组播基础

组播的服务模型有哪几种？
接收者主机接收数据时可以对源进行选择，因此产生了 ASM（Any-SourceMulticast）和 SSM（Source-Specific Multicast）两种服务模型。
这两种服务模型默认使用不同的组播组地址范围。

（1）ASM：任意源模式，接收者主机加入组播组以后可以接收到任意源发送到该组的数据。
1 判断条件：最后一跳路由器生成组播路由条目为（*，G）
2 缺点：可能会收到重复的组播流量；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。
（2）SSM：指定源模式，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。
1 判断条件：最后一跳路由器生成组播路由条目为（S，G）
2 优点：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。

扩展问题 1:SSM 服务模型的缺点是什么？
答：一个组播源对应一颗组播源树。
对于设备来说开销较大，需要耗费较多的开销去维护相应的组播表项；

组播 mac 地址是怎么生成的？组播 mac 地址的作用是什么？使用

过程中需要注意什么问题？
组播 mac 地址是一个虚拟的 mac 地址，组播 IP 地址无法配置在主机上或者某一个接口上，所以无法通过真实的 mac 地址承载组播流量。生成过程是通过 IP 地址和 mac 地址的映射形成：
1 加上 MAC 地址固定前缀（24bit）为：01-00-5E；
2 后面 24bit 由 IP 地址的后 23bit 构成；
3 第 25 bit 位固定为 0；

例如：238.128.128.128，生成的组播 mac 地址为：01-00-5E-00-10-10
此时第 25bit 位固定为 0，所以此时第四字节的结果为 00。

最笨的方法，也是最安全的方法，就是把 IP 地址，换成二进制数。再把二进制换成 16 进制，第 1 个 8 位（从左往右看）不用，因为有组播 MAC 的限制已经规定是 0100.5E 开头，第 2 个 8 位中的第 1 位规定为 0（从左往右看）.所以当不为 0 时，要改成 0 来换算。

源 IP	224.1.1.1
换算成二进制	11100000.00000001.00000001.00000001
再换成 16 进制	01 .01 .01
加上组播 MAC 头	0100.5E
最后形成：	0100.5E01.01.01

就 23 位?

有一个有趣的故事是关于为什么只有 23 位有价值的 MAC 地址空间分配给 IP 组播。回到 20 世纪 90 年代初，Steve Deering 取得了一些关于 IP 组播研究工作的成果，因此，他希望 IEEE 配置 16 个连续不断的组织机构惟一性标识符(OUI)作为 IP 组播 MAC 地址使用。因为一个 OUI 包含 24 位有价值的地址空间，16 个连续不断的 OUI 将提供全部 28 位有价值的 MAC 地址空间，并且允许一对一地把第 3 层 IP 组播地址映射到 MAC 地址。很遗憾，当时一个 OUI 的价格是 $1 000，Steve 的经理，Jon Postel，不愿花 $16 000 购买全部 28 位有价值 MAC 地址。相反，Jon 愿意在预算外花 $1 000 购买一个 OUI，并且拿出一半地址(23 位)给 Steve 供 IP 组播研究之用。

以上内容摘抄自《IP组播网络设计开发（第1卷）》Beau Williamson著

组播 mac 地址的作用是：
1 在组播源泛洪组播数据时，能以组播 mac 地址当成目的 MAC 地址进行数据的正常封装；
2 在接收端会自动生成一份组播 mac 地址。功能是当主机收到一份组播数据时，解封装时读取到数据链路层。即对比数据包的目的 mac 地址就能够判断这份组播数据是否为主机需要的组播数据，能够节省接收端设备的开销。

组播 mac 地址出现的问题---映射缺陷：
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。

会导致 32 个组播 IP 地址映射到相同的组播 MAC 地址，会消耗接收者处理性能
（解封装到三层的 IP 才知道才报文不是本设备需要接收的）
如何解决：
1、规划时避免
2、使用 IGMP-snooping 技术

扩展问题 1:为什么要有组播技术？或者：组播的优势是什么？
（1）与单播比较，可以实现一对多的通信，不用在源处复制多份组播流后发出；
（2）与广播比较，可以节省设备性能，因为组播接收者可以通过目的 MAC 地址就判断自己是否加入了对应的组。

组播路由协议有哪些？用于什么场景？作用是什么？
（1）IPv4 组播协议：
在 IP 组播传输模型中，发送者不关心接收者所处的位置，只要将数据发送到约定的目的地址，剩下的工作就交给网络去完成。网络中的组播设备必须收集接收者的信息，并按照正确的路径实现组播报文的转发和复制。在组播的发展过程中，形成了一套完整的协议来完成此任务。
（2）组播组管理协议 IGMP（Internet Group Management Protocol）：
IGMP 是负责 IPv4 组播成员管理的协议，运行在组播网络中的最后一段，即三层网络设备与用户主机相连的网段内。IGMP 协议在主机端实现组播组成员加入与离开，在上游的三层设备中实现组成员关系的维护与管理，同时支持与上层组播路由协议的信息交互。
到目前为止，IGMP 有三个版本：
IGMPv1、IGMPv2 和 IGMPv3。所有 IGMP 版本都支持 ASM 模型。IGMPv3 可以直接应用于 SSM 模型，而 IGMPv1 和 IGMPv2 则需

