

IP 组播基础

组播知识点：

组播地址分类，IGMP v1 v2 v3 的区别，查询器选举，IGMP Snooping 原理，PIM DM 基本配置，构建 SPT，5 种报文，hello 间隔，嫁接和断言的作用，DR 选举，PIM SM 基本配置，构建 SPT 和 RPT，接收侧 DR 和组播源侧 DR，RP 选举，(*,G) (S,G)，Switchover 机制

组播优点：

降低网络流量、减轻服务器和 CPU 负荷
减少冗余流量、节约网络带宽、降低网络负载
使多点应用成为可能

组播劣势：

基于 UDP
尽力而为
没有拥塞避免机制
报文重复
报文失序

224.0.0.0 ~ 224.0.0.255 保留给路由协议和其他网络维护功能

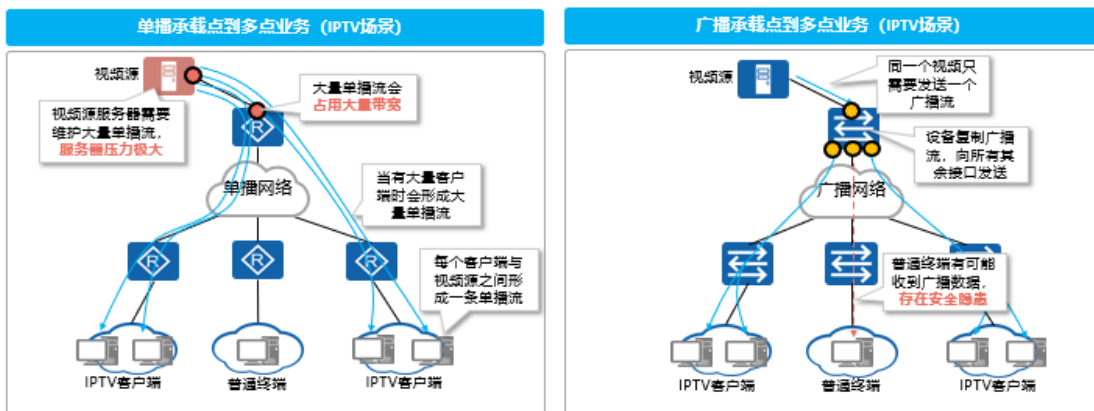
224.0.0.1	同一网络中的所有系统
224.0.0.2	同一网络中的所有路由器
224.0.0.5	所有 OSPF 路由器
224.0.0.6	OSPF 指定路由器
224.0.0.9	RIPv2 路由器

224.0.0.10 EIGRP 路由器
224.0.0.13 PIM 路由器
224.0.0.18 VRRP

- 网络中存在各种各样的业务，从流量模型看一般可以将业务分为两类：
- 点到点业务：比如 FTP，WEB 业务，此类业务主要特点是不同的用户有不同的需求，比如用户 A 需要下载资料 A，用户 B 需要下载资料 B。此类业务一般由单播承载，服务器对于不同用户发送不同的点到点数据流。
- 点到多点业务：比如 IPTV，视频会议等，此类业务的特点是用户对于业务有相同的需求，比如用户 A，B，C，D 都需要收看视频 X，此类业务可以使用单播，组播，广播承载。但使用单播或广播承载点到多点业务时存在一定问题。
- 组播技术能够较好的解决单播或广播在承载点到多点业务时存在的问题。
- 本课程主要介绍使用组播承载点到多点业务的优点，组播网络的基本概念与组播转发的基本流程。

点到多点业务的困境

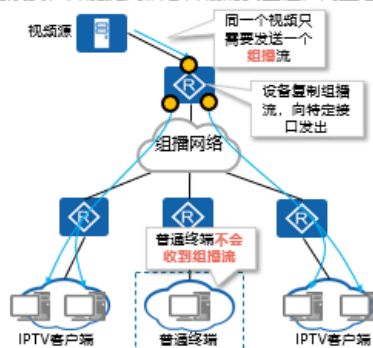
点到多点业务可以由单播，组播，广播进行承载，现网中也有各种各样的实现方式。但使用单播或者广播承载点到多点业务时存在一些固有的问题。



