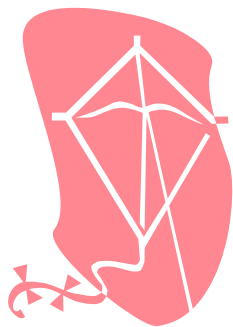


第8章

组播服务

组播服务

组播技术优点



单播

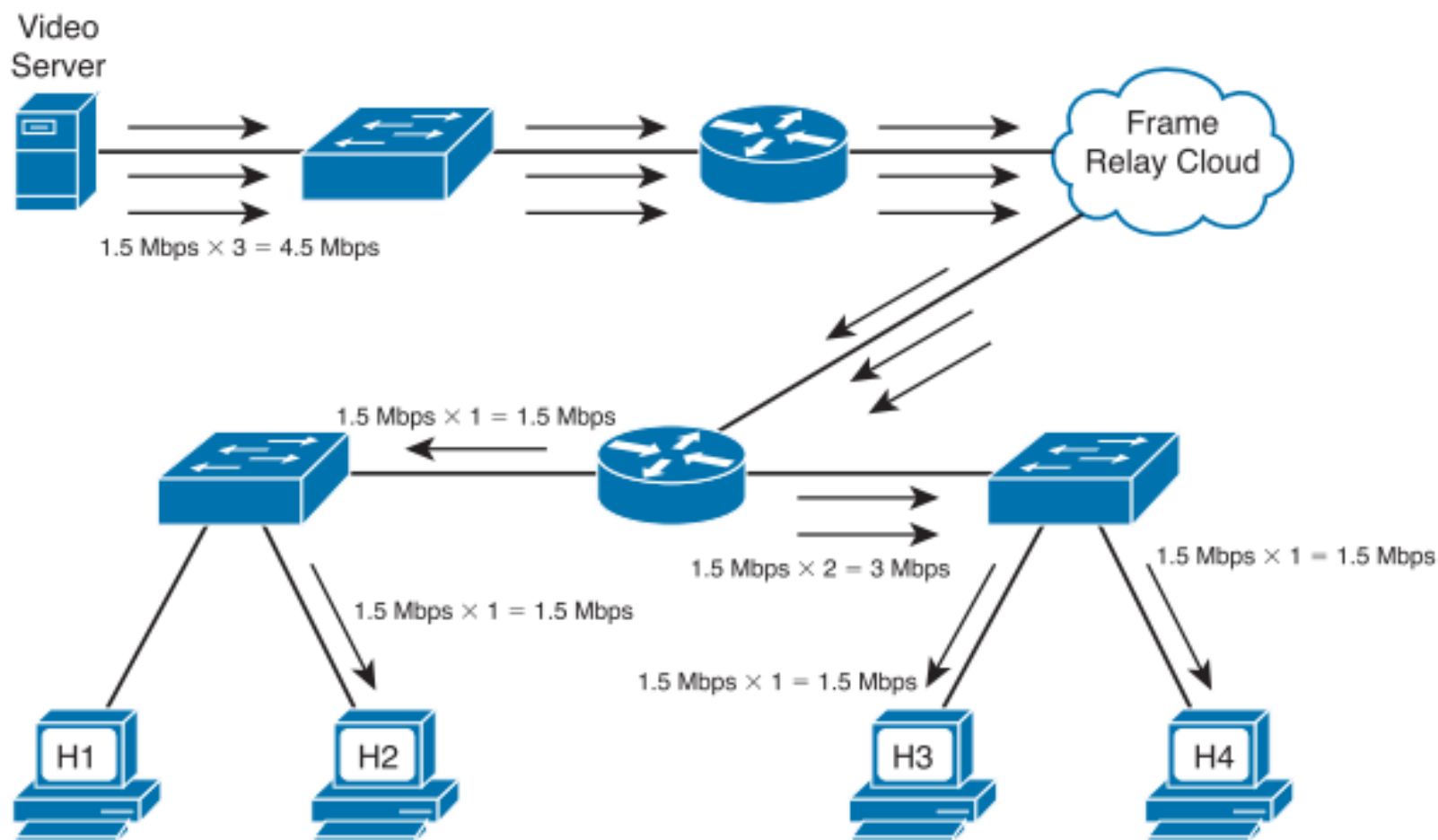


Figure 7-1 Unicast

单播 (可扩展性)

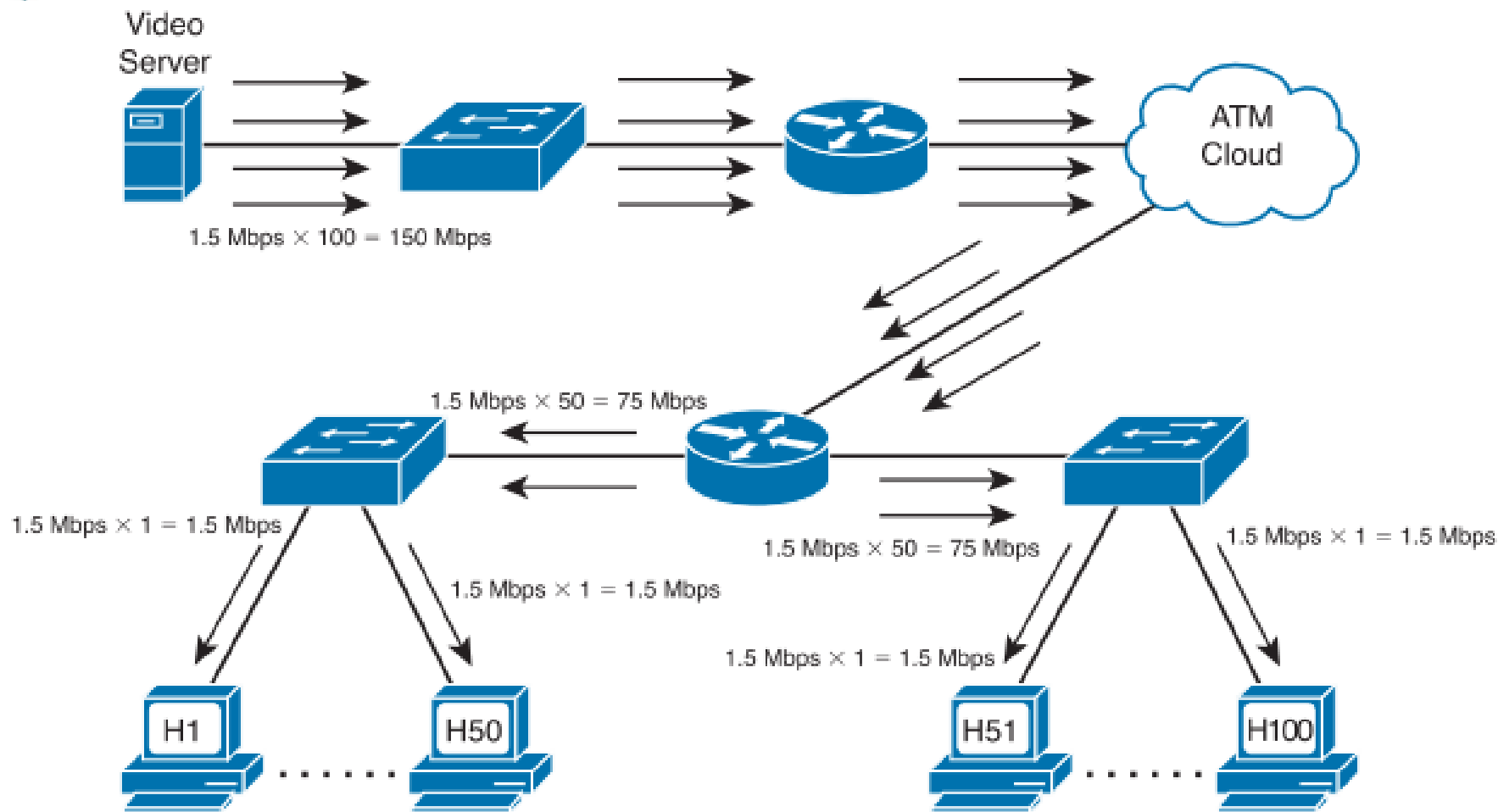


Figure 7-2 *Unicast Does Not Scale to Large Numbers of Receivers*

广播

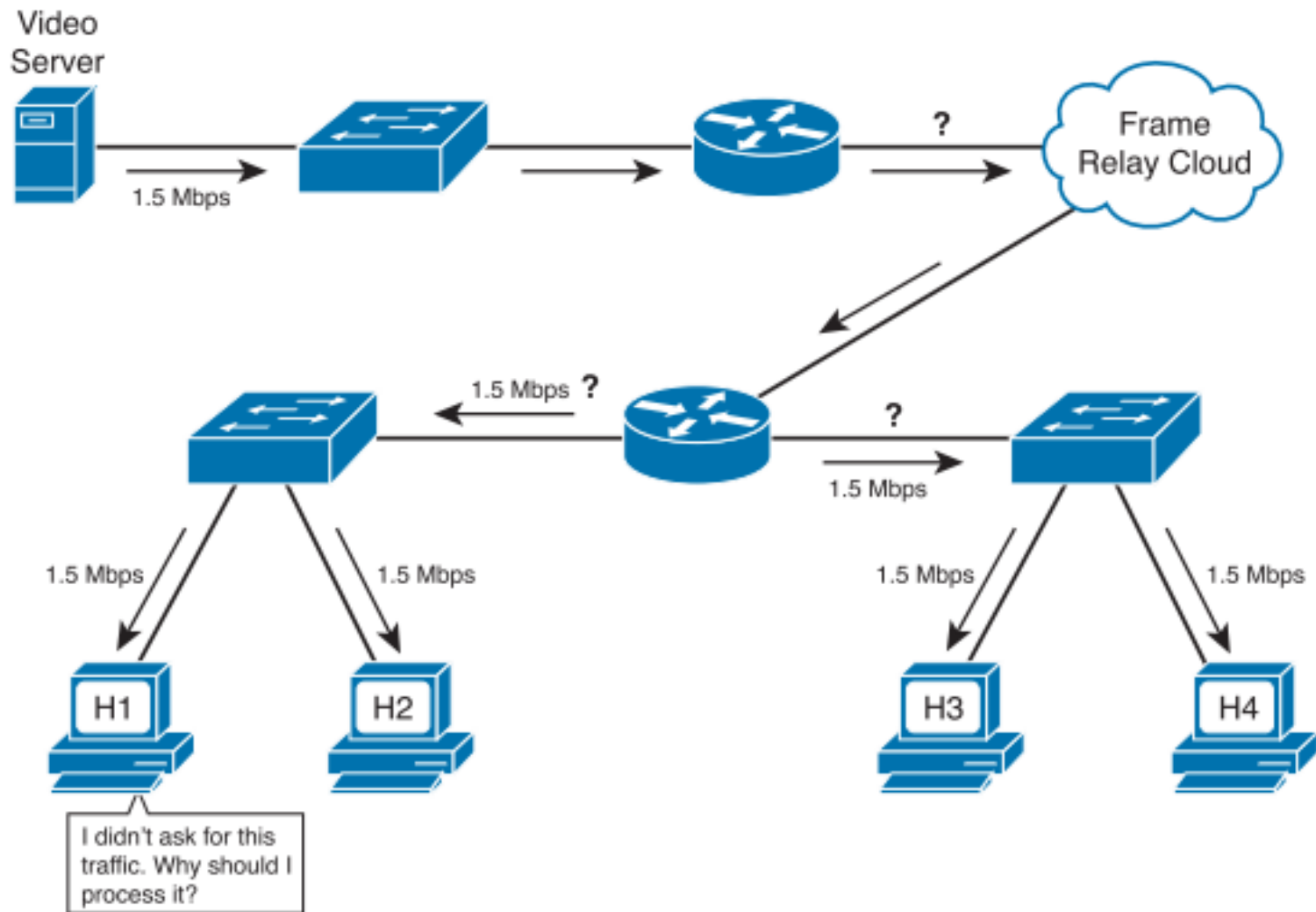


Figure 7-3 *Broadcast Wastes Bandwidth and Increases Processing Load on the CPU*

组播

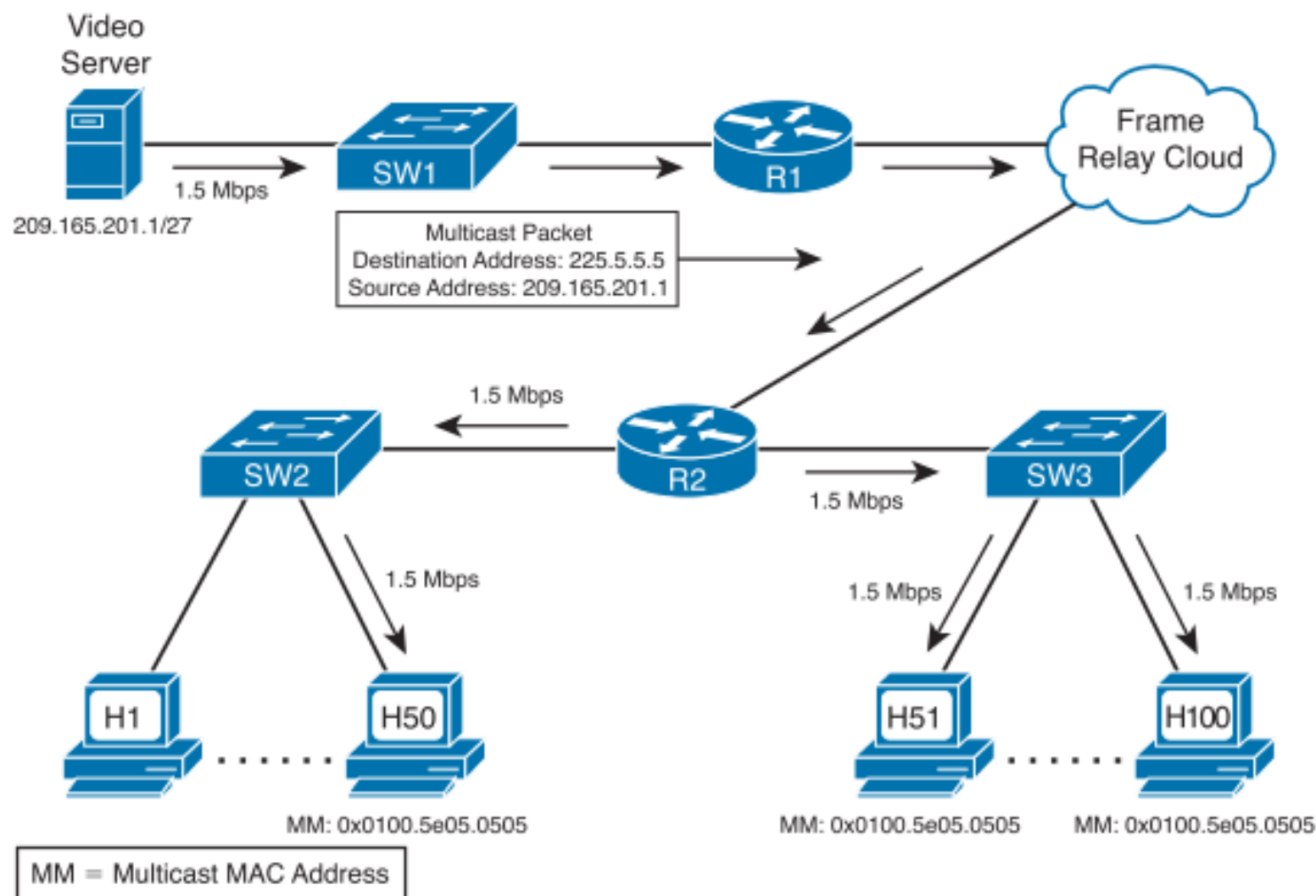


Figure 7-4 *How Multicast Delivers Traffic to Selected Users*

组播（可扩展性）

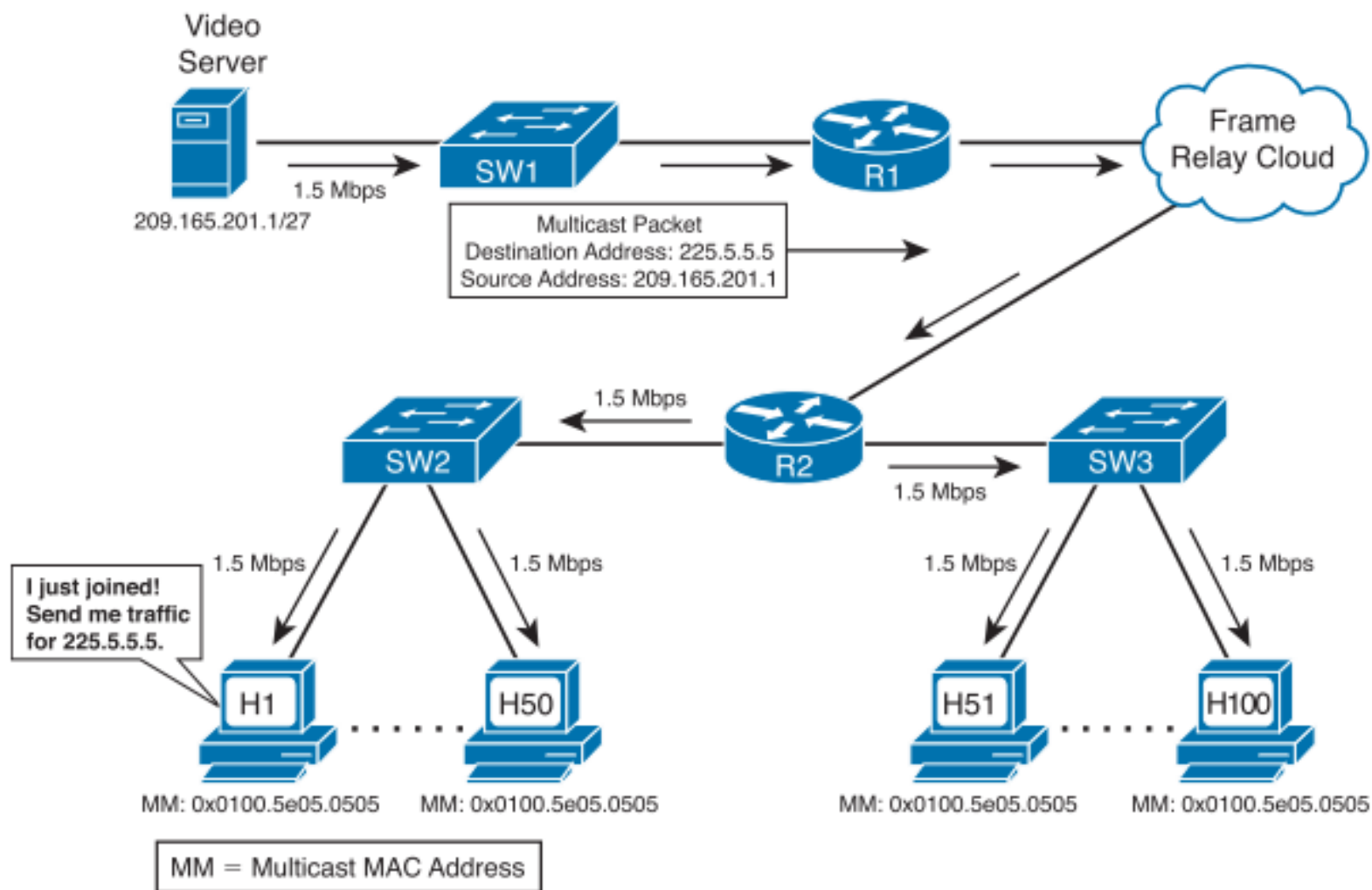
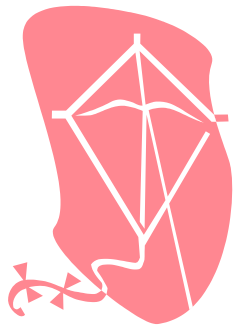