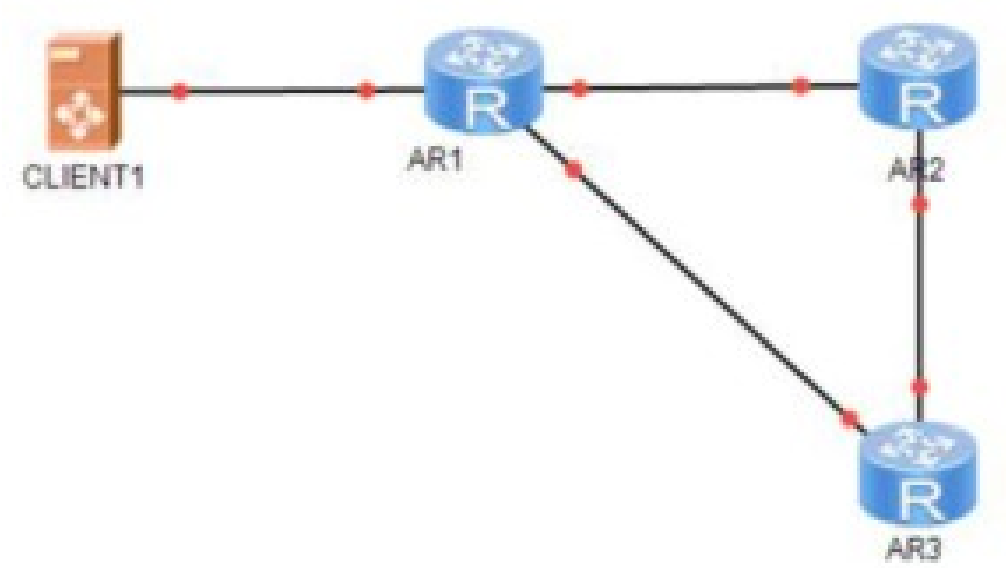
要 SSM Mapping 技术的支持。
（3）IGMP Snooping：
IGMP Snooping 功能可以使交换机工作在二层时，通过侦听上游的三层设备和用户主机之间发送的 IGMP 报文来建立组播数据报文的二层转发表，管理和控制组播数据报文的转发，进而有效抑制组播数据在二层网络中泛洪。
与 IGMP 对应，IGMP Snooping 就是 IGMP 协议在二层设备中的延伸协议，可以通过配置 IGMPSnooping 的版本使交换机可以处理不同 IGMP 版本的报文。
（4）PIM（协议无关组播）：
用于组播路由器和组播路由器之间，实现组播路由器构建组播路由表，协议无关组播指的是与网络层运行哪一种单播路由协议无关，但是网络层必须要运行一种单播路由协议，用于组播数据转发时执行 RPF 检查。

组播分发树有哪些？各自有什么特点？
MDT：组播分发树（指导组播数据流转发）
是组播数据流所经过的路由器的转发路径形成的一颗无环的树
（1）源树：SPT，组播接收者的路由器到组播源的路由器路径最短的树
1 特点：1、转发路径最短
2 根据（S,G）转发组播数据流，上下游接口的设置规则如下：
上游接口：组播流的入接口，只能存在一个（离组播源最近的接口）
下游接口：组播流的转发接口列表（离接收者最近的接口）
3 缺点：多个源存在时，存在多个（S,G）条目，对设备消耗大
优点：转发路径最短

（2）共享树：RPT（RP 汇聚点）组播接收者的路由器到 RP 路由器路径最短的树

1 特点：源到接收者的转发路径不是最短
2 根据（*，G）转发组播数据流，上下游接口的设置规则如下：
上游接口：组播流的入接口，只能存在一个（离 RP 最近的接口）
下游接口：组播流的转发接口列表（离接收者最近的接口）
3 缺点：转发路径不一定是最优路径；
优点：多个源存在时，共享一颗树，对设备消耗较小

RPF 检查是什么？作用是什么？
（1）RPF 检查：接收组播数据流之前或者接收组播协议报文时（例如 BSR 消息）必须执行 RPF 检查，检查通过接收组播流或者报文，检查不通过丢弃组播流或者报文。
RPF：反向路径转发检查
1 避免组播流环路
2 .避免重复组播流
3 .避免组播报文环路（不建议提）



R3 能通过 RPF 检查，确保只接收一个方向泛洪的组播流量；
R3 只接收一个方向的组播数据流量，不会出现 R1-R2-R3 或者 R1-R3-R2 的流量路径，所以不出现环路。

检查过程：
组播流的入接口和相应的 RPF 接口是否是一致，
同时检查组播流的源 IP 地址是否与 RPF neighbor 一致。

a）SPT 树上的 RPF 检查：
收到组播流，查看组播流的 S，根据 S 查找组播路由表，单播路由表中到达 S 的接口为 RPF 接口，对比 RPF 接口和组播流的入接口是否一致，一致则通过
b）RPT 树上的 RPF 检查：
是根据 RP 的地址来选择 RPF 的接口，对比组播流的入接口与到 RP 的接口是否一致

（2）RPF 接口选择的依据：
相应路由协议的优先级，如果优先级一致根据以下规则选择：
1 静态组播路由 ip rpf-route-static （默认优先级为 1）
2 根据 MP-BGP
3 单播路由表（如果单播路由表到达源存在两个下一跳地址，选择下一跳 ip 地址大的作为 RPF 接口）

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：
a）通过报文源地址，分别从单播路由表、MBGP 路由表和组播静态路由表中各选出一条最优路由。根据以下原则从这三条最优路由中选择一条作为 RPF 路由。
b）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。如果没有配置按照最长匹配选择路由，则从这三条路由中选出优先级最高的那条路由；如果它们的优先级相同，则按照组播静态路由、MBGP 路由、单播路由

的顺序进行选择。

最后，路由器会将报文的入接口与 RPF 路由的 RPF 接口进行比较。
如果一致则 RPF 检查通过，表明该报文来源路径正确，会将其向下游转发；
如果不一致即 RPF 检查失败，表明该报文来源路径错误，就将其丢弃。

（4）针对 BSR 消息的 RPF check
1 当一个接口收到 BSR 消息，会根据 BSR 消息中 BSR 的地址执行 RPF check。
当发现收到 BSR 消息的接口不是 RPF check 的接口，发送 BSR 消息的不是自己的 pim 邻居，会将 BSR 消息丢掉。
2 BSR 消息进行 RPF check 的作用：防止环路

扩展问题 1:RPF 的作用？
（1）防止环路
（2）防止重复报文

扩展问题 2:组播中哪些报文需要进行 RPF？
组播流（分为从 SPT 树流下来的组播流和 RPT 树流下来的组播流）
BSR 报文（检查 RPF 接口以及 RPF 邻居）

扩展问题 3:是不是收到每一份组播流量都会进行 RPF 检查？
不是的，因为每次都收到组播流都进行 RPF 检查很消耗设备的性能；
组播路由协议通过已有的单播路由、MBGP 路由或组播静态路由信息来确定上、下游邻居设备，创建组播路由表项。运用 RPF 检查机制，来确保组播数据流能够沿组播分发树（路径）正确的传输，