- 单播 (Unicast) 是在一台源 IP 主机和一台目的 IP 主机之间进行。网络上绝大部分的数据都是以单播的形式传输的，例如电子邮件收发、网上银行都是采用单播实现的。
- 在单播通信中每一个数据包都有确切的目的 IP 地址；对于同一份数据，如果存在多个接收者，Server 需发送与接收者数目相同的单播数据包；当接收者增加到成百上千时，将极大加重 Server 创建相同数据和发送多份相同拷贝后所产生的消耗，网络中的设备性能及链路带宽都会面临一定程度的浪费；单播方式较适合用户稀少的网络，当用户量较大时很难保证网络传输质量。
- 广播 (Broadcast) 是在一台源 IP 主机和网络中所有其它的 IP 主机之间进行，属于一对所有的通讯方式，所有主机都可以接收到 (不管是否需要) 。
- 广播数据包被限制在广播域中；一旦有设备发送广播数据，则广播域内所有设备都会收到这个数据包，并且不得不耗费资源去处理，大量的广播数据包将消耗网络的带宽及设备资源；广播方式只适合共享网段，且信息安全性和有偿服务得不到保障。

使用组播承载点到多点业务

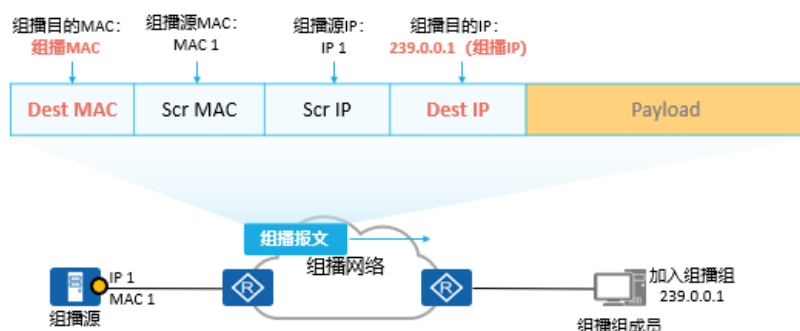
- 组播方式下，单一的信息流沿组播分发树被同时发送给一组用户，相同的组播数据流在每一条链路上最多仅有一份。相比单播和广播，使用组播的好处如下：
 - 相比单播，用户的增加不会导致信息源负载的加重，不会导致网络资源消耗的显著增加。
 - 相比广播，不会造成网络资源的浪费，并能提高信息传输的安全性，而且组播可以实现跨网段的传输。



- 组播（Multicast）是在一台源IP主机和多台（一组）IP主机之间进行，中间的网络设备根据接收者的需要，有选择性地对数据进行复制和转发。
- 组播技术有效地满足了单点发送、多点接收的需求，实现了IP网络中点到多点业务数据的高效传送，能够大量节约网络带宽、降低网络负载。
- 组播分发树：组播流量的转发路径。

组播数据报文结构

- 组播数据报文的结构与单播报文类似，但组播数据报文的目的MAC地址与目的IP地址与单播报文有很大差异。
 - 组播目的IP地址：目的IP地址为组播IP地址，地址范围从224.0.0.0到239.255.255.255
 - 组播目的MAC地址：目的MAC地址为组播MAC地址，组播MAC地址由组播IP地址映射而来



组播IP地址

- 在IPv4地址空间中，D类地址（224.0.0.0/4）被用于组播。一个组播地址就表示一个点到多点的数据流，比如IPTV数据流，语音会议数据流。
- 大多数情况下，同一个组播网络里不同的业务（比如，IPTV，语音会议）就需要使用不同的组播IP地址。
- IANA对D类地址做了进一步的定义，几种主要的组播地址如下表所示：

范围	含义
224.0.0.0—224.0.0.255	为路由协议预留的永久组地址
224.0.1.0—231.255.255.255	Any-Source临时组播组地址
233.0.0.0—238.255.255.255	Source-Specific临时组播组地址
232.0.0.0—232.255.255.255	本地管理的Any-Source临时组播组地址
239.0.0.0—239.255.255.255	

