

## 实验：组播

### HCIP 分解实验 - 组播

臧家林制作



组播实验 1：组播基本概念

组播实验 2：IGMP 实验

组播实验 3：PIM-DM

组播实验 4：PIM-SM

组播实验 5：PIM-SM 的 RP

=====

### 组播实验 1：IP 组播的基本概念

IANA(The Internet Assigned Numbers Authority，互联网数字分配机构)是负责协调一些使 Internet 正常运作的机构。IANA 将 IP 地址分成了 A、B、C、D、E 5 类，

其中的 D 类为组播 IP 地址，范围是 224.0.0.0-239.255.255.255。

IP 单播是一种点到点的通信模式，而 IP 组播则是一种点到多

点的通信模式。一个发送者需要同时向多个接收者发送完全相同的信息时，如果采用单播方式，则网络需要传输大量的报文，相比之下，采用组播方式可以大大减少需要传输的报文数量，从而可以节约大量的网络资源。随着网络的不断发展，电子商务、网络会议、视频点播、远程教学等服务大量兴起，这些服务大多符合点到多点的模式，特别适合于 IP 组播的应用。

## IP地址分类



### 私有 IP 地址

A 类 10.0.0.0 - 10.255.255.255

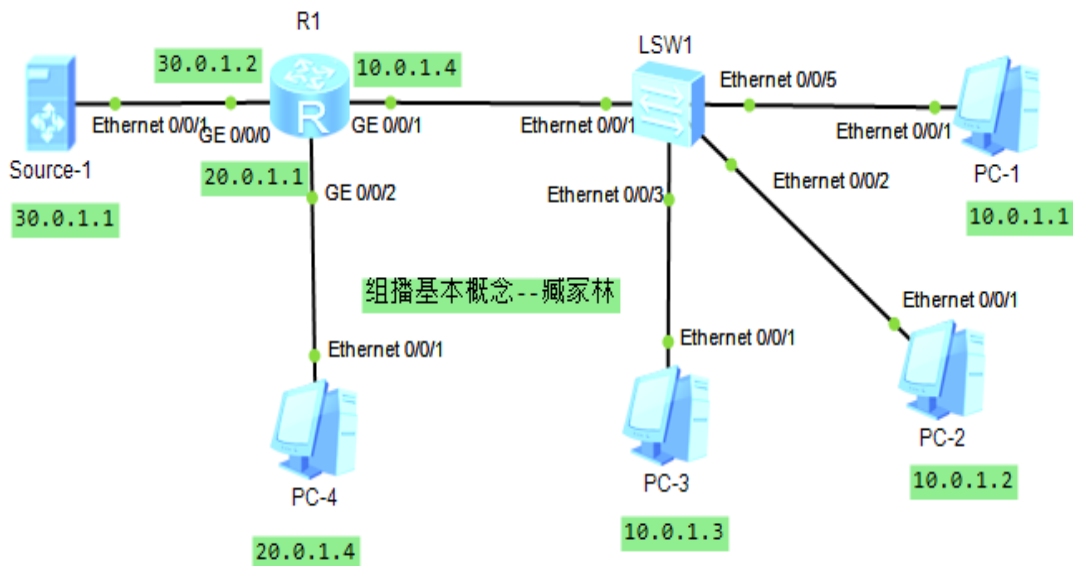
B 类 172.16.0.0 - 172.31.255.255

C 类 192.168.0.0 - 192.168.255.255

在组播方式下，组播报文将沿着组播路由协议建立的树

型路由从信息源传递到众多的终端用户。在这个过程中，只有该组播的成员才能收到并处理该组播组的报文，而对于不是该组播组的成员，要么不能收到该组播组的报文，要么收到后直接丢弃。

虽然 IP 广播也是一种点到点多的通信模式，但相比之下，IP 组播总的来说更具优势。例如，IP 组播是可以跨越网段的，而 IP 广播只能限制在一个网段内。另外，IP 组播也比 IP 广播有更好的安全性。



配置 R1 的 IP 地址

```
R1:
undo ter mo
sys
sysname R1
user-interface console 0
idle-timeout 0 0
int g0/0/0
```

```
ip add 30.0.1.2 24
int g0/0/1
ip add 10.0.1.4 24
int g0/0/2
ip add 20.0.1.1 24
q
```

## 观察单播方式

单播报文的目的 IP 地址只能标识一个唯一的接收者。

R1 : ping 10.0.1.1

```
[R1]ping 10.0.1.1
PING 10.0.1.1: 56 data bytes, press CTRL_C to break
  Reply from 10.0.1.1: bytes=56 Sequence=1 ttl=128 time=140 ms
  Reply from 10.0.1.1: bytes=56 Sequence=2 ttl=128 time=60 ms
  Reply from 10.0.1.1: bytes=56 Sequence=3 ttl=128 time=60 ms
  Reply from 10.0.1.1: bytes=56 Sequence=4 ttl=128 time=50 ms
  Reply from 10.0.1.1: bytes=56 Sequence=5 ttl=128 time=60 ms
```

可以把交换机的 stp 协议关闭掉，以免抓取到大量的 BPDU

SW1:

stp disable

在 PC -1 的接口 e0/0/1 抓包，可以看到有 icmp 报文

5	7.469000	10.0.1.4	10.0.1.1	ICMP	Echo (ping) request (id=0xcfab, seq(be/le)=256/1, ttl=255)
6	7.469000	10.0.1.1	10.0.1.4	ICMP	Echo (ping) reply (id=0xcfab, seq(be/le)=256/1, ttl=128)
7	8.000000	10.0.1.4	10.0.1.1	ICMP	Echo (ping) request (id=0xcfab, seq(be/le)=512/2, ttl=255)
8	8.000000	10.0.1.1	10.0.1.4	ICMP	Echo (ping) reply (id=0xcfab, seq(be/le)=512/2, ttl=128)
9	8.516000	10.0.1.4	10.0.1.1	ICMP	Echo (ping) request (id=0xcfab, seq(be/le)=768/3, ttl=255)
10	8.516000	10.0.1.1	10.0.1.4	ICMP	Echo (ping) reply (id=0xcfab, seq(be/le)=768/3, ttl=128)
11	9.047000	10.0.1.4	10.0.1.1	ICMP	Echo (ping) request (id=0xcfab, seq(be/le)=1024/4, ttl=255)
12	9.047000	10.0.1.1	10.0.1.4	ICMP	Echo (ping) reply (id=0xcfab, seq(be/le)=1024/4, ttl=128)

PC -2 PC -3 的接口上没有接收到报文

如果多个用户需要获得相同的信息，那么在单播方式下，网络中需要传输的报文数量将和用户数量成正比。用户数量越多，网络中包含相同信息的报文数量就越多，这样既浪费网络设备的 CPU 资源又浪费网络的带宽资源。

=====

### 观察广播方式

一个广播 IP 地址标识了某确定网段内的所有网络设备，该网段内每个网络设备都会收到并处理该网段的广播报文。

在 R1 上配置 RIPv1 协议，使 R1 通过广播方式发送 RIP 报文，然后在 PC 接口上抓包，查看报文的接收情况。

R1：

rip

network 10.0.0.0

在 PC 的接口上抓包，可以看到有 255.255.255.255 的 RIP 报文

Time	Source	Destination	Protocol	Info
8 13.922000	10.0.1.4	255.255.255.255	RIPv1	Response
21 41.234000	10.0.1.4	255.255.255.255	RIPv1	Response
31 58.359000	10.0.1.4	255.255.255.255	RIPv1	Response
36 66.437000	10.0.1.4	255.255.255.255	RIPv1	Response
41 71.531000	10.0.1.4	255.255.255.255	RIPv1	Response

可以看到 3 台 PC 都收到了 R1 通过广播方式发送的 RIP 报文。

广播方式只在同一个网段中才有效，不能跨越网段。另外，广播方式是无法区分接收者的，这对于信息的安全性和服务的有偿性而言是一个问题。

=====

## 观察组播方式

把 RIP 协议删除掉

R1:

```
undo rip 1
```

使用组播方式时，只有加入到该组播组的成员才能收到并处理该组播组的报文。

在 R1 上完成组播的基本配置。全局开启组播功能，在接口下开启组播功能及 IGMP 功能。

R1:

```
multicast routing-enable
```

```
int g0/0/0
```

```
pim dm
```

```
int g0/0/1
```

```
pim dm
```

```
igmp enable
```

```
int g0/0/2
```

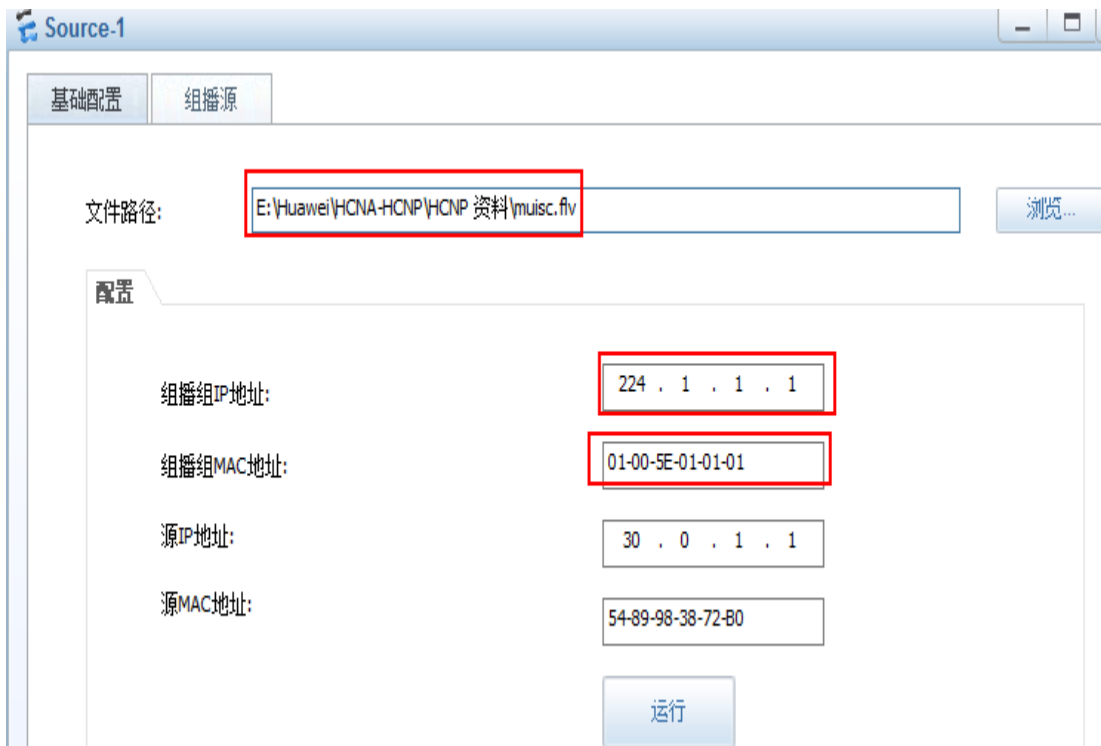
```
pim dm
```

```
igmp enable
```

组播服务器 source-1 需要使用 VLC 播放视频的方式来发送组播。



组播源设置，flv 文件的路径，组播组的 IP 地址为 224.1.1.1  
MAC 地址会自动改变。



让 PC 4 加入 224.1.1.1 组播组，使用 ICMP v2 协议（默认值）

基础配置

命令行

组播

UDP发包工具

数据包配置

IGMP 版本：

Version 1

Version 2

Version 3

源 IP：

20 . 0 . 1 . 4

源 MAC：

54-89-98-90-51-A8

启动 VLC

目的 IP：

224 . 1 . 1 . 1

目的 MAC：

01-00-5E-01-01-01

加入

离开

source-1 先不运行，让 PC-4 先加入，启动 VLC 之后 source-1 再运行。

在 PC-4 上抓包，有 IGMP 报文，而其它几个 PC 上没有

Time	Source	Destination	Protocol	Info
1 0.000000	20.0.1.1	224.0.0.1	IGMP	V2 Membership Query, general
2 0.000000	20.0.1.4	224.1.1.1	IGMP	V2 Membership Report / Join group 224.1.1.1
3 15.094000	20.0.1.1	224.0.0.13	PIMv2	Hello

