

HCIP-PIM-1

作者: [shuaiqijun](#)

原文链接: <https://ld246.com/article/1614852683614>

来源网站: [链滴](#)

许可协议: [署名-相同方式共享 4.0 国际 \(CC BY-SA 4.0\)](#)

PIM协议简介

定义

PIM (Protocol Independent Multicast) 称为协议无关组播。这里的协议无关指的是与单播路由协议无关，即PIM不需要维护专门的单播路由信息。作为组播路由解决方案，它直接利用单播路由表的路由信息，对组播报文执行RPF (Reverse Path Forwarding, 逆向路径转发) 检查，检查通过后创建组播路由表项，从而转发组播报文。

目前设备实际支持的PIM协议包括：PIM-DM (PIM-Dense Mode) 、PIM-SM (PIM-Sparse Mod) 。

目的

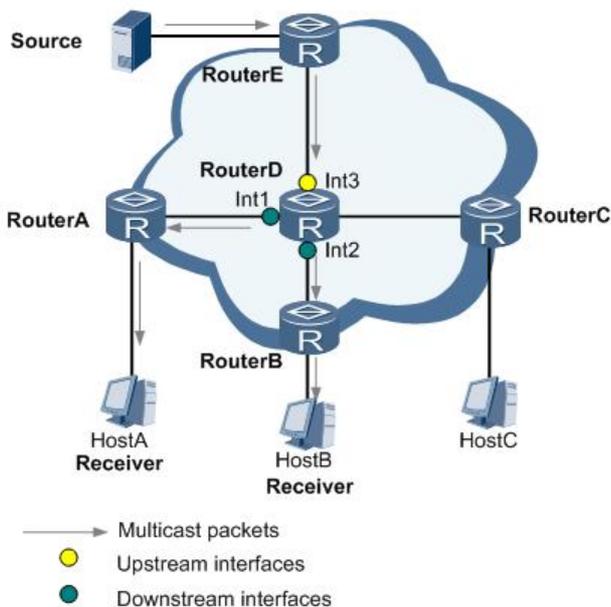
1992年，为了承载网络视频会议、音频会议，以MOSP和DVMRP为组播路由协议的虚拟IP组播骨干网——Mbone建立成功。这为组播技术应用和推广起到了积极作用，在随后的十几年，组播路由协议得到了很大发展。

但是随着多个组播路由协议的开发与应用，人们渐渐感觉到，如果像单播路由一样通过多种路由协议来动态生成组播路由，会带来不同路由协议间在互相引入路由时操作繁琐的问题。而且网络设备对单播和组播路由信息都需要维护。这也就催生了一种与单播路由协议无关的组播路由协议——PIM。PIM协议只专注于组成员和组播源状态相关的信息，而选取路径的信息直接从单播路由表获取。因此它需要维护庞大的路由信息，从而降低了PIM协议的复杂性。这使得PIM协议成为应用最广泛的域内组播协议。

说明：

由PIM路由器所组成的网络称为PIM网络。通常一个大的PIM网络可以划分为多个PIM域来管理和控制组播报文的转发，这里的域内组播协议即是指PIM域内组播协议。

图1 PIM网络示意图



组播分发树

PIM网络以组播组为单位在路由器上建立一点到多点的组播转发路径。由于组播转发路径呈现树型结构，也称为组播分发树MDT (Multicast Distribution Tree)。

