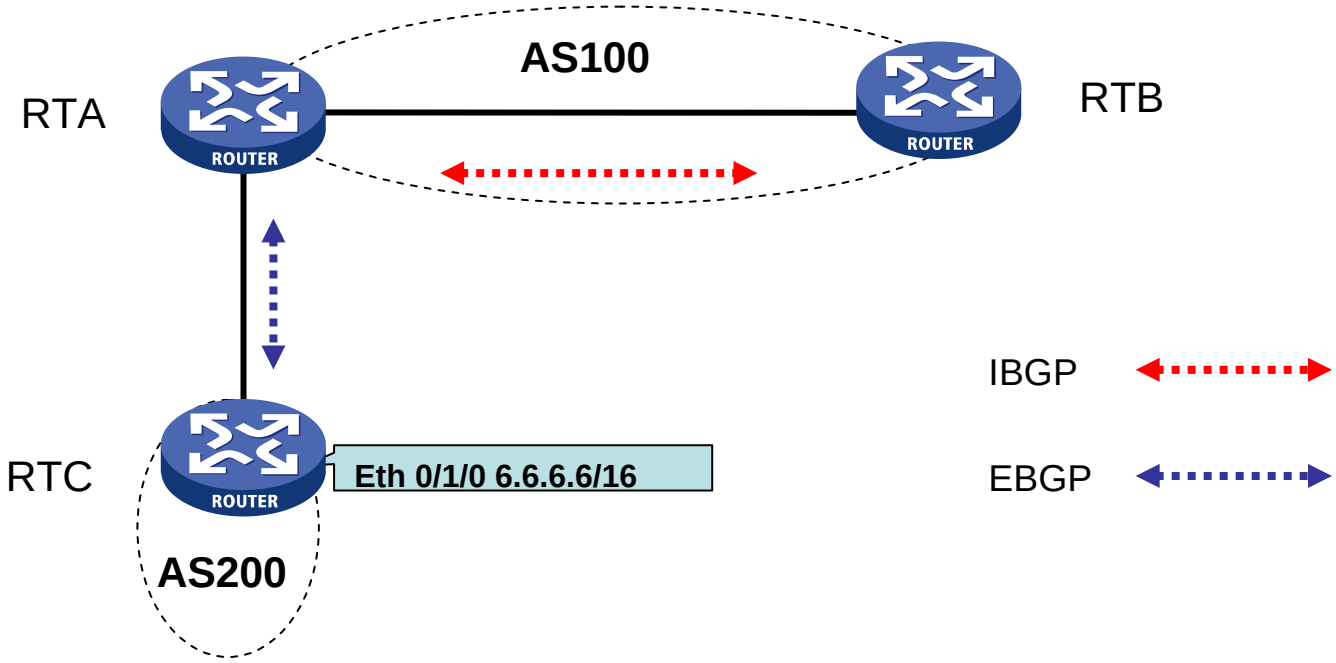
BGP典型案例三（3）

- 排障过程
  - → 在RTA上查看IP路由表及BGP路由表，发现路由194.186.23.0/25和194.186.23.128/25，指向RTB；
  - → 在RTB上查看路由表，发现有一条194.186.23.0/24的聚合路由，指回RTA。造成了路由环路。
- 原因分析
  - → 因为RTA与RTC之间是IBGP连接，所以RTA通过BGP可以学习到RTC的明细路由（194.186.23.0/25和194.186.23.128/25），因此将访问194.186.23.128/25的流量转发至RTB；
  - → 而RTB与RTA之间是IBGP连接，所以通过BGP学习到RTA发布的汇总路由信息，因此会把RTA发给RTB的流量再次发回RTA，在RTA和RTB形成环路。

BGP典型案例三（4）

- 解决方案
  - → 方案一：在RTB和RTC之间也建立IBGP连接，这样RTC发布的路由也可以被RTB学习到，到达194.186.23.128/25的流量可以被RTB正确转发。
  - → 方案二：如果RTB和RTC不参与域间选路，则不需要运行BGP，在RTA、RTB、RTC之间运行一种IGP，再通过RTA把学习到的路由通过BGP发布出去。
  - → 方案三：在RTB上配置一条到194.186.23.128/25的静态路由，下一跳指向RTC，使去往194.186.23.128/25的流量经过RTB和RTC之间的链路到达RTC。
- 建议和总结
  - → 在部署BGP的时候，建议所有的IBGP邻居间配置成全连接，以减少发生路由环路的可能。
  - → 引起路由环路的原因还有很多，如路由协议之间的引入设置不当、聚合路由配置不当、缺省路由配置不当等等，在组网方案的设计中应当注意，尽量避免路由环路。