source-1 运行  
在 PC-4 抓包，有大量的组播报文

Time	Source	Destination	Protocol	Info
716 56.400000	30.0.1.1	224.1.1.1	UDP	Source port: raven-rdp Destination port: 0
717 56.422000	30.0.1.1	224.1.1.1	UDP	Source port: raven-rdp Destination port: 0
718 56.438000	30.0.1.1	224.1.1.1	UDP	Source port: raven-rdp Destination port: 0
719 56.453000	30.0.1.1	224.1.1.1	UDP	Source port: raven-rdp Destination port: 0
720 56.453000	30.0.1.1	224.1.1.1	UDP	Source port: raven-rdp Destination port: 0
721 56.485000	30.0.1.1	224.1.1.1	UDP	Source port: raven-rdp Destination port: 0
722 56.500000	30.0.1.1	224.1.1.1	UDP	Source port: raven-rdp Destination port: 0
723 56.531000	30.0.1.1	224.1.1.1	UDP	Source port: raven-rdp Destination port: 0
724 56.531000	30.0.1.1	224.1.1.1	UDP	Source port: raven-rdp Destination port: 0
725 56.547000	30.0.1.1	224.1.1.1	UDP	Source port: raven-rdp Destination port: 0
726 56.563000	30.0.1.1	224.1.1.1	UDP	Source port: raven-rdp Destination port: 0
727 56.563000	30.0.1.1	224.1.1.1	UDP	Source port: raven-rdp Destination port: 0
728 56.581000	30.0.1.1	224.1.1.1	UDP	Source port: raven-rdp Destination port: 0



而在 PC-3 抓包，则没有

time	source	destination	protocol	info
6 50.500000	HuaweiTe_09:2a:01	Spanning-tree-(for-STP		MST. Root = 32768/0/4c:1f:cc:09:2a:01 Cost = 0
7 51.218000	10.0.1.4	224.0.0.1	IGMP	V2 Membership Query, general
8 52.672000	HuaweiTe_09:2a:01	Spanning-tree-(for-STP		MST. Root = 32768/0/4c:1f:cc:09:2a:01 Cost = 0
9 54.859000	HuaweiTe_09:2a:01	Spanning-tree-(for-STP		MST. Root = 32768/0/4c:1f:cc:09:2a:01 Cost = 0
10 57.062000	HuaweiTe_09:2a:01	Spanning-tree-(for-STP		MST. Root = 32768/0/4c:1f:cc:09:2a:01 Cost = 0
11 59.281000	HuaweiTe_09:2a:01	Spanning-tree-(for-STP		MST. Root = 32768/0/4c:1f:cc:09:2a:01 Cost = 0
12 61.468000	HuaweiTe_09:2a:01	Spanning-tree-(for-STP		MST. Root = 32768/0/4c:1f:cc:09:2a:01 Cost = 0
13 63.734000	HuaweiTe_09:2a:01	Spanning-tree-(for-STP		MST. Root = 32768/0/4c:1f:cc:09:2a:01 Cost = 0
14 65.890000	HuaweiTe_09:2a:01	Spanning-tree-(for-STP		MST. Root = 32768/0/4c:1f:cc:09:2a:01 Cost = 0
15 67.250000	10.0.1.4	224.0.0.13	PIMv2	Hello
16 68.187000	HuaweiTe_09:2a:01	Spanning-tree-(for-STP		MST. Root = 32768/0/4c:1f:cc:09:2a:01 Cost = 0
17 70.422000	HuaweiTe_09:2a:01	Spanning-tree-(for-STP		MST. Root = 32768/0/4c:1f:cc:09:2a:01 Cost = 0
18 72.734000	HuaweiTe_09:2a:01	Spanning-tree-(for-STP		MST. Root = 32768/0/4c:1f:cc:09:2a:01 Cost = 0
19 74.906000	HuaweiTe_09:2a:01	Spanning-tree-(for-STP		MST. Root = 32768/0/4c:1f:cc:09:2a:01 Cost = 0

=====

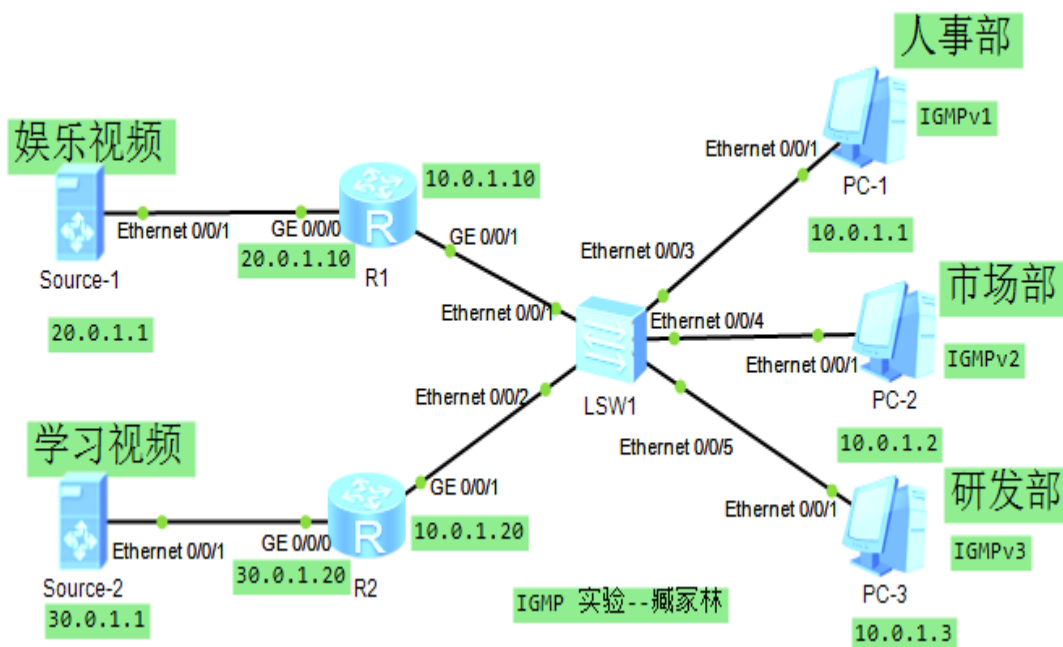
## 组播实验 2：IGMP 实验

IGMP ( Internet Group Management Protocol,因特网组管理协议 ) 是 TCP/IP 协议簇中负责组播成员管理的协议，其作用是在用户主机和与其直连的组播路由器之间建立和维护组播成员关系。通过在用户主机与其直连的组播路由器上配置和运行 IGMP，可以实现主机动态地加入和离开组播组，以及组播路由器对本地网络中组播成员信息的动态管理。IGMP 有 3 个版本，分别是 IGMPv1、IGMPv2、IGMPv3。

IGMPv1 主要基于查询和响应机制来完成组播组的管理。主机通过发送 report 消息加入到某组播组，主机离开组播时不发送离开报文，离开后再收到路由器发送的查询消息时不反馈 report 消息，待维护组成员关系的定时器超时后，路由器会自动删除该主机的成员记录。

IGMPv2 与 IGMPv1 基本相似，主机的不同点在于 IGMPv2 具有某些报文抑制机制，可以减少不必要的 IGMP 重复报文，从而节省网络带宽资源，另外，主机离开组播组时，会主动向路由器发送离开报文。

IGMPv1 和 IGMPv2 报文中都只能携带组播组的信息，不能携带组播源的信息，所以主机只能选择加入某个组，而不能选择组播源，这一问题在 IGMPv3 中得到了解决。运行 IGMPv3 时，主机不仅能够选择组，还能根据选择组播源。主机发送的 IGMPv3 报文中可以包含多个组记录，每个组记录中可以包含多个组播源。



## 两个组播服务器

source -1 存放一些娱乐视频，source -2 存放一些学习视频。

人事部只想看娱乐视频，且不在意带宽资源问题，所以在 PC-1 上配置的是 IGMPv1

市场部希望尽量节约带宽资源，所以需要在 PC-2 上配置 IG

MPv2

研发部不允许观看娱乐视频，所以需要在 PC-3 上配置 IGM  
Pv3

R1 :

```
undo ter mo
sys
sysname R1
user-interface console 0
idle-timeout 0 0
int g0/0/0
ip add 20.0.1.10 24
int g0/0/1
ip add 10.0.1.10 24
q
```

R2 :

```
undo ter mo
sys
sysname R2
user-interface console 0
idle-timeout 0 0
int g0/0/0
ip add 30.0.1.10 24
int g0/0/1
ip add 10.0.1.20 24
q
```

设置两个组播服务器的组播组 IP 地址为 224.1.1.1，MAC 地址自动变为 01-00-5E-01-01-01

两台服务器都需要设置

基础配置

组播源

文件路径:

配置

组播组IP地址:

224 . 1 . 1 . 1

组播组MAC地址:

01-00-5E-01-01-01

源IP地址:

20 . 0 . 1 . 1

源MAC地址:

54-89-98-FC-35-3E

### 配置组播协议

由于公司需要通过组播方式发送视频，所以路由器需要开启组播功能，相应接口下也需要配置组播协议。

R1:

```
multicast routing-enable
int g0/0/0
pim dm
int g0/0/1
pim dm
```

R2:

```
multicast routing-enable
int g0/0/0
pim dm
int g0/0/1
pim dm
```

### 配置 IGMPv1

在 PC 上使用 IGMP 加入相应的组播组，接收并观看组播视频，因此路由器用户侧的接口同样需要开启 IGMP 功能以处

理 IGMP 消息，并修改版本为 IGMPv1。（默认为 IGMPv2）

R1:

```
int g0/0/1
```

```
igmp enable
```

```
igmp version 1
```

<R1>display igmp interface 查看 IGMP 接口信息，版本为 1，  
查询消息的发送间隔为 60s

```
[R1]dis igmp interface
```

```
Interface information of VPN-Instance: public net
```

```
GigabitEthernet0/0/1(10.0.1.10):
```

```
IGMP is enabled
```

```
Current IGMP version is 1
```

```
IGMP state: up
```

```
IGMP group policy: none
```

```
IGMP limit: -
```

```
Value of query interval for IGMP (negotiated): -
```

```
Value of query interval for IGMP (configured): 60 s
```

```
Value of other querier timeout for IGMP: 0 s
```

