

而且不是可扩展的。Cisco 开发了它自己的组播路由协议，称为协议无关组播（Protocol Independent Multicast, PIM）。PIM 使用单播路由表作出转发决定。因此，路由器对单播路由协议的选择可以是本书中介绍的任何一个协议——RIP、IGRP、EIGRP、OSPF、BGP 或者 IS-IS。PIM 按照两种不同的模式工作——密集模式和稀疏模式。每一种模式都有它自己的优点和缺点，而且每一种模式都有不同的实现方法。

密集模式使用泛洪和剪枝（flood-and-prune）机制转发组播流量。密集模式非常容易实现，其复杂度比稀疏模式小；但是，在大型网络中，密集模式不是可扩展的。因此，密集模式更适合于一个小型的组播环境。

稀疏模式使用显式的组加入（group-join）机制转发组播流量。与密集模式不同，稀疏模式有很好的可扩展性，可以运行在一个大型的组播环境中。因为它有更好的可扩展性，所以它的实现比密集模式更复杂，这一点对诊断而言，意味着更大的难度。

12.1 IGMP 版本 1、IGMP 版本 2 和逆向路径转发的基本原理

在深入讨论复杂的 PIM 协议之前，首先需要理解支持因特网组管理协议（Internet Group Management Protocol, IGMP）和逆向路径转发（reverse path forwarding, RPF）的概念。

IGMP 是在主机（也称为接收者）和启用组播的路由器之间行使职责的协议。简而言之，IGMP 使该路由器能够知道一个主机对接收发往一个特定组的组播流量感兴趣。无论何时，只要 PIM 被启用，这一路由器上的 IGMP 就被启用。IGMP 消息的 TTL 值为 1 以后被发送出去。因此，IGMP 分组仅限于本地网络。

12.1.1 IGMP 版本 1

在 IGMP 版本 1（在 RFC 1112 中定义）中，路由器向组播地址为 224.0.0.1 的“所有主机”发送 IGMP 查询，以请求有活跃的组播接收者的组播组。这些组播接收者也可以向该路由器发送 IGMP 报告，以提醒路由器：它们对接收一个特殊的组播流感兴趣。主机可以异步发送该报告或者响应这些由该路由器发送的 IGMP 查询。如果对于同一个组播组而言不止一个组播接收者存在，那么这些主机中只有一个主机发送一条 IGMP 报告消息；其他主机取消它们的报告。

例如，在图 12-1 中，路由器 R1 向 224.0.0.1 地址定期发送 IGMP 查询。各子网的每一组中只有一个成员向这个路由器发送 IGMP 报告消息（在这个例子中是 H2），而其他主机 H1 和 H3 则取消它们的报告。

在 IGMP 版本 1 中，没有对 IGMP 查询者的选择。如果这个网段中存在不止一个路由器，那么所有路由器都定期发送 IGMP 查询。IGMP 版本 1 没有专门用于主机离开组的机制。如果这些主机不再对接收发往特定组的组播分组感兴趣，那么它们就简单地不响应从该路由器中发送的 IGMP 查询分组。但该路由器继续发送查询分组。如果这个路由器没有在一个 IGMP 查询中听到一个响应，那么这一组就被暂停，这个路由器停止在用于这一组的那个网段上发

送组播分
个路由器

12

IGMP
更有效率

●
●
●

版本
当一个主
它发送一
对这个组
路由器假
在图
由器 R1
感兴趣。

送组播分组。如果在这一暂停时间段之后，主机又想要接收组播分组，那么该主机只要向这个路由器发送一个新的 IGMP 加入请求，该路由器就开始再次转发组播分组。

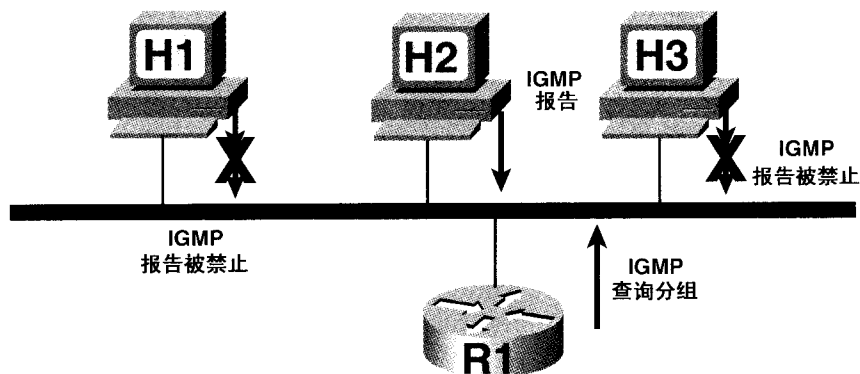


图 12-1 IGMP 版本 1 示例

12.1.2 IGMP 版本 2

IGMP 版本 2（在 RFC 2236 中定义）引入了几个改动，从而使 IGMP 在加入和离开组时更有效率。这里列出了其中一些重要的改动：

- **查询者选举机制**——在一个多接入网络中，一个 IGMP 查询者路由器的选举是基于 IP 地址的。因此，每个网段只有一个路由器发送 IGMP 查询。
- **离开组消息**——如果一个主机不再对一个组播组感兴趣，那么它发送一条离开消息。与版本 1 相比，这样做减少了离开等待时间。
- **特定组的查询**——路由器在它暂停一个组之前发送一个特定于组的查询。这样做确保该网段上不再有对接收该组播组感兴趣的主机。

版本 2 中的 IGMP 加入机制与版本 1 相同。但是，离开机制有一点不同。在版本 2 中，当一个主机希望离开这一组时，它发送一条 IGMP 离开消息。路由器在收到这一离开消息时，它发送一个 IGMP 查询分组，以查看是否有任何其他主机对该组感兴趣。如果一些主机仍然对这个组感兴趣，那么它们发送一条 IGMP 加入消息覆盖这条 IGMP 离开消息；否则，这个路由器假设没有其他主机对该组播组感兴趣，这个组被暂停。

在图 12-2 中，主机 H2 希望离开这个组播组，因为它发送了一条 IGMP 离开消息。当路由器 R1 收到这一离开消息时，它立即发送一个 IGMP 查询，以确保没有其他主机对这个组感兴趣。在这个例子中，主机 H1 仍然在这一组中，H1 通过发送一条 IGMP 报告消息覆盖了

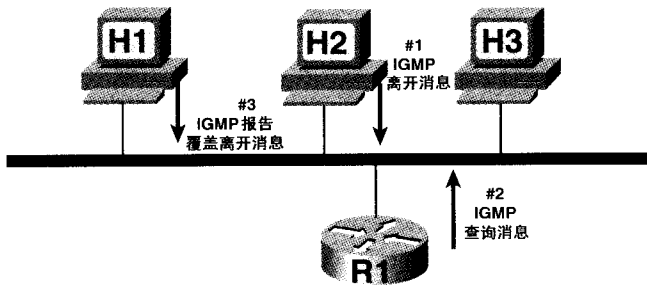


图 12-2 IGMP 版本 2 离开机制——组播组被保留

IGMP 离开消息。在这种情况下，这个组继续活跃，路由器保持在这个网段上转发组播分组。

图 12-3 描述了与图 12-2 中一样的情况。路由器 R1 在收到来自主机 H2 的 IGMP 离开消息之后发送一个 IGMP 查询。因为主机 H1 和 H3 都对这个组不感兴趣，所以它们简单地忽略从路由器发送的 IGMP 查询消息。当 R1 没有听到来自对其 IGMP 查询的任何响应时，它暂停这个网段上的该组，路由器 R1 停止发送那个组的组播分组。