BGP典型案例四（1）



- 故障现象
  - → RTA和RTB之间建立IBGP邻居关系，RTA和RTC之间建立EBGP邻居关系。RTC发布路由6.6.0.0/16给RTA，RTA把从EBGP邻居学来的路由发布给IBGP邻居RTB
  - → 在RTB上使用display bgp查看，发现学到的路由为无效路由

BGP典型案例四（2）

- RTA上的配置：

```
bgp 100
 undo synchronization
 group as100 internal
 peer 150.1.1.2 group as100
 group as200 external
 peer 12.110.150.2 group as200 as-number 200
```

- RTB上的配置：

```
bgp 100
 undo synchronization
 group as100 internal
 peer 150.1.1.1 group as100
```

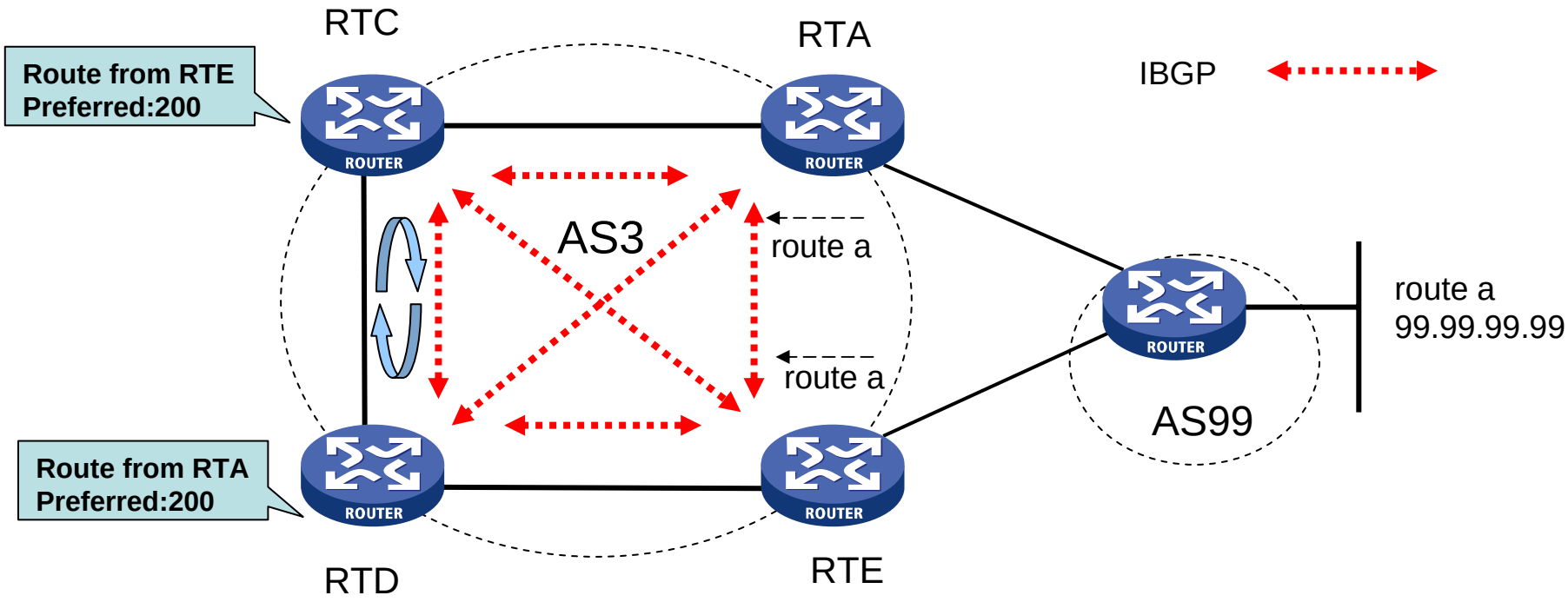
- RTC上的配置：

```
bgp 200
  network 6.6.0.0 255.255.0.0
  undo synchronization
  group as100 external
  peer 12.110.150.1 group as100 as-number 100
```

BGP典型案例四（3）

- 排障过程
  - → 查看RTB的BGP路由表，发现6.6.0.0/16的下一跳为12.110.150.2；
  - → 再查RTB的路由表，没有到达12.110.0.0/16网段的路由。
- 原因分析
  - → 在BGP中，对于从IBGP和EBGP学来的路由的下一跳的处理是不同的。EBGP邻居发布路由时会把路由的下一跳改为自己的出接口地址，而IBGP邻居在发布从EBGP学到的路由时默认情况下不会改变下一跳。
- 解决方案
  - → 在RTA上配置next-hop-local命令，将从EBGP邻居学到的路由下一跳改为自己的接口地址，这样就不会出现由于下一跳不可达导致的路由失效问题了。