同时可以避免转发路径上环路的产生。
在实际组播数据转发过程中，如果对每一份接收到的组播数据报文都通过单播路由表进行 RPF 检查，会给路由器带来很大负担。因此，路由器在收到一份来自源 S 发往组 G 的组播数据报文之后，首先会在组播转
发表中查找有无相应的（S，G）组播转发表项：
a）如果不存在（S，G）转发表项，则对该报文执行 RPF 检查，将检查到的 RPF 接口作为入接口，创建组播路由表项，下发到组播转发表中。其中，对 RPF 检查结果的处理方式为：
如果检查通过，表明接收接口为 RPF 接口，向转发表项的所有出接口转发；
如果检查失败，表明报文来源路径错误，丢弃该报文。
b）如果存在（S，G）转发表项，并且接收该报文的接口与转发表项的入接口一致，则向所有的出接口转发该报文。
c）如果存在（S，G）转发表项，但是接收该报文的接口与转发表项的入接口不一致，则对此报文进行 RPF 检查。

对 RPF 检查结果的处理方式为：
a）若 RPF 检查选取出的 RPF 接口与转发表项的入接口一致，则说明（S，G）表项正确，报文来源路径错误，将其丢弃。
b）若 RPF 检查选取出的 RPF 接口与转发表项的入接口不符，则说明（S，G）表项已过时，于是把表项中的入接口更新为 RPF 接口。

然后再根据 RPF 检查规则进行判断：如果接收该报文的接口正是其 RPF 接口，则向转发表项的所有出接口转发该报文，否则将其丢弃。

IGMP 协议

IGMP 是什么？IGMP 的版本有哪些？工作机制是如何的？

（1）IGMP 的作用：IGMP 是 Internet Group Management Protocol 的简称，又被称为互联网组管理协议，是 TCP/IP 协议族中负责 IPv4 组播成员管理的协议。IGMP 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。IGMP 通过在接收者主机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 报文之上，协议号为 2。
（2）应用场景：只要是网络中存在组播接收者的场景，都需要在路由器上开启 IGMP。
（3）版本与工作机制：
到目前为止，IGMP 有三个版本：
IGMPv1 版本（由 RFC 1112 定义）
IGMPv2 版本（由 RFC 2236 定义）---华为默认版本
IGMPv3 版本（由 RFC 3376 定义）

IGMPv1 的工作机制：

1 查询机制：普遍组查询

普遍组查询报文由组播路由器（在 IGMPv1 版本中，由 PIM 选出的 DR 周期性发出）发出，用于查询是否存在组播组每 60S 发送一次，报文发送的目的地址为 224.0.0.1。

2 响应机制：成员报告报文

接收者发出，用于表明加入哪个组播组中

主动发送:主动加入某个组

被动发送:收到普遍组查询报文，被动回复加入某个组

最大响应时间：10s，不可修改（因为报文中不包含此字段），单位为 1s。

3 组成员离开机制（静默离开）：没有特定的离开机制

组播路由只能根据成员报告报文判断是否存在组播组,组播组成员在 120S+10S（ 2 个普遍组查询时间加上 1 个最大响应时间）的时间内没有回复成员报告则认为该组播组已经不存在。

响应抑制机制：
组播路由器发出查询报文后，组播组成员会启动响应定时器（1-10 s 随机），响应时间先到的组成员先回复成员报告报文（同时该成员被选为 last-reporter），其他组成员收到加入组相同的成员报告报文就会取
消发送成员报告报文。

扩展问题 1:组成员是如何收到成员报告报文的？
答：因为成员报告报文的目的 IP 是加入组组播 IP 地址，所以处于同一个组的成员会接收同一组的成员报告报文。

扩展问题 2:响应抑制机制默认开启？
答：是的，默认开启。（IGMPv1、v2 版本默认开启，IGMPv3 无响应抑制机制）

扩展问题 3:IGMPv1 的查询器如果发生故障，另一个组播路由器需要多久才能成为新的查询器周期性发送查询报文？
答：105s，因为 IGMPv1 的查询器是 DR，如果依靠 DR 的老化时间进行收敛。

扩展问题 4:响应抑制机制的好处是什么？
答：组播路由器无需知道有多少成员，只需知道有该组播组存在成员，如有发出查询报文，该组的所有成员都回复成员报告报文，反而对路由器会造成负担，所以响应抑制机制能减少带宽资源消耗与节省设备开销。

扩展问题 5：IGMPv1 的离开机制有什么缺陷？
答：可能会造成在 130s 的时间内组播路由器会一直泛洪组播流量，造成设备开销与链路带宽被浪费。

扩展问题 6：运行了 IGMPv1 的接口是否一定要开启 PIM 协议？
答：是的，因为需要 PIM 协议选出 DR 进行周期性查询的工作。

扩展问题 7：IGMPv1 与 IGMPv2 的兼容性如何？
答：可以实现向下兼容，IGMPv1 会自动忽略掉 IGMPv2 的报文中那些不识别的字段，从而实现兼容。

扩展问题 8: IGMPv1 与 IGMPv2 具有什么限制/缺点？
答：不支持 SSM 组播流量模型。

IGMPV2 的工作机制：
（1）查询机制：
1 普遍组查询报文（同 V1 功能一致）
2 特定组查询报文（新增报文，用于组播路由器针对某一个组发送的查询报文）
特定组查询报文：最后一跳路由器收到成员离开报文时发出，用于查询某个组是否还存在成员，如果存在
成员会收到响应报文，不会删除该组播组，如果发出特定组查询报文在 1S 内没有得到回复，再发送一次（总共发送 2 次），如果还没有收到响应报文，则认为不存在该组播组的接受者。实际上相当于老化时间变为 2s
3 查询器选举机制：
在 IGMPv2 机制中，IGMPv2 新增本身的查询器选举机制，由普遍组查询报文进行选举，
选举条件为：IP 地址越小越优先。
过程：刚开始所有的最后一跳路由器都认为自己为查询器相互发送查询报文，当收到查询报文之后，比较查询报文的源 IP 与自己接口的 IP 地址。如果自己接口的 IP 地址比较小，继续充当查询器，如果自己接口的 IP 地址比较大，不再发送查询报文，被动监听查

询报文。如果 125S 没有收到，就自己充当查询器。

（2）成员报告机制：

与 V1 报文版本完全一致，存在响应抑制机制、last-reporter 等。

（3）成员离开机制

离开报文：由成员离开某个组播组时发出报文

只有 last-reporter 发送的离开报文才会触发特定组查询机制。

最大响应时间机制：

收到普遍组查询报文的主机会在 0-10s 的时间范围内随机挑选一个时间进行响应。例如：主机 A 随机选择时间，假设时间选择为 5s，则主机 A 会等待 5s 后再发送成员报告报文。

Last-reporter 机制：

Last-reporter 被称为：最后一个报告者，每一次普遍组查询报文第一个响应的主机被选为 last-reporter。组播路由器通过记录主机的 IP 地址而而记录 last-reporter。

