

第 5 篇 IP 组播

第 18 章 IP 组播概述

第 19 章 组播组管理协议

第 20 章 组播转发机制

第 21 章 组播路由协议

第 22 章 组播配置和维护

第18章 IP 组播概述

IP 组播技术实现了数据在 IP 网络中点到多点的高效传送，能够节约大量网络带宽、降低网络负载。通过 IP 组播技术可以方便地在 IP 网络之上提供一些增值业务，包括在线直播、网络电视、远程教育、远程医疗、IP 监控、实时视频会议等对带宽和数据交互的实时性要求较高的信息服务。

本章对比了组播和单播、广播的不同，并介绍了组播的体系架构和组播模型。

18.1 本章目标

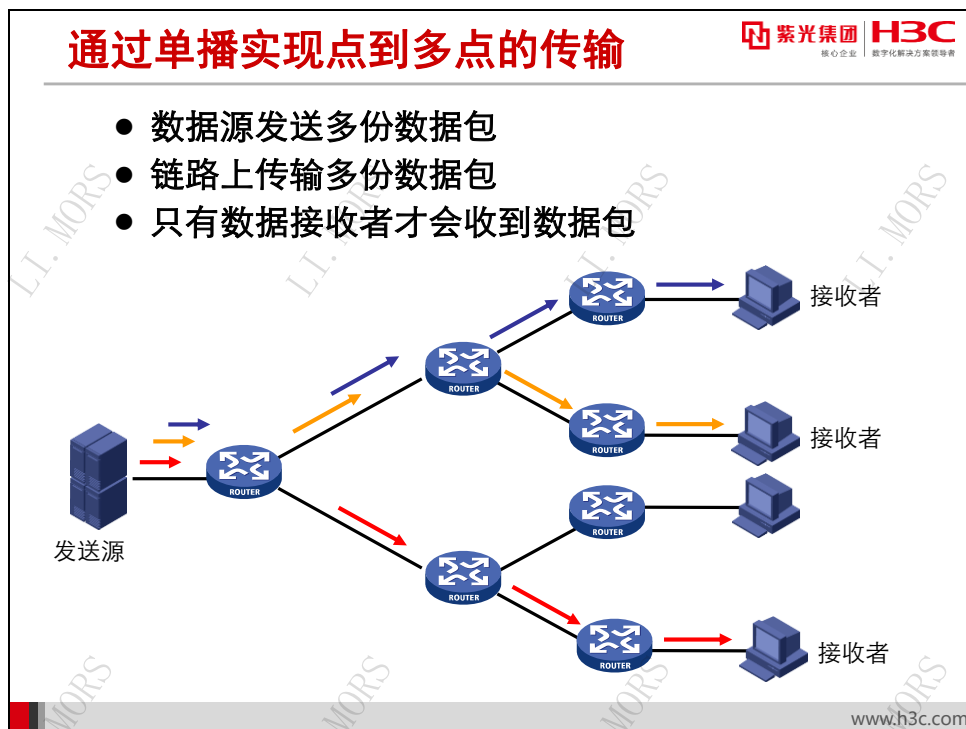
课程目标

学习完本课程，您应该能够：

- 了解组播概念
- 了解组播的优缺点及典型应用
- 掌握组播技术体系架构
- 了解组播模型的分类



18.2 组播介绍

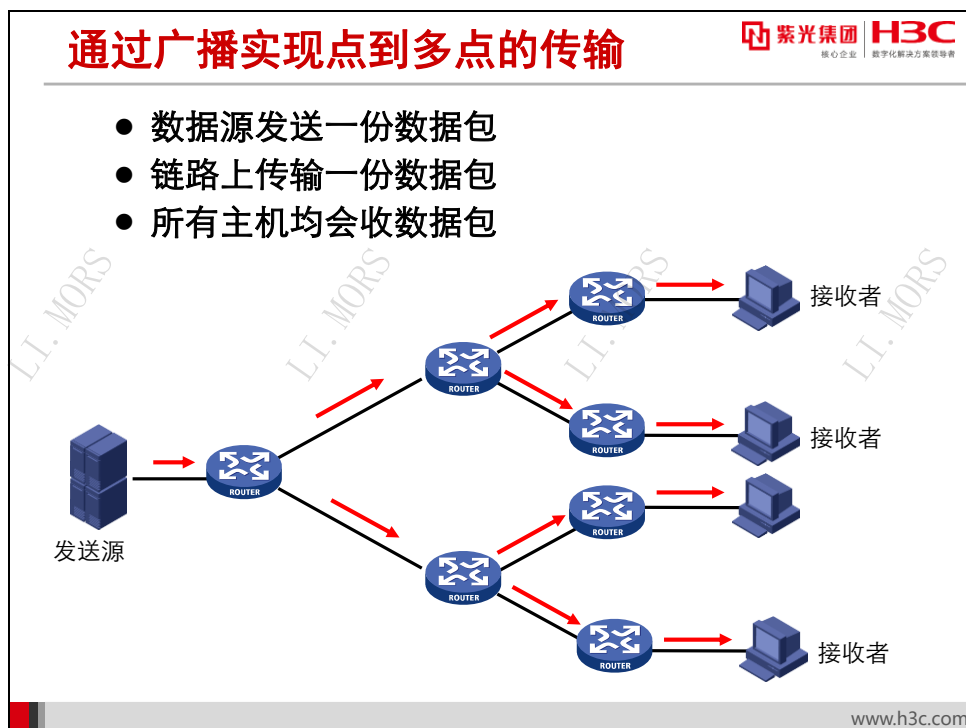


在 IP 网络中，节点之间的通信通常采用点到点的方式，即在同一时刻，一个发送源只能发送数据给一个接收者，这种通信方式称为单播。单播以其简洁、实用的通信方式在 IP 网络中得到广泛使用。

对于某些网络应用如多媒体会议、IP 视频监控，需要发送源将数据发送给网络中的部分接收者，不是一个也不是全部，这种传输方式称为点到多点的传输。

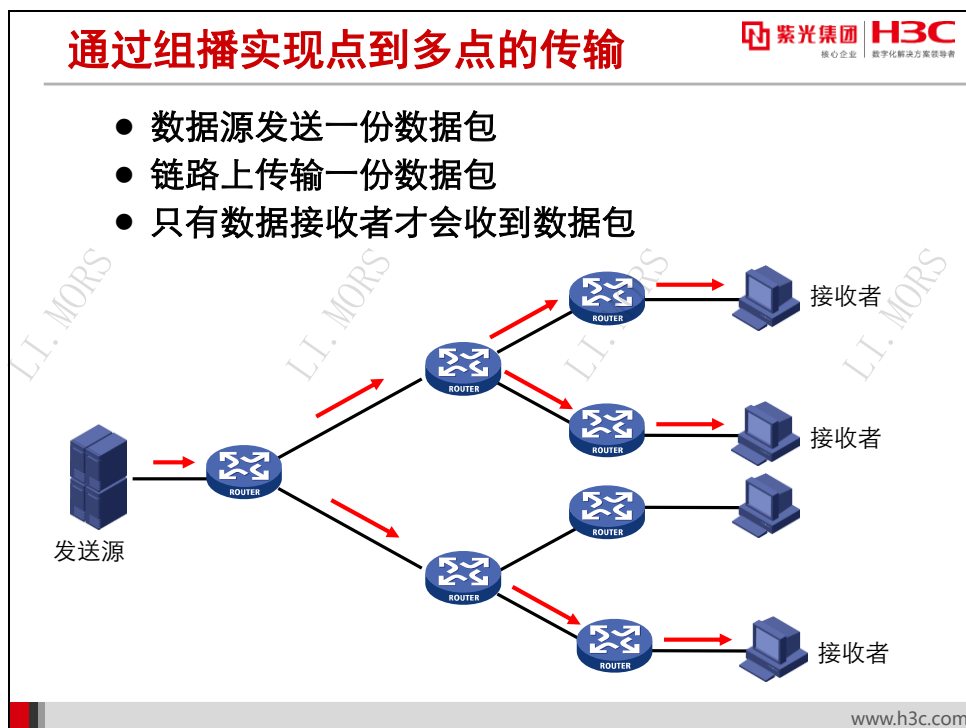
采用单播实现点到多点的传输时，发送源需要向每一个接收者单独发送一份数据，当接收者数量增加时，发送源复制数据的工作负荷也会成比例增加。此外，由于同一个时刻发送源只能发送数据给一个接收者，当接收者数量巨大时，一些接收者接收数据的延时会大大增加，这对于一些延时敏感的应用如多媒体会议、视频监控等，是不可接受的。

从图中还可以看到，链路上可能会传输大量目的地址不同但是内容完全相同报文，这些内容相同的报文占用了大量的链路带宽，降低了链路带宽的有效利用率。



通过广播方式也可以实现网络中多个接收者收到发送源发送的数据，并且不管接收者数目是多少，发送源只需要发送一份广播数据。

相对于单播，广播方式减少了发送源的处理，降低了发送源的负荷。但是采用广播方式，网络中的所有主机都会收到广播数据，而不管其是否需要接收，这样不仅数据的安全性得不到保障，而且会造成网络中信息的泛滥，浪费大量带宽资源。



单播和广播均不能以最小的网络开销实现数据的单点发送、多点接收，IP 组播（以下简称组播）技术的出现解决了这个问题。

组播指发送源将产生的单一 IP 数据包通过网络发送给一组特定接收者的网络传输方式。组播结合了单播和广播的优点，在进行点到多点传输时，发送源不需要关心接收者的数目，仅需要发送一份报文；路由器仅关心接口下是否有接收者，同样不需要关心接收者的数量，所以在路由器之间的链路上也仅传送一份报文。

和单播相比，组播减轻了发送源的负担，并且提高了链路的有效利用率。此外，发送源可以同时发送报文给多个接收者，可以满足低延时应用的需求。

和广播相比，组播方式下路由器仅在有接收者的接口复制报文，报文最终仅传递给接收者而非网络中的所有主机，可以节省大量网络带宽。另外，广播只能在同一网段中进行，而组播可以实现跨网段的传输。

利用组播技术可以方便地提供一些新的增值业务，包括在线直播、网络电视、远程教育、远程医疗、网络电台、实时视频会议等对带宽和数据交互的实时性要求较高的信息服务。

组播的优缺点

● 优点

- 增强效率，控制网络流量，减少服务器和CPU负载
- 优化性能，消除流量冗余
- 分布式应用，使多点传输成为可能

● 缺点

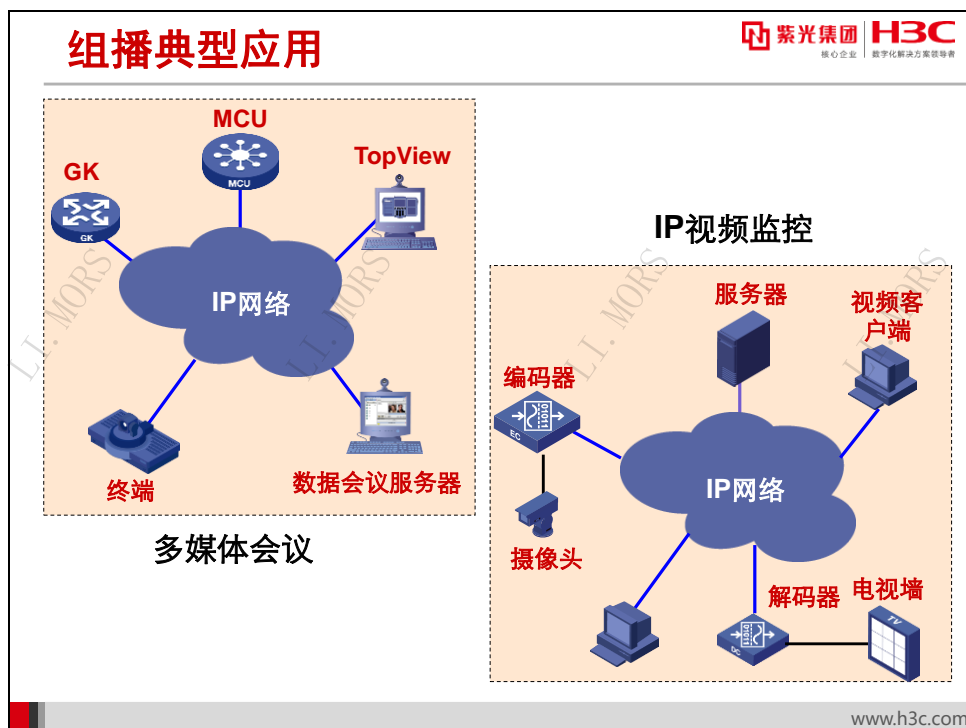
- 尽最大努力交付
- 无拥塞控制
- 数据包重复
- 数据包的无序交付

从组播和单播、广播的对比可以总结得出组播的优点：

- 组播可以增强报文发送效率，控制网络流量，减少服务器和 CPU 的负载；
- 组播可以优化网络性能，消除流量冗余；
- 组播可以适应分布式应用，当接收者数量变化时，网络流量的波动很平稳。

由于组播应用基于 UDP 而非 TCP，这就决定了组播应用存在 UDP 相应的缺点：

- 组播数据基于 Best Effort 发送，无法保证语音、视频等应用的优先传输，当报文丢失时，采用应用层的重传机制无法保证实时应用的低延时需求；
- 不提供拥塞控制机制，当网络出现拥塞时，无法为高优先级的应用保留带宽；
- 无法实现组播数据包重复检测，当网络拓扑发生变化时，接收者可能会收到重复的报文，需要应用层去剔除；
- 无法纠正组播数据包乱序到达的问题。



组播技术主要应用在多媒体会议、IP 视频监控、实时数据组播、游戏和仿真等方面。

多媒体会议是最早的组播应用。多媒体会议的工具最先在 UNIX 环境下被开发出来，这些工具允许通过组播实现多对多的多媒体会议。除多媒体工具外，还有基于 UNIX 的白板工具被开发出来，用于用户共享公共的电子白板，适合网络教学。

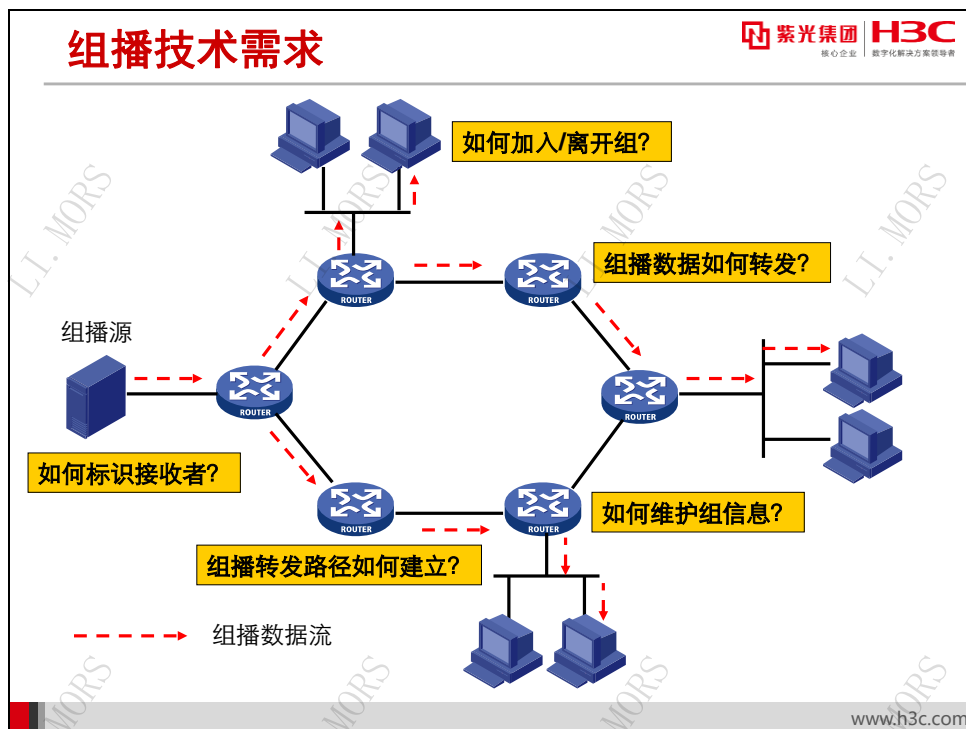
IP 视频监控是近年来发展迅猛的安防应用，IP 视频监控系统中摄像头发送模拟视音频信号给编码器，编码器进行模/数转换、编码压缩、IP 封装后将报文发送到 IP 网络。解码器收到 IP 报文后进行解封装、解码、解压缩、数/模转换，然后将模拟视音频流送到电视墙显示。视频客户端也可以接收 IP 报文进行软解码本地播放。

在 IP 视频监控中，编码器可以通过组播方式发送监控视频，组播方式可以极大的降低编码器端的负荷，减少网络中的冗余视音频流，并且可以满足实时监控的需求。

实时数据传送是组播很受欢迎的一个应用领域，例如股票、金融数据传送、实况足球播放、现场演唱会点播等。

组播还适合于网络游戏和仿真应用。如果采用单播方式，游戏玩家之间需要建立点到点的连接，对于 PC 机或服务器的处理能力而言是 N 平方数量级的负荷。当玩家数量上升，PC 机或服务器将不堪重负。采用组播方式，某一个玩家不需要和其他每一个玩家都建立连接，例如玩家 A 要发送数据给其他玩家，其他玩家仅需要加入组播组 A，则玩家 A 发送到组播组 A 的数据所有其他玩家都可以收到。

18.3 组播技术体系架构



组播的实现机制较单播复杂，要实现组播，首先需要解决如下几个问题：

- 组播的接收者是数目不定的一组接收者，无法像单播一样使用主机 IP 地址来进行标识，所以首先要解决如何在网络中标识一组接收者。
- 如果实现了对组的标识，还需要解决接收者如何加入和离开这个组，路由设备又如何维护组成员信息。
- 组播接收者可能分散在网络中的任何角落，那么组播源和组播接收者之间的转发路径基于什么模型，组播数据如何在路径上转发。
- 组播数据转发路径如何建立和维护。

上述技术需求通过组播架构中的一些重要机制来实现，包括：组播地址、组播组管理协议、组播分发树模型、组播转发机制和组播路由协议。

组播地址



- **组播地址范围**
→ 224.0.0.0 – 239.255.255.255
- **本地协议预留组播地址**
→ 224.0.0.0 – 224.0.1.255
- **本地管理组地址**
→ 239.0.0.0 – 239.255.255.255
- **用户组播地址**
→ 224.0.2.0 – 238.255.255.255
- **组播MAC地址:**
→ 以太网: 01-00-5e-xx-xx-xx

www.h3c.com

组播通信中使用组播地址来标识一组接收者，使用组播地址标识的接收者集合称为组播组。

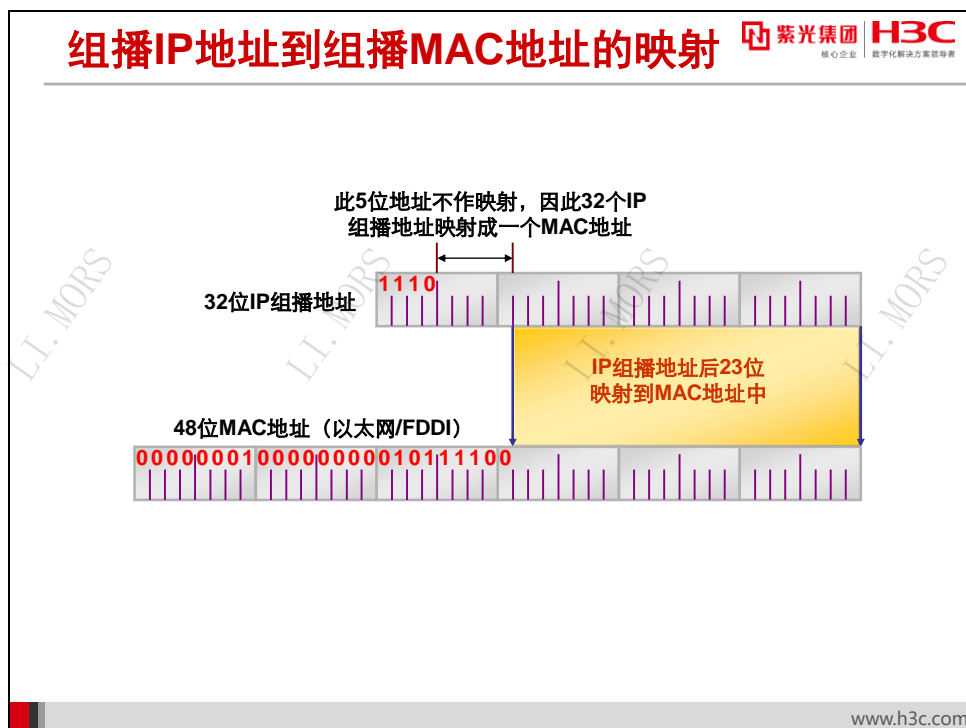
IANA (Internet Assigned Numbers Authority, 因特网编号分配委员会) 将 D 类地址空间分配给 IPv4 组播使用，地址范围为 224.0.0.0 到 239.255.255.255，组播地址的分类和具体含义如下：

- 224.0.0.0~224.0.1.255，协议预留组播地址，除 224.0.0.0 保留不做分配外，其它地址供路由协议、拓扑查找和协议维护等使用。
- 224.0.2.0~238.255.255.255，用户组地址，全网范围内有效。
- 239.0.0.0~239.255.255.255，本地管理组地址，仅在本地管理域内有效。

组播地址解决了 IP 报文在网络层寻址的问题，但通信最终还要依赖于数据链路层和物理层，因此和单播一样，组播也需要考虑数据在链路层如何寻址。

以太网传输单播 IP 报文的时候，目的 MAC 地址使用的是接收者的 MAC 地址。由于组播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，所以在链路层需要使用特定的组播 MAC 地址来标识一组接收者。

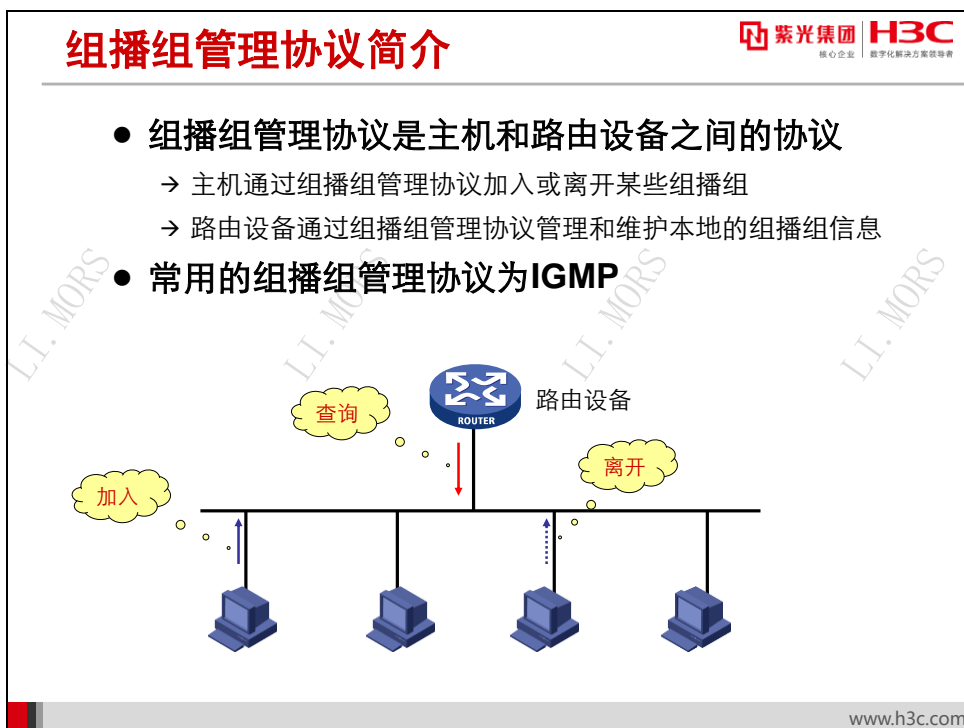
IANA 定义 IPv4 组播 MAC 地址格式为 01-00-5E-XX-XX-XX。



组播 MAC 地址中高 24 位固定为 0x01005E，第 25 位为 0，低 23 位来自于组播 IP 地址的低 23 位。

由于组播 IP 地址的高 4 位是 1110，代表组播标识，而低 28 位中只有 23 位被映射到组播 MAC 地址，这样组播地址中就有 5 位信息丢失。于是，就有 32 个组播 IP 地址映射到了同一个组播 MAC 地址上，因此在二层处理过程中，设备可能要接收一些本组播组以外的组播数据，而这些多余的组播数据就需要设备的上层进行过滤了。

例如组播 IP 地址为 228.128.128.128，其对应的组播 MAC 地址为 01-00-5E-00-80-80。组播 IP 地址为 229.128.128.128，其对应的组播 MAC 地址仍然为 01-00-5E-00-80-80。



解决了如何标识组播组的问题，还需要考虑接收者怎样加入组播组，如何维护组播组以及由谁来维护组播组等问题。在组播架构中使用组播组管理协议来实现上述需求。

组播组管理协议是运行于主机和路由器之间的协议。主机通过组播组管理协议通知路由器加入或离开某个组播组；路由器通过组播组管理协议响应主机加入请求，建立相应的组播表项，并通过查询消息维护组播组信息。

常用的组播组管理协议为 IGMP（Internet Group Management Protocol，因特网组管理协议）。

组播分发树模型简介

紫光集团 H3C
核心企业 数字化转型领导者

- 组播分发树是组播数据的转发路径
- 根据树根位置的不同，组播分发树模型分为
 - 最短路径树模型
 - 共享树模型

www.h3c.com

在单播通信中，发送源和接收者之间的路径是点到点的一条线，起点为发送源，目的地为一个接收者，该路径由单播路由协议建立。

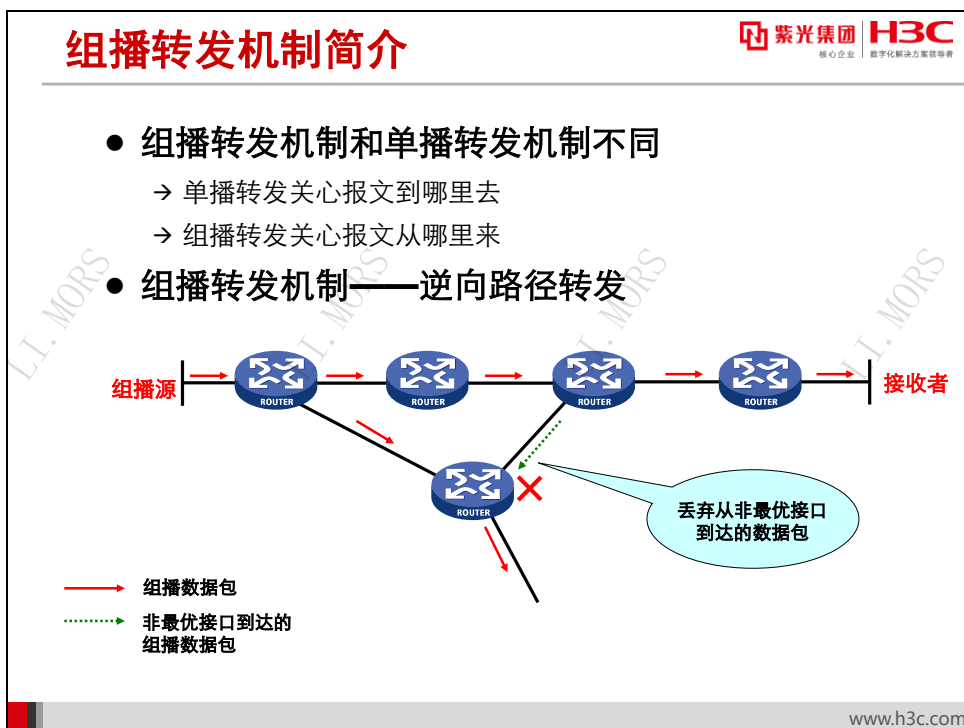
在组播中，目的地是数目不定的一组接收者，这就决定了组播报文转发路径和单播不同。

组播转发路径基于树形结构，称为组播分发树，接收者位于树形结构的叶子处，组播分发树由组播路由协议建立。

根据组播分发树树根位置的不同，组播分发树模型分为最短路径树（Shortest Path Tree，简称 SPT）模型和共享树（Rendezvous Point Tree，简称 RPT）模型。

SPT 树根为发送源，因此 SPT 也称为“源树”，其从发送源到每一个接收者的路径都是最优的。

RPT 模型树根为网络中的某一台设备，称为汇聚点，从发送源到接收者的组播数据必须首先经过汇聚点，然后再由汇聚点发送到每一个接收者，因此 RPT 模型中，从发送源到接收者之间的路径不一定是最优路径。



在单播通信中，IP 报文转发的依据是报文的目的 IP 地址，目的 IP 地址在网络中唯一的标识了一台主机，网络中的路由器收到单播 IP 报文后只需要通过目的 IP 地址查找单播路由表，确定报文对应的下一跳地址，得出报文的出接口，然后将报文从该出接口发出即可。沿途的每一台路由器都进行同样的操作，即可保证将报文准确无误的送到目的地，并且通过选择最优路径，可以消除网络中的路径环路。

组播不能简单的通过查看报文的目的 IP 地址就得到报文传送的最优路径以及对应的出接口，因为组播报文的目的地址不是一台明确的主机，有可能路由器每一个接口都存在接收者。

组播采用逆向路径转发的方式，判断组播报文是否从指向组播分发树树根的最短路径到达，只有来自于最优路径的组播报文才会被转发，来自于非最优路径的组播报文会被丢弃。沿途每一台组播路由器都进行同样的操作，可以保证报文从组播源到组播组中的每一个接收者所经过的路径都是最优的，并且可以消除组播路径环路。