组播分发树主要包括以下两种：

- 以组播源为根，组播组成员为叶子的组播分发树称为SPT (Shortest Path Tree)。SPT同时适用于PIM-DM网络和PIM-SM网络。

如[图1]中的RouterE→RouterD→RouterA (RouterB)，就是一棵以Source为根，以HostA、HostB为叶子的SPT。

- 以RP (Rendezvous Point) 为根，组播组成员为叶子的组播分发树称为RPT (RP Tree)。RPT适用于PIM-SM网络。

PIM路由器

在接口上使能了PIM协议的路由器即为PIM路由器。在建立组播分发树的过程中，PIM路由器又分为下几种：

- 叶子路由器：与用户主机相连的PIM路由器，但连接的用户主机不一定为组成员，如[图1]中的RouterA、RouterB、RouterC。
- 第一跳路由器：组播转发路径上，与组播源相连且负责转发该组播源发出的组播数据的PIM路由器如[图1]中的RouterE。
- 最后一跳路由器：组播转发路径上，与组播组成员相连且负责向该组成员转发组播数据的PIM路由器。如[图1]中的RouterA、RouterB。
- 中间路由器：组播转发路径上，第一跳路由器与最后一跳路由器之间的PIM路由器。如[图1]中的RouterD。

PIM路由表项

PIM路由表项即通过PIM协议建立的组播路由表项。PIM网络中存在两种路由表项：(S, G)路由表项或(*, G)路由表项。S表示组播源，G表示组播组，*表示任意。

- (S, G)路由表项主要用于在PIM网络中建立SPT。对于PIM-DM网络和PIM-SM网络适用。
- (*, G)路由表项主要用于在PIM网络中建立RPT。对于PIM-SM网络适用。

PIM路由器上可能同时存在两种路由表项。当收到源地址为S，组地址为G的组播报文，且RPF检查通的情况下，按照如下的规则转发：

- 如果存在(S, G)路由表项，则由(S, G)路由表项指导报文转发。
- 如果不存在(S, G)路由表项，只存在(*, G)路由表项，则先依照(*, G)路由表项创建(S, G)路由表项，再由(S, G)路由表项指导报文转发。

PIM路由表项中主要用于指导转发的信息如下：

- 组播源地址。
- 组播组地址。
- 上游接口：本地路由器上接收到组播数据的接口，如[图1]中的Int3。

- 下游接口：将组播数据转发出去的接口，如[图1]中的Int1、Int2。



PIM-DM基本概述

- 采用“推（Push）模式”转发组播报文。
- PIM-DM的关键任务：
 - 建立SPT（Shortest Path Tree，最短路径树）。
- PIM-DM的工作机制：
 - 邻居发现。
 - 扩散与剪枝。
 - 状态刷新。
 - 嫁接。
 - 断言。

PIM-DM使用“推（Push）模式”转发组播报文，一般应用于组播组成员规模相对较小、相对密集的网络。在实现过程中，它会假设网络中的组成员分布非常稠密，每个网段都可能存在组成员。当有活跃组播源出现时，PIM-DM会将组播源发来的组播报文扩散到整个网络的PIM路由器上，再裁剪掉不存组成员的分支。PIM-DM通过周期性的进行“扩散（Flooding）—剪枝（Prune）”，来构建并维护棵连接组播源和组成员的单向无环SPT（Source Specific Shortest Path Tree）。如果在下一次“扩-剪枝”进行前，被裁剪掉的分支由于其叶子路由器上有新的组成员加入而希望提前恢复转发状态，可通过嫁接（Graft）机制主动恢复其对组播报文的转发。

PIM-DM的关键工作机制包括邻居发现、扩散、剪枝、嫁接、断言和状态刷新。其中，扩散、剪枝、嫁接是构建SPT的主要方法。

邻居发现（Neighbor Discovery）

在PIM路由器每个使能了PIM协议的接口上，都会对外发送Hello报文。封装Hello报文的组播报文的地址是224.0.0.13（表示同一网段中所有PIM路由器）、源地址为接口的IP地址、TTL数值为1。

Hello报文的作用：发现PIM邻居、协调各项PIM协议报文参数、维持邻居关系。

● 发现PIM邻居

同一网段中的PIM路由器都必须接收目的地址为224.0.0.13的组播报文。这样直接相连的PIM路由器间通过交互Hello报文以后，就可以彼此知道自己的邻居信息，建立邻居关系。