<R1>debugging igmp report

<R1>debugging igmp event

<R1>terminal monitor

<R1>terminal debugging

PC-1 配置 IGMPv1 并加入组 224.1.1.1

```
Apr 25 2018 21:28:49.530.2-08:00 R1 IGMP/7/REPORT:(public net): Received v1 report for group 224.1.1.1 on interface GigabitEthernet0/0/1(10.0.1.10) (G081904)
Apr 25 2018 21:28:49.530.3-08:00 R1 IGMP/7/EVENT:(public net): Creating group(224.1.1.1) for interface GigabitEthernet0/0/1(10.0.1.10) (G014427)
```

<R1>display igmp group

```
<R1>display igmp group
Interface group report information of VPN-Instance: public net
GigabitEthernet0/0/1(10.0.1.10):
Total 1 IGMP Group reported
  Group Address    Last Reporter    Uptime           Expires
  224.1.1.1        10.0.1.1         00:00:56         00:01:14
<R1>
```

此时 source-1 播放视频，PC-1 运行 VLC 即可观看  
PC-1 先运行 VLC，source-1 再点运行

PC-1 点击离开



运行 IGMPv1 不会发送离组消息，130s 内如果没有收到某个组成员的 Report 消息，则认为该成员已经离组

=====

## 配置 IGMPv2

PC-2 使用 IGMPv2 , 所以 R1 和 R2 接口配置为 IGMPv2

R1:

```
int g0/0/1
```

```
igmp version 2
```

R2:

```
int g0/0/1
```

```
igmp enable
```

PC-2 版本 2 加入

The screenshot shows the 'PC-2' configuration window with the '数据配置' (Data Configuration) tab selected. The 'IGMP 版本' (IGMP Version) section has three radio buttons: 'Version 1', 'Version 2' (which is selected and highlighted with a red box), and 'Version 3'. Below this, the '源 IP' (Source IP) is set to '10 . 0 . 1 . 2' and the '源 MAC' (Source MAC) is '54-89-98-FB-56-8A'. The '目的 IP' (Destination IP) is set to '224 . 1 . 1 . 1' (highlighted with a red box) and the '目的 MAC' (Destination MAC) is '01-00-5E-01-01-01'. To the right of the '目的 IP' field is a '加入' (Add) button, also highlighted with a red box. Further right are buttons for '启动 VLC' (Start VLC) and '离开' (Leave).

<R1>display igmp interface

```
[R1]dis igmp interface
Interface information of VPN-Instance: public net
GigabitEthernet0/0/1(10.0.1.10):
  IGMP is enabled
  Current IGMP version is 2
  IGMP state: up
  IGMP group policy: none
  IGMP limit: -
  Value of query interval for IGMP (negotiated): -
  Value of query interval for IGMP (configured): 60 s
  Value of other querier timeout for IGMP: 0 s
  Value of maximum query response time for IGMP: 10 s
  Querier for IGMP: 10.0.1.10 (this router)
Total 1 IGMP Group reported
```

运行的是 IGMPv2 还显示了 IGMP 查询器为 10.0.1.10，即 R1

同一网段上有多个组播路由器时，每个组播路由器都能从别的组播路由器和主机那里收到成员关系报告消息，但是只需要一台路由器发送成员资格查询消息，所以，这就需要一个路由器选举机制来确定一台路由器作为查询器。

在 PC-2 的接口上抓包

3	3.422000	10.0.1.10	224.0.0.1	IGMP	V2 Membership Query, general
4	3.422000	10.0.1.2	224.1.1.1	IGMP	V2 Membership Report / Join group 224.1.1.1
35	63.453000	10.0.1.10	224.0.0.1	IGMP	V2 Membership Query, general
36	63.453000	10.0.1.2	224.1.1.1	IGMP	V2 Membership Report / Join group 224.1.1.1



```

# Frame 3: 60 bytes on wire (480 bits), 60 bytes captured (480 bits)
+ Ethernet II, Src: HuaweiTe_bf:33:f5 (54:89:98:bf:33:f5), Dst: IPv4mcast_00:00:01 (01:00:5e:00:00:01)
+ Internet Protocol, Src: 10.0.1.10 (10.0.1.10), Dst: 224.0.0.1 (224.0.0.1)
- Internet Group Management Protocol
  [IGMP Version: 2]
  Type: Membership Query (0x11)
  Max Response Time: 10.0 sec (0x64)
  Header checksum: 0xee9b [correct]
  Multicast Address: 0.0.0.0 (0.0.0.0)

```

R1:  
debugging igmp report  
debugging igmp leave

PC-2 使用 IGMPv2 加入 224.1.1.1

```

- -
Apr 25 2018 21:32:32.700.1-08:00 R1 IGMP/7/REPORT:(public net): Received v2 report
rt for group 224.1.1.1 on interface GigabitEthernet0/0/1(10.0.1.10) (G082859)

```

PC-2 离开

<R1>

```

Apr 25 2018 21:36:38.730.2-08:00 R1 IGMP/7/LEAVE:(public net): LEAVE received on
interface GigabitEthernet0/0/1(10.0.1.10) has destination address(224.1.1.1), i
s not equal to all-router-multicast-address(224.0.0.2) (G083040)

```

R1 收到 PC-2 发送的离组消息

=====

配置 IGMPv3

由于 PC-3 不允许接收 source-1 的娱乐视频，所以 PC-3 需要使用 IGMPv3 来选择组播源，拒绝来自 source-1 的组播流。

R1:  
int g0/0/1  
igmp version 3

R2:  
int g0/0/1  
igmp version 3

## 配置 PC-3

The screenshot shows the 'PC-3' configuration window with the 'UDP发包工具' (UDP Packet Tool) tab selected. Under the '数据包配置' (Data Packet Configuration) section, the following settings are visible:

- IGMP 版本: Version 3 (selected)
- 源 IP: 10.0.1.3
- 源 MAC: 54-89-98-38-1D-3C
- 组播组 IP 起始: 224.1.1.1
- 组数: 1
- 增量: 1
- 类型: MODE\_IS\_EXCLUDE
- 组播源 IP 起始: 20.0.1.1
- 地址数: 1
- 增量: 1

Buttons for '启动 VLC' (Start VLC) and '发送' (Send) are present.

在 R1 的 g0/0/1 接口抓包，PC-3 点击发送

Time	Source	Destination	Protocol	Info
1 0.000000	HuaweiTe_5f:6f:00	Spanning-tree-(for-STP		MST. Root = 32768/0/4c:1f:cc:5f:
2 2.437000	HuaweiTe_5f:6f:00	Spanning-tree-(for-STP		MST. Root = 32768/0/4c:1f:cc:5f:
3 4.625000	HuaweiTe_5f:6f:00	Spanning-tree-(for-STP		MST. Root = 32768/0/4c:1f:cc:5f:
4 7.000000	HuaweiTe_5f:6f:00	Spanning-tree-(for-STP		MST. Root = 32768/0/4c:1f:cc:5f:
5 7.109000	10.0.1.3	224.0.0.22	IGMP	V3 Membership Report / Join grou
6 9.297000	HuaweiTe_5f:6f:00	Spanning-tree-(for-STP		MST. Root = 32768/0/4c:1f:cc:5f:
7 11.515000	HuaweiTe_5f:6f:00	Spanning-tree-(for-STP		MST. Root = 32768/0/4c:1f:cc:5f:

```

] Frame 5: 58 bytes on wire (464 bits), 58 bytes captured (464 bits)
] Ethernet II, Src: HuaweiTe_38:1d:3c (54:89:98:38:1d:3c), Dst: IPv4mcast_00:00
] Internet Protocol, Src: 10.0.1.3 (10.0.1.3), Dst: 224.0.0.22 (224.0.0.22)
Internet Group Management Protocol
  [IGMP Version: 3]
  Type: Membership Report (0x22)
  Header checksum: 0xe5f9 [correct]
  Num Group Records: 1
  Group Record : 224.1.1.1 Mode Is Exclude
    Record Type: Mode Is Exclude (2)
    Aux Data Len: 0
    Num Src: 1
    Multicast Address: 224.1.1.1 (224.1.1.1)
    Source Address: 20.0.1.1 (20.0.1.1)

```

R1 收到了来自 PC-3 发送的 IGMPv3 消息，模式为 exclude，组播地址为 224.1.1.1，组播源地址为 20.0.1.1，其含义就是拒绝接收来自 20.0.1.1 发送的组播地址为 224.1.1.1 的信息。

=====

### 组播实验 3 : PIM-DM

树 ( Tree ) 和图 ( Graph ) 是计算机科学领域中两个常用的概念，前者具有层次化结构，而后者没有。从组播的角度来看，网络可以抽象为 Tree，也称为组播树 ( Multicast Tree )，从单播的角度来看，网络可以抽象为 Graph。组播树又可以分为两类，一类称为 Source-Based Tree，另一类称为 Group-Shared Tree。

一棵 Source-Based Tree 是由组播源和组播成员共同决定的。例如，对于一个特定的组播组，如果组成员的分布已经确定了，则组播源的位置不同将会导致形成不同的 Source-Based Tree。同样，对于一个特定的组播组，如果组播源的位置已经确定了，则组成员的不同分布也将会导致产生不同的 Source-Based Tree。一棵 Group-Shared Tree 仅仅由组成员的分布完全决定的，而与组播源的位置无关。

PIM-DM ( Protocol Independent Multicast Dense Mode ) 和 PIM-SM ( Protocol Independent Multicast Sparse Mode ) 是两个常见的组播路由协议，前者基于 Source-Based Tree，后者基于 Group-Shared Tree。