综上所述，单播转发的时候主要关心报文往哪里去，而组播转发的时候主要关心报文从哪里来。

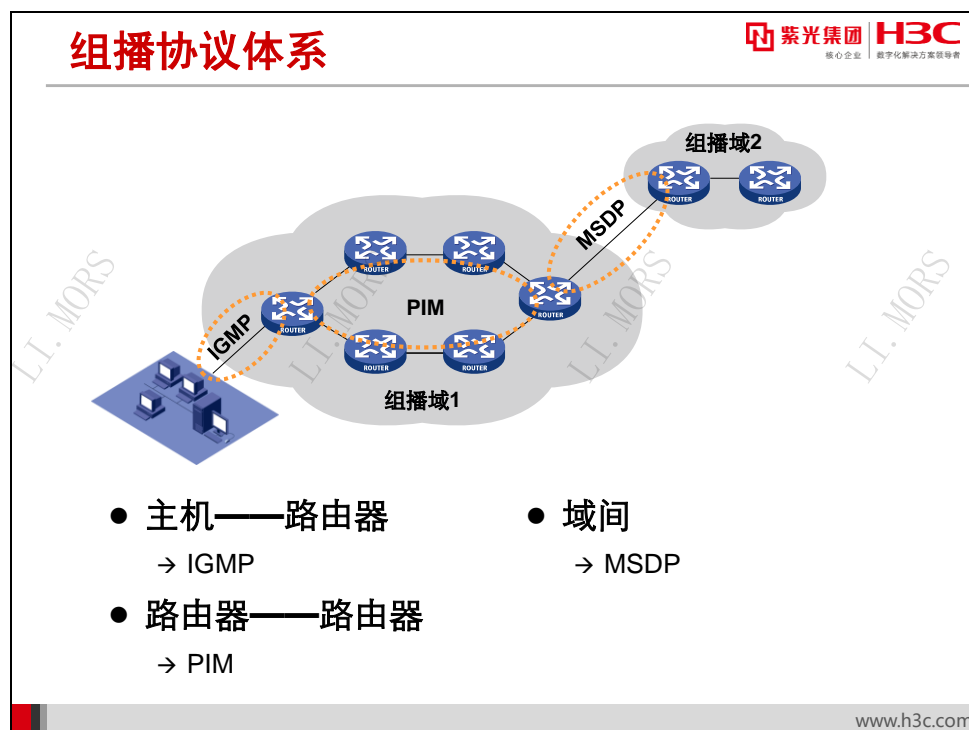
组播路由简介

- 组播路由协议和单播路由协议一样，用于建立数据转发的路径
- 根据作用范围组播路由协议可以分为域内组播路由协议和域间组播路由协议
 - 域内组播路由协议主要包括DVMRP、MOSPF和PIM
 - 域间组播路由协议主要包括MSDP、MBGP
- 域内组播路由协议根据建立的组播分发树的不同可以分为基于SPT的组播路由协议和基于RPT的组播路由协议
 - 基于SPT的组播路由协议包括PIM DM、DVMRP、MOSPF
 - 基于RPT的组播路由协议包括PIM SM

组播路由协议运行在三层组播设备之间，用于建立和维护组播路由，并正确、高效地转发组播数据包。组播路由协议建立了从一个数据源端到多个接收端的无环（loop-free）数据传输路径，即组播分发树。

组播路由协议根据作用范围可以分为域内组播路由协议和域间组播路由协议，其中域内组播路由协议主要包括 DVMRP（Distance Vector Multicast Routing Protocol，距离矢量组播路由协议）、MOSPF（Multicast Extensions to OSPF，组播 OSPF 协议）和 PIM（Protocol Independent Multicast，协议无关组播），域间组播路由协议主要包括 MSDP（Multicast Source Discovery Protocol，组播源发现协议）、MBGP（Multicast BGP，组播 BGP）。

域内组播路由协议根据建立的组播分发树的不同可以分为基于 SPT 的组播路由协议和基于 RPT 的组播路由协议，其中基于 SPT 的组播路由协议包括 PIM DM（Protocol Independent Multicast-Dense Mode，协议无关组播—密集模式）、DVMRP、MOSPF，基于 RPT 的组播路由协议包括 PIM SM（Protocol Independent Multicast-Sparse Mode，协议无关组播—稀疏模式）。



组播协议主要包含主机和路由器之间的协议，路由器和路由器之间的协议，以及组播域之间的协议。

主机和路由器之间的协议即组播组管理协议，IPv4 中通常使用 IGMP。通过 IGMP，路由器可以了解在本地网段中，哪些组播组存在接收者，并维护组成员信息。

路由器和路由器之间的协议为组播路由协议，常用的组播路由协议为 PIM。通过 PIM，可以将组成员信息扩散到整个网络，从而建立送发送源到接收者之间的组播分发树。

组播域的边界通常为单播域的边界，由于域之间组播路由信息有可能无法直接交互，导致接收者无法跨域接收组播数据，此时需要在域之间运行域间组播路由协议，解决域间组播通信的问题，常用的域间组播路由协议为 MSDP。

18.4 组播模型

组播模型分类

● 根据接收者对组播源处理方式的不同，组播模型分为以下两类

- ASM (Any-Source Multicast, 任意信源组播) 模型
- SSM (Source-Specific Multicast, 指定信源组播) 模型



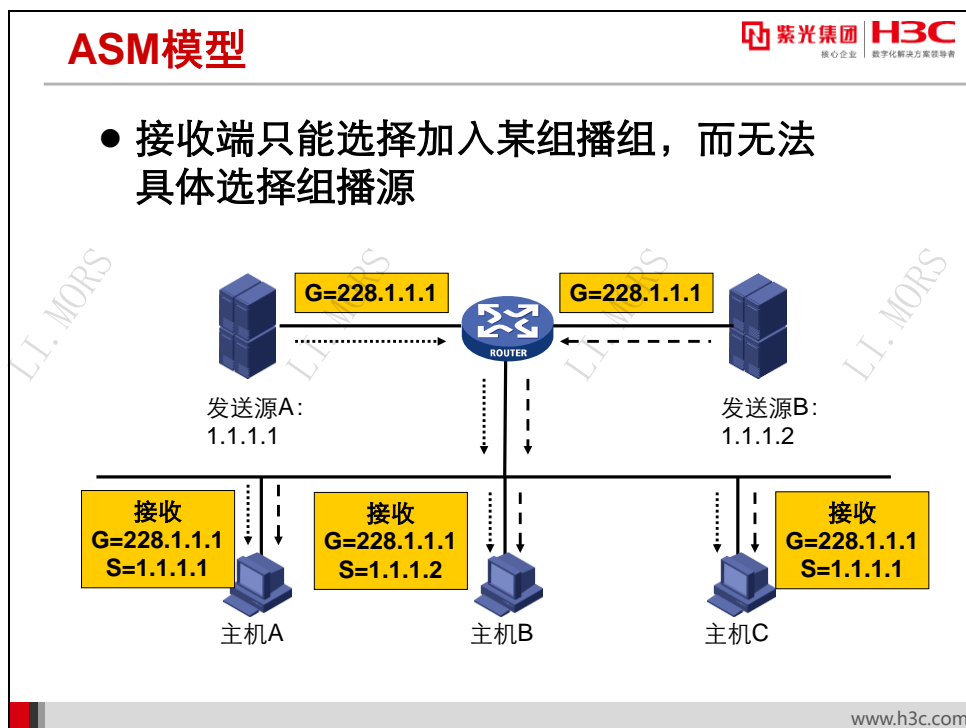
紫光集团 H3C
核心企业 数字化解决方案领导者

www.h3c.com

根据接收者对组播源处理方式的不同，组播模型分为以下两类：

- **ASM (Any-Source Multicast, 任意信源组播) 模型：**在 ASM 模型中，组播接收者无法指定组播源，任意组播源发送到同一个组播组的数据，都会被网络设备传送到组播接收者。
- **SSM (Source-Specific Multicast, 指定信源组播) 模型：**在现实生活中，用户可能只对某些组播源发送的组播信息感兴趣，而不愿接收其它源发送的信息。SSM 模型为用户提供了一种能够在客户端指定组播源的传输服务。

SSM 模型与 ASM 模型的根本区别在于：SSM 模型中的接收者已经通过其它手段预先知道了组播源的具体位置。

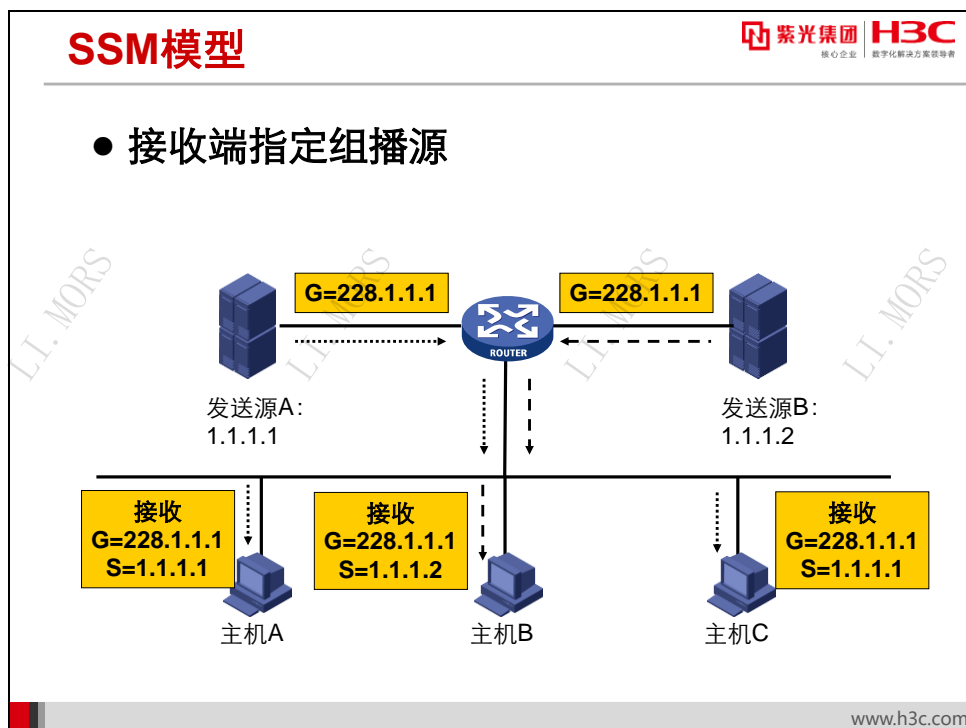


ASM 模型中，当接收者通过组播组管理协议加入某组播组时，并不区分组播数据的发送源。

上图中，发送源 A 发送组播数据，源地址为 1.1.1.1，目的组播地址为 228.1.1.1；发送源 B 发送组播数据，源地址为 1.1.1.2，目的组播地址也为 228.1.1.1。

主机 A、C 希望接收发送源 A 发送的组播数据，主机 B 希望接收发送源 B 发送的组播数据，因此主机 A、B、C 通过组播组管理协议加入了组播组 228.1.1.1。但由于 ASM 模型无法指定发送源，此后，路由器会将发送源 A 和发送源 B 发送的组播数据均发送到主机 A、B、C。

ASM 模型无法满足主机接收指定发送源发送的组播数据，如果主机收到多份来自不同发送源的相同组播组的数据，需要上层应用进行区分。



和 ASM 模型不同，在 SSM 模型中接收者指定组播组的同时还可以指定发送源。

例如，主机 A 和主机 C 要求接收发送源 A 发送的组播地址为 228.1.1.1 的组播数据，则仅有源地址为 1.1.1.1，目的地址为 228.1.1.1 的组播数据会发送到主机 A 和主机 C，发送源 B 发送的组播数据虽然组播地址也为 228.1.1.1，但是由于源地址不是主机 A 和主机 C 所指定的，因此发送源 B 发送的组播报文不会被主机 A 和主机 C 接收。

同样，主机 B 指定接收源地址为 1.1.1.2，目的地址为 228.1.1.1 的组播数据，则仅有发送源 B 发送的组播报文会被主机 B 接收。

18.5 本章总结

本章总结

- 引入组播，介绍组播概念并和单播、广播进行了对比
- 介绍组播的优缺点和主要应用
- 对组播的体系架构进行介绍
- 介绍了组播的模型

18.6 习题和解答

18.6.1 习题

1. 关于组播和单播、广播的对比，正确的有（ ）
 - A. 和单播相比组播可以减轻发送源的负担
 - B. 和广播相比组播可以减轻发送源的负担
 - C. 和单播相比组播可以减少链路负载
 - D. 和广播相比组播可以提升链路使用率
2. 228.129.129.129，对应的组播 MAC 地址为（ ）
 - A. 01-00-5E-00-01-01
 - B. 00-01-5E-11-81-81
 - C. 01-00-5E-01-81-81
 - D. 00-01-5E-00-81-81
3. 关于组播协议体系架构，说法正确的有（ ）
 - A. 路由器和主机之间通常运行组播路由协议
 - B. 路由器和路由器之间通常运行组播组管理协议
 - C. 主机和路由器之间通常运行组播组管理协议
 - D. 常用的域间组播路由协议为 MSDP
4. 关于组播体系架构，说法正确的有（ ）
 - A. 组播组管理协议负责组播数据的转发
 - B. 组播分发树由组播路由协议建立
 - C. 组播组管理协议用于主机加入和离开组播组
 - D. 组播组管理协议用于路由器维护组播组信息
5. 关于组播模型，正确的有（ ）
 - A. ASM 模型中，接收者无法通过组播组管理协议指定接收某组播源发送的组播数据
 - B. ASM 模型中，如果有多个组播源发送相同组播组数据，则会导致接收者收到多份组播数据，需要应用层进行区分
 - C. SSM 模型中，接收者指定组播组的同时还可以指定发送源
 - D. SSM 模型中的接收者已经通过其它手段预先知道了组播源的具体位置

18.6.2 习题答案

1. ACD 2. C 3. CD 4. BCD 5. ABCD

第19章 组播组管理协议

组播组管理协议运行在主机和与其直接相连的三层组播设备之间。常用的组播组管理协议为 IGMP，该协议规定了主机与三层组播设备之间建立和维护组播组成员关系的机制。

在实际组播应用中，通常还需要二层组播协议来配合组播组管理协议的工作。二层组播协议包含 IGMP Snooping（IGMP 窥探）和组播 VLAN。通过 IGMP Snooping 可以解决组播报文在链路层广播的问题，通过组播 VLAN 可以节省网络带宽，降低三层设备的负担。

本章介绍了 IGMP 各版本协议处理机制，并对二层组播协议 IGMP Snooping 和组播 VLAN 进行介绍。

19.1 本章目标

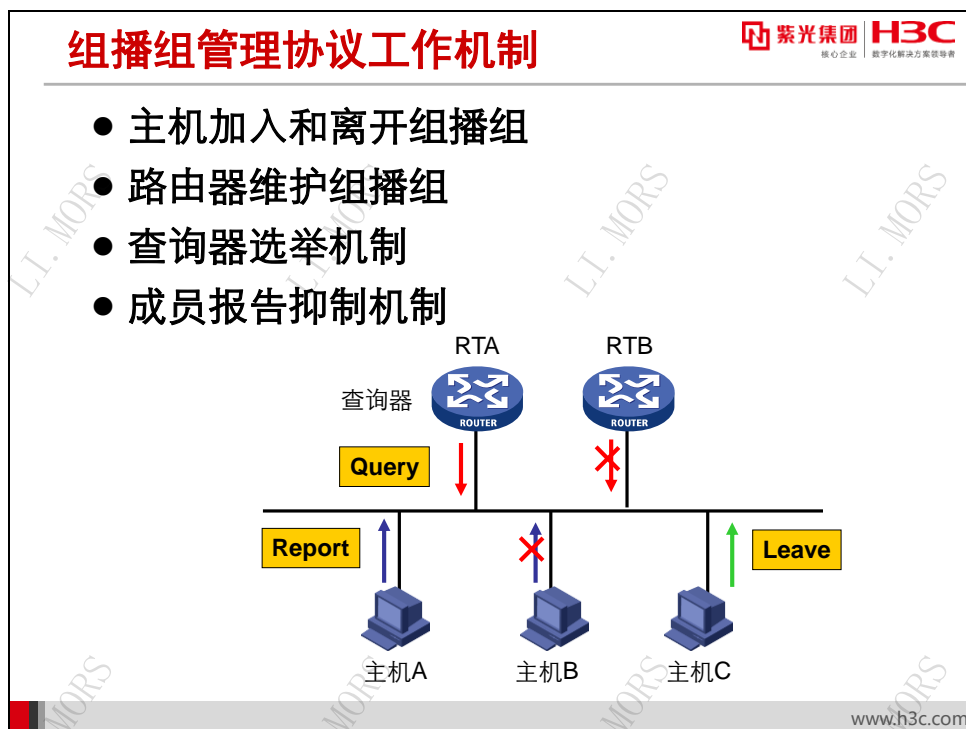
课程目标

● 学习完本课程，您应该能够：

- 掌握IGMP协议原理
- 掌握IGMP各版本间的互操作
- 掌握IGMP Snooping和组播VLAN



19.2 组播组管理协议概述



组播组管理协议是运行于主机和路由器之间的协议。组播组管理协议的工作机制包含：成员加入和离开组播组、路由器维护组播组、查询器选举机制以及成员报告抑制机制。

组播路由器周期性的发送查询报文，询问网段上是否有组播接收者。如果网段上有主机希望接收某个组播组的数据，则主机会向路由器回复成员报告报文，报告自己想要接收哪个组播组的数据。路由器收到主机发送的成员报告报文后，会为主机请求加入的组播组建立一个表项，表示该组播组在该网段有成员。

当主机需要接收某个组播组的数据时，也可以不必等待路由器发送查询报文，而直接发送成员报告报文请求加入某个组播组。同样，路由器收到成员报告报文后，会更新组播组信息。

当主机不再需要接收某个组播组的数据时，主机可以发送离开消息通知路由器离开该组播组，路由器会通过查询机制判断网段上该组播组是否有其他成员存在。如果还有其他成员存在，则路由器继续维护该组播组，如果没有成员存在，路由器会将该组播组信息删除，此后，路由器不再将该组播组的数据转发到该网段。

当局域网网段上有多台路由器时，网段上可能会出现重复的查询报文，也可能会有重复的组播数据发送到该网段，此时需要在这些路由器之间选举某台路由器负责该网段组播数据的转发以及组播组的维护，这台路由器即为该网段的查询器。查询器选举结束后，网段上只有查询器会发送查询报文，其他路由器仅在查询器发生故障，且经过新一轮的选举成为新的查询器后，才会发送查询报文。

对于某个网段，路由器仅需要知道是否有组播组成员存在，只要某个组播组有成员，路由器就会为该组播组维护相应信息，并将组播数据转发到该网段。路由器不需要确切的知道组播组在该网段有哪些组成员。因此如果一个网段有多台主机想要接收同一个组播组的数据，只需要有一台主机发送成员报告报文即可，这就需要一种机制来抑制网段上同一个组播组的其他主机发送成员报告报文，这个机制就是成员报告抑制机制。

当主机需要发送成员报告报文时，首先会等待一段随机的延时。如果在这段延时期间，主机收到了其他主机发送的成员报告报文，并且和自身要求加入的组播组信息相同，则该主机会取消自身成员报告报文的发送。如果在延时期间没有收到要求加入相同组播组的成员报告报文，则主机会正常发送成员报告报文。通过成员报告抑制机制，可以减少网段上重复成员报告报文的发送。

组播组管理协议不负责通知主机加入哪些组播组，这是由主机上层应用决定的，组播组管理协议也不负责组播报文在路由器之间的转发，这由组播路由协议实现。

常用组播组管理协议

紫光集团 H3C
核心企业 数字化转型领导者

- IPv4使用IGMP，包含三个版本
 - IGMPv1在RFC1112中定义
 - IGMPv2在RFC2236中定义
 - IGMPv3在RFC3376中定义

www.h3c.com

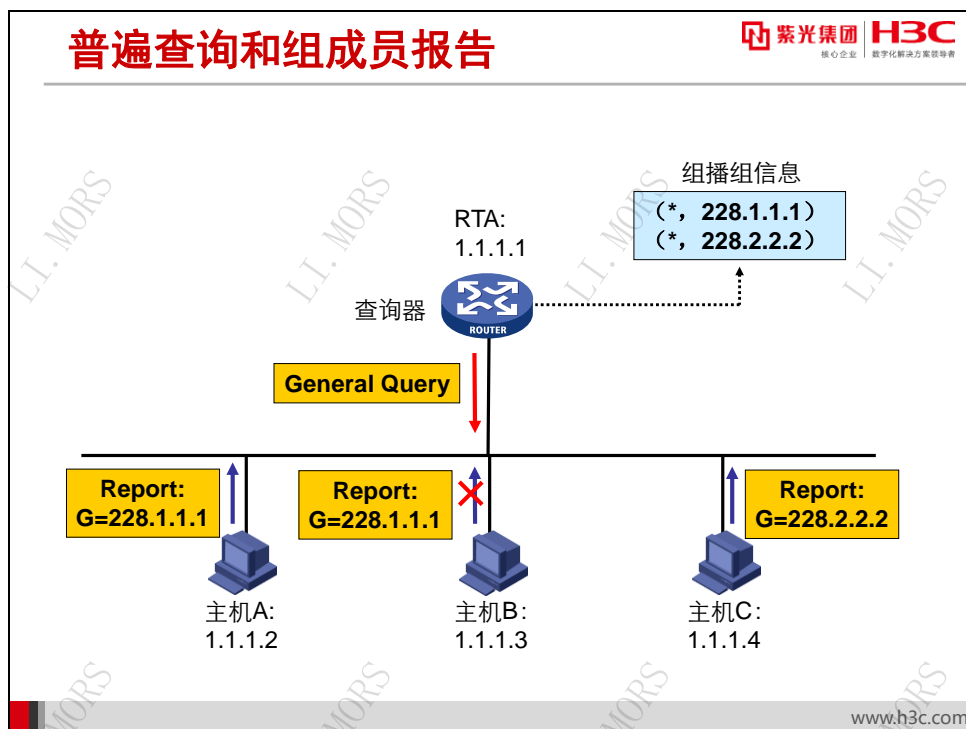
IPv4 常用的组播组管理协议为 IGMP。到目前为止，IGMP 有三个版本：

- IGMPv1 在 RFC 1112 中定义
- IGMPv2 在 RFC 2236 中定义
- IGMPv3 在 RFC 3376 中定义

IGMPv1 定义了基本的查询和成员报告过程，IGMPv2 在此基础上添加了组成员快速离开机制和查询器选举机制，IGMPv3 又在 IGMPv2 的基础上增加了指定组播源的功能。

所有版本的 IGMP 都支持 ASM 模型；IGMPv3 可以支持 SSM 模型。

19.3 IGMPv2



IGMP 查询器 RTA 周期性地以组播方式向本地网段内的所有主机发送 IGMP General Query 报文，目的地址为 224.0.0.1，TTL 为 1。

收到 General Query 报文后，由于主机 A 和主机 B 希望接收组播组 228.1.1.1 的数据，所以会对 General Query 报文进行响应。为了防止 Membership Report 报文发生冲突，主机 A 和主机 B 会等待一段时间后，才会发送 Membership Report 报文，延迟时间为[0, Max Reps Time]之间的一个随机值，Max Reps Time 为最大响应延时，由路由器在 General Query 报文中发布。

假设主机 B 延迟时间大于主机 A 的延迟时间，则主机 A 会首先发送 Membership Report 报文，目的地址为主机 A 要加入的组播组地址 228.1.1.1。主机 B 收到主机 A 发送的 Membership Report 报文后，经过判断发现是关于同一个组 228.1.1.1 的 Membership Report 报文，则主机 B 会取消自己 Membership Report 报文的发送，从而减少了本地网段的信息流量。

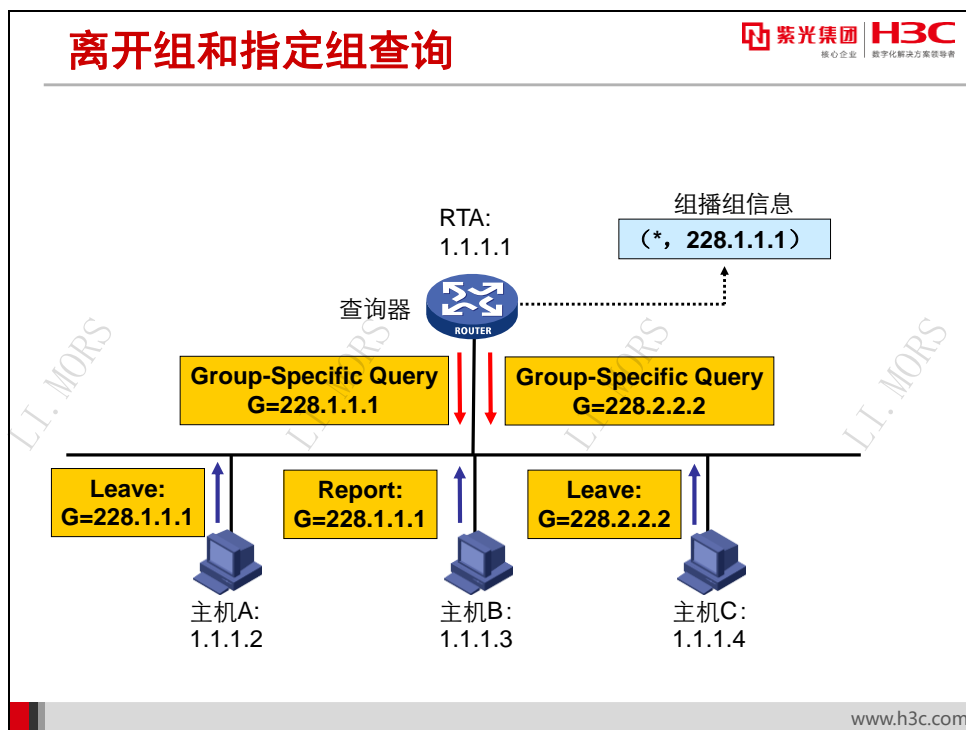
与此同时，由于主机 C 关注的是组播组 228.2.2.2，所以它仍将会发送关于组播组 228.2.2.2 的 Membership Report 报文，以向 RTA 宣告其属于组播组 228.2.2.2。

当然，如果主机有新的组播接收需求，可以主动向 RTA 发送 Membership Report 报文，而不必等待 RTA 周期性的发送 General Query 报文。从而节省了主机加入某组播组的时间。

经过以上的查询和响应过程，RTA 了解到本地网段中有组 228.1.1.1 和组 228.2.2.2 的成员，于是创建组播转发表项（*, 228.1.1.1）和（*, 228.2.2.2），作为组播数据的转发依据，其中*代表任意组播源。

组播转发表项还包含组播报文的入接口和组播报文的出接口列表等信息，用于指导组播报文在本地路由器的转发，此时 RTA 会将连接主机 A、B、C 的接口加入到对应组播组的出接口列表中。

当由组播源发往 228.1.1.1 或 228.2.2.2 的组播报文到达 RTA 时，由于 RTA 上存在（*, 228.1.1.1）和（*, 228.2.2.2）组播转发项，于是将该组播报文会根据出接口信息转发到本地网段，接收者主机便能收到该组播数据了。



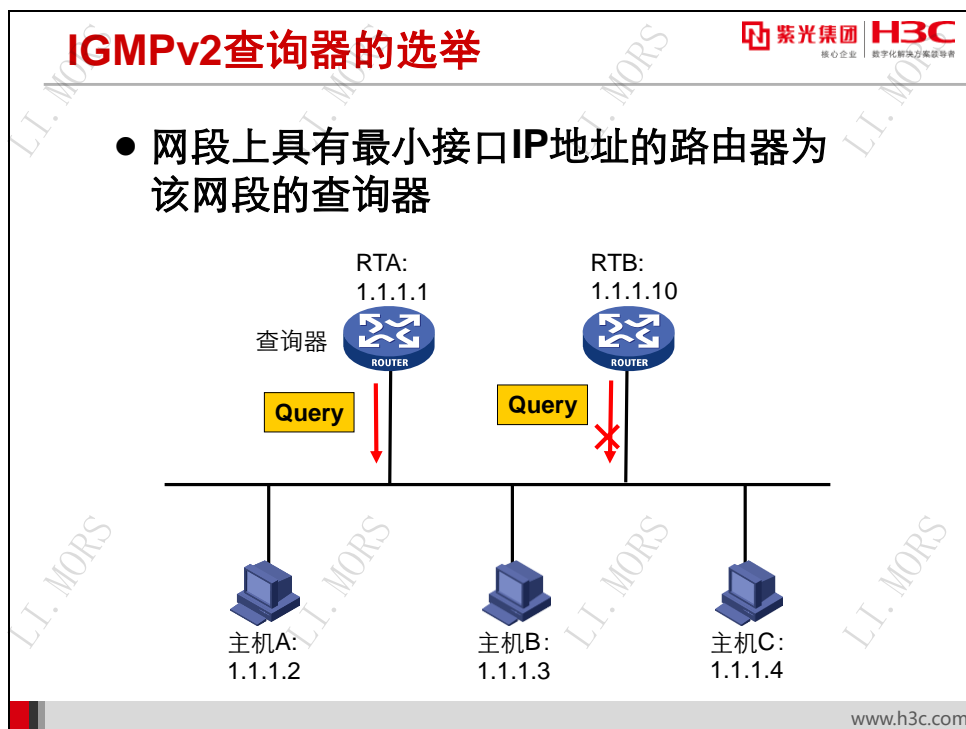
在 IGMPv2 中，当一个主机离开某组播组时会向本地网段内的所有组播路由器发送 Leave Group 报文，目的地址为 224.0.0.2，报文中包含要离开的组播组地址信息。

图中主机 A 和主机 C 要分别离开组播组 228.1.1.1 和 228.2.2.2，则会向 RTA 发送 Leave Group 报文。

当 IGMP 查询器 RTA 收到 Leave Group 报文后，会发送另外一种查询报文——Group-Specific Query 报文，用于确认该组播组在网段内是否还有成员存在。Group-Specific Query 报文的地址为所要查询的指定组播组地址。