只有邻居关系建立成功后，PIM路由器才能接收其他PIM协议报文，从而创建组播路由表项。

● 协调各项PIM协议报文参数

Hello报文中携带多项PIM协议报文参数，主要用于PIM邻居之间PIM协议报文的控制。具体如下：

- DR_Priority：表示各路由器接口竞选DR的优先级，优先级越高越容易获胜。
- Holdtime：表示保持邻居为可达状态的超时时间。

- LAN_Delay: 表示共享网段内传输Prune报文的延迟时间。
 - Neighbor-Tracking: 表示邻居跟踪功能。
 - Override-Interval: 表示Hello报文中携带的否决剪枝的时间间隔。
- 维持邻居关系

PIM路由器之间周期性地发送Hello报文。如果Holdtime超时还没有收到该PIM邻居发出的新的Hello文，PIM路由器就认为该邻居不可达，将其从邻居列表中清除。

PIM邻居的变化将导致网络中组播拓扑的变化。如果组播分发树上的某上游邻居或下游邻居不可达，导致组播路由重新收敛，组播分发树迁移。

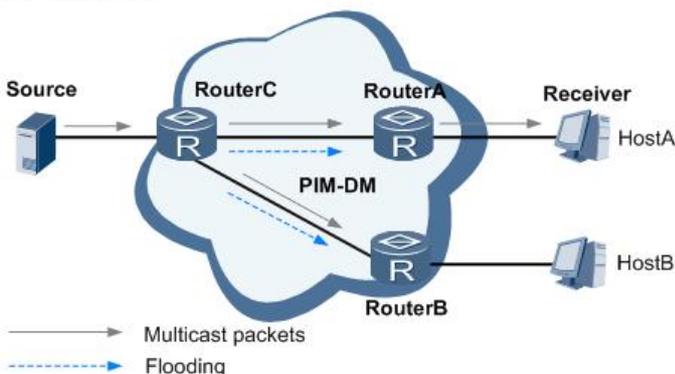
扩散 (Flooding)

当PIM-DM网络中出现活跃的组播源之后，组播源发送的组播报文将在全网内扩散。当PIM路由器接到组播报文，根据单播路由表进行RPF检查通过后，就会在该路由器上创建 (S, G) 表项，下游接口表中包括除上游接口之外与所有PIM邻居相连的接口，后续到达的组播报文将从各个下游接口转发出。

最后组播报文扩散到达叶子路由器，会出现以下两种情况：

- 若与该叶子路由器相连用户网段上存在组成员，则将与该网段相连的接口加入 (S, G) 表项的下游接口列表中，后续的组播报文会向组成员转发。
- 若与该叶子路由器相连用户网段上不存在组成员，且不需要向其下游PIM邻居转发组播报文，则执行剪枝]机制。

图1 扩散示意图



扩散过程如下：

1. 组播源S开始向组播组G发送组播报文。
2. RouterC接收到源发送的组播报文后，根据单播路由表进行RPF检查。RPF检查通过，创建 (S, G) 表项，下游接口列表包括与RouterA和RouterB相连的接口，后续到达的报文向RouterA和RouterB发。
3. RouterA接收到来自RouterC的组播报文，RPF检查成功，在本地创建对应 (S, G) 表项，下游接口列表添加与组成员HostA相连的接口，后续到达的报文向HostA转发。
4. RouterB接收到来自RouterC的组播报文，由于与RouterB相连下游网段不存在组成员和PIM邻居执行剪枝。

剪枝 (Prune)

当PIM路由器接收到组播报文后，RPF检查通过，但是下游网段没有组播报文需求。此时PIM路由器向上游发送剪枝报文，通知上游路由器禁止相应下游接口的转发，将其从(S, G)表项的下游接口列表中删除。剪枝操作由叶子路由器发起，逐跳向上，最终组播转发路径上只存在与组成员相连的分支。

