本機多點傳送流 — 任意來源多點傳送模型

目錄

簡介

本檔案介紹任何來源多點傳送(ASM)模型的封包流量。

背景資訊

本文提供了本地組播資料包流的詳細資料包流及其輸出分析，描述了控制平面和轉發平面中的詳細分析輸出和資料包流。

ASM是接收者不瞭解傳送者的模型。這意味著它可以接收來自任何來源的流量。接收方只知道傳送方使用的組播組和網際網路組管理協定(IGMP)，以便訂閱接收所有發往此地址的流量。

所有這一切都將在本文檔中介紹：

1.當接收器處於活動狀態時發生的情況。

2.當源處於活動狀態時發生的情況。

3.在集結點(RP)接收註冊時會發生的情況。

4.(S，G)如何形成。直到第一跳路由器(FHR)。

5.第一個組播流採用的路徑。

6.在最後一跳路由器(LHR)上收到兩個流時會發生的情況。

7.最短路徑樹(SPT)在共用樹上的形成方式。到底發生了什麼以及為什麼進行切換。

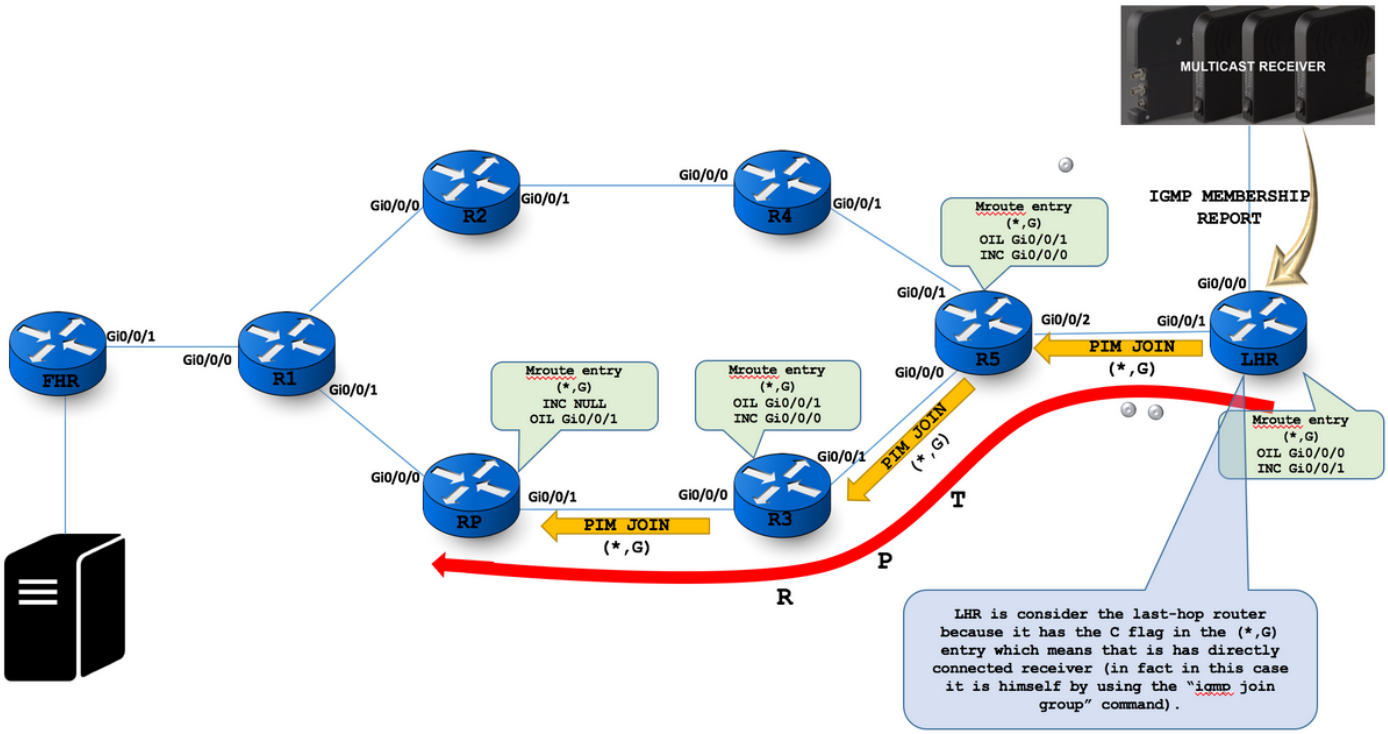
協定無關組播(PIM)用作源和接收器之間的組播路由協定，以建立組播樹。在ASM中，使用(*,G)組播條目，其中*表示任何源，G是接收流量時感興趣的組播組地址接收器。