由于主机 B 仍然需要接收组播组 228.1.1.1 的数据，所以主机 B 收到 Group-Specific Query 报文后，会在[0, Max Reps Time]时间内回复 Membership Report 报文。

RTA 在最大响应时间内收到了主机 B 的 Membership Report 报文，认为组播组 228.1.1.1 仍然有接收者，所以继续维护该组播组。由于主机 C 离开后，网段内不再有组播组 228.2.2.2 的成员，所以在最大响应时间内 RTA 无法收到关于组播组 228.2.2.2 的 Membership Report 报文，RTA 将删除组播组 228.2.2.2 的表项，此后不会再将目的地址为 228.2.2.2 的报文发送到该网段。



在 IGMPv1 中没有定义查询器选举机制，当某共享网段上存在多个组播路由器时，需要通过组播路由协议（如 PIM）选举的 DR（Designate Router，指定路由器）充当查询器。

在 IGMPv2 中，增加了独立的查询器选举机制，其选举过程如下：

- 所有 IGMPv2 路由器在初始时都认为自己是查询器，并向本地网段内的所有主机和路由器发送 IGMP General Query 报文，目的地址为 224.0.0.1。
- 本地网段中的其它 IGMPv2 路由器在收到该报文后，将报文的源 IP 地址与自己的接口地址作比较。通过比较，IP 地址最小的路由器将成为本网段的查询器，其它路由器成为非查询器（Non-Querier）。
- 所有非查询器上都会启动一个定时器，在该定时器超时前，如果收到了来自查询器的 IGMP 查询报文，则重置该定时器；否则，就认为当前查询器失效，从而发起新的查询器选举过程。

IGMPv2协议报文

紫光集团 H3C
核心企业 数字化转型领导者

0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2

Type	Max Reps Time	Checksum
Group Address		

- **Type**为IGMP报文类型，包含**Membership Query**报文、版本1的**Membership Report**报文、版本2的**Membership Report**报文和**Leave Group**报文
- **Max Reps Time**为最大响应时间，只有**Membership Query**报文使用该字段
- 不同报文组地址有不同含义
 - General Query报文的组地址字段值为0
 - Group-Specific Query报文中，组地址为被查询的组播组地址
 - Membership Report报文和Leave Group报文中，组地址为主机要加入或离开的组播组地址

www.h3c.com

IGMPv2 是在 IGMPv1 的基础之上发展而来的，其报文格式和 IGMPv1 报文格式基本相同。不同之处在于 IGMPv2 将 IGMPv1 中的 4 位 Version 字段与 4 位 Type 字段整合成为一个 8 位的 Type 字段。将 IGMPv1 预留没有使用的 8 位 Unused 字段设置为 8 位 Max Reps Time 字段。

IGMPv2 的 Type 字段定义了四种报文类型：

- Membership Query 报文，Type 字段值为 0x11。
- 版本 1 的 Membership Report 报文，Type 字段值为 0x12，该 Report 报文用于和 IGMPv1 兼容。
- 版本 2 的 Membership Report 报文，Type 字段值为 0x16。
- Leave Group 报文，Type 字段值为 0x17。

Membership Query 报文又包含两个类型：General Query 和 Group-Specific Query。这两种查询报文由组地址字段进行区分：General Query 报文的组地址字段为 0，而 Group-Specific Query 报文的组地址字段为要查询的组播组的地址。

最大响应时间（Max Resp Time）字段仅在 Membership Query 报文中使用。该字段规定了主机在发送一个 Membership Report 报文时最大的延时时间，单位为 0.1 秒，默认值为 100（即 10 秒）。在 Membership Report 报文和 Leave Group 报文中，会由主机设置为 0。

校验和（Checksum）字段是 IGMP 报文长度的 16 位检测。

组地址（Group Address）字段在不同的报文类型中有不同的含义。在 General Query 报文中该字段设置为 0；在 Group-Specific Query 报文中该字段为被查询的组播组地址；在

Membership Report 报文和 **Leave Group** 报文中，组地址字段为主机想要加入或离开的组播组地址。

19.4 IGMPv3

IGMPv3概述

- **IGMPv3增加了对源过滤的支持**
→ IGMPv3主机不仅可以接收某个组播组的数据，还可以选择接收或拒绝某些源发送到这个组播组的组播数据
- **IGMPv3定义了新的报文类型和格式**
- **IGMPv3 Report报文目的组播地址变为 224.0.0.22**
- **IGMPv3取消成员报告抑制机制**

www.h3c.com

IGMPv3 在兼容和继承 IGMPv1 和 IGMPv2 的基础上，进一步增强了主机的控制能力，并增强了 Membership Query 报文和 Membership Report 报文的功能。

IGMPv3 增加了对组播源过滤的支持，IGMPv3 主机不仅可以接收某个组播组的数据，还可以根据喜好选择接收或拒绝某些源发送到这个组播组的数据。

例如，网络中有两个频道都在播放 NBA 比赛，频道 1 的节目用组播流（1.1.1.1，228.1.1.1）表示，其中单播地址 1.1.1.1 代表频道 1 组播源，组播地址 228.1.1.1 代表 NBA 比赛节目。同样频道 2 用组播流（2.2.2.2，228.1.1.1）表示。如果网络中的设备仅支持 IGMPv1/v2，就无法做到只接收频道 1 的节目而不接收频道 2 的节目。因为 IGMPv1/v2 无法区分组播源，只能区分组播组。而如果用户设备支持 IGMPv3 协议，就可以通知路由器只接收组播源为 1.1.1.1 的组播流，而不想接收组播源为 2.2.2.2 的组播流，这样路由器就可以只把频道 1 的 NBA 比赛转发给用户。

IGMPv3 增加了对特定源组查询的支持，在 Group-and-Source-Specific Query 报文中，既携带组地址，还携带一个或多个源地址。IGMPv3 取消了 Leave Group 报文类型，通过在 Membership Report 报文中申明不再接收任何源发送给某组播组的数据，即可实现离开这个组播组的功能。Membership Report 报文的地址使用组播地址 224.0.0.22，不再使用具体组播组的地址。

IGMPv3 中一个 Membership Report 报文可以携带多个源组信息，不同于 IGMPv1/v2 仅能包含一个组信息，因而大量减少了 Membership Report 报文的数量，不再需要成员报告抑

制机制。取消成员报告抑制机制后，IGMPv3 主机不需要对收到的 Membership Report 报文进行解析，可以大量减少主机的工作量。

IGMPv3过滤模式和源列表

紫光集团 H3C
核心企业 数字化解决方案领导者

- **IGMPv3主机为接口上每一个组播组都维护一个表项信息，其格式为：**
 - （组地址，过滤模式，源列表）
- **过滤模式包含INCLUDE和EXCLUDE两种类型**
 - INCLUDE模式表示只接收来自于在源列表中列出的组播源发送的组播数据包
 - EXCLUDE模式表示只接收来自于不在源列表中列出的组播源发送的组播数据包
- **源列表包含0个或多个IP单播地址，通常用集合形式来表示**

www.h3c.com

IGMPv3 主机为接口上每一个组播组都维护一个表项信息，其格式为（组地址，过滤模式，源列表）。

组播组的过滤模式包含 INCLUDE 和 EXCLUDE 两种类型：

- INCLUDE 模式表示只接收来自于在源列表中列出的组播源发送的组播数据。
- EXCLUDE 模式表示只接收来自于不在源列表中列出的组播源发送的组播数据。

源列表包含 0 个或多个 IP 单播地址，通常用集合形式来表示。通常使用 INCLUDE（S，G）表示接收来自源 S 的组播地址为 G 的数据，使用 EXCLUDE（S，G）表示不希望接收来自源 S 的组播地址为 G 的数据。

例如，主机 A 接收来自于组播源 1.1.1.1 发送的目的地址为 228.1.1.1 的组播流，则主机 A 维护的组播组 228.1.1.1 的表项为 INCLUDE（1.1.1.1，228.1.1.1）。

如果主机 B 不希望接收来自于组播源 2.2.2.2 发送的目的地址为 228.1.1.1 的组播流，则主机 B 维护的组播组 228.1.1.1 的表项为 EXCLUDE（2.2.2.2，228.1.1.1）。

IGMPv3主机侧维护的组状态

紫光集团 H3C
核心企业 数字化转型领导者

- **IGMPv3组播组有当前状态、过滤模式改变以及源列表改变三种状态，对应的组记录也有三种类型**
 - 当前状态记录
 - 过滤模式变化记录
 - 源列表变化记录
- **当主机接口维护的组状态发生变化时，会主动发送组记录类型为过滤模式变化或源列表变化的 Membership Report 报文**
- **当接收到 Membership Query 报文时，会响应组记录类型为当前状态的 Membership Report 报文**

www.h3c.com

IGMPv3 主机还为接口上每一个组播组维护状态信息。当收到路由器发送的 Membership Query 报文时，主机向路由器回应组播组的当前状态，当组播组的状态改变时主机主动向路由器发送 Membership Report 报文通知组播组的状态发生了变化。

IGMPv3 组播组有当前状态、过滤模式改变状态、源列表改变状态三种状态，这些状态信息在 Membership Report 报文的组记录字段（Group Record）中表示：

- **当前状态记录：**用于响应接口上收到的 Membership Query 报文，向路由器报告组播组的当前状态。组播组的当前状态包括 MODE_IS_INCLUDE 和 MODE_IS_EXCLUDE。同一时刻一个组播组只能处于一种当前状态。

MODE_IS_INCLUDE (S, G) 表示接口对于组播组 G 的过滤模式是 INCLUDE，简写为 IS_IN (S, G)。源列表 S 表示对于该组播组，主机所有感兴趣的组播源。

MODE_IS_EXCLUDE (S, G) 表示接口对于组播组 G 的过滤模式是 EXCLUDE，简写为 IS_EX (S, G)。源列表 S 表示对于该组播组，主机所有不希望接收的组播源。

- **过滤模式改变记录：**当接口维护的某个组播组的过滤模式发生变化时，主机会主动向路由器发送包含过滤模式改变记录的 Membership Report 报文。过滤模式改变记录包括 CHANGE_TO_INCLUDE_MODE 和 CHANGE_TO_EXCLUDE_MODE 两类。

CHANGE_TO_INCLUDE_MODE (S, G) 表示对于组播组 G 而言，接口的过滤模式已经从 EXCLUDE 模式变为 INCLUDE 模式。该组记录简写为 TO_IN (S, G)。

CHANGE_TO_EXCLUDE_MODE (S, G) 表示对于组播组 G 而言, 接口的过滤模式已经从 INCLUDE 模式变为 EXCLUDE 模式。该组记录简写为 TO_EX (S, G)。

- **源列表变化记录:** 当接口维护的某个组播组的源列表发生变化时, 主机会主动向路由器发送包含源列表变化记录的 Membership Report 报文。源列表变化记录包括 ALLOW_NEW_SOURCES 和 BLOCK_OLD_SOURCES 两类。

ALLOW_NEW_SOURCES (S, G) 表示对于组播组 G 而言, 主机希望接收一些新的组播源 S 发送的组播数据。如果当前过滤模式为 INCLUDE, S 将被添加到源列表中; 反之, 如果当前过滤模式为 EXCLUDE, S 将被从源列表中删除。该组记录简写为 ALLOW (S, G)。

BLOCK_OLD_SOURCES (S) 表示对于组播组 G 而言, 主机不再希望接收 S 发送的组播数据。如果当前过滤模式为 INCLUDE, 则 S 被从源列表中删除; 反之, 如果当前过滤模式为 EXCLUDE, 则 S 将被添加到源列表中。该组记录简写为 BLOCK (S, G)。

如果源列表发生了两种变化, 一种是允许了新的组播源, 另一种是阻塞旧的组播源, 则主机需要为同一个组播地址发送两个组记录, 一个是 ALLOW(S1, G), 另一个是 BLOCK (S2, G)。

IGMPv3 路由器侧维护的组状态

紫光集团 H3C
核心企业 数字化解决方案领导者

- 路由器为接口上的每一个组播组维护一个状态, 该状态包含组过滤模式和源列表以及定时器
 - 组状态格式为 (组地址, 组定时器, 过滤模式, 源记录列表)
 - 源记录格式为 (源地址, 源定时器)
- 每个组播组只对应一种过滤模式
 - 对于 INCLUDE 模式, 源记录列表包含该接口网段的主机需要接收的组播源列表, 表示为: INCLUDE (S, G)
 - 对于 EXCLUDE 模式, 源记录列表包含两类源列表, 第一类与过滤模式相反, 是主机需要接收的组播源列表 S1; 第二类是主机不需要接收的组播源列表 S2, 表示为: EXCLUDE (S1, S2, G)

www.h3c.com

IGMPv3 路由器侧也为接口上的每一个组播组维护状态信息, 和 IGMPv3 主机侧相比, 路由器还为每一个组播组以及源列表中的每一个组播源维护状态定时器。

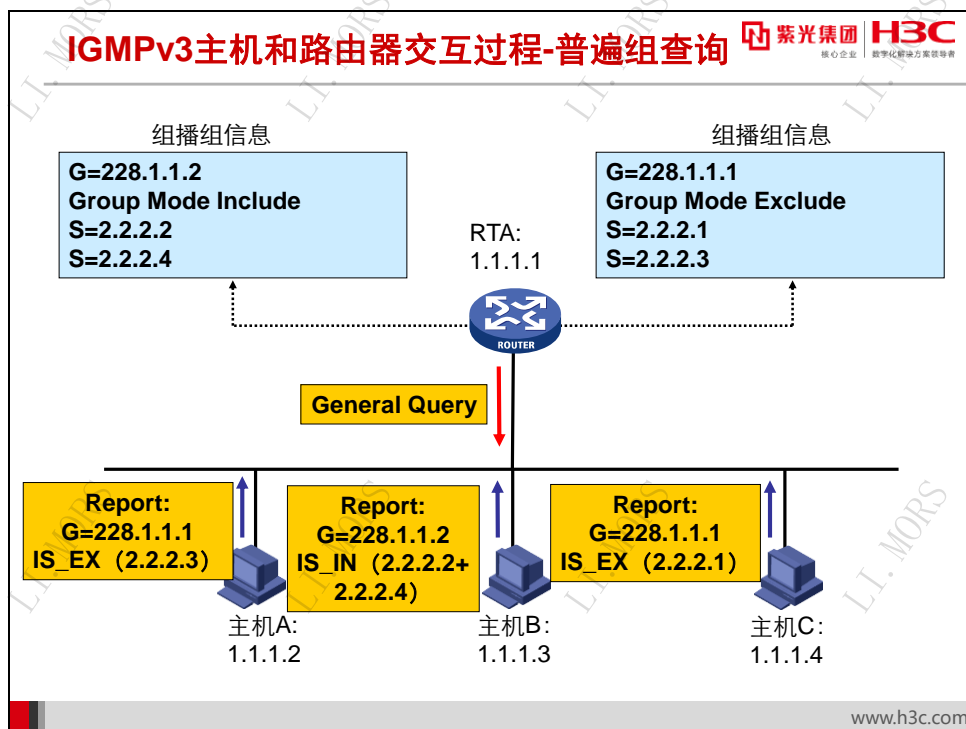
路由器维护的组状态格式为（组地址，组定时器，过滤模式，源记录列表），其中源记录的格式为（源地址，源定时器）。

路由器维护的每一个组播组只对应一种过滤模式：

- 对于 **INCLUDE** 模式，源记录列表包含该接口所属网段内的主机需要接收的所有组播源，表示为 **INCLUDE (S, G)**；
- 对于 **EXCLUDE** 模式，源记录列表包含两类源列表，第一类与过滤模式相反，是主机需要接收的组播源列表 **S1**；第二类是主机不需要接收的组播源列表 **S2**，表示为：**EXCLUDE (S1, S2, G)**。

例如，对于组播组 **228.1.1.1**，过滤模式为 **EXCLUDE**，接口所属网段上有主机希望接收组播源 **1.1.1.1** 的数据，有主机不希望接收组播源 **2.2.2.2** 和 **3.3.3.3** 的数据，则路由器维护的组播组状态为 **EXCLUDE (1.1.1.1, 2.2.2.2+3.3.3.3, 228.1.1.1)**。也就是说，当路由器维护的组播组为 **EXCLUDE** 模式时，路由器既要记录主机不想接收的组播源也同时要记录主机需要接收的组播源。

组定时器只在 **EXCLUDE** 模式时有用。当组定时器老化时，如果所有源定时器也老化则路由器会删除该组记录；如果此时仍有源定时器运行，则组过滤模式从 **EXCLUDE** 变为 **INCLUDE**。



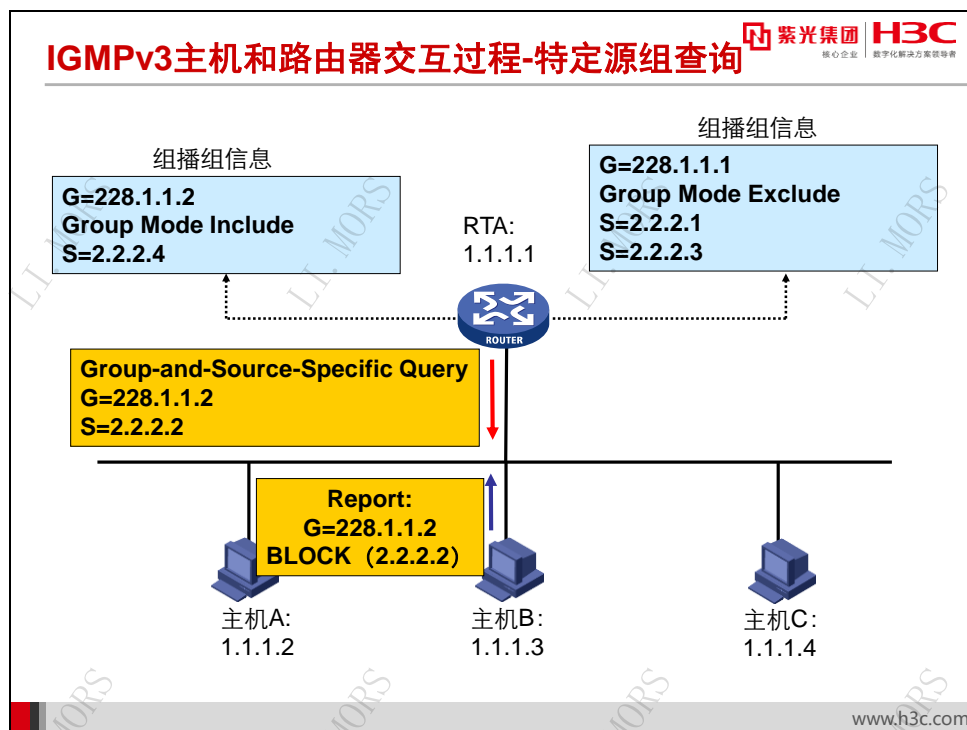
IGMPv3 路由器 RTA 周期性发送 General Query 报文，IGMPv3 主机收到 General Query 报文后，在接口设置定时器，定时器超时后发送 Membership Report 报文。

主机 A 希望接收组播组 228.1.1.1 的数据，但是不希望接收来自组播源 2.2.2.3 的数据，因此主机 A 回复的 Membership Report 报文中的组记录为 IS_EX (2.2.2.3, 228.1.1.1)。

主机 B 希望接收组播组 228.1.1.2 的数据，且希望接收来自组播源 2.2.2.2 和 2.2.2.4 的数据，因此主机 B 回复的 Membership Report 报文中的组记录为 IS_IN (2.2.2.2+2.2.2.4, 228.1.1.2)。

主机 C 希望接收组播组 228.1.1.1 的数据，但是不希望接收来自组播源 2.2.2.1 的数据，因此主机 C 回复的 Membership Report 报文中的组记录为 IS_EX (2.2.2.1, 228.1.1.1)。

路由器收到 Membership Report 报文后，会为组播组 228.1.1.1 和 228.1.1.2 建立表项维护状态信息。其中，组播组 228.1.1.1，过滤模式为 EXCLUDE，排除的组播源为 2.2.2.1 和 2.2.2.3，表示为 EXCLUDE (NULL, 2.2.2.1+2.2.2.3, 228.1.1.1)。组播组 228.1.1.2 过滤模式为 INCLUDE，需要接收的组播源为 2.2.2.2 和 2.2.2.4，表示为 INCLUDE (2.2.2.2+2.2.2.4, 228.1.1.2)。

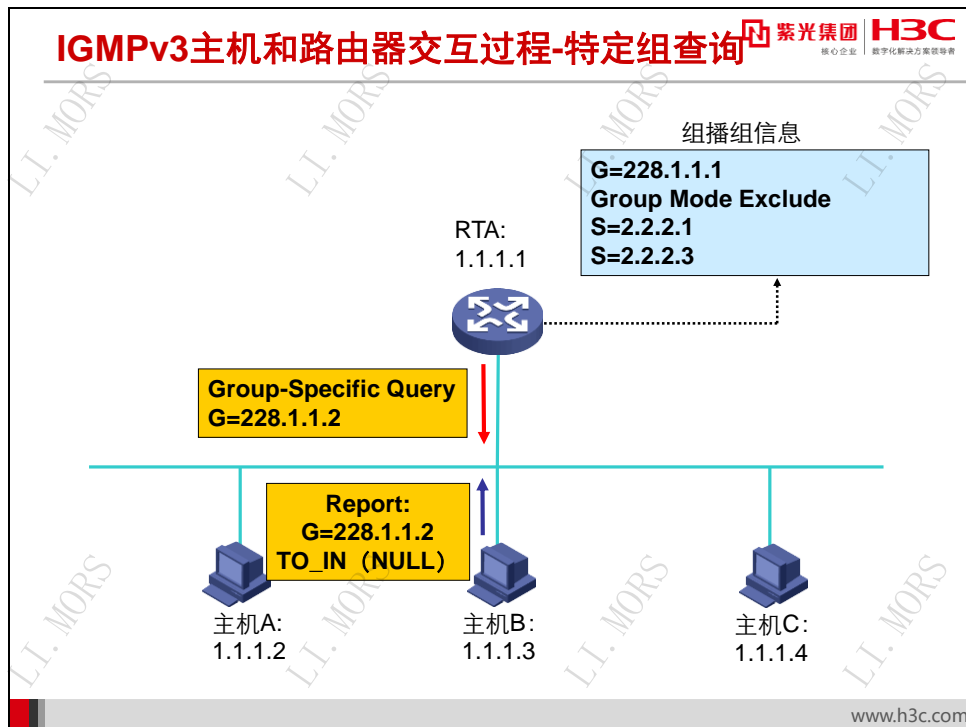


当主机 B 不再希望接收来自组播源 2.2.2.2 发送的组播数据时，会主动发送 Membership Report 报文，报文中包含的组记录为 BLOCK (2.2.2.2, 228.1.1.2)。

路由器收到组记录为 BLOCK (2.2.2.2, 228.1.1.2) 的 Membership Report 报文后，会发送 Group-and-Source-Specific Query 报文，询问网段上还有没有主机希望接收组播源 2.2.2.2 发送的组播地址为 228.1.1.2 的组播数据。

如果网段上仍然有主机愿意接收该特定源发送的组播流，主机会回应 Membership Report 报文，报文中包含的组记录为 IS_IN (2.2.2.2, 228.1.1.2)。

如果在 Max Reps Time 时间内，路由器没有收到任何对于该 Group-and-Source-Specific Query 报文的回应，则路由器会在组播组 228.1.1.2 对应的源列表中，删除组播源 2.2.2.2，此时该组播组状态信息为 INCLUDE (2.2.2.4, 228.1.1.2)。组播组 228.1.1.1 的源列表信息保持不变。



当主机不愿意接收某组播组的数据时，可以发送离开组报文。IGMPv3 取消了 Leave 报文，使用特殊的组记录方式来表示主机离开某个组。

例如，主机 B 希望离开组播组 228.1.1.2，则主机 B 会主动发送 Membership Report 报文，报文中包含的组记录为 TO_IN (NULL, 228.1.1.2)，表示不想接收任何组播源发送到组播组 228.1.1.2 的数据。

路由器收到组记录为 TO_IN (NULL, 228.1.1.2) 的报文，会为组播组 228.1.1.2 发送一个 Group-Specific Query 报文，如果在 Max Reps Time 时间内路由器没有收到任何关于该 Group-Specific Query 报文的回应，则路由器会删除整个组播组 228.1.1.2 的表项信息。

IGMPv3 还可以通过一种特殊的组记录来实现 IGMPv1/v2 中主机加入某组播组的功能。例如，当 IGMPv3 主机想要加入组播组 228.1.1.1 且不关心组播源时，可以发送组记录为 IS_EX (NULL, 228.1.1.1) 的 Membership Report 报文，表示不排除任意源发送到组播组 228.1.1.1 的数据。

IGMPv3 报文类型



- **IGMPv3 定义的报文类型**

- 0x11: Membership Query 报文

- 0x22: Membership Report 报文

- **兼容 IGMPv1/v2 版本的报文类型**

- 0x12: IGMPv1 Membership Report 报文

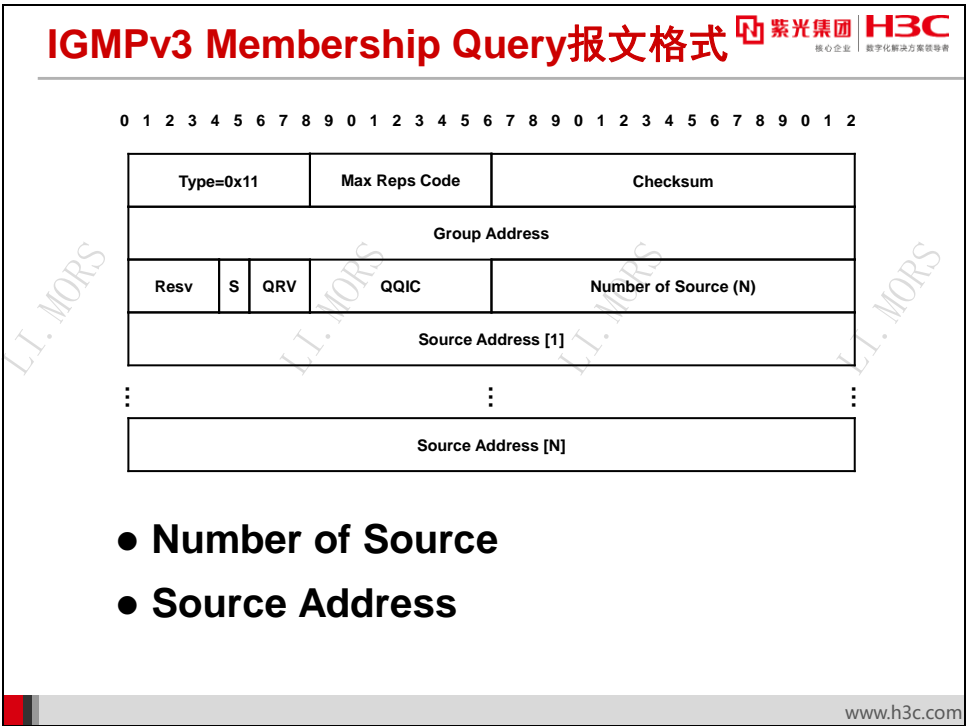
- 0x16: IGMPv2 Membership Report 报文

- 0x17: IGMPv2 Leave Group 报文

www.h3c.com

IGMPv3 有 Membership Query 报文和 Membership Report 报文两种类型，Membership Query 报文类型号为 0x11，包含 General Query 报文、Group-Specific Query 报文和 Group-and-Source-Specific Query 报文。Membership Report 报文类型号为 0x22。

此外，IGMPv3 可以兼容 IGMPv1/v2 的报文，包含类型号为 0x12 的 IGMPv1 Membership Report 报文、类型号为 0x16 的 IGMPv2 Membership Report 报文和类型号为 0x17 的 IGMPv2 Leave Group 报文。



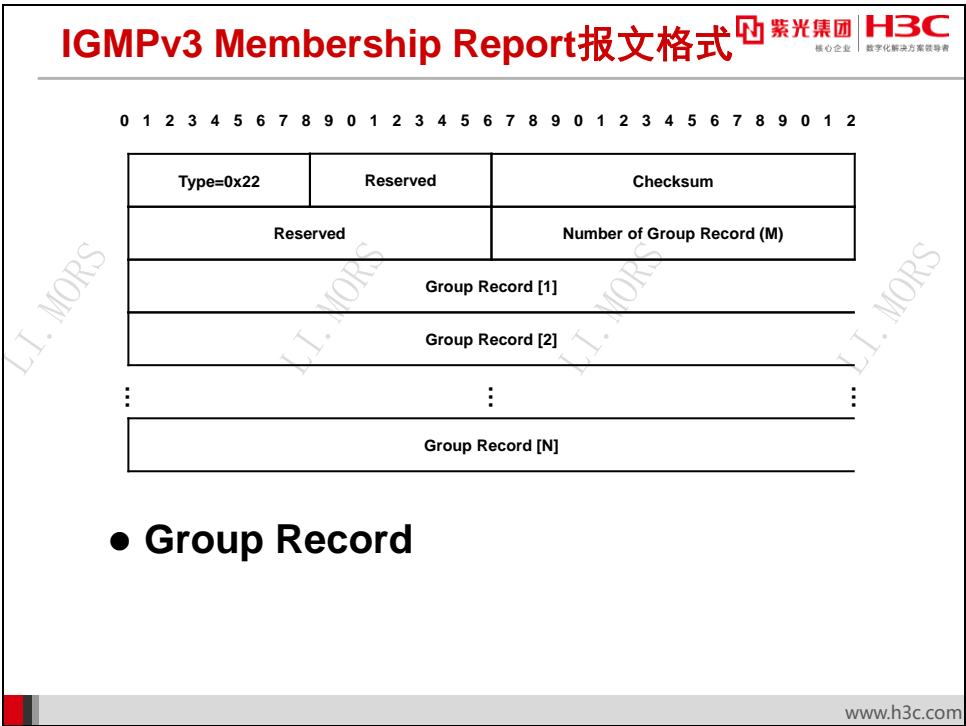
IGMPv3 的查询报文在 IGMPv1/v2 报文的基础上增加了源地址列表，源地址数以及一些标志位。