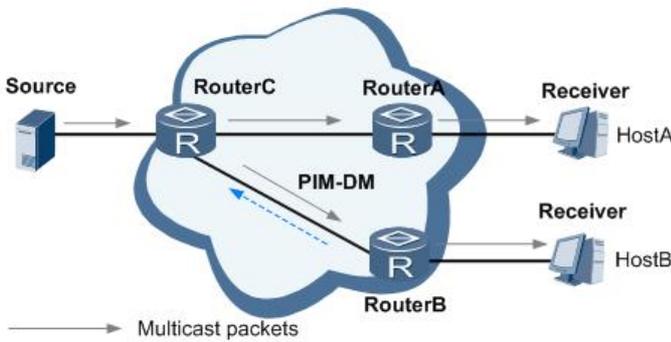
路由器为被裁剪的下游接口启动一个剪枝定时器，定时器超时后接口恢复转发。组播报文重新在全网范围内扩散，新加入的组成员可以接收到组播报文。随后，下游不存在组成员的叶子路由器将向上发起剪枝操作。通过这种周期性的扩散-剪枝，PIM-DM周期性的刷新SPT。

当下游接口被剪枝后：

- 如果下游叶子路由器有组成员加入，并且希望在下次“扩散-剪枝”前就恢复组播报文转发，则执行[嫁接]机制。
- 如果下游叶子路由器一直没有组成员加入，希望该接口保持抑制转发状态，则执行[状态刷新]机制。

如[图2]所示，RouterB上未连接组成员，这种情况下，RouterB向上游发起剪枝。

图3 嫁接示意图



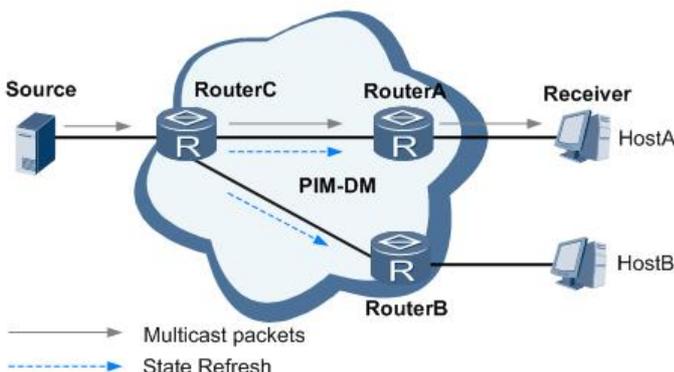
嫁接过程如下：

1. RouterB希望在下一次“剪枝-扩散”来临前恢复对HostB组播报文的转发，向上游路由器RouterC送Graft报文，请求恢复相应出接口转发组播报文。
2. RouterC收到Graft报文后，恢复该接口转发，将该接口添加到(S, G)表项中的下游接口列表中。由于RouterC上有组播报文到达，嫁接过程停止。后续到达的报文向RouterB转发。

状态刷新 (State Refresh)

在PIM-DM网络中，为了避免被裁剪的接口因为“剪枝定时器”超时而恢复转发，离组播源最近的第跳路由器会周期性地触发State Refresh报文在全网内扩散。收到State Refresh报文的PIM路由器会重新剪枝定时器的状态。被裁剪接口的下游叶子路由器如果一直没有组成员加入，该接口将一直处于抑制转发状态。

如[图4]所示，与RouterC上被裁剪接口相连的叶子路由器上一一直没有组成员加入。图4状态刷新示意图



状态刷新过程如下：

1. RouterC触发状态刷新机制，将State Refresh报文向RouterA和RouterB扩散。
2. RouterC上存在被裁剪接口，刷新该接口的“剪枝定时器”的状态。下一次“扩散-剪枝”来临时由于RouterB上仍然没有组成员加入，RouterC上被裁剪的接口将被抑制转发组播报文。

断言 (Assert)

当一个网段内有多个相连的PIM路由器RPF检查通过向该网段转发组播报文时，则需要通过断言机制保证只有一个PIM路由器向该网段转发组播报文。PIM路由器在接收到邻居路由器发送的相同组播报文后，会以组播的方式向本网段的所有PIM路由器发送Assert报文，其中目的地址为永久组地址224.0.0.3。其它PIM路由器在接收到Assert报文后，将自身参数与对方报文中携带的参数做比较，进行Assert竞选。竞选规则如下：