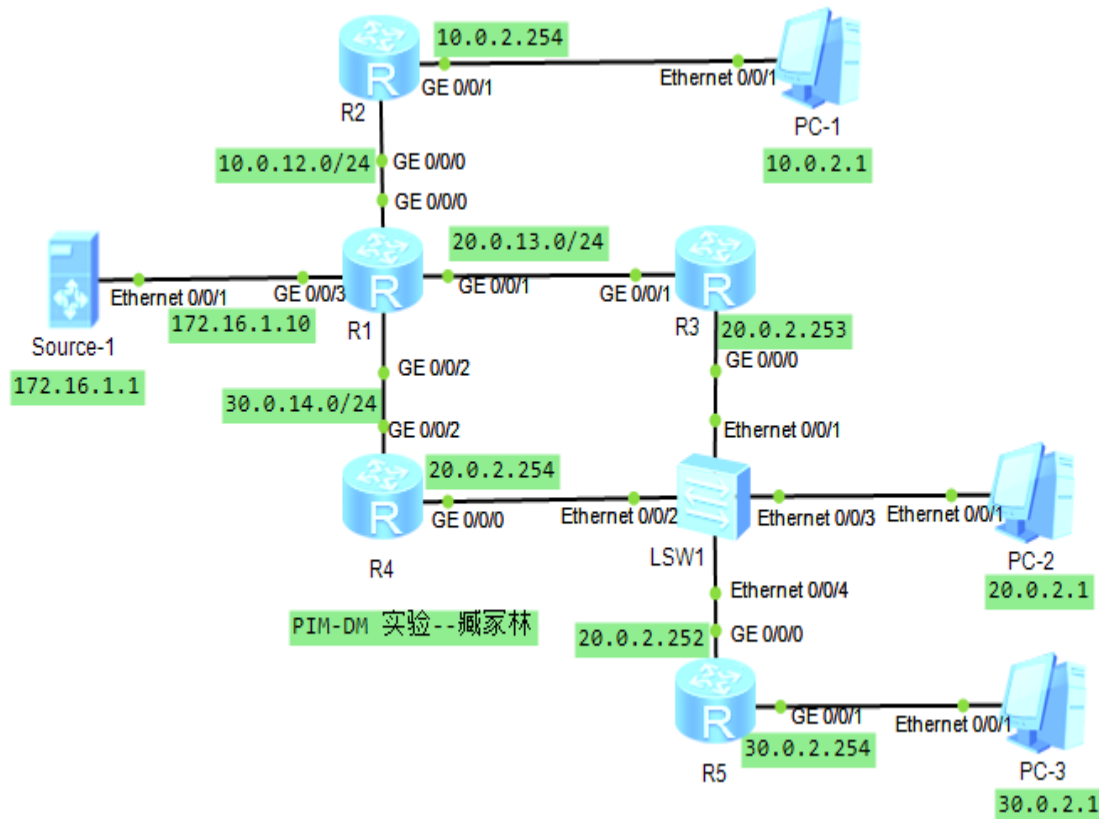
PIM-DM 主要采用扩散--剪枝的方式来转发组播数据流。对于组播成员稀少的网络，PIM-DM 会产生大量的剪枝报文，而如果网络规模较大，则扩散--剪枝的周期就会比较长，因此，PIM-DM 一般适合于规模较小、组播成员比较密集的网络。

PIM-DM 首先假设网络中的每个子网都存在至少一个组成员，并将组播数据包从组播源扩散到网络中的所有路由器，然后，对于实际上没有组成员的分支进行剪枝操作。所谓剪枝 ( Prune )，就是路由器向上游节点发送剪枝消息，通知上游节点不用再转发组播数据到该分支上。上游节点收到剪枝消息后，会将相应的接口从其组播转发表项 ( S,G ) 中删除，只保留包含组成员的分支，这样便可减少网络资源的消耗。另外，各个被剪枝的节点同时还提供了超时机制，当剪枝超时后 ( 默认为 210s ) 将重新开始扩散--剪枝过程。被裁剪的分支如果临时有组播数据转发需求，也可以使用嫁接 ( Graft ) 机制主动请求恢复组播数据的转发。

周期性的扩散--剪枝行为是 PIM-DM 的一个重要特征，通过这样的行为，PIM-DM 可以构建并动态地维护一棵从组播源到组成员的单向无环的 SPT ( Shortest Path Tree )。SPT 是以组播源为根、组播组成员为枝叶的从组播源到组成员的一棵最短路径树，此树也就是组播数据的转发路径。组播数据的转发中会出现上游接口和下游接口这两个概念，[路由器收到组播数据的接口称为上游接口](#)，[转发组播数据的接口称为下游接口](#)。

在 PIM-DM 网络中，路由器需要周期性地发送 Hello 消息一发

现邻居并维护 PIM 邻居关系。此外，Hello 消息还有一个重要的作用，路由器会通过比较 Hello 消息中携带的优先级和 IP 地址，为多路由器网段选举出 DR，并以它作为 IGMPv1 中的查询器。



### 基本配置

R1:

```
undo ter mo
sys
sysname R1
user-interface console 0
idle-timeout 0 0
int g0/0/3
ip add 172.16.1.10 24
int g0/0/0
ip add 10.0.12.1 24
```

```
int g0/0/1
ip add 20.0.13.1 24
int g0/0/2
ip add 30.0.14.1 24
q
```

```
R2:
undo ter mo
sys
sysname R2
user-interface console 0
idle-timeout 0 0
int g0/0/1
ip add 10.0.2.254 24
int g0/0/0
ip add 10.0.12.2 24
q
```

```
R3:
undo ter mo
sys
sysname R3
user-interface console 0
idle-timeout 0 0
int g0/0/1
ip add 20.0.13.3 24
int g0/0/0
ip add 20.0.2.253 24
q
```

```
R4:
undo ter mo
sys
sysname R4
```

```
user-interface console 0
idle-timeout 0 0
int g0/0/2
ip add 30.0.14.4 24
int g0/0/0
ip add 20.0.2.254 24
q
```

```
R5:
undo ter mo
sys
sysname R5
user-interface console 0
idle-timeout 0 0
int g0/0/0
ip add 20.0.2.252 24
int g0/0/1
ip add 30.0.2.254 24
q
```

### 配置 IGP 协议

在每台路由器上配置 OSPF 路由协议，并通告直连网段

```
R1 :
ospf
area 0
net 172.16.1.10 0.0.0.0
net 10.0.12.1 0.0.0.0
net 20.0.13.1 0.0.0.0
net 30.0.14.1 0.0.0.0
q
```

```
R2 :
ospf
area 0
```

```
net 10.0.2.254 0.0.0.0
net 10.0.12.2 0.0.0.0
q
```

```
R3 :
ospf
area 0
net 20.0.13.3 0.0.0.0
net 20.0.2.253 0.0.0.0
q
```

```
R4 :
ospf
area 0
net 30.0.14.4 0.0.0.0
net 20.0.2.254 0.0.0.0
q
```

```
R5 :
ospf
area 0
net 20.0.2.252 0.0.0.0
net 30.0.2.254 0.0.0.0
q
```

查看 OSPF 邻居已经建立好，并且学习到路由

### 配置 PIM-DM

在所有路由器上开启组播功能，并在每台路由器的每个接口下配置命令 pim dm

```
R1:
multicast routing-enable
int g0/0/0
```



```
pim dm
int g0/0/1
pim dm
int g0/0/2
pim dm
int g0/0/3
pim dm
```

```
R2:
multicast routing-enable
int g0/0/0
pim dm
int g0/0/1
pim dm
```

```
R3:
multicast routing-enable
int g0/0/0
pim dm
int g0/0/1
pim dm
```

```
R4:
multicast routing-enable
int g0/0/0
pim dm
int g0/0/2
pim dm
```

```
R5:
multicast routing-enable
int g0/0/0
pim dm
int g0/0/1
```

pim dm

配置完成后，在 R1 查看 PIM 邻居的建立情况 <R1>display pim neighbor

```
[R1]display pim neighbor
VPN-Instance: public net
Total Number of Neighbors = 3
```

Neighbor	Interface	Uptime	Expires	Dr-Priority	BFD-Session
10.0.12.2	GE0/0/0	00:01:12	00:01:33	1	N
20.0.13.3	GE0/0/1	00:00:48	00:01:27	1	N
30.0.14.4	GE0/0/2	00:00:37	00:01:38	1	N

```
[R1]
```

在 R2 R3 R4 R5 的用户侧接口下开启 IGMP

```
R2 :
int g0/0/1
igmp enable
R3 :
int g0/0/0
igmp enable
R4 :
int g0/0/0
igmp enable
R5 :
int g0/0/1
igmp enable
```

在 R2 上查看 IGMP 接口 <R2>display igmp interface

```
[R2]dis igmp int
Interface information of VPN-Instance: public net
GigabitEthernet0/0/1(10.0.2.254):
  IGMP is enabled
  Current IGMP version is 2
  IGMP state: up
  IGMP group policy: none
  IGMP limit: -
```

服务器 source-1 设置组播地址 224.1.1.1



当路由器接收到组播源发送的组播数据后便会自动生成组播路由，查看 R1 的组播路由表

```
<R1>display pim routing-table
```

```
[R1]dis pim routing-table
VPN-Instance: public net
Total 0 (*, G) entry; 1 (S, G) entry

(172.16.1.1, 224.1.1.1)
  Protocol: pim-dm, Flag: LOC ACT
  UpTime: 00:00:16
  Upstream interface: GigabitEthernet0/0/3
    Upstream neighbor: NULL
    RPF prime neighbor: NULL
  Downstream interface(s) information: None
```

R1 的上游接口为 g0/0/3

由于 PC-1 并没有第一时间加入该组播组，所以 R2 没有收到 IGMP 加入消息，于是 R2 会认为自己没有连接任何组成员。R2 在收到 R1 发送的组播数据后，会向 R1 发送剪枝消息。R1 收到剪枝消息后，会立即停止向 R2 发送组播数据包。

=====

观察 PIM-DM 中的 graft 机制

让 PC-1 使用 IGMPv2 加入组播组 224.1.1.1

PC-1

基础配置 | 命令行 | 组播 | UDP发包工具

数据包配置

IGMP 版本: ☐ Version 1 ☒ Version 2 ☐ Version 3

源 IP: 10.0.2.1 源 MAC: 54-89-98-67-13-36 启动 VLC

目的 IP: 224.1.1.1 目的 MAC: 01-00-5E-01-01-01 加入 离开

R2 收到 PC-1 的加入消息后，会立即发送 Graft（嫁接）消息给 R1，R1 收到后会立即开始重新转发组播数据包给 R2，R2 再转发给 PC-1

R2 的 g0/0/0 抓包

55	198.938000	10.0.12.2	224.0.0.5	OSPF	Hello Packet
56	203.844000	10.0.12.2	10.0.12.1	PIMv2	Graft
57	203.844000	10.0.12.1	10.0.12.2	PIMv2	Graft-Ack
58	208.328000	10.0.12.2	224.0.0.5	OSPF	Hello Packet
59	209.109000	10.0.12.1	224.0.0.5	OSPF	Hello Packet
60	215.625000	10.0.12.1	224.0.0.13	PIMv2	Hello
61	217.609000	10.0.12.2	224.0.0.13	PIMv2	Hello

=====

观察 PIM-DM 中 DR 的选举

PC-2 使用 IGMPv1 加入组播组 224.1.1.1，在 R3 R4 上修改 IGMP 为版本 1

R3:  
 int g0/0/0  
 igmp version 1

R4:  
 int g0/0/0  
 igmp version 1

由于 PC-2 通过交换机同时连接了 R3 R4，为了避免收到重复的 IGMP 查询消息，需要选举出一个查询器，查询报文仅由查询器发送。IGMPv1 中查询器的选举由组播路由协议决定，PIM-DM 选举出来的 DR 即为 IGMPv1 的查询器。

