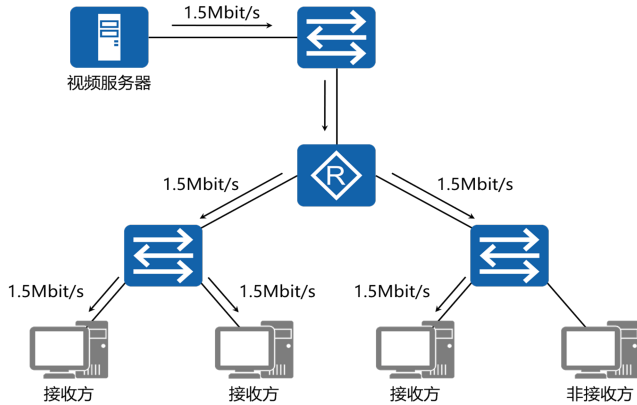


部署多层交换网络中的组播

一、组播的概念

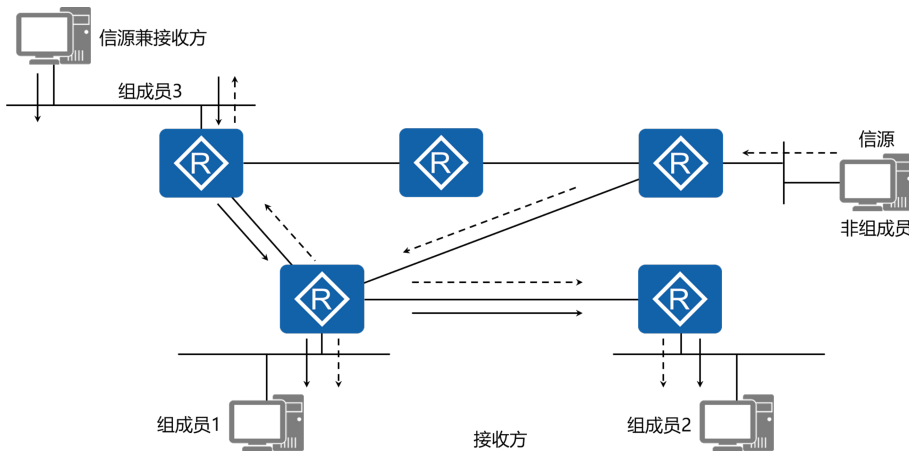
在当今互联网多媒体通信流中，其数据会以下列方式之一进行传输，这些方式占用的网络带宽各不相同

- 1、单播：数据包副本发送给每个客户的单播地址，因此单播存在验证的局限，若信息很大，则需传输多次
- 2、广播：所有客户端都将接收数据包，而无论其是否需要该数据，存在安全隐患，且浪费带宽
- 3、组播：介于单播与广播之间，数据包副本发送给一个特殊地址，加入该组的客户端会接收该数据，其余则不会，客户端可同时属于多个组



二、组播组成员资格

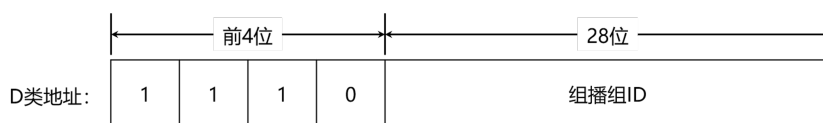
- 1、IP组播依赖于一个名为组播IP地址的虚拟组地址【D类地址】
- 2、接收方加入组播组，接收发送给该组的数据包
- 3、非成员客户端也能够将数据包发送给其他组播组客户端
- 4、同一个客户端可同时属于多个组播组



- 5、IP组播通信使用UDP作为传输层协议，因此组播继承UDP的所有特点【不保证可靠性，传输效率高】
- 6、不提供可靠性、流控和错误恢复
- 7、数据包报头短，开销少
- 8、可靠性由QoS负责提供

三、组播IP地址结构

- 1、组播使用D类地址224.0.0.0 — 239.255.255.255



- 2、组播共包含228个IP地址【超过2.68亿个】
- 3、D类地址只能用作组地址或IP组播通信流的目标地址
- 4、组播数据报的源地址总是单播地址

四、组播IP地址范围

1、组播IP地址可分为【动态】与【静态】两大类：

1.1、【动态】组播地址根据需要分配给相应的应用程序，动态组播地址具有固定的生命周期，应用程序仅当需要时才请求这种地址

1.2、【静态】组播地址：

1.2.1、保留的链路本地地址

IANA将224.0.0.0 — 224.0.0.255的地址保留给本地网段上的网络协议使用

其生存时间【TTL】值为1

路由器不转发这些地址的数据包

例如：

OSPF: 224.0.0.5、224.0.0.6

RIP: 224.0.0.9

VRRP: 224.0.0.18

EIGRP: 224.0.0.10

所有主机: 224.0.0.1【每台支持组播的主机都必须加入这个组播组，向该地址发送ICMP请求，所有主机都会回应】

所有路由器: 224.0.0.2【组播路由器中所有启用了组播功能的接口都将加入这个组播组】

1.2.2、全局范围地址

224.0.1.0 — 238.255.255.255的地址被称为全局范围地址

在组织之间和通过Internet以组播方式传输数据

例如：

NTP【Network Time Protocol】224.0.1.1

1.2.3、GLOP地址

233.0.0.0 — 233.255.255.255的地址让拥有自治系统号的组织静态地定义其用途，这种做法被称为GLOP编址
域的自治系统号被用作GLOP地址的第2个字节和第3个字节

例如：

自治系统号62010，其16进制表示为F23A

将其拆成两个字节后，分别为F2和3A，其对应的十进制表示为242和58

因此，233.242.58.0/24被全局保留，供自主系统62010使用

1.2.4、有限范围地址

239.0.0.0 — 239.255.255.255被称为有限范围地址

仅可在本地组或组织内使用这些组播地址

Internet边界路由器不会转发这些数据包【类似于私有IP地址】

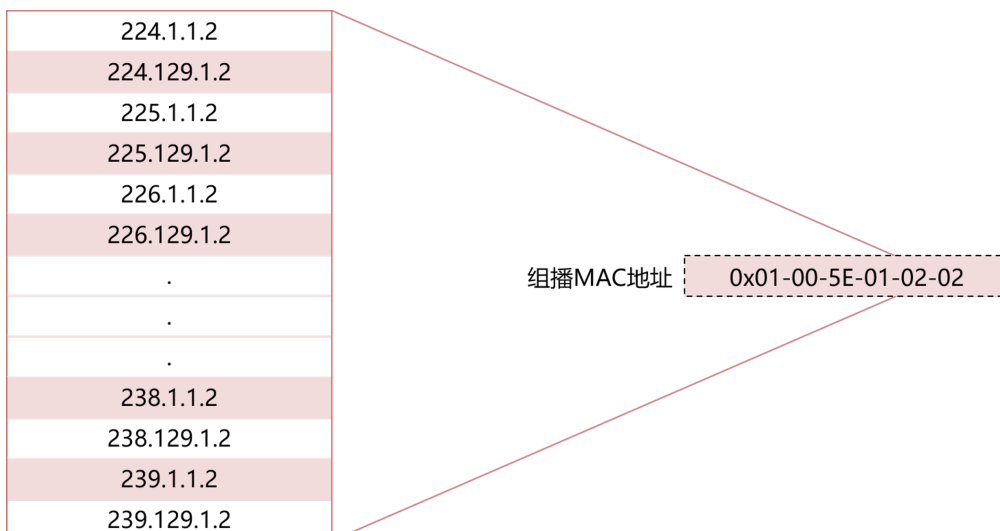
2、组播MAC地址的结构

2.1、IP组播数据包的目标IP地址对应于一个唯一的组播MAC地址

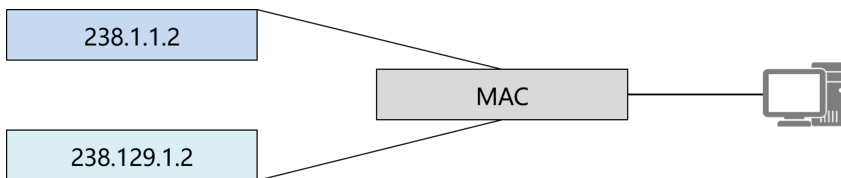
2.2、组播MAC地址的前24位为0x01-00-5E，第25位为0【前25位固定不变】

2.3、所有IP组播地址的前4位都是1110，需要将其余的28位映射到MAC地址的最后23位

2.4、在将组播IP地址映射到MAC地址的时候，没有使用其余的28位中的前5位，因此得到的映射关系为【32IP : 1MAC】



2.5、组播组中的主机，其网络接口监听IP地址映射的MAC地址，若该MAC地址还对应于另一个IP组播地址，则CPU将同时处理这两组IP组播帧



五、反向路径转发【RPF】

- 1、在单播网络环境中，路由通信流从单台源被路由到单台目的地
- 2、默认情况下，路由器不关心源地址，只在路由表中查找目标地址
- 3、路由器接收数据包，查找路由表，将数据包从正确的接口发送，若无法找到目的地则丢弃数据包
- 4、具备组播功能的路由器和多层交换机可以创建分发树，控制IP组播通信流在网络中的传输路径
- 5、反向路径转发【RPF】通过执行入站接口检查决定转发还是丢弃入站组播帧
- 6、RPF的主要功能是确保组播分发树不存在环路，且沿正确的方向转发组播通信流
- 7、RPF的工作原理：
 - 7.1、在组播转发中，信源将通信流发送给一个主机组，该组由一个组播地址表示
 - 7.2、组播路由器收到组播数据包后，确定上游方向【去往信源的方向】后，再确定下游方向【去往接收方的方向】
 - 7.3、当组播数据包通过RPF进程所确定的正确上游接口接收到时才转发
- 8、对于沿源树传输的通信流，RPF检查过程如下：
 - 8.1、为确定数据包到达的接口是否位于返回信源的反向路径上，路由器在单播路由选择表中查找源地址
 - 8.2、若数据包已经到达位于返回信源的反向路径上的接口，则RPF检查成功，路由器将复制数据包，并将其转发到出站接口
 - 8.3、若RPF检查失败，路由器丢弃数据包，并将这种丢弃标记为RPF失败丢弃

六、组播转发树

- 1、通过创建组播分发树，具备组播功能的路由器能够控制IP组播通信流在网络中的传输路径
- 2、组播分发树共有2大类型：
 - 2.1、源树【SPT | 最短路径树】
 - 2.1.1、最简单的组播分发树
 - 2.1.2、信源为根节点，通过网络去往接收方的分支组成一棵树
 - 2.1.3、源树使用通过网络去往接收方的最短路径【根据网络中运行的单播路由协议判断何为最短路径】
注：使用 (S, G) 表示法为：(192.168.1.1, 239.1.1.10)
 - 2.2、共享树【RPT】
 - 2.2.1、共享树将网络中某个点作为共享根，该共享根被称为RP【Rendezvous Point | 集结点，汇聚点】
 - 2.2.2、信源将数据包发送至RP，然后通信流沿共享树【RPT】被转发到所有接收方
 - 2.2.3、若接收方位于信源与RP之间，将直接收到该组播数据包
注：使用 (S, G) 表示法为：(*, 239.1.1.10)

3、SPT与RPT的比较

- 3.1、相同点：
 - 3.1.1、均不存在环路
 - 3.1.2、因为组播组成员能够随时加入或离开，因此分发树将动态更新分支上所有活动接收方都不再请求接收发送给组播组的通信流时，路由器将把该分支从树中剪除，不再沿这个分支转发通信流
 - 3.1.3、若该分支的某个接收方重新激活，并请求组播通信流，路由器将动态的修改分支，并再次转发数据包
- 3.2、源树的优点：
 - 3.2.1、在信源和接收方之间创建一条最优路径
 - 3.2.2、可最大限度降低转发延迟
- 3.3、源树的缺点：
 - 3.3.1、增加了额外的开销
 - 3.3.2、在包含数千个信源和数千个组播组的网络中带来资源不足的问题
 - 3.3.3、排错困难

3.4、共享树的优点：

- 3.4.1、每台路由器需要存储的状态信息少
- 3.4.2、降低了内存需求和复杂度

3.5、共享树的缺点：

- 3.5.1、选择的信源到接收方的路径不是最优的
- 3.5.2、造成大的延迟
- 3.5.3、过度使用某些链路，另其它链路空闲

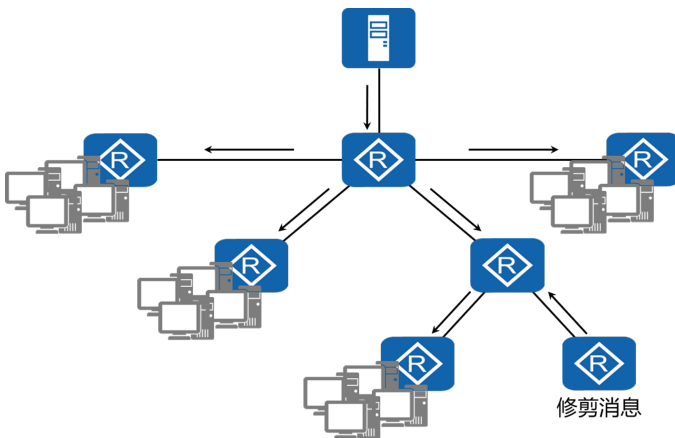
注：需仔细考虑RP的位置

七、IP组播路由协议

- 1、单播网络环境中存在6大动态路由协议：RIP、OSPF、IGRP、EIGRP、IS-IS、BGP
- 2、组播网络环境中亦存在诸多组播路由协议：PIM【协议无关组播】、MOSPF【组播OSPF】、DVMRP【距离矢量组播路由协议】、CBT【核心树】、IGMP【互联网组管理协议】、CGMP【思科组管理协议】、MSDP【组播源发现协议】等
- 3、HCIP重点学习和介绍PIM及IGMP

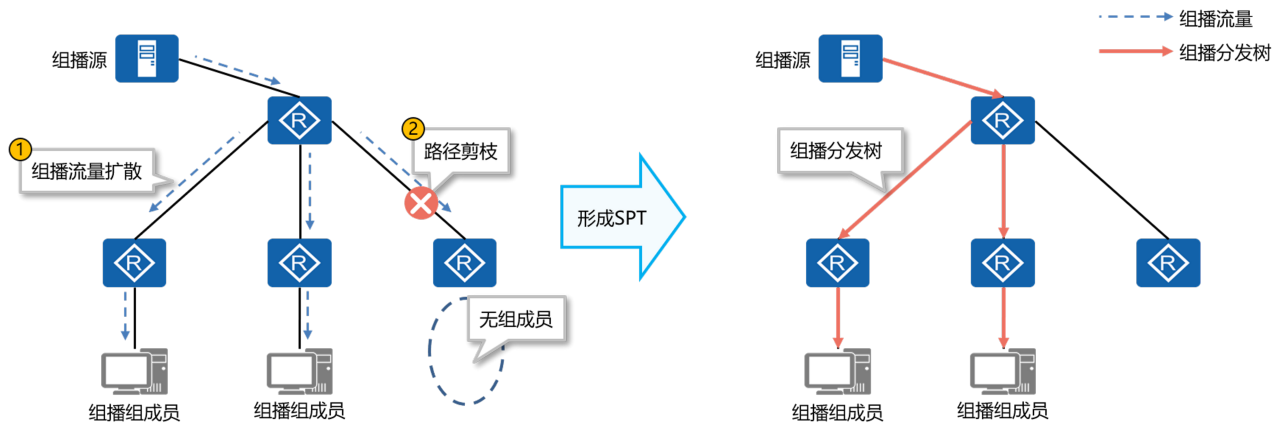
八、PIM【协议无关组播】

- 1、单播路由选择协议负责将网络联通，而组播路由选择协议负责建立组播递送树和转发组播数据包
- 2、路由器使用PIM邻居发现机制来确定PIM邻居
- 3、PIM邻居使用Hello消息来建立和维护PIM组播分发树
- 4、Hello消息通过组播方式发送，使用组播地址224.0.0.13【所有PIM路由器】
- 5、PIM拥有2大类运行模式：
 - 5.1、PIM-DM【密集模式】
 - 5.1.1、密集模式使用SPT来转发组播通信流
 - 5.1.2、密集模式适用于：发送方与接收方距离较近；发送方数目少，接收方数目多；组播流量高；组播流量稳定