步驟1.當接收器處於活動狀態時，它將傳送IGMP報告消息

- 收到接收者的興趣表達式後，指定路由器(DR)將向該組播組的RP傳送PIM加入消息。

- 此連線消息稱為(*,G)連線，因為它將所有源的組G連線到該組。
- (*,G)連線逐跳流向組的RP，並且在其通過的每個路由器中，組G的組播樹狀態被例項化。

LHR被視為最後一跳路由器，因為它在(*,G)條目中具有C標誌，這表示它具有直連線收器(實際上在此情況下使用**igmp join group**命令是它自己)。



Step 1 : On receiving the receiver's expression of interest, the DR then sends a PIM Join message towards the RP for that multicast group.
This Join message is known as a (*,G) Join because it joins group G for all sources to that group.
The (*,G) Join travels hop-by-hop towards the RP for the group, and in each router it passes through, multicast tree state for group G is instantiated.

```
LHR#sh ip igmp groups
IGMP Connected Group Membership
Group Address    Interface                Uptime    Expires   Last Reporter   Group Accounted
224.1.1.1        GigabitEthernet1/0       00:37:30  00:02:02  10.0.108.8
224.0.1.40       FastEthernet0/0          01:21:01  00:02:43  10.0.78.8

LHR#sh ip mroute
(*, 224.1.1.1), 00:00:29/00:02:30, RP 4.4.4.4, flags: SJCL
  Incoming interface: GigabitEthernet2/0/1, RPF nbr 10.0.78.7
  Outgoing interface list:
    GigabitEthernet0/0/0, Forward/Sparse

C flag in the (*,G) entry which means that is has directly connected receiver.

RP #sh ip mroute
(*, 224.1.1.1), 00:10:39/00:02:39, RP 4.4.4.4, flags: S
  Incoming interface: Null, RPF nbr 0.0.0.0
  Outgoing interface list:
    FastEthernet0/0, Forward/Sparse

S flag Sparse mode created.

(*, 224.0.1.40), 01:56:40/00:02:58, RP 4.4.4.4, flags: SJPCL
  Incoming interface: FastEthernet0/0, RPF nbr 10.0.78.7
  Outgoing interface list: Null(*, 224.0.1.40), 01:56:40/00:02:58, RP 4.4.4.4, flags: SJPCL
  Incoming interface: FastEthernet0/0, RPF nbr 10.0.78.7
  Outgoing interface list: Null

There is a single (*,G) entry for the group 224.0.1.40 which is Auto-RP Discovery group address.
```

The value of "0.0.0.0" means self, and it appears in the output if the router is the RP itself

NOTE : To prevent a stale PIM-SM forwarding state from getting stuck in the routers, it is given a finite lifetime (3 minutes), after which it is deleted. Routers refresh shared trees by periodically (once a minute) sending (*, G) Joins to the upstream neighbor in the direction of the RP.

Actually the PIM register message encapsulates the multicast packet sent by the source into a unicast packet.