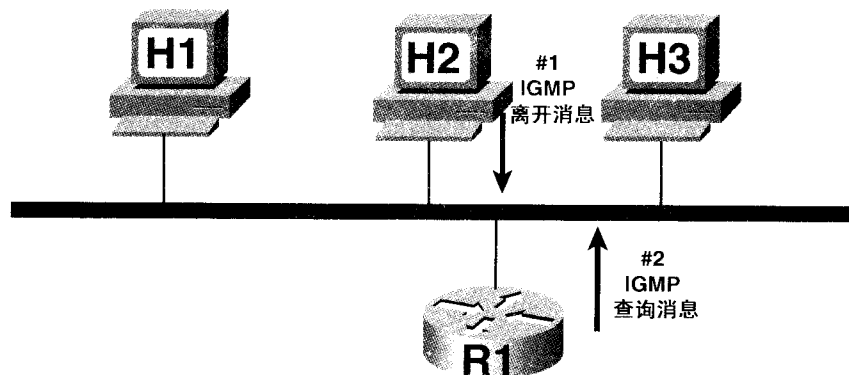


图 12-3 IGMP 版本 2 离开机制——组播组暂停

12.1.3 组播转发 (RPF)

在理解 PIM 如何工作之前，必须理解组播转发机制的工作方式。组播路由与单播路由相反。单播路由关心这一分组发往哪里，而组播路由关心的是这些分组来自哪里。组播路由使用 RPF 的概念确定是否存在一个转发循环。简而言之，RPF 允许该路由器只有在在一个组播分组是在源地址的上游接口上接收的情况下才转发该分组。具体地说，如果该路由器在某接口上接收到一个组播分组，则使用单播路由表来检查这个组播分组的源地址。如果这个分组是在单播路由表中指定用于源地址的接口上接收的，那么 RPF 检查成功；否则，RPF 检查失败，这些组播分组被丢弃。

在图 12-4 中，路由器从接口 S1 收到一个组播分组，源地址是 192.168.3.1。路由器检查其单播路由表，查找地址 192.168.3.1。单播路由表指出它是从接口 S1 处获知网络 192.168.3.0/24 的。在这个例子中，RPF 检查是成功的，因为该组播接口和单播路由表是一致的。一旦 RPF 检查成功，该路由器就向其输出接口转发这些组播分组——本例中这些输出接口是接口 E0 和 S2。输出接口是那些符合下列条件之一的接口：

- 在这一接口上发现了一个 PIM 邻居。
- 这一接口上的一个主机已经加入该组。
- 这一接口已经手动配置为加入该组。

在图 12-5 中，路由器从接口 S0 处接收到源地址为 192.168.3.1 的组播分组。如前例所示，该路由器检查其单播路由表并发现单播路由表是从接口 S1 处获知网络 192.168.3.0/24 的。在这个例子中，单播路由表与组播分组来自不同的接口。因此，RPF 检查失败，这些组播分组被丢弃。

转发组播分

GMP 离开消
简单地忽略
响应时，它暂

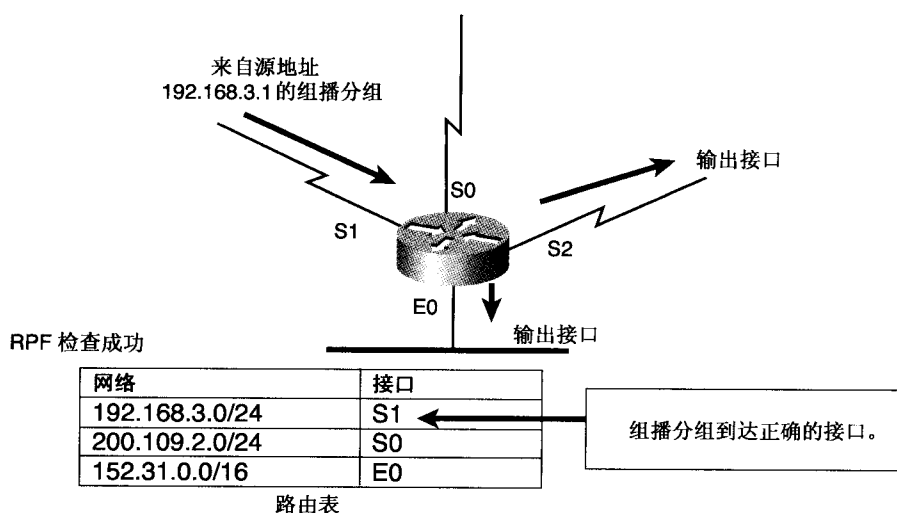


图 12-4 成功的 RPF 检查

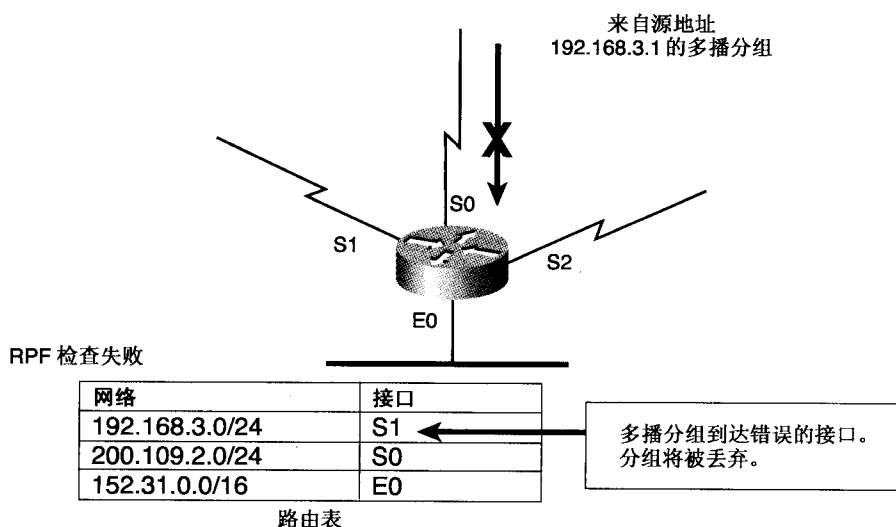


图 12-5 失败的 RPF 检查

12.2 PIM 密集模式

PIM 有两种操作模式：密集模式和稀疏模式。密集模式使用泛洪和剪枝机制转发组播分组。路由器假设每一个组播接口都对组播分组感兴趣，除非它被告知不是这样。路由器首先向所有接口转发组播分组。不想接收组播分组的网段从那些相邻路由器中收到剪枝消息，然后这一分支被剪除。

当路由器最初被配置为组播的时候，它向组播地址 224.0.0.2（这一子网地址上的所有路由器）定期发送 PIM 查询分组以发现 PIM 邻居。这些 PIM 查询分组被发送给那些配置用于 PIM 的接口。当在某个接口上接收到 PIM 查询时，就通过那个接口建立 PIM 邻居关系。

PIM 密集模式在其输出接口列表（也被称为 *olist*）上泛洪组播分组。如果下面这些条件为真，那么 PIM 密集模式把一个接口放入其 *olist* 中：

- 该接口有一个已建立的 PIM 邻居。
- 该接口通过 IGMP 把主机加入该组播组。
- 已经通过 `ip igmp join-group` 命令把该接口手动配置为加入这一组。

当一个运行密集模式的路由器首次收到组播分组的时候，它向在该 *olist* 中列出的所有接口泛洪组播分组。如果一个路由器从其邻居处收到一个剪枝分组，那么它就停止在某个接口上转发组播分组。

在图 12-6 中，路由器 R1 在接口 S0 上接收到达的组播分组。因为 R1 正在运行密集模式，所以它向其所有 *olist* 接口（E0 和 S1）泛洪组播分组。因为路由器 R2 没有任何对组播流量感兴趣的主机，所以它向 R1 发送一条 PIM 剪枝消息。当 R1 收到该 PIM 剪枝消息的时候，它在停止向接口 E0 转发该组的组播分组之前等待三秒钟。这三秒钟的延迟使该网段上的其他路由器可以使用一个 PIM 加入来覆盖这一剪枝消息。