Figure 7-5 Multicasting Is Scalable

组播服务

组播地址



组播IP地址

■ 组播源IP地址：

◆ A、B、C类地址

■ 组播目标IP地址：

◆ D类地址

◆ 224.0.0.0 --239.255.255.255

	4 bytes (32 bits)				Values
	1st	2nd	3rd	4th	
A	0NNNNNNN	Host	Host	Host	0~127
B	10NNNNNNN	Network	Host	Host	128~191
C	110NNNNN	Network	Network	Host	192~223
D	1110MMMM	MGroup	MGroup	MGroup	224~239
E	1111XXXX				240~255

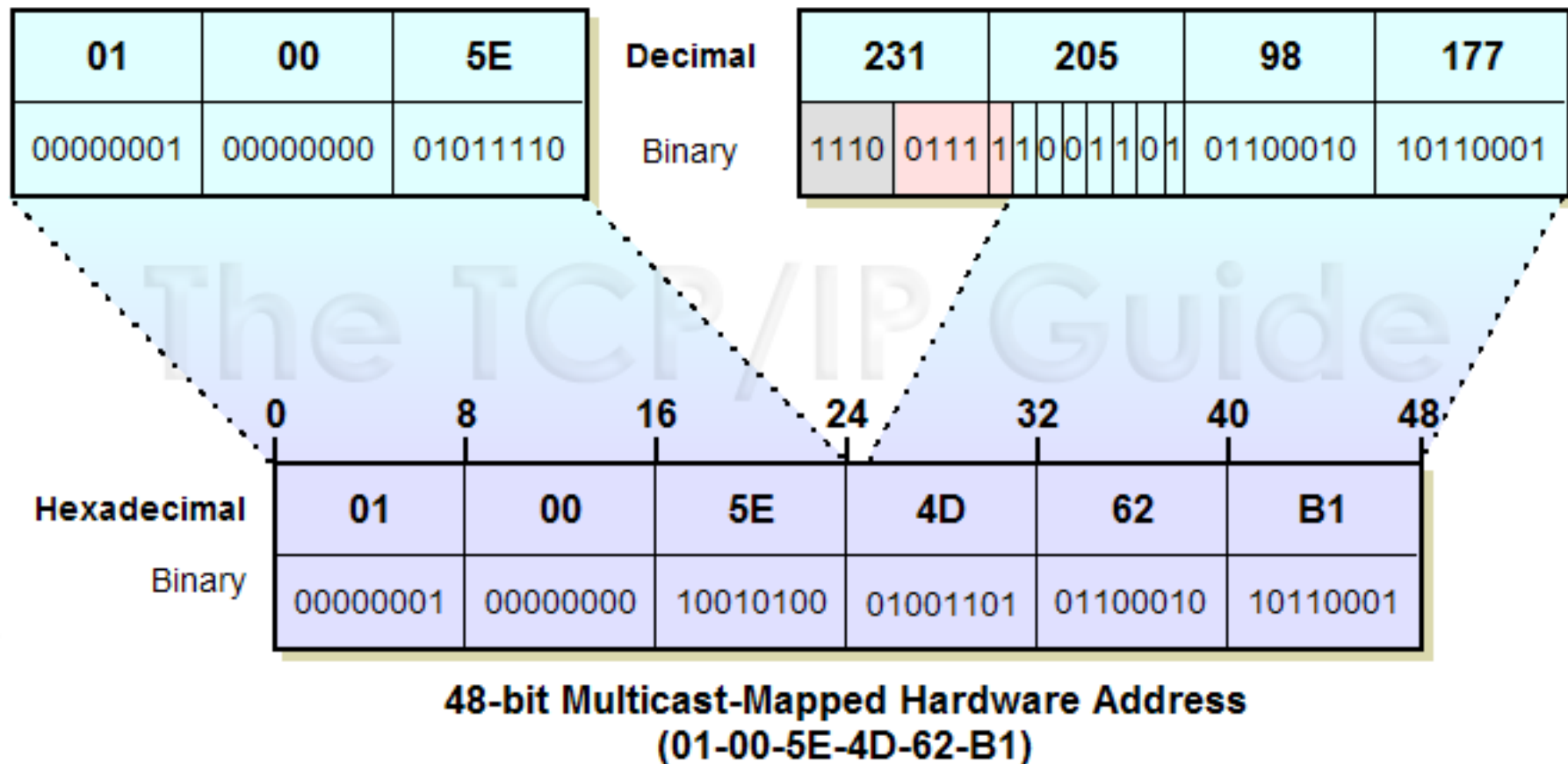
目的MAC直接由IP映射，无需ARP

[illegible][illegible]

组播 MAC 地址--实例

24-bit IANA Multicast OUI
(01-00-5E)

32-bit Multicast IP Address
(231.205.98.177)



对应关系

- 32个IP组播地址对应同一个MAC组播地址

32-IP Multicast Addresses

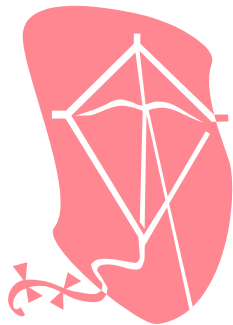
224.1.1.1
224.129.1.1
225.1.1.1
225.129.1.1
:
:
:
238.1.1.1
238.129.1.1
239.1.1.1
239.129.1.1

1-Multicast MAC Address

0x0100.5E01.0101

组播服务

组播模型

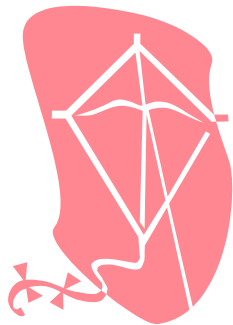


组播模型

- 根据对组播源的控制程度的不同，组播分为以下三种模型。
 - ◆ ASM 模型：Any-Source Multicast，任意信源组播
 - ◆ SFM 模型：Source-Filtered Multicast，信源过滤组播
 - ◆ SSM 模型：Source-Specific Multicast，指定信源组播

组播服务

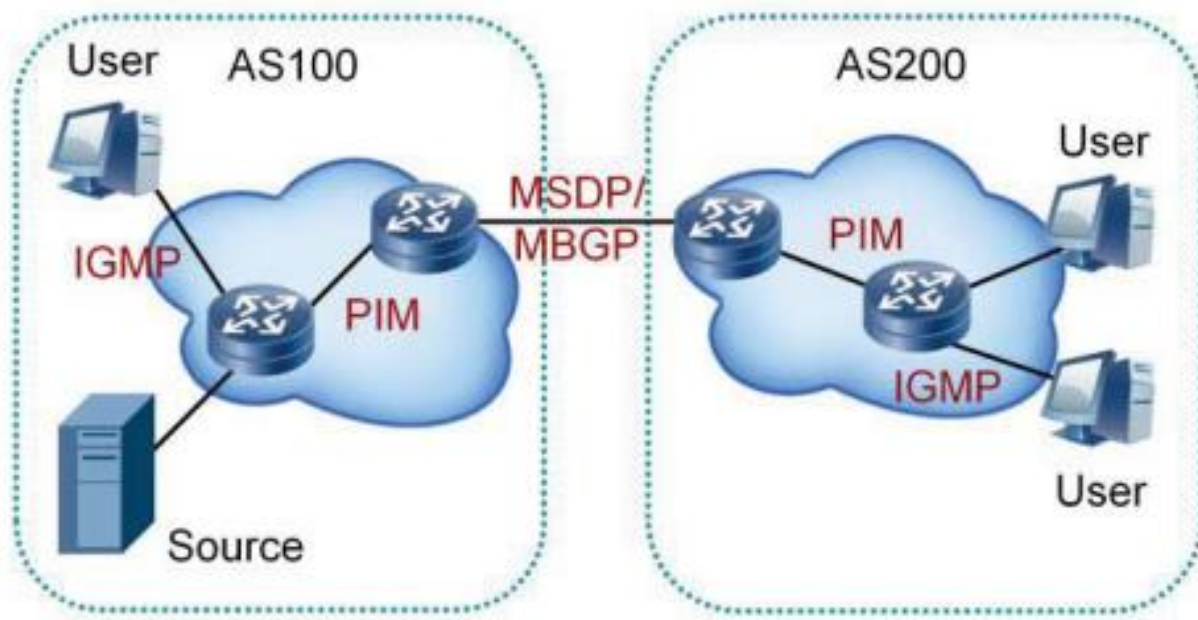
组播协议



组播协议

■ 组播协议包括

- ◆ 用于主机注册的**组播组管理**协议
- ◆ 用于组播选路转发的**组播路由**协议。



组播协议

■ 组播组管理协议

- ◆ IGMP (Internet Group Management Protocol)

■ 组播路由协议，对于ASM模型：

- ◆ 域内组播路由协议包括：

- DVRMP (Distance Vector Multicast Routing Protocol)
- MOSPF是OSPF路由协议的扩展协议
- PIM (Protocol Independent Multicast)

- ◆ 域间组播路由：

- MSDP (Multicast Source Discovery Protocol)
- MPBGP (MultiProtocol Border Gateway Protocol) 的 MBGP (Multicast BGP)

组播服务



组管理协议

IGMP

■ IGMP 协议：

- ◆ Internet Group Management Protocol
(Internet 组管理协议)
- ◆ 在主机和相邻路由器之间使用，
- ◆ 是公有三层协议，
- ◆ 用来实现主机加入和退出组播组。

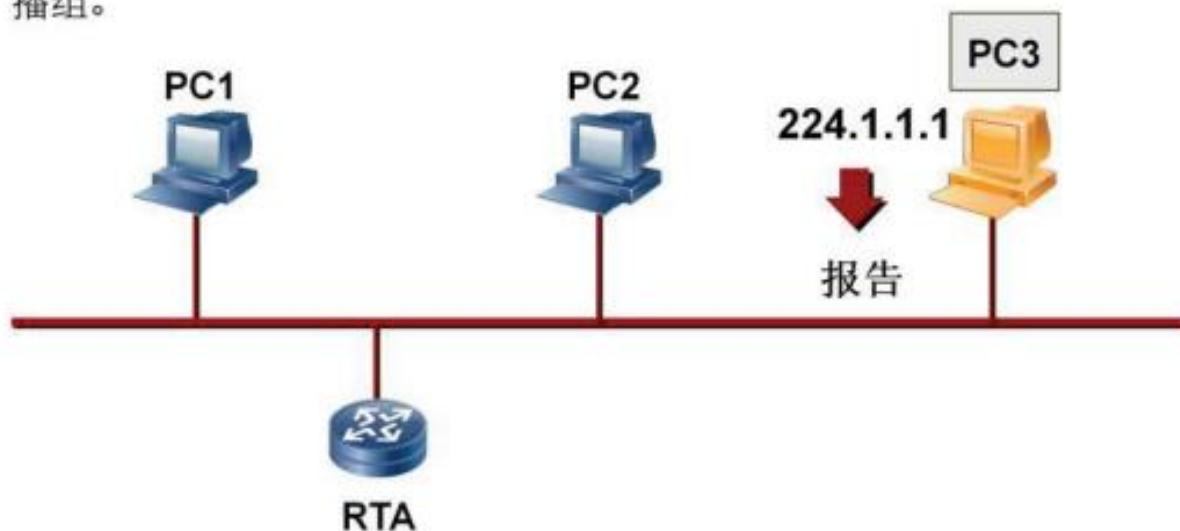
IGMP版本

- 到目前为止，IGMP有三个版本：
 - ◆ IGMPv1 (RFC1112)定义了基本的组成员查询和报告过程。
 - ◆ IGMPv2 (RFC2236)在IGMPv1的基础上添加了组成员快速离开的机制
 - ◆ IGMPv 3增加的主要功能是成员可以指定接收或指定不接收某些组播源的报文。
 - SSM (Source-Specific Multicast)模型