九、PIM-DM基本概念

- 1、PIM-DM主要用在组成员较多且相对密集的网络中，通过“扩散-剪枝”的方式形成组播转发树【SPT】
- 2、PIM-DM在形成SPT的过程中，除了扩散【Flooding】，剪枝【Prune】机制外，还会涉及邻居发现【Neighbor Discovery】，嫁接【Graft】，断言【Assert】和状态刷新【State Refresh】机制



3、PIM-DM协议报文

3.1、PIM协议报文直接采用IP封装，目的地址224.0.0.13，IP协议号103

3.2、PIM-DM与PIM-SM使用的协议报文类型有所不同

3.3、PIM-DM使用报文主要是以下几类：

| 报文类型 | 报文功能 |
|-------------------|--|
| Hello | 用于PIM邻居发现，协议参数协商，PIM邻居关系维护等 |
| Join/Prune【加入/剪枝】 | 加入报文用于加入组播分发树，剪枝报文则用于修剪组播分发树。加入及剪枝报文在PIM中使用相同的报文格式，只不过报文载荷中的字段内容有所不同 |
| Graft【嫁接】 | 用于将设备所在的分支嫁接到组播分发树 |
| Graft-ACK【嫁接确认】 | 用于对邻居发送的Graft报文进行确认 |
| Assert【断言】 | 用于断言机制 |

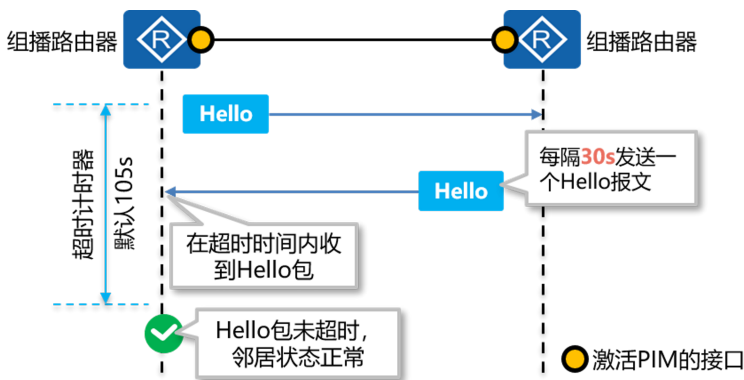
4、邻居发现机制

4.1、组播转发路径只能在PIM邻居之间建立，因此邻居发现是形成组播分发树的先决条件

4.2、邻居发现主要通过PIM Hello包完成

4.3、当路由器的接口激活PIM后，接口便周期性发送PIM Hello数据包，目的地址224.0.0.13。交互Hello报文后，组播路由器之间就能知道邻居信息，建立PIM邻居关系

4.4、PIM邻居关系依靠Hello包维持，邻居超时时间默认105s，如果超时时间内收不到邻居发来的Hello包，则删除邻居关系



5、首次形成组播分发树

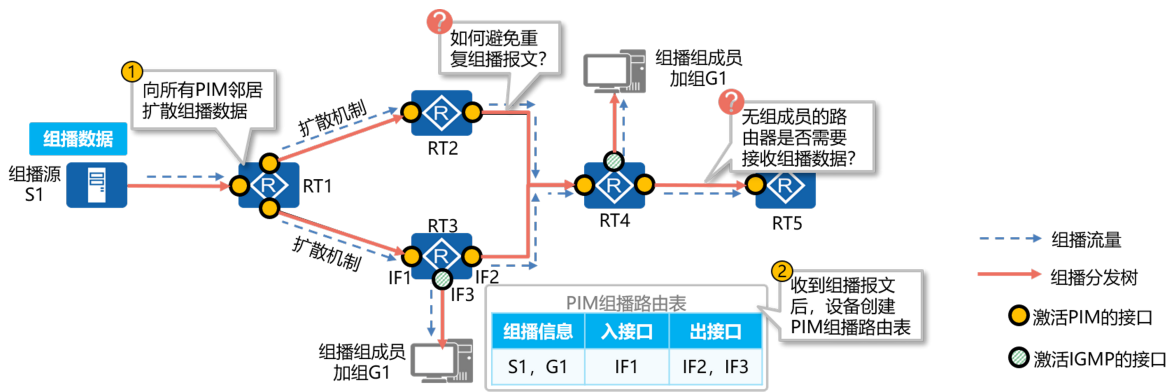
PIM-DM模式首次形成组播分发树主要依赖扩散机制、剪枝机制、断言机制与DR选举机制

5.1、扩散机制：组播数据包向所有的PIM邻居泛洪，同时组播路由器产生组播路由表项

5.1.1、组播源发送的组播报文会在全网内扩散。当PIM路由器接收到组播报文，先进行RPF检查，通过后会在该路由器上创建（S，G）表项，之后会向所有PIM邻居发送

5.1.2、PIM-DM形成的（S，G）表项有老化时间【默认210s】，如果老化时间超前没有收到新的组播报文，则删除（S，G）表项

5.1.3、扩散具体过程如下：



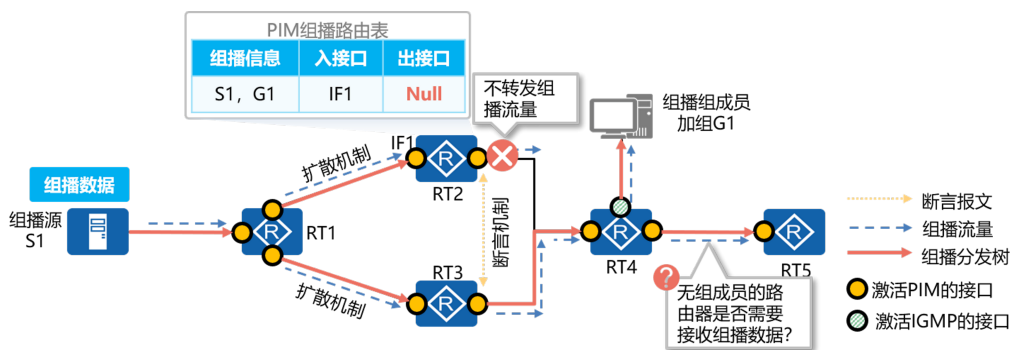
注: 扩散机制会周期性(默认180s)全网扩散组播数据, 周期性扩散的主要目的是探测是否有新成员加组, 但是由于全网扩散组播数据会浪费大量带宽, 所以现在的组播网络一般使用“状态刷新机制”加上“嫁接机制”来实现周期性全网扩散感知新成员加组的目的

5.2、断言机制: 当组播转发过程中存在多路访问网络, 则需要选举出一个组播转发路由器, 避免重复组播报文

5.2.1、当一个网段内有多个相连的PIM路由器向该网段转发组播报文时, 需要通过断言机制【Assert】来保证只有一个PIM路由器向该网段转发组播报文

5.2.2、通过断言机制的选举规则将决定组播路由器的转发行为:

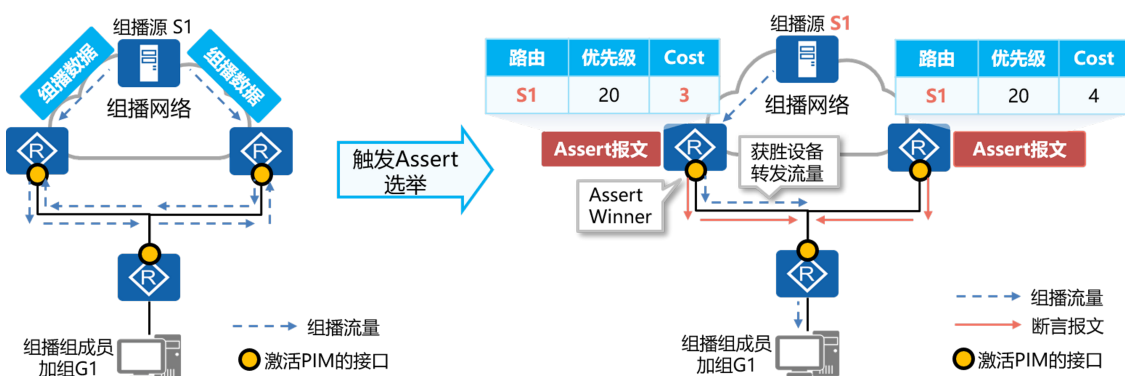
- 获胜一方的下游接口称为Assert Winner, 将负责后续对该网段组播报文的转发
- 落败一方的下游接口称为Assert Loser, 后续不会对该网段转发组播报文, PIM路由器也会将其从(S, G)表项下游接口列表中删除



5.2.3、断言机制选举规则

PIM路由器在接收到邻居路由器发送的相同组播报文后, 会向该网段发送断言【Assert】报文, 进行Assert选举。Assert报文内会携带到组播源的单播路由前缀, 路由优先级与开销。选举规则如下:

- 单播路由协议优先级较高者获胜
- 若优先级相同, 则到组播源的开销较小者获胜
- 若以上都相同, 则下游接口IP地址最大者获胜

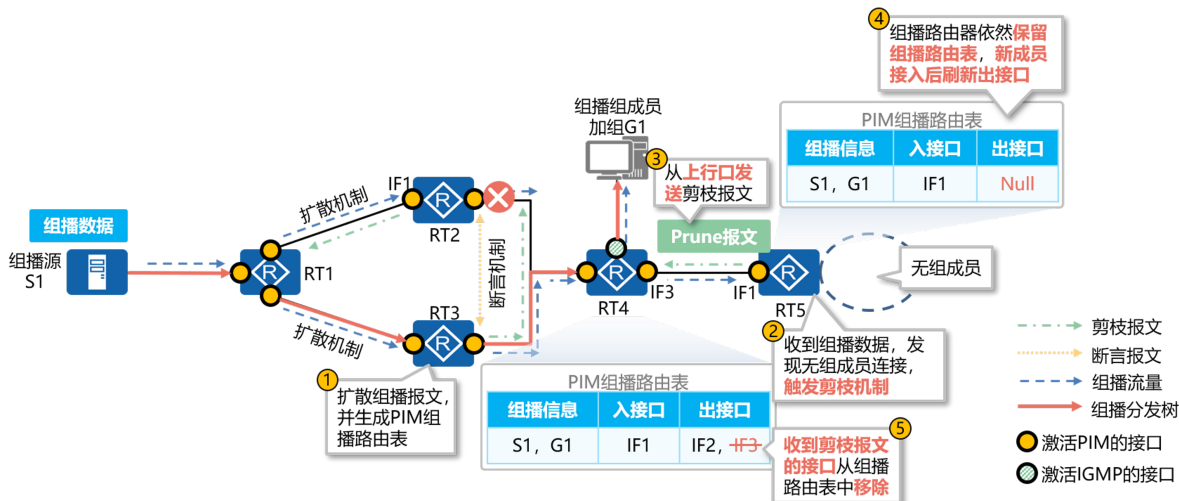


注: Assert选举失败的设备会抑制转发, 并将这种抑制转发的状态保持一段时间, 这段时间就被称为Assert保持时间, 默认180s。Assert保持时间超时后, 竞选失败的设备会恢复转发从而触发新一轮竞选

5.3、剪枝机制: 若组播路由器下没有组成员, 则将源到该组播路由器的组播转发路径剪枝

5.3.1、对于没有组成员连接的组播路由器，组播网络无需再将组播流量继续发往该设备。通过剪枝机制，组播网络可以将此类路径剪枝

5.3.2、剪枝机制工作原理如下：



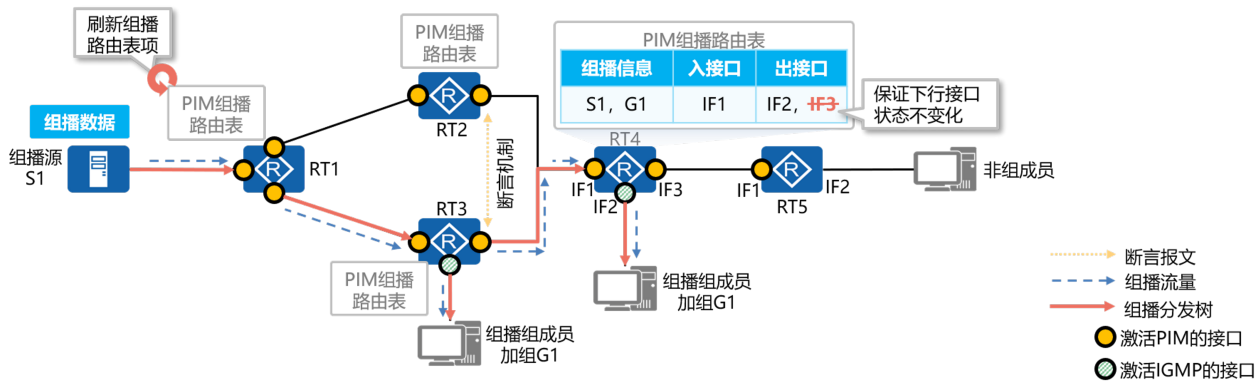
注：路由器为被裁剪的下游接口启动一个剪枝定时器（默认210s），定时器超时后接口恢复转发。组播报文重新在全网范围内扩散，新加入的组成员可以接收到组播报文。随后，下游不存在组成员的叶子路由器将向上发起剪枝操作。通过这种周期性的扩散-剪枝，PIM-DM周期性的刷新SPT

5.4、维护组播分发树

5.4.1、组播分发树形成后不会一直存在，也不会一直不变

5.4.2、在PIM邻居关系稳定，组成员没有变化的情况下，维护组播分发树一般有两种方式：

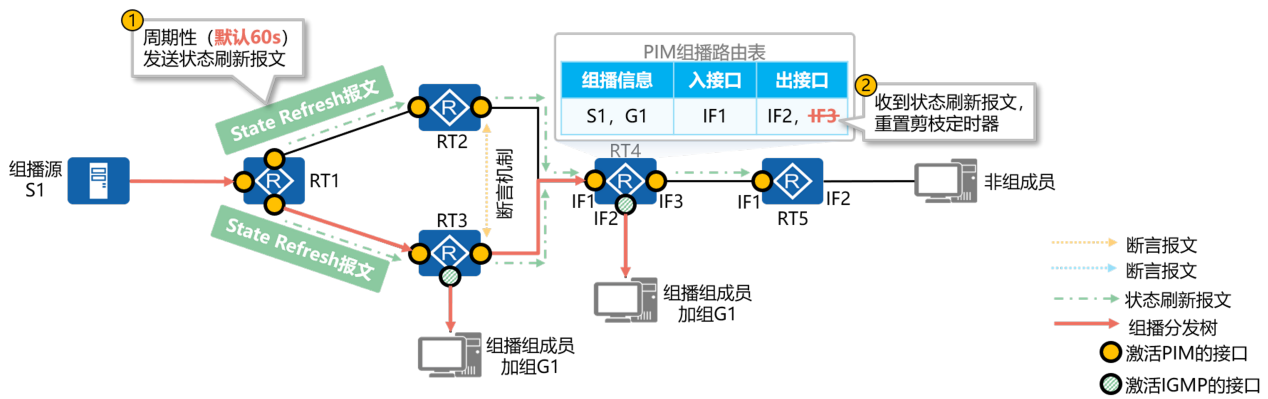
- a、持续发送组播报文，保证组播路由表项能一直存在
- b、发送状态刷新报文，保证组播路由表项的下行接口状态不发生变化