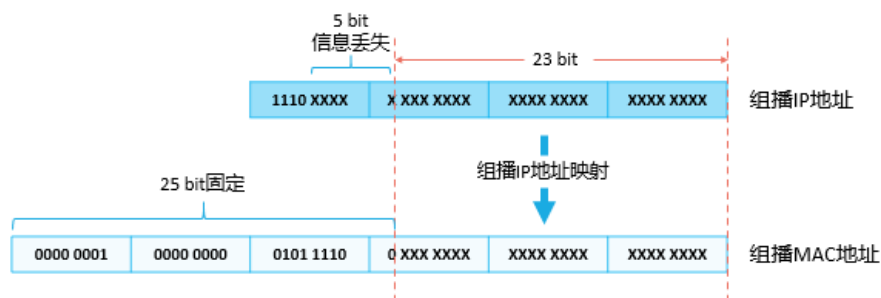
- IPv4 组播地址：
- IPv4 地址空间分为五类，即A类、B类、C类、D类和

E 类。D 类地址为 IPv4 组播地址，范围是从 224.0.0.0 到 239.255.255.255，用于标识组播组，且仅能作为组播报文的目的地址使用，不能作为源地址使用。

- IPv4 组播报文的源地址字段为 IPv4 单播地址，可使用 A、B 或 C 类地址，不能是 D 类、E 类地址。
- 在网络层上，加入同一组播组的所有用户主机能够识别同一个 IPv4 组播组地址。一旦网络中某用户加入该组播组，则此用户就能接收以该组地址为目的地址的 IP 组播报文。

组播MAC地址

- 以太网传输IPv4单播报文的时候，目的MAC地址使用的是接收者的MAC地址。但是在传输组播数据时，其目的地不再是一个具体的接收者，而是一个成员不确定的组，所以要使用IPv4组播MAC地址。
- IANA规定，IPv4组播MAC地址的高24位为0x01005e，第25位为0，低23位为IPv4组播地址的低23位，例如组播组地址224.0.1.1对应的组播MAC地址为01-00-5e-00-01-01。



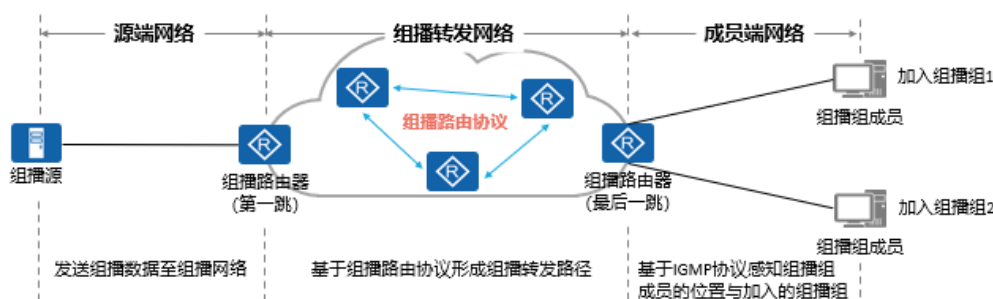
- IPv4 组播地址的前 4 位是固定的 1110，对应组播 MAC 地址的高 25 位，后 28 位中只有 23 位被映射到 MAC 地址，因此丢失了 5 位的地址信息，直接结果是有 32 个 IPv4 组播地址映射到同一 MAC 地址上。例如 IP 地址为 224.0.1.1、224.128.1.1、225.0.1.1、239.128.1.1 等组播组的组播 MAC 地址都为 01-00-5e-00-01-01。网络管理员在分配地址时必须考虑这种情况。
- IETF 认为同一个局域网中两个或多个组地址生成相同的 MAC 地址的几率非常低，不会造成太大的影响。
- 组播 MAC 地址标识了一组设备，这种 MAC 地址第 1 个字节的最低比特位为 1，例如 0100-5e-00ab。
- 一个组播 MAC 地址所标识的一组设备有着共同的特点，

那就是它们都加入了相同的组播组，这些设备将会侦听目的 MAC 地址为该组播 MAC 地址的数据帧。只有单播 MAC 地址才能够被分配给一个以太网接口，组播或广播 MAC 地址是不能被分配给任何一个以太网接口的，换句话说，这两种类型的 MAC 地址不能作为数据帧的源 MAC 地址，而只能作为目的 MAC 地址。