扩展问题 1： V2 的普遍组查询报文相比较于 V1 的报文格式上，有什么变化？

增加了一个最大响应时间字段，默认为 10.0 秒，可以进行调整，调整范围为 1.0-25.0 秒的范围。

好处：

1.可以调整最大响应时间的范围（1-25s）

2.最大响应时间的单位为 0.1S，最大响应时间相同的概率小

3.主机数量多时调大，可以减少最大响应时间相同的概率，主机数量少时可以调小，加快收敛。

扩展问题 2：特定组查询报文跟普遍组查询报文的区别有哪些？

1.目的 IP 地址不同；

2.最大响应时间不同，并且特定组查询报文的最大响应抑制时间为

1s，无法修改
3.group address 字段填充内容不同。

扩展问题 3：为什么只有 last-reporter 会触发特定组查询机制？
因为组播路由器没有收到 last-reporter 报文，则可以确定该组播组存在相应的组播组成员；

扩展问题 4： 如果主机是组播组的最后一个成员，但是由于主机掉电，无法发送离开报文时，组播路由器怎么办？
组播路由器会沿用 IGMPv1 版本的机制，等待老化时间（2 个查询时间+1 个最大响应时间）后删除该组播组表项。

IGMPv3 工作机制

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。在工作机制上，与 IGMPv2 相比，IGMPv3 增加了主机对组播源的选择能力

（1）新增报文：

1 特定源组加入：
IGMPv3 的成员报告报文的目的地址为 224.0.0.22（表示同一网段所有使能 IGMPv3 的路由器）。通过在报告报文中携带组记录，主机在加入组播组的同时，能够明确要求接收或不接收特定组播源发出的组播数据
2 特定源组查询：
当接收到组成员发送的改变组播组与源列表的对应关系的报告时（比 CHANGE_TO_INCLUDE_MODE，CHANGE_TO_EXCLUDE_MODE），IGMP 查询器会发送特定源组查询报文。如果组成员希望接收其中任意一个源的组播数据，将反馈报告报文。IGMP 查询器根据反馈的组成员报告更新该组对应的源列表

（2）与 IGMPv2 相比，IGMPv3 报文的变化有哪些？

1 IGMPv3 报文包含两大类：查询报文和成员报告报文。

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。igmpv3 成员离开通过发送加入组播源地址列表为空的成员报告报文来表示离开

2 查询报文中不仅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）。该报文由查询器向共享网段内特定组播组成员发送，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 成员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据。IGMPv3 增加了针对组播源的过滤模式（INCLUDE/EXCLUDE），将组播组与源列表之间的对应关系简单的表示为（G，INCLUDE，(S1、S2...)），表示只接收来自指定组播源 S1、S2……发往组 G 的数据；或（G，EXCLUDE，(S1、S2...)），表示接收除了组播源 S1、S2……之外的组播源发给组 G 的数据。当组播组与组

播源列表的对应关系发生了变化，IGMPv3 报告报文会将该关系变化存放于组记录（Group Record）字段，发送给 IGMP 查询器。

4 在 IGMPv3 中一个成员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报告只能携带一个组播组。这样在 IGMPv3 中报文数量大大减少

总结：1.报文类型 2.查询报文 3.成员加组的方式 4.报文格式

扩展问题 5：IGMP SSM Mapping 的作用和实现机制？

为 IGMPv1 主机和 IGMPv2 主机提供 SSM 服务

通过在路由器上静态配置 SSM 地址的映射规则，将 IGMPv1 和 IGMPv2 报告报文中的（*，G）信息转化为对应的（S，G）信息

扩展问题 6：为什么 IGMPv3 没有被广泛应用？（重要）

（1）IGMPV3 移除了响应抑制机制，而 v3 的成员报告报文中加入一个组时，可能携带很多的源 IP 地址，这样的话成员报告报文的大小会比较大，查询器处理起来更加消耗设备的性能；
（2）SSM 模型没有被广泛应用；

扩展问题 7：IGMPv2/IGMPv3 的查询器选举除了比较 ip 地址外，还有什么比较规则吗？
没有，因为在普遍组查询报文中，没有 priority 之类的字段；
IGMPV2 只通过 IP 地址比较出查询器，由 ip 地址小的充当查询器；

扩展问题 8：IGMPv3 与 IGMPv1/v2 的版本的兼容性如何？
兼容性较差，原因是 IGMPv3 使用的组播 IP 地址为 224.0.0.22 为目的进行报文的发送，但是当组播路由器为 IGMPv3 版本时，主机使用 IGMPv1/v2 版本时，可以使用 IGMP-SSM mapping 技术进行兼容；

IGMP SSM mapping：
如果接收者想要使用 SSM 模式接收组播源，但是接收者并不支持 IGMPV3 时使用
如何实现：
通过在路由器上静态配置 SSM 地址的映射规则，将 IGMPv1 和 IGMPv2 报告报文中的（*，G）信息转化为对应的（S，G）信息使用 SSMmaping 功能，成员加入的组播组地址必须为 232.0.0.0-232.255.255.255（可以通过命令进行修改）

注意： 组播路由协议必须是 PIM-SM 模式，路由器的 IGMP 版本必须为 IGMPV3 版本

配置：
igmp

ssm-mapping 232.1.1.1 255.255.255.255 200.1.1.1
手工建立源 IP 地址与 GROUP 组地址的映射关系。

interface GigabitEthernet0/0/0
ip address 192.168.1.254 255.255.255.0
igmp enable
igmp version 3
igmp ssm-mapping enable

检查：
display igmp group ssm-mapping
display igmp routing-table

IGMP Snooping

IGMP Snooping 的机制？同 proxy 有什么不同？snooping 有什么缺点？snooping 检查所有组播报文吗？如何做到的？
（1）IGMP Snooping 的作用
IGMP Snooping 是二层组播的基本功能，可以实现组播数据在数据链路层的转发和控制。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

扩展问题 1：IGMP snooping 的好处是什么？
默认情况下，交换机收到组播报文后立刻执行泛洪操作，无法保证组播流量的安全性与有偿性（费用统计），同时也会浪费设备的带宽资源与开销。所以使用 IGMP snooping 技术创建二层组播转发表，组播流量按照二层组播转发表去转发组播流量，所以能够节省设备开销与链路带宽资源。