5.5、状态刷新机制

5.5.1、在PIM-DM网络中，为了避免被裁剪的接口因为“剪枝定时器”超时而恢复转发，离组播源最近的第一跳路由器会周期性地触发State Refresh报文在全网内扩散

5.5.2、收到State Refresh报文的PIM路由器会刷新剪枝定时器的状态。被裁剪接口的下游叶子路由器如果一直没有组成员加入，该接口将一直处于抑制转发状态



5.6、新成员加入

5.6.1、当有新成员加入组播组后，组播网络需要更新组播分发树，才能将组播数据发往组成员。PIM-DM模式在使用“扩散-剪枝”的方式建立组播分发树后，通过状态刷新机制，使下行接口一旦被抑制就无法自动恢复

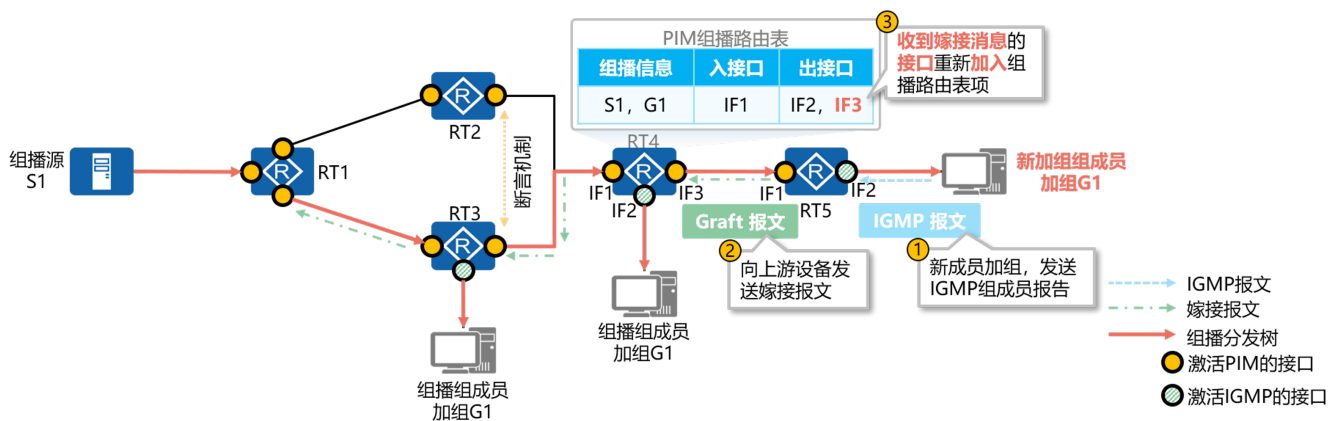
5.6.2、因此需要一些机制来更新组播分发树，一般PIM-DM模式更新组播分发树的方法有两种：

- a、等待组播路由表超时后，全网重新泛洪。该方法不可控，在现网中无法实现
- b、使用嫁接【Graft】机制，当新成员加组后，主动反向建立组播分发路径。现网中一般使用嫁接机制来实现新成员加组

5.7、嫁接机制

5.7.1、PIM-DM通过嫁接机制，使有新组成员加入的网段快速得到组播报文

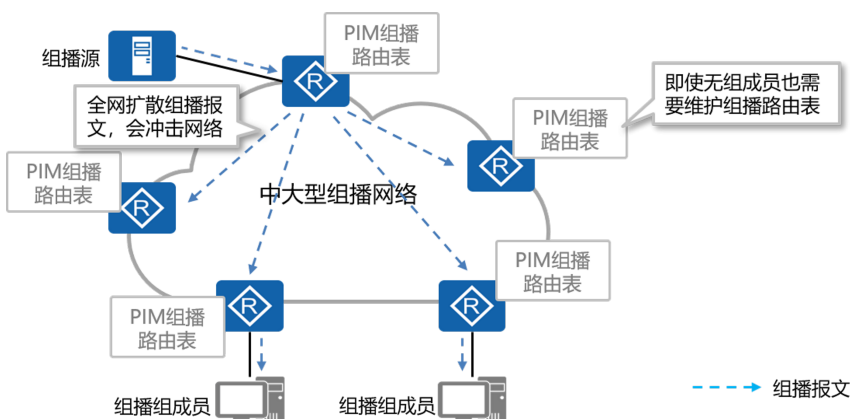
5.7.2、叶子路由器通过IGMP了解到与其相连的用户网段上，组播组G有新的组成员加入。随后叶子路由器会基于本地的组播路由表向上游发送Graft报文，请求上游路由器恢复相应出接口转发，将其添加在（S，G）表项下游接口列表中



5.8、PIM-DM的局限性

中大型组播网络中由于网络较大，若依然使用PIM-DM会遇到诸多问题：

- 5.8.1、使用“扩散-剪枝”方式需要全网扩散组播报文，对于网络有一定冲击
- 5.8.2、所有组播路由器均需要维护组播路由表，即使该组播路由器无需转发组播数据
- 5.8.3、对于组成员较为稀疏的组播网络，使用“扩散-剪枝”形成组播分发树的效率不高

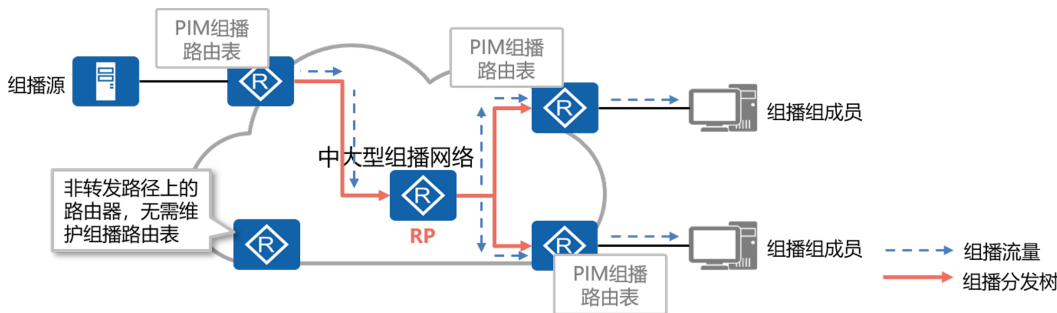


十、PIM-SM【ASM】

1、PIM-DM模型使用“扩散-剪枝”形成组播分发树的原因是：组播网络中大部分组播路由器无法得知组成员的位置

2、PIM-SM【ASM】模型形成组播分发树的方法是：

- 2.1、将组成员的位置事先告知某台组播路由器【Rendezvous Point | RP】，形成RPT【RP Tree】
- 2.2、组播源在发送组播数据时，组播网络先将组播数据发送至RP，然后由RP再将组播数据转发给组成员
- 2.3、对于部分次优的组播转发路径，PIM-SM【ASM】能自动优化为最优路径【SPT】



注：通过PIM-SM【ASM】模式形成组播分发树有如下好处：只有组播转发路径上的组播路由器需要维护组播路由表、通过RP可以让所有组播路由器获知组成员的位置、避免“扩散-剪枝”机制，提高组播分发树的形成效率

3、PIM — SM【ASM】协议报文

PIM协议报文直接采用IP封装，目的地址224.0.0.13，IP协议号103

PIM-SM使用报文主要是以下几类：

| 报文类型 | 报文功能 |
|------------------------------------|---|
| Hello | 用于PIM邻居发现，协议参数协商，PIM邻居关系维护等 |
| Register【注册】 | 用于事先源的注册过程。这是一种单播报文，在源的注册过程中，组播数据被第一跳路由器封装在单播注册报文中发往RP |
| Register-Stop【注册停止】 | RP使用该报文通知第一跳路由器停止通过注册报文发送组播流量 |
| Join/Prune【加入/剪枝】 | 加入报文用于加入组播分发树，剪枝则用于修剪组播分发树 |
| Assert【断言】 | 用于断言机制 |
| Bootstrap【自举】 | 用于BSR选举。另外BSR也使用该报文向网络中扩散C-RP【Candidate-RP 候选RP】的汇总信息 |
| Candidate-RP-Advertisement【候选RP通告】 | C-RP使用该报文向BSR发送通告，报文中包含该C-RP的IP地址及优先级等信息 |

4、RP简介

4.1、集结点 | 汇聚点RP【Rendezvous Point】为网络中一台重要的PIM路由器，用于处理源端DR注册信息及组成员加入请求，网络中的所有PIM路由器都必须知道RP的地址，类似于一个供求信息的汇聚中心

4.2、目前可以通过以下方式配置RP：

- 4.2.1、静态RP：在网络中的所有PIM路由器上配置相同的RP地址，静态指定RP的位置
- 4.2.2、动态RP：通过选举机制在多个C-RP【Candidate-RP | 候选RP】之间选举出RP

5、动态选举RP

动态选举RP会涉及两类角色C-BSR【Candidate-Bootstrap Router】与C-RP【Candidate-RP】：

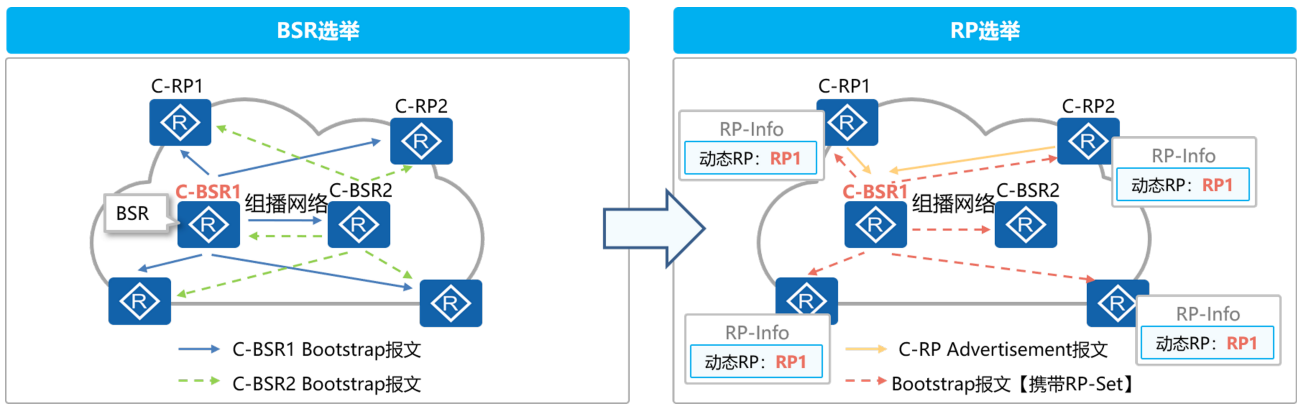
5.1、C-BSR通过竞选能选举出一个唯一的BSR

注：BSR竞选规则如下：优先级较高者获胜（优先级数值越大优先级越高）；若优先级相同，IP地址较大者获胜

5.2、BSR的作用是收集C-RP的信息并形成RP-Set信息，BSR通过PIM报文将RP-Set信息扩散给所有PIM路由器

5.3、PIM路由器收到RP-Set消息后，根据RP选举规则选举出合适的RP

RP竞选规则如下：与用户加入的组地址匹配的C-RP服务的组范围掩码最长者获胜；若以上比较结果相同，则C-RP优先级较高者获胜（优先级数值越小优先级越高）；若以上比较结果都相同，则执行Hash函数，计算结果较大者获胜；若以上比较结果都相同，则C-RP的IP地址较大者获胜



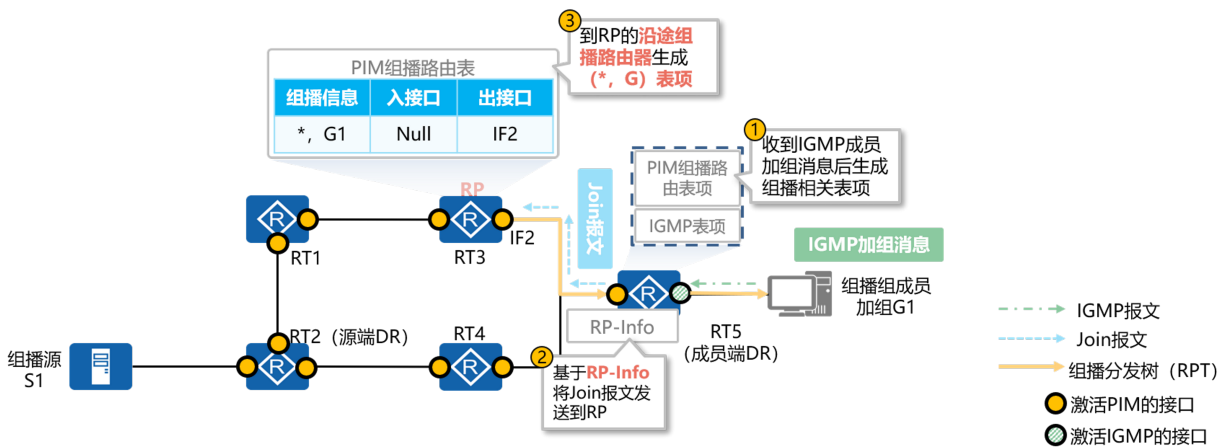
6、首次形成组播分发树

PIM-SM【ASM】模式首次形成组播分发树主要依赖RPT构建机制、组播源注册机制与DR选举机制

6.1、RPT构建机制：组播叶子路由器主动建立到RP的组播分发树【RPT】

6.1.1、RPT【RP Tree】是一棵以RP为根，以存在组成员关系的PIM路由器为叶子的组播分发树

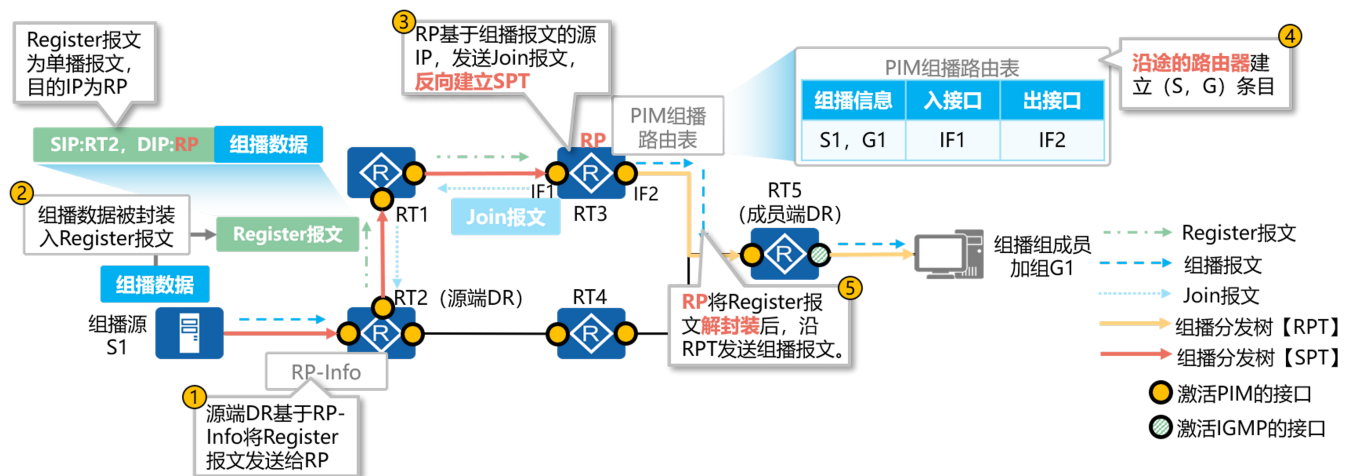
6.1.2、当网络中出现组成员【形成IGMP表项】时，组成员端DR向RP发送Join报文，在通向RP的路径上逐跳创建(*, G)表项，生成一棵以RP为根的RPT