<R3>display pim interface

[R3]dis pim interface

VPN-Instance: public net

Interface	State	NbrCnt	HelloInt	DR-Pri	DR-Address
GE0/0/0	up	2	30	1	20.0.2.254
GE0/0/1	up	1	30	1	20.0.13.3 (local)

[R4]dis pim interface

VPN-Instance: public net

Interface	State	NbrCnt	HelloInt	DR-Pri	DR-Address
GE0/0/0	up	2	30	1	20.0.2.254 (local)
GE0/0/2	up	1	30	1	30.0.14.4 (local)

PIM-DM 选举出来的 DR 为 R4，优先级一样时，IP 地址较大者为 DR

验证一下 IGMPv1 的查询路由器是否就是 PIM-DM 选举出来的 DR，查看关于查询器的情况。

```
[R3]dis igmp interface
Interface information of VPN-Instance: public net
GigabitEthernet0/0/0(20.0.2.253):
  IGMP is enabled
  Current IGMP version is 1
  IGMP state: up
  IGMP group policy: none
  IGMP limit: -
  Value of query interval for IGMP (negotiated): -
  Value of query interval for IGMP (configured): 60 s
  Value of other querier timeout for IGMP: 0 s
  Value of maximum query response time for IGMP: -
  Querier for IGMP: 20.0.2.254
Total 1 IGMP Group reported
```

```
[R4]dis igmp int
Interface information of VPN-Instance: public net
GigabitEthernet0/0/0(20.0.2.254):
  IGMP is enabled
  Current IGMP version is 1
  IGMP state: up
  IGMP group policy: none
  IGMP limit: -
  Value of query interval for IGMP (negotiated): -
  Value of query interval for IGMP (configured): 60 s
  Value of other querier timeout for IGMP: 0 s
  Value of maximum query response time for IGMP: -
  Querier for IGMP: 20.0.2.254 (this router)
Total 1 IGMP Group reported
```

=====

## 观察 PIM-DM 中的 Assert 机制

R3 R4 从上游接收到组播报文后，都向下转发该组播报文，这样就会导致下游节点 R5 收到两份完全相同的组播报文。为了避免这种情况的发生 PIM-DM 采用了 Assert 机制来选定一个唯一的转发者。

即对于一个特定的组播组，如果同一个网段上存在多个上游路由器，则这些上游路由器中到组播源的路径开销最小者将被选举为转发者。如果相同，则 IP 地址最大的路由器将被选定为转发者。只有转发者才能向该网段转发相应的组播数据报文。

R3:  
int g0/0/0  
igmp version 2

R4:  
int g0/0/0  
igmp version 2

PC-3 使用 IGMPv2 加入 224.1.1.1  
**源播放视频**，在 R5 的 g0/0/0 开始抓包





23	23.391000	20.0.2.252	224.0.0.5	OSPF	Hello Packet
24	24.172000	HuaweiTe_b0:68:db	Spanning-tree-(for-STP	MST. Root = 32768/0/4c:1f:cc:b0:68:db	Cost :
25	25.969000	172.16.1.1	224.1.1.1	UDP	Source port: 7817 Destination port: 0
26	25.969000	172.16.1.1	224.1.1.1	UDP	Source port: 7817 Destination port: 0
27	25.969000	20.0.2.252	224.0.0.13	PIMv2	Join/Prune
28	26.000000	20.0.2.254	224.0.0.13	PIMv2	Assert
29	26.000000	20.0.2.253	224.0.0.13	PIMv2	Assert
30	26.016000	20.0.2.254	224.0.0.13	PIMv2	Assert
31	26.016000	20.0.2.253	224.0.0.13	PIMv2	Join/Prune
32	26.047000	172.16.1.1	224.1.1.1	UDP	Source port: 7817 Destination port: 0
33	26.047000	172.16.1.1	224.1.1.1	UDP	Source port: 7817 Destination port: 0
34	26.047000	172.16.1.1	224.1.1.1	UDP	Source port: 7817 Destination port: 0

<R5>display pim routing-table 可以看到 R5 的上游节点地址为 20.0.2.254 ，即 R4

```
(172.16.1.1, 224.1.1.1)
  Protocol: pim-dm, Flag:
  UpTime: 00:00:14
  Upstream interface: GigabitEthernet0/0/0
    Upstream neighbor: 20.0.2.254
    RPF prime neighbor: 20.0.2.254
  Downstream interface(s) information:
  Total number of downstreams: 1
    1: GigabitEthernet0/0/1
      Protocol: pim-dm, UpTime: 00:00:14, Expires:
```

[R5]

=====

## 组播实验 4 : PIM-SM

PIM-SM 是一种基于 Group-Shared Tree 的组播路由协

议，与 PIM-DM 不同，它适合于组播成员分布广泛而稀疏的大型网络。Group-Shared Tree 分为两种，一种被称为 Steinet Tree,另一种被称为 Rendezvous Point Tree ( 简称 RPT )，PIM-SM 采用的组播树是 RPT.

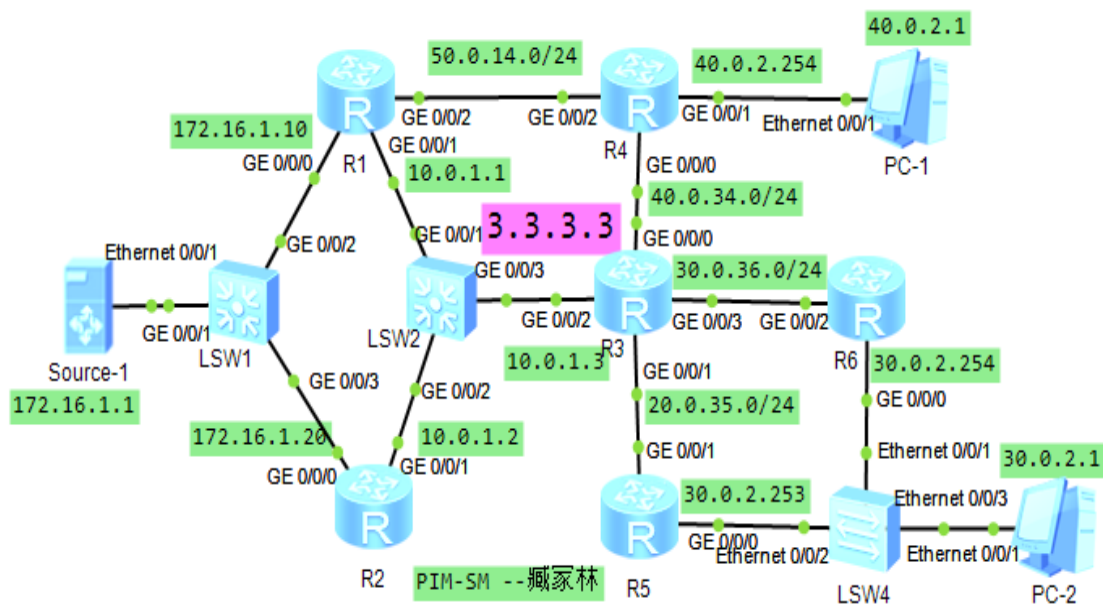
RPT 是一棵以汇聚点 RP ( Rendezvous Point ) 路由为根，以直连有组成员的路由器为叶子的组播树。RP 是一个组播供求信息的汇聚中心，它需要处理组播源端 DR 发送的组播注册消息及用户端 DR 发送的组播加入消息。

RP 是 PIM-SM 网络中的一台至关重要的路由器，网络中的所有 PIM 路由器都必须知道谁是 DR。当网络中出现活跃的组播源 ( 组播源向某组播组发送第一个组播数据 ) 时，组播源端 DR 会将此组播数据封装在注册消息中并以单播形式发往 RP，RP 收到此消息后会立即创建相应的 ( S,G ) 组播路由表项。当网络中出现活跃的组播用户 ( 用户主机通过 IGMP 加入某组播组 G ) 时，用户端 DR 会向 RP 发送加入组播组 G 的消息，在该消息去往 RP 的路径上的每台路由器都创建 ( \*, G ) 表项，由此便生成了一棵以 RP 为根的 RPT。当网络中有活跃的组播用户时，组播报文先被封装在单播报文中从组播源发往 RP，然后 RP 再将组播报文沿 RPT 转发给组播用户。若网络中没有活跃的组播用户时，组播数据到达 RP 后就停止了，不会再向下转发。

显然，RPT 并非是一棵 STP ( Shortest Path Tree )，经由 RP 的转发路径可能不是从组播源到组播用户之间的最短路径。为了提高组播转发效率，PIM-SM 在实际部署时，通常都会允许从 RPT 切换到 SPT.

PIM 路由会通过 PIM-Hello 消息来发现 PIM 邻居、协

调各项协议参数、维护邻居关系。PIM-Hello 消息的目的 IP 地址是组播地址 224.0.0.13（表示同一网段中的所有 PIM 路由器），源地址为发送接口的 IP 地址，TTL 值为 1。另外，PIM-Hello 的一个重要作用就是用来选举 DR。在 PIM-SM 中，DR 分为两种，组播源网段中的 DR 称为组播源端 DR，它负责向 RP 发送组播注册消息，组成员网络中的 DR 称为用户端 DR，它负责向 RP 发送组播加入消息。



## 基本配置

R1:

```
undo ter mo
```

```
sys
```

```
sysname R1
```

```
user-interface console 0
```

```
idle-timeout 0 0
```

```
int g0/0/0
```

```
ip add 172.16.1.10 24
```

```
int g0/0/1
```

```
ip add 10.0.1.1 24
```

```
int g0/0/2
ip add 50.0.14.1 24
q
```

```
R2:
undo ter mo
sys
sysname R2
user-interface console 0
idle-timeout 0 0
int g0/0/0
ip add 172.16.1.20 24
int g0/0/1
ip add 10.0.1.2 24
q
```

```
R3:
undo ter mo
sys
sysname R3
user-interface console 0
idle-timeout 0 0
int loo 0
ip add 3.3.3.3 32
int g0/0/2
ip add 10.0.1.3 24
int g0/0/1
ip add 20.0.35.3 24
int g0/0/3
ip add 30.0.36.3 24
int g0/0/0
ip add 40.0.34.3 24
q
```

R4:  
undo ter mo  
sys  
sysname R4  
user-interface console 0  
idle-timeout 0 0  
int g0/0/2  
ip add 50.0.14.4 24  
int g0/0/0  
ip add 40.0.34.4 24  
int g0/0/1  
ip add 40.0.2.254 24  
q

R5:  
undo ter mo  
sys  
sysname R5  
user-interface console 0  
idle-timeout 0 0  
int g0/0/1  
ip add 20.0.35.5 24  
int g0/0/0  
ip add 30.0.2.253 24  
q