对于普遍组查询报文，组地址字段置位 0，源地址数字段也为 0，报文中不包含任何源地址信息。

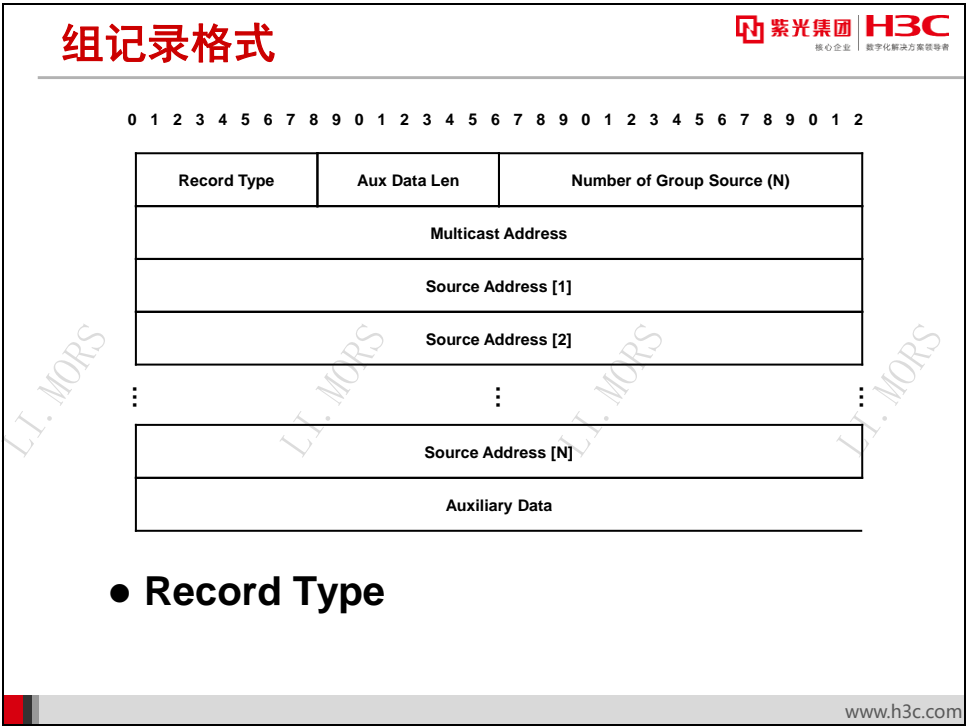
对于特定组查询报文，组地址字段为被查询的组播组地址，源地址数字段也为 0，报文中不包含任何源地址信息。

对于特定源组查询报文，组地址字段为被查询的组播组地址，源地址数字段为被查询的源地址的个数，在源地址字段列出被查询的各个源地址。

IGMPv3 普遍组查询报文的地址为 224.0.0.1，特定组查询报文和特定源组查询报文的地址为被查询的组播组地址。



IGMPv3 的 Report 报文比 IGMPv1/v2 的 Report 报文增加了组记录数字段和组记录字段。一个 Report 报文可以携带多个组记录，每一个组记录单独的记录了某个组的记录类型以及对应的组播源信息。



组记录格式主要包括：记录类型、组播组地址、组播源个数和组播源地址列表。组记录类型包含当前状态记录、过滤模式改变记录和源列表变化记录三种类型。其中当前状态记录又可以分为 IS_IN 和 IS_EX 两类，过滤模式改变记录分为 TO_IN 和 TO_EX 两类，源列表变化记录分为 ALLOW 和 BLOCK 两类。

19.5 IGMP不同版本间的操作

IGMP版本判断及版本功能差异



- 通过比较类型字段及报文长度可以判断IGMP的版本
 - IGMPv1: 报文长度为8字节且Max Resp Code字段为0
 - IGMPv2: 报文长度为8字节且Max Resp Code字段不为0
 - IGMPv3: 报文长度大于8字节

	IGMPv1	IGMPv2	IGMPv3
查询器选举	依靠上层路由协议	自己选举	自己选举
离开组方式	默默离开	主动发出离开报文	主动发出离开报文
特定组查询	无	有	有
特定源、组加入	无	无	有

www.h3c.com

当运行不同版本的 IGMP 路由器和主机协同工作时，通过比较 IGMP 报文中的类型字段以及报文的长度即可判断 IGMP 的版本。

IGMP 的 Membership Report 报文的版本可以直接通过类型字段进行区分，不同版本的 Membership Report 报文使用了不同的类型号。IGMP 的 Membership Query 报文版本可以通过比较报文长度以及 Max Reps Code 字段的值来确定。

- **IGMPv1:** Membership Query 报文长度为 8 字节，且 Max Reps Code 字段值为 0。IGMPv1 中 Membership Report 报文的最大响应延时不是通过 Membership Query 报文获得，而是使用固定值 10 秒。收到 Membership Query 报文后，主机会在[0, 10]秒间随机选择一个延时发送 Membership Report 报文。
- **IGMPv2:** Membership Query 报文长度为 8 字节，且 Max Reps Code 字段值不为 0。
- **IGMPv3:** Membership Query 报文长度大于 8 字节。

IGMP 各版本功能有较大差异，对于查询器选举、离开组方式、指定组查询、指定源组加入等功能，IGMP 各版本差别如下：

IGMPv1 本身不支持查询器选举，这个工作需要组播路由协议配合支持；IGMPv1 中没有 Leave Group 报文，当主机离开某个组时，不会发送任何消息选择默默的离开，如果该主机为组播组的最后一位接收者，则路由器默认情况下需要最多两个普遍查询间隔即两分钟的时间才能得知组播组没有成员，在此期间路由器会继续将组播数据发送到该网段，浪费大量网络带宽；IGMPv1 无法支持特定组查询和特定源组的加入。

IGMPv2 在 IGMPv1 的基础上，增加了查询器选举机制，并且通过使用 Leave Group 报文和 Group-Specific Query 报文大大减少了路由器感知主机离开某个组的时间。IGMPv2 不支持针对特定源组的加入。

IGMPv3 在 IGMPv2 的基础上增加了对特定源组加入的支持，路由器也可以针对特定源组进行查询。

IGMPv2和IGMPv1的兼容

- 与IGMPv1路由器的兼容
 - 当IGMPv2主机发现IGMPv1路由器，则必须发送版本1的Report报文，且不再发送Leave报文
 - 当网段上同时存在IGMPv1路由器和IGMPv2路由器，则查询器的IGMP版本必须使用较低版本即IGMPv1
- 与IGMPv1主机的兼容
 - IGMPv2路由器如果发现IGMPv1主机，则必须忽略后续收到的Leave报文
 - IGMPv2主机的Report报文会被IGMPv1主机的Report报文抑制

www.h3c.com

运行新 IGMP 版本的主机和路由器可以和运行老 IGMP 版本的主机和路由器兼容。当运行 IGMPv2 的主机或路由器和运行 IGMPv1 的主机或路由器协同工作时，分为如下几种场景：

- 与 IGMPv1 路由器的兼容

当 IGMPv2 主机和 IGMPv1 路由器一起运行时，由于 IGMPv1 路由器无法识别版本 2 的 Membership Report 报文，所以 IGMPv2 的主机必须发送版本 1 的 Membership Report 报文，且不再发送 Leave Group 报文。

当 IGMPv2 路由器和 IGMPv1 路由器一起运行时，将由 IGMPv1 路由器充当查询器的角色。

- 与 IGMPv1 主机的兼容

IGMPv2 路由器如果发现 IGMPv1 主机，则必须忽略网段上收到的任何 Leave Group 报文。因为 IGMPv1 主机不能识别 Group-Specific Query 报文，如果发送 Group-Specific Query 报文，IGMPv1 主机将不作任何回应，可能导致路由器错误的删除组播表项。

IGMPv2 主机的 Membership Report 报文会被 IGMPv1 主机发送的关于同一个组的 Membership Report 报文抑制。

IGMPv3和IGMPv1/v2的兼容-主机操作



紫光集团 核心企业 数字化解决方案领导者

● IGMPv3主机与IGMPv1/v2路由器的兼容

IGMPv3主机工作模式	定时器状态
IGMPv3（默认）	IGMPv2/v1查询器状态定时器没有运行
IGMPv2	IGMPv2查询器状态定时器运行，IGMPv1查询器状态定时器没有运行
IGMPv1	IGMPv1查询器状态定时器运行

● IGMPv3主机与IGMPv1/v2主机的兼容

→ 遵循IGMPv1/v2的成员报告抑制机制

www.h3c.com

当 IGMPv3 主机和 IGMPv1/v2 的路由器一起运行时，IGMPv3 主机会为网段上的 IGMPv1 或 IGMPv2 查询器维护一个定时器，IGMPv3 主机的工作模式由这些定时器状态决定。

当 IGMPv1 查询器状态定时器运行时，说明网段上有 IGMPv1 路由器且充当查询器的角色，则 IGMPv3 主机将工作在 IGMPv1 模式；当 IGMPv1 查询器状态定时器没有运行，而 IGMPv2 查询器状态定时器运行时，说明网络上有 IGMPv2 路由器且充当查询器的角色，则 IGMPv3 主机将工作在 IGMPv2 模式；当没有 IGMPv1/IGMPv2 查询器状态定时器运行时，IGMPv3 主机工作在其默认模式即 IGMPv3 模式。

当 IGMPv3 主机和 IGMPv1/v2 主机一同运行时，IGMPv3 主机的 Report 报文可以被 IGMPv1/v2 主机发送的关于同一个组的 Membership Report 报文抑制。

IGMPv3和IGMPv1/v2的兼容-路由器操作

紫光集团 H3C
核心企业 | 数字化解决方案领导者

● IGMPv3路由器与IGMPv1/v2路由器的兼容

→ 网段中的查询器需要运行较低的IGMP版本

● IGMPv3路由器与IGMPv1/v2主机的兼容

→ 维护IGMPv1/v2主机状态定时器

→ 收到IGMPv1/v2主机报文的处理

收到IGMP报文	IGMPv3路由器操作
IGMPv1 Report	转换成组记录类型为IS_EX (NULL, G) 的IGMPv3报文
IGMPv2 Report	转换城组记录类型为IS_EX (NULL, G) 的IGMPv3报文
IGMPv2 Leave	转换城组记录类型为TO_IN (NULL, G) 的IGMPv3报文

www.h3c.com

当 IGMPv3 路由器和 IGMPv1/v2 路由器一起运行时，网段中的查询器必须为较低 IGMP 版本的路由器，即 IGMPv1/IGMPv2 路由器。

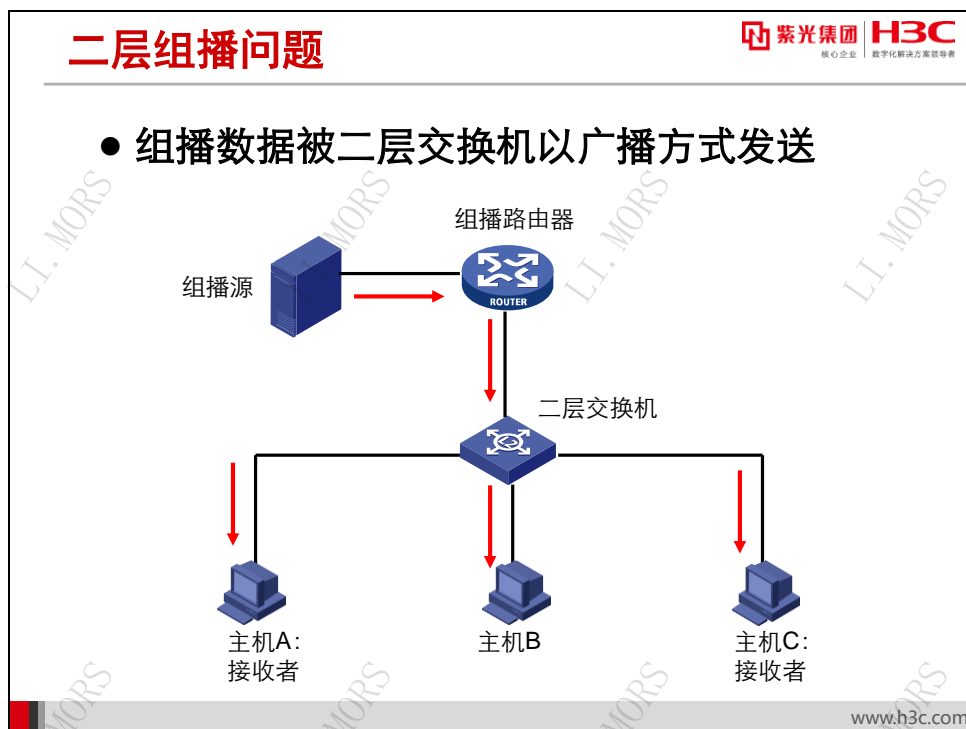
当 IGMPv3 路由器和 IGMPv1/v2 主机一起运行时，IGMPv3 路由器的工作模式由 IGMPv1/v2 主机状态定时器决定。

当 IGMPv1 主机状态定时器运行时，说明网段上有 IGMPv1 主机，则 IGMPv3 路由器将工作在 IGMPv1 模式；当 IGMPv1 主机状态定时器没有运行，而 IGMPv2 主机状态定时器运行时，说明网络上有 IGMPv2 主机，则 IGMPv3 路由器将工作在 IGMPv2 模式；当没有 IGMPv1/IGMPv2 主机状态定时器运行时，IGMPv3 路由器工作在其默认模式即 IGMPv3 模式。

当 IGMPv3 路由器发送 Membership Query 报文，收到 IGMPv1/v2 的 Membership Report 报文时，会将其转换成包含组记录 IS_EX (NULL, G) 的 IGMPv3 Membership Report 报文来处理；

当 IGMPv3 路由器收到 IGMPv2 的 Leave Group 报文时，会将其转换成包含组记录 TO_IN (NULL, G) 的 IGMPv3 Membership Report 报文来处理。

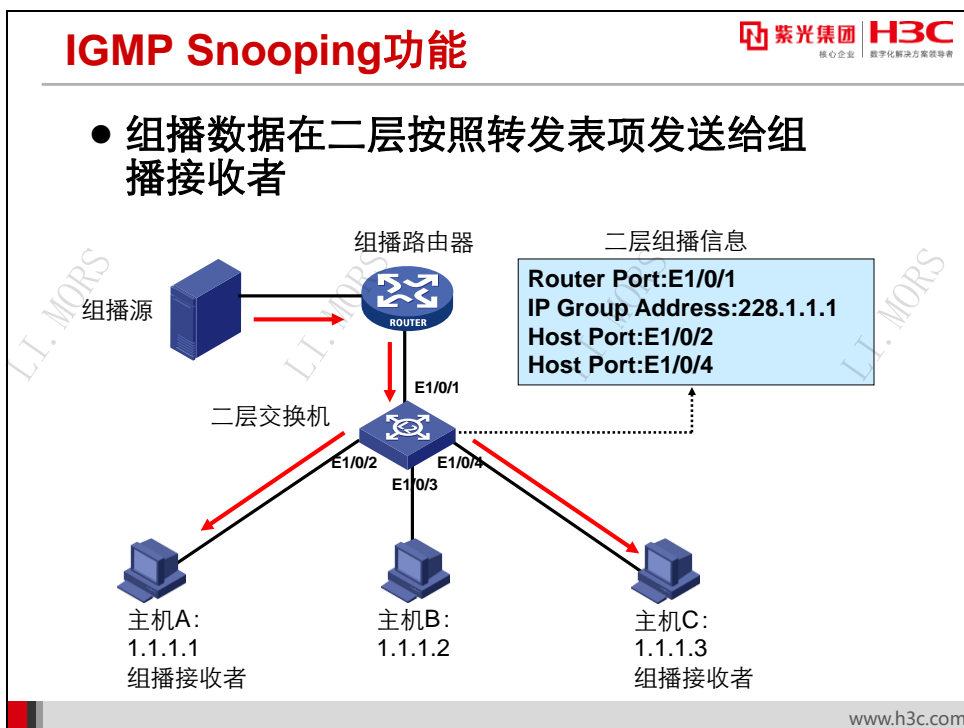
19.6 IGMP Snooping



在实际网络中，由于路由器接口数有限，所以主机通常都是通过交换机连接到路由器。

当主机和路由器之间的交换机为二层交换机时，其无法识别路由器发来的组播报文，因此会作为未知报文在网段内广播，导致不属于该组播组成员的主机也收到了组播报文。如图所示，主机 B 也将收到发往主机 A 和主机 C 的组播报文，由于主机 B 不是接收者，因此会将组播报文丢弃。

可以看到这种情况浪费了网络带宽，并增加了非接收者主机的处理负担，偏离了组播设计的初衷。



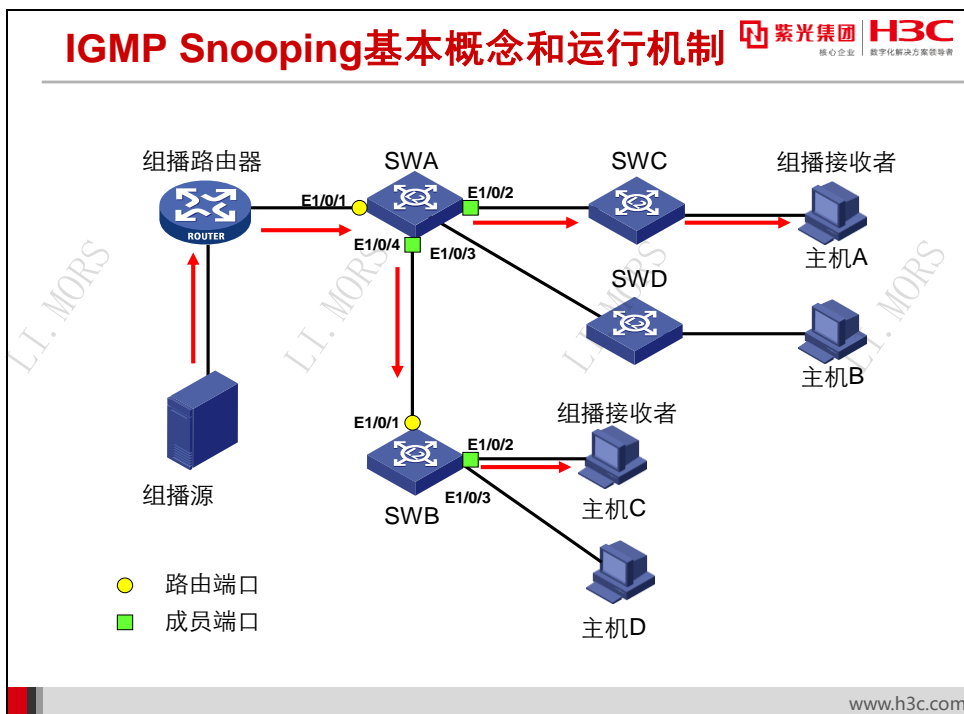
通过在二层交换机上实现 IGMP Snooping 功能可以解决组播报文在二层被广播发送的问题。

IGMP Snooping 是运行在二层设备上的组播约束机制，用于管理和控制组播组。运行 IGMP Snooping 的二层设备通过对收到的 IGMP 报文进行分析，为端口和 MAC 组播地址建立起映射关系，并根据这样的映射关系转发组播数据。

图中，主机 A 和主机 C 是组播组 228.1.1.1 的接收者，则二层交换机会通过 IGMP Snooping 为组播组 228.1.1.1 记录一个表项，表项中包含连接路由器的端口、组播组地址以及连接接收者的端口。当有目的地址为 228.1.1.1 的组播报文到达时，二层交换机只会将组播报文从连接接收者的端口 E1/0/2 和 E1/0/4 发送出去，从而避免了二层的广播。

IGMP Snooping 通过二层组播将信息只转发给有需要的接收者，可以带来以下好处：

- 减少了二层网络中的广播报文，节约了网络带宽；
- 增强了组播信息的安全性；
- 为实现对每台主机的单独计费带来了方便。



在路由器和主机间运行 IGMP Snooping 的交换机上存在两种端口：路由器端口和成员端口。

路由器端口（Router Port）指交换机上朝向三层组播设备（DR 或 IGMP 查询器）一侧的端口，如 SWA 和 SW B 各自的 Ethernet1/0/1 端口。交换机将本设备上的所有路由器端口都记录在路由器端口列表中。

成员端口（Member Port）又称组播组成员端口，表示交换机上朝向组播组成员一侧的端口，如 SWA 的 Ethernet1/0/2 和 Ethernet1/0/4 端口，以及 SWB 的 Ethernet1/0/2 端口。交换机将本设备上的所有成员端口都记录在 IGMP Snooping 转发表中。

IGMP Snooping 的工作机制包含：Membership Query 报文处理、Membership Report 报文处理和 Leave Group 报文处理。

IGMP 查询器定期向本地网段内的所有主机与路由器（224.0.0.1）发送 IGMP General Query 报文，以查询该网段有哪些组播组的成员。在收到 IGMP Membership Report 报文时，IGMP Snooping 交换机将其从 VLAN 内的所有路由器端口转发出去，并从该报文中解析出主机要加入的组播组地址，并对报文的接收端口做如下处理：

- 如果不存在该组播组所对应的转发表项，则创建转发表项，将该端口作为动态成员端口添加到出端口列表中，并启动其老化定时器；
- 如果已存在该组播组所对应的转发表项，但其出端口列表中不包含该端口，则将该端口作为动态成员端口添加到出端口列表中，并启动其老化定时器；

- 如果已存在该组播组所对应的转发表项，且其出端口列表中已包含该动态成员端口，则重置其老化定时器。

IGMP Snooping 交换机不会将 IGMP Membership Report 报文通过非路由器端口转发出去，因为根据主机上的 IGMP 成员报告抑制机制，如果非路由器端口下还有该组播组的成员主机，则这些主机在收到该 Membership Report 报文后便抑制了自身 Membership Report 报文的发送，从而使 IGMP Snooping 交换机无法获知这些端口下还有该组播组的成员主机。

运行 IGMPv1 的主机离开组播组时不会发送 IGMP 离开组报文，因此交换机无法立即获知主机离开的信息，只能等待成员端口的老化定时器超时后，交换机会将该端口从转发表中删除。

运行 IGMPv2 或 IGMPv3 的主机离开组播组时，会通知路由器自己离开了某个组播组。当交换机从某动态成员端口上收到 IGMP 离开组报文时，首先判断要离开的组播组所对应的转发表项是否存在，以及该组播组所对应转发表项的出端口列表中是否包含该接收端口：

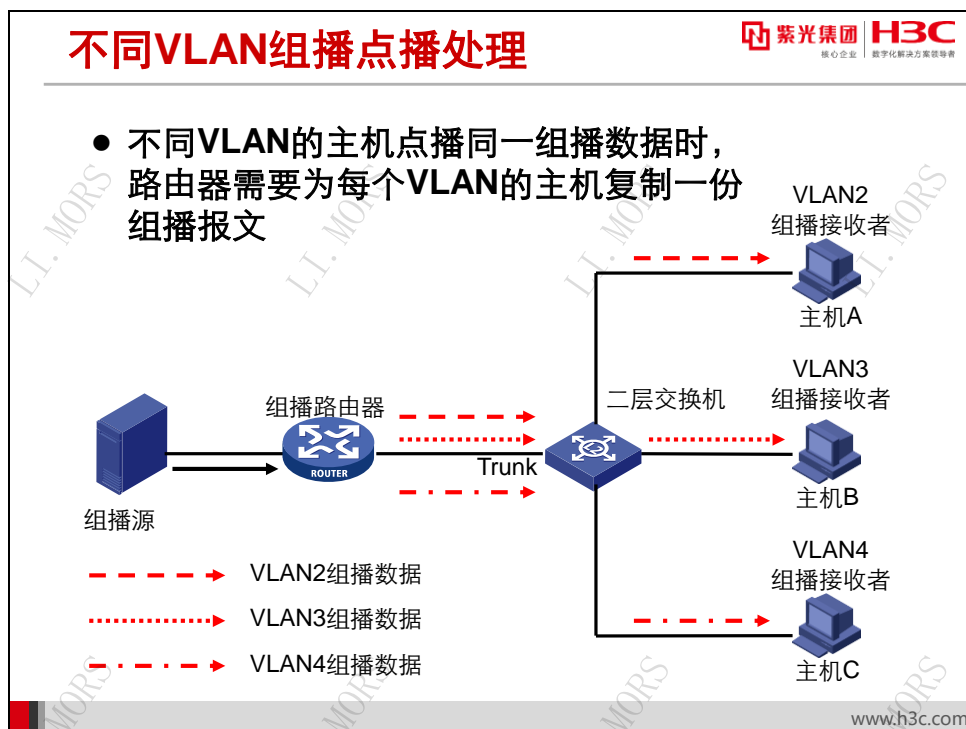
- 如果不存在该组播组对应的转发表项，或者该组播组对应转发表项的出端口列表中不包含该端口，交换机将该报文直接丢弃；
- 如果存在该组播组对应的转发表项，且该组播组对应转发表项的出端口列表中包含该端口，交换机会将该报文从 VLAN 内的所有路由器端口转发出去。同时，由于并不知道该接收端口下是否还有该组播组的其它成员，所以交换机不会立刻把该端口从该组播组所对应转发表项的出端口列表中删除，而是重置其老化定时器。

当 IGMP 查询器收到 IGMP Leave Group 报文后，从中解析出主机要离开的组播组的地址，并通过接收端口向该组播组发送 IGMP Group-Specific Query 报文。交换机在收到 IGMP Group-Specific Query 报文后，将其从 VLAN 内的所有路由器端口和该组播组的所有成员端口转发出去。

对于 IGMP Leave Group 报文的接收端口（假定为动态成员端口），交换机在其老化时间内：

- 如果从该端口收到了主机响应该特定组查询的 IGMP Membership Report 报文，则表示该端口下还有该组播组的成员，于是重置其老化定时器；
- 如果没有从该端口收到主机响应特定组查询的 IGMP Membership Report 报文，则表示该端口下已没有该组播组的成员，则在其老化时间超时后，将其从该组播组所对应转发表项的出端口列表中删除。

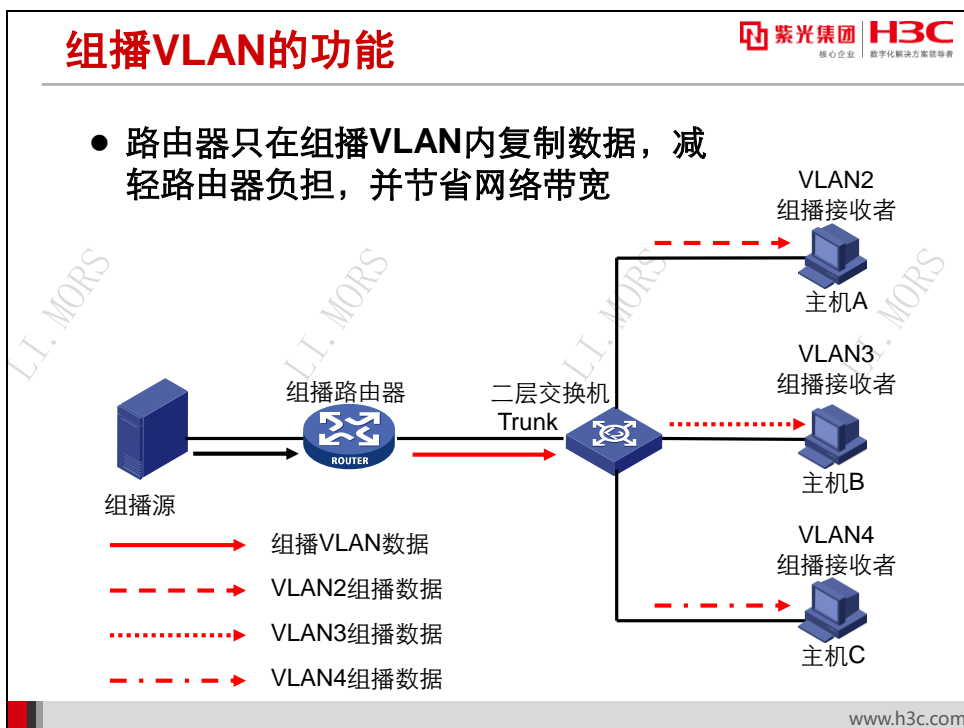
19.7 组播VLAN



在传统的组播点播方式下，当属于不同 VLAN 的主机同时点播同一组播组时，路由器需要把组播数据在每个用户 VLAN 内都复制一份发送给二层交换机。这样既造成了带宽的浪费，也给路由器增加了额外的负担。

图中，主机 A 位于 VLAN2，主机 B 位于 VLAN3，主机 C 位于 VLAN4，三台主机都希望接收同一组播组的数据，则路由器在收到组播报文后，需要在 VLAN2、VLAN3 和 VLAN4 内各复制一份组播报文，然后发送给二层交换机。二层交换机从 Trunk 端口收到组播数据后，分别在对应的 VLAN 内发送。

当存在大量属于不同 VLAN 的接收者时，路由器和二层交换机之间的链路上将会传送大量内容相同的组播报文，而路由器也会增加很多处理负担。



使用组播 VLAN 功能可以解决这个问题。在二层交换机上配置了组播 VLAN 后，路由器只需把组播数据在组播 VLAN 内复制一份发送给二层交换机，而不必在每个用户 VLAN 内都复制一份，从而节省了网络带宽，也减轻了路由器的负担。

图中，接收者主机 A、B 和 C 分属不同的用户 VLAN。在二层交换机上配置 VLAN 10 为组播 VLAN，将所有的用户 VLAN 都配置为该组播 VLAN 的子 VLAN，并在组播 VLAN 内使能 IGMP Snooping。

配置完成后，IGMP Snooping 将在组播 VLAN 中对路由器端口进行维护，而在各子 VLAN 中对成员端口进行维护。这样，路由器只需把组播数据在组播 VLAN 内复制一份发送给二层交换机，二层交换机会将其分发给该组播 VLAN 内那些有接收者的子 VLAN。