6.2、组播源注册机制：通过该机制形成组播源到RP的组播分发树【SPT】

6.2.1、PIM-SM【ASM】模型中，源端DR到RP的组播分发树无法使用Join报文创建，因此需要组播源注册机制帮助形成源端DR到RP的组播分发树【SPT】

6.2.2、形成SPT需要基于Register报文与Join报文，具体过程如下：



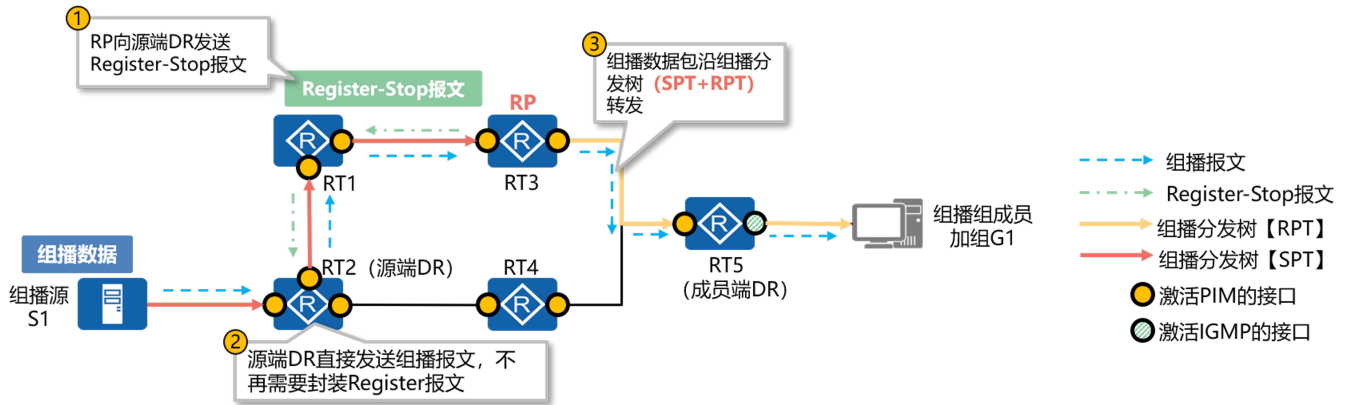
6.2.3、组播源与源端DR之间没有IGMP协议，无法通过IGMP生成PIM(*, G)表项，进而无法发送Join消息形成组播分发树

6.2.4、组播源信息注册到RP后，便形成组播源到RP的SPT，但源端DR此时仍然会将组播数据包封装入Register报文，该方式会造成一些问题：

a、源端DR最初发送的是单播Register报文，但是该方式会加重源端DR与RP的工作量

b、源端DR形成到RP的SPT后，会同时发送单播Register报文和组播报文，造成重复组播包的问题

6.2.5、SPT建立后，RP使用Register-Stop报文通知源端DR后续报文可以以组播报文形式发送



6.3、DR选举机制：DR负责源端或组成员端组播报文的收发，避免重复组播报文，同时成员端DR还负责发送Join加组消息

6.3.1、在源端网络或者成员端网络中，有可能有多台组播路由器转发组播流量，从而造成重复组播报文的问题

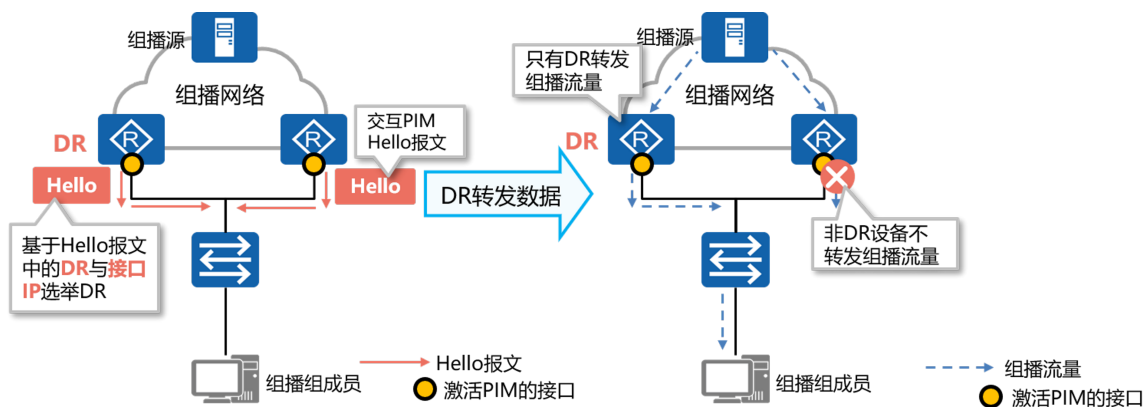
6.3.2、PIM DR【Designated Router】是源端网络或者成员端网络的唯一组播转发者，由于不存在别的组播转发路由器就避免了重复组播报文的问题

6.3.3、PIM DR的选举：

a、在PIM-SM【ASM】中各路由器通过比较Hello消息上携带的优先级和IP地址，为多路访问网络选举指定路由器DR

b、接口DR优先级高的路由器将成为该MA网络的DR，DR优先级默认为1；在优先级相同的情况下，接口IP地址大的路由器将成为DR

c、当DR出现故障后，邻居路由器之间会重新选举DR

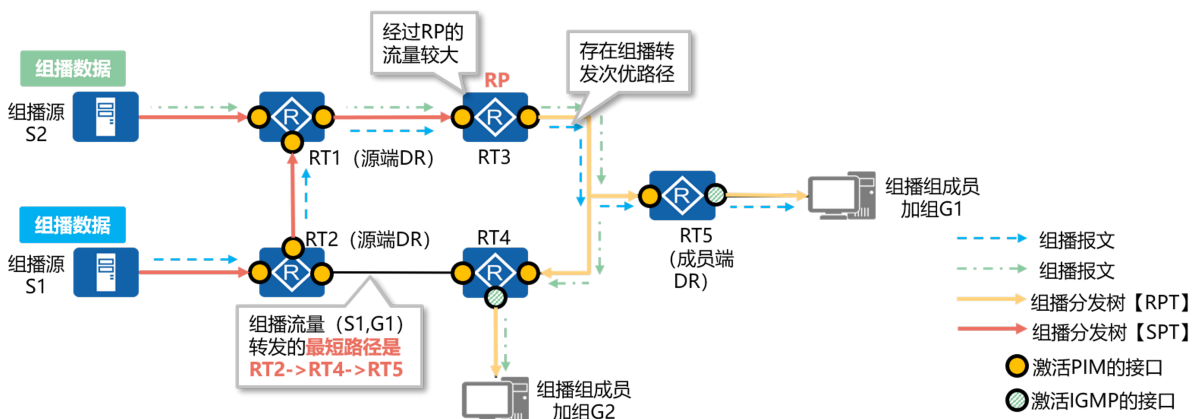


6.4、RPT次优路径问题

在PIM-SM网络中，一个组播组只对应一个RP。因此组播数据最初都会发往RP，由RP进行转发，这会导致两个问题：

6.4.1、过大的组播流量会对RP形成巨大的负担

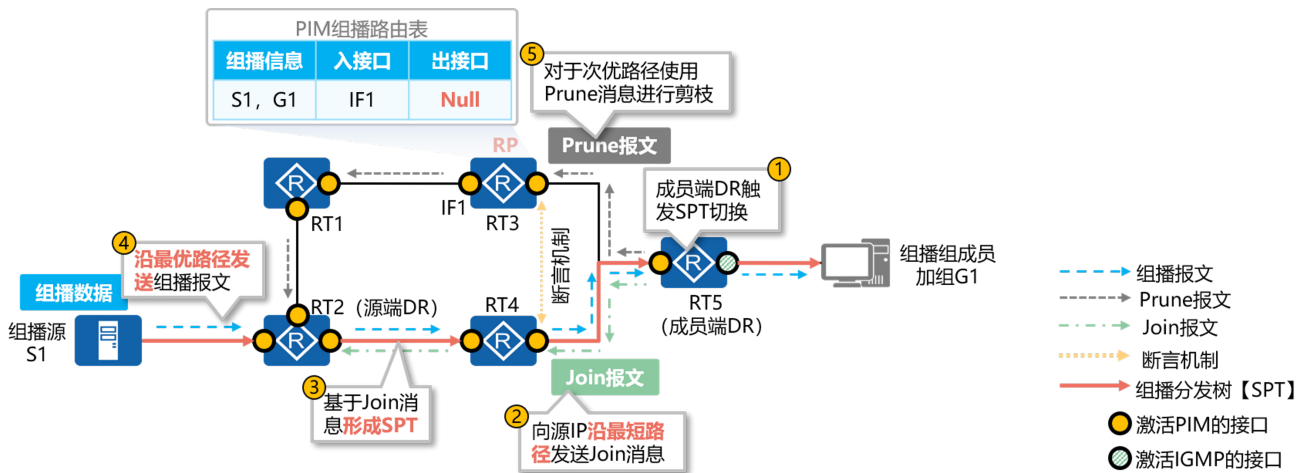
6.4.2、组播转发路径有可能是次优路径



7、SPT切换机制

7.1、当数据发送至RP后，RP会沿RPT将数据发送给成员端DR。为了解决RPT潜在的次优路径问题，成员端DR会基于组播数据包中的源IP，反向建立从成员端DR到源的SPT

7.2、具体过程如下：



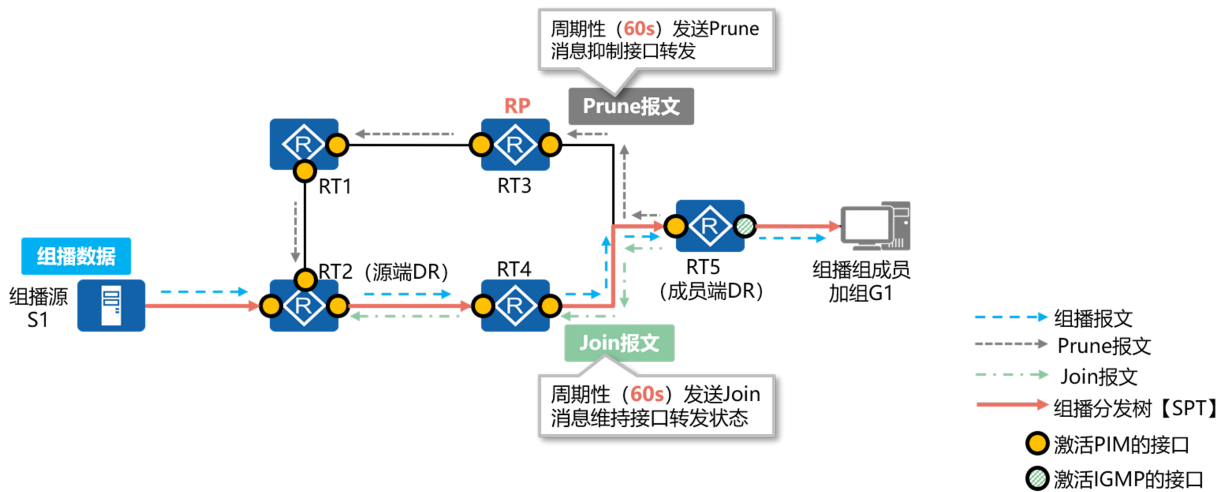
注1：多路访问网络在SPT切换的过程中可能会存在重复报文，需要利用断言机制快速选定下行接口

注2：SPT切换的触发条件：缺省情况下，当RP或者组成员端DR收到第1个组播数据包之后，就会向源发起SPT切换

8、维护组播分发树

当组播分发树【SPT或RPT】稳定后，成员端DR会周期性发送Join/Prune报文，用于维护组播分发树

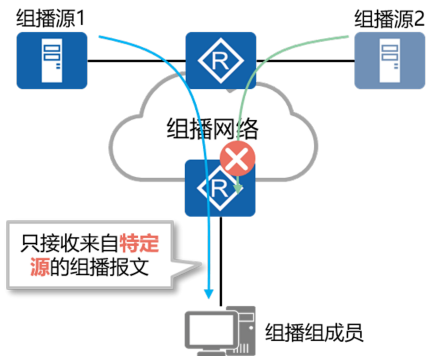
如果组播在一段时间后【默认210s】没有流量则SPT树会消失，成员端DR恢复到RP的RPT树



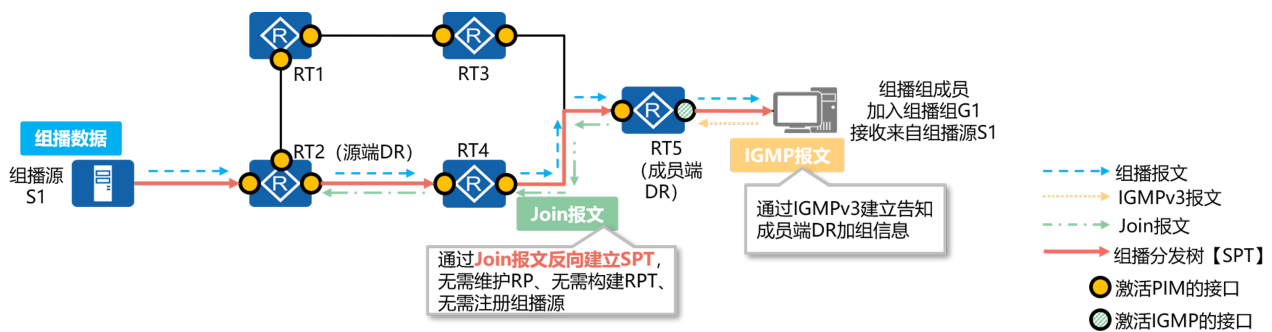
十一、PIM — SM【SSM】

1、SSM模型针对特定源和组的绑定数据流提供服务，接收者主机在加入组播组时，可以指定只接收哪些源的数据或指定拒绝接收来自哪些源的数据。加入组播组以后，主机只会收到指定源发送到该组的数据

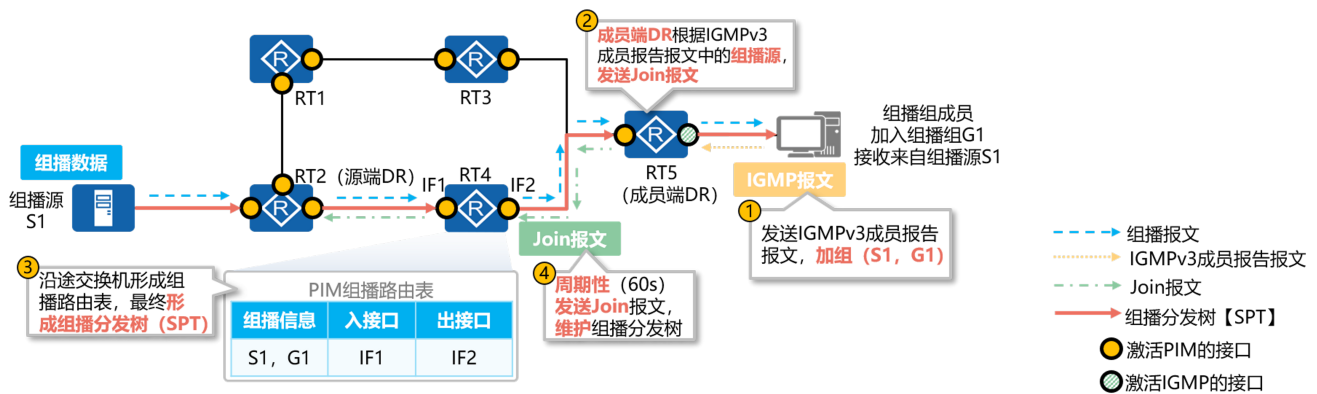
2、SSM模型对组地址不再要求全网唯一，只需要每个组播源保持唯一。这里的“唯一”指的是同一个源上不同的组播应用必须使用不同的SSM地址来区分。不同的源之间可以使用相同的组地址，因为SSM模型中针对每一个（源，组）信息都会生成表项。这样一方面节省了组播组地址，另一方面也不会造成网络拥塞