R6:  
undo ter mo  
sys  
sysname R6  
user-interface console 0  
idle-timeout 0 0  
int g0/0/0  
ip add 30.0.2.254 24

```
int g0/0/2
ip add 30.0.36.6 24
q
```

配置 IGP 协议 OSPF

```
R1 :
ospf
area 0
net 172.16.1.10 0.0.0.0
net 10.0.1.1 0.0.0.0
net 50.0.14.1 0.0.0.0
q
```

```
R2 :
ospf
area 0
net 172.16.1.20 0.0.0.0
net 10.0.1.2 0.0.0.0
q
```

```
R3 :
ospf
area 0
net 3.3.3.3 0.0.0.0
net 10.0.1.3 0.0.0.0
net 20.0.35.3 0.0.0.0
net 30.0.36.3 0.0.0.0
net 40.0.34.3 0.0.0.0
q
```

```
R4 :
ospf
area 0
```

```
net 50.0.14.4 0.0.0.0
net 40.0.34.4 0.0.0.0
net 40.0.2.254 0.0.0.0
q
```

```
R5 :
ospf
area 0
net 20.0.35.5 0.0.0.0
net 30.0.2.253 0.0.0.0
q
```

```
R6 :
ospf
area 0
net 30.0.2.254 0.0.0.0
net 30.0.36.6 0.0.0.0
q
```

在 R3 上查看 OSPF 邻居的建立 <R3>display ospf peer brief

```
[R3]dis ospf peer bri
```

```
OSPF Process 1 with Router ID 3.3.3.3  
Peer Statistic Information
```

Area Id	Interface	Neighbor id	State
0.0.0.0	GigabitEthernet0/0/2	172.16.1.10	Full
0.0.0.0	GigabitEthernet0/0/2	172.16.1.20	Full
0.0.0.0	GigabitEthernet0/0/1	20.0.35.5	Full
0.0.0.0	GigabitEthernet0/0/3	30.0.2.254	Full
0.0.0.0	GigabitEthernet0/0/0	50.0.14.4	Full

```
=====
```

## 配置 PIM-SM

在所有路由器上开启组播功能，并在每台路由器的接口下配置 pim sm，在 R5 和 R6 的 g0/0/0，R4 的 g0/0/1 接口下开启 IGMP。

```
R1:  
multicast routing-enable  
int g0/0/0  
pim sm  
int g0/0/1  
pim sm  
int g0/0/2  
pim sm
```

```
R2:  
multicast routing-enable
```



```
int g0/0/0
pim sm
int g0/0/1
pim sm
```

```
R3:
multicast routing-enable
int g0/0/0
pim sm
int g0/0/1
pim sm
int g0/0/2
pim sm
int g0/0/3
pim sm
```

R4 R5 R6 连接 PC , 需要开启 igmp 协议

```
R4:
multicast routing-enable
int g0/0/0
pim sm
int g0/0/1
pim sm
igmp enable
int g0/0/2
pim sm
```

```
R5:
multicast routing-enable
int g0/0/0
pim sm
igmp enable
int g0/0/1
```

```
pim sm
```

```
R6:  
multicast routing-enable  
int g0/0/0  
pim sm  
igmp enable  
int g0/0/2  
pim sm
```

### 配置静态 RP

选择 R3 为 RP 节点，并在每台路由器上手工配置 R3 为静态 RP

```
R1:  
pim  
static-rp 3.3.3.3  
R2:  
pim  
static-rp 3.3.3.3  
R3:  
pim  
static-rp 3.3.3.3  
R4:  
pim  
static-rp 3.3.3.3  
R5:  
pim  
static-rp 3.3.3.3  
R6:  
pim  
static-rp 3.3.3.3
```

配置完成后，在 R3 上查看 PIM 邻居信息 <R3>display pim

## neighbor

```
[R3]display pim neighbor
```

```
VPN-Instance: public net
```

```
Total Number of Neighbors = 5
```

Neighbor	Interface	Uptime	Expires	Dr-Priority	BFD-Session
40.0.34.4	GE0/0/0	00:02:16	00:01:29	1	N
20.0.35.5	GE0/0/1	00:01:45	00:01:30	1	N
10.0.1.2	GE0/0/2	00:02:41	00:01:20	1	N
10.0.1.1	GE0/0/2	00:02:40	00:01:22	1	N
30.0.36.6	GE0/0/3	00:01:32	00:01:42	1	N

```
[R3]
```

=====

## 用户端 DR

拓扑中，PC-2 R5 R6 处于同一网段，如果 PC-2 希望加入组播组，则 PIM-SM 需要在 R5 R6 之间选举出一台用户端 DR 来发送组播加入消息，从而避免 RP 收到重复的加入消息。

DR 选举首先比较优先级，默认为 1,再比较 IP 地址，较大者成为 DR

```
[R5]display pim interface
```

```
VPN-Instance: public net
```

Interface	State	NbrCnt	HelloInt	DR-Pri	DR-Address
GE0/0/0	up	1	30	1	30.0.2.254
GE0/0/1	up	1	30	1	20.0.35.5 (local)

```
[R6]dis pim interface
```

VPN-Instance: public net

Interface	State	NbrCnt	HelloInt	DR-Pri	DR-Address	
GE0/0/0	up	1	30	1	30.0.2.254	(local)
GE0/0/2	up	1	30	1	30.0.36.6	(local)

可以看 R6 (30.0.2.254)现在是网段的 DR

PC-2 加入 224.1.1.1



在 R3 上查看 PIM 路由表 <R3>display pim routing-table

```
[R3]dis pim routing-table
VPN-Instance: public net
Total 1 (*, G) entry; 0 (S, G) entry

(*, 224.1.1.1)
  RP: 3.3.3.3 (local)
  Protocol: pim-sm, Flag: WC
  UpTime: 00:00:10
  Upstream interface: Register
    Upstream neighbor: NULL
    RPF prime neighbor: NULL
  Downstream interface(s) information:
    Total number of downstreams: 1
    1: GigabitEthernet0/0/3
        Protocol: pim-sm, UpTime: 00:00:10, Expires: 00:03:20
```

可以看到 R3 在收到加入消息后，其组播路由表中生成了 ( \*, G ) 的组播路由条目，下游接口为 g0/0/3，形成了从 RP R3 到 R6 的一棵 RPT

=====

## 组播源端 DR

组播源 source-1，R1，R2 处于同一网段，PIM-SM 需要在 R1 和 R2 之间选举出组播端 DR 来向 RP 发送注册消息。

```
<R1>display pim interface
```

```
[R1]dis pim int
VPN-Instance: public net
Interface          State NbrCnt HelloInt  DR-Pri  DR-Address
GE0/0/0            up    1     30        1       172.16.1.20
GE0/0/1            up    2     30        1       10.0.1.3
GE0/0/2            up    1     30        1       50.0.14.4
```

```
[R2]dis pim interface
VPN-Instance: public net
Interface          State NbrCnt HelloInt  DR-Pri  DR-Address
GE0/0/0            up    1     30        1       172.16.1.20 (local)
GE0/0/1            up    2     30        1       10.0.1.3
```

可以看到 IP 地址较大的 R2 172.16.1.20 目前是组播源端 DR。

修改优先级，让 R1 成为组播源端 DR

R1:

```
int g0/0/0
pim hello-option dr-priority 9
```

```
[R1]dis pim interface
VPN-Instance: public net
Interface          State NbrCnt HelloInt  DR-Pri  DR-Address
GE0/0/0            up    1     30        9       172.16.1.10 (local)
GE0/0/1            up    2     30        1       10.0.1.3
GE0/0/2            up    1     30        1       50.0.14.4
```

=====

从 RPT 切换到 SPT

切换过程大体如下：最后一跳组播路由器 R4 收到来自上游路由器 R3 转发的组播数据包后，会查看自己的单播路由表，发现去往组播源 172.16.1.1 的最短路径的下一跳不是上游路由器 R3，而是路由器 R1，因此 R4 会发起由 RPT 到 SPT 的切换。

在 PC-1 加入组播组 224.1.1.1



PC-1

基础配置 命令行 组播 UDP发包工具

数据配置

IGMP 版本: ☐ Version 1 ☒ Version 2 ☐ Version 3

源 IP: 40 . 0 . 2 . 1 源 MAC: 54-89-98-6D-6D-36 启动 VLC

目的 IP: 224 . 1 . 1 . 1 目的 MAC: 01-00-5E-01-01-01 加入 离开

在 R4 上查看组播路由表 <R4>display pim routing-table

```
[R4]dis pim routing-table
VPN-Instance: public net
Total 1 (*, G) entry; 0 (S, G) entry

(*, 224.1.1.1)
  RP: 3.3.3.3
  Protocol: pim-sm, Flag: WC
  UpTime: 00:00:06
  Upstream interface: GigabitEthernet0/0/0
    Upstream neighbor: 40.0.34.3
    RPF prime neighbor: 40.0.34.3
  Downstream interface(s) information:
  Total number of downstreams: 1
    1: GigabitEthernet0/0/1
```

R4 生成了 ( \*, 224.1.1.1 ) 的组播路由，上游接口是连接到 RP 路由器的 g0/0/0

在组播源端发送组播地址为 224.1.1.1 的组播报文





当 PC-1 接收到组播报文后，在 R4 上查看组播路由表

(172.16.1.1, 224.1.1.1)

RP: 3.3.3.3

Protocol: pim-sm, Flag: SPT ACT

UpTime: 00:00:08

Upstream interface: GigabitEthernet0/0/2

Upstream neighbor: 50.0.14.1

RPF prime neighbor: 50.0.14.1

Downstream interface(s) information:

Total number of downstreams: 1

1: GigabitEthernet0/0/1

Protocol: pim-sm, UpTime: 00:00:08, Expires:

[R4]

R4 在接收到组播组数据后生成了 ( 172.16.1.1,224.1.1.1 ) 的组播路由，且上游接口切换到 g0/0/2。切换过程中，R4 会向 R1 发送组播加入消息，要求从 R1 接收组播数据，同时 R4 也会向 R3 发送剪枝消息，使 R3 停止向自己转发组播组数据。

默认，PIM-SM 路由器会在收到第一个组播数据包后立即进行从 RPT 到 SPT 的切换。如果不希望发生切换，则可使用配置命令来禁止切换。也可以设定切换阈值，实现有条件的切换。

### PC-1 离开组

```
R4 :  
pim  
spt-switch-threshold infinity
```

配置完成后，在 PC-1 上使用 IGMP 加入组播组 224.1.1.1 查看 R4 的组播路由表

```
(172.16.1.1, 224.1.1.1)
  RP: 3.3.3.3
  Protocol: pim-sm, Flag: ACT
  UpTime: 00:02:37
  Upstream interface: GigabitEthernet0/0/0
    Upstream neighbor: 40.0.34.3
    RPF prime neighbor: 40.0.34.3
  Downstream interface(s) information:
  Total number of downstreams: 1
    1: GigabitEthernet0/0/1
      Protocol: pim-sm, UpTime: 00:00:12, Expires: -
```