（2）IGMP Snooping 的两种端口角色：

路由器端口（router port）

1 .通过监听 IGMP 普遍组查询消息或 PIM Hello 消息来感知哪些接口是路由器端口

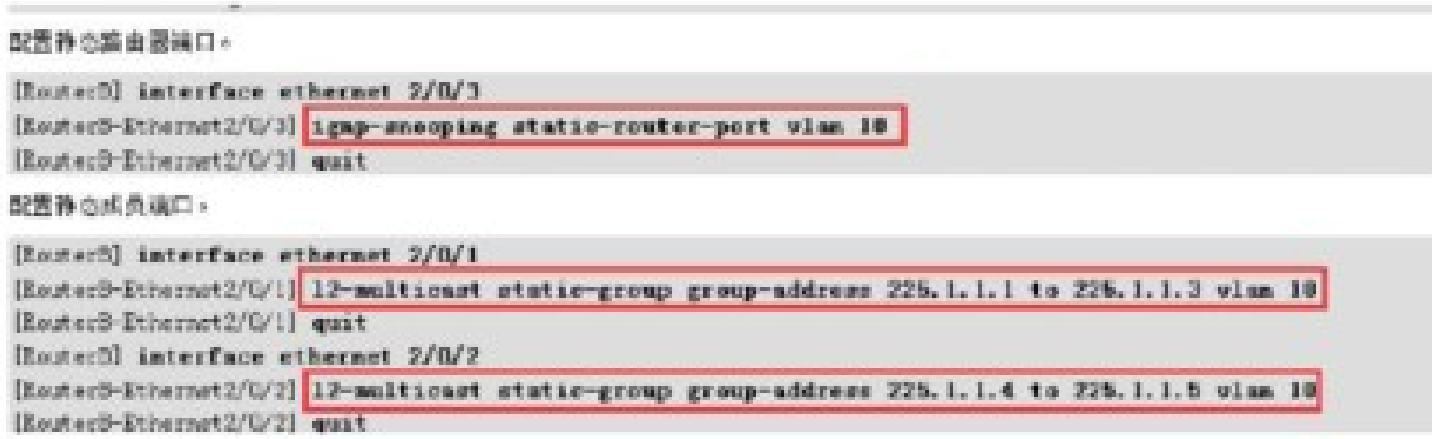
2. 通过手工指定路由器端口

（180s 内如果没有收到 igmp 查询消息或者 pim hello 消息，会从组播表中消失）

成员端口（Member port）

1 收到 IGMP Report 报文的接口将被认为是动态成员端口

2 手工配置成员端口



（3）IGMP Snooping 的工作机制

1 普遍组查询

当二层组播设备收到普遍组查询消息时，向 VLAN 内除接收接口外的其他所有接口转发，并对接收接口做如下处理：

a）如果路由器端口列表中尚未包含该接口，则将其添加进去，并启动老化定时器；

b) 如果路由器端口列表中已包含该动态路由器端口，则重置老化定时器。

说明： 收到 IGMP 普遍组查询报文时，动态路由器端口的老化定时

器缺省为 180 秒，可以通过命令行配置。收到 PIM Hello 报文时，动态路由器端口的老化时间为 Hello 报文中 Holdtime 字段的值(105 s)

2 成员报告

当二层组播设备收到成员报告消息时，向 VLAN 内所有路由器端口。从报文中解析出主机要加入的组播组地址，并对接收接口做如下处理：

a)如果不存在该组对应的转发表项，则创建转发表项，将该接口作为动态成员端口添加到出接口列表中，并启动老化定时器。

b)如果已存在该组对应的转发表项，但出接口列表中未包含该接口，则将该接口作为动态成员端口添加到出接口列表，并启动老化定时器。

c)如果已存在该组所对应的转发表项，且出接口列表中已包含该动态成员端口，则重置其老化定时器。

说明： 收到 IGMP 报告报文后，动态成员端口的老化定时器（2x60+10=130s）= 健壮系数 x 普遍组查询间隔+ 最大响应时间

总结：如有组成员对应表项，则更新老化计时器，
　　如果没有对应表项，则新增。

3 成员离开组播组

当二层组播设备收到 IGMP 离开消息时，判断离开的组是否存在的转发表项，以及转发表项出接口列表是否包含报文的接收接口：

a)如果不存在该组对应的转发表项，或者该组对应转发表项的出接口列表中不包含接收接口，二层组播设备不转发该报文，将其直接丢弃。

b)如果存在该组对应的转发表项，且转发表项的出接口列表中包含该接口，二层组播设备会将报文向 VLAN 内所有路由器端口转发路由器收到该离开消息（last-report）后，会回一个特定组查询消息，二层组播设备收到后会将特定组查询消息向有特定组成员的接口转

发.对于 IGMP 离开报文的接收接口（假定为动态成员端口），二层组播设备在其老化时间内：
c)如果从该接口收到了主机响应 IGMP 特定组/源组查询的报告报文，表示接口下还有该组的成员，于是重置其老化定时器。
d)如果没有从该接口收到主机响应 IGMP 特定组/源组查询的报告报文，则表示接口下已没有该组成员，则在老化时间超时后，将接口从该组的转发表项出接口列表中删除
说明： 收到 IGMP 离开报文后，动态成员端口的老化定时器（2x1=2s） = 健壮系数 x 特定组查询间隔

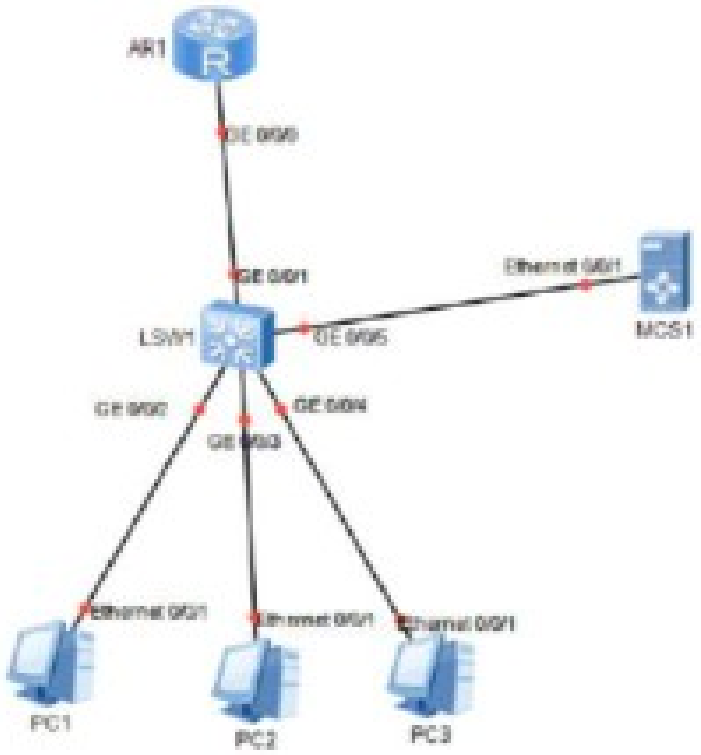
总结：如果收到离开报文中的 group 不在表项内，直接丢弃；特定组查询报文才是影响表项的关键。