- 3、由于SSM提前定义了组播的源地址，所以PIM-SM【SSM】可以在成员端DR上基于组播源地址直接反向建立SPT
- 4、PIM-SM【SSM】无需维护RP、无需构建RPT、无需注册组播源，可以直接在组播源与组成员之间建立SPT
- 5、在PIM-SM【SSM】模型中，关键机制包括邻居发现、DR竞选、构建SPT
- 6、PIM-SM【SSM】无需Assert机制



- 7、PIM-SM【SSM】模型构建组播分发树的形成主要依赖IGMPv3报文与Join报文
- 8、PIM-SM【SSM】模型形成的组播分发树会一直存在，不会因为组播流量而消失

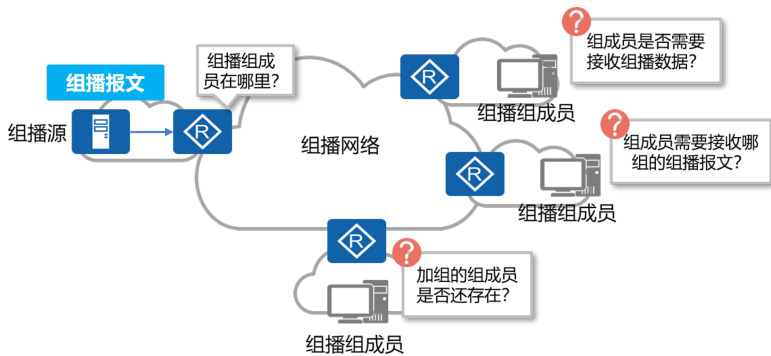


PIM总结 — PIM模型比较:

| 协议 | 模型分类 | 适用场景 | 工作机制 |
|--------|-------|---------------------------------------|--|
| PIM-DM | ASM模型 | 适合规模较小、组播组成员相对比较密集的局域网 | 通过周期性“扩散-剪枝”维护一棵连接组播源和组成员的单向无环SPT |
| PIM-SM | ASM模型 | 适合网络中的组成员相对比较稀疏，分布广泛的大型网络 | 采用接收者主动加入的方式建立组播分发树，需要维护RP、构建RPT、注册组播源 |
| | SSM模型 | 适合网络中的用户预先知道组播源的位置，直接向指定的组播源请求组播数据的场景 | 直接在组播源与组成员之间建立SPT，无需维护RP、构建RPT、注册组播源 |

十二、组播网络的转发困境

- 1、IP组播通信的特点是报文从一个源发出，被转发到一组特定的组播组成员。在组播通信模型中，组播源不关注接收者的位置信息，组播数据转发需要依赖组播网络才能将数据发送至组播组成员
- 2、组播数据在进行传递时，组播网络为了将组播数据转发至组播组成员，需要知道组播组成员的位置与所加组播组



3、组播网络感知组播组成员有两种方法：

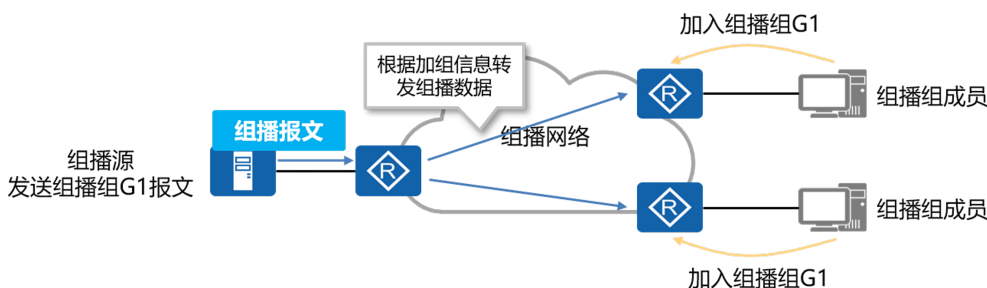
3.1、手工静态配置：在组播路由器上静态指定连接组播组成员的接口，静态配置组成员加组信息

3.1.1、手工静态方式灵活性差，配置工作量大，但相对比较稳定，对于新上线的组成员能够快速建立组播转发通路

3.2、动态感知：通过IGMP协议通知组播网络，组播网络根据IGMP消息感知组播组成员所在接口，以及组成员加组信息

3.2.1、动态感知方式较为灵活，且配置简单，现网一般使用动态感知方式

4、当组播网络获得组成员位置与加组信息后，可以基于这些信息转发组播报文



十三、IGMP【Internet Group Management Protocol | 互联网组管理协议】

1、IGMP是TCP/IP协议族中负责IPv4组播成员管理的协议，用来在接收者主机和与其直接相邻的组播路由器之间建立和维护组播组成员关系

2、IGMP通过在组播组成员和组播路由器之间交互IGMP报文实现组成员管理功能，IGMP报文封装在IP报文中

3、IGMP共有3个版本：

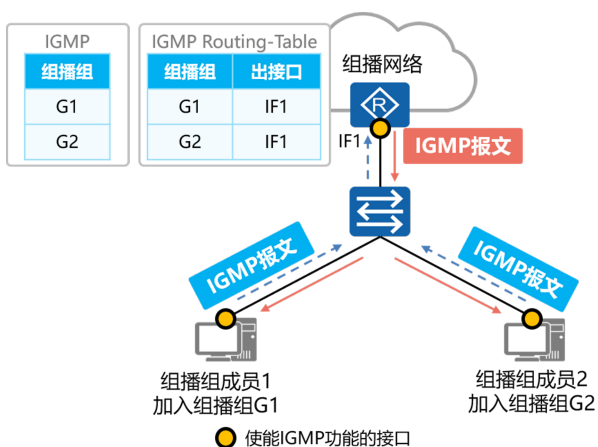
3.1、IGMPv1

3.2、IGMPv2

3.3、IGMPv3

4、组播路由器与组成员间交互报文后会生成IGMP路由表项与IGMP组表项

5、IGMP路由表项与IGMP组表项将帮助设备生成组播路由表项

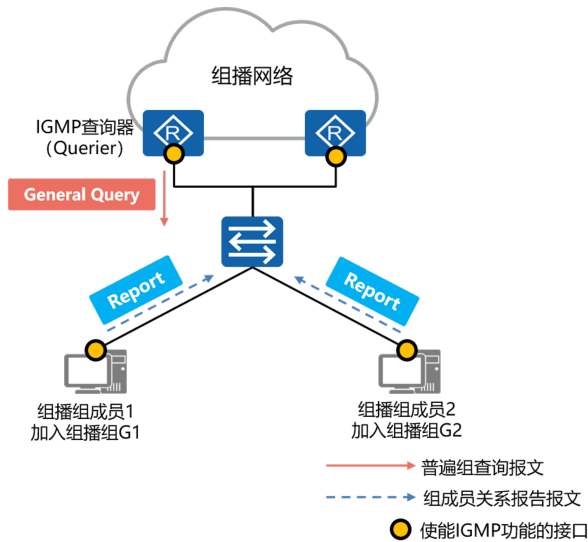


6、再最后一跳组播路由器【组播叶子路由器】上，组播路由表可以基于IGMP路由表项，IGMP组表项与组播协议路由表【PIM路由表】汇总后形成

7、IGMP路由表项与IGMP组表项能为组播协议路由表提供组播组地址信息与出接口信息

十四、IGMPv1

- 1、每个LAN上必须有一台组播路由器定期的传输主机成员资格查询消息，确定哪些主机组包含位于路由器直接相连的LAN上的成员
- 2、该查询消息发送给组播组“所有主机”【224.0.0.1】，TTL为1
- 3、IGMPv1主要基于查询和响应机制完成组播组管理
- 4、查询和响应机制由两种报文实现：
 - 4.1、普遍组查询报文【General Query】：查询器向共享网络上所有主机和路由器发送的查询报文，用于查询哪些组播组存在成员
 - 4.2、成员关系报告报文【Report】：主机向查询器发送的报告报文，用于申请加入某个组播组或者应答查询报文
- 5、由于IGMP报文是组播报文，因此一个多路访问网络里只需要一个组播路由器发送查询报文即可，该组播路由器被称为IGMP查询器【Querier】



- 6、终端使用主机成员资格报告进行应答，为其所属的每个组播组都应答一个这样的报告
- 7、IGMPv1数据包格式如下：

| | | | | |
|-----|----|-----|-----|----|
| | 4 | 7 | 15 | 31 |
| 版本 | 类型 | 未使用 | 校验和 | |
| 组地址 | | | | |

版本 = 1
 类型：
 — 1 = 主机成员资格查询
 — 2 = 主机成员资格报告
 组地址：
 — 组播组地址

8、IGMPv1数据包中的字段：

| 字段 | 具体阐述 |
|-----|--|
| 版本 | 1 |
| 类型 | 对主机而言，IGMP消息有2种： 1：主机成员资格查询 2：主机成员资格报告 |
| 未使用 | 未使用字段，发送时为0，接收时忽略它 |
| 校验和 | 用于验证数据包的完整性 |
| 组地址 | 在主机成员资格查询消息中，发送时将其设置为0，接收时忽略它；在主机成员资格报告消息中，该字段存储了要报告的组播组地址 |

9、IGMPv1组成员加组机制

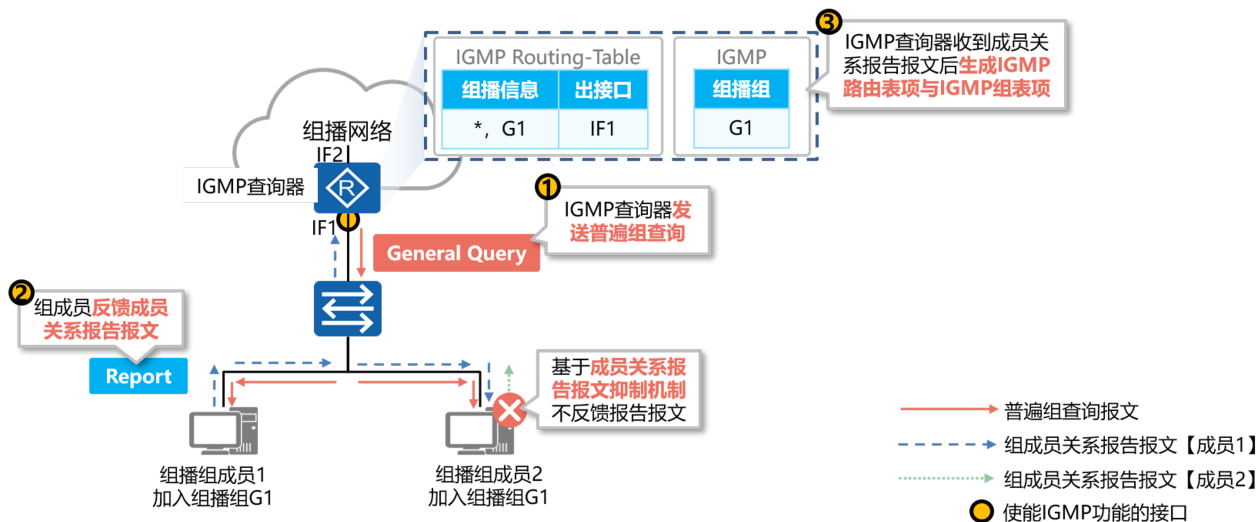
- 9.1、通过普遍组查询报文与成员关系报告报文，IGMP查询器可以了解到该网段内哪些组播组存在成员

9.2、IGMPv1 普遍组查询及响应过程如下：

9.2.1、IGMP 查询器发送目的地址为 224.0.0.1（表示同一网段内所有主机和路由器）的普遍组查询报文；收到该查询报文的组成员启动定时器。普遍组查询报文是周期性发送的，发送周期可以通过命令配置，缺省情况下每隔 60 秒发送一次。组成员 1 和组成员 2 是组播组 G1 的成员，则在本地启动定时器 Timer-G1。缺省情况下，定时器的范围为 0 ~ 10 秒之间的随机值

9.2.2、第一个定时器超时的组成员发送针对该组的报告报文；假设组成员 1 上的 Timer-G1 首先超时，组成员 1 向该网段发送目的地址为 G1 的报告报文。也想加入组 G1 的组成员 2 收到此报告报文，则停止定时器 Timer-G1，不再发送针对 G1 的报告报文。这样报告报文被抑制，可以减少网段上的流量

9.2.3、IGMP 查询器接收到组播组成员 1 的报告报文后，了解到本网段内存在组播组 G1 的成员，则由生成 IGMP 组表项与 (*, G1) IGMP 路由表项，“*”代表任意组播源。网络中一旦有组播组 G1 的数据到达路由器，将向该网段转发



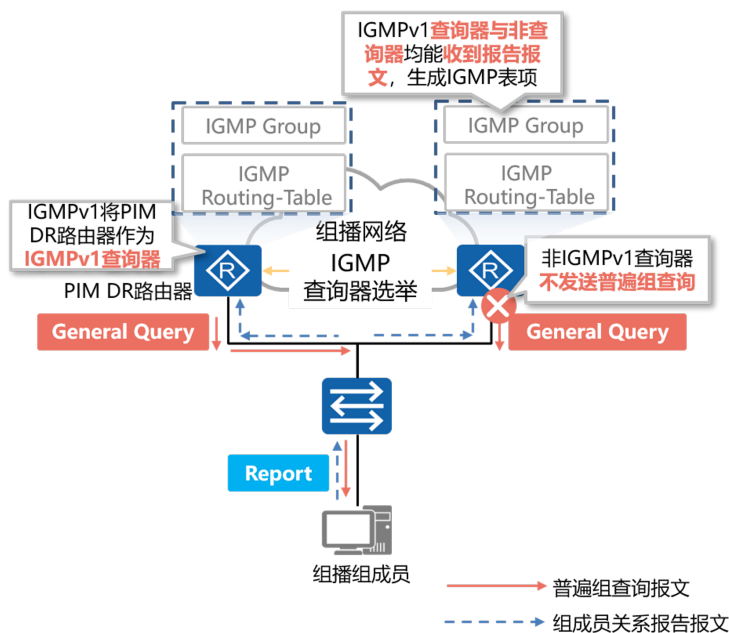
10、IGMPv1 查询器选举机制

10.1、普遍组查询是组播报文，因此同一网段内只需要一台查询器即可查询所有组成员的加组信息

10.2、IGMPv1 没有基于 IGMP 的查询器选举机制，所以需要依赖组播路由协议【PIM】进行 IGMP 查询器选举

10.3、IGMPv1 将组播路由协议【PIM】选举出唯一的组播信息转发者【Assert Winner 或 DR】作为 IGMPv1 的查询器，负责该网段的组成员关系查询

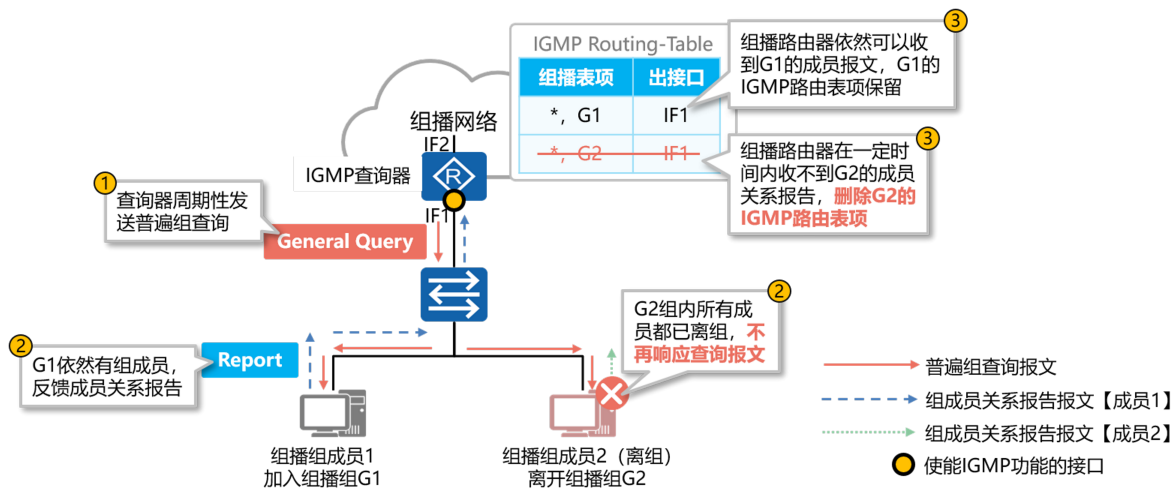
10.4、查询器和非查询器均能收到成员关系报告【目的地址 224.0.0.1】，因此均能形成 IGMP 路由表与 IGMP 组表项