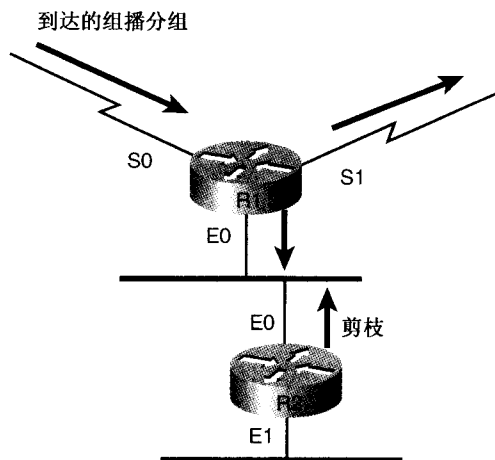


图 12-6 PIM 密集模式剪枝

在这个接口已经被剪枝的情况下，路由器停止向该接口转发该组的组播分组。在图 12-6 中，假设路由器 R2 现在有一个主机希望在接口 E1 上接收组播分组。路由器 R2 向路由器 R1 发送一个 PIM 加入分组。当路由器 R1 收到这一 PIM 加入消息的时候，它把其接口 E0 置为转发状态，组播流量流向路由器 R2。

如果在同一个 LAN 接口上有两个路由器，并且这两个路由器都有到该组播流的源地址的连接，那么这两个路由器都有可能转发组播分组，从而导致在这一 LAN 接口上出现重复分组。PIM 密集模式有一个断言机制（*assert mechanism*）可以避免出现这种情况。图 12-7 描述了 PIM 断言机制。

在缺少 PIM 断言过程的情况下，两个路由器都在接口 E0 上转发组播分组。这导致在该 LAN 接口上出现重复分组。通过使用断言机制，这两个路由器都向该组播流的源地址发送带有单播路由距离和度量的 PIM 断言分组。拥有对这一源地址而言最好的单播路由的路由器将获胜，并开始转发这些组播分组。在断言比较过程中失败的那个路由器剪枝它的输出接口。

如果按照单播路由的度量比较出现了平局，那么具有最高 IP 地址的那个路由器赢得断言比较。通过这种方法，只有一个路由器可以活跃地转发组播流量。

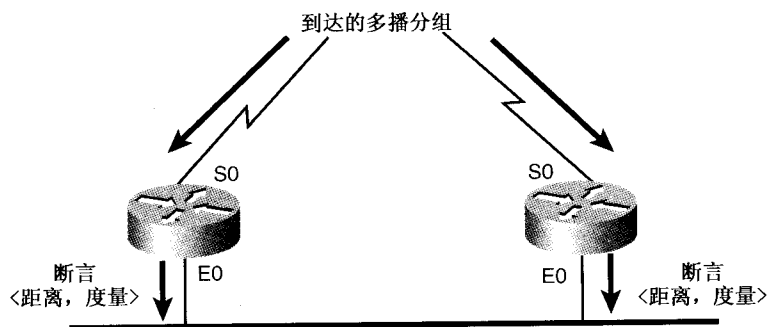


图 12-7 PIM 断言机制

12.3 PIM 稀疏模式

PIM 稀疏模式使用与密集模式相反的方式工作。PIM 密集模式假设所有组播接口都对组播分组感兴趣，除非被告知不是这样。而在 PIM 稀疏模式中，路由器假设组播接口没有对一个接收组播分组感兴趣，除非在该接口上收到一个 PIM 加入消息。PIM 稀疏模式比 PIM 密集模式有更好的可扩展性，但是其概念更复杂。PIM 稀疏模式使用集合点（rendezvous point, RP）的思想。RP 是其中发送者和接收者在建立最短路径树之前第一次相遇的地方。最短路径树是在组播发送者和接收者之间的最短路径。对一个特定的组播组而言，只有一个 RP 被选择。RP 的选择通过静态配置完成或者通过自动 RP 机制动态获知。

PIM 稀疏模式发现邻居的方式与 PIM 密集模式所采用的方式是一样的。PIM 路由器发送 PIM 查询分组来发现这一链路上的 PIM 邻居。在 PIM 稀疏模式中，一个 LAN 网段上的最高 IP 地址被选为指定路由器（DR）。这个指定路由器用于向该网段的 RP 发送 PIM 加入消息。

在稀疏模式中，组播流有两个部分：

1. 接收者向 RP 发送 PIM 加入。
2. 组播源向 RP 发送 PIM 注册。

在 PIM 稀疏模式的加入机制中，最靠近接收站的路由器向 RP 发送 PIM 加入消息。如果在这一 LAN 网段上存在不止一个路由器，那么这个 PIM 的 DR 向该 RP 发送加入。然后 PIM 加入被逐跳地发往该 RP。图 12-8 描述了 PIM 稀疏模式的加入机制。

在图 12-8 中，PC 向其 Ethernet 接口发送 IGMP 加入消息。因为路由器 R2 在接口 E1 上有一个较高的 IP 地址，所以它是 DR。路由器 R2 向 RP 发送 PIM 加入消息，因此 R2 向路由器 R1 发送 PIM 加入消息。而路由器 R1 又向 RP 发送 PIM 加入。因此，这些 PIM 加入被逐跳地从最靠近该接收者的叶路由器沿路发送到 RP。该叶路由器被定义为最外层的边缘路由器，它只连接了接收者。

稀疏模式操作的第二个部分是注册过程，该过程可以分解为下面的事件序列：

1. 稀疏模式注册消息是由离组播流的发送者最近的路由器发送的。如果在这个 LAN 网段上存在不止一个路由器，那么 PIM 的 DR 负责向 RP 发送 PIM 注册消息。

图 12-10 是 PIM 稀疏模式注册过程的继续。当路由器 R1 收到来自 RP 的加入消息时，R1 开始向 RP 发送组播流量。来自 R1 的注册消息仍然被转发。然后 RP 收到这个组播分组的两个副本，一个来自该注册消息，另一个来自本地的组播流。当 RP 看到组播分组的两个副本时，它向第一跳路由器（在这个例子中是路由器 R1）发送一条注册停止消息。当 R1 收到这条注册停止消息时，它停止把这些组播分组封装成注册消息，流量现在按原样从 R1 流向 R2，然后流向 RP。最后 RP 把组播流转发到其下游的邻居。

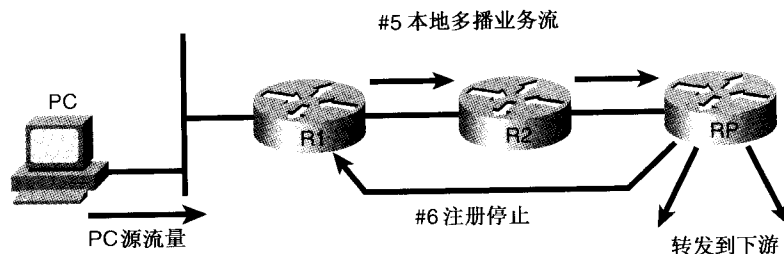


图 12-10 PIM 稀疏模式注册过程，第二部分

稀疏模式剪枝机制与密集模式中使用的机制是一样的。如果路由器对组播流量不感兴趣，那么它向在源地址上游的邻居发送一条 PIM 剪枝消息。关于 PIM 操作的更详细的描述，请参考 Beau Williamson *Developing IP Multicast Networks* 第 1 卷。

12.4 IGMP 分组格式和 PIM 分组格式

对于理解 PIM 的操作而言，IGMP 分组格式和 PIM 分组格式是很有用的。倘若需要查看嗅探轨迹的话，理解这些分组格式还有助于 PIM 问题的诊断。本节讨论 IGMP 的重要分组格式和 PIM 的重要分组格式。

12.4.1 IGMP 分组格式

IGMP 消息始终以 TTL 被置为 1 后被发送，而且使用协议号为 2 的 IP 封装。图 12-11 显示了 IGMP 版本 2 的分组格式。IGMP 版本 1 的分组格式与版本 2 的格式相比有一点不同。IGMP 版本 1 分组把版本 2 中的 Type 字段划分为两个部分，从而同时包含版本号 and 类型。

Type 字段指明 IGMP 分组的不同类型：

- Type 11 是指 IGMP 成员关系查询。
- Type 12 是指 IGMP 版本 1 的成员关系报告。
- Type 16 是指 IGMP 版本 2 的成员关系报告。
- Type 17 指明 IGMP 版本 2 的离开组。