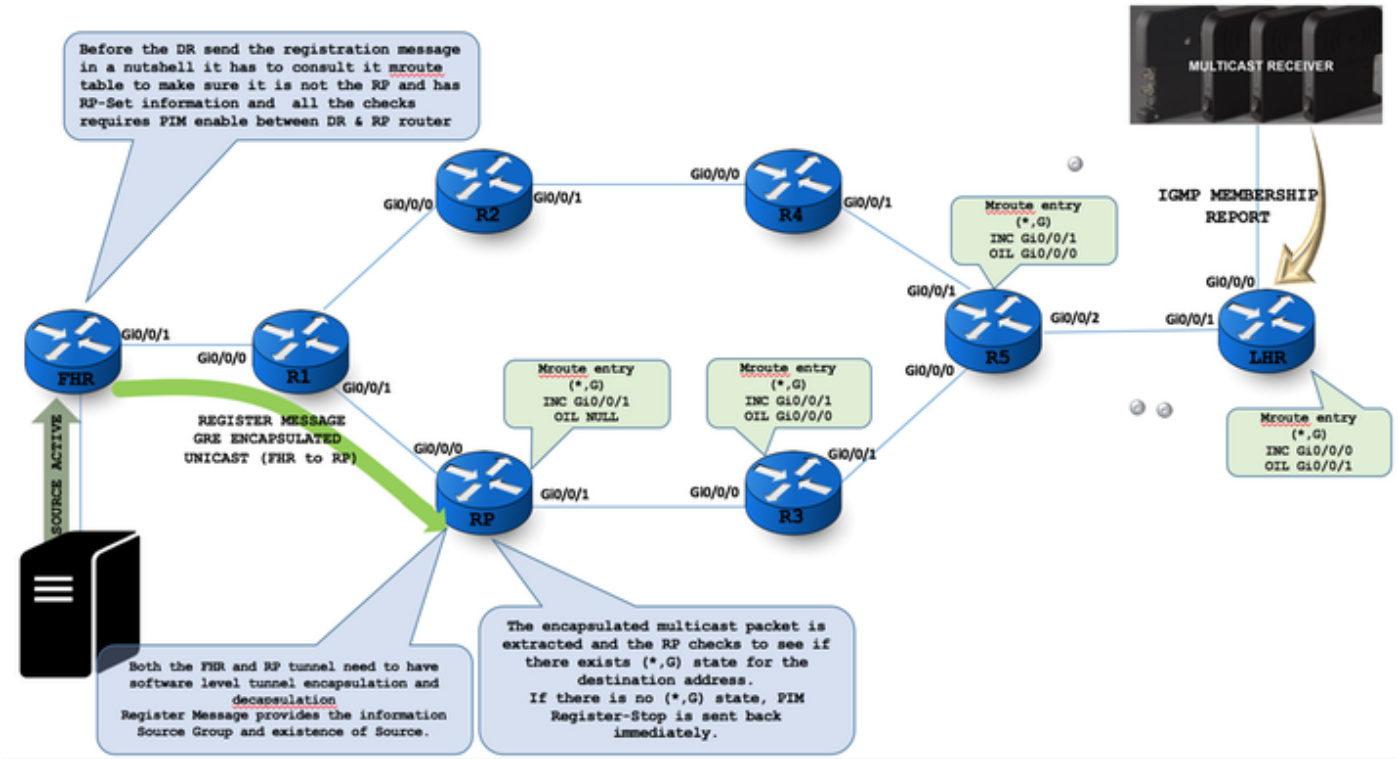
```
▶ Frame 59: 68 bytes on wire (544 bits), 68 bytes captured (544 bits) on interface 0
▶ Ethernet II, Src: ca:08:fa:92:00:00 (ca:08:fa:92:00:00), Dst: IPv4mcast_0d (01:00:5e:00:00:0d)
▼ Internet Protocol Version 4, Src: 10.0.78.8, Dst: 224.0.0.13
    0100 .... = Version: 4
    .... 0101 = Header Length: 20 bytes
  ▶ Differentiated Services Field: 0xc0 (DSCP: CS6, ECN: Not-ECT)
    Total Length: 54
    Identification: 0x0b27 (2855)
  ▶ Flags: 0x00
    Fragment offset: 0
    Time to live: 1
    Protocol: PIM (103)
  ▶ Header checksum: 0x7565 [validation disabled]
    Source: 10.0.78.8
    Destination: 224.0.0.13
    [Source GeoIP: Unknown]
    [Destination GeoIP: Unknown]
▼ Protocol Independent Multicast
    0010 .... = Version: 2
    .... 0011 = Type: Join/Prune (3)
    Reserved byte(s): 00
    Checksum: 0x87c7 [correct]
  ▼ PIM Options
      Upstream-neighbor: 10.0.78.7
      Reserved byte(s): 00
      Num Groups: 1
      Holdtime: 210
    ▼ Group 0: 224.10.10.10/32
      ▶ Num Joins: 1
        Num Prunes: 0
```

TTL is always 1. Which means it's a RP/RE destined packet.