11、IGMPv1 组成员离组机制

11.1、IGMPv1 没有专门定义离开组消息，当组播组成员离开组播组时，将不会再对普遍组查询报文做出回应

11.2、当网段内不存在特定组的组成员，IGMP 查询器不会收到特定组成员的报告报文，则在一定时间【缺省值为 130s】后，删除特定组所对应的组播转发表项



十五、IGMPv2

1、IGMPv1在离组机制与查询器选举机制上有一定缺陷：

- 1.1、IGMPv1离组使用超时机制，组成员只能静默离组。在未超时的时间内，组播流量依然会被组播路由器转发
- 1.2、IGMPv1查询器选举必须要依赖PIM协议，导致查询器选举不够灵活

2、IGMPv2改善了IGMPv1的缺陷：

- 2.1、IGMPv2组成员加组机制与IGMPv1基本相同
- 2.2、IGMPv2增加了离开组机制
- 2.3、IGMPv2增加了查询器选举机制
- 2.4、IGMPv2增加了特定组查询机制
- 2.5、IGMPv2能与IGMPv1兼容

3、IGMPv2报文格式

3.1、为了改善组成员离开机制，IGMPv2新增了两种报文：

- 3.1.1、成员离开报文【Leave】：成员离开组播组时主动向查询器发送的报文，用于宣告自己离开了某个组播组。成员离开报文目的地址为224.0.0.2
- 3.1.2、特定组查询报文【Group-Specific Query】：查询器向共享网段内指定组播组发送的查询报文，用于查询该组播组是否存在成员。特定组查询报文目的地址为所查询组播组的组地址

3.2、IGMPv2对普遍组查询报文格式也做了改进，添加了最大响应时间【Max Response Time】字段。此字段取值可以通过命令配置，用于控制成员对于查询报文的响应速度

3.3、IGMPv2报文格式如下：

| 类型 | 最长响应时间 | 校验和 | |
|----|--------|-----|-----|
| | | | 组地址 |

组播消息类型

最长响应时间：

- 多长时间内必须发送响应报告
- 对于普遍组查询，最大响应时间默认为10秒
- 对于特定组查询，最大响应时间默认为1秒

组地址：

- 组播组地址【常规查询为0.0.0.0】
- 特定组查询报文中，组地址为需要查询的组地址
- 在成员关系报告或离开组的消息中，组地址为需要报告或离开的组地址

3.4、IGMPv2数据包中的字段：

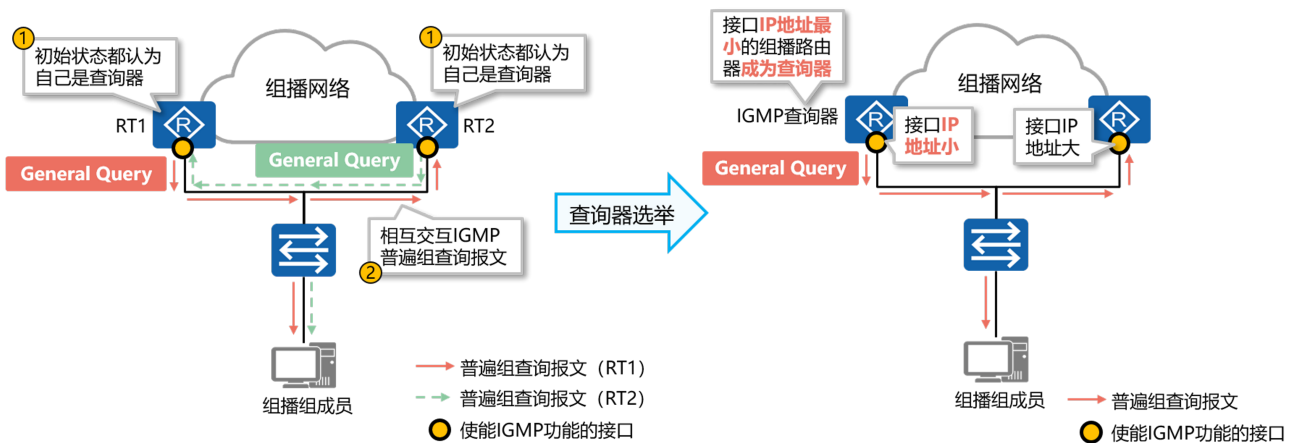
| 字段 | 具体阐述 |
|--------|--|
| 类型 | 0x11: 成员资格查询 0x12: 第1版成员资格报告 0x16: 第2版成员资格报告 0x17: 脱离报告 |
| 最长响应时间 | 针对成员资格查询而言有意义，它指定多长时间之内必须发送响应报告，单位为1/10秒，默认值为10s 在其它消息中，取值为0 |
| 校验和 | 计算得到的校验和，与IGMP校验和相同 |
| 组地址 | 在常规查询中为0.0.0.0 在指定组查询中为查询的组地址 在报告中为组播组地址 |

3.5、IGMPv2查询器选举机制

3.5.1、IGMPv2组成员加组机制与IGMPv1一致

3.5.2、IGMPv2查询器选举机制与IGMPv1有较大差异。IGMPv2使用独立的查询器选举机制，当共享网段上存在多个组播路由器时，IP地址最小的路由器成为查询器

3.5.3、非查询器上都会启动一个定时器【即其它查询器存在时间定时器Other Querier Present Timer】，在该定时器超时前，如果收到了来自查询器的查询报文，则重置该定时器；否则，就认为原查询器失效，并发起新的查询器选举过程



3.6、IGMPv2组成员离开机制

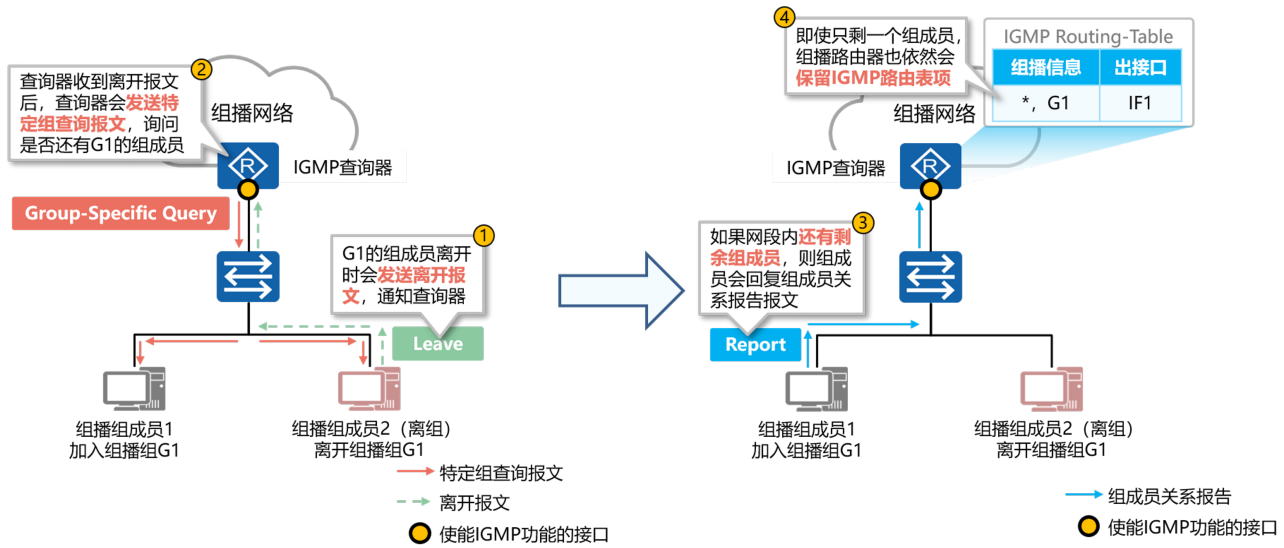
3.6.1、IGMPv2使用成员离开报文与特定组查询报文加速感知IGMPv2组成员离开

3.6.2、组播组成员向本地网段内的所有组播路由器（目的地址为224.0.0.2）发送针对组G1的离开报文

3.6.3、查询器收到离开报文，会发送针对组G1的特定组查询报文。发送间隔和发送次数可以通过命令配置，缺省情况下每隔1秒发送一次，共发送两次。同时查询器启动组成员关系定时器（Timer-Membership=发送间隔x发送次数）

3.6.4、若该网段内还存在组G1的其他成员，这些成员在收到查询器发送的特定组查询报文后，会立即发送针对组G1的报告报文。查询器收到针对组G1的报告报文后将维护该组成员关系

3.6.5、若该网段内不存在组G1的其它成员，查询器将不会收到针对组G1的报告报文。在Timer-Membership超时后，查询器将删除(*, G1)对应的IGMP组表项。当有组G1的组播数据到达查询器时，查询器将不会向下游转发



十六、SSM模型带来的挑战

- 出于安全考虑，组播组成员可以只选择接收从特定组播源发来的组播数据。组成员需要告知组播网络，接收来自哪些特定组播源的组播流量
- IGMPv1与IGMPv2的报文中均无法携带组播源的信息，因此无法配合SSM使用【可使用SSM Mapping功能解决这个问题】
- IGMPv3主要是为了配合SSM【Source-Specific Multicast】模型发展起来的，提供了在报文中携带组播源信息的能力，即主机可以对组播源进行选择
- SSM模型的组播地址范围：232.0.0.0~232.255.255.255

十七、IGMPv3

1、IGMPv3大部分工作机制与IGMPv2类似：

- 1.1、查询器选举机制一致：IP地址小的为查询器
- 1.2、使用普遍组查询报文查询组成员加组信息
- 1.3、使用特定组查询报文查询特定组播的成员存活情况

2、IGMPv3需要支持上报组播源信息，与IGMPv2相比IGMPv3的变化如下：

- 2.1、IGMPv3查询报文除了包含普遍组查询报文和特定组查询报文，还新增了特定源组查询报文【Group-and-Source-Specific Query】
- 2.2、IGMPv3成员关系报告报文不仅包含主机想要加入的组播组，而且包含主机想要接收来自哪些组播源的数据
- 2.3、由于同个组播组的不同成员可能希望接收来自不同源的组播，因此IGMPv3无需成员关系报告报文抑制机制
- 2.4、IGMPv3没有定义专门的成员离开报文，成员离开通过特定类型的报告报文来传达

3、IGMPv3报文格式 — 查询报文

IGMPv3的查询报文共有三类：

- 3.1、普遍组查询报文【General Query】：该报文作用与IGMPv1,IGMPv2中的普遍组查询报文作用一致
- 3.2、特定组查询报文【Group-Specific Query】：该报文作用与IGMPv2中的特定组查询报文作用一致
- 3.3、特定源组查询报文【Group-and-Source-Specific Query】：该报文用于查询该组成员是否愿意接收特定源发送的数据。特定源组查询通过在报文中携带一个或多个组播源地址来达到这一目的
- 3.4、IGMPv3查询报文格式如下：

| 类型 = 0x11 | | | 最大响应编码 | 校验和 |
|-----------|---|-----|--------|--------|
| 组地址 | | | | |
| Resv | S | QRV | QQIC | 信源数【N】 |
| 信源地址【1】 | | | | |
| 信源地址【2】 | | | | |
| . | | | | |
| . | | | | |
| 信源地址【N】 | | | | |

3.5、IGMPv3查询数据包中的字段：

| 字段 | 具体阐述 |
|-------------|---|
| 类型 | 取值为0x11，表示IGMPv3查询 |
| 最长响应编码 | 指定多长时间之内必须发送响应报告，单位为秒 |
| 组地址 | 组播组地址，在常规查询中为0.0.0.0；在特定组查询报文和特定源组查询报文中，该字段为要查询的组播组地址 |
| S | S标记，指出禁止路由器对其进行处理 |
| QRV | 查询方的健壮性值，影响定时器和重传次数，默认为2 |
| QQIC | 查询方的查询间隔编码【单位为秒】，指定查询方使用的查询间隔 |
| 信源数【N】 | 报文中包含的组播源的数量；对于普遍组查询报文和特定组查询报文，该字段为0；对于特定源组查询报文，该字段非0 |
| 信源地址【1...N】 | 信源的地址 |

4、IGMPv3报文格式 — 成员关系报告报文

4.1、IGMPv3成员关系报告报文除了通告组成员的加组信息外，还能通告组成员希望接收的组播源信息。通告组播源主要有两种模式：

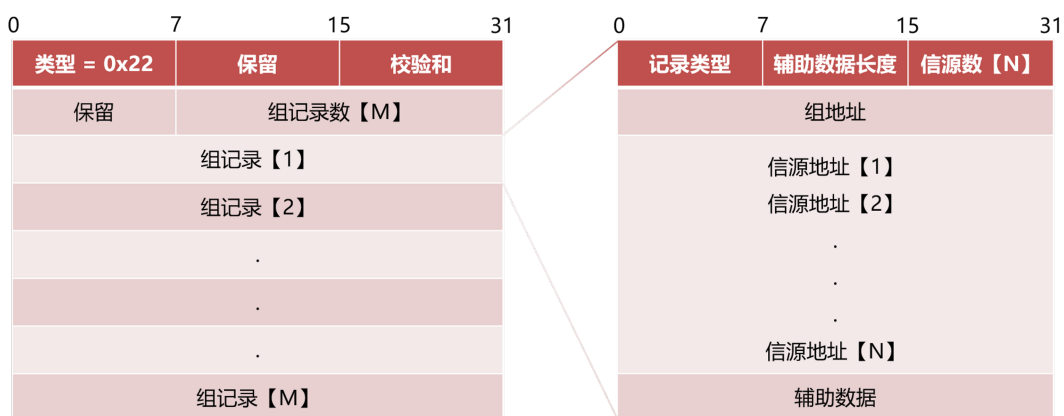
4.1.1、INCLUDE：希望接收来自特定组播源的组播流量

4.1.2、EXCLUDE：希望过滤来自特定组播源的组播流量

4.2、成员关系报告报文中的组播组信息和组播源信息的关系会记录在组记录【Group Record】字段，发送给IGMP查询器

4.3、IGMPv3成员关系报告报文的目的地地址为224.0.0.22

4.4、IGMPv3成员报告报文格式：



4.5、IGMPv3成员报告数据包中的字段：

| 字段 | 具体阐述 |
|-------------|---|
| 类型 | 取值为0x22，表示IGMPv3报告消息 |
| 组记录数【N】 | 报告中包含的组记录数；在IGMPv3中一个成员关系报告报文可以携带多个组播组信息，而之前的版本一个成员关系报告只能携带一个组播组，因此IGMPv3中报文数量大大减少 |
| 组记录【1...M】 | 组记录【一系列包含有关发送方成员资格消息的字段】 |
| 记录类型 | 组记录的类型，共分为3大类： 当前状态报告 ：用于对查询报文进行响应，通告自己目前的状态，共包含2种：【MODE-IS-INCLUDE 与 MODE-IS-EXCLUDE】 过滤模式改变报告 ：当组和源的关系在INCLUDE和EXCLUDE之间切换时，会通告过滤模式发生变化，共包含2种：【CHANGE_TO_INCLUDE_MODE 与 CHANGE_TO_EXCLUDE_MODE】 源列表改变报告 ：当指定源发生改变时，会通告源列表发生变化，共包含2种：【ALLOW_NEW_SOURCES 与 BLOCK_OLD_SOURCES】 |
| 信源数 | 记录中包含的信源数 |
| 信源地址【1...N】 | 信源的地址 |