19.8 本章总结

本章总结

- 对IGMP的协议机制进行介绍
- 介绍了IGMP不同版本间的互操作
- 介绍二层组播协议IGMP Snooping
- 介绍了组播VLAN功能

www.h3c.com

19.9 习题和解答

19.9.1 习题

1. 组播组管理协议的机制主要包含 ()
 - A. 主机加入和离开组播组
 - B. 路由器维护组播组
 - C. 查询器的选举
 - D. 成员报告抑制机制
2. IGMPv2 协议报文包含如下哪几类 ()
 - A. General Query
 - B. Group-Specific Query
 - C. Membership Report
 - D. Leave Group
3. IGMPv3 Membership Report 报文中的组记录字段有几种类型 ()
 - A. 当前状态记录
 - B. 过滤模式变化记录
 - C. 源列表变化记录
 - D. 组变化记录
4. 以下关于 IGMP Snooping, 正确的是 ()
 - A. IGMP Snooping 是三层组播协议
 - B. 路由器端口指交换机上朝向三层组播设备一侧的端口
 - C. 成员端口表示交换机上朝向组播组成员一侧的端口
 - D. 在收到 Membership Report 报文时, IGMP Snooping 交换机将其从 VLAN 内除接收端口以外的其它所有端口转发出去
5. 以下关于组播 VLAN, 正确的是 ()
 - A. 在传统的组播点播方式下, 当属于不同 VLAN 的主机同时点播同一组播组时, 路由器需要把组播数据在每个用户 VLAN 内都复制一份发送给二层交换机
 - B. 在二层交换机上配置了组播 VLAN 后, 路由器只需把组播数据在组播 VLAN 内复制一份发送给二层交换机
 - C. IGMP Snooping 将在组播 VLAN 中对路由器端口进行维护
 - D. IGMP Snooping 将在各子 VLAN 中对成员端口进行维护

19.9.2 习题答案

1. ABCD 2. ABCD 3. ABC 4. BC 5. ABCD

第20章 组播转发机制

与单播转发路径不同，组播数据转发路径基于树形结构，不同的组播路由协议计算生成的组播分发树不同。

由于组播报文的目的地址是一组接收者，无法从发送方向判断组播转发的最优路径，组播采用逆向路径转发检查机制确保组播数据沿正确路径传输。

本章介绍了组播报文分发树模型以及组播报文转发时的 RPF 检查机制。

20.1 本章目标

课程目标

学习完本课程，您应该能够：

- 掌握组播分发树模型
- 掌握RPF机制



20.2 组播分发树模型

组播分发树概念

- 组播分发树指组播数据在网络中的转发路径，由组播路由协议建立
- 根据树根节点的不同，组播分发树可分为最短路径树和共享树
 - SPT（Shortest Path Tree，最短路径树）：树根为组播源所连接的指定路由器
 - RPT（Rendezvous Point Tree，共享树）：树根为 RP（Rendezvous Point，汇聚点）



核心企业 | 数字化解决方案领导者

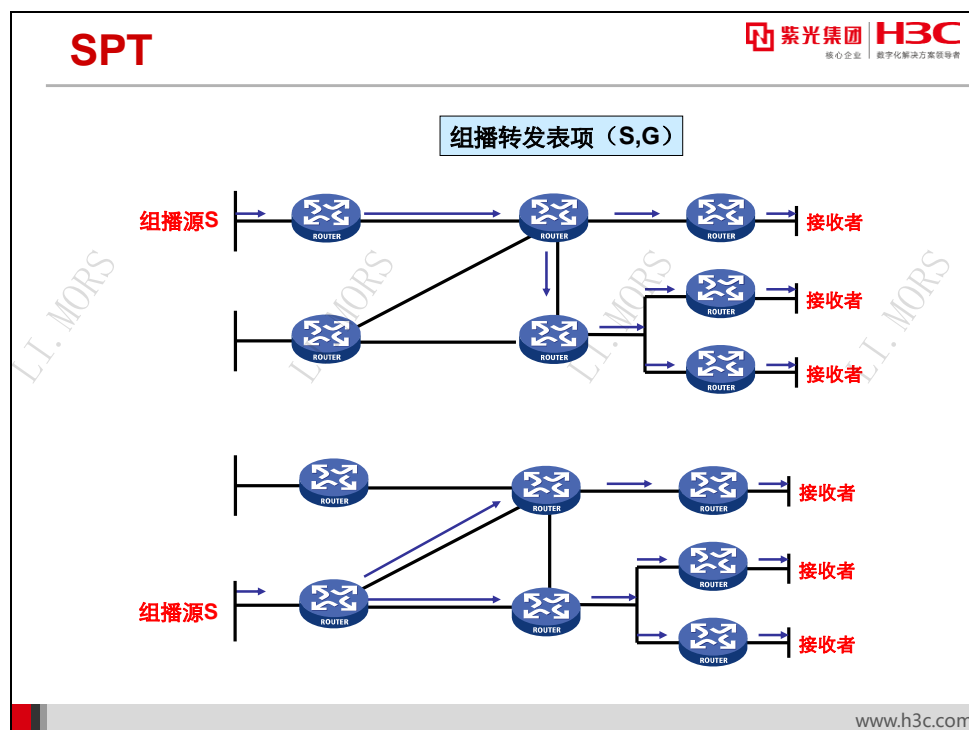
www.h3c.com

单播传输中，接收者是唯一的，所以从发送源到接收者之间的路径是点到点的一条线。

组播传输中，接收者可能是一组分布于网络中任何角落的主机，这就决定了组播数据的转发路径是点到多点的一棵树，因此组播报文的转发路径称为组播分发树。

组播分发树由组播路由协议建立。根据树根节点的不同，组播分发树模型可分为：

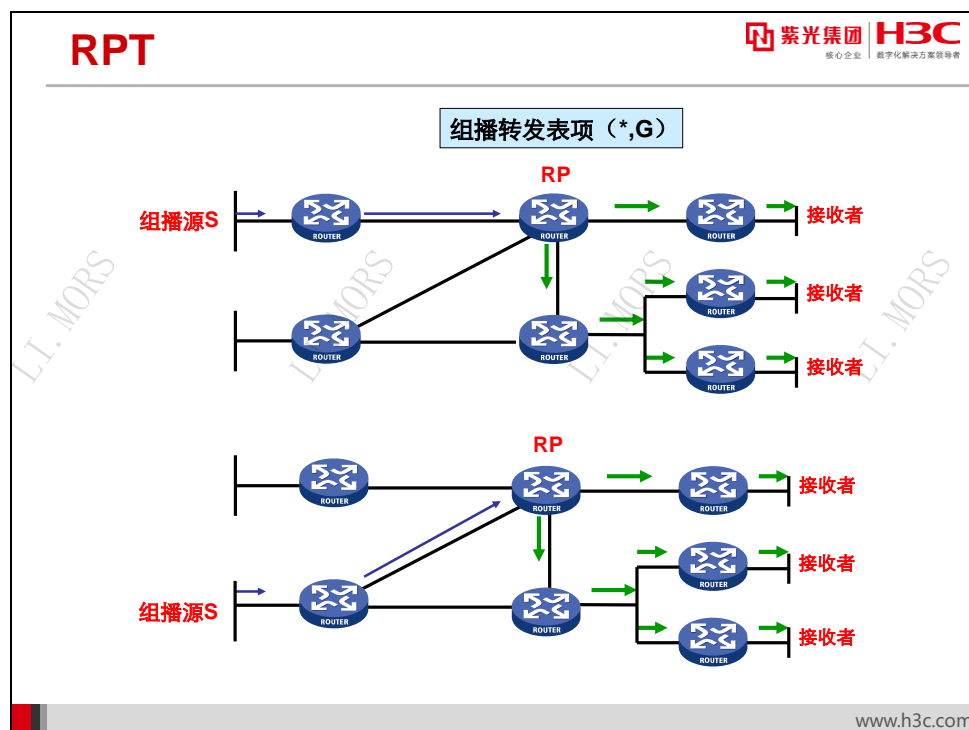
- **SPT（Shortest Path Tree，最短路径树）**：树根为组播源，从树根出发到达每一个接收者所经过的路径都是最优的。
- **RPT（Rendezvous Point Tree，共享树）**：树根是网络中的某一台路由器，称为 RP（Rendezvous Point，汇聚点）。网络中所有接收者共享 RP，作为组播分发树的树根。组播源发出的组播数据必须首先到达 RP，再由 RP 分发给各个组播接收者，这样就无法保证组播数据从组播源到组播接收者经过的是最优路径。



在 SPT 模型中，组播源到达任何一个接收者所经过的路径都是最优的。

SPT 上的每一个路由器都会维护 (S, G) 表项，用于组播报文的转发，其中 S 表示 SPT 的根即组播源，G 指该 SPT 的组播组地址。

图中可以看到，从组播源到每一个接收者所经过的路径都是最优路径。不同组播源以自己为根，独立建立 SPT。沿 SPT 传送组播报文可以将报文转发的路径延迟降到最低。



RPT 由接收者端发起建立，由于接收者不了解组播源的位置，所以需要在网络中指定一个特殊的节点，作为所有接收者共享的树根，这个根节点就是 RP（Rendezvous Point，汇聚点）。

RPT 上的每一个路由器都会维护 $(*, G)$ 表项，用于组播报文的转发，其中*表示任意源，G 指组播组地址。

图中，组播源分别发送组播报文给各自的接收者，在组播报文到达接收者之前首先需要经过 RP，然后再由 RP 分发给不同的接收者。可以看到组播报文在组播源和部分接收者之间没有走最优路径。

和 SPT 相比，RPT 的路径不是最优的，组播报文转发时引入了额外的路径延迟。但是路由器维护的表项信息比较简单，可以节省路由器内存空间。

20.3 RPF机制

RPF介绍



核心企业 | 数字化解决方案领导者

- **RPF (Reverse Path Forwarding, 逆向路径转发)**
 - 确保组播数据沿正确路径传输
 - 避免组播路径环路产生
- **RPF检查的过程如下:**
 - 如果数据包是在到达组播源的最优路径上到达, 则RPF检查成功, 数据包被转发
 - 如果RPF检查失败, 弃数据包
- **RPF检查基于的是单播路由表**

www.h3c.com

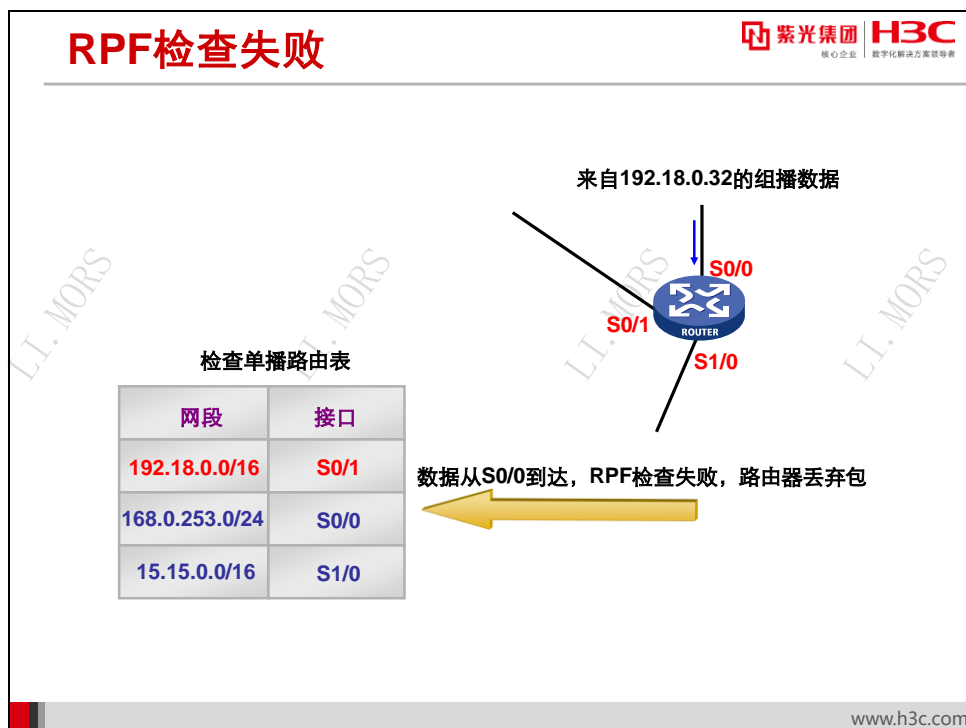
单播数据包转发的依据是目的 IP 地址, 该目的地址在网络中唯一地标识了一个节点的位置, 所以根据目的地址即可以确定一条到达目的节点的最优路径。沿着确定的路径转发还可以避免环路的产生。

组播数据包的目的地址是组播地址, 该组播地址标识了网络中的一组接收者, 这些接收者可能处于网络中的任意位置, 仅通过目的地址无法确保组播报文沿着正确的路径转发。此外, 组播接收者位置的不确定可能会导致路径环路的产生。

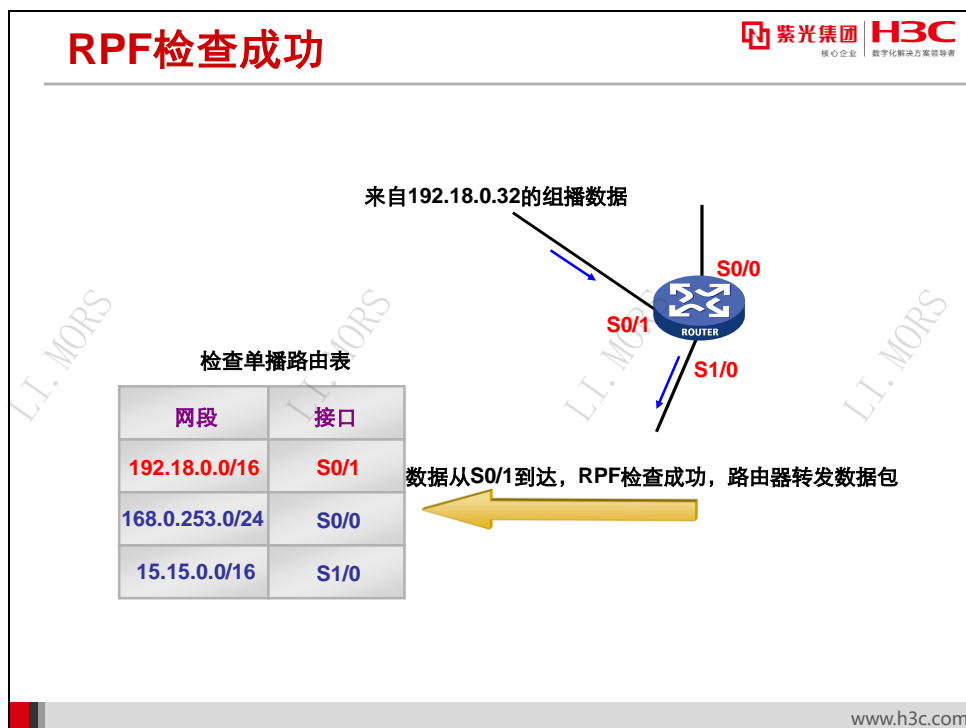
通过 RPF (Reverse Path Forwarding, 逆向路径转发) 检查机制可以解决上述问题。

RPF 检查依据的是组播数据包的源地址。进行 RPF 检查时, 以组播数据包的源 IP 地址为目的地址查找单播路由表, 选取一条最优单播路由。对应表项中的出接口为 RPF 接口, 下一跳为 RPF 邻居。路由器认为来自 RPF 邻居且由该 RPF 接口收到的组播包所经历的路径是从“报文源”到本地的最短路径。如果 RPF 检查失败即组播包不是从到达“报文源”的下一跳出接口到达, 则丢弃收到的组播包。沿途每一个路由器都进行同样的检查, 可以保证组播包从“报文源”到接收者之间的路径是最优的。

根据组播分发树模型的不同, “报文源”所代表的含义也不同。如果当前组播包沿着从组播源到接收者或组播源到 RP 的 SPT 进行传输, 则以组播源为“报文源”进行 RPF 检查; 如果当前报文沿着从 RP 到接收者的 RPT 进行传输, 则以 RP 为“报文源”进行 RPF 检查。

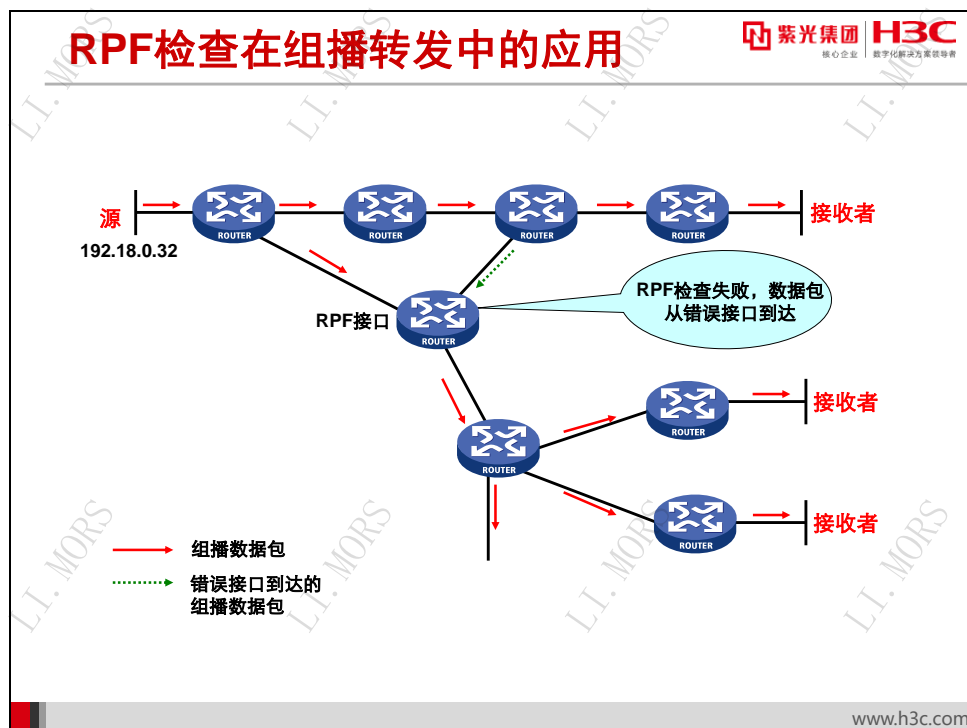


图中组播报文从路由器的接口 S0/0 到达，报文的源 IP 地址为 192.18.0.32。路由器将报文的源 IP 地址作为目的地址查找单播路由表，发现从路由器出发到达报文源的单播路由下一跳的出接口为 S0/1，不同于报文当前到达路由器的接口，说明该报文不是沿着最优的路径到达路由器，路由器丢弃该报文。



图中组播报文从路由器的接口 S0/1 到达，报文的源 IP 地址为 192.18.0.32。路由器将报文的源 IP 地址作为目的地址查找单播路由表，发现从路由器出发到达报文源的单播路由由下一跳的出接口为 S0/1，和报文当前到达路由器的接口相同，说明该报文沿着正确的接口到达路由器，路由器将该报文从所有对应的出接口发送出去。

组播报文的出接口列表中具体包含哪些接口，由组播路由协议确定。



对每一个收到的组播数据报文都进行 RPF 检查会给路由器带来较大负担，利用组播转发表可以解决这个问题。

路由器转发组播报文基于的是组播转发表项，表项包含报文源、组播组地址、入接口和出接口列表等信息。对于 SPT，此处的报文源为组播源；对于 RPT，此处的报文源为 RP。为便于表示，此处统一使用 (S, G) 代表组播转发表项，S 代表报文源，G 代表组播组。

当路由器收到组播数据报文后，查找组播转发表：

- 如果组播转发表中不存在对应的 (S, G) 表项，则对该报文执行 RPF 检查，将得到的 RPF 接口作为 (S, G) 表项的入接口，并结合路由信息得到 (S, G) 表项的出接口列表。

若该报文实际到达的接口正是其 RPF 接口，则 RPF 检查通过，向出接口列表中的所有出接口转发该报文；若该报文实际到达的接口不是其 RPF 接口，则 RPF 检查失败，丢弃该报文。

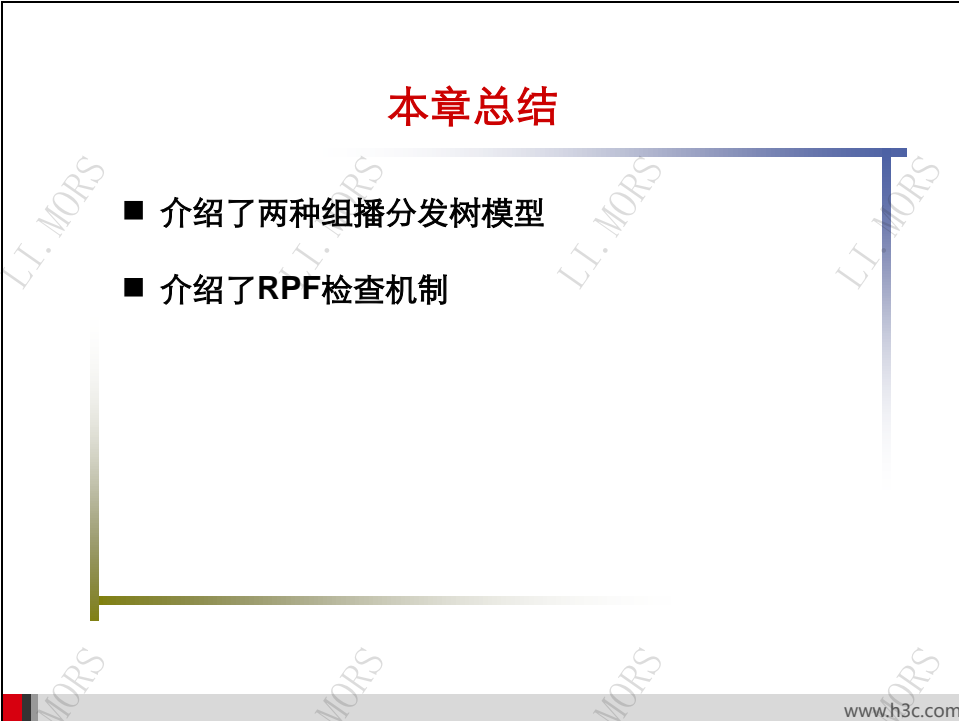
- 如果组播转发表中已存在 (S, G) 表项，且该报文实际到达的接口与入接口相匹配，则向出接口列表中的所有出接口转发该报文。

- 如果组播转发表中已存在（S，G）表项，但该报文实际到达的接口与表项中的入接口不匹配，则对此报文执行 RPF 检查。

如果检查得到的 RPF 接口与表项中的入接口一致，则说明当前（S，G）表项正确，该报文 RPF 检查失败，丢弃这个来自错误路径的报文。

如果其 RPF 接口与表项中的入接口不符，则说明（S，G）表项已过时，于是把表项中的入接口更新为检查得到的 RPF 接口。如果该报文实际到达的接口正是该 RPF 接口，则向出接口列表中的所有出接口转发该报文，否则将其丢弃。

20.4 本章总结



本章总结

- 介绍了两种组播分发树模型
- 介绍了RPF检查机制

www.h3c.com

20.5 习题和解答

20.5.1 习题

1. 关于组播分发树模型，下列哪些说法是正确的？（ ）
 - A. 组播分发树模型分为 SPT 和 RPT 两种
 - B. SPT 树根为组播源，从树根出发到达每一个接收者所经过的路径都是最优的
 - C. 树根是网络中的 RP，从树根出发到达每一个接收者所经过的路径不一定是最优
 - D. RPT 中路由器维护的表项信息比较简单，可以节省路由器内存空间。
2. 关于 SPT，正确的是（ ）
 - A. SPT 模型中，组播源到达任何一个接收者所经过的路径都是最优的
 - B. SPT 上的每一个路由器都会维护（*, G）表项
 - C. SPT 上的每一个路由器都会维护（S, G）表项
 - D. 不同组播源以自己为根，独立建立 SPT。
3. 关于 RPT，正确的是（ ）
 - A. RPT 由接收者端发起建立
 - B. RPT 上的每一个路由器都会维护（*, G）表项
 - C. RPT 上的每一个路由器都会维护（S, G）表项
 - D. 组播源分别发送组播报文给各自的接收者，在组播报文到达接收者之前首先需要经过 RP，然后再由 RP 分发给不同的接收者。
4. 关于 RPF，下列哪些说法是正确的？（ ）
 - A. RPF 检查依据的是组播数据包的源地址
 - B. 进行 RPF 检查时，以组播数据包的源 IP 地址为目的地址查找单播路由表，选取一条最优单播路由
 - C. 如果当前组播包沿着从组播源到接收者或组播源到 RP 的 SPT 进行传输，则以组播源为“报文源”进行 RPF 检查
 - D. 如果当前报文沿着从 RP 到接收者的 RPT 进行传输，则以 RP 为“报文源”进行 RPF 检查。
5. 当路由器收到组播数据报文后，查找组播转发表，如下哪些说法是正确的？（ ）
 - A. 路由器转发组播报文基于的是组播转发表项，表项包含报文源、组播组地址、入接口和出接口列表等信息

- B. 如果组播转发表中不存在对应的（S，G）表项，则对该报文执行 RPF 检查，将得到的 RPF 接口作为（S，G）表项的入接口
- C. 如果组播转发表中已存在（S，G）表项，且该报文实际到达的接口与入接口相匹配，则向出接口列表中的所有出接口转发该报文
- D. 如果组播转发表中已存在（S，G）表项，且该报文实际到达的接口与入接口相匹配，则向出接口列表中的所有出接口转发该报文

20.5.2 习题答案

- 1. ABCD
- 2. ACD
- 3. ABD
- 4. ABCD
- 5. ABCD

第21章 组播路由协议

路由器在进行组播报文转发时，需要查找组播路由表，依照出接口列表将报文复制并发送。组播路由表是通过组播路由协议得到的。组播路由协议包含 DVMRP、MOSPF、PIM 等。

本章首先对组播路由协议进行了概述，然后对常用的 PIM-DM、PIM-SM、PIM-SSM 协议的机制进行详细介绍。

21.1 本章目标

课程目标


学习完本课程，您应该能够：

- 熟悉常用的组播路由协议
- 掌握PIM-DM协议的原理
- 掌握PIM-SM协议的原理
- 掌握PIM-SSM协议的原理



21.2 组播路由协议概述

组播路由协议概念和分类



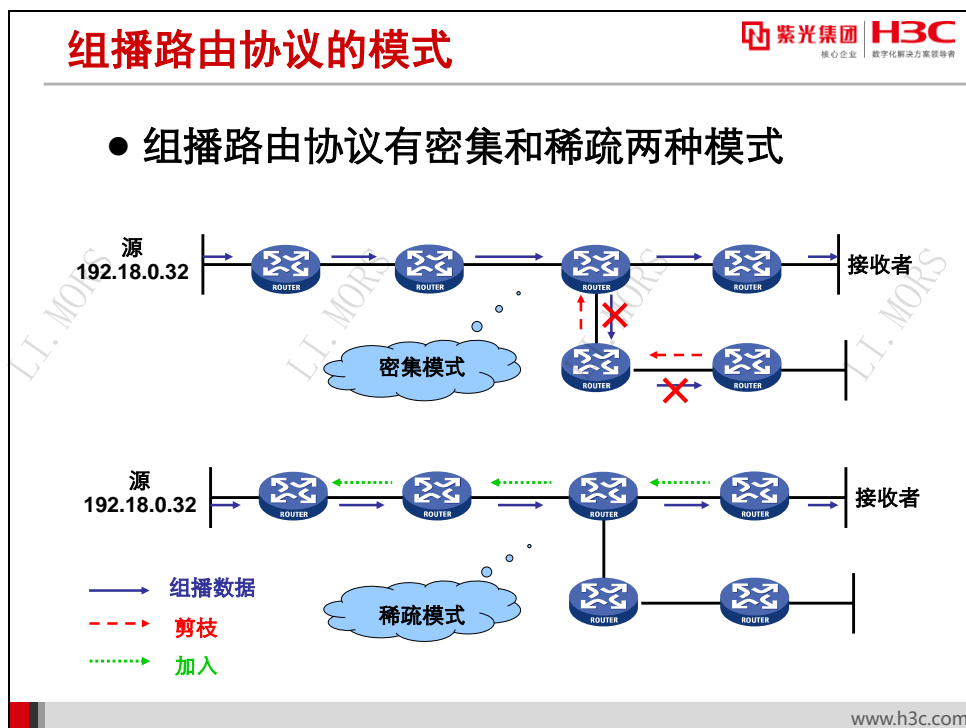
- 组播路由协议运行在三层组播设备之间，用于建立和维护组播路由，并正确、高效地转发组播报文。
- 组播路由协议建立了从一个数据源端到多个接收端的无环数据传输路径，即组播分发树。
- 对于ASM组播模型，组播路由协议可以分为域内和域间两类
 - 域内组播路由协议用于在AS内部发现组播源并构建组播分发树
 - 域间组播路由协议用于实现组播信息在AS之间的传递

www.h3c.com

组播路由协议运行在三层组播设备之间，用于建立和维护组播路由表，并正确、高效地转发组播报文。

组播路由协议为组播源和组播接收者建立了无环的传输路径，即组播分发树，不同的组播路由协议基于的组播分发树可能不同。

组播路由协议根据不同的组播模型可以分为基于 ASM（任意信源模型）的组播路由协议和基于 SSM（指定信源模型）的组播路由协议。对于 ASM 组播模型，组播路由协议又可以分为域内组播路由协议和域间组播路由协议。域内组播路由协议用于在 AS 内部发现组播源并构建组播分发树；域间组播路由协议用于实现组播信息在 AS 间的传递。



根据组播应用环境中组播接收者疏密程度的不同，组播路由协议有密集和稀疏两种模式。

- 在密集模式下，组播数据流采用“推”的方式从组播源泛洪发送到网络中的每一个角落，组播接收者采用被动接收的方式接收组播报文。如果某路由器所负责的所有网段均不存在接收者，则该路由器会向上游路由器发送请求，停止上游向自己泛洪组播报文。