PIM JOIN Message carries the active group address

步驟2.當源處於活動狀態時

- 在DR傳送註冊消息之前，簡而言之，它必須查閱mroute表以確保它不是RP並具有RP-Set資訊，並且所有檢查都要求在DR和RP路由器之間啟用PIM。
- FHR和RP隧道都需要軟體級隧道封裝和解除封裝。
- Register Message提供Source Group和Source存在的資訊。
- 將提取封裝的組播資料包，RP檢查目標地址是否存在(*,G)狀態。
- 如果沒有(*,G)狀態，PIM註冊停止將立即傳送回。



Once Source is active :

```
FHR #
(1.1.1.1, 224.22.22.44), 00:03:15/00:00:02, flags: PFT
  Incoming interface: Loopback0, RPF nbr 0.0.0.0, Registering
  Outgoing interface list: Null
```

Register flag (F) is enabled for registration process in the FHR.

```
F flag: Source is directly connected and the register process must be used to notify the RP to this source.
P flag: Outgoing interface is null as no one has joined the SPT tree yet for this source
T flag: traffic is being received from the source.
```

PIM must enable between DR & RP router to send and receive the Register message.

```
▶ Frame 442: 142 bytes on wire (1136 bits), 142 bytes captured (1136 bits) on interface 0
▶ Ethernet II, Src: ca:01:c1:46:00:1c (ca:01:c1:46:00:1c), Dst: ca:02:c1:6a:00:00 (ca:02:c1:6a:00:00)
▶ Internet Protocol Version 4, Src: 10.0.12.1, Dst: 4.4.4.4
▼ Protocol Independent Multicast
    0010 .... = Version: 2
    .... 0001 = Type: Register (1)
    Reserved byte(s): 00
    Checksum: 0xdeff [correct]
  ▼ PIM Options
    ▶ Flags: 0x00000000
      0100 .... = IP Version: IPv4 (4)
▶ Internet Protocol Version 4, Src: 1.1.1.1, Dst: 224.10.10.10
▶ Internet Control Message Protocol
```

If no active receiver present at RP, then RP sends REGISTER STOP DR will be silent for default 60 seconds may result in the so-called "join latency" where a newly Joined listener may have to way for almost a minute before it can discover a multicast source. This is why in many practical deployments with dynamic listeners you see PIM SSM being used in favor of complicated PIM SM mechanics.

1.1.1.1	224.22.22.44	PIMv2	142	Register
4.4.4.4	10.0.91.1	PIMv2	52	Register-stop

```
RP #
(1.1.1.1, 224.22.22.44), 00:00:43/00:02:16, flags: P
  Incoming interface: FastEthernet0/0, RPF nbr 10.0.24.2
  Outgoing interface list: Null
```

Prune Flag (P) is set as no active receiver (*,G) entry present in RP.