(4)IGMP Snooping 的缺点有哪些？
1 CPU 和内存开销问题，CPU 需要处理 IGMP 报文来维护二层组播转发表，如果交换机下的组很多的话，二层组播转发表需要占用的内存也是相当大的
2 如果在不能识别 IGMP 报文的二层交换机上打开 igmp snooping 会监听所有的组播报文，并且会把这些组播报文全部送给 CPU 处理，会加重 cpu 的负载

扩展问题 2：IGMP Snooping 检查所有组播报文吗？如何做到的？
IGMP Snooping 不一定是检查所有组播报文的，要看交换引擎是否识别 igmp 报文。如果交换引擎识别 igmp 报文，只会把 igmp 报文送给 cpu 处理，其他组播报文就会按照二层组播转发表进行转发。如果交换引擎不识别 igmp 报文，则会把所有的组播报文都送给 cpu 处理。通过 IP 包头的协议字段来识别的（pim：103，IGMP：2，UDP：17，TCP：6）

此场景，具有什么缺点？

如果交换机 SW1 开启了 igmp-snooping，在正常的情况下，G0/0/5 接口不为路由端口也不为成员端口，所以当 MCS1 泛洪组播流量时，SW1 会执行将组播流量泛洪的动作，同时 SW1 收到大量的组播流量并且无法识别出这些组播流量是否属于 IGMP 报文，所以 SW1 会将组播流量上送给 CPU，进行识别与判断，造成 CPU 开销的大量浪费。



总结：igmp-snooping 机制在该场景下无法生效的同时，会造成链路带宽与设备 CPU 的大量消耗。
解决方案：将 G0/0/5 接口手工指定为路由端口。

扩展问题 4：IGMP snooping 产生的背景？
答：针对交换机直接泛洪组播流量带来的问题：
1.浪费链路带宽
2.安全性
3.有偿性---对应 ISP 来说
4.消耗设备性能（针对不同组播 IP 映射到相同 MAC 带来的问

题：）

扩展问题 5：二层组播转发表项，如何删除对应的表项？
（1）2S，特定组查询；
（2）130S，两次普遍组查询+响应抑制；

扩展问题 6：如果没有手动的配置路由器端口角色的功能会怎么样？
场景：SW 上连着一台组播源
（1）浪费带宽等；
（2）交换机接口根据 IP 报文的 protocol 字段无法判断是否为 IGMP 报文，只能转交给 CPU 处理，但是 CPU 也无法识别，导致交换机性能大量消耗；

注意只有开启了 multicast routing-table 的路由器才会接收组播报文，否则收到组播报文都是直接丢弃；

PIM 的基本概念

（1）PIM（协议无关组播路由协议）是什么协议？作用是什么？
作用：用于组播路由器与组播路由器之间构建组播路由表
协议无关组播：与网络层运行哪一种单播路由协议无关，但是网络层必须要运行一种单播路由协议，用于组播数据转发时进行 RPF 检查。
特点：
1 工作在一个 AS 内组播路由协议
2 PIM 分为两种模式，PIM-SM,PIM-DM

使用 PIM 的位置：
1 连接组播源的接口
2 需要转发组播路径经过的接口

3 连接接收者的接口（可选，运行 IGMPV1 的接口必须要配置 PIM）

PIM 协议的角色：
1 第一跳路由器：连接组播源的路由器（DR）
2 最后一跳路由器：连接接收者的路由器（DR）
3 中间路由器：负责转发组播数据流
4 叶子路由器：与用户主机相连的 PIM 路由器，但连接的用户主机不一定为组成员
5 MDT：组播分发树，组播数据流所经过的路径形成的一颗无环的树

DR 主要作用如下：
1 在连接组播源的共享网段，由 DR 负责向 RP 发送 Register 注册报文。
2 在连接组成员的共享网段，由 DR 负责向 RP 发送 Join 加入报文。

注 1：在 DM 模式中也选 DR，但不发挥作用
注 2：每个中间网段也会选举 DR，SM 模式中，DR 只在头一条和最后一跳发挥作用。

扩展问题 1：如何判断网络中使用 PIM-DM 还是 PIM-SM 作为组播路由协议？
通过网络管理员对网络中接收者的分布情况与组播路由器上（*，G）条目的数量而决定的。

扩展问题 2：如果接收者的数量很多，是不是一定会使用 PIM-DM 模式？
不是，主要是通过组播路由器上（*，G）条目的数量而决定的，如果出现一个组播路由器下挂着 100 多个接收者。那么基于整体考虑，

还是使用 PIM-SM 模式会比较节省组播网络中组播路由器的设备开销。
注意：组播路由协议的选择主要是看接收者分布的密集程度

扩展问题 3：有两台路由器在同一网段，哪台路由器负责向组播接收者发送组播数据流？如果是在组播源处，谁负责接收组播流？
在 PIM-DM 中，由 Assert Winner 负责转发组播流，组播源处都会接受组播流，转不转发组播流需要的组播转发表（看<S，G>或者<*，G>的下游接口是否为空，如果为空，不转发，如果不是空的，朝下游接口转发）
在 PIM-SM 中，由 DR 负责转发组播流，在组播源处都会接收组播源，只由 DR 负责向 RP 发送注册消息，RP 朝着源的发送 join 消息，根据（S，G）条目来进行组播流转发

PIM-DM 工作原理

PIM-DM 属于密集模式的组播路由协议，使用“推（Push）模式”传送组播数据，通常适用于组播组成员相对比较密集的小型网络，其基本原理如下：
PIM-DM 假设网络中的每个子网都存在至少一个组播组成员，因此组播数据将被扩散（Flooding）到网络中的所有节点。然后，PIM-DM 对没有组播数据转发的分支进行剪枝（Prune），只保留包含接收者的分支。这种“扩散—剪枝”现象周期性地发生，被剪枝的分支也可以周期性地恢复成转发状态。
当被剪枝分支的节点上出现了组播组的成员时，为了减少该节点恢复成转发状态所需的时间，PIM-DM 使用嫁接（Graft）机制主动恢复其对组播数据的转发。
一般说来，密集模式下数据包的转发路径是有源树（Source Tree，即以组播源为“根”、组播组成员为“枝叶”的一棵转发树）。由于有源树使用的是从组播源到接收者的最短路径，因此也称为最短路径树（Shortest Path Tree，SPT）。