主动加入

IGMPv2组成员加入

当一个主机加入一个组播组，则应该立即发送一个或多个成员关系报告给组播组。

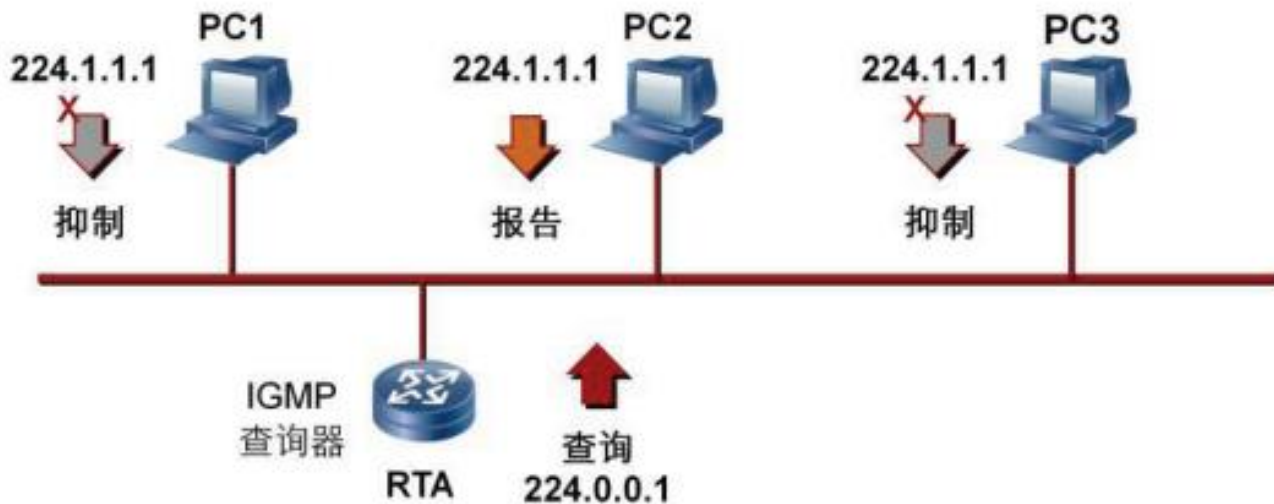


查询与相应

IGMPv2查询与响应

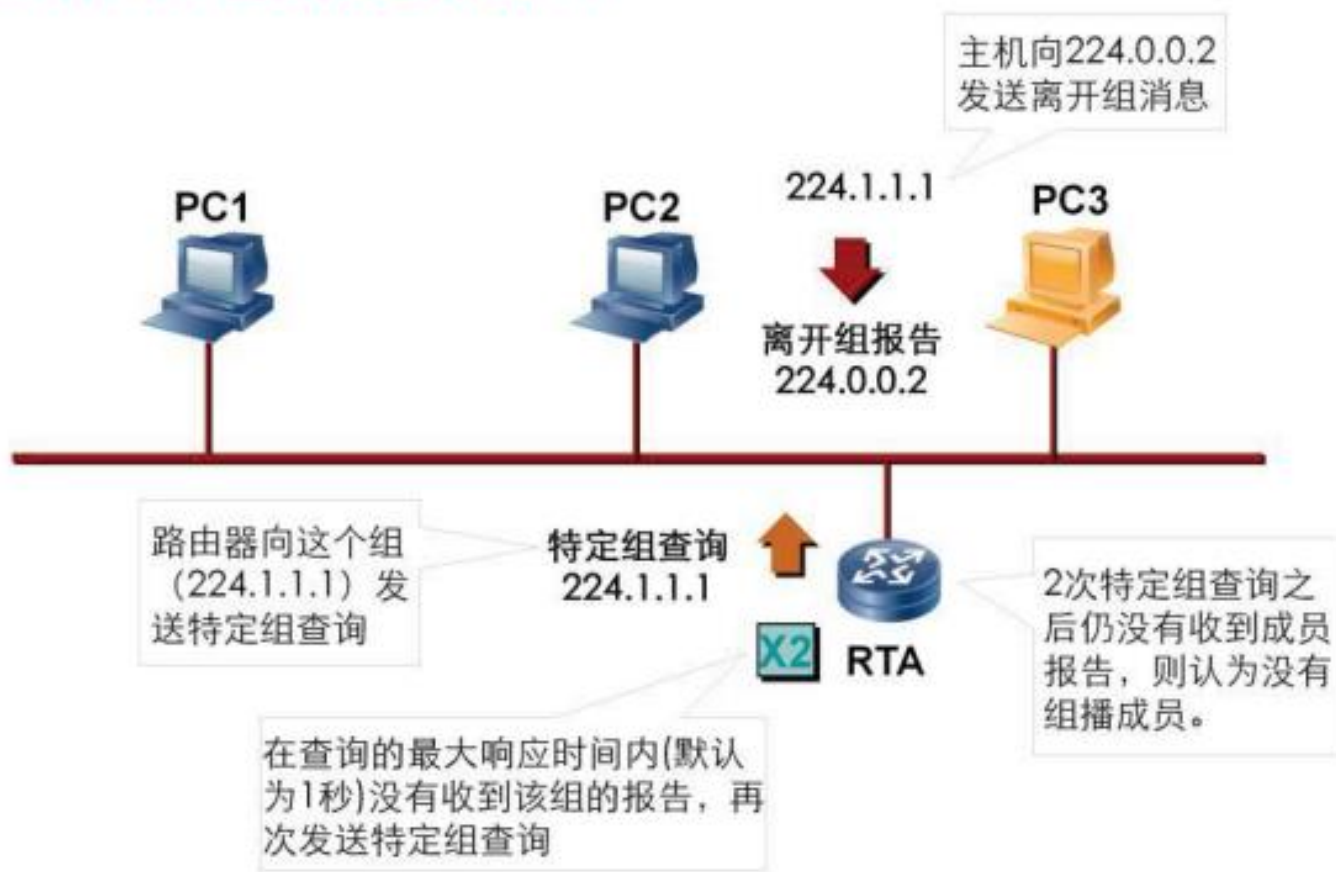
与IGMPv1相同，存在抑制机制。

增加了最大响应时间。



退出操作

IGMPv2组成员离开



比较

IGMP版本比较

	IGMPv1	IGMPv2	IGMPv3
查询器选举	依靠上层路由协议	自己选举	自己选举
成员离开方式	默默离开	主动发出离开报文	主动发出离开报文
指定组查询	不支持	支持	支持
指定源、组加入	不支持	不支持	支持



IGMP Snooping

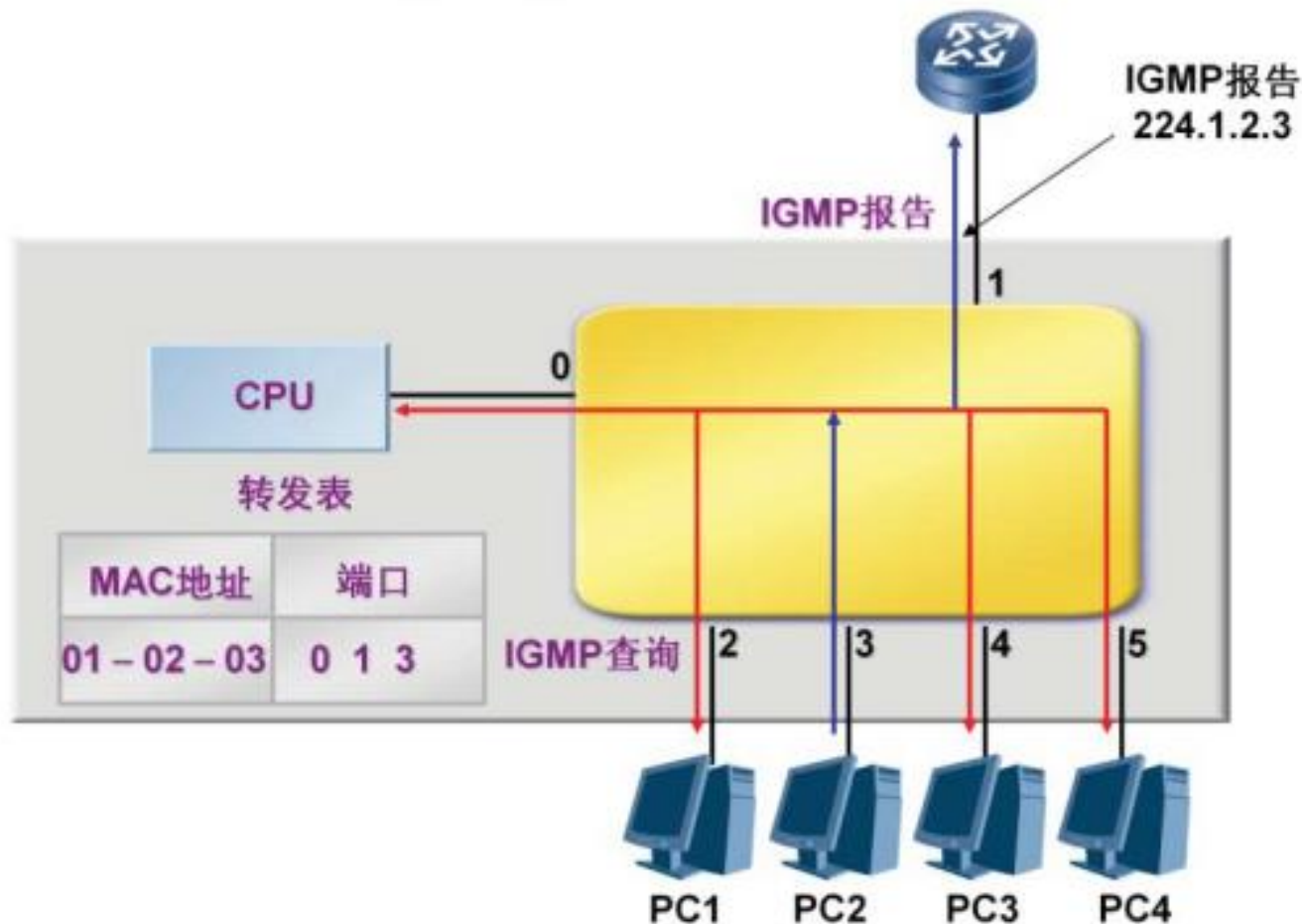
L2组播数据泛洪问题

■ 链路层组播数据泛洪问题：

- ◆ 数据链路交换机只支持二层地址，因此不参与IGMP
- ◆ 组播MAC地址不会做为源地址，因此无MAC表项，从而使用泛洪(flooding)方式发送。

IGMP Snooping 建立MAC表

IGMP Snooping建立和维护组



组播服务



组播路由协议

协议无关组播路由协议

- Protocol-Independent Multicast (PIM) 是一种组播路由协议，用来在路由器和路由间进行组播数据包的发送。
 - ◆ 支持所有的单播路由协议。
 - ◆ 它使用单播路由协议创建的路由表来进行组播路由计算。
- PIM 不进行路由信息交换。

协议无关组播路由

■ PIM有两种模式：

◆ 密集模式，PIM-DM (Dense mode)

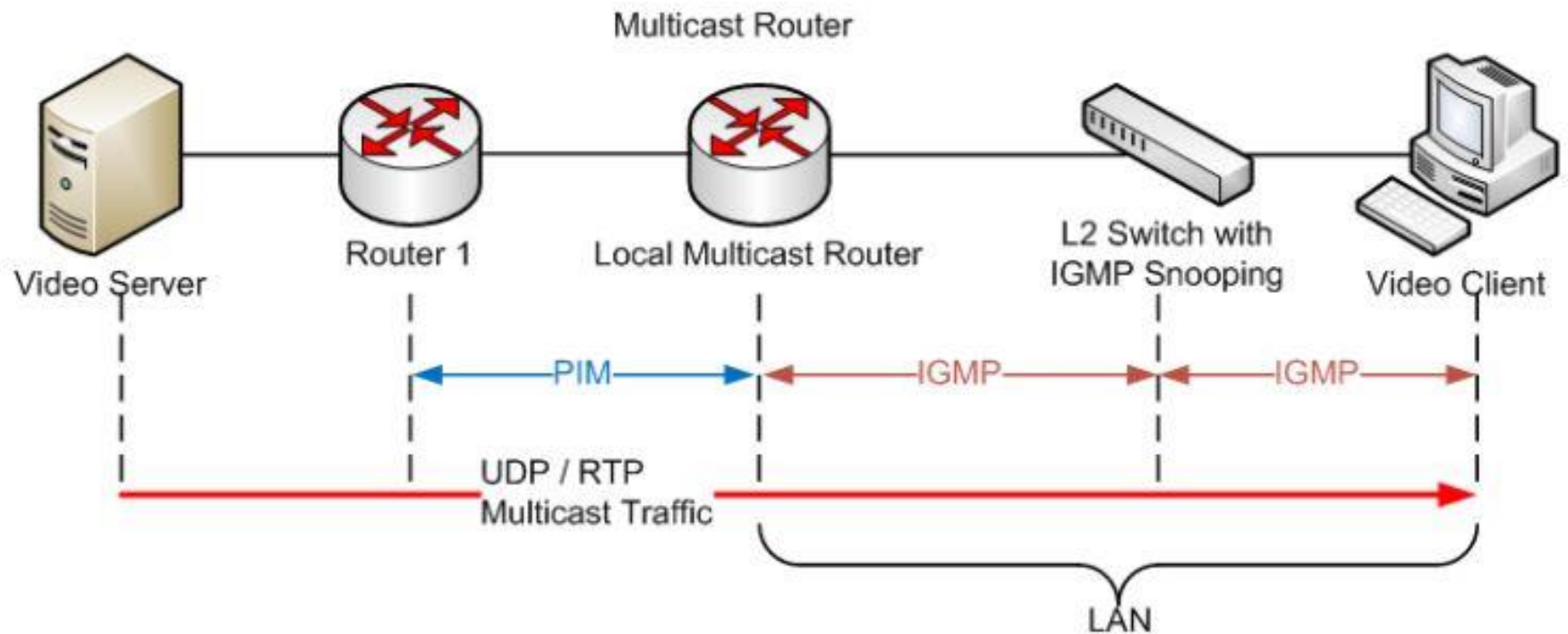
- 适用于小型网络。

- 组播数据包被扩散到网络中的所有点，然后再裁剪

◆ 稀疏模式，PIM-SM (Sparse mode)

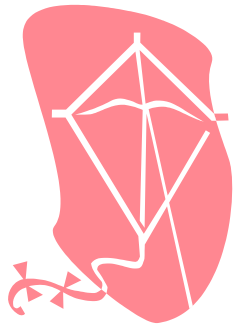
- 适用于组成员分布相对分散、范围较广、大规模的网络

组播协议



组播服务

组播树



组播树分类

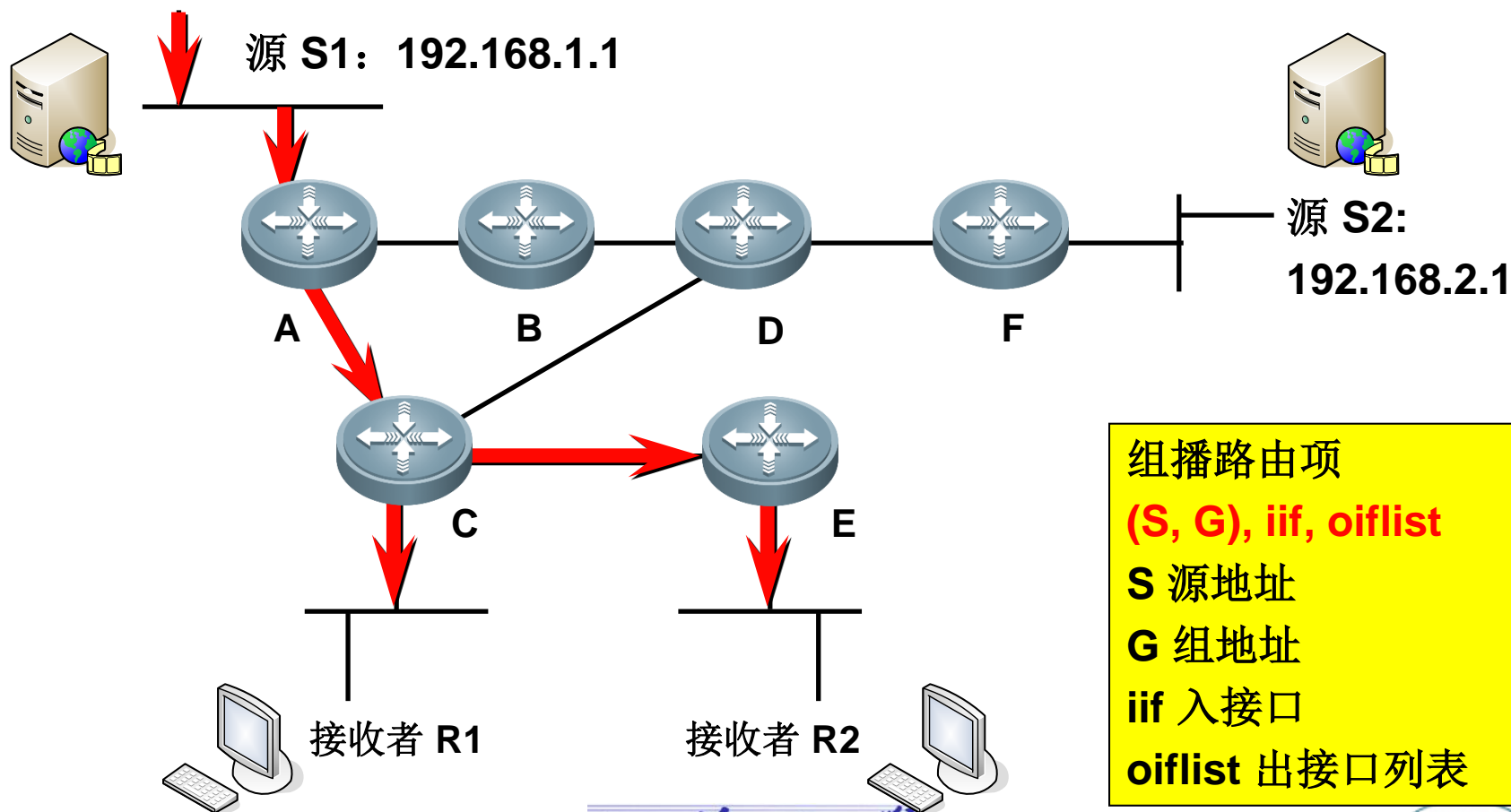
- 启用组播的路由器动态创建组播树 (distribution trees) 来进行组播转发路径控制。
- 组播树分类：
 - ◆ 有源树, Source tree: 每个组播源向一个组播组发送时创建一个有源树。
 - ◆ 共享树, Shared tree: 共享树被一个组播组的所有组播源源所共享。

⑩有源树, 也称为SPT (shortest path tree)

⑩共享树, 也称为RPT(Rendezvous PointTree)