[R4]

可以看到，当配置禁止切换后，R4 生成的（172.16.1.1, 224.1.1.1）的组播路由的上游接口没有发生切换。

=====

## 组播实验 5：PIM-SM 的 RP

一个 PIM-SM 网络中可以存在一个或多个 RP。一个 RP 可以对应若干个组播组，负责这些组播组的注册消息的处理、加入消息的处理以及组播数据的转发，但是同一个组播组只能对应一个 RP。RP 是 PIM-SM 网络的核心，网络中的路由器必须知道 RP 的地址。

RP 有两种形式：静态 RP 和动态 RP。

静态 RP 是由人工选定的，PIM 网络中的所有 PIM 路由器上都需要逐一进行配置。通过配置，每台路由便知道了静态 RP

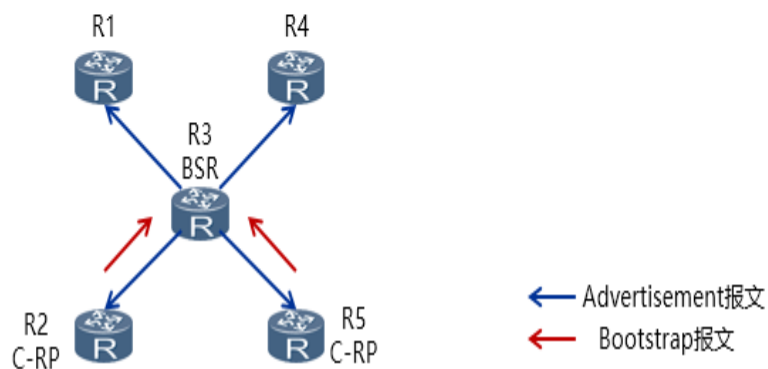
的地址。

动态 RP 的确定过程相对比较复杂一些，在 PIM 网络中人工选定并配置若干台 PIM 路由器，使得它们成为 C-RP( Candidate-RP ),RP 将从 C-RP 中选举产生。如果 C-RP 只有一个，则 RP 就是这个 C-RP。如果有多个 C-RP，则优先级最高者（优先级数值越小优先级越高，缺省值是 0）将会被选举为 RP，如果通过优先级比较无法选举出 RP，则依靠 Hash 算法算出的数值来决定 RP，数值最大者将成为 RP（Hash 算法参数：组地址、掩码长度、C-RP 地址），如果通过 Hash 数值也无法确定出 RP，则拥有最高 IP 地址的 C-RP 将成为 RP。

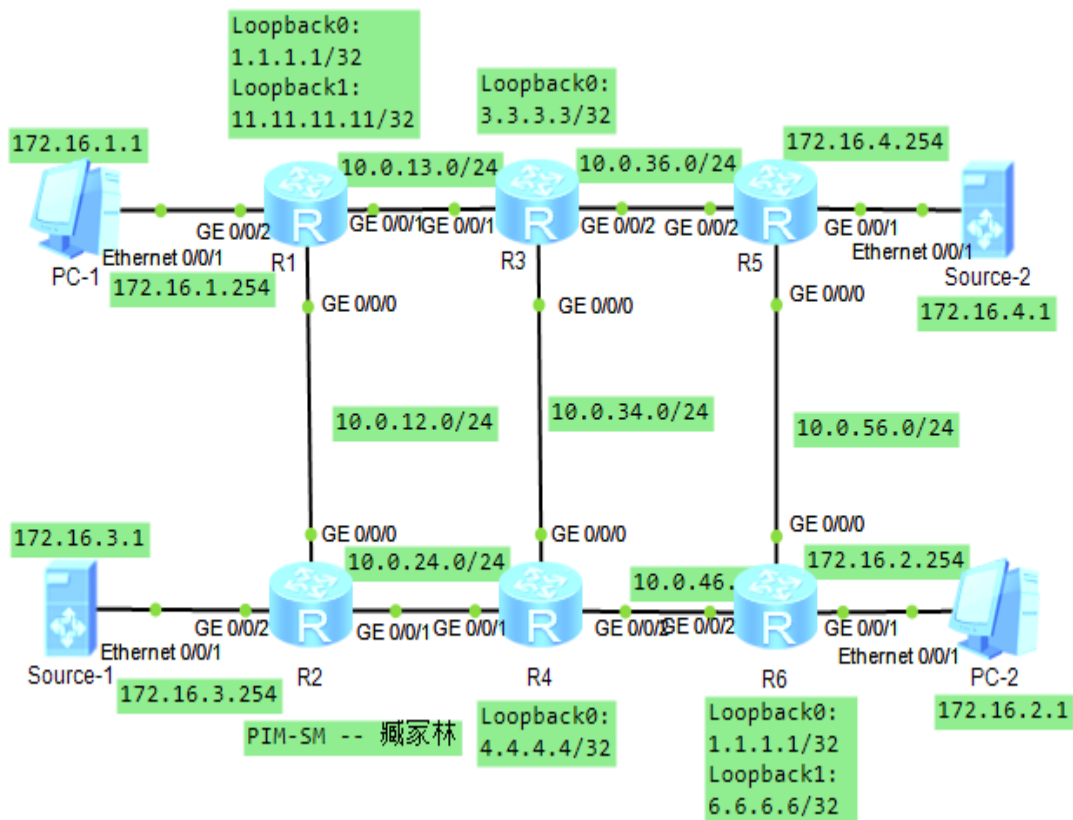
选定和配置 C-RP 时，还必须同时选定和配置 C-BSR（Candidate-Bootstrap Router），并由 C-BSR 选举产生出一个 BSR。如果有多个 C-BSR，则拥有最高 IP 地址的 C-BSR 将成为 BSR。BSR 是 PIM-SM 网络的管理核心，它负责收集网络中 C-RP 发出的 Advertisement 宣告信息，并计算出与每个组播组对应的 RP，然后将 RP 的信息发布到整个 PIM-SM 网络中。[BSR 动态映射组播组与 RP 的关系](#)

## 动态RP

- 选举BSR
  - C-BSR组播发送Bootstrap报文，比较Bootstrap报文选举BSR
- 选举RP
  - C-RP向BSR单播发送Advertisement报文
  - BSR将信息汇总为RP-Set，封装在Bootstrap报文中，发布给全网的每一台PIM-SM路由器
  - 全网PIM-SM路由器根据Bootstrap报文包含的信息选举RP



在传统的 PIM-SM 网络中，每个组播组只能映射到一个 RP，当网络负载较大以及流量分布不合理时，可能导致 RP 拥塞或者网络资源严重浪费的情况。解决上述问题的一个方案便是配置 Anycast RP，在同一个 PIM-SM 网络中设置多个具有相同环回地址的 RP，组播源和组播用户分别选择距离自己最近的 RP 进行 RPT 的创建，从而实现分担和优化组播流量的目的。



## 基本配置

R1:

```
undo ter mo
```

```
sys
```

```
sysname R1
```

```
user-interface console 0
```

```
idle-timeout 0 0
```

```
int loop 0
```

```
ip add 1.1.1.1 32
```

```
int loop 1
```

```
ip add 11.11.11.11 32
```

```
int g0/0/2
```

```
ip add 172.16.1.254 24
```

```
int g0/0/1
```

```
ip add 10.0.13.1 24
```

```
int g0/0/0
```

```
ip add 10.0.12.1 24
```

q

R2:

```
undo ter mo
sys
sysname R2
user-interface console 0
idle-timeout 0 0
int g0/0/2
ip add 172.16.3.254 24
int g0/0/1
ip add 10.0.24.2 24
int g0/0/0
ip add 10.0.12.2 24
q
```

R3:

```
undo ter mo
sys
sysname R3
user-interface console 0
idle-timeout 0 0
int loop 0
ip add 3.3.3.3 32
int g0/0/1
ip add 10.0.13.3 24
int g0/0/2
ip add 10.0.36.3 24
int g0/0/0
ip add 10.0.34.3 24
q
```

R4:

```
undo ter mo
```

```
sys
sysname R4
user-interface console 0
idle-timeout 0 0
int loop 0
ip add 4.4.4.4 32
int g0/0/1
ip add 10.0.24.4 24
int g0/0/2
ip add 10.0.46.4 24
int g0/0/0
ip add 10.0.34.4 24
q
```

```
R5:
undo ter mo
sys
sysname R5
user-interface console 0
idle-timeout 0 0
int g0/0/2
ip add 10.0.36.5 24
int g0/0/1
ip add 172.16.4.254 24
int g0/0/0
ip add 10.0.56.5 24
q
```

```
R6:
undo ter mo
sys
sysname R6
user-interface console 0
idle-timeout 0 0
```



```
int loop 0
ip add 1.1.1.1 32
int loop 1
ip add 6.6.6.6 32
int g0/0/2
ip add 10.0.46.6 24
int g0/0/1
ip add 172.16.2.254 24
int g0/0/0
ip add 10.0.56.6 24
q
```

### 配置 IGP 协议 OSPF

```
R1:
ospf
area 0
net 1.1.1.1 0.0.0.0
net 11.11.11.11 0.0.0.0
net 172.16.1.254 0.0.0.0
net 10.0.13.1 0.0.0.0
net 10.0.12.1 0.0.0.0
q
```

```
R2:
ospf
area 0
net 172.16.3.254 0.0.0.0
net 10.0.24.2 0.0.0.0
net 10.0.12.2 0.0.0.0
q
```

```
R3:
ospf
area 0
```

```
net 3.3.3.3 0.0.0.0
net 10.0.13.3 0.0.0.0
net 10.0.36.3 0.0.0.0
net 10.0.34.3 0.0.0.0
q
```

```
R4:
ospf
area 0
net 4.4.4.4 0.0.0.0
net 10.0.24.4 0.0.0.0
net 10.0.46.4 0.0.0.0
net 10.0.34.4 0.0.0.0
q
```

```
R5:
ospf
area 0
net 10.0.36.5 0.0.0.0
net 172.16.4.254 0.0.0.0
net 10.0.56.5 0.0.0.0
q
```

```
R6:
ospf
area 0
net 1.1.1.1 0.0.0.0
net 6.6.6.6 0.0.0.0
net 10.0.46.6 0.0.0.0
net 172.16.2.254 0.0.0.0
net 10.0.56.6 0.0.0.0
q
```

查看 OSPF 邻居的建立情况 <R1>display ospf peer brief

```
[R1]dis ospf peer bri
```

```
OSPF Process 1 with Router ID 1.1.1.1
Peer Statistic Information
```

Area Id	Interface	Neighbor id	State
0.0.0.0	GigabitEthernet0/0/1	3.3.3.3	Full
0.0.0.0	GigabitEthernet0/0/0	172.16.3.254	Full