PIM-DM 的工作机制可以概括如下：
邻居发现

构建 SPT

嫁接

断言

PIM-SM（稀疏模式）

想象网络的接收者分布比较稀疏，接收者需要流量需要向 RP（汇聚点）进行请求，当组播源活跃需要向 RP 注册，发现是否存在接收者，如果存在才会发送组播流量。
特点：
1 .使用“拉（Pull）模式”转发组播报文
2 .存在 RP
3 .存在 SPT 以及 RPT 树，流量根据 SPT 以及 RPT 树转发
4. 适用于大型组播网络

工作机制：
（1）邻居发现
建立邻居的过程同 PIM-DM 一致，协商参数、报文格式类型也完全一致。DR 竞选规则也一致。
DR 作用：
1 .在最后一跳路由器给 IGMPV1 充当查询器；
2 .在最后一跳路由器只有 DR 路由器才会向 RP 建立 RPT 树（避免收到重复组播流）；
3 .在最后一跳路由器由 DR 发出 RPT 树向 SPT 树切换；
4 .在第一跳路由器由 DR 向 RP 发出注册；

（2）RP 发现

注意问题：
1 每一个组播组只能存在一个 RP；
2 .RP 的地址必须要全网可达能通过 IGP 协议实现访问（建立 rpt 树）；
3 .RP 所在的接口地址必须要运行 PIM SM。

成为 RP 的方式：
手工指定
a)每台组播路由器都需要配置
b）没有冗余备份（对于一个组播组）
基于 msdp 协议，静态 RP 可以实现备份。

动态发现
a）不需要每台组播路由器配置 RP，可以自动学习到 RP
b）有冗余备份，一个组播组可以配置多个 C-RP，一个主多个备份

实现：
BSR：负责收集以及通告 RP 的信息,网络中可以存在多台 C-BSR
RP：为一个或者多个组提供 RP 服务。网络中可以存在多台 C-RP

工作流程
1 网络中所有的 C-BSR 都会以自己为 BSR 向所有的 PIM 邻居发出 BSR 报文（组播发送）;
2 PIM 路由器收到 C-BSR 消息之后会选出最优的一台作为 BSR:

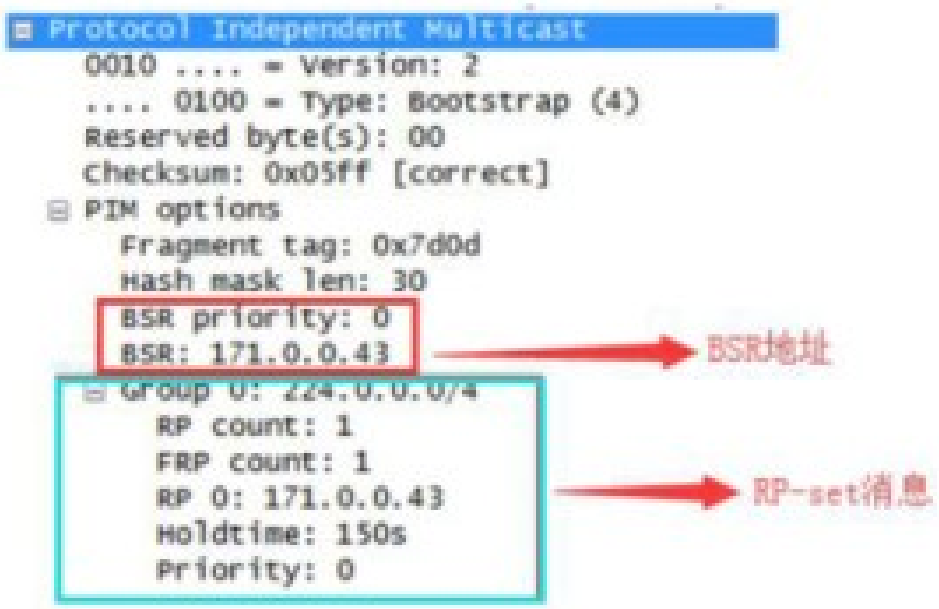
BSR 选举规则：
a）比较 BSR 的优先级（默认为 0，越大越优先）；
b）比较 BSR 的 IP 地址，越大越优；

3 .网络中选出最优的 BSR 之后，BSR 会周期性每 60s 泛洪 BSR

消息。所有的 C-RP 会向 BSR 单播发送 RP 的通告报文。RP 通告包含（RP 服务的组范围，RP 的优先级，RP 的 IP 地址）；
4 .当 BSR 收到所有 RP 的通告报文之后，汇总成 RP-set 发送给所有的 PIM 邻居；
5 .PIM 路由器根据收到的 BSR 报文中的 RP-SET 信息，选出最优的 RP

选举规则如下：
a).RP 服务组范围的精确度，范围越精确越优；
b)RP 的优先级，越小越优,RP 优先级默认为 0；
c)如果优先级相同，则执行 Hash 函数，计算结果较大者获胜；
d).如果以上都相同，则 C-RP 地址较大者获胜。

Bsr 消息报文



（3）RPT 构建:

1 .由最后一跳 DR 路由器发起，当收到 IGMP 的报告消息，会产生相应的（*,G）条目，然后向去往 RP 的上游路由发出（*,G）的 join 消息

2 .上游路由器收到（*,G）的 jion 消息之后，会检查（*,G）的 jion 消息的上游邻居接口和接收报文的接口 IP 地址是否是一致，如果一致创建（*,G）的路由条目，并且继续往指向 RP 的上游路由器发送（*,G）的
jion 消息
上游接口：指向 RP 的接口
下游接口：收到（*,G）的 jion 消息接口
3 .上游路由逐跳往 RP 发送（*,G）的 jion 消息构建 RPT 树，直达 RP 路由器为止
上游接口：NULL
下游接口：收到（*,G）的 jion 消息接口