上面列出的类型是最重要的类型，可以在 RFC 2236 中找到其他 Type 字段信息。

最大响应时间字段仅用于成员关系查询消息中，它指定一个主机要等待响应一条查询消息的最长时间。

0	8	16	24	31
类型	最大响应 时间	校验和		
组地址				

图 12-11 IGMP 分组格式

校验和字段是 IGMP 消息的校验和，以验证分组的完整性。

组地址字段包含该接收者有兴趣接收的组播组。当发送一般的 IGMP 查询时，这个组地址字段设置为 0。

12.4.2 PIM 分组/消息格式

PIM 版本 1 的分组被封装成 IGMP 类型 14 分组。PIM 版本 2 使用它自己的协议编号 103，不被封装成 IGMP。PIM 版本 2 的分组发送到组播地址 224.0.0.13。图 12-12 显示了 PIM 的 hello 消息的格式。

0	8	16	24	31
版本	类型=0	保留	校验和	
选项类型			选项长度	
选项值				
• • •				
选项类型			选项长度	
选项值				

图 12-12 PIM 的 hello 分组格式

PIM 的 hello 消息属于 PIM 类型 0 分组。选项类型始终设置为 1。PIM 的 hello 分组用于建立 PIM 的邻居关系。

图 12-13 显示了 PIM 稀疏模式的 PIM 注册消息。

PIM 注册消息属于 PIM 的类型 1 分组。DR 把这个组播分组封装成注册分组并发送到 RP。根据这个分组的格式，如果路由器是这个发送者的一个 PIM 组播边界路由器，那么 B 位被设

置为 1。如果发送路由器是一个直接连接发送者的 DR，那么 B 位被设置为 0。N 位是 null 注册，它用于预防在接收 REGISTER STOP 消息间隔内发往 RP 的不必要突发流量。图 12-14 显示了 REGISTER STOP 消息的格式。

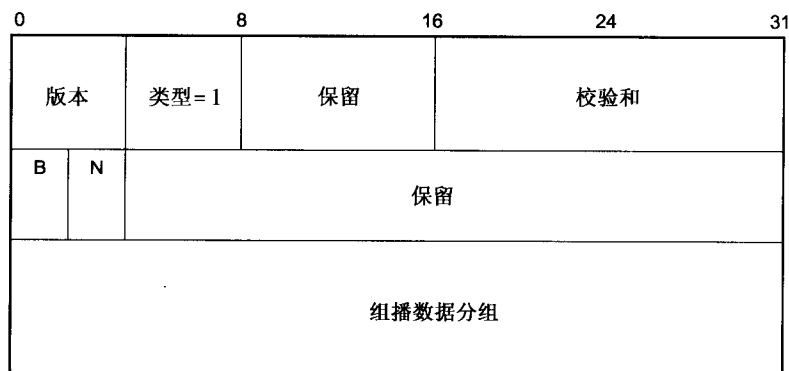


图 12-13 PIM 注册分组

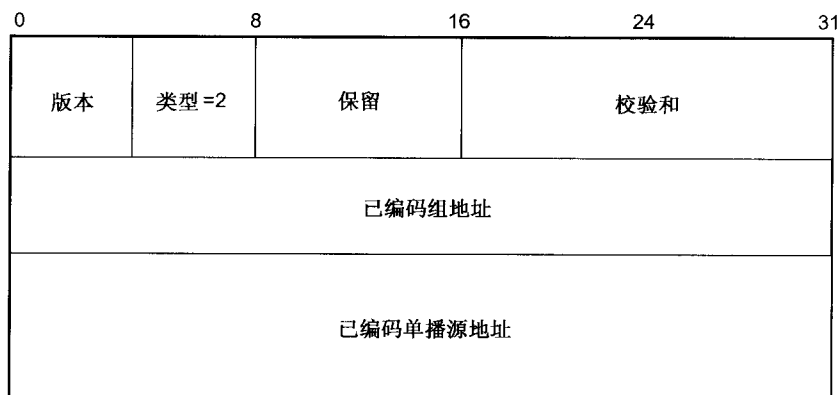


图 12-14 PIM 的 REGISTER STOP 分组格式

编码的 (encoded) 组地址包含那些被封装为 REGISTER 消息的组播地址的组。已编码的单播源地址包含该单播流的发送者的单播地址。根据前面的讨论，REGISTER STOP 消息被发送到发送该 REGISTER 消息的路由器，以指明已经收到一个本地的组播流，同时 DR 可以停止发送 REGISTER 分组。

图 12-15 描述了 PIM JOIN/PRUNE 消息，它同时用于密集模式和稀疏模式中。

在 PIM JOIN/PRUNE 分组中，已编码的单播上游邻居地址是执行这一加入或者剪枝的 RPF 邻居的 IP 地址。组数是包含在这条消息中的组播组集合的数目。每一个集合都由一个已编码的组播组地址后面跟着一个用于加入或剪枝的一系列已编码的源地址组成。

图 12-16 显示了 PIM 断言消息的格式。断言机制使路由器可以选择一个活跃的路由器转发组播分组，以避免 LAN 上分组的重复。

编码的组地址是该组播组的地址。编码的单播源地址是该组播流发送者的 IP 地址。度量是与使用的路由协议相关的组播流发送者的路由度量。如果使用的路由协议是 RIP，那

么该度量可以是跳的数目。R 位是 RPT 位，如果触发这个断言分组的组播分组沿着这一共享树向下发送，那么 R 位被设置为 1。如果该组播分组沿着最短路径树向下发送，那么 R 位被设置为 0。

0	8	16	24	31
版本	类型=3	保留	校验和	
已编码的单播上游邻居地址				
保留		组数	保持时间	
已编码的组播组地址 1				
已加入源地址数			已剪枝的源地址数	
已编码的加入源地址 1				
⋮				
已编码的加入源地址 n				
已编码的剪枝源地址 1				
⋮				
已编码的剪枝源地址 n				
⋮				
已编码的组播组地址 n				
已加入源地址数			已剪枝的源地址数	
⋮				
已编码的加入源地址 n				
已编码的剪枝源地址 1				
⋮				
已编码的剪枝源地址 n				

图 12-15 PIM 的 JOIN/PRUNE 分组格式

PI
机之间
本质，
好的可
理论来

- 1.
- 2.
- 3.
- 4.
- 5.
- 6.
- 7.
- 8.

组沿着这一共
发送, 那么 R

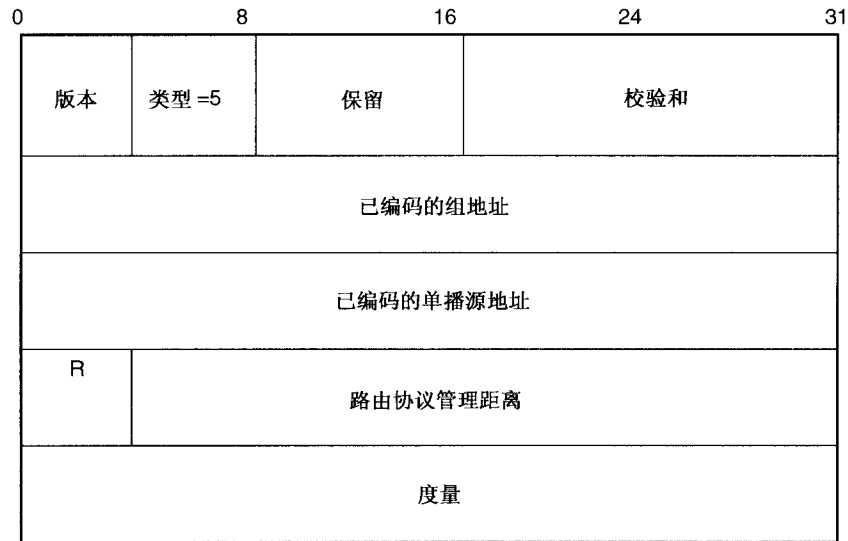


图 12-16 PIM 的断言分组格式

12.5 小 结

PIM 协议提供独立于底层单播路由协议的组播路由。IGMP 机制实现了路由器和组播主机之间的通信。PIM 密集模式提供一种容易的组播路由实现, 但是由于泛洪——剪枝机制的本质, 它不是一个可扩展的组播解决方案。但是, PIM 稀疏模式提供了大型网络中需要的很好的可扩展性, 不过与 PIM 密集模式相比, 它的操作更复杂一些。下一章将基于本章讲述的理论来诊断 PIM 的常见问题。

12.6 复 习 题

1. 单播、广播和组播之间的区别是什么?
2. PIM 的不同模式是什么?
3. PIM 密集模式的操作机制是什么?
4. PIM 稀疏模式操作机制是什么?
5. IGMP 版本 1 和版本 2 的组离开机制有什么不同?
6. IGMP 使用什么组播地址进行 IGMP 查询?
7. RPF 检查如何工作?
8. 什么是集合点 (RP)?