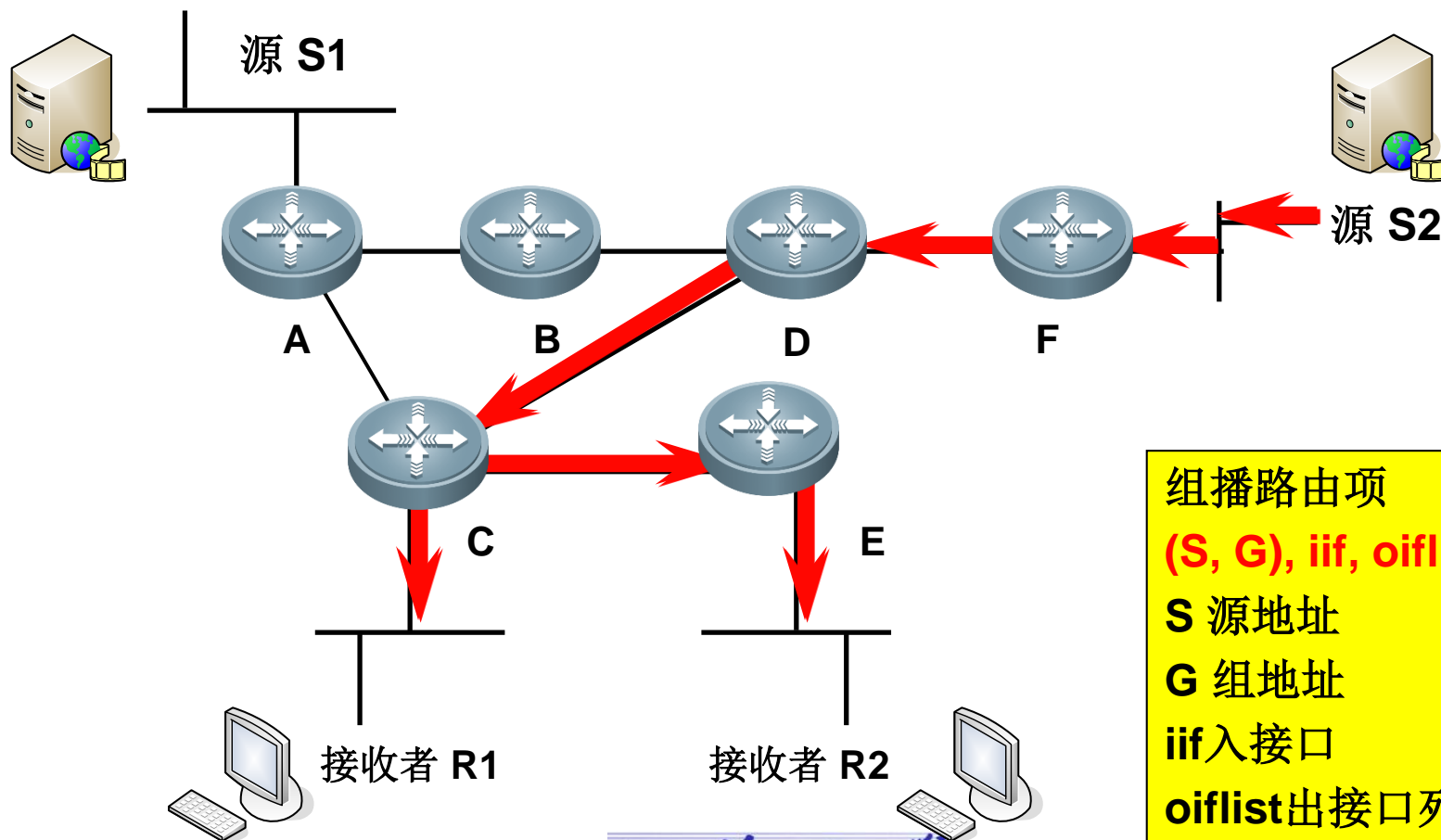
有源树

有源树的根是组播数据流的来源，有源树的分支形成了通过网络到达接收站点的分布树。



有源树

因为有源树以最短的路径贯穿网络，所以也被成为最短路径树。（**SPT: shortest path tree**）



组播路由项

(S, G), iif, oiflist

S 源地址

G 组地址

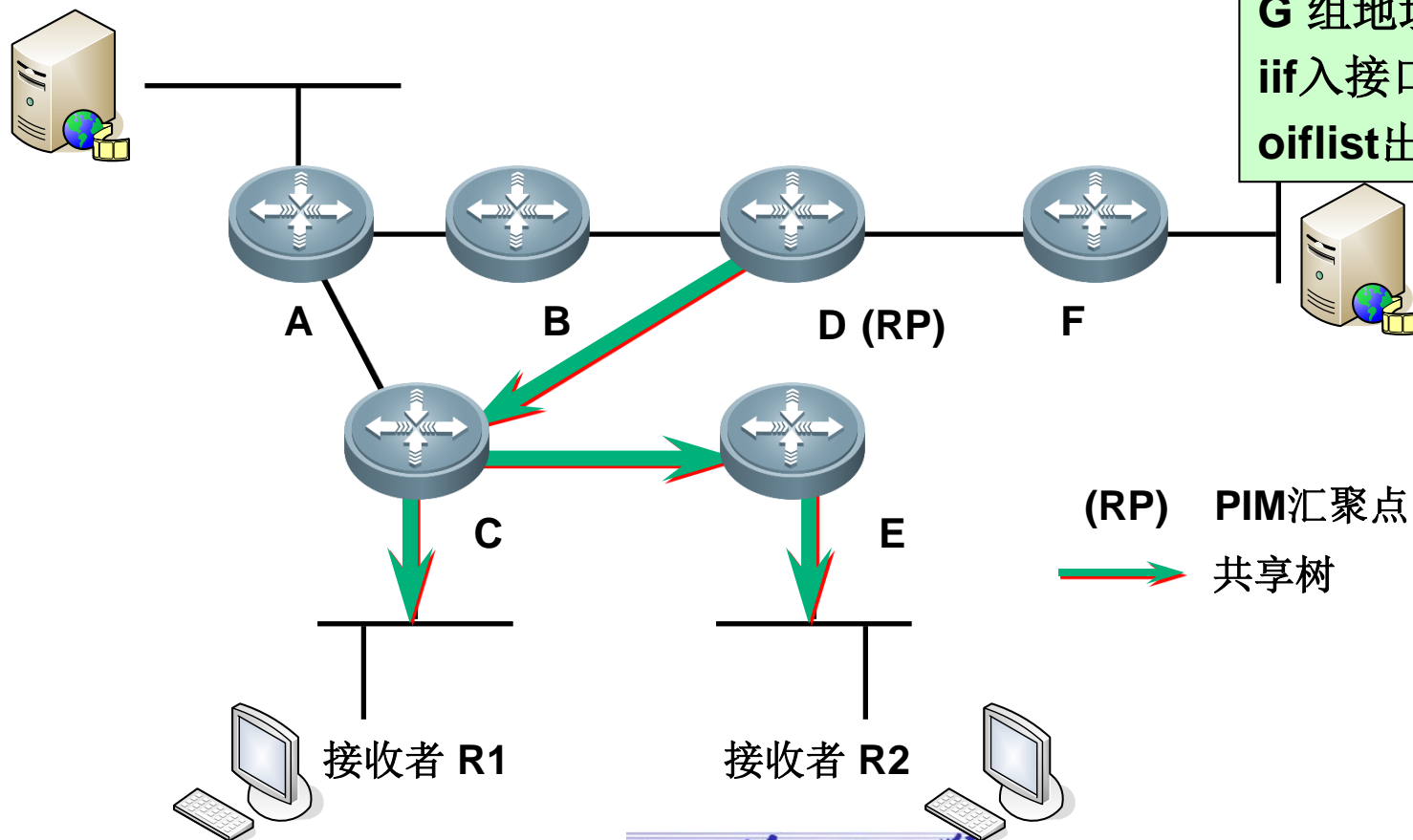
iif 入接口

oiflist 出接口列表

共享树

共享树的根不在组播数据流的源头，而位于网络中管理员指定的**汇聚点**。

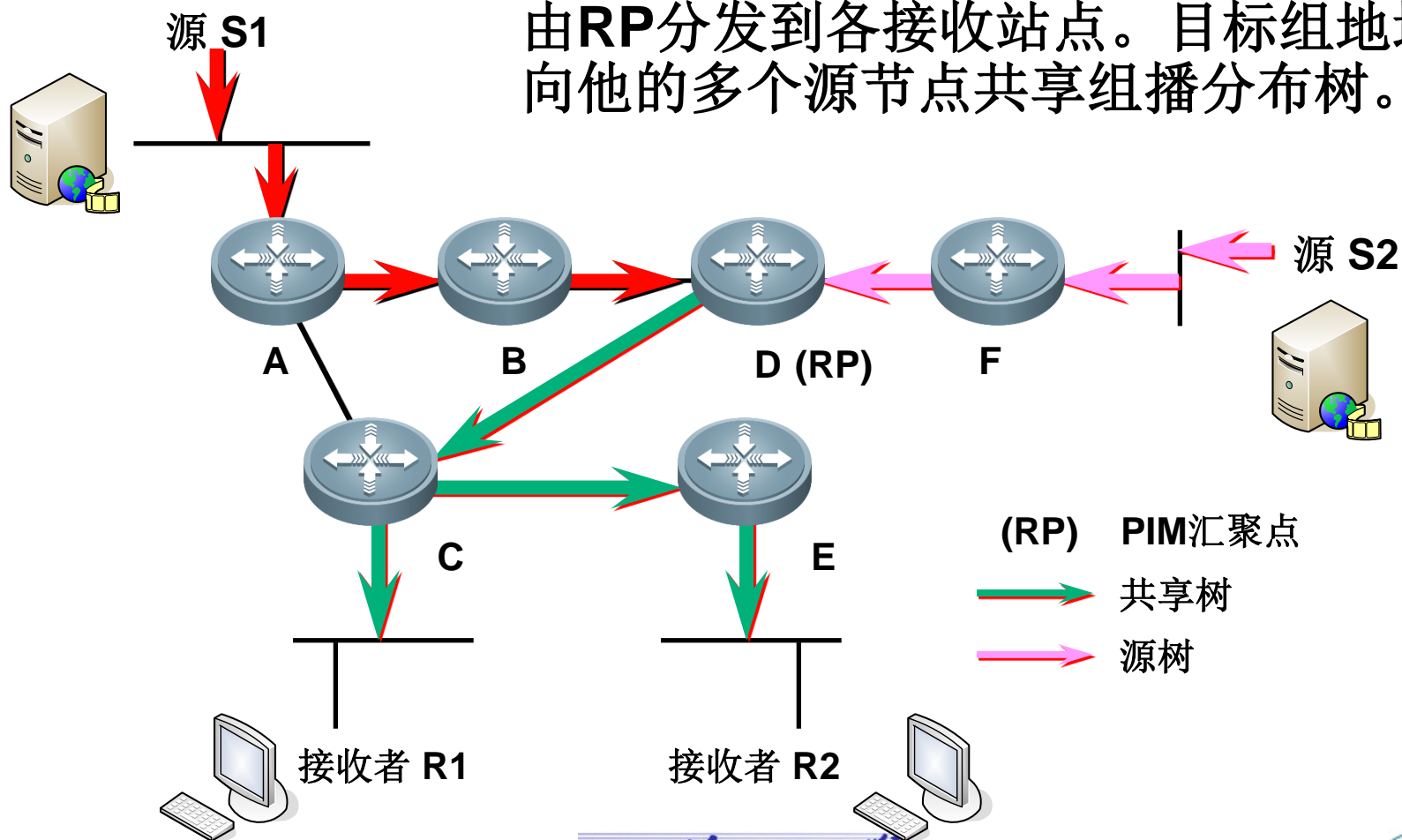
(RP: rendezvous point)



组播路由项
(*, G), iif, oiflist
* 任何源地址
G 组地址
iif入接口
oiflist出接口列表

组播分发树

所有组播源先把组播流发送到**RP**，再由**RP**分发到各接收站点。目标组地址向他的多个源节点共享组播分布树。



组播树总结

■ 有源树

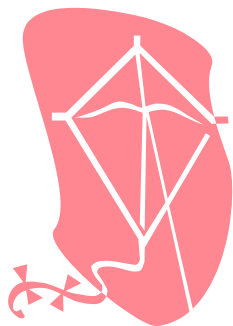
- ◆ 组播源作为根，分支指向接收者。
- ◆ 占用内存较多 (S,G)，但路径最优，接收者到组播源是最佳路径，延迟最小。

■ 共享树

- ◆ 该树的根称为聚合点 (rendezvous point, RP)。
- ◆ 占用内存较少 (*,G)，路径不一定是最优的，组播流转发必须先经过RP，引入额外的延迟。

组播服务

反向路径转发



组播转发

- 组播路由转发和单播路由转发：
 - ◆ 单播转发关心报文到哪里去。
 - ◆ 组播转发关心报文从哪里来。
- 组播路由转发使用 “反向路径转发” 机制防止循环

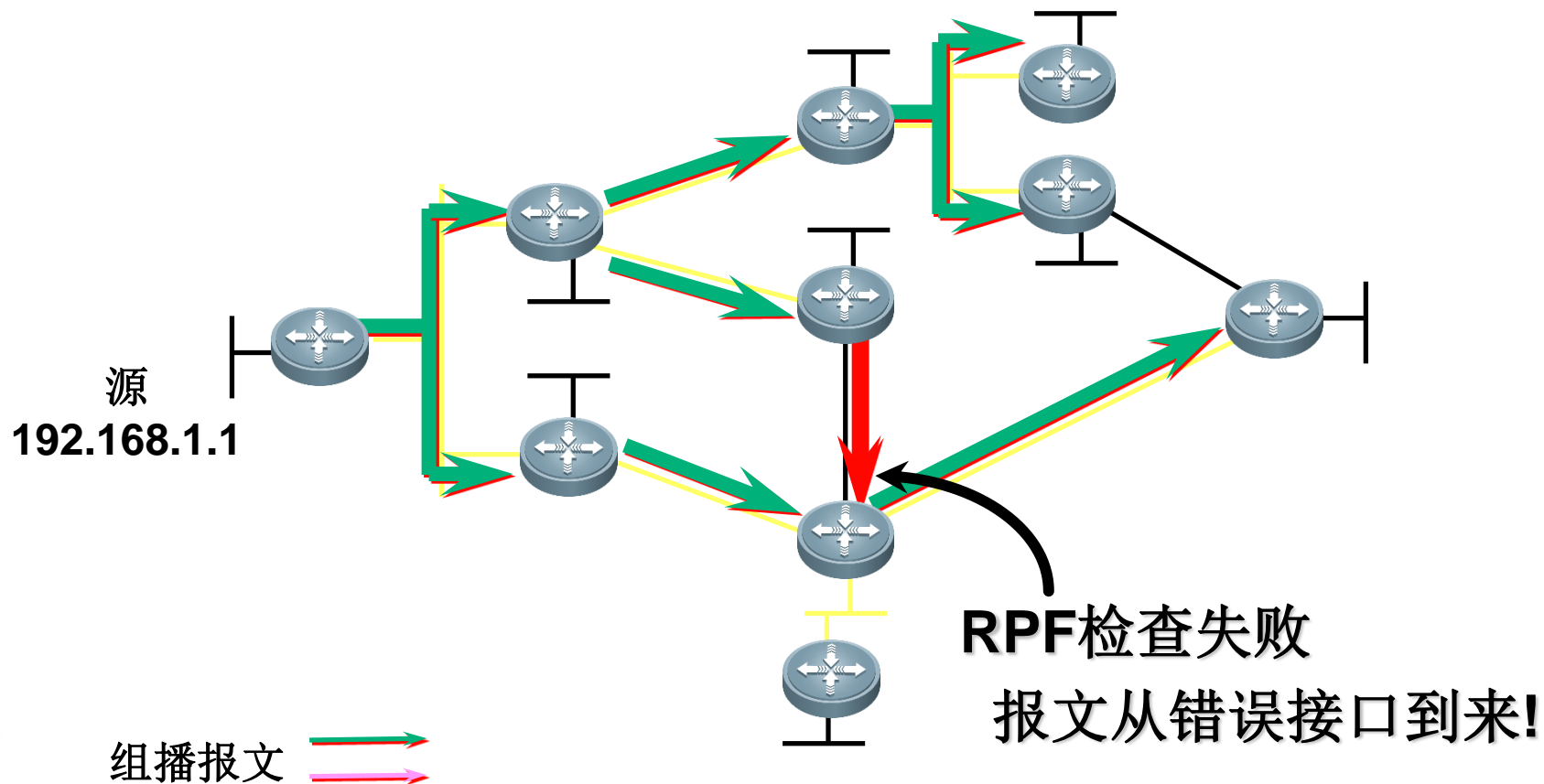
反向路径转发(RPF, Reverse Path Forwarding)

RPF

- 路由器收到组播数据后，进行RPF检查：
 - ◆ 在单播路由表中查找组播报文源地址的路由表项
 - ◆ 比较路由表项的出接口和组播报文的入接口：
 - 相符：RPF成功，向分发树下游转发
 - 不符：RPF失败，报文丢弃

反向路径转发(RPF, Reverse Path Forwarding)

举例: RPF检查



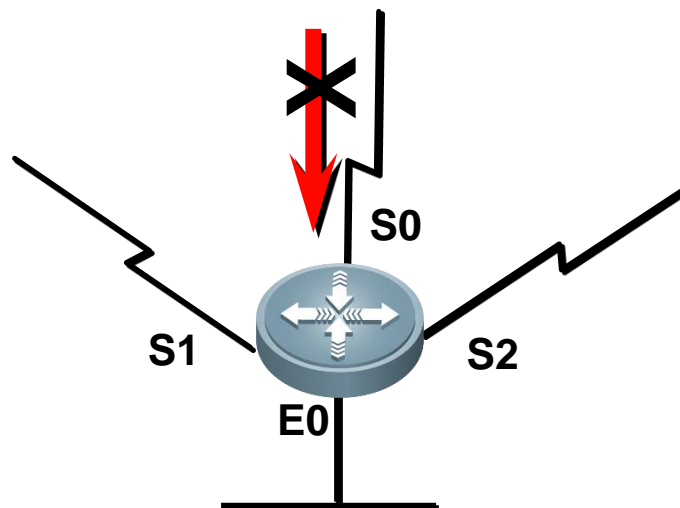
组播转发

看得更仔细点: RPF检查失败

源192.168.1.1
发出的组播数据报文

RPF检查失败!