（4）组播源注册

1 .当第一跳路由器 DR 收到来自组播源的流量，在第一跳路由器生成(S,G),并且会向 RP 发出单播的注册报文，单播的注册报文包含组播的流量
2 .当 RP 收到注册报文，会检查注册报文中的(S,G)是否在 RP 上存在相应（*，G）；
a) 如果不存在，向组播源发送注册停止报文，组播源收到注册停止消息，停止发送注册报文 60S
b)如果存在，RP 会将组播报文沿着 RPT 树发送给接收者，并且在 RPT 树上会生成相应的（S,G）条目
3 .RP 会向组播源的方向发送（S,G）的 jion 报文，上游的路由器收到（S,G）的 jion 报文也会创建（S,G）的路由条目，并且继续向组播源的方向发送（S,G）的 jion 报文，直至组播源的第一跳路由器为止；
4 .RP 从 SPT 树收到组播流之后，会单播向第一跳 DR 路由器发送注册停止报文。

（5）SPT 切换

1 好处:
a)组播源到接收者走为最优的路径
b）.减轻 RP 的负担

2 切换条件：由最后一跳路由器的 DR 发起，当接收的组播流达到一定的阈值（默认为 0）

3 切换的过程：
a)当最后一跳路由器的 DR 组播流达到一定的阈值，会向组播源的方向发出（S,G）的 jion 消息；
b).当上游路由器收到（S,G）的 jion 消息，会将（S,G）的 flag 位切换为 SPT 置位，上游接口为离组播源最近的接口，下游接口为接收（S,G）的 jion 消息接口，并且继续往组播源的方向发（S,G）的 jion 消息，建立 SPT 树；
c).当分叉点的组播路由器收到两份组播流，会触发（S,G）rpt 置位的剪枝报文，朝 RP 的方向进行发送，作用是为了删除掉 RPT 树上（S,G）的下游接口。

扩展问题 1：静态 RP 与动态 RP 的优缺点分别是什么？
静态 RP 的优点：配置简单、快捷，无周期性发送报文机制，设备开销小；
缺点：无备份机制、负载均衡机制；
动态 RP 的优点：能实现备份，能根据网络拓扑的变化，进行收敛；
缺点：会额外地消耗设备的开销，存在报文的周期性发送，收敛慢

扩展问题 2：动态 RP 的 C-BSR 的收敛时间是多少？RP 的呢？
c-bsr 的收敛时间为 180s，c-rp 的收敛时间为 150s。

扩展问题 3:如果网络中的 RP 设备发生故障，其他的路由器需要多久才能得知 RP 发生故障？

180s，因为 bsr 需要 150s 确定 RP 发生故障，但是要通知其他路由器，需要第三份 BSR 消息才能通知全网的路由器。

扩展问题 4: 在 PIM-SM 中，存在断言场景吗？
答：在一般情况下，PIM-SM 不存在断言场景，因为 RPF 树的构建，发送的 join 报文是只能发送给一个上游邻居的，所以无法触发断言。
特殊场景：PIM-SM 中可能出现断言：

扩展问题 5:注册报文是不间断发送的吗？为什么？
答：因为当组播源活跃时，注册报文已经会携带组播流量发送给 RP，RP 一旦收到之后就立刻沿着 RPT 树泛洪，所以需要保证组播流量的泛洪，需要不间断地发送注册报文。

扩展问题 6：RP 什么时候会发送注册停止报文？
<1>收到 SPT 树过来的组播流之后；
<2>收到的注册报文中包含的（S，G）与 RP 上建立的 RPT 树不相符，说明 RP 不需要接收此组播流，触发注册停止报文；

扩展问题 7：RP 如何判断收到的组播流量是从 SPT 树流过来的？
答：通过组播流量中的 flag 位去进行判断。

扩展问题 8：注册消息是单播发送的吗？
答：是的，因为注册报文前面封装 PIM 报文，在 PIM 报文前面在封装单播 IP 报文，路由器收到后将注册报文当成普通的单播报文进行转发。

扩展问题 9:SPT 切换的作用是什么？
（1）防止次优路径---使得源到接收者的组播流量走最优的路径；
（2）RP 负荷过重，减轻 RP 的负担

扩展问题 10:何时发生 Switchover？
在最后一跳路由器上收到第一份组播报文时，会检查组播报文是否超过接口设置的阈值，如果超过阈值，就会发生 SPT 切换（默认开启）

扩展问题 11:pim-sm 中，生成 RPT 树后，成员如何离组？
成员侧 DR 确定底下没有加入相应组的成员时，删除对应的（*，G）表项，并向上游发送 prune 报文，上游邻居收到 prune 报文后，直接删除（*，G）表项，然后继续向上游发送 prune 报文；

扩展问题 12:（*，G）join 报文和（S,G）join 报文的区别是什么？
（*，G）join 报文中的 ip address 为 RP
（S,G） join 报文中的 ip address 为组播源

扩展问题（*，G）prune 报文和（S,G）prune 报文的区别？
（*，G）join 报文中 ip address 为 RP
（S,G） join 报文中 ip address 为组播源

PIM-SSM（指定源模式）

（1） 应用场景：在接收者想指定接收来之哪一个组播源的流量时使用

1 .属于一种特殊的 SM 模式，使用“拉（Pull）模式”转发组播报文
2 .无 RP
3 .无 RPT 树，只存在 SPT 树
4 .必须要运行 IGMPV3

（2）工作机制：
1. 邻居发现
2 .接收者加入，运行 IGMPV3 的最后一跳路由器会生成相应的（S,

G）条目，并且根据(S,G)中的 S 找到组播源对应的出接口，然后向该接口发出（S,G）的 jion 消息
3 .上游路由器收到（S,G）的 jion 消息，会生成相应的（S,G）条目，并且继续向源的方向继续发送（S,G）的 jion 消息，建立 SPT 树。

配置：
1 、每台路由器的接口运行 PIM-SM
2 、接收者和最后一跳路由器运行 IGMPV3

SSM-maping 在接收者不支持 IGMPV3 时使用
igmp
ssm-mapping 231.2.2.0 255.255.255.0 200.1.1.1
ssm-mapping 232.2.2.2 255.255.255.255 200.2.2.2

接口下
igmp ssm-mapping enable
默认只有 232.0.0.0-232.255.255.255 支持 SSM，可用通过 SSM-policy 修改，让更多的组支持 SSM

扩展问题 1:运行 PIM-SSM 的条件？
1、接收者侧运行 IGMPV3；
2、运行 PIM-SM；
3、组播成员加入的组范围为 232.0.0.0~232.255.255.255；