密集模式实现非常简单，但是泛洪机制会浪费大量网络带宽。密集模式适用于网络环境中成员众多的场合，如股票交易大厅、学校网上教学等。

- 在稀疏模式下，组播数据流采用“拉”的方式从组播源发送到组播接收者，组播接收端路由器主动向组播源发送接收请求，组播报文只会发送到真正有接收需求的网段。

稀疏模式实现较为复杂，但是节省了大量网络带宽。稀疏模式适用于网络环境中接收成员较少的场合，如小区音/视频点播、IP 智能监控等。

域内组播路由协议

- **DVMRP (Distance Vector Multicast Routing Protocol, 距离矢量组播路由协议)**
- **MOSPF (Multicast Extensions to OSPF, 组播OSPF协议)**
- **PIM (Protocol Independent Multicast, 协议无关组播)**
 - PIM-DM
 - PIM-SM
 - PIM-SSM

域内组播路由协议主要包含 DVMRP、MOSPF 和 PIM。

DVMRP (Distance Vector Multicast Routing Protocol, 距离矢量组播路由协议) 基于距离矢量算法, 即 “RIP 模式”。DVMRP 路由协议运行时首先通过交互组播路由信息, 得到一颗经过裁剪的泛洪广播树, 树根为组播源, 所有 DVMRP 路由器都为叶子节点。所有路由器都拥有唯一的上游路由器指向组播源, 组播报文沿着这颗经过裁剪的广播树, 通过泛洪/剪枝机制到达组播接收者。DVMRP 具有 RIP 的所有缺点并且需要交互特定的组播路由信息, 实际应用中使用很少。

MOSPF (Multicast Extensions to OSPF, 组播 OSPF 协议) 是 OSPF 支持组播的扩展模式。MOSPF 基于链路状态算法, 通过交互特殊的 LSA——组成员 LSA, 使得路由器了解各组播组成员的位置, 当收到组播报文时可以沿最短路径树将组播报文发送到每一个组播接收者。MOSPF 采用链路状态算法, 可以清楚的了解每一个接收者的具体位置, 所以不会产生大量的泛洪组播流, 但是其运行环境受限, 只能运行在 OSPF 网络中, 实际应用中使用较少。

PIM (Protocol Independent Multicast, 协议无关组播) 需要使用单播路由表进行 RPF 检查, 但是不依赖于某种具体的单播路由协议来生成单播路由表。PIM 具有 DVMRP 和 MOSPF 无法比拟的优点, 相对于 DVMRP, PIM 不需要交互组播路由信息; 相对于 MOSPF, PIM 灵活简单, 因此 PIM 在实际组播应用中使用非常广泛。

PIM 协议报文基于 UDP 协议, 端口号是 103, 使用专门的组播 IP 地址 224.0.0.13。根据实现机制的不同, PIM 协议可以分为 PIM-DM、PIM-SM 和 PIM-SSM 三种。

21.3 PIM-DM

PIM-DM概述

- **PIM-DM**是密集模式的组播路由协议，该协议假定网络中组播接收者较多，且分布于大部分网络设备上，因此采用“推”的方式将组播流量周期性的扩散到网络中的所有邻居设备上
- **PIM-DM**在**RFC3973**中定义，适用于小型组播网络

紫光集团 H3C
核心企业 数字化解决方案领导者
www.h3c.com

PIM-DM（Protocol Independent Multicast-Dense Mode，协议无关组播—密集模式）在 RFC3973 中定义，通常适用于组播组成员相对比较密集的小型网络。

PIM-DM 使用“推（Push）模式”传送组播数据，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。

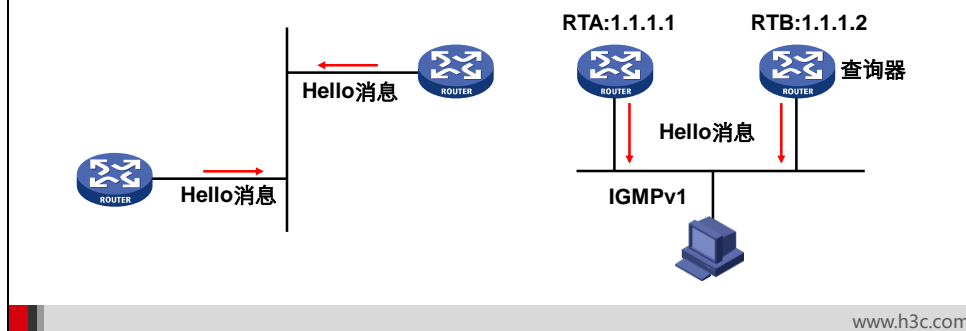
PIM-DM 还包含邻居发现、断言（Assert）、状态刷新等处理机制。

PIM-DM 基于 SPT 模型，通过 PIM-DM 可以构建以组播源为根、接收者为叶子的组播分发树，组播报文沿最优路径到达每一个接收者。

邻居发现机制

紫光集团 H3C
核心企业 数字化解决方案领导者

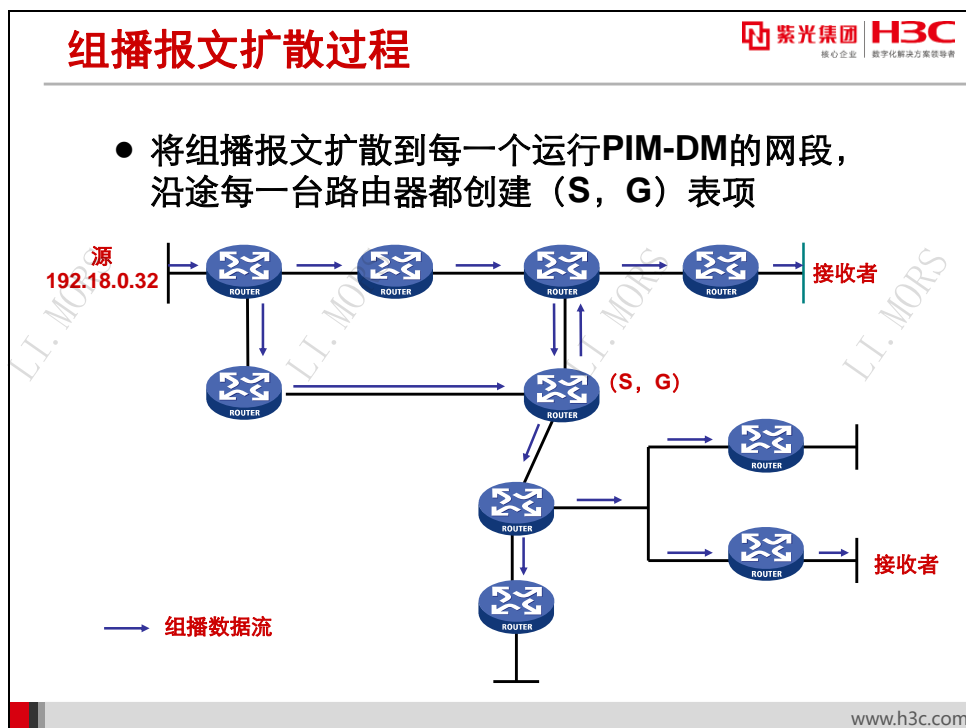
- PIM路由器之间周期性发送Hello消息，用于发现邻居、建立并维护邻居关系
- 当共享网段路由器运行的IGMP版本为v1时，可以通过Hello消息为该网段选举IGMP查询器



PIM 路由器周期性地以组播方式发送 PIM Hello 消息，目的组播地址为 224.0.0.13，所有 PIM 路由器都是该组播组的成员。通过 Hello 消息，路由器可以发现 PIM 邻居并建立和维护各路由器之间的 PIM 邻居关系。

当连接组播接收者的共享网段上有多台路由器，且路由器运行 IGMPv1 时，路由器可以通过 PIM 的 Hello 消息为该网段的接收者主机选举 DR，并由 DR 充当 IGMPv1 的查询器。PIM-DM 只有此时才需要选举 DR。

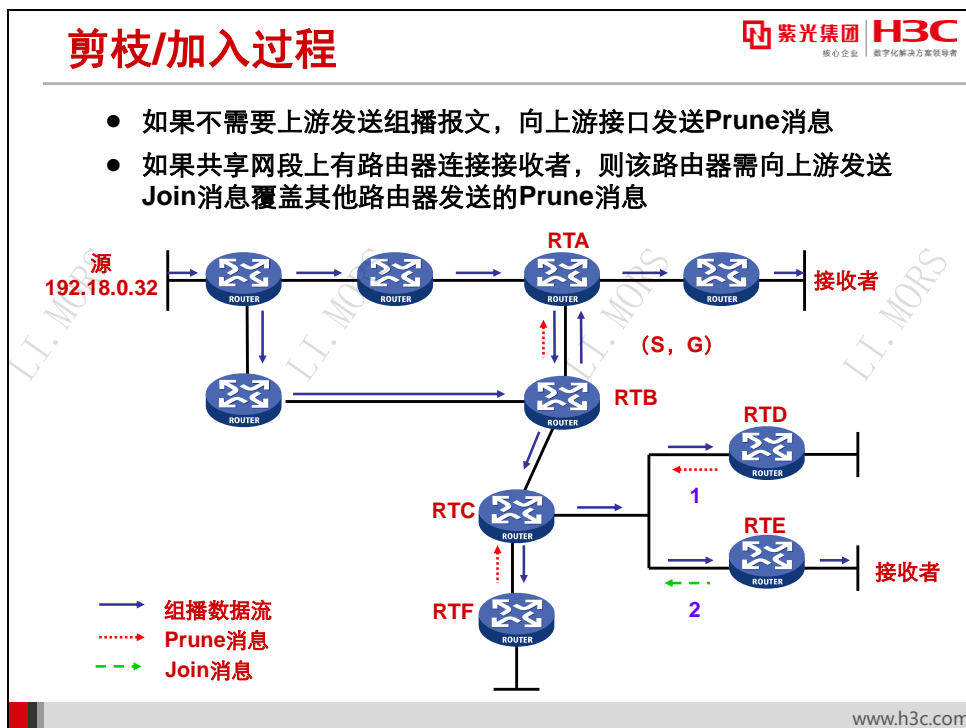
选举 DR 时，首先比较 Hello 消息的优先级，拥有最高 Hello 消息优先级的路由器将称为该网段的 DR；如果 Hello 消息的优先级相同，则通过比较路由器接口 IP 地址来竞选 DR，IP 地址最大的路由器将成为 DR。当 DR 出现故障时，其余路由器在超时后没有收到 DR 发送的 Hello 消息，则会触发新的 DR 选举过程。



当收到组播报文时，PIM-DM 路由器首先进行 RPF 检查，确定报文是否从正确的接口到达。如果 RPF 检查正确，路由器会为该组播报文建立组播转发表项 (S, G)，S 代表发送源，G 代表组播组。该组播转发表项还包含一个入接口和出接口列表，组播转发表项的入接口即为 RPF 接口，出接口列表包含除 RPF 接口外的所有连接 PIM-DM 邻居的接口以及直接连接组播组 G 成员主机的接口。

网络中的每一台 PIM 路由器都进行同样的处理，经过扩散 (Flooding) 过程，PIM-DM 域内的每台路由器都会收到来自源 S 的组播报文，并创建 (S, G) 表项。

如图所示，假设所有路由器均运行 PIM-DM，组播源 192.168.0.32 发送组播报文到直连的路由器，路由器经过 RPF 检查后，将组播报文从非 RPF 接口发送出去，沿途每一台路由器都进行了同样的处理过程。最后，网络中每一台路由器都收到了来自 192.168.0.32 的组播报文，并建立 (192.168.0.32, G) 表项。如果路由器本地有组播组 G 的成员，则路由器会将组播报文发送给接收者。



如果路由器不需要 PIM 邻居发送组播报文给自己，则发送 Prune 消息给 PIM 邻居，PIM 邻居收到 Prune（剪枝）消息后将不再转发该组播组的报文到本路由器。

路由器有两种情况需要发送 Prune 消息：

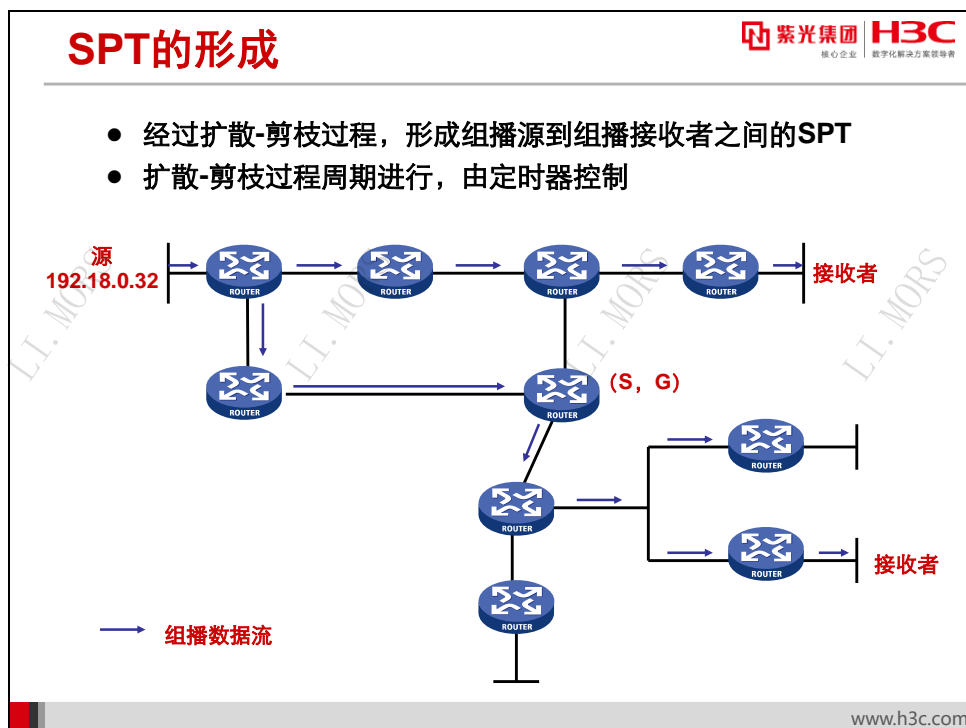
- 如果路由器维护的 (S, G) 表项中出接口列表为空，表示路由器下游没有组播接收者，则路由器会向上游 RPF 邻居发送 Prune 消息。

例如，RTD 和 RTF 会向 RPF 邻居 RTC 发送 Prune 消息，剪掉不需要的组播流。

- 如果路由器从非 RPF 接口收到组播报文，则会触发 Assert（断言）过程，Assert 失败的一方也会向获胜的一方发送 Prune 消息。例如，RTB 从非 RPF 接口（连接 RTA 的接口）收到组播报文，会向 RTA 发送 Prune 消息。Assert 机制在后续内容中会详细介绍。

如果有多台路由器通过共享网段互联，则一台路由器发送的 Prune 消息可能会错误的导致其他路由器无法收到组播报文。例如，RTD 下游没有接收者，所以 RTD 向 RPF 邻居 RTC 发送 Prune 消息，但是同一网段的路由器 RTE 的下游存在接收者，如果 RTC 停止向该网段发送组播报文，将导致 RTE 无法收到并转发报文给接收者。

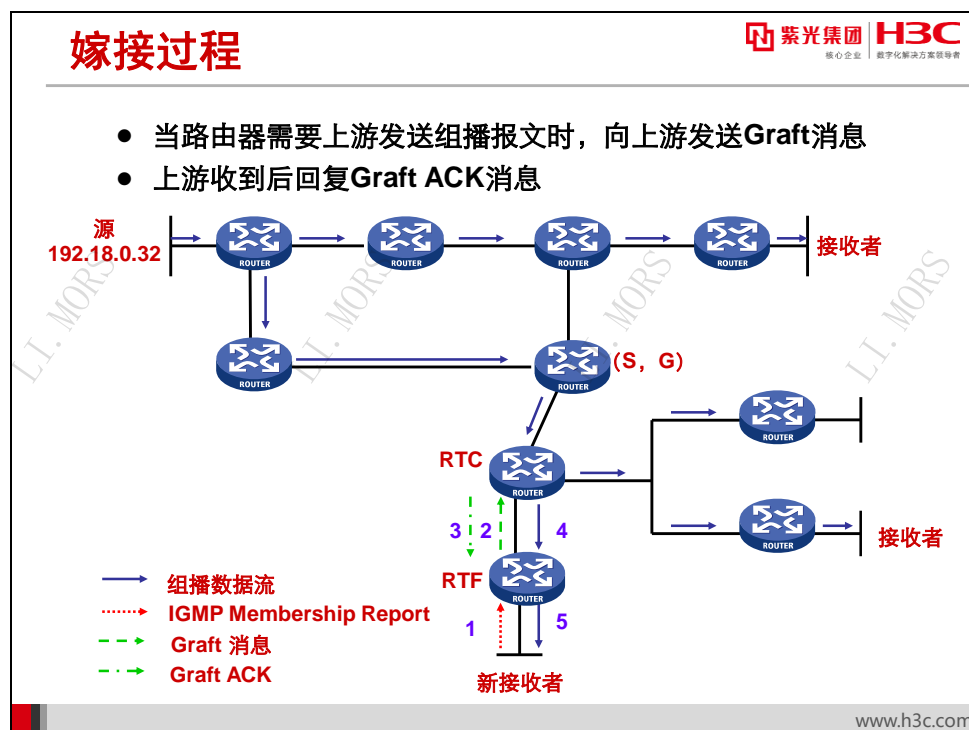
针对这种问题，PIM-DM 定义了加入机制：如果路由器从 RPF 接口收到 Prune 消息，且路由器下游存在接收者时，路由器要向 RPF 邻居发送 Join 消息，使得上游 RPF 邻居继续维持组播报文在该共享网段的转发。在本例中，RTE 从 RPF 接口收到 RTD 发送的 Prune 消息，则 RTE 会从 RPF 接口发送 Join 消息给 RTC，使得 RTC 继续转发组播报文到共享网段。这是 PIM-DM 中唯一使用到 Join 消息的地方。



经过“扩散—剪枝”过程，最终形成组播源和组播接收者之间的 SPT。

网络中的所有路由器都会一直维护 (S, G) 表项，只有当组播源 S 不再发送组播流之后，路由器才会删除该组播源对应的 (S, G) 表项信息。

PIM-DM 的“扩散—剪枝”过程是周期性发生的。各个被剪枝的接口提供超时机制，当剪枝超时后接口会重新向下游发送组播报文，这个“扩散—剪枝”过程会一直持续下去。



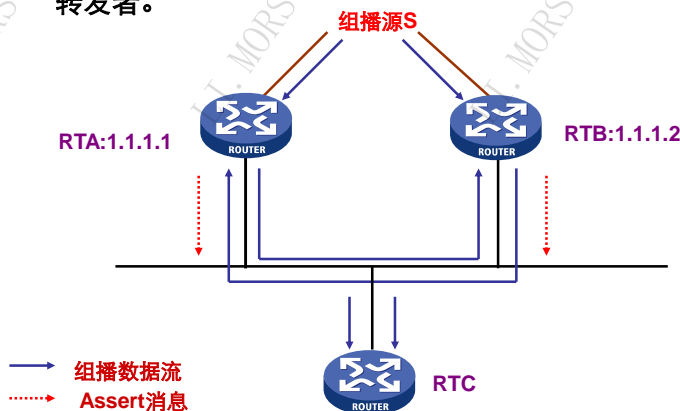
当被剪枝的路由器上出现组播接收者时，路由器不必等待下一次扩散的到来，而可以通过嫁接（Graft）机制快速的恢复上游 RPF 邻居对组播报文的转发。

例如，RTF 下有接收者想要加入组播组 G，则该接收者会向 RTF 发送 IGMP Membership Report 报文，RTF 收到 Membership Report 报文后会建立组播组成员表项，并向 RPF 邻居 RTC 发送 Graft 消息，请求 RTC 发送组播流。RTC 收到 RTF 发送的 Graft 消息后，向 RTC 回复 Graft ACK，告诉 RTC 已经收到了 Graft 消息。然后，RTC 会将收到 Graft 消息的接口添加到 (S, G) 表项的出接口列表中。之后，组播数据就会被 RTC 转发到 RTF，而接收者也就能从 RTF 收到组播数据了。

嫁接过程中，Graft 消息和 Graft ACK 都是单播方式发送的，其中 Graft 消息的单播目的地址为 RPF 邻居的 IP 地址，RPF 邻居回复的 Graft ACK 的目的地址为本路由器发送嫁接消息的源 IP 地址。

断言机制

- 在一个网段内如果存在多台组播路由器，则相同的组播报文可能会被重复发送到该网段。
- 通过断言（Assert）机制可以选定网段上唯一的组播数据转发者。



在一个网段内如果存在多台组播路由器，则相同的组播报文可能会被重复发送到该网段。为了避免出现这种情况，需要通过断言（Assert）机制来选择网段上唯一的组播数据转发者。

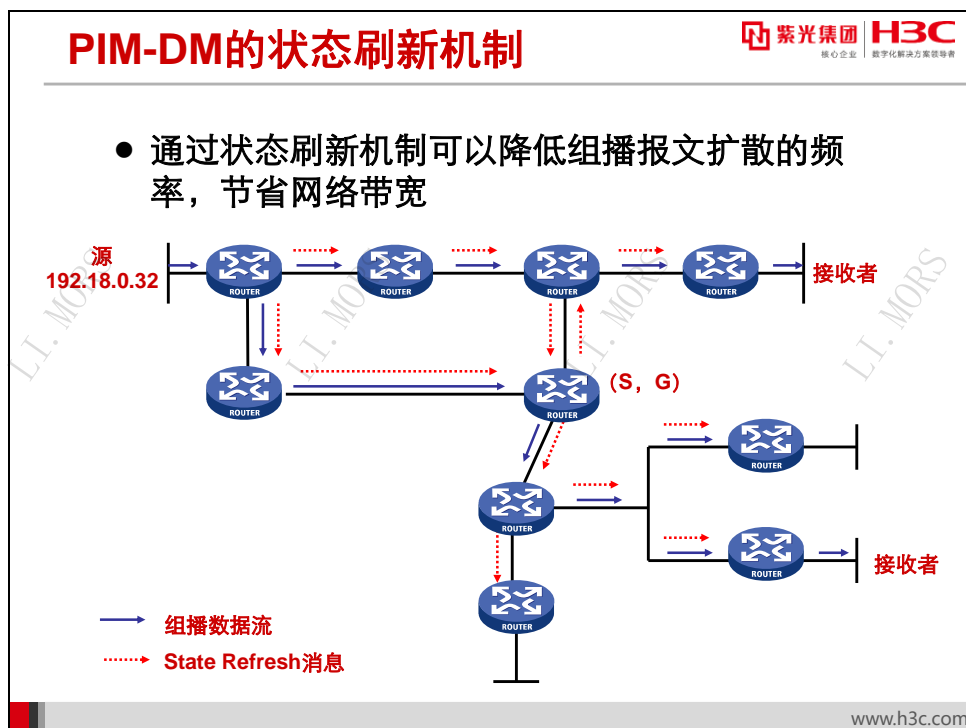
图中，当 RTA 和 RTB 从上游节点收到来自源 S 的组播报文后，都会向本地网段转发该报文，于是处于下游的节点 RTC 就会收到两份相同的组播报文，RTA 和 RTB 也会从各自的本地接口收到对方转发来的该组播报文。

此时，RTA 和 RTB 会在收到重复组播报文的接口上发送 Assert 消息，Assert 消息中携带组播源地址 S、组播组地址 G、到达组播源的单播路由的优先级和 Metric。通过比较 Assert 消息后，RTA 和 RTB 中的获胜者将成为组播报文在该网段的唯一转发者。Assert 消息的比较规则如下：

- 到组播源 S 的单播路由的优先级较高者获胜；
- 如果到组播源 S 的单播路由的优先级相等，那么到组播源的 Metric 值较小者获胜；
- 如果到组播源 S 的 Metric 值也相等，则本地接口 IP 地址较大者获胜。

假设 RTA 和 RTB 到达组播源 S 的单播路由由优先级和度量值都相同，因为 RTB 的 IP 地址大于 RTA 的 IP 地址，RTB 为断言获胜者。作为断言失败者，RTA 会向 RTB 发送 Prune 消息，由于该网段上存在接收者 RTC，当 RTC 收到 Prune 消息后，会发送 Join 消息，使得 RTB 继续向该网段转发组播报文。

当 RTB 上游接口中断时，为避免组播业务长期中断，RTB 会立即发送 Metric 值为无穷大的 Assert 消息，从而引发新一轮的断言过程，使得 RTA 成为该网段新的断言获胜者。



通过“扩散—剪枝”过程可以快速、简便的建立从组播源到组播接收者之间的 SPT。但是由于该过程是周期进行的，每隔一段时间，网络中就会充斥组播报文，而不需要接收组播报文的路由器就需要再次申请剪枝，重复的“扩散—剪枝”过程浪费了大量的网络带宽，增加了路由器的处理负担。

针对上述问题，PIM-DM 提供了状态刷新机制。和组播源直接相连的路由器发出 State Refresh 消息，其他路由器收到 State Refresh 消息后会重置剪枝超时定时器，同时向入接口之外的所有连接 PIM 邻居的接口发送 State Refresh 消息。其中，对于当前处于转发状态的接口，State Refresh 消息中的剪枝标志位置 0；对于当前处于剪枝状态的接口，State Refresh 消息中的剪枝标志位置 1。

通过周期性的发送 State Refresh 消息，可以使得处于剪枝状态的接口维持在当前状态，从而减少不必要的扩散。

状态刷新机制，使用周期性的协议报文替代周期性的组播数据扩散，较少网络消耗，优化了网络资源。

21.4 PIM-SM

PIM-SM概述



- **PIM-SM是稀疏模式的组播路由协议。**
PIM-SM采用“拉”的方式，根据接收者的需求，在组播接收者和组播源之间建立组播转发树
- **RFC4601中定义，实际应用中适用于任何形式的网络**
- **PIM-SM使用RP（Rendezvous Point，汇聚点）作为共享树的根**
 - 组播源通过DR注册到RP
 - 接收者通过DR向RP发起加入

www.h3c.com

PIM-SM（Protocol Independent Multicast-Sparse Mode，协议无关组播—稀疏模式）在RFC4601中定义，在实际应用中适用于任何形式的网络。

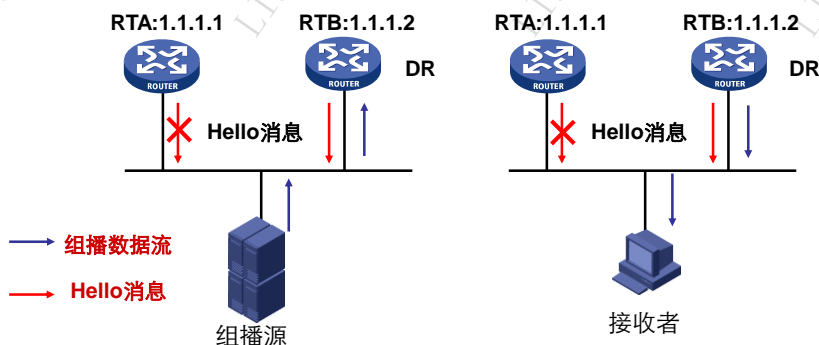
PIM-SM使用“拉（Pull）模式”传送组播数据，其基本原理如下：

- PIM-SM 假设所有主机都不需要接收组播数据，只向明确提出需要组播数据的主机转发。
- 使用 RP 作为共享树的根。连接接收者的 DR 向某组播组对应的 RP 发送加入消息（Join），该消息被逐跳送达 RP，所经过的路径就形成了 RPT 的分支。
- 组播源如果要向某组播组发送组播数据，首先由组播源侧 DR 负责向 RP 进行注册，然后组播源把组播数据发向 RP，当组播数据到达 RP 后，被复制并沿着 RPT 发送给接收者。

邻居发现和DR选举

紫光集团 H3C
核心企业 数字化转型领导者

- PIM-SM邻居发现过程和PIM-DM相同
- 共享网段需要选举DR，作为该网段的唯一转发者



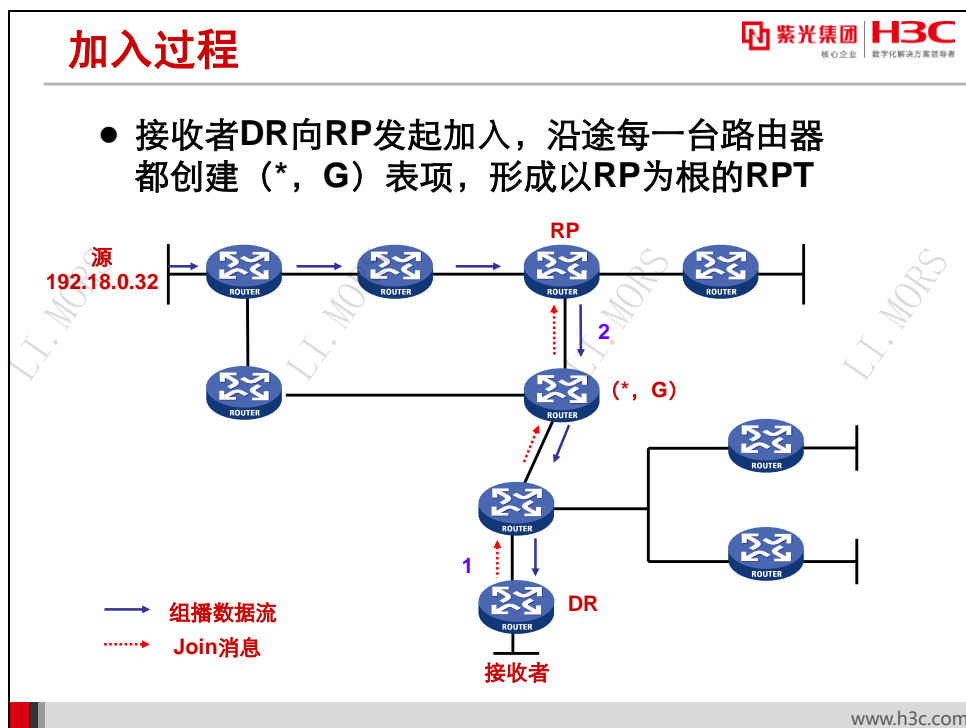
www.h3c.com

PIM-SM 邻居发现过程和 PIM-DM 相同，都是通过以组播方式发送 Hello 报文，从而建立和维护 PIM 邻居关系。

在 PIM-DM 中，只有当接收者侧有多台路由器连接到共享网段，且路由器运行 IGMPv1 时，才需要进行 DR 的选举。在 PIM-SM 中，无论路由器运行什么样的 IGMP 版本，接收者侧或发送源侧的路由器都需要进行 DR 的选举，选举出的 DR 将作为该共享网络中组播数据的唯一转发者。其中，接收者侧的 DR 负责向 RP 发送加入报文；组播源侧的 DR 负责向 RP 发送注册报文。