- 对于组播 MAC 地址，相信大家并不会太陌生，例如 STP 协议的 BPDU 载荷便是被直接封装在以太网数据帧中的，并且数据帧的目的 MAC 地址为 0180-c200-0000，这就是一个组播 MAC 地址，类似这样的例子还有很多，此处不再一一列举，这些组播 MAC 地址并不与组播 IP 地址存在关联。
- 除此之外，还有一类组播 MAC 地址是我们需要格外关注的，那就是与组播 IP 地址存在映射关系的组播 MAC 地址。本课程介绍的组播 MAC 地址对应该类型。

组播网络基本架构

- 组播网络大体可以分为三个部分：
 - 源端网络：将组播源产生的组播数据发送至组播网络。
 - 组播转发网络：形成无环的组播转发路径，该转发路径也被称为组播分发树（Multicast Distribution Tree）。
 - 成员端网络：让组播网络感知组播组成员位置与加入的组播组。



- 组播源（Source）：组播流量的发送者，例如多媒体服务器。组播源无需运行任何组播协议，只需简单地将组播数据发送出来即可。
- 组播接收者（Receiver）：也被称为组播组成员，是期望接收特定组播组流量的设备，例如运行多媒体直播客户端

软件的 PC。

- 组播组（Multicast Group）：用 IP 组播地址进行标识的一个集合。任何用户主机（或其他接收设备），加入一个组播组，就成为了该组成员，可以识别并接收发往该组播组的组播数据。
- 组播路由器（Multicast Router）：支持组播、运行组播协议的网络设备，实际上不仅仅路由器能够支持组播，交换机、防火墙等设备也能够支持组播（取决于设备型号），路由器仅是一个代表。
- 第一跳路由器（First-Hop Router）：组播转发路径上，与组播源相连且负责转发该组播源发出的组播数据的 PIM 路由器。
- 最后一跳路由器（Last-Hop Router）：组播转发路径上，与组播组成员相连且负责向该组成员转发组播数据的 PIM 路由器。
- IGMP（Internet Group Management Protocol，因特网组管理协议），是 TCP/IP 协议族中负责 IP 组播成员管理的协议，它用来在接收者和与其直接相邻的组播路由器之间建立、维护组播组成员关系。

组播服务模型

- 组播组成员在接收组播数据时可以对于组播数据源进行选择，因此产生了ASM（Any-Source Multicast，任意源组播）和SSM（Source-Specific Multicast，指定源组播）两种组播服务模型。
 - ASM：组成员加入组播组以后，组成员可以接收到任意源发送到该组的数据。
 - SSM：组成员加入组播组以后，组成员只会收到指定源发送到该组的数据。