第 13 章

PIM 协议疑难解析

本章讨论 PIM 组播协议的疑难解析方法。本章分为三部分：

- IGMP 加入问题。
- PIM 密集模式问题。
- PIM 稀疏模式问题。

每一部分都有进行疑难解析的流程图和案例研究，以帮助读者分析和诊断常见的 PIM 组播问题。如果需要复习一些概念，第 12 章，“理解 PIM 协议”，讨论了 PIM 协议的一般操作。欲了解更多关于组播和 PIM 协议的细节，请参考 Beau Williamson 所著的 *Developing IP Multicast Networkst*，第一卷。

13.1 IGMP 加入问题疑难解析

正如在第 12 章“理解 PIM 协议”中所讨论的，IGMP 加入是组播接收者（主机）和路由器之间的一种通信机制。组播接收者使用 IGMP 加入消息通知组播路由器：本地网段上的主机对某些组播组感兴趣。这使得路由器可以向这个网段转发组播分组。本节讨论 IGMP 加入的问题。请参考图 13-1，其中有一个关于 IGMP 加入问题的疑难解析流程图。

图 13-2 给出了诊断 IGMP 加入问题的网络配置。

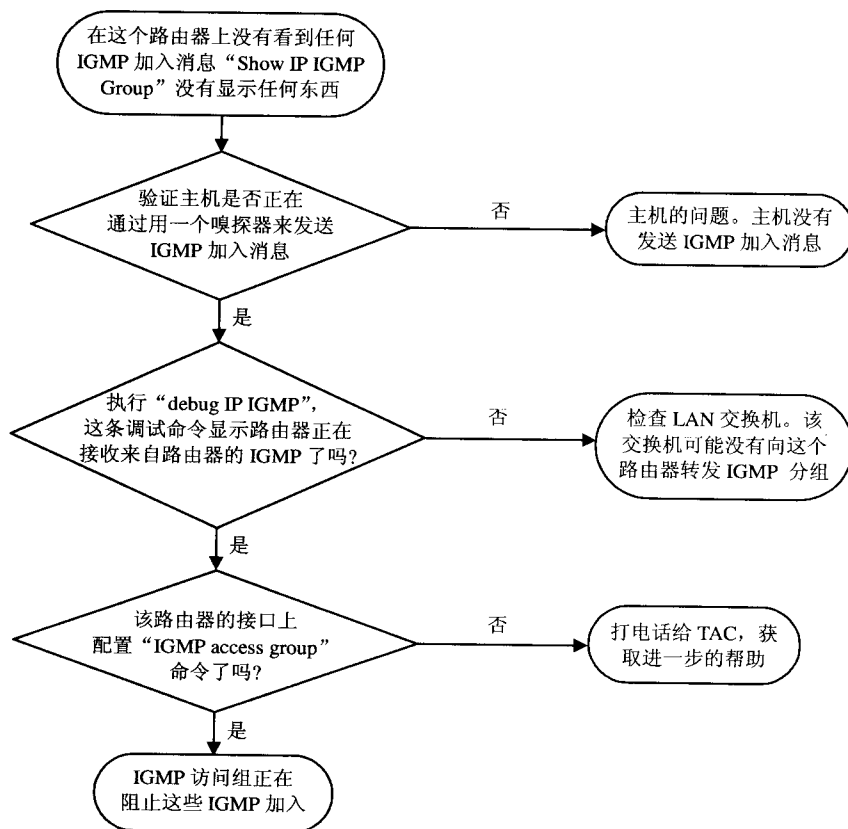


图 13-1 IGMP 加入问题疑难解析流程图

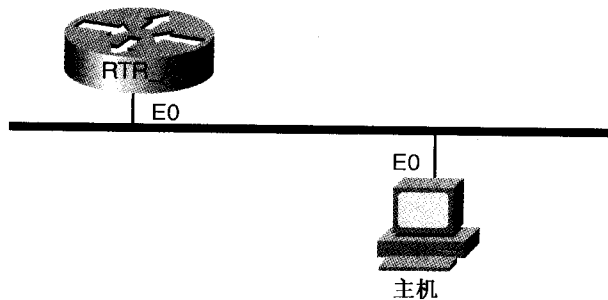


图 13-2 用于案例研究的针对 IGMP 加入问题的网络图

在图 13-2 中, 组播主机对加入组播组 225.1.1.1 感兴趣。该主机向组播地址 224.0.0.2 (这是“所有路由器”组播地址) 发送 IGMP 加入消息。问题是路由器 A 没有看到这些 IGMP 加入消息。为了解哪些组播组通过 IGMP 被加入到路由器上, 需要使用命令 `show ip igmp group`。例 13-1 显示了路由器 A 的 `show ip igmp group` 命令的输出。

例 13-1 路由器 A 上 `show ip igmp group` 命令的输出

```
RTR_A#show ip igmp group
```

```
IGMP Connected Group MembershipGroup Address Interface Uptime Expires
```

(待续)

```
Last Reporter
RTR_A#
```

例 13-1 中的输出表明路由器 A 没有看到来自组播主机的 IGMP 加入分组。如果路由器 A 没有看到来自该组播主机的 IGMP 加入分组，那么疑难解析的下一个逻辑步骤是询问该组播主机是否正在发送 IGMP 加入分组，以及该路由器对这些 IGMP 加入分组实施了什么操作。为查看一个路由器上的 IGMP 事务，需要使用 **debug ip igmp** 命令，如例 13-2 中路由器 A 的输出所示。

例 13-2 路由器 A 上 debug ip igmp 命令的输出

```
RTR_A#debug ip igmp

IGMP: Received v2 Report from 150.150.2.2 (Ethernet0) for 225.1.1.1
IGMP: Group 225.1.1.1 access denied on Ethernet0
```

debug 命令的输出显示这个路由器正在接收来自该主机的 IGMP 加入消息，但是它们被路由器抛弃了。例 13-3 显示了路由器 A 的配置。

例 13-3 路由器 A 的配置

```
RTR_A#show run
ip multicast-routing
interface ethernet 0
    ip address 150.150.2.1 255.255.255.0
    ip pim dense-mode
    ip igmp access-group 1
access-list 1 deny 224.0.0.0 3.255.255.255
access-list 1 permit any
```

例 13-3 中的配置揭示了 225.1.1.1 的 IGMP 加入被抛弃的原因：在接口 Ethernet 0 上有一条配置用于 IGMP 的 **access-group** 语句。该接入组被加入到 **access-list 1**，它拒绝范围从 224.0.0.0 到 227.255.255.255 的任何组加入。网络管理员希望仅拒绝从 224.0.0.0 到 224.255.255.255 的组，这是路由器上的一个配置错误。

解决 IGMP 加入问题的方法

解决 IGMP 加入问题的方法是修改 **access-list 1**，使之与例 13-4 中所示的相匹配。

例 13-4 用于允许加入组播组 225.1.1.1 的路由器 A 配置的替代方案

```
RTR_A#access-list 1 deny 224.0.0.0 0.255.255.255
access-list 1 permit any
```

使用这个新的配置，该路由器不会再拒绝组 225.1.1.1 的 IGMP 加入消息。例 13-5 显示了路由器 A 上的 **debug** 命令，以及在完成这些改动之后 **show ip igmp group** 命令的输出。

0.0.0.2 (这
IGMP 加
mp group。

持续)

例 13-5 路由器 A 上 debug ip igmp 命令和 show ip igmp group 命令的输出

```
RTR_A#debug ip igmp
```

```
IGMP: Received v2 Report from 150.150.2.2 (Ethernet0) for 225.1.1.1
```

```
RTR_A#show ip igmp group
```