注：关于记录类型的详细说明：

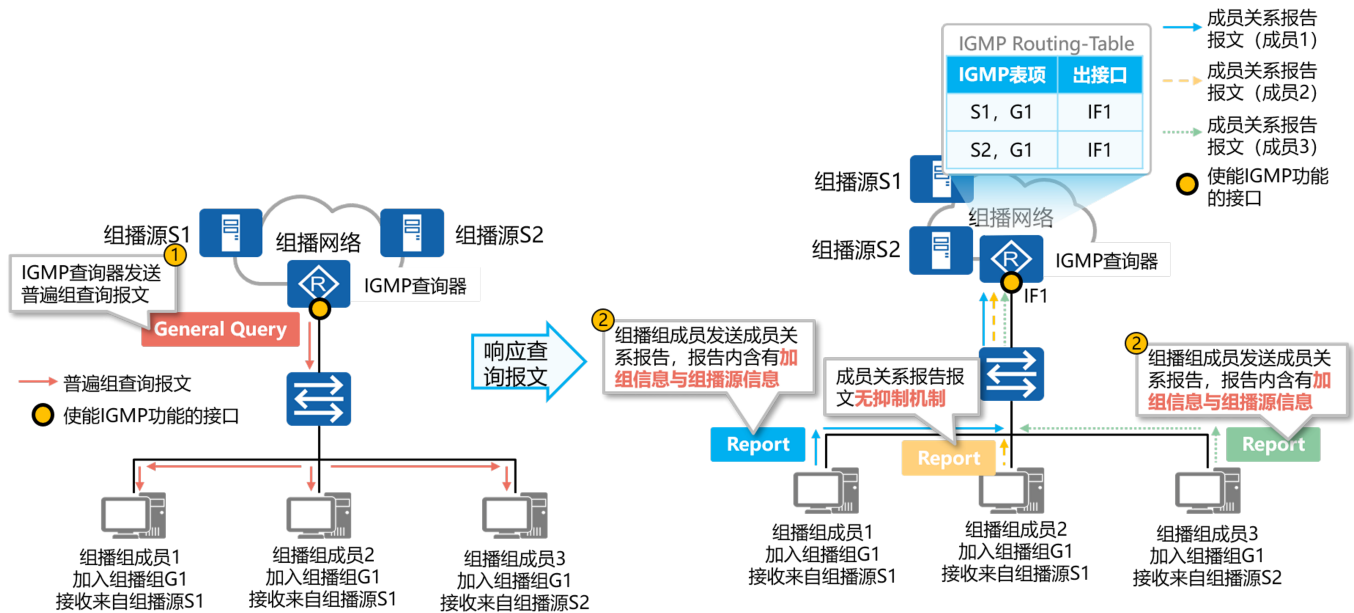
- 1、当前状态报告：用于对查询报文进行响应，通告自己目前的状态，共两种：一种是MODE_IS_INCLUDE，表示接收源地址列表包含的源发往该组的组播数据。如果指定源地址列表为空，该报文无效；另一种是MODE_IS_EXCLUDE，表示不接收源地址列表包含的源发往该组的组播数据
- 2、过滤模式改变报告：当组和源的关系在INCLUDE和EXCLUDE之间切换时，会通告过滤模式发生变化，共两种：一种是CHANGE_TO_INCLUDE_MODE，表示过滤模式由EXCLUDE转换到INCLUDE，接收源地址列表包含的新组播源发往该组播组的数据。如果指定源地址列表为空，主机将离开组播组；另一种是CHANGE_TO_EXCLUDE_MODE，表示过滤模式由INCLUDE转换到EXCLUDE，拒绝源地址列表包含的新组播源发往该组的组播数据
- 3、源列表改变报告：当指定源发生改变时，会通告源列表发生变化，共两种：一种是ALLOW_NEW_SOURCES，表示在现有的基础上，需要接收源地址列表包含的组播源发往该组播组的组播数据。如果当前对应关系为INCLUDE，则向现有源列表中添加这些组播源；如果当前对应关系为EXCLUDE，则从现有阻塞源列表中删除这些组播源；另一种是BLOCK_OLD_SOURCES，表示在现有的基础上，不再接收源地址列表包含的组播源发往该组播组的组播数据。如果当前对应关系为INCLUDE，则从现有源列表中删除这些组播源；如果当前对应关系为EXCLUDE，则向现有源列表中添加这些组播源

4.6、IGMPv3组成员加组机制

4.6.1、IGMPv3组成员加组机制与IGMPv2类似，但有以下不同：

- a、IGMPv3的成员关系报告报文能够携带组播源信息
- b、IGMPv3成员关系报告报文没有成员关系报告报文抑制机制

4.6.2、IGMPv3组成员加组流程如下：

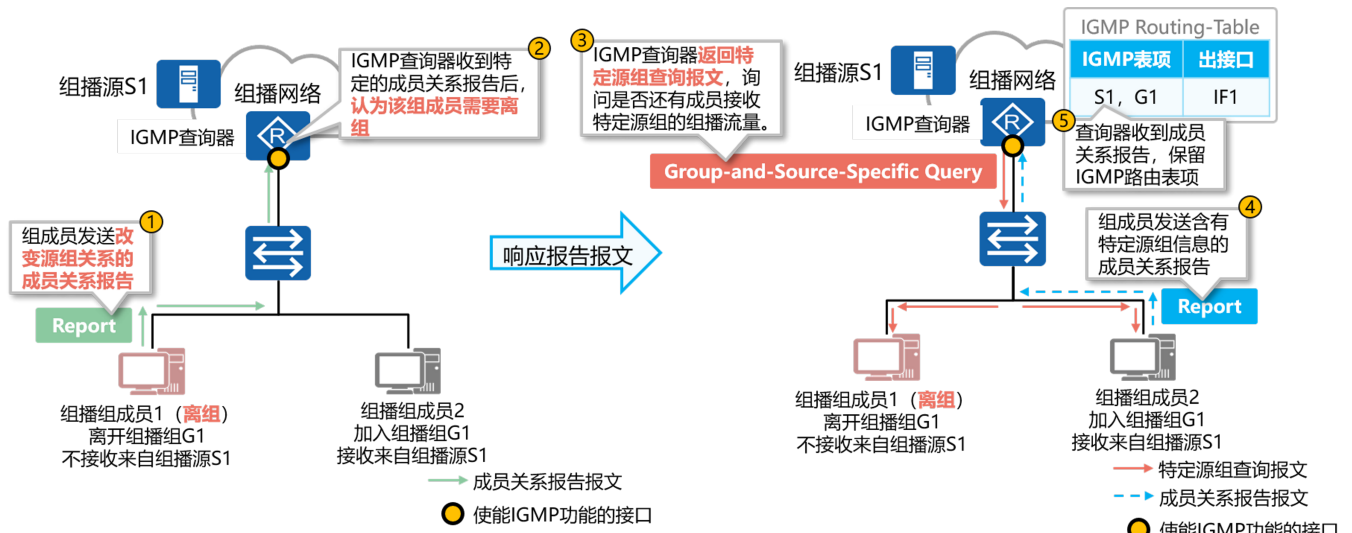


4.7、IGMPv3组成员离组机制

4.7.1、IGMPv3没有专门的成员离开报文，成员离开需要借助组成员关系报告实现

注：改变源组关系的成员关系报告报文很多，如设备原来接收来自S1的组播数据，那么通过发布 (G1, EXCLUDE, S1) 报文或 (G1, CHANGE_TO_EXCLUDE_MODE, S1) 报文都可以改变源组关系

4.7.2、IGMP查询器在收到改变源组对应关系的成员关系报告后，会发送特定源组查询报文，确认是否还有组成员存在



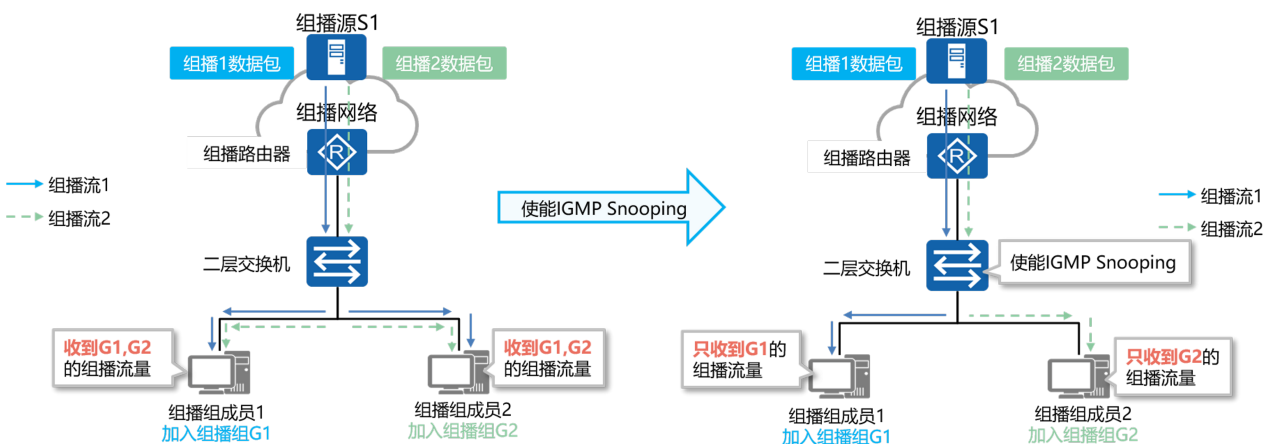
IGMP总结 — IGMP各版本间的差异

| 机制 | IGMPv1 | IGMPv2 | IGMPv3 |
|--------|--------|--------|---------------|
| 查询器选举 | 依靠其它协议 | 自己选举 | 自己选举 |
| 成员离开方式 | 静默离开 | 主动离开 | 主动离开 |
| 特定组查询 | 不支持 | 支持 | 支持 |
| 指定源、组 | 不支持 | 不支持 | 支持 |
| 版本兼容性 | | IGMPv1 | IGMPv1、IGMPv2 |

十八、以太网的组播转发问题

1、当Router将组播报文转发至Switch以后，Switch负责将组播报文转发给组播用户。由于组播报文的地址为组播地址，在二层设备上学习不到这一类MAC表项的，因此组播报文就会在所有接口进行广播，和它在同一广播域内的组播成员和非组播成员都能收到组播报文。这样不但浪费了网络带宽，而且影响了网络信息安全

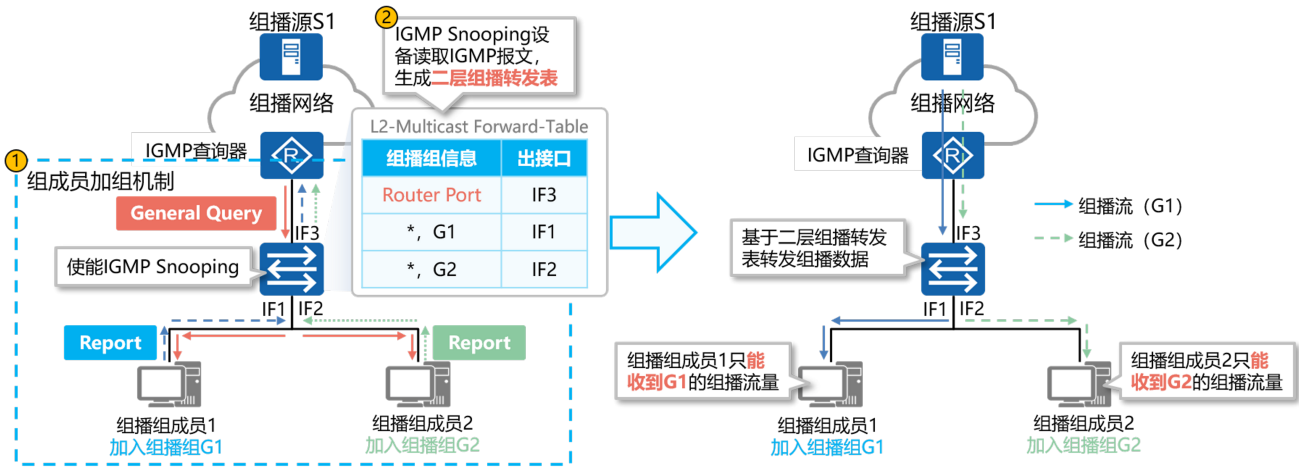
2、IGMP Snooping有效地解决了这个问题。配置IGMP Snooping后，二层组播设备可以侦听和分析组播用户和上游路由器之间的IGMP报文，根据这些信息建立二层组播转发表项，控制组播数据报文转发。这样就防止了组播数据在二层网络中的广播



十九、IGMP Snooping介绍

1、IGMP Snooping可以实现组播数据在数据链路层的转发和控制

2、当主机和上游三层设备之间传递的IGMP协议报文通过二层组播设备时，IGMP Snooping分析报文携带的信息，根据这些信息建立和维护二层组播转发表，从而指导组播数据在数据链路层按需转发



3、二层组播转发表项中存在两类接口：

3.1、路由器端口（Router Port）：二层组播设备上朝向三层组播设备（DR或IGMP查询器）一侧的接口，二层组播设备从此接口接收组播数据报文

注：路由器端口生成过程：

- a、由协议生成的路由器端口叫做动态路由器端口。收到源地址不为0.0.0.0的IGMP普遍组查询报文或PIM Hello报文（三层组播设备的PIM接口向外发送的用于发现并维持邻居关系的报文）的接口都将被视为动态路由器端口
- b、手工配置的路由器端口叫做静态路由器端口

3.2、成员端口（Member Port）：又称组播组成员端口，表示二层组播设备上朝向组播组成员一侧的端口，二层组播设备往此接口发送组播数据报文

注：成员端口生成过程：

- a、由协议生成的成员端口叫做动态成员端口。收到IGMP Report报文的接口，二层组播设备会将其标识为动态成员端口
- b、手工配置的成员端口叫做静态成员端口

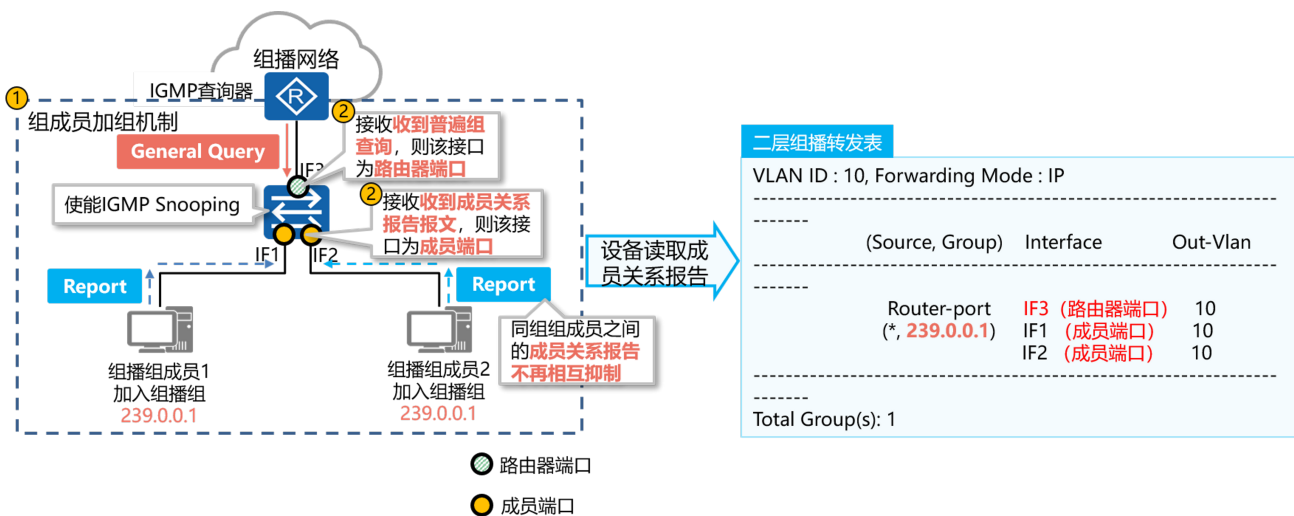
4、IGMP Snooping设备通过监听IGMP报文，形成二层组播转发表，并决定接口类型