DR 的选举方式和 PIM-DM 中介绍的选举方式相同。首先，比较 Hello 报文中的优先级，拥有最高 Hello 报文优先级的路由器将成为 DR；如果优先级相同，则需要比较路由器接口 IP 地址的大小，IP 地址最大的路由器将成为 DR。

接收者侧的 DR 必须运行 IGMP，否则连接在该 DR 上的接收者将不能通过该 DR 加入组播组。

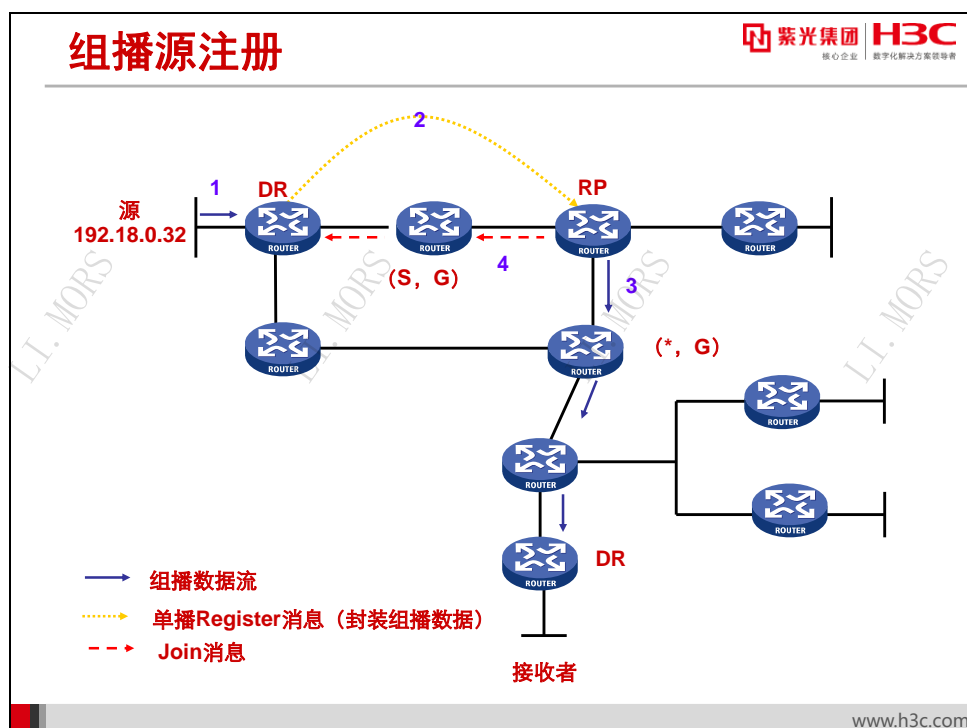


接收者首先通过 IGMP Report 报文通知 DR 加入某组播组，DR 会在本地建立 (*, G) 表项，然后向 RPF 邻居发送 Join 消息，此时 DR 会将连接接收者的接口加入到 (*, G) 表项的出接口列表中，同时将指向 RP 的单播路由的下一跳出接口作为该 (*, G) 表项的入接口。由于接收者并没有指定组播源，因此 (*, G) 表项中的组播源为任意组播源。

Join 消息将沿着 DR 指向 RP 的单播路由逐跳发送，最后到达 RP。沿途每一台路由器包括 RP 都会建立对应的 (*, G) 表项。网络中每一个接收者侧的 DR 都进行相同的操作，最终形成以 RP 为根的 RPT。

当 RP 收到组播组 G 的报文时，从 (*, G) 表项的出接口将报文转发出去，组播报文沿 RPT 传送，最终由接收者侧的 DR 转发给接收者。

只要 DR 本地有组播组 G 的接收者，DR 就会周期性的向上游发送加入消息，保证上游路由器正常发送组播报文到本地。当 DR 相连的所有网段都没有组播组 G 的接收者时，DR 会向上游路由器发送剪枝消息，通知上游停止向自己发送组播组 G 的报文。

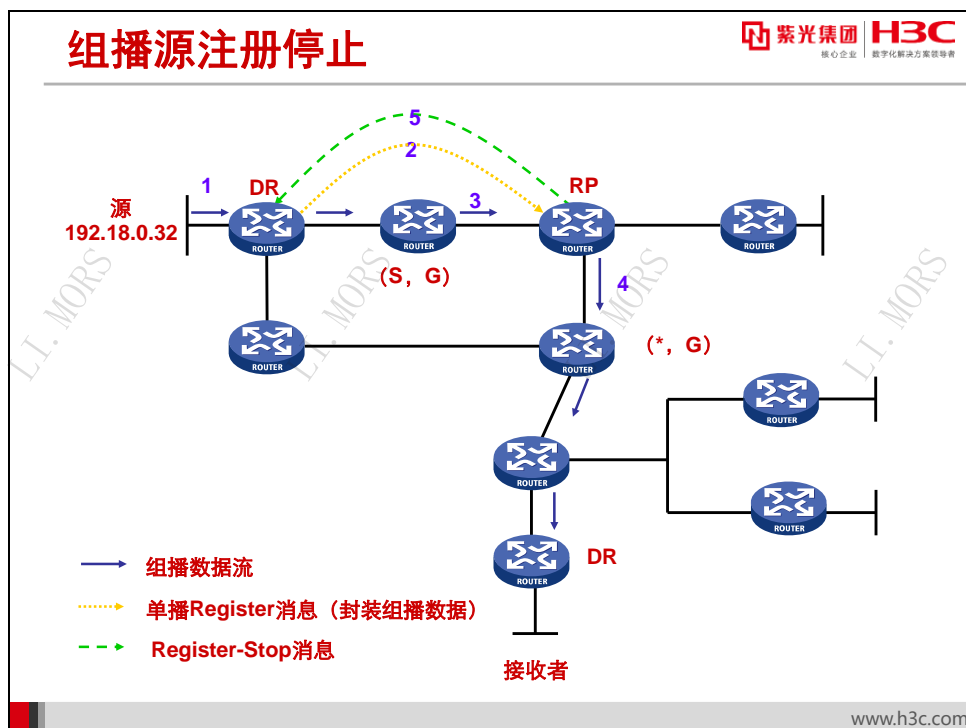


组播源不关心有没有接收者，会一直向某组播组 G 发送组播数据。当组播源侧的 DR 收到组播组 G 的报文时，会建立 (S, G) 表项，同时将连接组播源的接口作为 (S, G) 表项的入接口。

DR 将收到的组播报文封装在 Register 消息中，单播发送到对应的 RP。RP 收到单播 Register 消息后，判断本地是否存在对应的 $(*, G)$ 表项。如果存在 $(*, G)$ 表项，RP 将 Register 消息中封装的组播报文取出，然后从 $(*, G)$ 表项的出接口发送出去。

RP 会为该组播源创建对应的 (S, G) 表项，入接口为收到单播 Register 报文的接口，出接口和 $(*, G)$ 表项中的出接口相同。然后，RP 向组播源方向发送特定源和组的 Join 消息，Join 消息将沿着 RP 指向组播源的单播路由逐跳发送，最后到达组播源侧的 DR。沿途每一台路由器都会建立 (S, G) 表项。

组播源侧的 DR 收到 Join 消息后，将接收 Join 消息的接口添加到 (S, G) 表项的出接口列表中，此时，形成了组播源到 RP 之间的一颗 SPT。

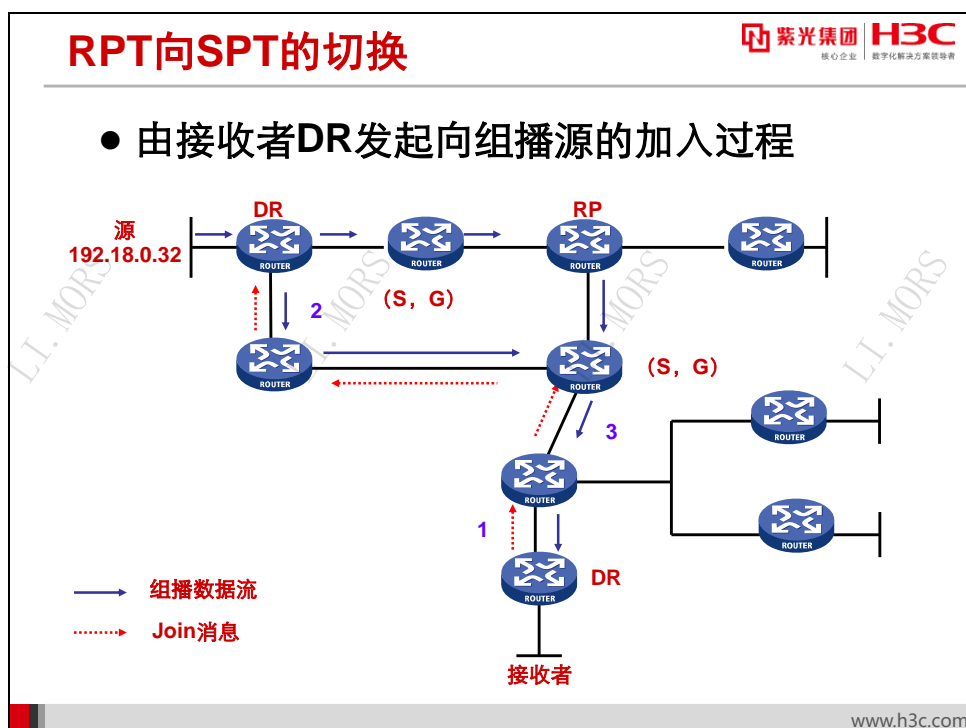


此时，组播源发出的组播报文会通过两种方式发送到 RP。一种是由 DR 进行注册封装，单播发送到 RP；一种是通过 DR 本地的 (S, G) 表项沿着 SPT 组播发送到 RP。

当单播和组播发送的报文到达 RP 时，RP 将丢弃单播注册封装的报文，而将以组播方式接收的报文发送到 RPT 上。同时，RP 会向组播源侧的 DR 单播发送 Register-Stop 消息，通知 DR 停止单播注册封装过程。此后，组播报文将仅沿着 SPT 发送到 RP，再由 RP 发送到 RPT 上。

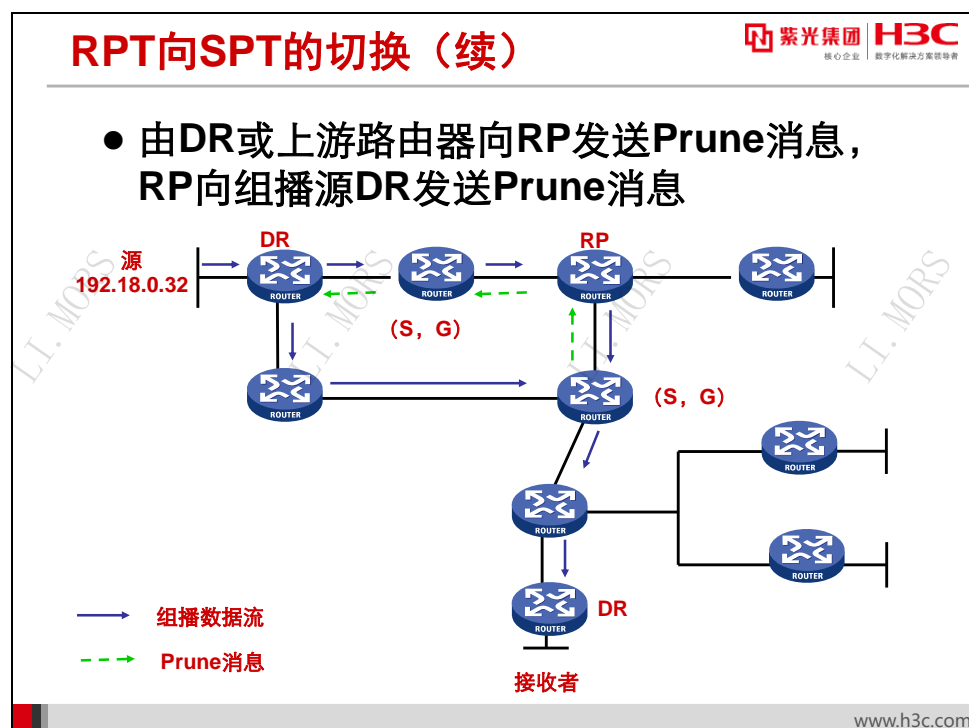
组播源的 DR 路由器会维护一个注册抑制定时器，在定时器时间范围内，DR 都以组播方式发送数据，而不去封装注册报文。在定时器超时前，DR 会发送若干空 Register 消息，该消息没有封装任何组播数据，仅仅用来提醒 RP：该发送 Register-Stop 消息了，否则 DR 会重新通过单播发送注册封装报文。RP 收到空 Register 消息后会发送 Register-Stop 消息，同时更新本地的 (S, G) 表项，而 DR 收到 Register-Stop 消息后会更新注册抑制定时器，并继续以组播方式发送数据。

至此为止，网络中形成两棵树，一颗是组播源到 RP 的 SPT，一颗是从 RP 到组播接收者的 RPT。可以看到组播源到 RP 所经过的路径是最优的，但是从组播源到接收者之间的路径不一定是最优的，因为必须通过 RP 转发。当网络中组播源和接收者数量众多时，RP 将会成为网络性能的瓶颈，并且降低了网络带宽使用率，增加了报文转发的路径延迟。



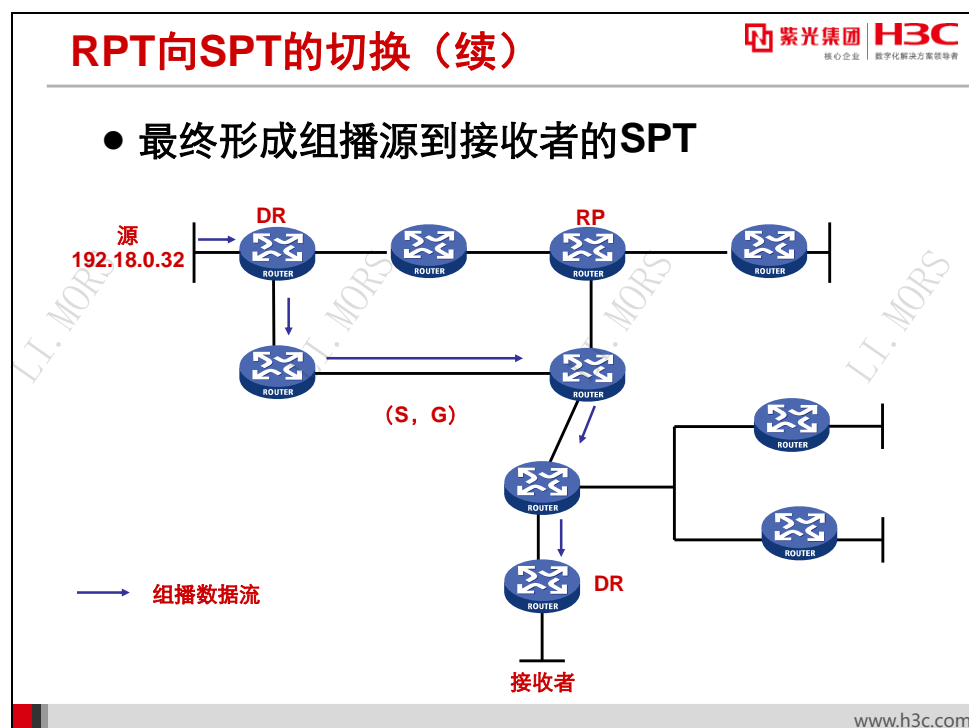
为了获得更小的报文延迟，接收者侧的 DR 会发起从 RPT 到 SPT 的切换。

当 DR 收到第一个组播报文时，就可以获取组播源的地址，然后向组播源方向发送特定源组的 Join 消息，Join 消息将沿着 DR 指向组播源的单播路由逐跳发送，最后到达组播源侧的 DR 或已经存在 (S, G) 表项的路由器。沿途每一台路由器都会建立 (S, G) 表项。



此后，组播源发送的报文将会沿两个方向到达接收者侧的 DR 或 DR 上游的某台路由器。其中，一封报文沿 SPT 到达，一封报文沿 RPT 到达，收到两封报文的 router 会丢弃从 RP 方向收到的组播报文，并向 RP 发送 Prune 消息。

如果 RP 本地 (S, G) 表项的出接口列表为空，则 RP 向组播源侧的 DR 发送 Prune 消息，剪掉不需要的组播流。



最终，网络中形成组播源和接收者之间的 SPT。

和 PIM-DM 相比，PIM-SM 中的 SPT 基于接收者主动的加入，所以不需要泛洪报文到网络中的每一个角落，节省了大量的网络带宽。和 RPT 相比，组播报文经过最优路径到达接收者，不需要 RP 的中转，提高了报文转发的效率。

RP发现

- **RP是PIM-SM运行的关键，RP可以通过静态指定也可以通过动态选举**
- **静态指定**
 - 网络中所有PIM路由器上静态指定RP，无法实现RP的冗余备份
- **自动选举**
 - 需在网络中使用Boot Strap协议，自动发现、选择RP，易于扩展，组网更灵活

PIM-SM 中，组播组加入请求是由接收者端主动发起的，而接收者初始时候，并不知道组播源的位置，因此需要 RP 作为接收者加入某组播组的汇聚点；组播源起始的时候也不知道接收者在什么地方，所以也将组播报文发给 RP，由 RP 中转。由此可见 RP 是 PIM-SM 中的核心设备，RP 的选择是 PIM-SM 运行的关键。

在结构简单的小型网络中，组播信息量少，整个网络仅依靠一个 RP 进行组播信息的转发即可，此时可以在 PIM-SM 域中的各路由器上静态指定 RP 的位置；但是在更多的情况下，PIM-SM 域的规模都很大，通过 RP 转发的组播信息量巨大。

为了缓解 RP 的负担并优化 RPT 的拓扑结构，可以在 PIM-SM 域中配置多个 C-RP（Candidate-RP，候选 RP），通过 Bootstrap 机制来动态选举 RP，使不同的 RP 服务于不同的组播组，此时需要配置 BSR（Bootstrap Router，自举路由器）。

BSR 是 PIM-SM 域的管理核心，一个 PIM-SM 域内只能有一个 BSR，但可以配置多个 C-BSR（Candidate-BSR，候选 BSR）。这样，一旦 BSR 发生故障，其余 C-BSR 能够通过自动选举产生新的 BSR，从而确保业务免受中断。

BSR机制



- BSR负责收集网络中由C-RP发来的宣告报文，汇总为RP-Set，通过自举消息发送到PIM-SM域
- 网络中的各路由器将依据RP-Set提供的信息，使用相同的规则从众多C-RP中为特定组播组选择其对应的RP
 - 首先比较C-RP的优先级，优先级较高者获胜
 - 若优先级相同，则使用哈希（Hash）函数计算哈希值，该值较大者获胜
 - 若优先级和哈希值都相同，则C-RP地址较大者获胜

www.h3c.com

C-RP 路由器周期性的以单播方式发送宣告报文（Advertisement）给 PIM-SM 域中的 BSR，该报文中携带有 C-RP 的地址和优先级以及其服务的组播组范围。BSR 收到宣告报文后，将这些信息汇总为 RP-Set（RP 集，即组播组与 RP 的映射关系数据库），封装在自举报文（Bootstrap）中并以组播方式发布到整个 PIM-SM 域。

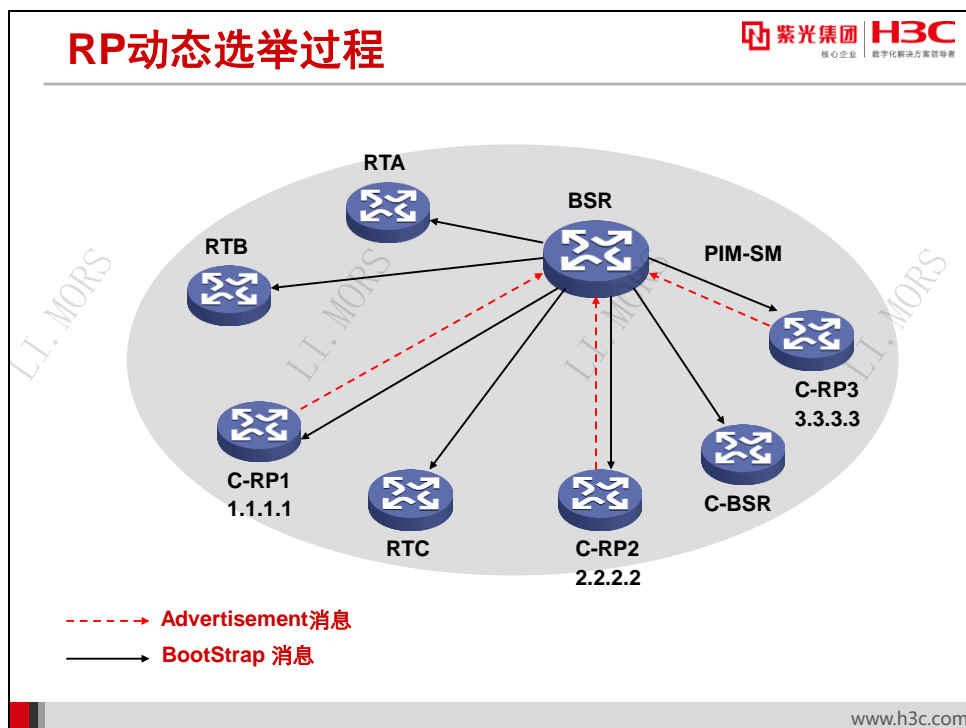
网络中的各路由器将依据 RP-Set 提供的信息，使用相同的规则从众多 C-RP 中为特定组播组选择其对应的 RP，具体规则如下：

- 首先比较 C-RP 的优先级，优先级较高者获胜；
- 若优先级相同，则使用哈希（Hash）函数计算哈希值，该值较大者获胜；
- 若优先级和哈希值都相同，则 C-RP 地址较大者获胜。

在一个 PIM-SM 域中只能有一个 BSR，由 C-BSR 通过自动选举产生。C-BSR 间的自动选举机制描述如下：

- 最初，每个 C-BSR 都认为自己是本 PIM-SM 域的 BSR，并使用接口的 IP 地址作为 BSR 地址，发送自举报文；
- 当某 C-BSR 收到其它 C-BSR 发来的自举报文时，首先比较自己与后者的优先级，优先级较高者获胜；在优先级相同的情况下，再比较自己与后者的 BSR 地址，拥有较大 IP 地址者获胜。

实际应用中，RP 和 BSR 可以是同一台路由器。




图中，PIM-SM 域有三台路由器作为 C-RP，C-RP 周期性的向网络中的 BSR 单播发送宣告报文。BSR 收到宣告报文后，将宣告报文中的 C-RP 地址、优先级以及其服务的组播组范围等信息汇总为 RP-Set，通过自举报文以组播方式发布给 PIM-SM 中的所有路由器。

其他路由器收到自举报文后为组播组选择对应的 RP。一个 RP 可以为多个组播组服务，而一个组播组只能选择一个 RP。

最后 RTA 选择 C-RP1 作为组播组 228.1.1.1 的 RP，RTB 选择 C-RP2 作为组播组 228.2.2.2 的 RP，RTC 选择 C-RP3 作为组播组 228.3.3.3 的 RP，实现了 RP 间的负载分担。

21.5 PIM-SSM

PIM-SSM概述


核心企业 | 数字化解决方案领导者

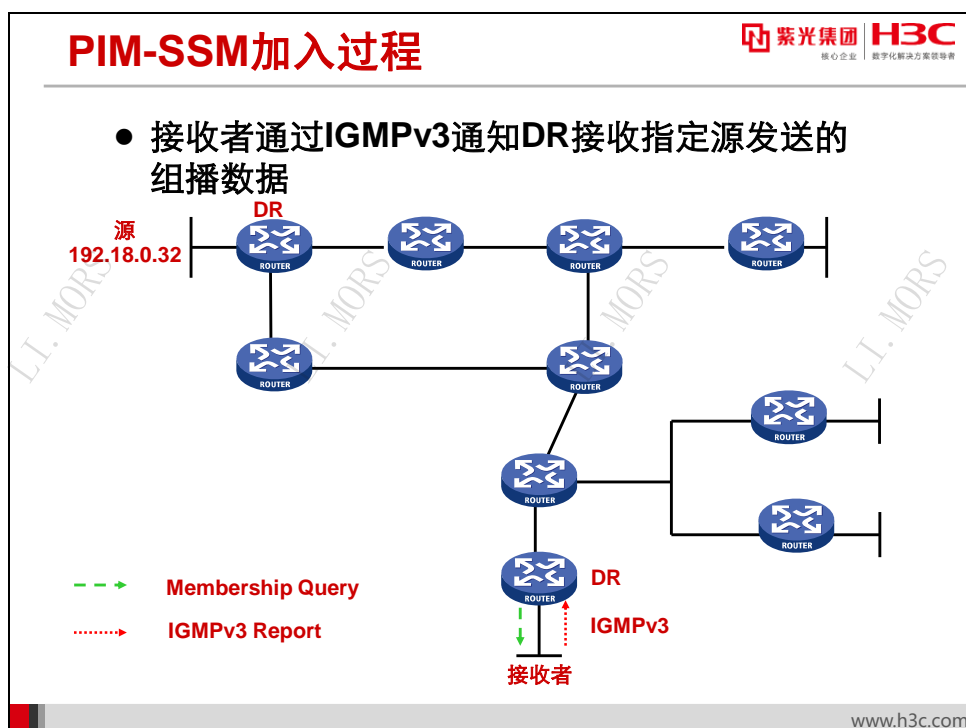
- **SSM模型在RFC4607中定义，为指定源组播提供了解决方案**
- **IANA为SSM分配了特定的组播地址段：232.0.0.0-232.255.255.255**
- **采用PIM-SM的一部分技术可以实现PIM的SSM模型即PIM-SSM**

www.h3c.com

组播的 SSM 模型在 RFC4607 中定义，SSM 模型为指定源组播提供了解决方案。IANA 为 SSM 分配了特定的组播地址段：232.0.0.0-232.255.255.255。

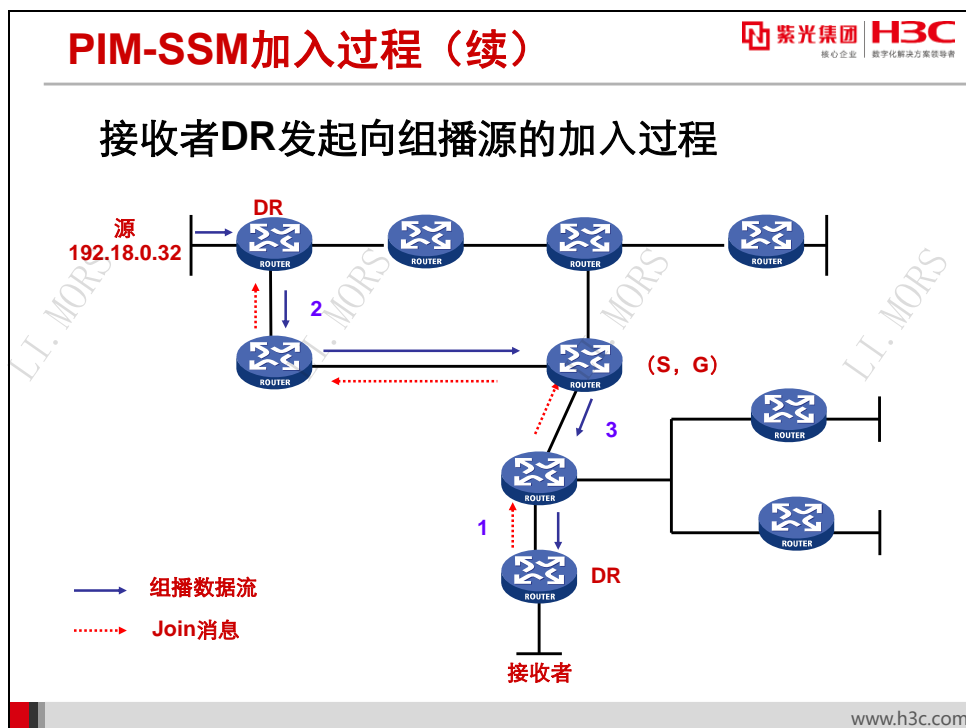
PIM-SSM 通过 PIM-SM 的一部分机制来实现，可以看作是 PIM-SM 的子集。PIM-SSM 的机制包含邻居发现和 DR 选举以及加入过程。

PIM-SSM 的邻居发现和 DR 选举机制与 PIM-SM 的对应机制相同。由于接收者事先已经通过某种途径知道了组播源的 IP 地址，所以 PIM-SSM 的加入过程和 PIM-SM 的加入过程相比有所区别。



主机端的 DR 周期性的发送 IGMPv3 查询报文，如果网段上有主机想要接收某特定源的组播报文，则主机回复 IGMPv3 Report 报文，包含想要接收的组播源地址 S 以及组播组地址 G。