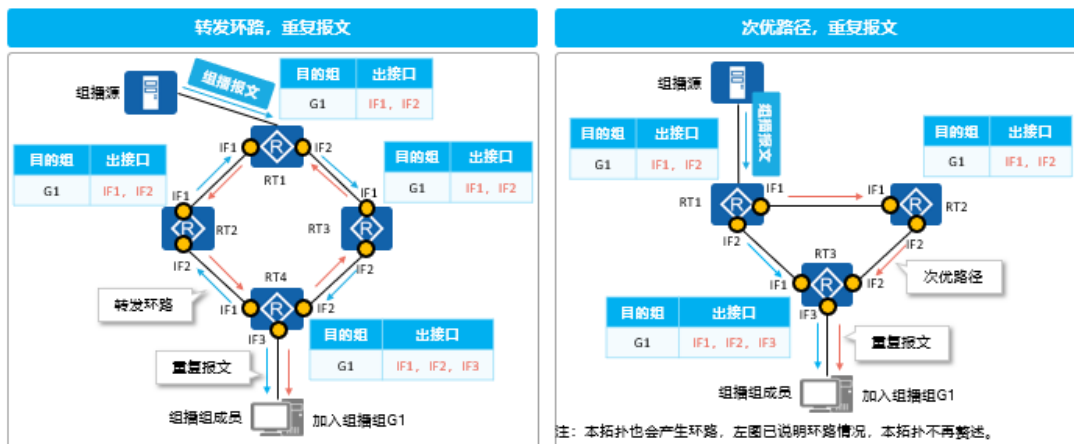


- ASM 特点：
- ASM 为了提高安全性，可以在路由器上配置针对组播源的过滤策略，允许或禁止来自某些组播源的报文通过。最终从接收者角度看，数据是经过筛选的。
- ASM 模型要求组地址必须整个组播网络中唯一。“唯一”指的是同一时刻一个 ASM 地址只能被一种组播应用使用。如果有两种不同的应用程序使用了同一个 ASM 组地址发送数据，它们的接收者会同时收到来自两个源的数据。这样一方面会导致网络流量拥塞，另一方面也会给接收者主机造成困扰。
- SSM 特点：
- SSM 模型对组地址不再要求全网唯一，只需要每个组播源保持唯一。这里的“唯一”指的是同一个源上不同的组播应用必须使用不同的 SSM 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。



组播数据转发的困局

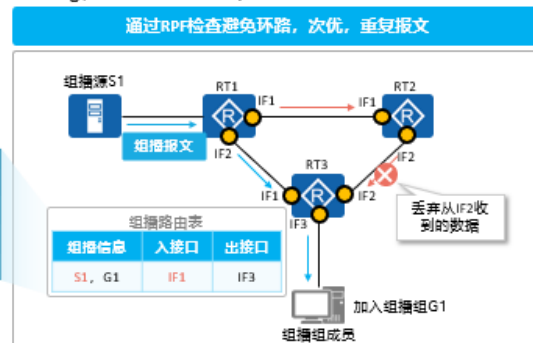
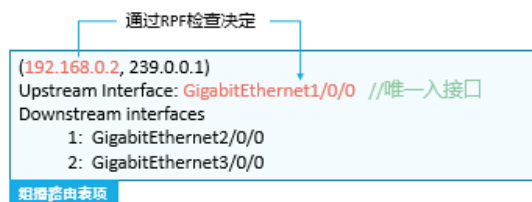
组播数据转发需要依赖路由表项。但是基于目的网络的路由表在转发组播数据时存在一定问题：





组播路由与RPF检查

- 由于组播转发容易产生环路，次优，重复报文，所以组播路由表项除了目的网络和出接口外还需要添加组播源和入接口的信息。设备**仅转发从特定唯一的入接口**收到的组播数据，从而**避免**组播转发时产生**环路，次优，重复报文（部分解决）**等问题。
- 对于相同的组播源，设备通过RPF（Reverse Path Forwarding，反向路径转发）检查可以确定设备上**唯一的组播流量入接口**。
- 组播路由表项以及与RPF检查的关系如下：

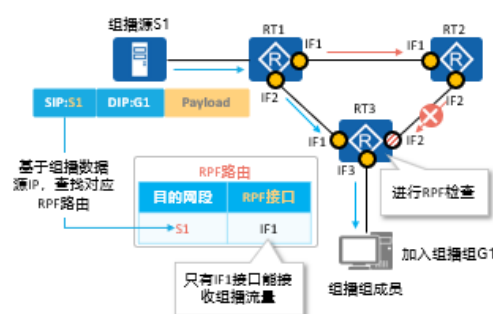
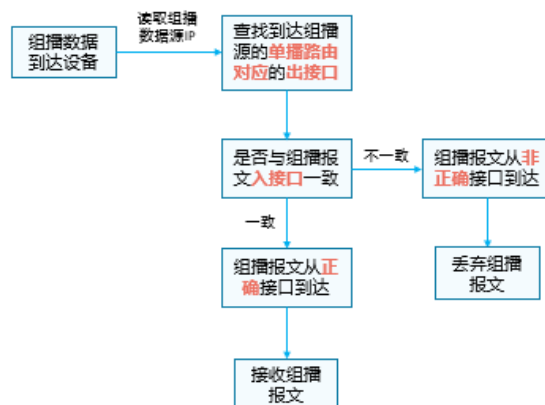


- 组播路由表项出接口一般需要通过组播路由协议确定。
- 组播路由协议将在《PIM 原理与配置》课程介绍。
- 组播路由表项包含组播源与组播组，因此有事又被称为 (S, G) 表项。