BGP典型案例五（1）



- 故障现象
  - → 所有设备均属于同一AS，RTD设置Preferred-value属性优选RTA发布的路由，RTC设置Preferred-value属性优选RTE发布的路由；
  - → 发现外部路由99.99.99.99在RTC和RTD之间形成环路。

BGP典型案例五（2）

- RTA上的配置：

```
bgp 3
 import-route static
 undo synchronization
 group ibgp internal
 peer ibgp connect-interface LoopBack0
 peer 3.3.3.3 group ibgp
 peer 11.1.1.1 group ibgp
 peer 13.1.1.1 group ibgp
#
ospf 1
 area 0.0.0.0
  network 10.0.0.0 0.255.255.255
  network 12.1.1.1 0.0.0.0

 ip route-static 99.99.99.99
255.255.255.255 NULL0
```

- RTC上的配置：

```
bgp 3
 router-id 3.3.3.3
 undo synchronization
 group ibgp internal
 peer ibgp connect-interface LoopBack0
 peer 11.1.1.1 group ibgp
 peer 12.1.1.1 group ibgp
 peer 13.1.1.1 group ibgp
 peer 13.1.1.1 preferred-value 200
#
ospf 1
 area 0.0.0.0
  network 10.0.0.0 0.255.255.255
  network 3.3.3.3 0.0.0.0
```

BGP典型案例五（3）

- RTD上的配置：

```
bgp 3
 router-id 11.1.1.1
 undo synchronization
 group ibgp internal
 peer ibgp connect-interface LoopBack0
 peer 12.1.1.1 group ibgp
 peer 12.1.1.1 preferred-value 200
 peer 3.3.3.3 group ibgp
 peer 13.1.1.1 group ibgp
#
ospf 1
 area 0.0.0.0
  network 10.0.0.0 0.255.255.255
  network 11.1.1.1 0.0.0.0
```

- RTE上的配置：

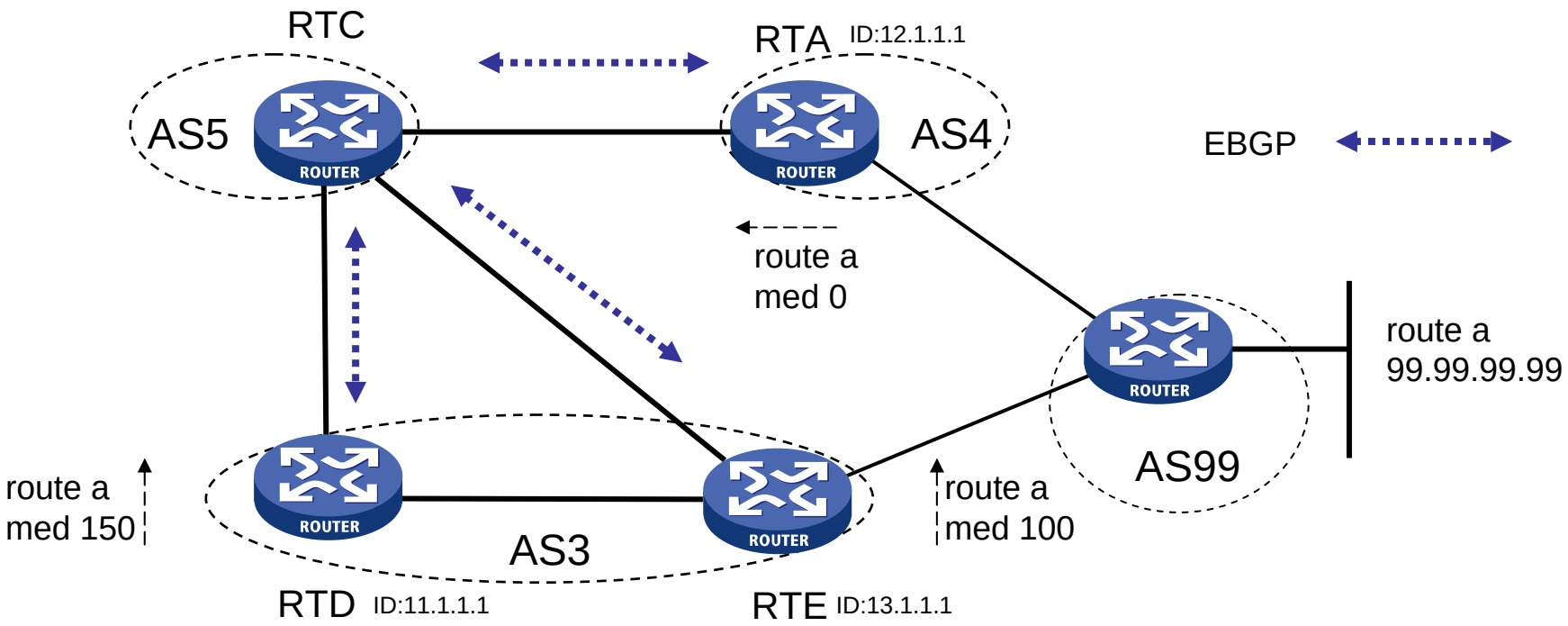
```
bgp 3
 router-id 13.1.1.1
 import-route static
 undo synchronization
 group ibgp internal
 peer ibgp connect-interface LoopBack0
 peer 11.1.1.1 group ibgp
 peer 12.1.1.1 group ibgp
 peer 3.3.3.3 group ibgp
#
ospf 1
 area 0.0.0.0
  network 10.0.0.0 0.255.255.255
  network 13.1.1.1 0.0.0.0
#
route-policy setmed permit node 10
 apply cost 100

 ip route-static 99.99.99.99
255.255.255.255 NULL0
bgp 4
 router-id 12.1.1.1
 network 99.99.99.99 255.255.255.255
 undo synchronization
 peer 10.2.1.2 as-number 5
```