RP SENDS REGISTER STOP WHEN NO ACTIVE RECIEVER FOR THE GROUP AND DISCARD THE MULTICAST PACKET

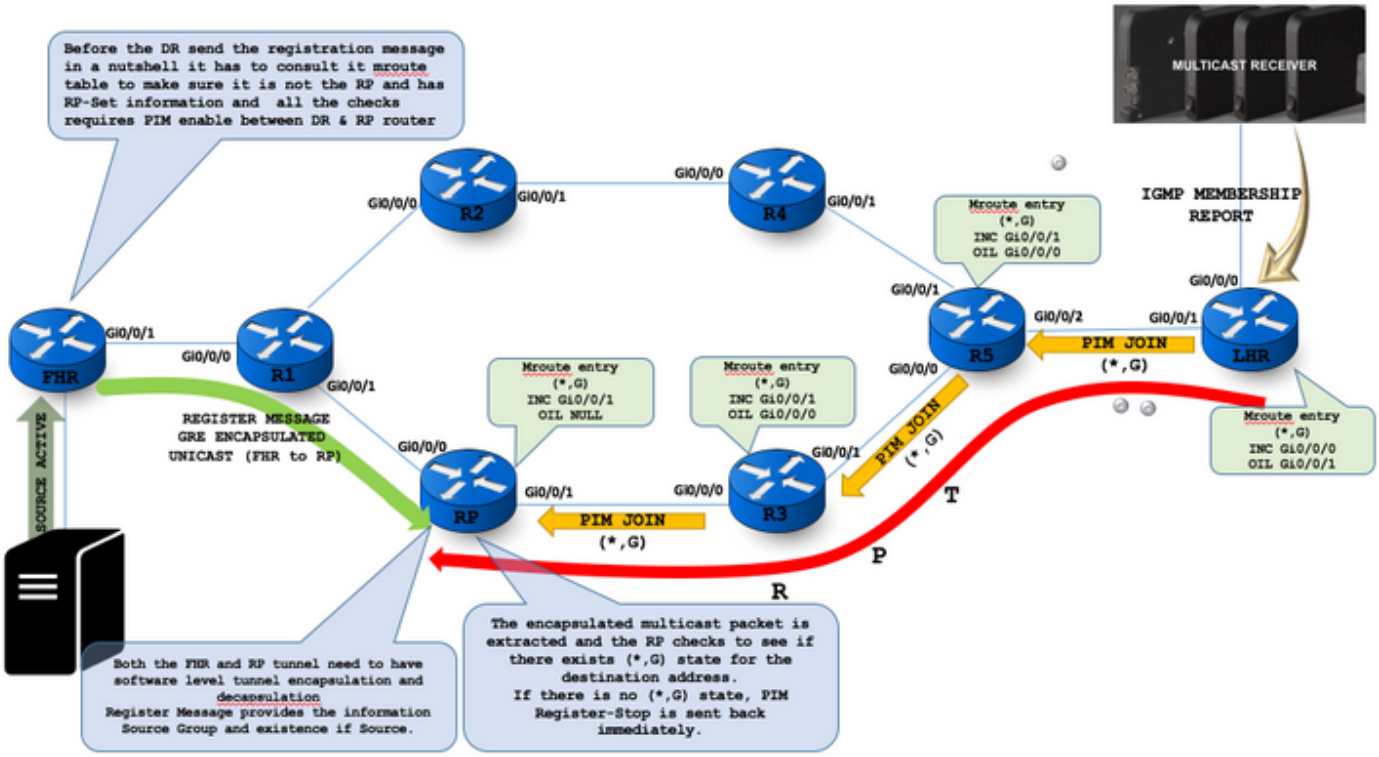
```
▶ Frame 973: 52 bytes on wire (416 bits), 52 bytes captured (416 bits) on interface 0
▶ Ethernet II, Src: ca:02:c1:6a:00:00 (ca:02:c1:6a:00:00), Dst: ca:01:c1:46:00:1c (ca:01:c1:46:00:1c)
▶ Internet Protocol Version 4, Src: 4.4.4.4, Dst: 10.0.91.1
▼ Protocol Independent Multicast
    0010 .... = Version: 2
    .... 0010 = Type: Register-stop (2)
    Reserved byte(s): 00
    Checksum: 0xe39a [correct]
  ▼ PIM Options
      Group: 224.22.22.44/32
      Source: 1.1.1.1
```

步驟3.表單共用樹

- 在DR傳送註冊消息之前，簡而言之，它必須查閱mroute表以確保它不是RP並具有RP-Set資訊

，並且所有檢查都要求在DR和RP路由器之間啟用PIM

- FHR和RP隧道都需要軟體級隧道封裝和解除封裝
- Register Message提供Source Group和existence if Source的資訊。
- 將提取封裝的組播資料包，RP檢查目標地址是否存在(*,G)狀態。
- 如果沒有(*,G)狀態，PIM註冊停止將立即傳送回。



The RP also sees that an active shared tree with a nonempty outgoing interface list exists and therefore sends the de-encapsulated packet down the shared tree.