Group Address	Interface	Uptime	Expires	Last Reporter
225.1.1.1	Ethernet0	00:00:40	00:02:18	150.150.2.2

现在路由器 A 显示, 组 225.1.1.1 与 IGMP 报告者一样被加入到 IP 地址为 150.150.2.2 的 Ethernet 0 接口。

13.2 PIM 密集模式疑难解析

组播密集模式的操作非常简单——它使用泛洪——剪枝机制形成一棵组播转发树。因为其操作的简单性, PIM 密集模式的疑难解析也非常简单。大多数 PIM 密集模式问题都与 RPF 检查失败问题以及 TTL 值问题有关。图 13-3 显示了诊断组播密集模式常见问题的流程图。

接下来的案例研究说明了一个典型的 PIM RPF 检查问题。图 13-4 显示这个案例的网络配置。

例 13-6 显示在图 13-4 中描述的那些路由器上的相关配置。

例 13-6 组播路由器配置

```
Multicast Source#interface ethernet 0
ip address 172.16.1.1 255.255.255.0
RTR_A#ip multicast-routing
interface ethernet 0
ip address 172.16.1.2 255.255.255.0
ip pim dense mode
interface serial 0
ip address 172.16.3.1 255.255.255.0
interface serial 1
ip address 172.16.2.1 255.255.255.0
ip pim dense mode
router eigrp 1
network 172.16.0.0
```

```
RTR_B#ip multicast-routing
ip multicast-routing
interface ethernet 0
ip address 172.16.5.1 255.255.255.0
ip pim dense mode
```

(待续)

```

interface serial 0
ip address 172.16.3.2 255.255.255.0
interface serial 1
ip address 172.16.4.1 255.255.255.0
ip pim dense mode
router eigrp 1
network 172.16.0.0

```

```

RTR_C#ip multicast-routing
interface serial 0
ip address 172.16.2.2 255.255.255.0
ip pim dense mode
interface serial 1
ip address 172.16.4.2 255.255.255.0
ip pim dense mode
router eigrp 1
network 172.16.0.0

```

```

Multicast Receiver#ip address 172.16.5.2 255.255.255.0

```

0.150.2.2 的

转发树。因
模式问题都
常见问题的

个案例的网络

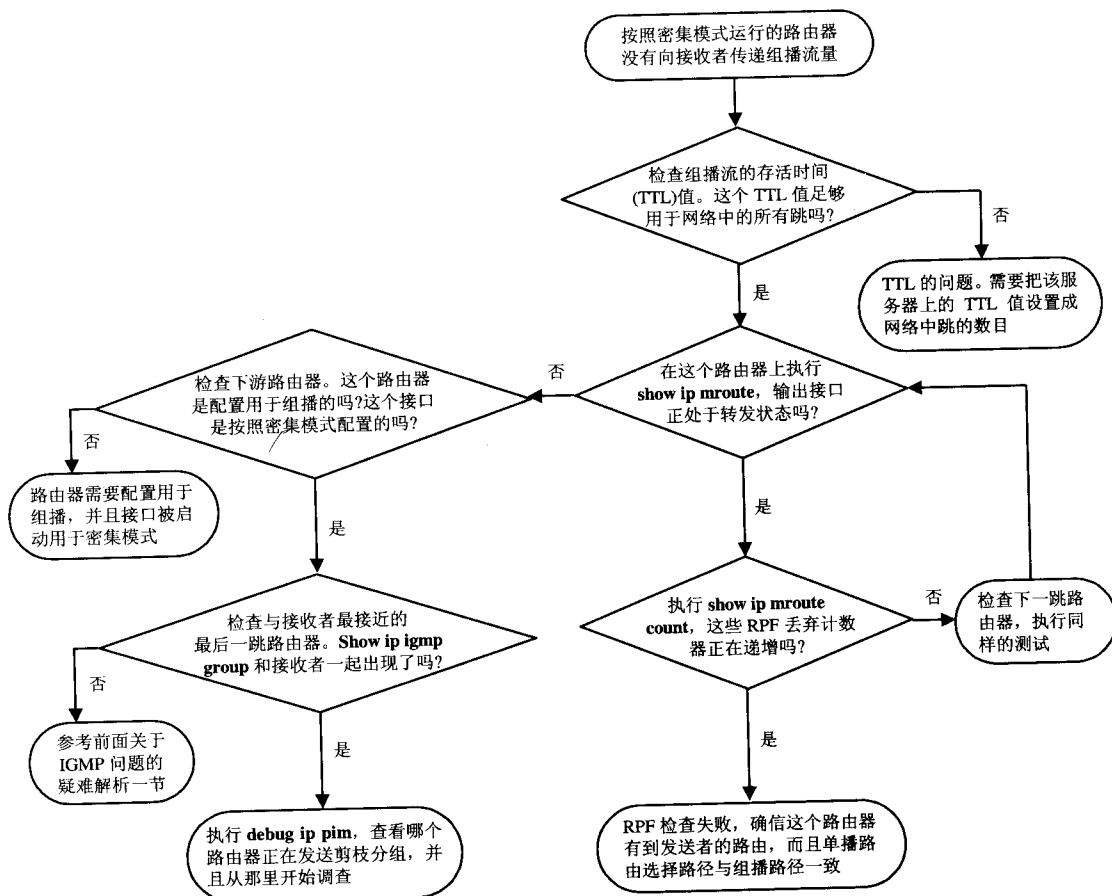


图 13-3 诊断组播密集模式常见问题的流程图

(待续)

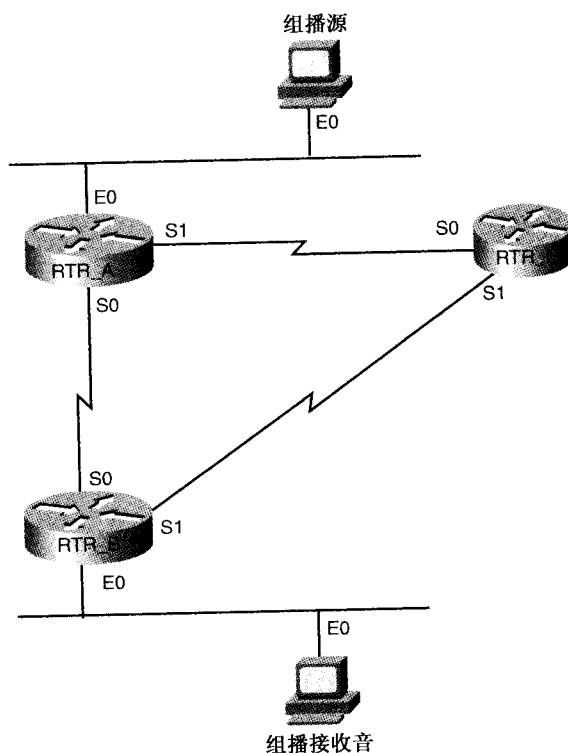


图 13-4 用于案例研究的针对 PIM RPF 检查问题的网络图

这个案例研究中出现的问题是组播服务器正在向组 225.1.1.1 发送一个组播流，但是组播接收者没有收到这个分组。当类似这样的问题第一次出现时，首先要做的事情是确保来自该服务器的组播流的 TTL 值大于 1，并且该 TTL 值大于该网络中跳的总数。这个 TTL 值是当发送该流的时候在服务器中的组播应用程序上设置的。为了检查分组的 TTL 值，需要使用一个嗅探器 (sniffer) 来捕获组播分组，并查看分组的内部以确定 TTL 值。在这个案例研究中，组播服务器的 TTL 值设置为 15 跳。诊断的下一步是查看组播路由表。例 13-7 显示了路由器 A 上 `show ip mroute` 命令的相关输出。

例 13-7 路由器 A 的 `show ip mroute` 命令的输出

```
RTR_A#show ip mroute 225.1.1.1

IP Multicast Routing Table
(*,225.1.1.1),
Incoming interface: Null, RPF nbr 0.0.0.0
Outgoing interface list:
Ethernet 0, Forward/Dense,
Serial 1, Forward/Dense

(172.16.1.1/32, 225.1.1.1),
Incoming interface: Ethernet 0, RPF nbr 0.0.0.0
Outgoing interface list:
Serial 1, Forward/Dense
```