单播路由表	
网络	接口
192.168.1.0/24	S1
198.14.32.0/24	S0
204.1.16.0/24	E0

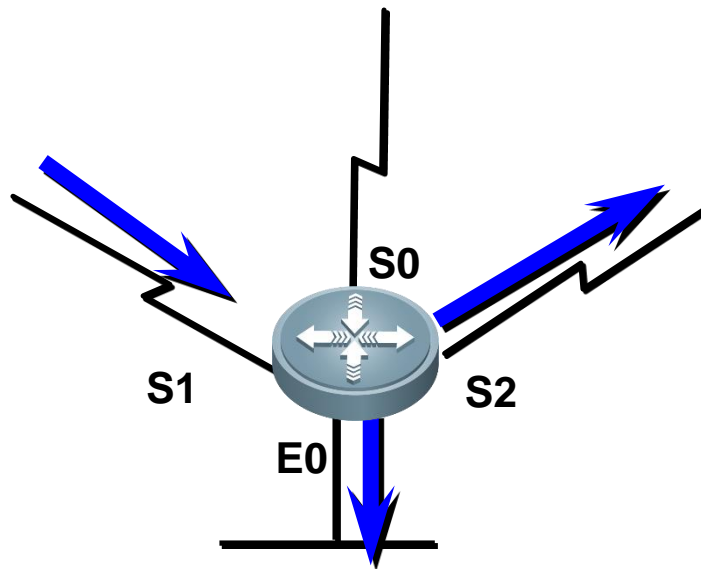


报文从错误接口到达
丢弃数据报文!

组播转发

看得更仔细点: **RPF**检查成功

源192.168.1.1
发出的组播数据报文

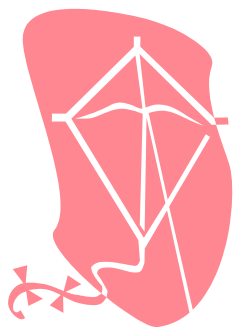


RPF检查成功!

单播路由表	
网络	接口
192.168.1.0/24	S1
198.14.32.0/24	S0
204.1.16.0/24	E0

数据报文从正确的接口到达!
向所有出接口 (即分发树的下游)转发

组播服务

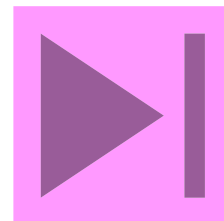


PIM

组播路由协议



密集模式 PIM-DM



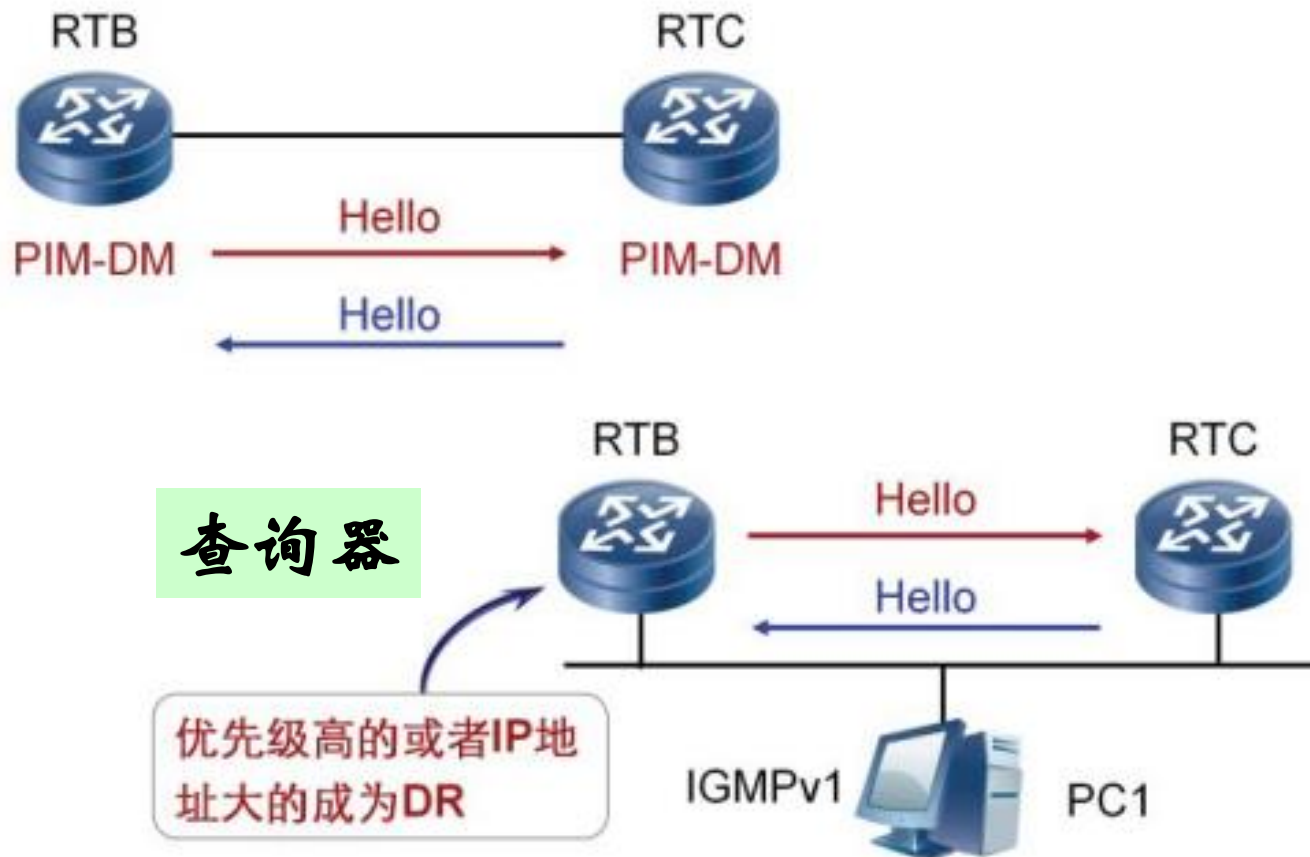
PIM-DM工作过程

- PIM-DM假设网络中的每个子网都存在至少一个对组播源感兴趣的接收站点
- PIM-DM的工作过程可以概括为：
 - ◆ 邻居发现、
 - ◆ 扩散、 Flooding
 - ◆ 剪枝、 Prune
 - ◆ 嫁接、 Graft
 - ◆ Assert 机制。