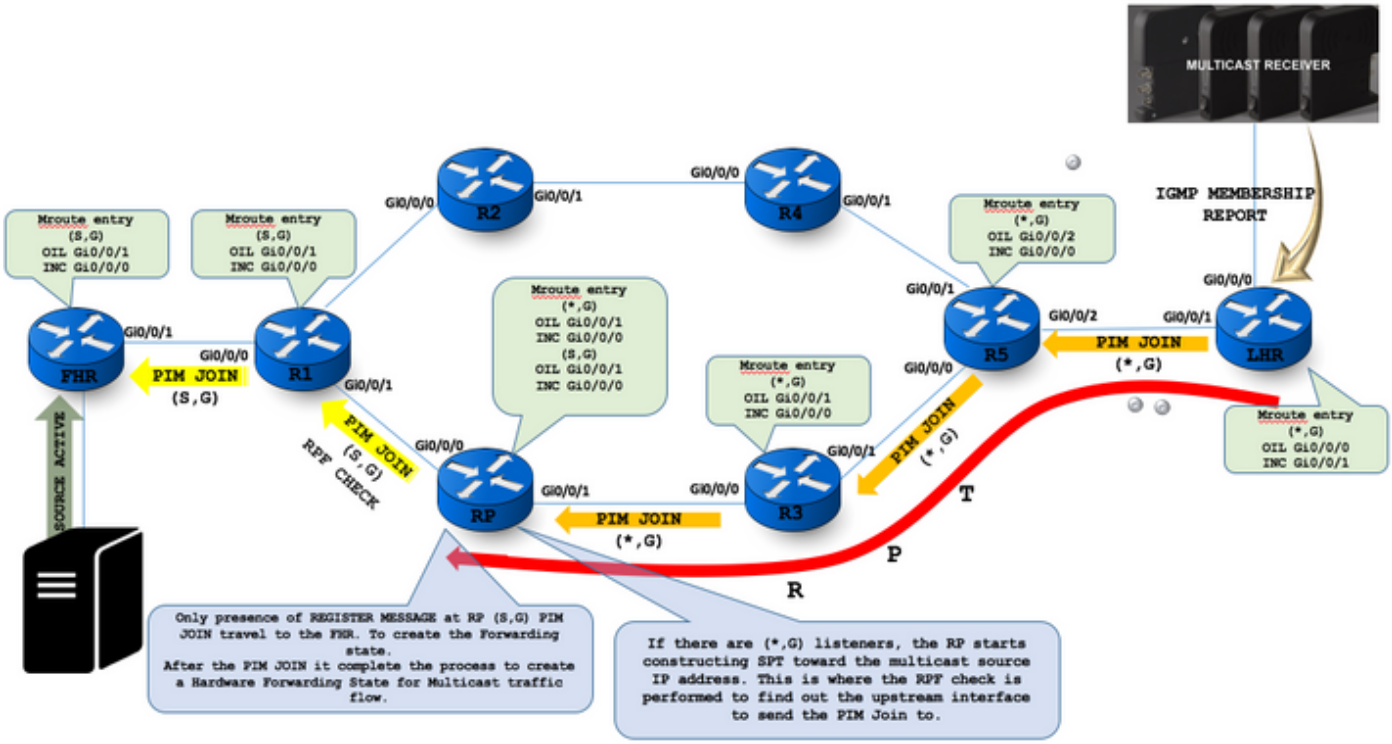
```
RP #
(*, 224.1.1.1), 02:45:12/00:03:11, RP 4.4.4.4, flags: S
  Incoming interface: Null, RPF nbr 0.0.0.0
  Outgoing interface list:
    FastEthernet0/0, Forward/Sparse, 02:45:12/00:03:11

(10.0.12.1, 224.1.1.1), 00:02:42/00:00:21, flags: T
  Incoming interface: FastEthernet0/0, RPF nbr 10.0.24.2
  Outgoing interface list: Null

T Flag set for the shared tree.
```

Presence of (*,G) at RP means active receiver.

```
▶ Frame 29: 76 bytes on wire (608 bits), 76 bytes captured (608 bits) on interface 0
▶ Ethernet II, Src: ca:04:f1:9c:00:00 (ca:04:f1:9c:00:00), Dst: IPv4mcast_0d (01:00:5e:00:00:0d)
▶ Internet Protocol Version 4, Src: 10.0.24.4, Dst: 224.0.0.13
▼ Protocol Independent Multicast
    0010 .... = Version: 2
    .... 0011 = Type: Join/Prune (3)
    Reserved byte(s): 00
    Checksum: 0xb4c2 [correct]
  ▼ PIM Options
      Upstream-neighbor: 10.0.24.2
      Reserved byte(s): 00
      Num Groups: 1
      Holdtime: 210
    ▼ Group 0: 224.1.1.1/32
      ▼ Num Joins: 2
          IP address: 1.1.1.1/32 (S)
          IP address: 10.0.12.1/32 (S)
        Num Prunes: 0
```