当主机想要接收某组播源发送的组播报文时，也可以主动发送 IGMPv3 Report 报文，向 DR 发起请求。



DR 收到主机发送的 Report 报文，判断主机想要加入的组播组 G 是否在 SSM 模型规定的 232.0.0.0-232.255.255.255 范围之内，如果组播组 G 在 SSM 模型范围之内，则 DR 会维护组成员信息，并建立 (S, G) 表项。

DR 向组播源 S 发起加入请求，加入消息将沿着 DR 指向组播源 S 的单播路由逐跳发送，最后到达组播源侧的 DR。沿途每一台路由器都会建立 (S, G) 表项。从而构建了一颗从组播源到接收者的 SPT。

此后，组播报文就会沿着组播源到接收者之间的 **SPT** 到达接收者。

PIM-SSM 模式中，由于主机事先知道了组播源的位置，可以直接向组播源发起加入，所以不需要在网络中选举 RP，简化了协议的处理。并且组播报文开始就沿着 SPT 传送，不需要 RPT 到 SPT 的切换，提高了报文转发的效率。

21.6 本章总结

本章总结

- 对组播路由协议进行概述
- 详细介绍了PIM-DM协议原理
- 详细介绍了PIM-SM协议原理
- 详细介绍了PIM-SSM协议原理

www.h3c.com

21.7 习题和解答

21.7.1 习题

1. 关于组播路由协议分类，正确的有（ ）
 - A. 组播路由协议有密集和稀疏两种模式
 - B. 根据应用范围，组播路由协议分为域内组播路由协议和域间组播路由协议
 - C. 域内组播路由协议有 DVMRP、MOSPF、PIM
 - D. 域间组播路由协议有 MSDP
2. PIM-DM 的机制包含（ ）
 - A. 扩散-剪枝机制
 - B. 嫁接机制
 - C. 断言机制
 - D. 状态刷新机制
3. PIM-SM 的机制包含（ ）
 - A. 加入机制
 - B. 组播源注册机制
 - C. RPT 向 SPT 切换机制
 - D. RP 发现机制
4. RP 动态选举规则为（ ）
 - A. 首先比较 C-RP 的优先级，优先级较高者获胜
 - B. 若优先级相同，则使用哈希（Hash）函数计算哈希值，该值较大者获胜
 - C. 若优先级和哈希值都相同，则 C-RP 地址较大者获胜
 - D. 若优先级和哈希值都相同，则 C-RP 地址较小者获胜
5. 关于 PIM-SSM，正确的有（ ）
 - A. 组播的 SSM 模型在 RFC4607 中定义，SSM 模型为指定源组播提供了解决方案
 - B. PIM-SSM 的机制包含邻居发现和 DR 选举以及加入过程
 - C. PIM-SSM 的邻居发现和 DR 选举机制和 PIM-SM 的对应机制相同
 - D. PIM-SSM 需要和 IGMPv3 配合使用

21.7.2 习题答案

1. ABCD 2. ABCD 3. ABCD 4. ABC 5. ABCD

第22章 组播配置和维护

本章介绍了组播的配置命令、维护命令并给出配置实例。

22.1 本章目标

课程目标

● 学习完本课程，您应该能够：

- 掌握组播基本配置命令
- 掌握组播维护命令



www.h3c.com

22.2 组播配置命令

组播基本配置

- 全局启动组播应用
`[H3C] multicast routing-enable`
- 进入IGMP视图
`[H3C] igmp`
- 配置IGMP版本
`[H3C-Vlan-interface100] igmp version version-number`



紫光集团 H3C
核心企业 数字化解决方案领导者

www.h3c.com

在配置各项组播功能之前，应该首先使能 IP 组播路由，否则部分其他组播配置不生效，设备不转发任何组播报文。

全局使能组播的命令为：

```
[H3C] multicast routing
```

igmp 命令用来进入 IGMP 视图。IGMP 视图下的全局配置对所有接口都有效，接口视图下的配置只对当前接口有效，后者的配置优先级较高。

进入 IGMP 视图的命令为：

```
[H3C] igmp
```

IGMP 缺省使能的版本为 IGMPv2，可在接口视图下指定 IGMP 版本。在接口视图下指定 IGMP 版本的命令为：

```
[H3C-Vlan-interface100]igmp version version-number
```

IGMP配置命令（续）



- 接口视图使能IGMP

```
[H3C-Vlan-interface100] igmp enable
```

- 配置发送普遍查询报文的时间间隔

```
[H3C-Vlan-interface100] igmp query-interval  
interval
```

- 配置普遍查询报文最大响应时间

```
[H3C-igmp] max-response-time interval
```

www.h3c.com

只有在接口上使能了 IGMP，在该接口上对其它 IGMP 特性所作的配置才能生效，所以首先需要使能接口的 IGMP 功能。

在接口使能 IGMP 的命令如下，只有先使能了 IP 组播路由，本命令才能生效。

```
[H3C-Vlan-interface100]igmp enable
```

当 IGMP 查询器启动后，会周期性地发送 IGMP 普遍组查询报文，以判断网络上是否有组播组成员，可以根据网络的实际情况来修改周期性发送 IGMP 普遍组查询报文的时间间隔。配置普遍查询报文发送时间间隔的命令为：

```
[H3C-Vlan-interface100]igmp query-interval interval
```

在收到 IGMP 普遍查询报文后，主机会为其所加入的每个组播组都启动一个延迟定时器，其值在 0 到最大响应时间（该时间值从 IGMP 查询报文的最大响应时间字段获得）中随机选定，当定时器的值减为 0 时，主机就会向该定时器对应的组播组发送 IGMP 成员关系报告报文。

合理配置 IGMP 查询的最大响应时间，既可以使主机对 IGMP 查询报文做出快速响应，又可以减少由于定时器同时超时，造成大量主机同时发送报告报文而引起的网络拥塞。

配置普遍查询报文的最大响应时间的命令为：

```
[H3C-igmp]max-response-time interval
```

 紫光集团 H3C
核心企业 数字化转型领导者

IGMP Snooping配置命令

- 全局和VLAN视图使能IGMP Snooping

```
[H3C ] igmp-snooping
```

```
[H3C-igmp-snooping]enable vlan vlan-id
```
- 配置端口快速离开功能

```
[H3C-Ethernet1/0/1] igmp-snooping fast-leave  
vlan vlan-list
```
- 配置模拟主机加入

```
[H3C-Ethernet1/0/1] igmp-snooping host-join group-  
address vlan vlan-id
```

www.h3c.com

配置 IGMP Snooping 时，必须首先全局使能 IGMP Snooping，然后在 IGMP Snooping 视图下指定使能 IGMP Snooping 功能的 VLAN，对应的配置命令为：

```
[H3C]igmp-snooping
```

```
[H3C-igmp-snooping]enable vlan vlan-id
```

在交换机上，如果端口下只连接有一个接收者，则可以通过使能端口快速离开功能以节约带宽和资源。

端口快速离开是指当交换机从某端口收到主机发送的离开某组播组的 IGMP 离开组报文时，直接把该端口从对应转发表项的出端口列表中删除。此后，当交换机收到对该组播组的 IGMP 特定组查询报文时，交换机将不再向该端口转发。

端口快速离开功能配置命令为：

```
[H3C-Ethernet1/0/1]igmp-snooping fast-leave
```

通常情况下，运行 IGMP 的主机会对 IGMP 查询器发出的查询报文进行响应。如果主机由于某种原因无法响应，就可能导致组播路由器认为该网段没有该组播组的成员，从而取消相应的转发路径。

为避免这种情况的发生，可以将交换机的某个端口配置成为组播组成员（即配置模拟主机加入）。当收到 IGMP 查询报文时由模拟主机进行响应，从而保证该交换机能够继续收到组播报文。

模拟主机加入的配置命令为：

```
[H3C-Ethernet1/0/1]igmp-snooping host-join group-address [ source-ip
source-address ] vlan vlan-id
```

IGMP Snooping配置命令（续）

 紫光集团 **H3C**
核心企业 | 数字化解决方案领导者

- 配置未知组播丢弃

[H3C-vlan2] igmp-snooping drop-unknown

- 配置查询器

[H3C-vlan2] igmp-snooping querier

- 配置普遍查询报文的源地址

[H3C-vlan2] igmp-snooping general-query source-ip ip-address

- 配置特定组查询报文的源地址

[H3C-vlan2] igmp-snooping special-query source-ip ip-address

www.h3c.com

未知组播数据报文是指在 IGMP Snooping 转发表中不存在对应转发表项的那些组播数据报文，交换机收到未知组播报文时将在未知组播数据报文所属的 VLAN 内广播该报文，这样会占用大量的网络带宽，影响转发效率。

可以在交换机上启动丢弃未知组播数据报文功能，当交换机收到未知组播数据报文时，只向其路由器端口转发，不在 VLAN 内广播。如果交换机没有路由器端口，数据报文会被丢弃，不再转发。

配置位置组播丢弃的命令为：

```
[H3C-vlan2]igmp-snooping drop-unknown
```

在运行了 IGMP 的组播网络中，会有一台三层组播设备充当 IGMP 查询器，负责发送 IGMP 查询报文，使三层组播设备能够在网络层建立并维护组播转发表项，从而在网络层正常转发组播数据。

但是，在一个没有三层组播设备的网络中，由于二层设备并不支持 IGMP，因此无法实现 IGMP 查询器的相关功能。为了解决这个问题，可以在二层设备上使能 IGMP Snooping 查询器，使二层设备能够在数据链路层建立并维护组播转发表项，从而在数据链路层正常转发组播数据。

二层设备查询器的配置包含配置查询器、配置普遍查询报文的源地址、配置特定组查询报文的源地址等配置，对应的配置命令为：

配置查询器：

[H3C-vlan2]igmp-snooping querier


配置普遍查询报文的源地址：

[H3C-vlan2]igmp-snooping general-query source-ip ip-address

配置特定组查询报文的源地址：

[H3C-vlan2]igmp-snooping special-query source-ip ip-address

组播VLAN命令



- 配置组播VLAN

[H3C] multicast-vlan vlan-id

- 向组播VLAN内添加子VLAN

[H3C-mvlan-100] subvlan vlan-list

www.h3c.com

配置组播 VLAN 时，首先需要把某个 VLAN 配置为组播 VLAN，再将用户 VLAN 添加到该组播 VLAN 内，使其成为组播 VLAN 的子 VLAN。

配置组播 VLAN 的命令为：

[H3C]multicast-vlan vlan-id

向组播 VLn 内添加子 VLAN 的命令为：

[H3C-mvlan-100]subvlan vlan-list

PIM配置命令



- 进入PIM视图

```
[H3C] pim
```

- 在接口启动PIM-DM

```
[H3C-Vlan-interface100] pim dm
```

- 在接口启动PIM-SM

```
[H3C-Vlan-interface100] pim sm
```

- 配置DR优先级

```
[H3C-Vlan-interface100] pim hello-option dr-  
priority priority
```

www.h3c.com

配置 PIM-DM 通常只需要在接口下使能 PIM-DM 即可，配置命令为：

```
[H3C-Vlan-interface100]pim dm
```

PIM-SM 的配置包含使能 PIM-SM、RP 的相关配置、C-BSR 的相关配置。如果需要指定某台设备作为共享网段上的 DR 设备，还需要配置 Hello 报文的优先级选项。

使能 PIM-SM 的配置命令为：

```
[H3C-Vlan-interface100]pim sm
```

配置 DR 优先级的命令为：

```
[H3C-Vlan-interface100]pim hello-option dr-priority priority
```

在 PIM 视图下配置 RP 和 C-BSR 的相关信息，进入 PIM 视图的配置命令为：

```
[H3C] pim
```

配置RP



● PIM视图配置静态RP

```
[H3C-pim] static-rp rp-address [ acl-number | bidir | preferred ]*
```

● 配置候选rp

```
[H3C-pim] c-rp ip-address [ advertisement-interval adv-interval | group-policy acl-number | holdtime hold-time | priority priority ] *[ bidir ]
```

● 配置候选BSR

```
[H3C-pim] c-bsr ip-address [ scope group-address { mask-length | mask } ] [ hash-length hash-length | priority priority ]*
```

www.h3c.com

PIM-SM 中 RP 可以通过手工指定，也可以通过动态选举。

当网络内仅有一个动态 RP 时，通过手工配置静态 RP 可以避免因单一节点故障而引起的通信中断，同时也可以避免 C-RP 与 BSR 之间频繁的信息交互而占用带宽。当网络中同时存在动态 RP 和静态 RP 时，如果指定了 preferred 参数，则优先选择静态 RP，只有当静态 RP 失效时，动态 RP 才能生效。如果未指定 preferred 参数，则表示优先选择动态 RP，只有当未配置动态 RP 或动态 RP 失效时，静态 RP 才能生效。

手工指定 RP 的配置命令为：

```
[H3C-pim] static-rp rp-address [ acl-number | bidir | preferred ]
```

配置 RP 通过动态选举时，可以把有意成为 RP 的路由器配置为 C-RP。建议在骨干网路由器上配置 C-RP。

配置候选 RP 的命令为：

```
[H3C-pim] c-rp ip-address [ advertisement-interval adv-interval | group-policy acl-number | holdtime hold-time | priority priority ] [ bidir ]
```

在一个 PIM-SM 域中只能有一个 BSR，但需要配置至少一个 C-BSR。任意一台路由器都可以被配置为 C-BSR。在 C-BSR 之间通过自动选举产生 BSR，BSR 负责在 PIM-SM 域中收集并发布 RP 信息。C-BSR 应配置在骨干网的路由器上，C-BSR 的 IP 地址必须有对应的本地接口，且该接口上必须使能 PIM。

配置候选 BSR 的命令为：

[H3C-pim] **c-bsr** *ip-address* [**scope** *group-address* { *mask-length* | *mask* }]
[**hash-length** *hash-length* | **priority** *priority*]

22.3 组播维护命令

查看组播路由和转发信息

- 查看组播路由表
`<H3C>display multicast routing-table`
- 查看组播转发表
`<H3C>display multicast forwarding-table`
- 查看组播源的RPF信息
`<H3C>display multicast rpf-info source-address`

紫光集团 H3C
核心企业 数字化解决方案领导者
www.h3c.com

组播路由和转发信息的查看主要包括查看组播路由表、查看组播转发表和查看组播源的RPF信息。

组播路由表是进行组播数据转发的基础，通过查看该表可以了解（S，G）表项等的建立情况。查看组播路由表的命令为：

<H3C> display multicast routing-table

组播转发表直接用于指导组播数据的转发，通过查看该表可以了解组播数据的转发状态。查看组播转发表命令为：

<H3C> display multicast forwarding-table

通过查看组播源的RPF信息可以了解RPF接口、RPF邻居等信息，配置命令为：

<H3C> display multicast rpf-info source-address

查看IGMP和IGMP Snooping信息



紫光集团 H3C
核心企业 数字化转型领导者

- 查看IGMP组信息

```
<H3C>display igmp group group-address
```

- 查看IGMP Snooping组信息

```
<H3C>display igmp-snooping group
```

www.h3c.com

通过查看 IGMP 组信息可以了解接口维护的某组播组的相关信息如 Report 报文个数、最后一个发送 Report 报文的主机地址、组播组超时时间等。配置命令为：

```
<H3C>display igmp group group-address
```

通过查看 IGMP Snooping 组信息可以了解路由端口、成员端口、组地址等信息。配置命令为：

```
<H3C>display igmp-snooping group
```

查看PIM信息

● 查看PIM路由表项

<H3C>display pim routing-table

● 查看RP信息

<H3C>display pim rp-info

www.h3c.com

通过查看 PIM 路由表项，可以了解（S，G）或（*，G）表项的入接口、（S，G）或（*，G）表项的上游邻居、（S，G）或（*，G）表项下游接口的信息。配置命令为：

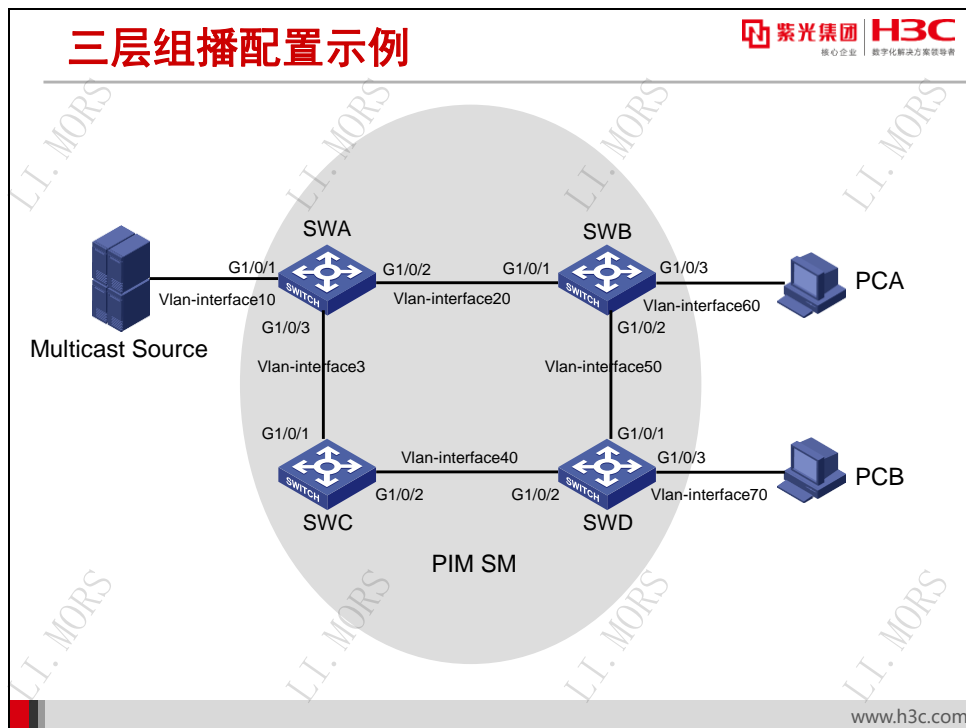
<H3C> display pim routing-table

使用 PIM-SM 时，通过查看 RP 信息，可以了解当前 RP 的地址、RP 所服务的组播组、RP 优先级、RP 超时时间等信息。配置命令为：

<H3C> display pim rp-info

22.4 组播配置示例

22.4.1 三层组播配置示例



本例要求配置组播路由协议 PIM-SM，IGMP 使用版本 2。配置过程如下：

配置 IP 地址和单播路由协议

配置 VLAN 接口以及各接口的 IP 地址和子网掩码，具体配置过程略。

配置 PIM-SM 域内的各交换机之间采用 OSPF 协议进行互连，确保在网络层互通，并且各交换机之间能够借助单播路由协议实现动态路由更新，具体配置过程略。

使能三层组播功能

在各交换机上使用命令 **multicast routing** 使能三层组播。

```
[SWA]multicast routing
[SWB]multicast routing
[SWC]multicast routing
[SWD]multicast routing
```

配置 IGMP

在交换机连接组播接收者的 VLAN 接口上使用命令 **igmp enable** 使能 IGMP。

```
[SWB-Vlan-interface10]igmp enable
[SWD-Vlan-interface20]igmp enable
```

配置 PIM-SM

在各交换机的三层接口上配置 PIM-SM，使用命令 **pim sm**。

```
[SWA-Vlan-interface1]pim sm
[SWA-Vlan-interface2]pim sm
[SWA-Vlan-interface3]pim sm
```

SWB、SWC、SWD 的配置和 SWA 相同，配置过程略。

指定 SWC 作为 PIM 域中的 C-RP 和 C-BSR

```
[SWC-pim]c-rp 3.3.3.3
[SWC-pim]c-bsr 3.3.3.3
```

在 SWB 和 SWD 上使用命令 **display igmp group** 查看组播组信息。

```
<SWB>display igmp group
IGMP groups in total: 2
Vlan-interface10(10.10.10.1):.
  IGMP groups reported in total: 2
  Group address   Last reporter   Uptime          Expires
  228.1.1.1       10.10.10.2      00:07:16        00:01:13
  239.255.255.250 10.10.10.2      16:41:32        00:01:13

[SWD]display igmp group
IGMP groups in total: 2
Vlan-interface20(20.20.20.1):.
  IGMP groups reported in total: 2
  Group address   Last reporter   Uptime          Expires
  228.1.1.1       20.20.20.2      00:35:45        00:01:54
  239.255.255.250 20.20.20.2      00:36:34        00:01:56
```

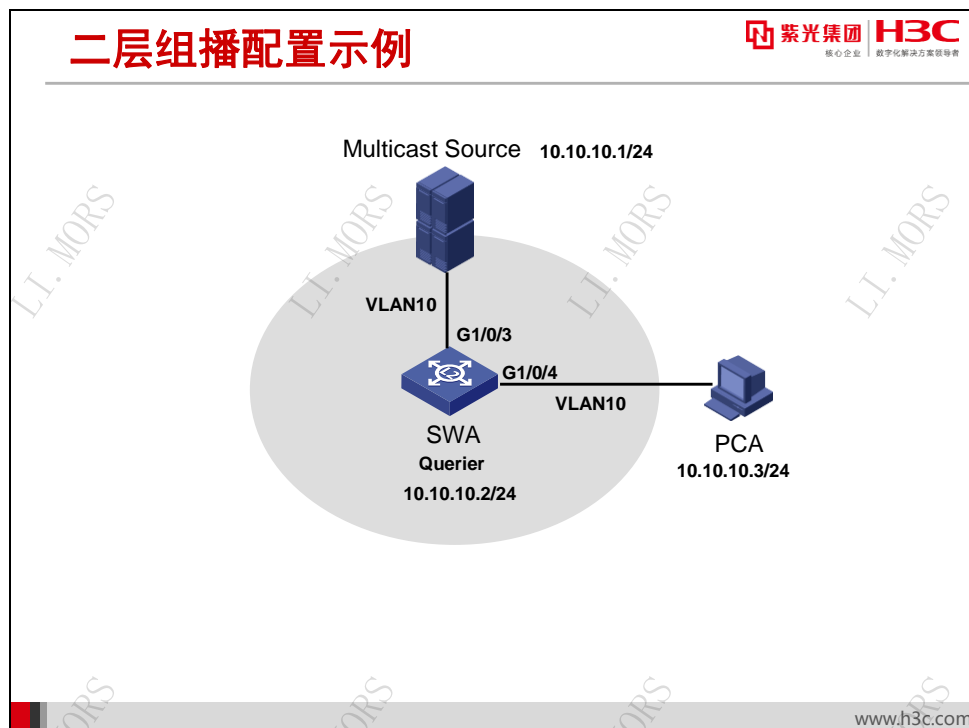
使用命令 **display multicast routing-table** 可以查看组播路由表的内容，组播路由表的内容和选择的组播路由协议无关，所以使用 PIM-DM 和使用 PIM-SM 时，组播路由表内容相同。

例如，SWA 的组播路由表如下，其余路由器相似。

```
[SWA]display multicast routing-table
Multicast routing table
Total 1 entry

00001. (1.1.1.2, 228.1.1.1)
  Uptime: 00:18:56
  Upstream Interface: Vlan-interface1
  List of 2 downstream interfaces
    1: Vlan-interface2
    2: Vlan-interface3
```


22.4.2 二层组播配置示例



本例要求配置 IGMP Snooping，并在 SWA 上配置 IGMP Snooping 查询器。配置过程如下：

在 SWA 上配置 VLAN10 并按照组网图将对应接口加入 VLAN10。配置 SWA 的 VLAN10 虚接口 IP 地址为 10.10.10.2，作为 IGMP Snooping 查询器的源 IP。

在 SWA 的全局和 VLAN10 内使能 IGMP Snooping。

```
[SWA]igmp-snooping
[SWA-vlan10]igmp-snooping enable
```

在 SWA 的 VLAN10 内使能 IGMP Snooping 查询器，并配置查询报文的源地址。

```
[SWA-vlan10]igmp-snooping querier
[SWA-vlan10]igmp-snooping general-query source-ip 10.10.10.2
[SWA-vlan10]igmp-snooping special-query source-ip 10.10.10.2
```

在 SWA 的 VLAN10 内配置未知组播丢弃。

```
[SWA-vlan10]igmp-snooping drop-unknown
```

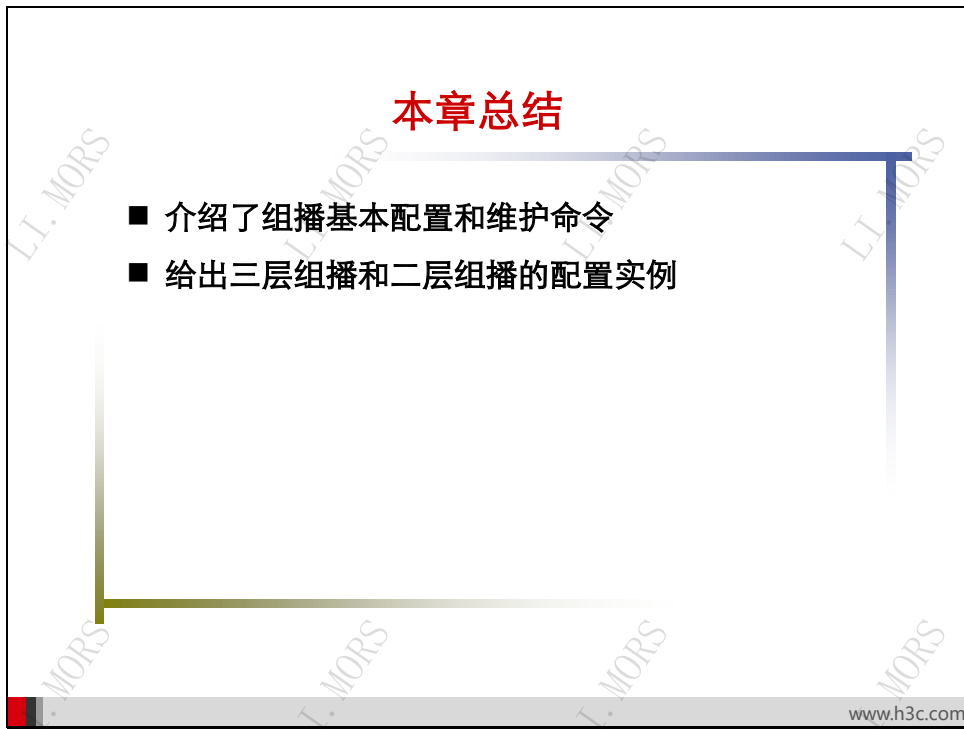
在 SWA 上使用命令 **display igmp-snooping group** 查看二层组播组信息，包含路由端口以及组成员端口等信息。

```
[H3C]dis igmp-snooping group
Total 3 entries.

VLAN 10: Total 3 entries.
(0.0.0.0, 225.0.0.2)
Host slots (0 in total):
Host ports (1 in total):
GE1/0/4 (00:04:02)
```

```
(0.0.0.0, 226.81.9.8)
Host slots (0 in total):
Host ports (1 in total):
    GE1/0/3                                (00:02:28)
(0.0.0.0, 239.255.255.250)
Host slots (0 in total):
Host ports (2 in total):
    GE1/0/3                                (00:02:30)
    GE1/0/4                                (00:02:26)
```

22.5 本章总结



本章总结

- 介绍了组播基本配置和维护命令
- 给出三层组播和二层组播的配置实例

www.h3c.com

22.6 习题和解答

22.6.1 习题

1. 进行三层组播配置之前，首先需要进行什么配置？（ ）
 - A. 进入 IGMP 视图
 - B. 配置 IGMP 协议版本
 - C. 配置 PIM 协议
 - D. 通过 multicast routing 全局启用组播
2. 关于 IGMP Snooping 配置，正确的有（ ）
 - A. 同一个 VLAN 下既可以配置 IGMP，也可以配置 IGMP Snooping
 - B. 端口快速离开是指当交换机从某端口收到主机发送的离开某组播组的 IGMP 离开组报文时，直接把该端口从对应转发表项的出端口列表中删除
 - C. 可以在交换机上启动丢弃未知组播数据报文功能，当交换机收到未知组播数据报文时，只向其路由器端口转发，不在 VLAN 内广播
 - D. 二层设备上使能 IGMP Snooping 查询器，使二层设备能够在数据链路层建立并维护组播转发表项，从而在数据链路层正常转发组播数据
3. 关于 PIM 协议配置，正确的有（ ）
 - A. 配置 PIM-DM 通常只需要在接口下使能 PIM-DM 即可
 - B. PIM-SM 的配置包含使能 PIM-SM、RP 的相关配置、C-BSR 的相关配置。如果需要指定某台设备作为共享网段上的 DR 设备，还需要配置 Hello 报文的优先级选项
 - C. PIM 协议在 VLAN 视图配置
 - D. PIM 协议在 VLAN 接口视图配置
4. 关于 RP 和 BSR，下列说法正确的有（ ）
 - A. PIM-SM 中 RP 可以通过手工指定，也可以通过动态选举
 - B. 配置 RP 通过动态选举时，可以把有意成为 RP 的路由器配置为 C-RP。建议在骨干网路由器上配置 C-RP 本
 - C. 在一个 PIM-SM 域中只能有一个 BSR
 - D. BSR 负责在 PIM-SM 域中收集并发布 RP 信息
5. 关于组播维护命令，正确的有（ ）
 - A. 组播路由表是进行组播数据转发的基础，通过查看该表可以了解（S，G）表项等的建立情况

- B. 组播转发表直接用于指导组播数据的转发，通过查看该表可以了解组播数据的转发状态
- C. 通过查看 IGMP 组信息可以了解接口维护的某组播组的相关信息如 Report 报文个数、最后一个发送 Report 报文的主机地址、组播组超时时间等
- D. 通过查看 PIM 路由表项，可以了解 (S, G) 或 (*, G) 表项的入接口、(S, G) 或 (*, G) 表项的上游邻居、(S, G) 或 (*, G) 表项下游接口的信息

22.6.2 习题答案

- 1. D
- 2. BCD
- 3. ABD
- 4. ABCD
- 5. ABCD