⑩周期性的扩散和剪枝现象是密集模式协议的特征。

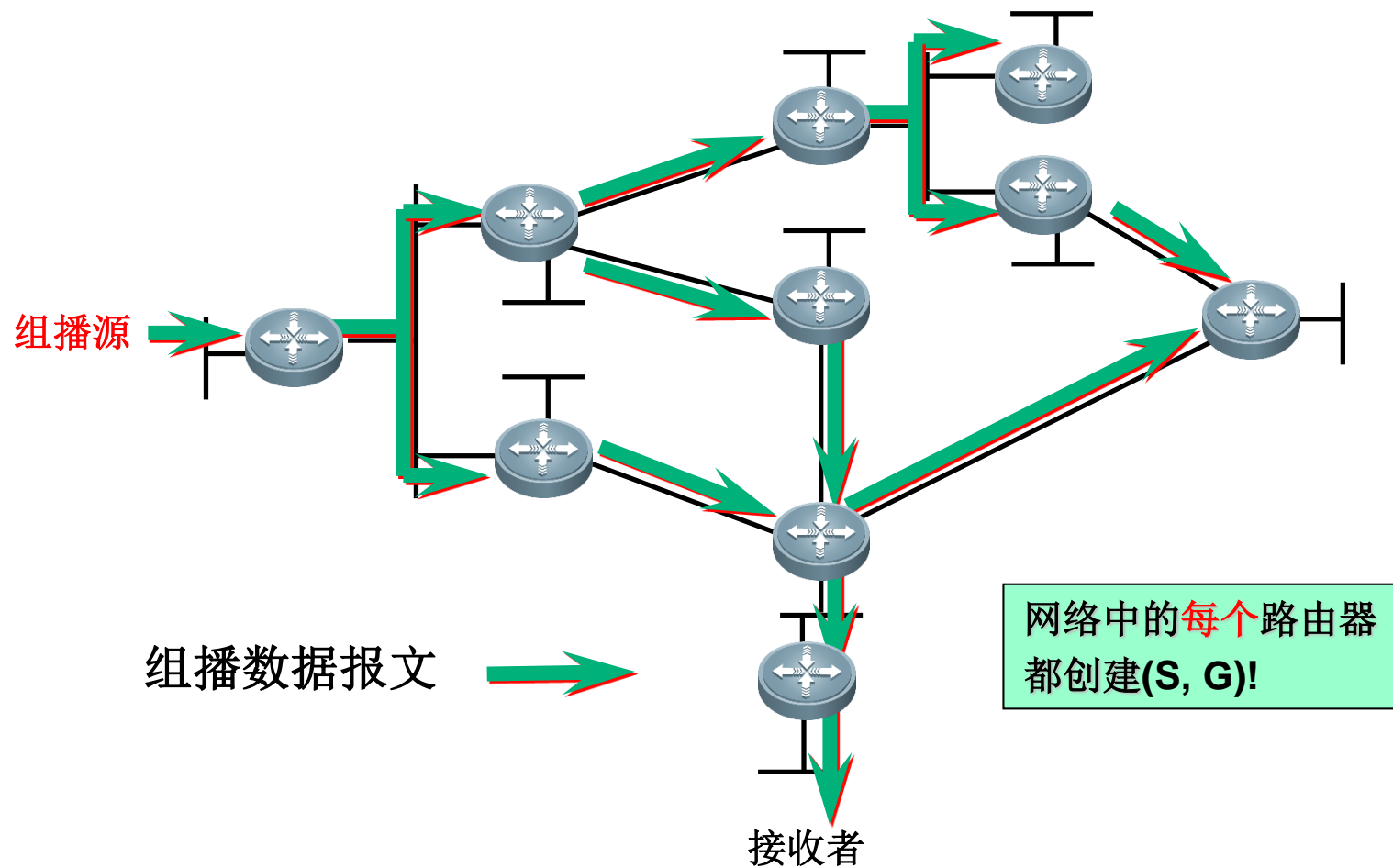
邻居发现

邻居发现Hello报文



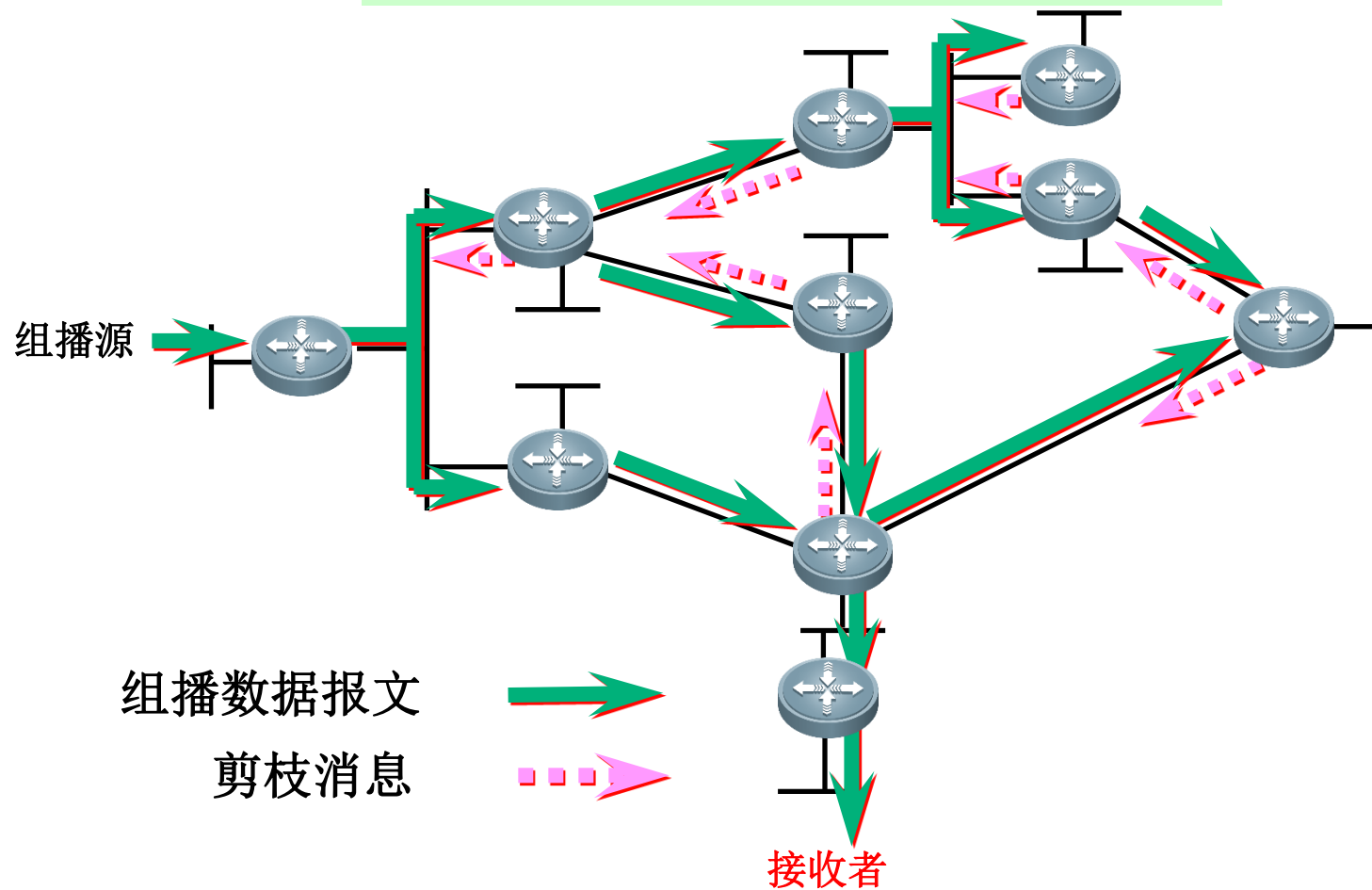
PIM-DM 泛滥

泛滥（扩散）

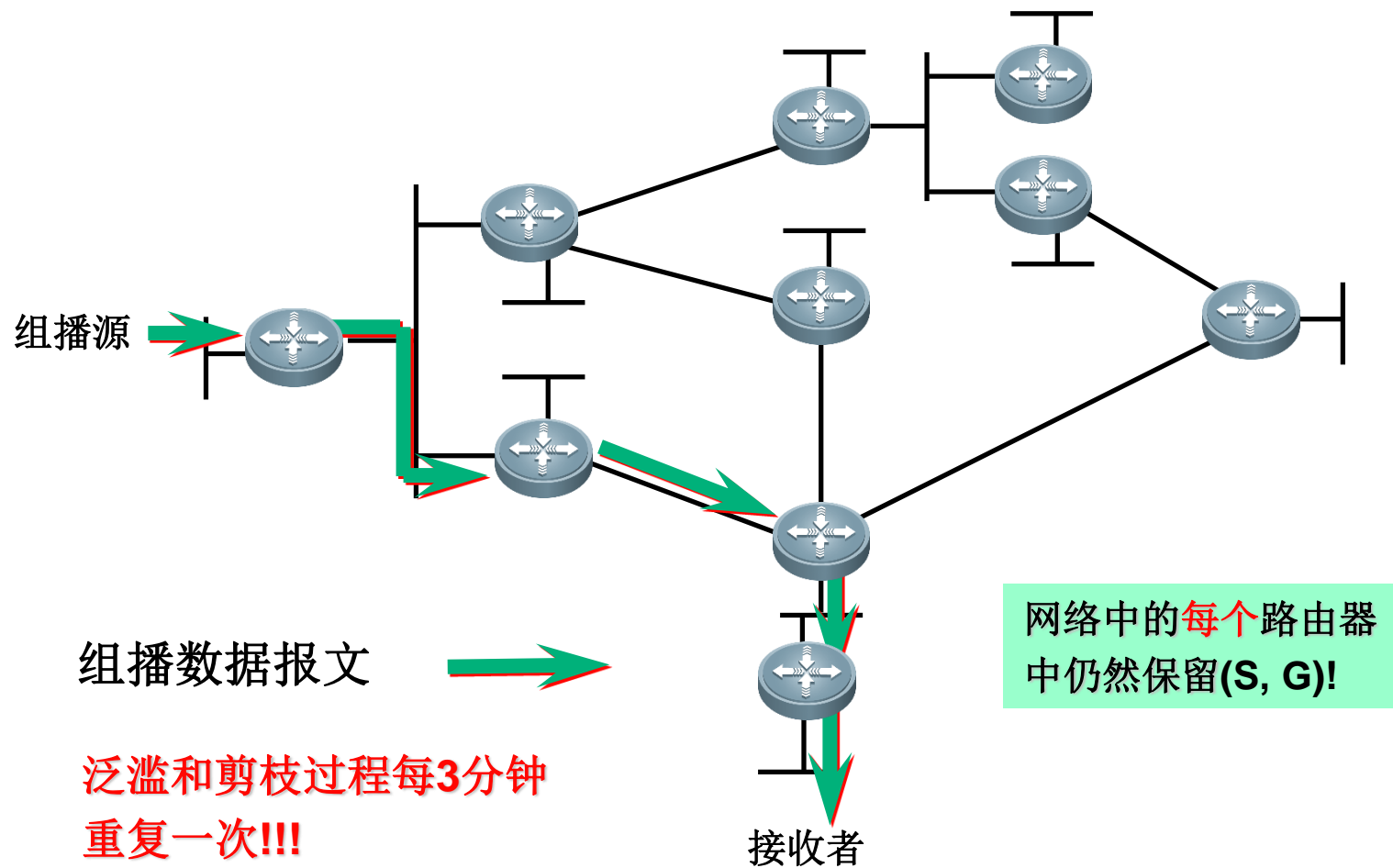


PIM-DM 剪枝

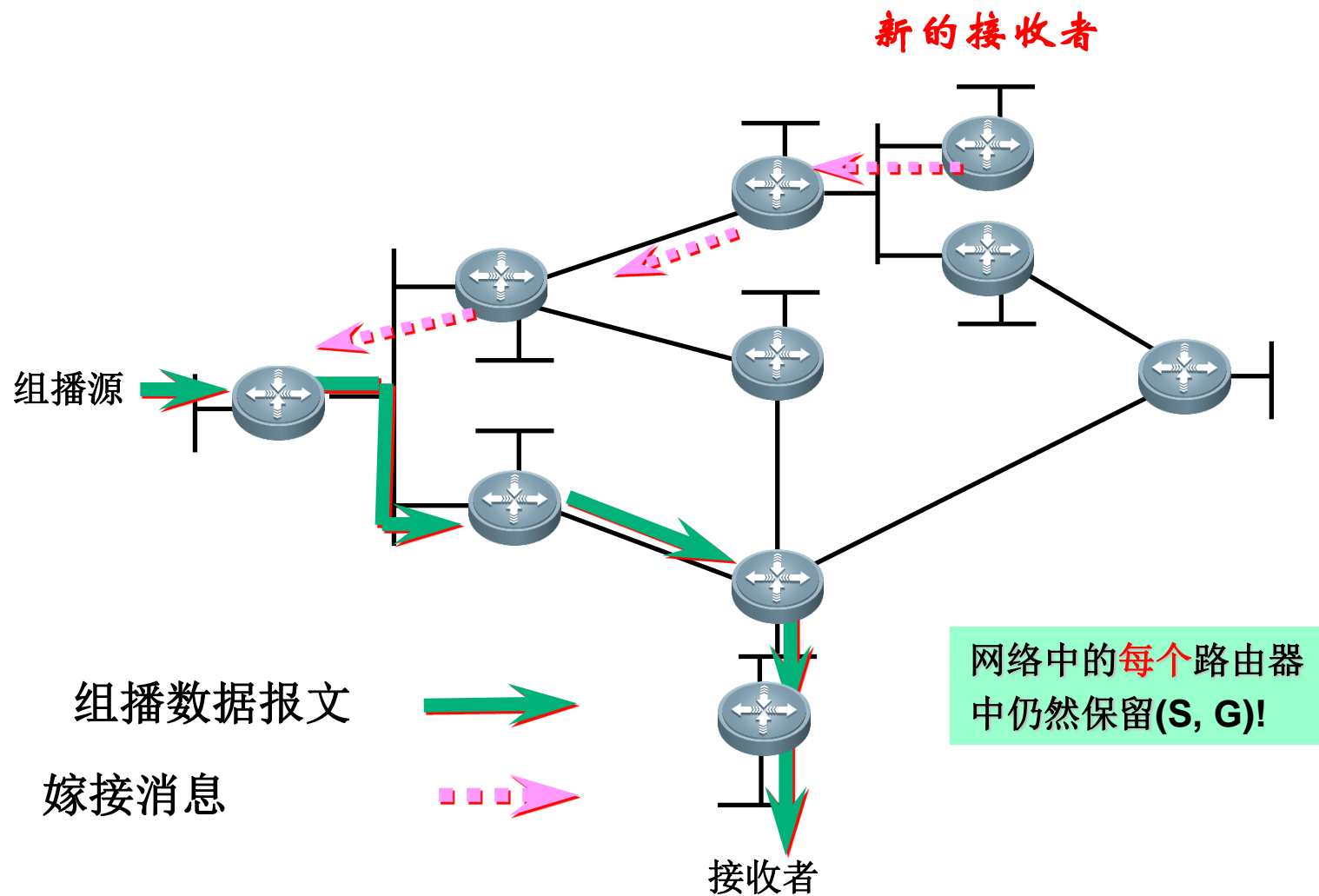
剪枝不需要的数据流



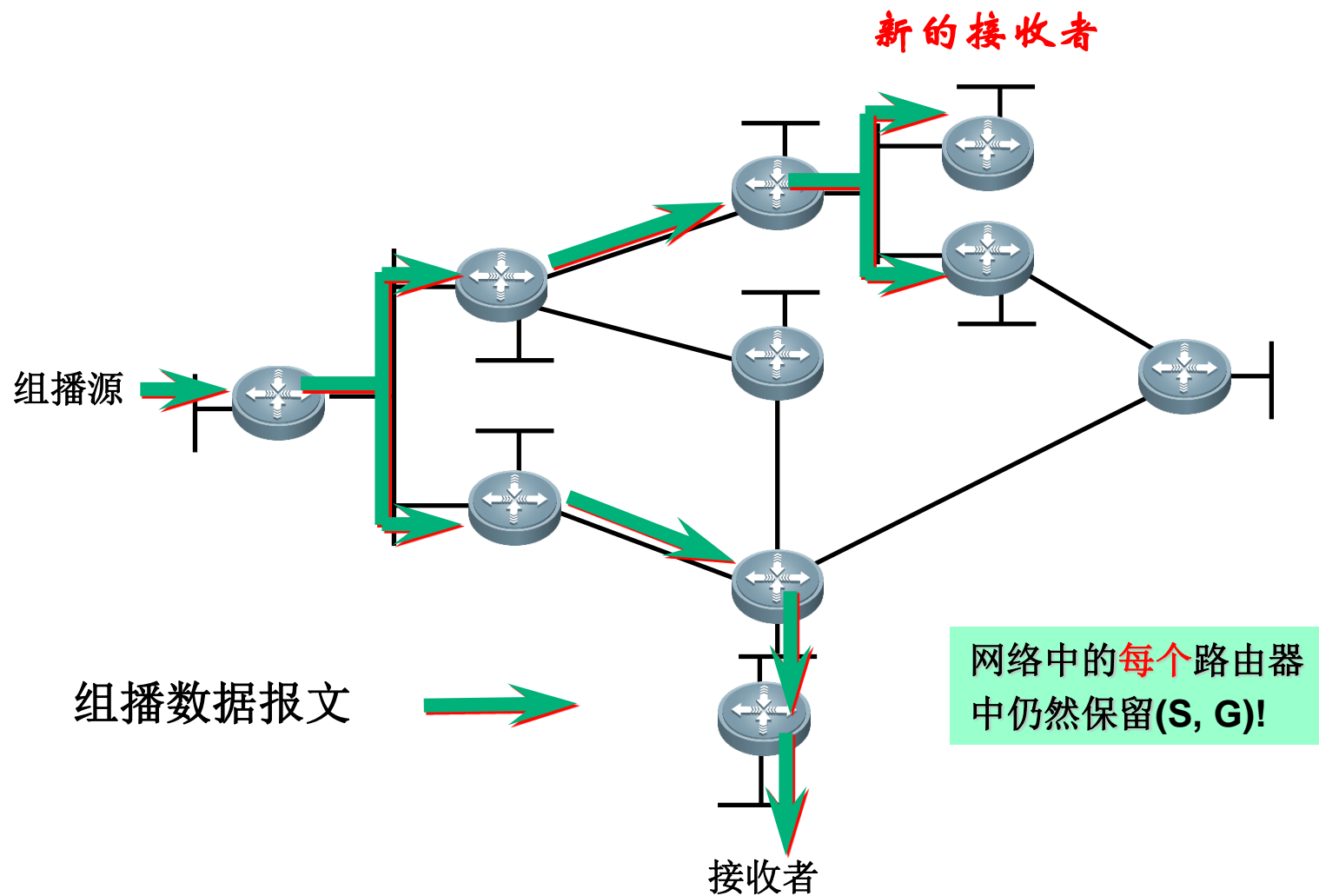
PIM-DM 剪枝完成



PIM-DM 嫁接



PIM-DM 嫁接

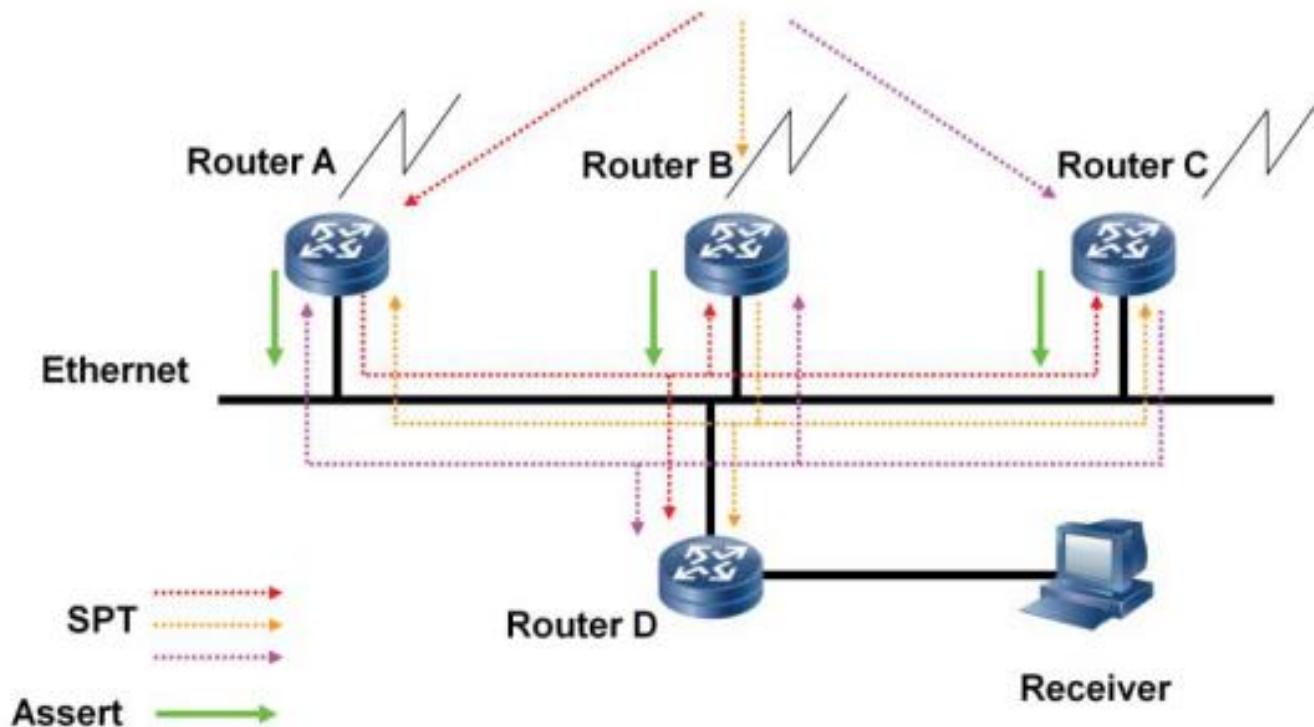


Assert 机制

⑩在共享网络（如Ethernet）中会出现相同报文的重复发送。

⑩需要通过Assert机制来选定一个唯一的转发者。

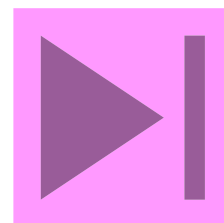
Assert机制



组播路由协议



稀疏模式 PIM-SM



PIM-SM

- **PIM-SM假设网络中的组成员分布非常稀疏，几乎所有网段均不存在组成员。**
 - ◆ 直到某网段出现组成员时，才构建组播路由，向该网段转发组播数据。
- **PIM-SM模型实现组播转发的核心任务是构造并维护一棵单向共享树。**
 - ◆ 共享树选择PIM中某一路由器作为公用根节点，称为汇聚点RP (Rendezvous Point)。
 - ◆ 组播数据通过RP沿共享树向接收者转发。