根据例 13-7 中路由器 A 的组播路由表, 可以看到路由器 A 接收了来自 Ethernet 0 的组播流, 并把它转发到 Serial 1 接口。**show ip mroute count** 命令验证了路由器 A 的确正在转发该组播流, 如例 13-8 中所示。

例 13-8 show ip mroute 命令的输出

```
RTR_A#show ip mroute 225.1.1.1 count

Forwarding Counts: Pkt Count/Pkts per seconds/Avg Pkt Size/Kilobits per second
Other Counts: Total/RPF failed/Other drops(OIF-null, rate-limit, etc)

Group: 225.1.1.1, Source count: 1, Group pkt count: 29543
Source: 172.16.1.1/32, forwarding: 29543/195/1043/203, other: 0/0/0
```

如例 13-8 中的输出所示, 路由器 A 转发分组的总数为 29,543 个, 速度是每秒 195 个。**show ip mroute count** 是一条重要的 **show** 命令, 因为它确认路由器正在传输组播分组, 还揭示是否发生了由于 RPF 失败而发生组播分组被丢弃的情况。根据例 13-7 和 13-8 中这条 **show** 命令的输出, 路由器 A 看上去是好的。下一步是行进到下一跳, 也就是路由器 B。例 13-9 显示在路由器 B 上 **show ip mroute 225.1.1.1** 命令的输出。

例 13-9 路由器 B 的 show ip mroute 225.1.1.1 命令的输出

```
RTR_B#show ip mroute 225.1.1.1
IP Multicast Routing Table
(*,225.1.1.1),
Incoming interface: Null, RPF nbr 0.0.0.0
Outgoing interface list:
Ethernet 0, Forward/Dense, Serial 0, Forward/Dense,
Serial 1, Forward/Dense

(172.16.1.1/32, 225.1.1.1),
Incoming interface: Serial 1, RPF nbr 172.16.3.1
Outgoing interface list:
Ethernet 0, Forward/Dense
```

在例 13-9 中, 路由器 B 的输出接口是 Ethernet 0, 它处于转发状态; 但是, 组播接收者没有收到任何组播分组。该接收者已经向路由器 B 发送了 IGMP 加入消息, 并且路由器 B 确认了该加入, 如例 13-10 所示。

例 13-10 路由器 B 确认 IGMP 加入

```
RTR_B#show ip igmp group

Group Address      Interface      Uptime        Expires        Last Reporter
225.1.1.1          Ethernet0      00:00:40      00:02:18      172.16.5.2
```

下一步是通过执行 **show ip mroute 225.1.1.1** 命令查看路由器 B 是否正在发送该组播分组, 如例 13-11 所示。

例 13-11 路由器 B 上 show ip mroute 225.1.1.1 命令的输出

```
RTR_B# show ip mroute 225.1.1.1 count
```

```
Forwarding Counts: Pkt Count/Pkts per seconds/Avg Pkt Size/Kilobits per second
Other Counts: Total/RPF failed/Other drops(OIF-null, rate-limit, etc)
Group: 225.1.1.1, Source count: 1, Group pkt count: 29543
Source: 172.16.1.1/32, forwarding: 29543/0/0/0, other: 29543/29543/0
```

要注意的是, 路由器 B 没有在 Ethernet 0 接口上转发任何分组。路由器 B 仍然把 Ethernet 0 置为转发状态是因为 Ethernet 0 上有一个活跃接收者。RPF 失败计数器一直在不断增长。看上去路由器 B 是由于 RPF 检查失败问题而没有向 Ethernet 0 转发分组的。RPF 检查失败表示单播路由路径与组播路径不一致。为了验证这一点, 路由器 B 上到源网络 172.16.1.0/24 的路由表如例 13-12 所示。

例 13-12 路由器 B 上 show ip route 命令的输出

```
RTR_B# show ip route 172.16.1.0 255.255.255.0
```

```
Routing entry for 172.16.1.0/24
Known via "EIGRP 1", distance 90, metric 2195456, type internal
  Redistributing via eigrp 1
  Last update from 172.16.3.1 on Serial 0, 00:10:30 ago
  Routing Descriptor Blocks:
  * 172.16.3.1, from 172.16.3.1, 00:10:30 ago, via Serial 0
    Route metric is 2195456, traffic share count is 1
  Total delay is 21000 microseconds, minimum bandwidth is 1544 Kbit
  Reliability 255/255, minimum MTU 1500 bytes
  Loading 1/255, Hops 1
```

路由器 B 上的组播路由表显示它正在使用 Serial 0 作为到达网络 172.16.1.0/24 的下一跳; 但是, 根据路由器 B 上 **show ip mroute 225.1.1.1** 的输出, 组播路径指向 Serial 1。注意 RPF 邻居被列为 172.16.3.1 (它的单播路径)。这样的分歧导致 RPF 检查失败, 路由器只会丢弃这个分组。

解决 PIM 密集模式问题的方法

查看配置, 路由器 A 和路由器 B 的 Serial 0 接口没有启用密集模式; 但是这是客户的需求。该客户不希望组播流充满路由器 A 和路由器 B 之间的 WAN 链路, 他还希望组播流在路由器 A、路由器 C 和路由器 B 的方向上流动。为了修正这个 RPF 问题并满足客户的需求, 需要在路由器 B 上配置一个静态组播路由, 以使 RPF 检查使用该静态组播路由而不是单播路由表。例 13-13 显示了用于路由器 B 的静态组播路由配置。

例 13-13 路由器 B 的静态组播路由配置

```
RTR_B# ip mroute 172.16.1.0 255.255.255.0 172.16.4.2
```

正确地使用这个静态的组播路由, 路由器 B 中的 RPF 检查就会成功, 组播分组转发到组播接收者而不是被丢弃。例 13-14 显示了在配置改动之后在路由器 B 上命令 `show ip mroute 225.1.1.1` 和 `show ip mroute 225.1.1.1 count` 的输出。

例 13-14 路由器 B 上 show ip mroute 命令的输出

```
RTR_B# show ip mroute 225.1.1.1
IP Multicast Routing Table
(*,225.1.1.1),
  Incoming interface: Null, RPF nbr 0.0.0.0
  Outgoing interface list:
Ethernet 0, Forward/Dense,
Serial 0, Forward/Dense,
Serial 1, Forward/Dense

(172.16.1.1/32, 225.1.1.1),
  Incoming interface: Serial 1, RPF nbr 172.16.4.2
  Outgoing interface list:
Ethernet 0, Forward/Dense

RTR_B# show ip mroute 225.1.1.1 count
Forwarding Counts: Pkt Count/Pkts per seconds/Avg Pkt Size/Kilobits per second
Other Counts: Total/RPF failed/Other drops(OIF-null, rate-limit, etc)

Group: 225.1.1.1, Source count: 1, Group pkt count: 26721
Source: 172.16.1.1/32, forwarding: 26721/254/1253/318, other: 0/0/0
```

例 13-14 中的输出揭示了路由器 B 正在 Ethernet 0 上转发组播分组, 而组播接收者现在正在接收该组播流。

13.3 PIM 稀疏模式疑难解析

PIM 稀疏模式的操作与密集模式有所不同, 它比密集模式更复杂。疑难解析步骤类似于密集模式中的情况。图 13-5 显示了 PIM 稀疏模式问题的疑难解析流程图。

接下来的案例研究检查如何诊断一个 PIM 稀疏模式的问题, 其中叶路由器向 RP 发送加入消息。图 13-6 给出了本案例研究的网络图。

根据图 13-6 可以看出, 路由器 A 和路由器 B 组成从组播服务器到接收者的连接。路由器 C 连接到网络 X, 而该网络没有直接连接到组播服务器的网络上。组播服务器在组播组 225.0.0.1 上发送组播流。例 13-15 显示了图 13-6 中那些路由器的配置。这些路由器上的稀疏/密集模式配置允许这些路由器在 RP 存在的情况下以稀疏模式运行。如果 RP 不可用, 那么这些路由器将以密集模式运行。