4.1、路由器端口形成后会启动老化计时器（默认180s），当路由器端口收到新的普遍组查询后刷新该计时器

4.2、成员端口形成后会启动老化计时器（默认180s），当成员端口收到新的成员关系报告报文后刷新该计时器

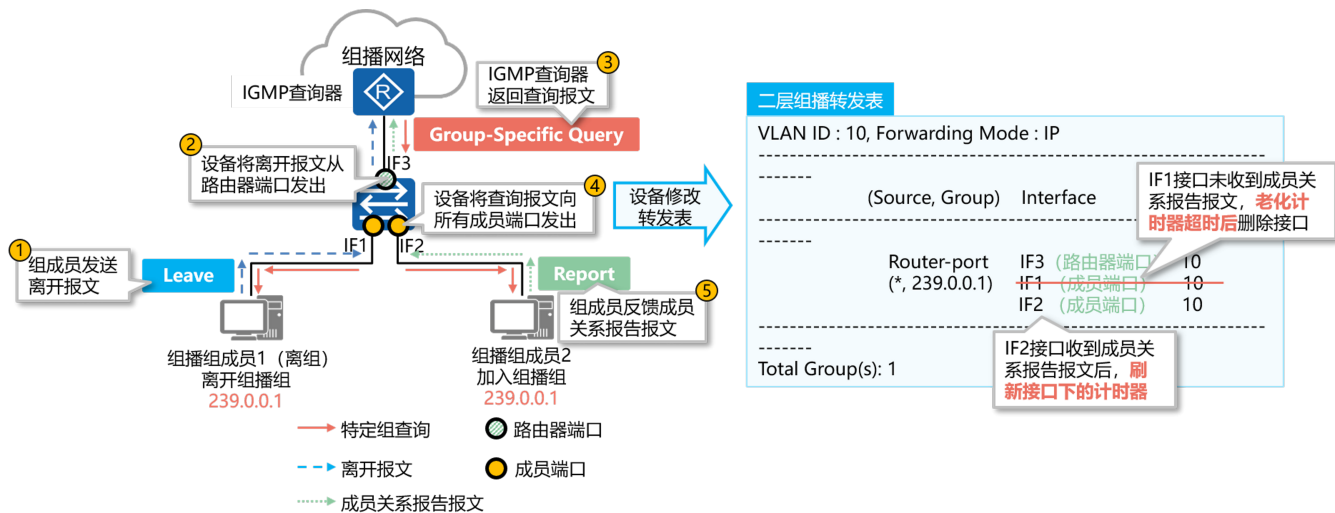
4.3、IGMP Snooping不再使用成员关系报告报文抑制机制：

- 4.3.1、由于IGMP Snooping需要监听IGMP报文才能决定端口角色，进而指导转发，所以所有组成员都需要发送IGMP组成员关系报告报文
- 4.3.2、当IGMP Snooping设备收到成员关系报告报文后，只将成员关系报告报文从路由器接口发送出去，从而避免其余组成员收到成员关系报告报文，不触发成员关系报告报文抑制机制



5、IGMP Snooping设备通过监听IGMP离开报文，IGMP成员关系报告报文决定特定端口是否还需要发送特定组播

注：收到IGMP离开报文后，成员端口的老化定时器 = 健壮系数（默认2）x 特定组查询间隔（默认1s）



二十、IGMP SSM Mapping介绍

1、现网中存在部分只能运行IGMPv1与IGMPv2的老旧终端，在部署SSM模式的组播时，由于IGMPv1与IGMPv2报文中无法携带组播源信息，因此无法使用SSM模式的组播网络

2、IGMP SSM Mapping通过静态的将组播源与组播组进行绑定，使得IGMPv1与IGMPv2的组成员也能接入SSM组播网络

3、无论使用IGMPv1，IGMPv2或IGMPv3加入SSM组播组，SSM组播组的地址依然是232.0.0.0~232.255.255.255

4、IGMP SSM Mapping工作原理：

配置了SSM Mapping规则后，当IGMP查询器收到来自成员主机的IGMPv1或IGMPv2报告报文时，首先检查该报文中所携带的组播组地址G，然后根据检查结果的不同分别进行处理

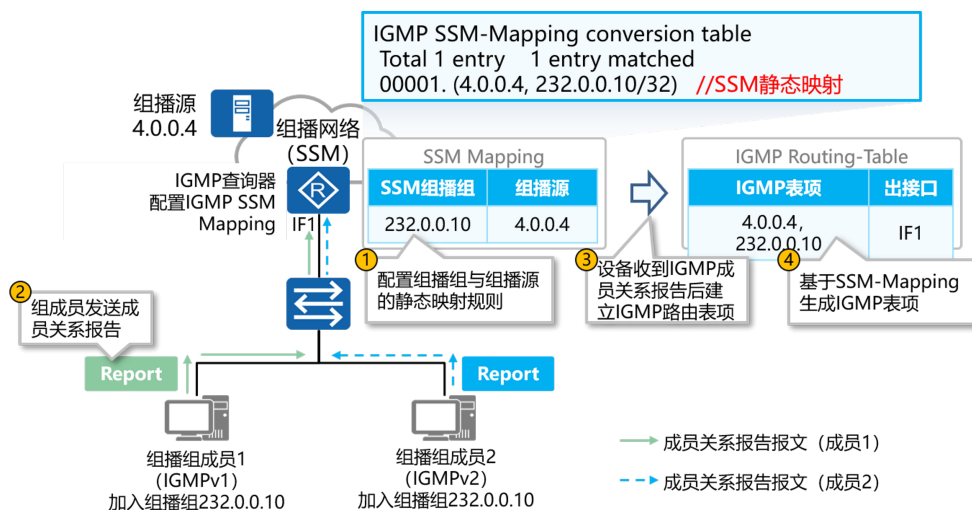
4.1、若G在ASM (Any-Source Multicast) 范围内，则只提供ASM服务

4.2、若G在SSM组地址范围内（缺省情况下为232.0.0.0~232.255.255.255）：

4.2.1、若路由器上没有G对应的SSM Mapping规则，则无法提供SSM服务，丢弃该报文

4.2.2、若路由器上有G对应的SSM Mapping规则，则依据规则将报告报文中所包含的(*, G)信息映射为(G, INCLUDE, (S1, S2...))信息，提供SSM服务

5、IGMP SSM Mapping不处理IGMPv3的报告报文。为了保证同一网段运行任意版本IGMP的主机都能得到SSM服务，需要在与成员主机所在网段相连的组播路由器接口上运行IGMPv3

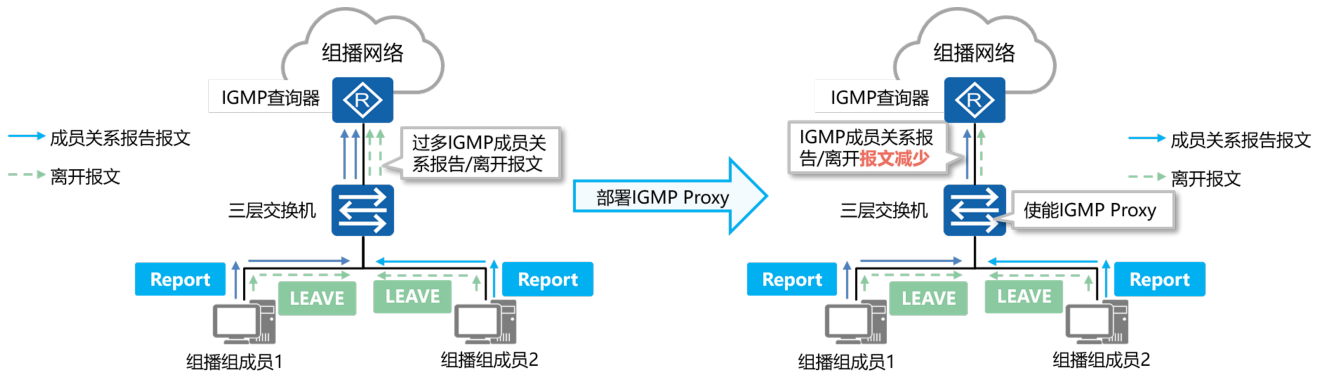


二十一、IGMP Proxy

1、现网中可能存在一台IGMP查询器需要管理大量组成员的情况，大量成员主机频繁加入/离开组播组时，会产生大量的IGMP成员关系报告/离开报文，从而给IGMP查询器带来较大的处理压力

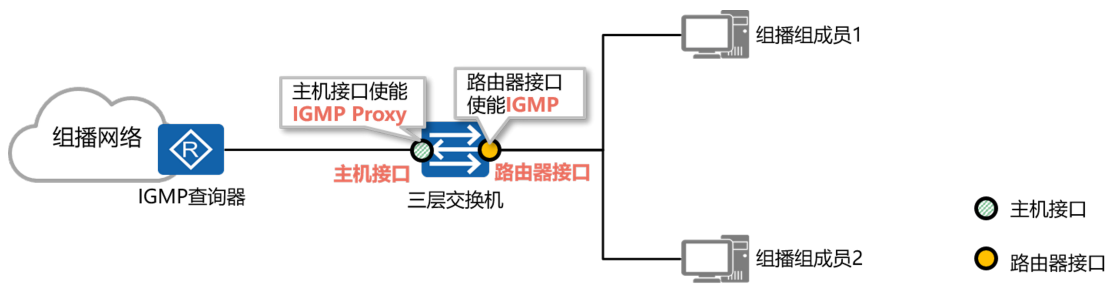
2、通过IGMP Proxy功能可减少IGMP查询器接收IGMP成员关系报告/离开报文的数量，减轻IGMP查询器压力

3、IGMP Proxy通常被部署在IGMP查询器和成员主机之间的三层设备上



4、IGMP Proxy基本概念

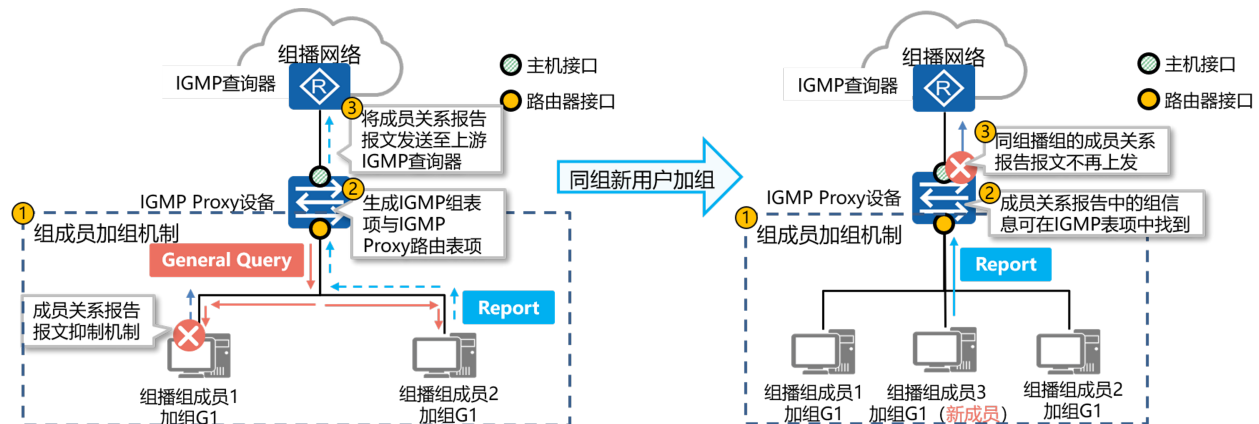
- 4.1、为了缓解IGMP查询器压力，IGMP Proxy设备将成员关系报告/离开报文汇聚后统一上送给IGMP查询器
- 4.2、IGMP Proxy设备也可以代理IGMP查询器向成员主机发送查询报文，维护组成员关系，基于组成员关系进行组播转发
- 4.3、为了实现以上功能，IGMP Proxy定义了两类接口：
 - 4.3.1、主机接口【Host Interface】：IGMP Proxy设备上配置IGMP Proxy功能的接口，该接口一般面向IGMP查询器
 - 4.3.2、路由器接口【Router Interface】：IGMP Proxy设备上配置IGMP功能的接口，该接口一般面向组成员



5、IGMP Proxy工作机制 — 成员加组

IGMP Proxy减少成员关系报告报文的工作机制如下：

- 5.1、路由器接口作为IGMP接口，对下呈现为IGMP查询器，发送查询报文，处理成员关系报告报文，形成IGMP表项，并将成员关系报告从主机接口发送给上游的IGMP查询器
- 5.2、当新用户加入同一个组播组时，IGMP Proxy设备不会向IGMP查询器反馈成员关系报告报文，因此减少了成员关系报告报文数量



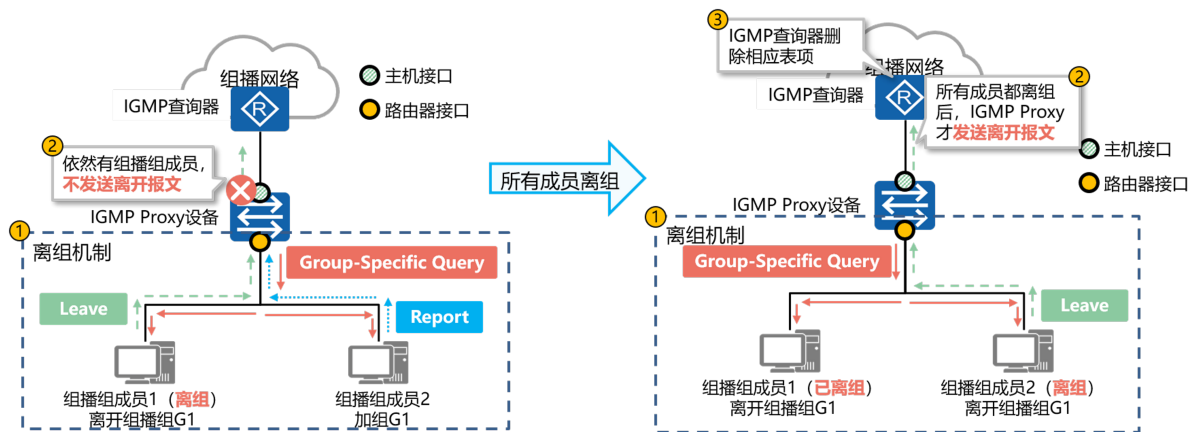
5.3、IGMP代理设备收到某组播组的报告报文后，会在IGMP组表项中查找该组播组：

- 5.3.1、若没有找到相应的组播组，IGMP代理设备会向接入设备发送针对该组播组的报告报文，并在组播转发表中添加该组播组；
- 5.3.2、若找到相应的组播组，IGMP代理设备就不需要向接入设备发送报告报文

6、IGMP Proxy工作机制 — 成员离组

IGMP Proxy减少离组报文的工作机制如下：

- 6.1、当组成员离开时，IGMP Proxy通过IGMP离组机制确定是否有特定组播组的组成员，当确定已经没有组成员后才发送离开报文给上游IGMP查询器



6.2、IGMP代理设备收到某组播组G1的离开报文后，会向接收到该离开报文的接口发送一个特定组查询报文，检查该接口下是否还存在组播组G1的其它成员：

6.2.1、若没有其它成员，在组播转发表中将该接口删除，然后判断组播组G1是否还有其它接口

a、若没有，IGMP代理设备再会向接入设备发送针对该组播组的离开报文

b、若有，IGMP代理设备不向接入设备发送针对该组播组的离开报文

6.2.2、若有其它成员，IGMP代理设备会继续向该接口转发组播数据

二十二、组播的配置

详细配置见实验手册