PIM-SM

■ 发送者和接收者在RP处进行汇聚

◆ 发送者

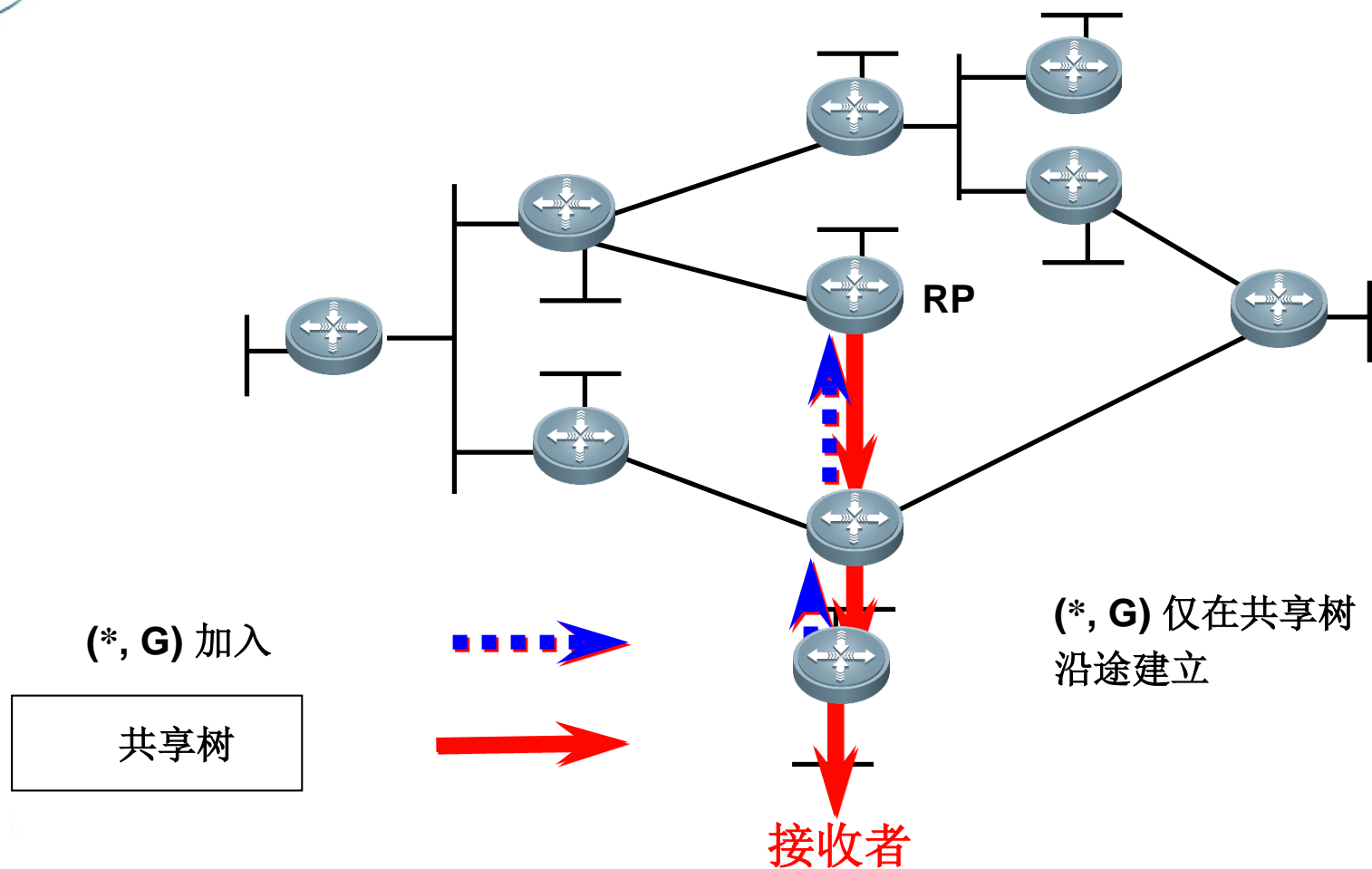
- 第一跳路由器把发送者注册到RP上

◆ 接收者

- DR（直连网络上的负责人）为接收者加入到共享树（树根在RP）

⑩ 共享网络（如Ethernet）选举DR（Designated Router），DR 将作为本网段中组播信息的唯一转发者。

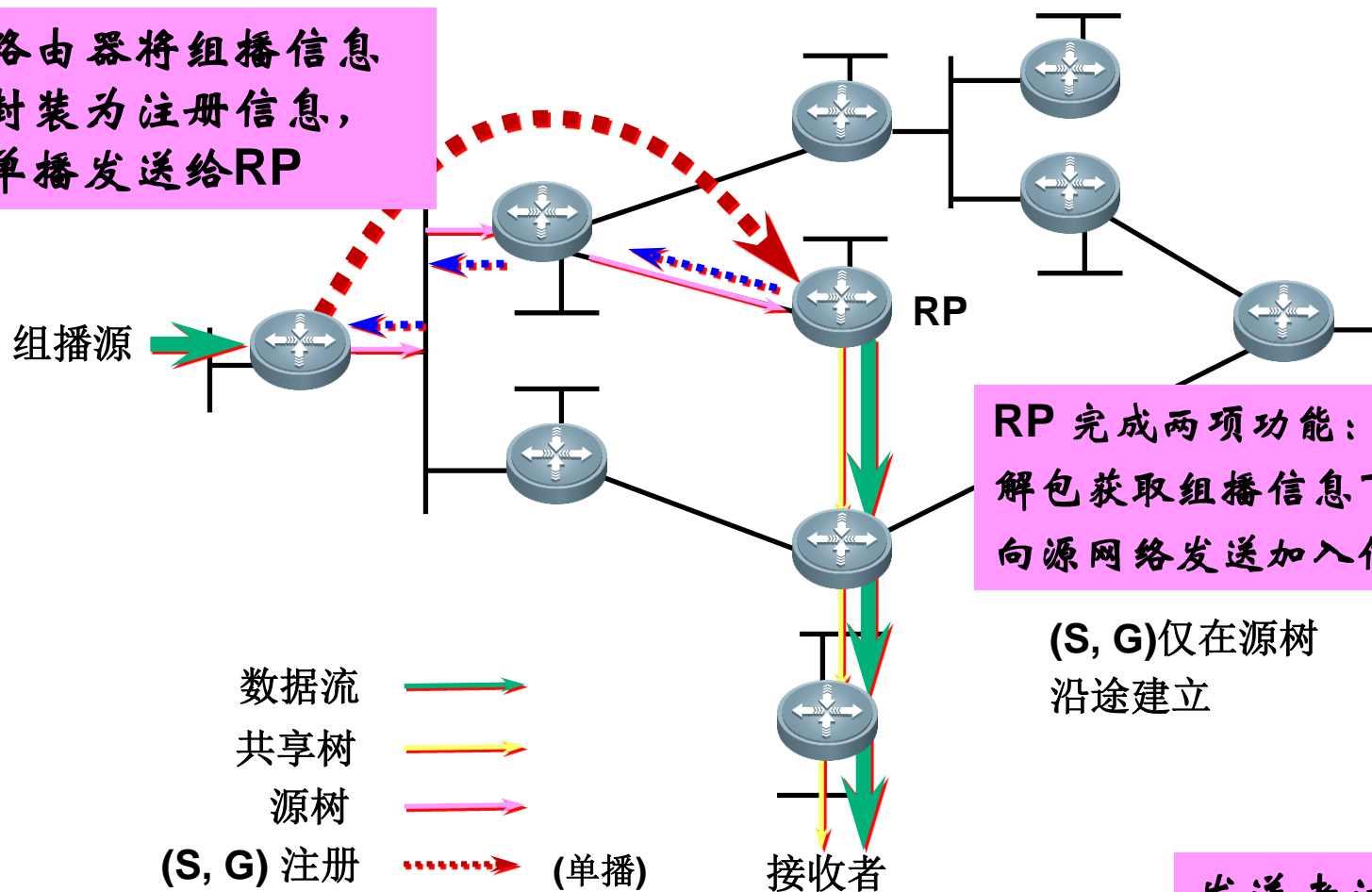
PIM-SM 接收者共享树加入



接收者加入

PIM-SM 发送者开始注册

路由器将组播信息封装为注册信息，单播发送给RP

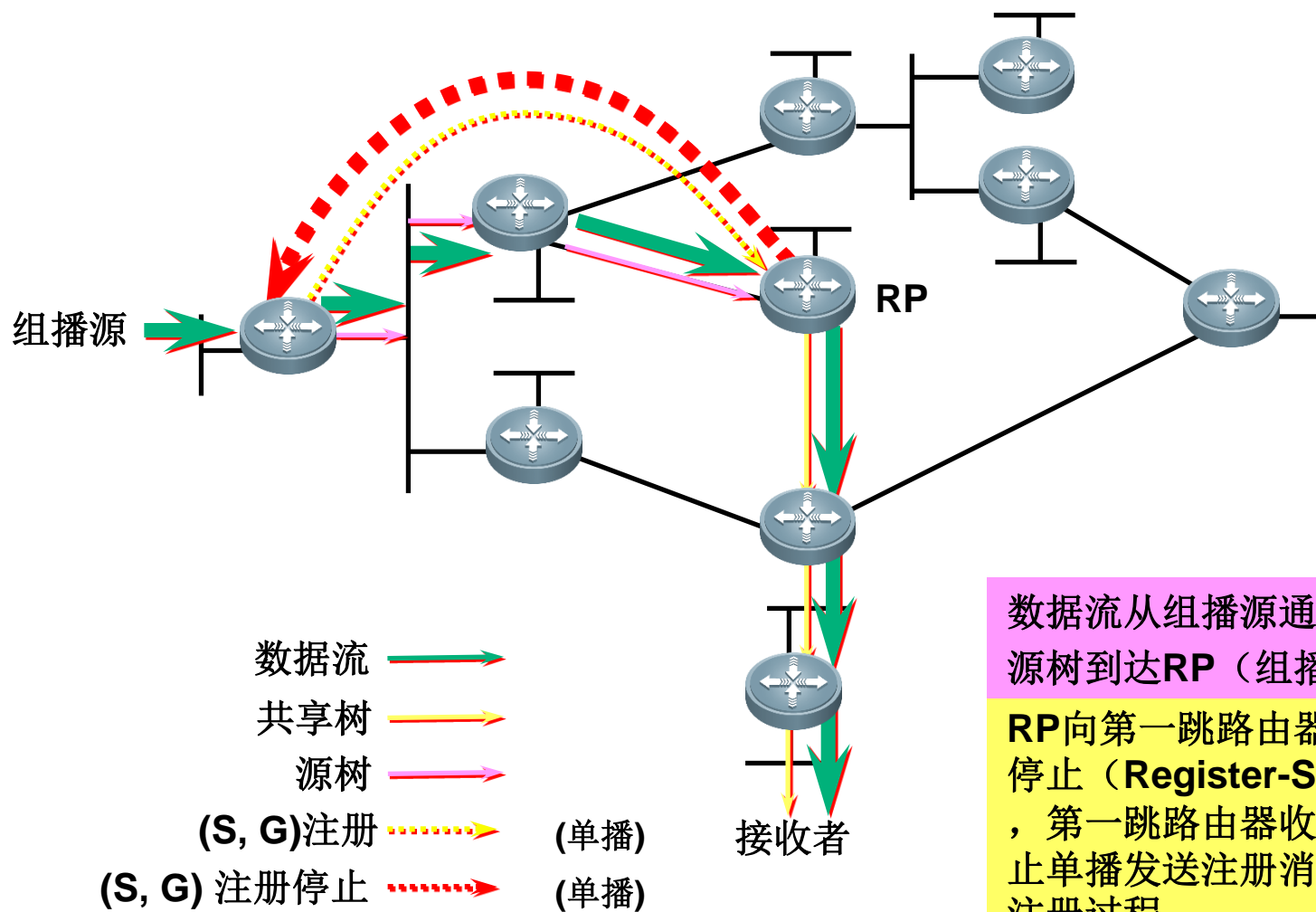


RP 完成两项功能：
解包获取组播信息下传。
向源网络发送加入信息，构造源树

(S, G) 仅在源树
沿途建立

发送者注册

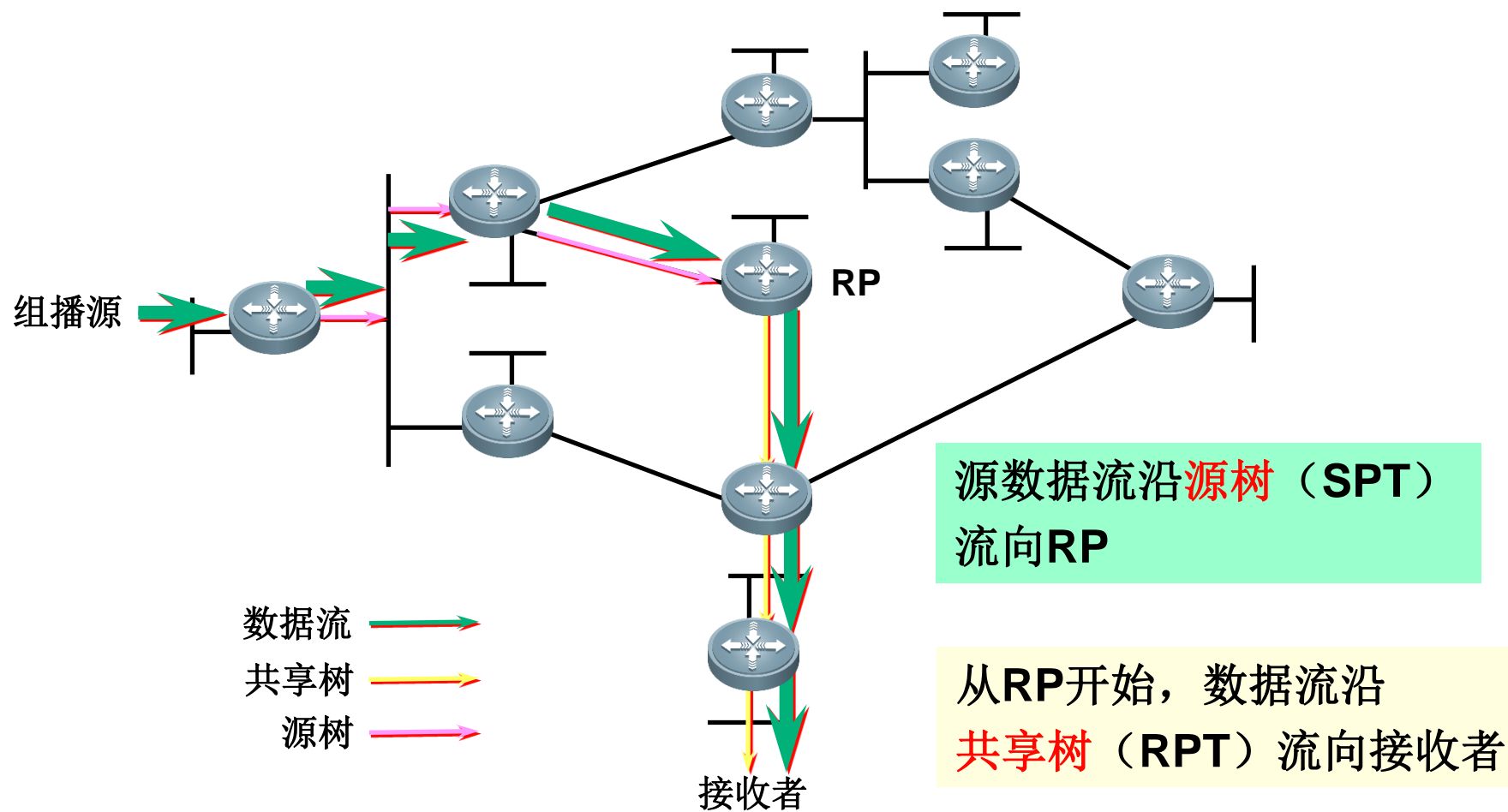
PIM-SM 发送者完成注册



数据流从组播源通过源树到达RP（组播发送）

RP向第一跳路由器发送注册停止（**Register-Stop**）消息，第一跳路由器收到后，停止单播发送注册消息，完成注册过程。

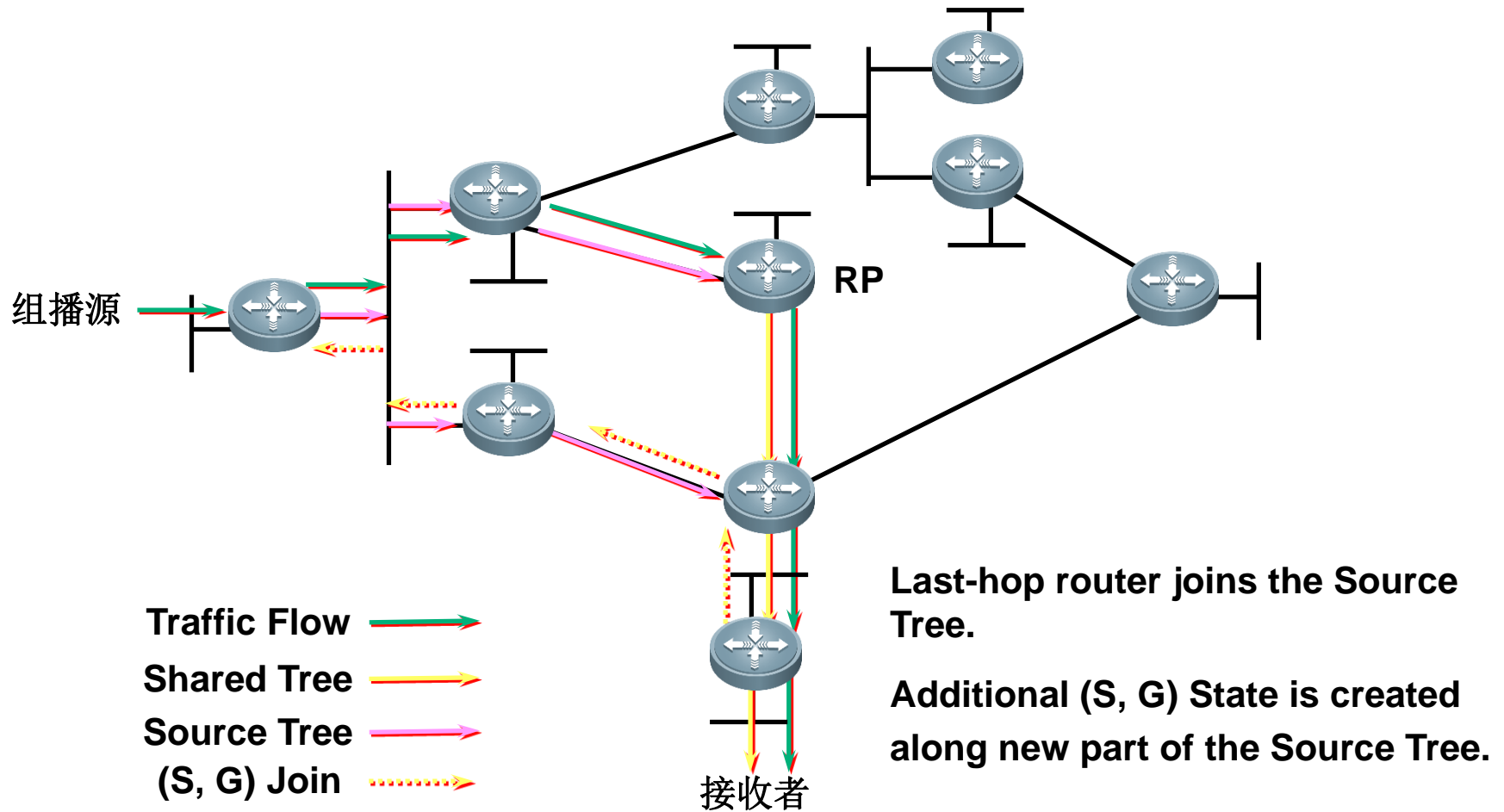