=====

## 配置 PIM-SM 和静态 RP

在所有路由器上开启组播功能，并在接口下配置 pim sm  
R1 g0/0/2 和 R6 g0/0/1 接口下开启 IGMP

R1:

```
multicast routing-enable
```

```
int g0/0/0
```

```
pim sm
```

```
int g0/0/1
```

```
pim sm
```

```
int g0/0/2
```

```
pim sm
```

```
igmp enable
```

```
q
```

R2:

```
multicast routing-enable
int g0/0/0
pim sm
int g0/0/1
pim sm
int g0/0/2
pim sm
q
```

```
R3:
multicast routing-enable
int g0/0/0
pim sm
int g0/0/1
pim sm
int g0/0/2
pim sm
q
```

```
R4:
multicast routing-enable
int g0/0/0
pim sm
int g0/0/1
pim sm
int g0/0/2
pim sm
q
```

```
R5:
multicast routing-enable
int g0/0/0
pim sm
int g0/0/1
```

```
pim sm
int g0/0/2
pim sm
q
```

```
R6:
multicast routing-enable
int g0/0/0
pim sm
int g0/0/1
pim sm
igmp enable
int g0/0/2
pim sm
q
```

配置完成后，查看 R1 的 PIM 邻居信息 <R1>display pim neighbor

```
[R1]dis pim neighbor
VPN-Instance: public net
Total Number of Neighbors = 2
```

Neighbor	Interface	Uptime	Expires	Dr-Priority	BFD-Session
10.0.12.2	GE0/0/0	00:01:37	00:01:38	1	N
10.0.13.3	GE0/0/1	00:01:17	00:01:28	1	N

手工配置 R1 的 loopback1 : 11.11.11.11 为静态 RP

```
R1:
int LoopBack 1
pim sm
pim
```

```
static-rp 11.11.11.11
```

R2:

```
pim
```

```
static-rp 11.11.11.11
```

R3:

```
pim
```

```
static-rp 11.11.11.11
```

R4:

```
pim
```

```
static-rp 11.11.11.11
```

R5:

```
pim
```

```
static-rp 11.11.11.11
```

R6:

```
pim
```

```
static-rp 11.11.11.11
```

配置完成后，在 R1 上查看 RP 信息 <R1>display pim  
rp-info

```
[R1]dis pim rp-info
```

```
VPN-Instance: public net
```

```
PIM SM static RP Number:1
```

```
Static RP: 11.11.11.11 (local)
```

```
[R1]
```

=====

## 配置动态 RP

选定并配置 R1 和 R6 为 C-RP，R1 使用 LoopBack 0 为 RP  
接口，R6 使用 LoopBack 1 为 RP 接口.同时，选定并配置 R

3 和 R4 为 C-BSR , R3 和 R4 都使用 LoopBack 0 作为 C-B SR 接口。

R1 使用 LoopBack 0 IP 地址 : 1.1.1.1

R6 使用 LoopBack 1 IP 地址 : 6.6.6.6

R1 :  
int LoopBack 0  
pim sm  
pim  
c-rp LoopBack 0

R6 :  
int LoopBack 1  
pim sm  
pim  
c-rp LoopBack 1

R3 :  
int LoopBack 0  
pim sm  
pim  
c-bsr LoopBack 0

R4 :  
int LoopBack 0  
pim sm  
pim  
c-bsr LoopBack 0

配置 source-1 的 IP 地址为 224.1.1.1

Source-1

基础配置 组播源

文件路径:  浏览...

配置

组播组IP地址:	<input type="text" value="224 . 1 . 1 . 1"/>
组播组MAC地址:	<input type="text" value="01-00-5E-01-01-01"/>
源IP地址:	<input type="text" value="172 . 16 . 3 . 1"/>
源MAC地址:	<input type="text" value="54-89-98-4B-03-AA"/>

配置 source-2 的 IP 地址为 225.1.1.1 , MAC 地址为 01-00-5E-02-02-02

Source-2

基础配置 组播源

文件路径:  浏览...

配置

组播组IP地址:	<input type="text" value="225 . 1 . 1 . 1"/>
组播组MAC地址:	<input type="text" value="01-00-5E-02-02-02"/>
源IP地址:	<input type="text" value="172 . 16 . 4 . 1"/>
源MAC地址:	<input type="text" value="54-89-98-80-2B-BF"/>



配置完成后，在 R1 上查看 RP 和 BSR 信息

```
<R1>display pim rp-info
```

```
[R1]dis pim rp-info
```

```
VPN-Instance: public net
```

```
PIM-SM BSR RP Number:2
```

```
Group/MaskLen: 224.0.0.0/4
```

```
RP: 1.1.1.1 (local)
```

```
Priority: 0
```

```
Uptime: 00:01:26
```

```
Expires: 00:02:14
```

```
Group/MaskLen: 224.0.0.0/4
```

```
RP: 6.6.6.6
```

```
Priority: 0
```

```
Uptime: 00:01:26
```

```
Expires: 00:02:14
```

```
PIM SM static RP Number:1
```

```
Static RP: 11.11.11.11 (local)
```

```
<R1>display pim rp-info 224.1.1.1
```

```
[R1]dis pim rp-info 224.1.1.1
```

```
VPN-Instance: public net
```

```
BSR RP Address is: 6.6.6.6
```

```
Priority: 0
```

```
Uptime: 00:02:30
```

```
Expires: 00:02:10
```

```
Static RP Address is: 11.11.11.11
```

```
RP mapping for this group is: 6.6.6.6
```

可以看到 R1 和 R6 的 RP 优先级的值默认都为 0，当静态 RP 和动态 RP 同时存在时，动态 RP 优先。

在优先级相同时，IP 地址较大的 C-RP R6 被选为了 RP。

<R1>display pim bsr-info

```
[R1]dis pim bsr-info
VPN-Instance: public net
Elected AdminScoped BSR Count: 0
Elected BSR Address: 4.4.4.4
    Priority: 0
    Hash mask length: 30
    State: Accept Preferred
    Scope: Not scoped
    Uptime: 00:03:24
    Expires: 00:01:47
    C-RP Count: 2
```

另外 C-BSR 优先级相同时，IP 地址较大的 R4 成为了 BSR

通过修改优先级，可以控制 RP 的选举

R6:

pim

c-rp priority 9

配置完成后，重新在 R1 上查看 RP 的信息

<R1>display pim rp-info 224.1.1.1

要等一会才生效

```
[R1]dis pim rp-info 224.1.1.1
VPN-Instance: public net
BSR RP Address is: 1.1.1.1
    Priority: 0
    Uptime: 00:06:14
    Expires: 00:02:27
Static RP Address is: 11.11.11.11
RP mapping for this group is: 1.1.1.1 (local host)
```

可以看到，当 RP 优先级的值调整为 9，数值越小优先级越高，优先较高的 R1，默认为 0，成为了 RP

=====

## 配置 Anycast RP

若 R1 为 RP，那么当 source-2 发送组播数据，PC-2 接收时，组播源端 DR（R5）产生的注册消息和用户端 DR（R6）产生的加入消息都要发送给远处的 R1，另外，组播数据也要经历 R5-R1-R6 的绕行路径，浪费了链路带宽和路由器 CPU 资源。在这种情况下，配置 Anycast RP 便是一个不错的解决方案。

静态配置的灵活性最差，不适合大型的组播网络环境，而且没有备用 RP，一旦当前 RP 失效，组播流量就会中断。

动态 RP、BSR 优势是可以配置多个候选 RP，但是同一时间只能有一个 RP 生效，其它候选 RP 只是起到了备份的作用，而不能参与流量转发实现负载均衡。

为了能够在网络中配置多个 RP 既能实现备份功能，又能让所

有 RP 同时工作实现流量负载功能，Anycast RP 是一个很好的选择。Anycast RP 采用了一种很新颖的思路来实现多个 RP 的负载与冗余，它通过在组播网络中将 2 个或者 2 个以上的 RP 配置成相同的 RP 地址来实现冗余和备份功能，每个 RP 路由器上都创建一个 loopback 接口来充当 RP，并且该接口地址的掩码必须是/32 位的，然后将该地址通过动态路由协议发布到网络中，因为所有 RP 的这个 loopback 接口地址全部是一样的，所以能够实现冗余功能，当其中某一台 RP 出现故障后，流量可以很平滑地被转移到另一台 RP 上，这都是因为所有 RP 的地址是同一个地址。

```
R1 :  
pim  
anycast-rp 1.1.1.1  
local-address 11.11.11.11  
peer 6.6.6.6
```

```
R6:  
int LoopBack 0  
pim sm  
pim  
undo c-rp LoopBack 1  
undo c-rp priority  
c-rp LoopBack 0  
anycast-rp 1.1.1.1  
local-address 6.6.6.6  
peer 11.11.11.11
```

配置完成后，在 R1 R6 上查看 RP 信息 <R1>display pim  
rp-info 224.1.1.1

```
[R1]dis pim rp-info 224.1.1.1
VPN-Instance: public net
BSR RP Address is: 1.1.1.1
Priority: 0
Uptime: 00:11:56
Expires: 00:01:44
Static RP Address is: 11.11.11.11
RP mapping for this group is: 1.1.1.1 (local host)
```

```
<R6>display pim rp-info 225.1.1.1
[R6]display pim rp-info 225.1.1.1
VPN-Instance: public net
BSR RP Address is: 1.1.1.1
Priority: 0
Uptime: 00:12:50
Expires: 00:01:51
Static RP Address is: 11.11.11.11
RP mapping for this group is: 1.1.1.1 (local host)
```

可以看到，RP 为 1.1.1.1，R1 和 R6 都可充当 RP。

=====

## 组播实验 6：RPF 检验

所谓 RPF（Reverse Path Forwarding）检验，就是指在基于 Source-Based Tree 的组播网络（例如 PIM-SM 网络）中路由器通过查找去往组播源的最优单播路由来判断所收到的组播数据是否来源于“正确的”上游接口。某一路由器运往某一组播源的最优单播路由所对应的出接口称为该路由器上关

于该组播源的 RPF 接口。一台路由器从某一接口收到一个组播数据后，如果发现该接口不是相应的组播源的 RPF 接口，就意味着 RPF 校验失败，所收到的组播数据将丢弃，如果发现该接口正是相应组播源的 RPF 接口，就表明 RPF 校验通过，所收到的组播数据将进行后续处理。

正是因为有了 RPF 校验机制，基于 Source-Based Tree 的组播网络中所生成的组播树才能是一棵 SPT，同时 RPF 校验机制也防止了组播数据在转发过程中出现重复报文及流量环路的情况。另外，RPF 检验过程中所使用的单播路由可以来源于任何一种单播路由协议，并不依赖于某一特定的单播路由协议。

当然，为了某些特殊的需要，RPF 接口也是可以被人修改的，如果路由器上配置了组播静态路由，则 RPF 校验将首先依据组播静态路由而非单播路由。通过配置组播静态路由，可以在当前路由器上为特定的组播源人为指定一个 RPF。组播静态路由只在所配置的路由器上才有效，不会以任何方式传递给其他路由器。