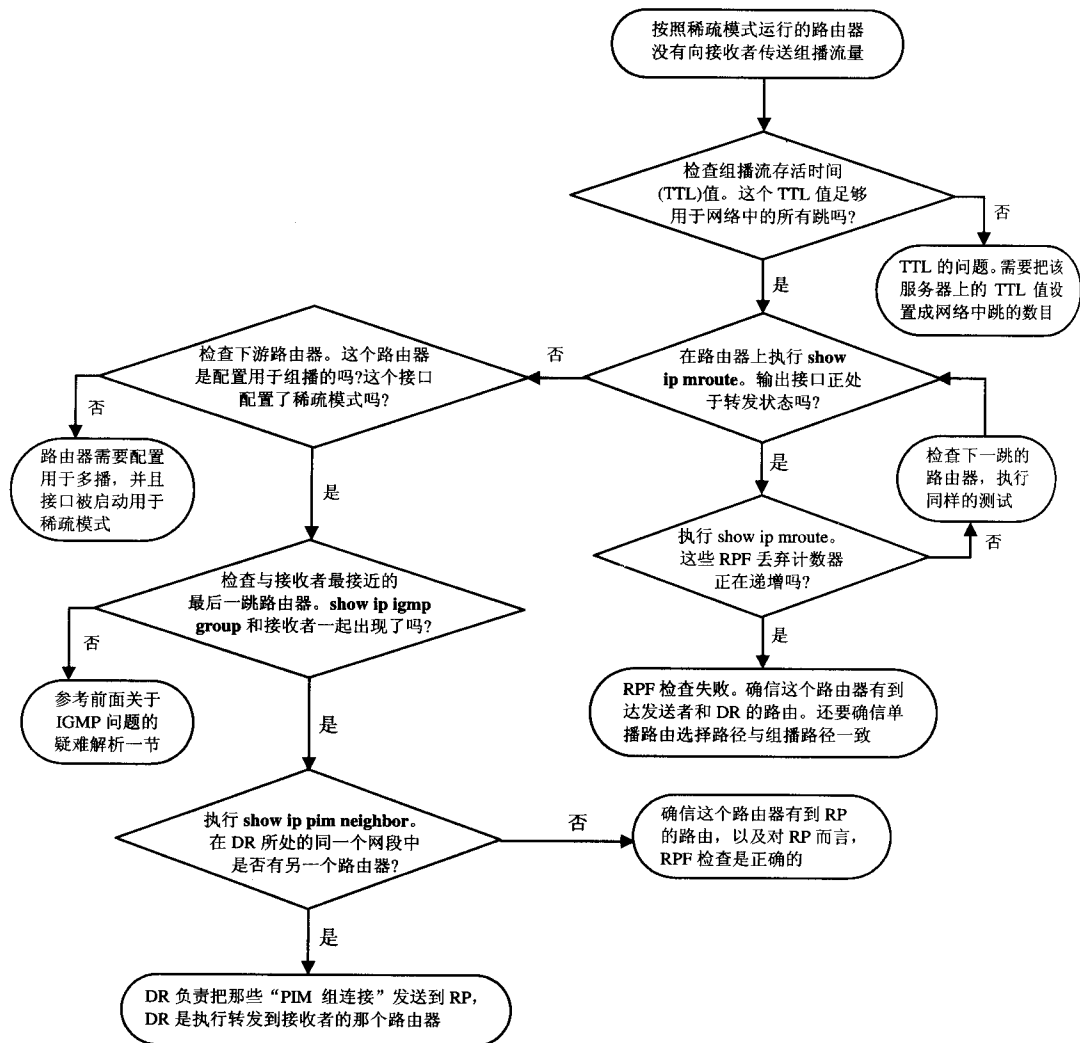


图 13-5 PIM 稀疏模式问题疑难解析流程图

例 13-15 路由器组播配置

```

Multicast Source# interface ethernet 0
ip address 172.16.1.1 255.255.255.0

RTR_A# ip multicast-routing
interface ethernet 0
ip address 172.16.1.2 255.255.255.0
ip pim sparse-dense mode
interface serial 0
ip address 172.16.2.1 255.255.255.0
ip pim sparse-dense mode
router eigrp 1
network 172.16.0.0
ip pim send-rp-announce ethernet 0 scope 16
  
```

(待续)

```

ip pim send-rp-discovery scope 16
RTR_B# ip multicast-routing
interface ethernet 0
ip address 172.16.3.1 255.255.255.0
ip pim sparse-dense mode
interface serial 0
ip address 172.16.2.2 255.255.255.0
ip pim sparse-dense mode
router eigrp 1
network 172.16.0.0

```

```

RTR_C# ip multicast-routing
interface ethernet 0
ip address 172.16.3.3 255.255.255.0
ip pim sparse-dense mode
interface serial 1
ip address 172.16.4.1 255.255.255.0
ip pim sparse-dense mode
router eigrp 1
network 172.16.0.0

```

```

Multicast Receiver# interface ethernet 0
ip address 172.16.3.2 255.255.255.0

```

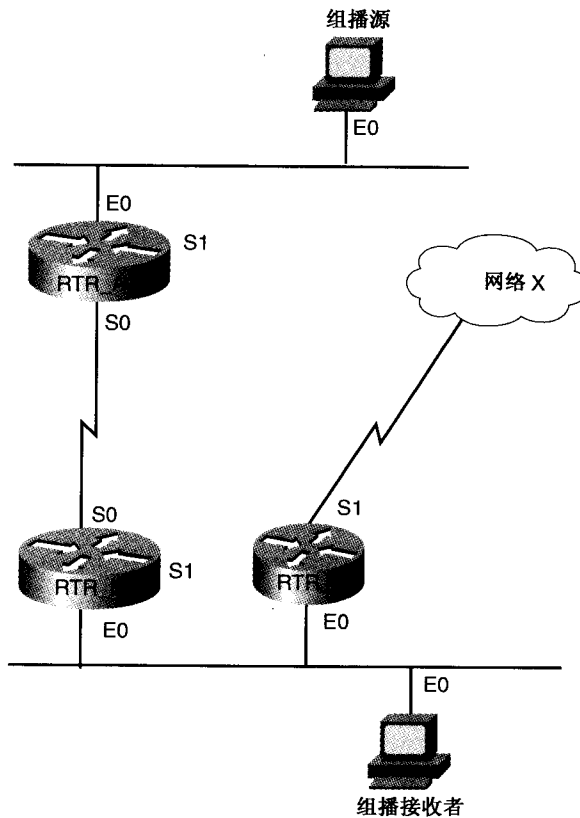


图 13-6 用于实例研究的针对 PIM 稀疏模式问题的网络图

(待续)

第 12 章

理解 PIM 协议

本章涵盖以下关于 PIM 协议的关键主题：

- IGMP 版本 1、IGMP 版本 2 和反向路径转发（RPF）的基本原理。

- PIM 密集模式。
- PIM 稀疏模式。
- IGMP 和 PIM 分组格式。

主机对主机传输已经成为学术界经常讨论的话题。随着技术的发展，出现了为这种传输提供支持的新的方法学。从一个特定的主机到另一个特定主机的传输称作单播（unicast）。

点对点的传输是容易的。目前，一个重要的推动力是想出如何从一个主机传输到多个主机，而且不会中断业务流。直到现在，如果想让一个主机与多个主机同时会话，那么必须求助于广播。最近几年，组播作为一种更有效的替代方案出现了。

组播、广播和单播之间的不同点在于，单播分组仅仅定位到一个主机。广播分组定位到同一网段上的所有主机，不管该主机是否对这一分组感兴趣。组播是只向对这些分组感兴趣的主机传递该分组的一种有效方法。组播分组的地址范围位于 D 类地址 224.0.0.0 到 239.255.255.255 之间。组播发送者只发送这一分组的一个副本，只有那些对这一组播分组感兴趣的主机才处理这一分组。

因为组播分组可能在到达预定目的地之前经过几个路由器，所以这些路由器需要启用一个路由协议，以确保组播分组被有效地传递并且不发生循环。

已经开发出几种组播路由协议用于这种情况。其中第一个协议是距离矢量组播路由协议（Distance Vector Multicast Routing Protocol, DVMRP）；但是，DVMRP 的收敛速度慢，