RPF检查工作原理

- RPF检查过程如下：



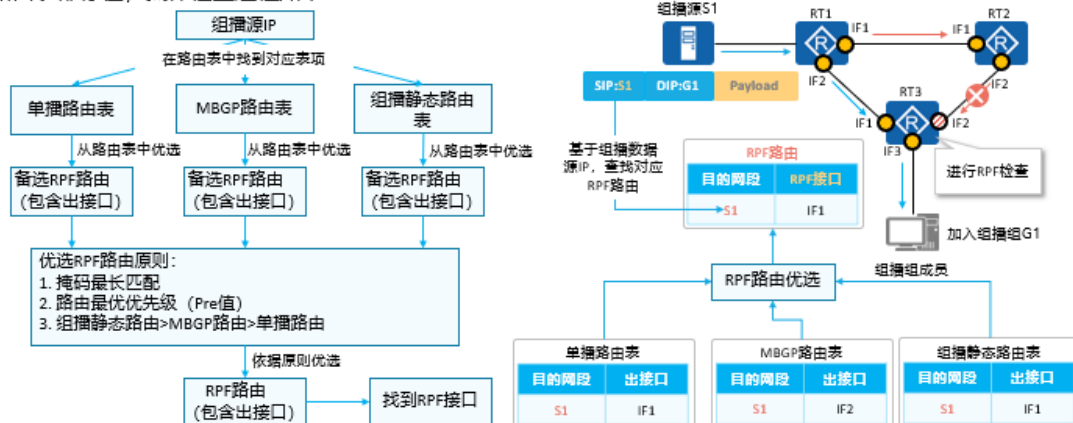
所有组播路由表项都需要进行RPF检查，本示例以RT3为例

- 组播路由器根据报文的源地址通过路由表（单播路由表、MBGP 路由表或组播静态路由表）查找到达“报文源”的路由，查看到“报文源”的路由表项的出接口是否与收到组播报文的入接口一致。如果一致，则认为该组播报文从正确的接口到达，

从而保证了整个转发路径的正确性和唯一性。这个过程就被称为 RPF 检查。

RPF路由选举规则

RPF路由可以从单播路由、MBGP路由、组播静态路由中选举产生。当路由器收到一份组播报文后，如果这三种路由表都存在，具体检查过程如下：

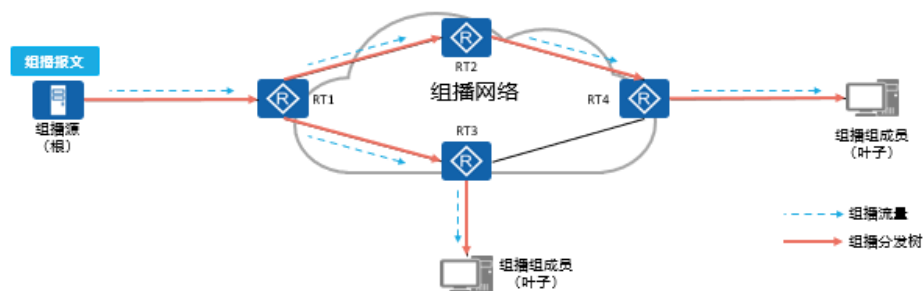


- 根据以下原则从这三条最优路由中选择一条作为 RPF 路由：
- 如果配置了按照最长匹配选择路由，则从这三条路由中选出最长匹配的那条路由；
- 如果这三条路由的掩码一样，则选择优先级最高的那条路由；
- 如果它们的优先级也相同，则按照组播静态路由、MBGP 路由、单播路由的顺序进行选择。
- MBGP：
- MBGP (Multicast BGP , 组播 BGP) 主要用于传递组播源相关的路由条目。
- 组播静态路由表：
- 手工配置组播源与出接口的对应关系。