BGP典型案例五（4）

- 排障过程
  - → 在RTC和RTD上查看BGP路由表。在RTC上，路由99.99.99.99的属性带有Preferred-value值200，下一跳指向RTE；在RTD上，路由99.99.99.99的属性带有Preferred-value值200，下一跳指向RTA。
  - → 再查看IP路由表。在RTC上，路由99.99.99.99的下一跳指向RTD；在RTD上，路由99.99.99.99的下一跳指向RTC，由此形成了环路。
  - → 导致RTC与RTD之间出现路由环路是由于对BGP的选路属性设置不当引起的
- 解决方案
  - → 删除RTC与RTD上的Preferred-value配置后，网络恢复正常。

BGP典型案例六（1）



- 故障现象
  - → AS3和AS4都能够从AS99学习99.99.99.99的路由并发布给AS5。RTD与RTE向AS5中的RTC发布路由时设置了不同的MED属性，分别为150和100。
  - → 在RTC的EBGP邻居中断并恢复后，RTC优选出来的BGP路由会发生变化

BGP典型案例六（2）

- RTA上的配置：

```
bgp 4
 router-id 12.1.1.1
 network 99.99.99.99 255.255.255.255
 undo synchronization
 peer 10.2.1.2 as-number 5
```

- RTC上的配置：

```
bgp 5
undo synchronization
 peer 10.2.1.1 as-number 4
 peer 10.4.1.2 as-number 3
 peer 10.6.1.2 as-number 3
```

- RTD上的配置：

```
 bgp 3
  router-id 11.1.1.1
  network 99.99.99.99 255.255.255.255
  undo synchronization
  peer 10.4.1.1 as-number 5
  peer 10.7.1.2 as-number 3
  peer 10.4.1.1 route-policy setmed export
 #
 route-policy setmed permit node 10
  apply cost 150
```

- RTE上的配置：

```
bgp 3
 router-id 13.1.1.1
 network 99.99.99.99 255.255.255.255
 undo synchronization
 peer 10.7.1.1 as-number 3
 peer 10.6.1.1 as-number 5
 peer 10.6.1.1 route-policy setmed export
#
route-policy setmed permit node 10
 apply cost 100
bgp 3
 network 99.99.99.99 255.255.255.255
 undo synchronization
 peer 3.3.3.3 as-number 3
 peer 3.3.3.3 connect-interface
LoopBack0
#
ospf 1
 area 0.0.0.0
  network 10.0.0.0 0.255.255.255
  network 12.1.1.1 0.0.0.0
```

BGP典型案例六（3）

- 排障过程

→在邻居关系稳定的情况查看RTC的BGP路由表，发现RTC优选了来自RTE的路由；

→断开RTC与RTD的EBGP邻居关系然后再恢复，发现RTC优选了来自RTD的路由。BGP前后选路不一致。

- 原因分析

→本例中初始状态：

-RTA和RTD的路由先到RTC，由于默认情况下BGP不比较来自不同AS路由的MED属性，所以RTB优选RTD路由（Router ID小）；

-在RTE路由到达时，由于RTD和RTE属于同一个AS，因此BGP会优选RTE路由（MED小）。

→RTD路由消失：

-BGP会从RTE和RTA中进行优选；同理，因RTE和RTA是不同的AS，BGP优选来自RTA的路由（Router ID小）。

→RTD路由恢复：

-因RTA与RTD是不同的AS，所以BGP优选来自RTD的路由（Router ID小）。

BGP典型案例六（4）

- 解决方案
  - → 调整RTC的配置，在BGP下使用bestroute compare-med，确保路由器根据路由来自的AS进行分组对MED排序优选。

本章总结

- RIP协议相关知识与故障排除
- OSPF协议相关知识与故障排除
- BGP协议相关知识与故障排除

ITоIP解决方案专家

杭州华三通信技术有限公司

www.h3c.com