1. 单播路由协议优先级较高者获胜。
2. 如果优先级相同，则到组播源的开销较小者获胜。
3. 如果以上都相同，则下游接口IP地址最大者获胜。

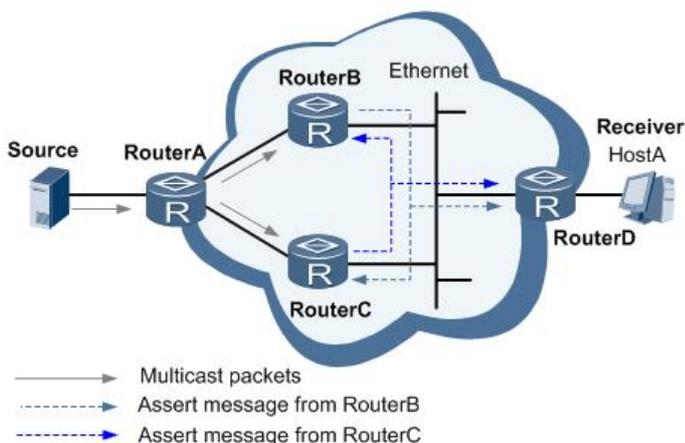
根据Assert竞选结果，路由器将执行不同的操作：

- 获胜一方的下游接口称为Assert Winner，将负责后续对该网段组播报文的转发。
- 落败一方的下游接口称为Assert Loser，后续不会对该网段转发组播报文，PIM路由器也会将其从（S，G）表项下游接口列表中删除。

Assert竞选结束后，该网段上只存在一个下游接口，只传输一份组播报文。所有Assert Loser可以周期性地恢复组播报文转发，从而引发周期性的Assert竞选。

如[图5]所示，RouterB和RouterC均通过了RPF检查，创建了（S，G）表项。并且两者的下游接口连在同一网段，RouterB和RouterC都向该网段发送组播报文。

图5 断言示意图



断言过程如下：

1. RouterB和RouterC从各自下游接口接收到对方发来的组播报文，RPF检查都失败，报文被丢弃。时，RouterB和RouterC分别向该网段发送Assert报文。
2. RouterB将自身的路由信息与RouterC发来的Assert报文中携带的路由信息进行比较，由于自身到播源的开销较小而获胜。于是后续组播报文仍然向该网段转发，RouterC在接收到组播报文后仍然由RPF检查失败而丢弃。

3. RouterC将自身的路由信息与RouterB发来的Assert报文中携带的路由信息进行比较，由于自身到播源的开销较大而落败。于是禁止相应下游接口向该网段转发组播报文，将其从(S, G)表项的下游接口列表中删除。

在ASM (Any-Source Multicast) 模型中，PIM-SM使用“拉 (Pull) 模式”转发组播报文，一般应于组播组成员规模相对较大、相对稀疏的网络。基于这一种稀疏的网络模型，它的实现方法是：

- 在网络中维护一台重要的PIM路由器：汇聚点RP (Rendezvous Point) ，可以为随时出现的组成员或组播源服务。网络中所有PIM路由器都知道RP的位置。
- 当网络中出现组成员 (用户主机通过IGMP加入某组播组G) 时，最后一跳路由器向RP发送Join报，逐跳创建 (*, G) 表项，生成一棵以RP为根的RPT。
- 当网络中出现活跃的组播源 (组播源向某组播组G发送第一个组播数据) 时，第一跳路由器将组播数据封装在Register报文中单播发往RP，在RP上创建 (S, G) 表项，注册源信息。

在ASM模型中，PIM-SM的关键机制包括邻居发现、DR竞选、RP发现、RPT构建、组播源注册、SP切换、断言；同时也可通过配置BSR (Bootstrap Router) 管理域来实现单个PIM-SM域的精细化管理