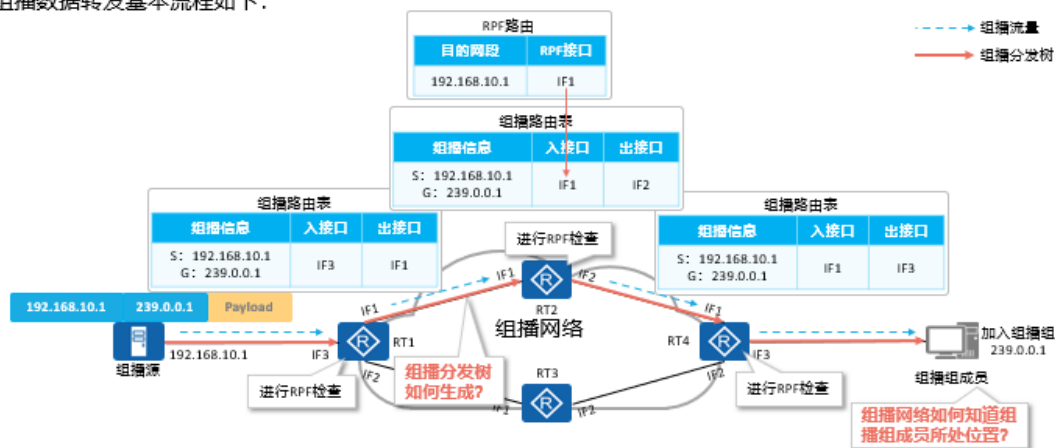
组播分发树

- 组播数据转发需要保证转发路径无环，无次优路径且无重复包。
- 通过RPF机制与组播路由协议，组播网络可以最终形成无环、无次优且无重复包的组播转发路径，该路径可以被称为**组播分发树**。
- 组播分发树以**组播源为根**，以**组成员为叶子**形成转发路径，组播数据在转发时都基于组播分发树进行转发。



组播数据转发流程

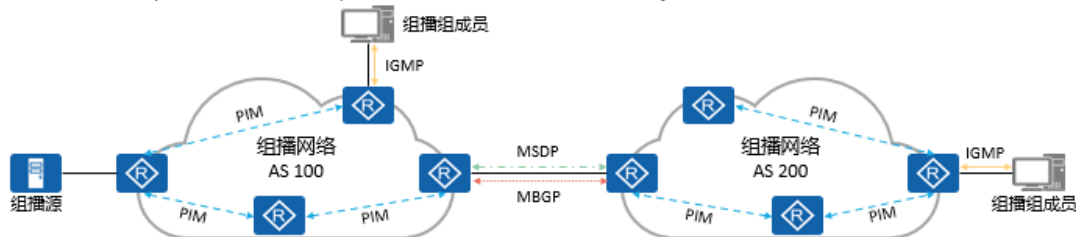
组播数据转发基本流程如下：



- 组播路由表项出接口与组播转发路径由组播路由协议决定。
- 组播路由协议主要有：PIM，MBGP，MSDP。
- 关于组播路由协议的内容将在《PIM 原理与配置》课程讲解。
- 组播组成员位置由IGMP 通告。
- 关于 IGMP 的内容将在《IGMP 原理与配置》课程讲解。

组播协议介绍

- 组播网络需要基于多种组播协议才能建立转发路径：
 - 工作在成员端网络的主要是IGMP (Internet Group Management Protocol, 因特网组管理协议) 协议, 用于告知组播网络, 组成员的位置与所加组播组。
 - 工作在组播转发网络的协议主要有PIM, MSDP, MBGP。
 - PIM (Protocol Independent Multicast, 协议无关组播) 协议主要作用是生成AS域内的组播分发树。
 - MSDP (Multicast Source Discovery Protocol, 组播源发现协议) 主要作用是帮助生成AS域间的组播分发树。
 - MBGP (Multicast BGP, 组播BGP) 主要作用是帮助跨域组播流进行RPF校验。



思考题：

- (单选题) IPv4 组播地址的范围是什么？
- 192.168.0.0~192.168.255.255
- 172.21.0.0~172.21.255.255
- 224.0.0.0~239.255.255.255
- 240.0.0.0~255.255.255.255
- (多选题) 下面描述 RPF 作用正确的是？
- 避免重复的组播报文
- 加速组播流量转发速度
- 防环

答案：

- C
- IANA (Internet Assigned Numbers Authority , 互联网编号分配委员会) 将 D 类地址空间分配给 IPv4 组播使用。IPv4 地址一共 32 位 , D 类地址最高 4 位为 1110 , 因此地址范围从 224.0.0.0 到 239.255.255.255。
- A , C