PIM-SM 组播正常



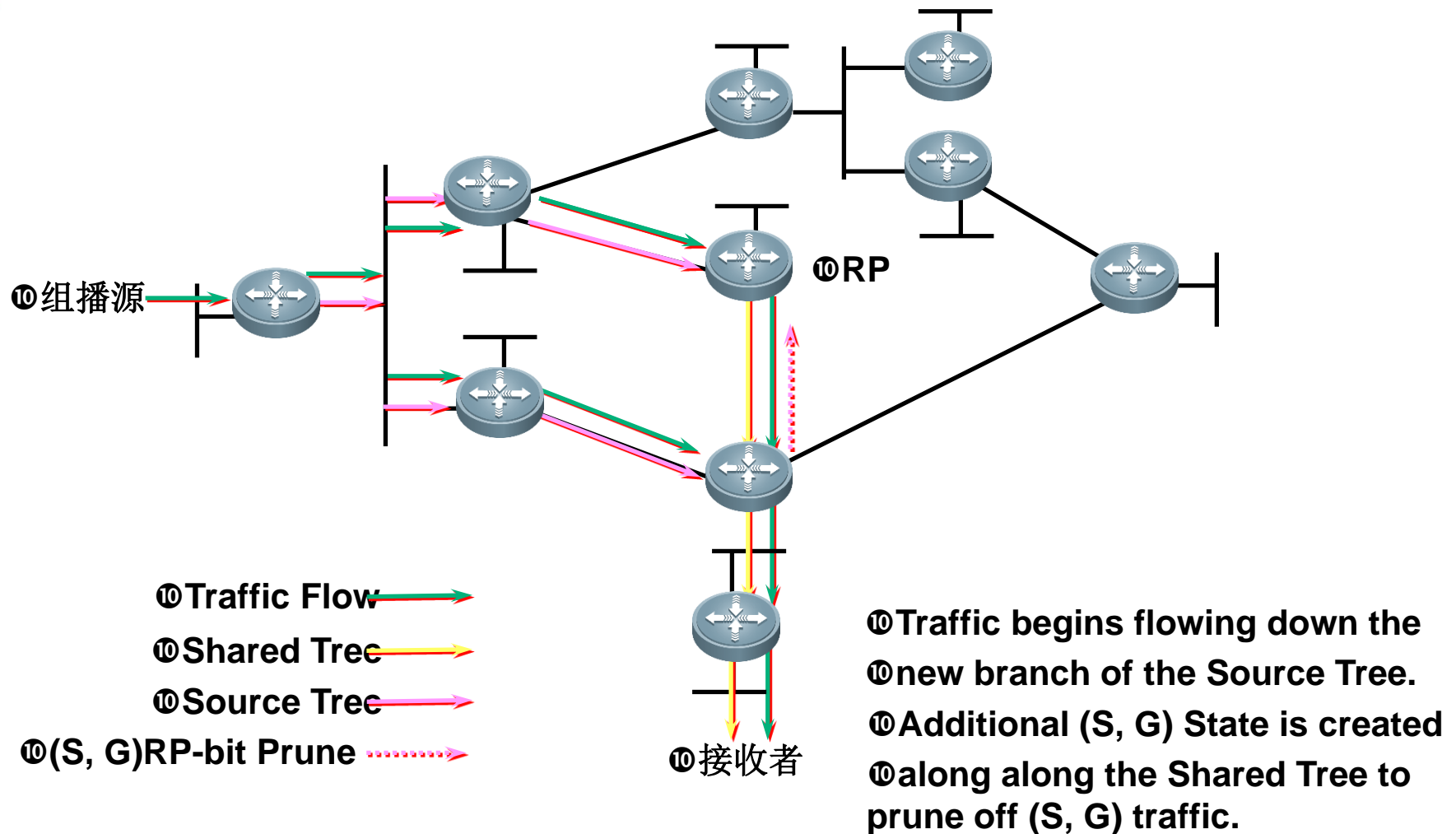
PIM-SM SPT 切换

- 针对特定的源，PIM-SM通过指定一个利用带宽的SPT阈值可以实现将最后一跳路由器（即离接收者最近的DR）从RPT切换到SPT。

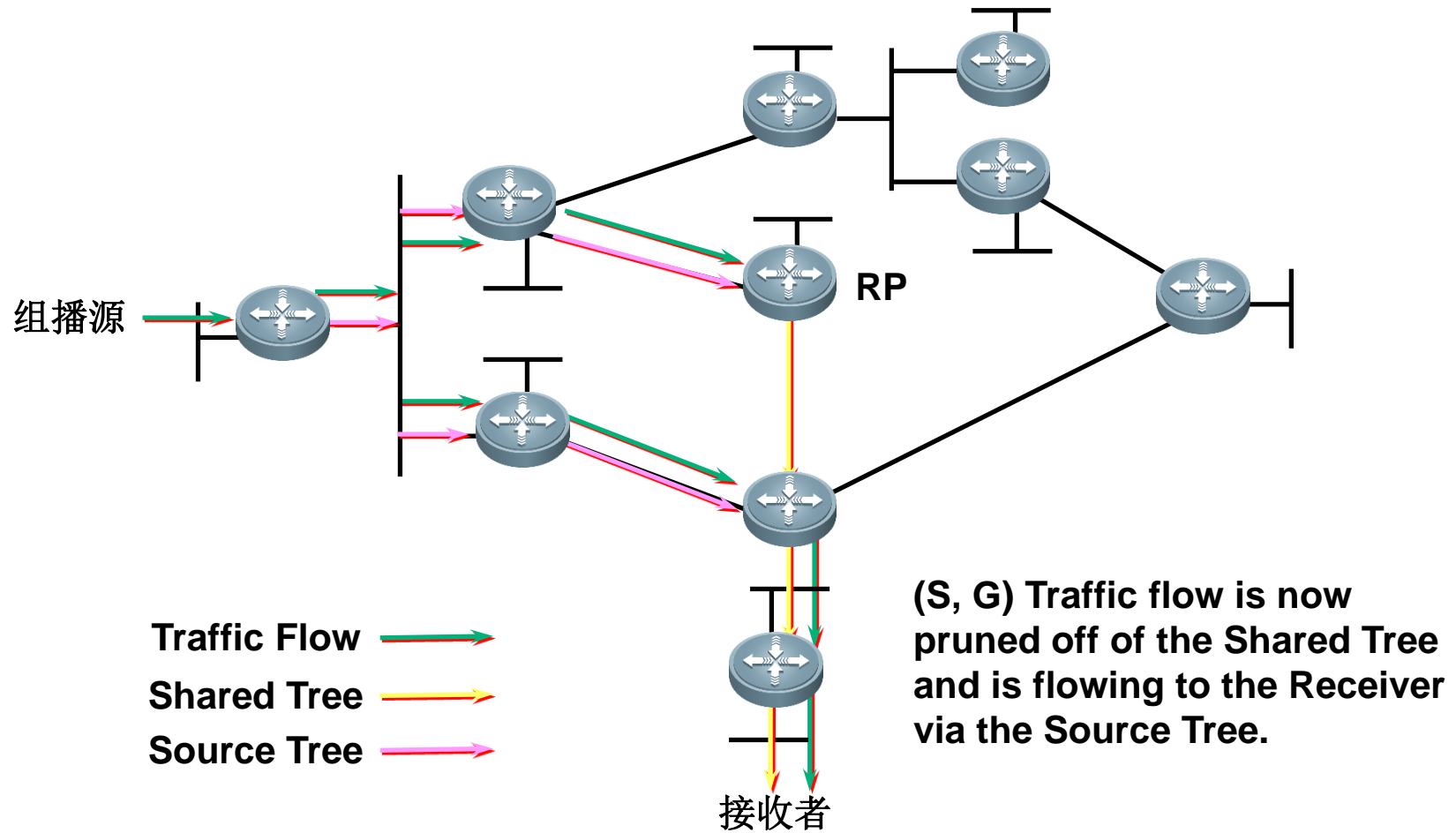
PIM-SM SPT 切换



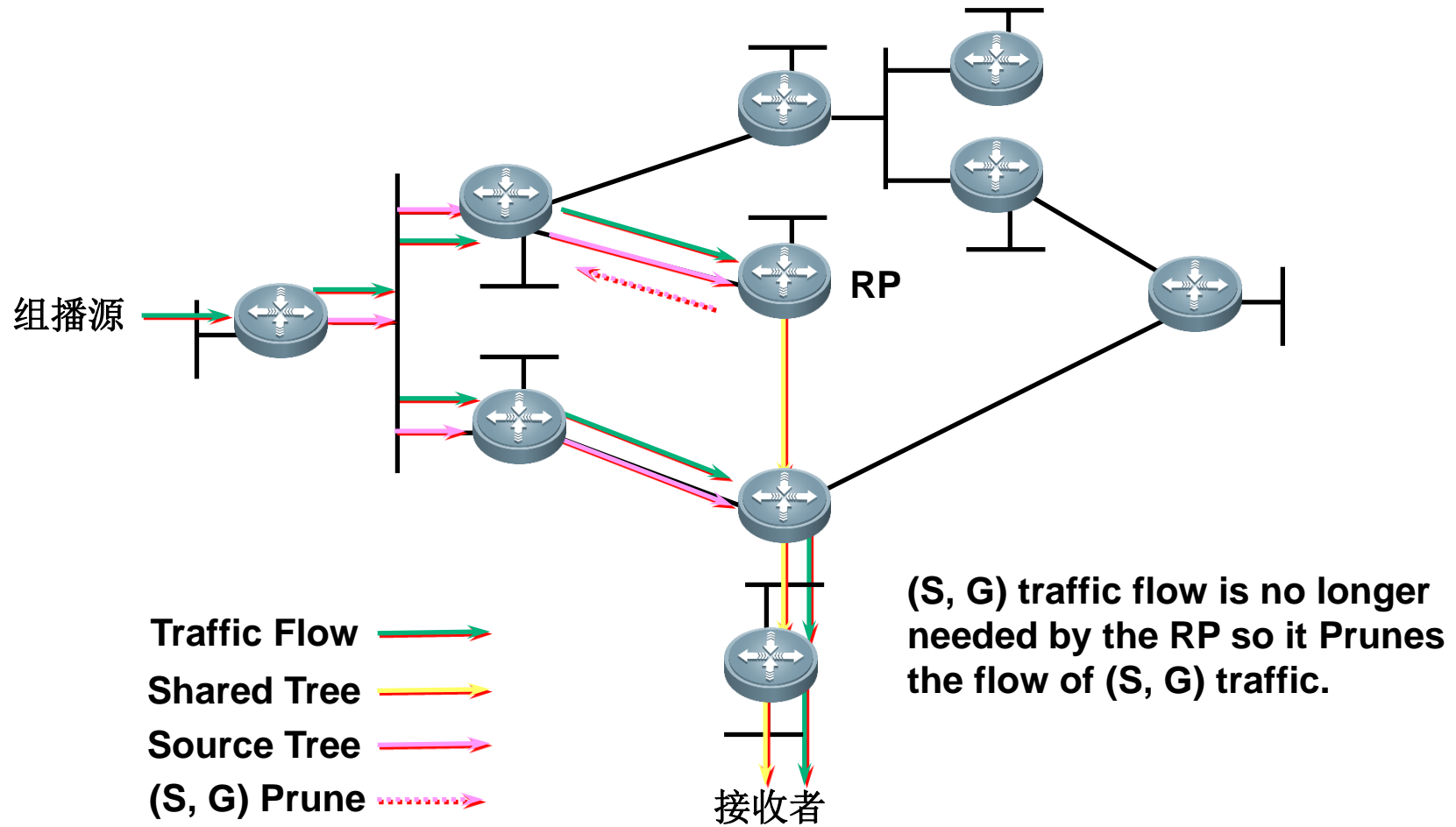
PIM-SM SPT 切换



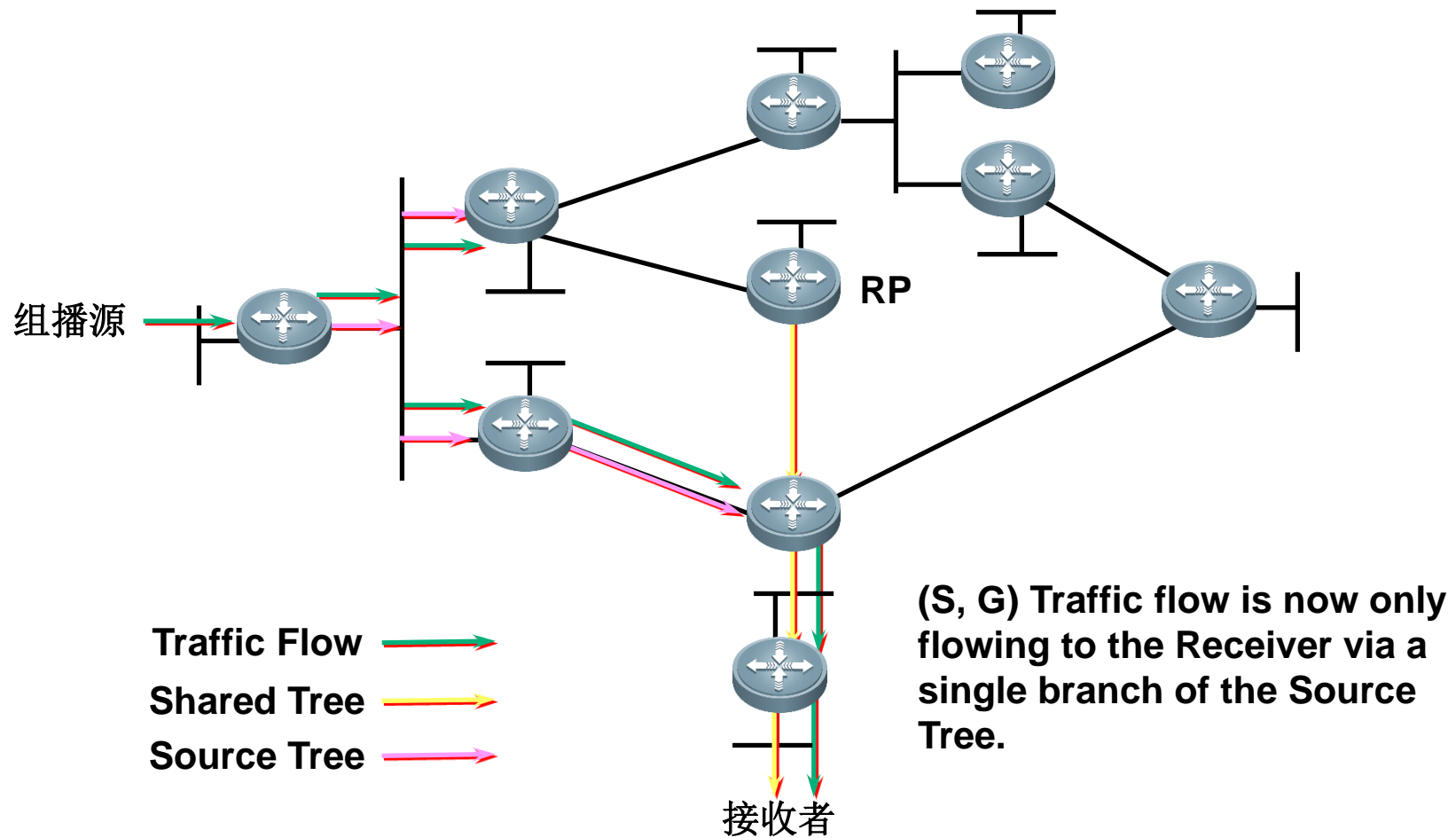
PIM-SM SPT 切换



PIM-SM SPT 切换



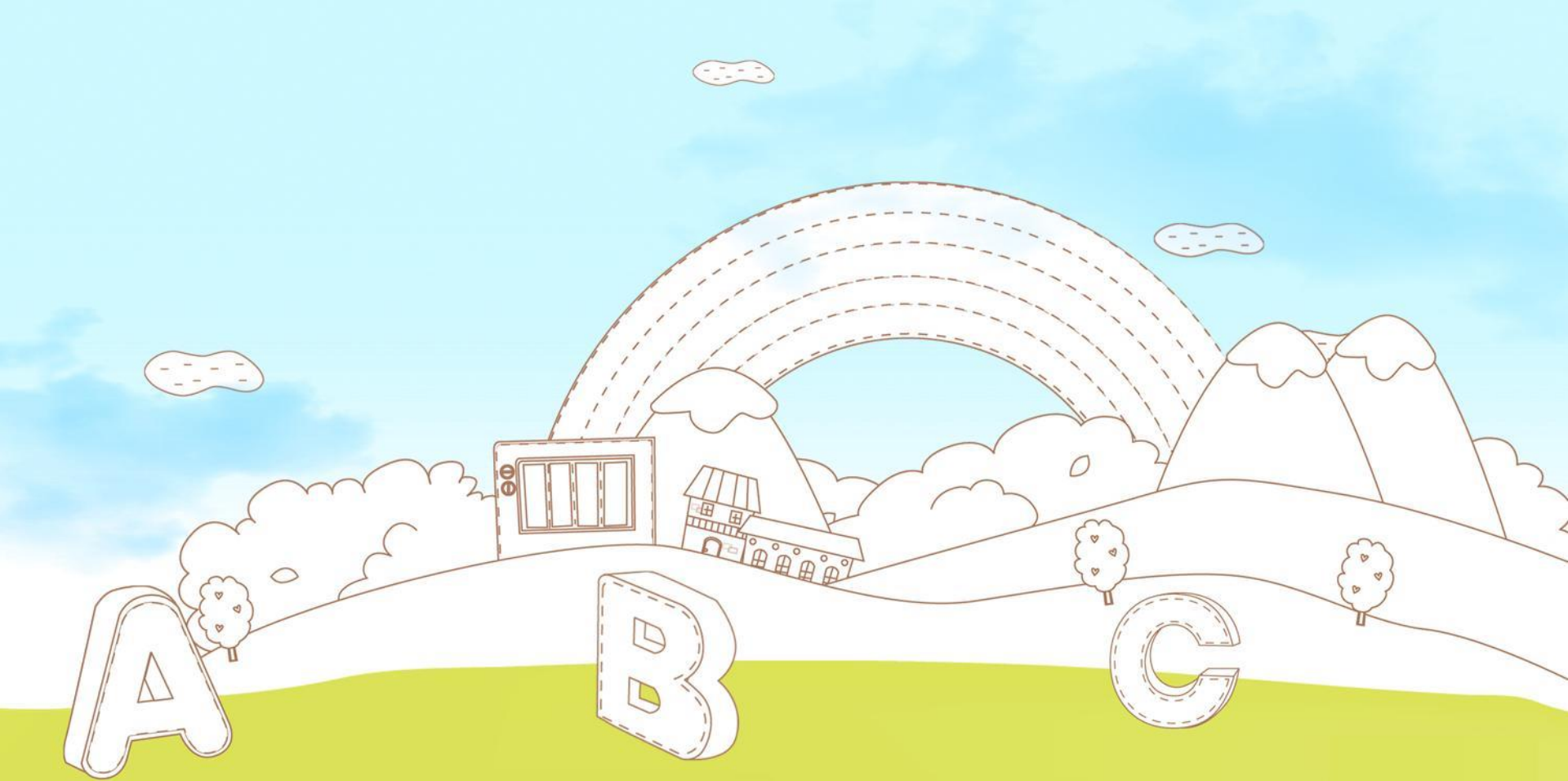
PIM-SM SPT 切换



PIM-SM工作过程

■ PIM-SM工作过程

- ◆ 邻居发现
- ◆ DR选举
- ◆ RP发现：手工指定选举产生
- ◆ 加入 (Join)：接收者叶子路由器到RP
- ◆ 剪枝 (Prune)：下游组播组成员全部离开
- ◆ 注册 (Register)：向RP通知组播源S的存在。
- ◆ SPT切换：实现DR从RPT切换至 SPT。



Thank You !