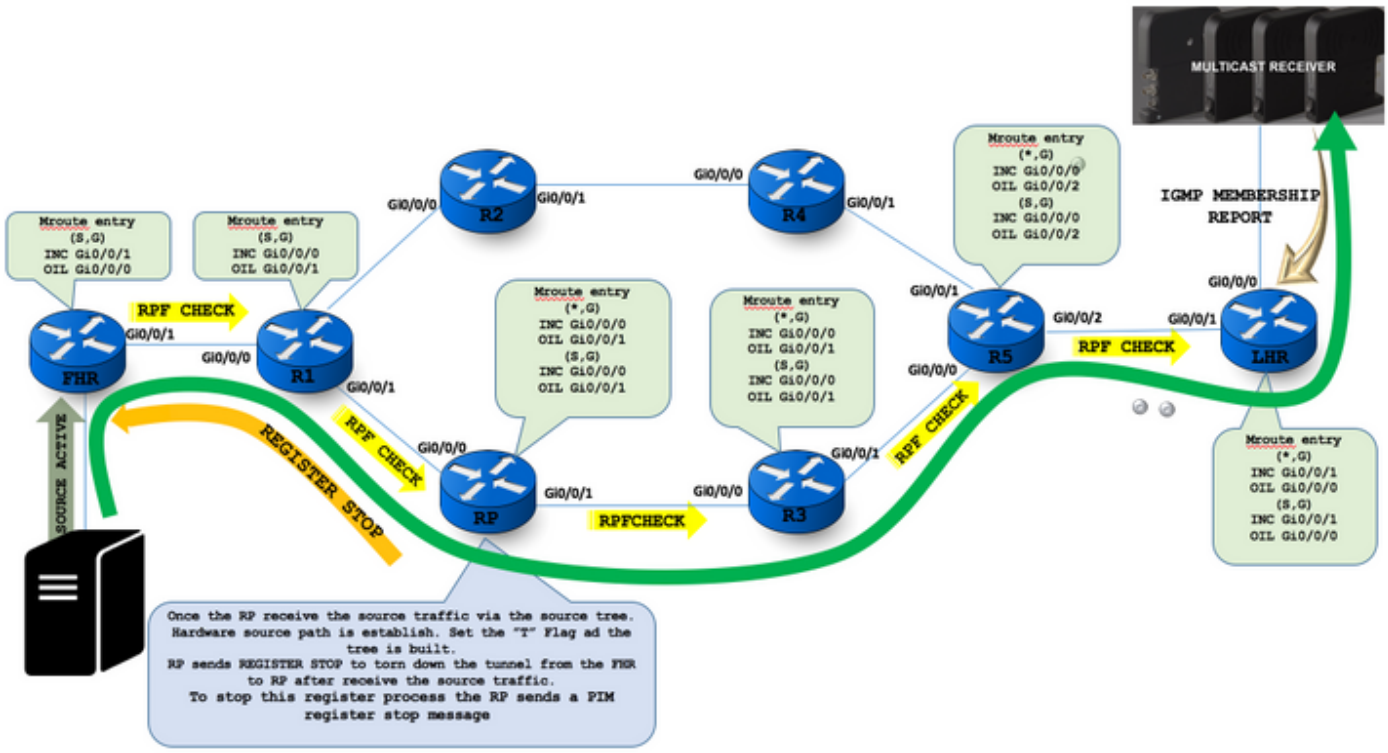
步驟4.(S，G)資料包到達FHR

- 只有在RP(S，G)PIM處存在註冊報文才能加入FHR。建立轉發狀態。
- PIM加入後，它完成為組播流量建立硬體轉發狀態的過程。
- 如果存在(*,G)偵聽器，RP開始向組播源IP地址構建SPT。在此執行RPF檢查以找出要向其傳送PIM加入的上游介面。



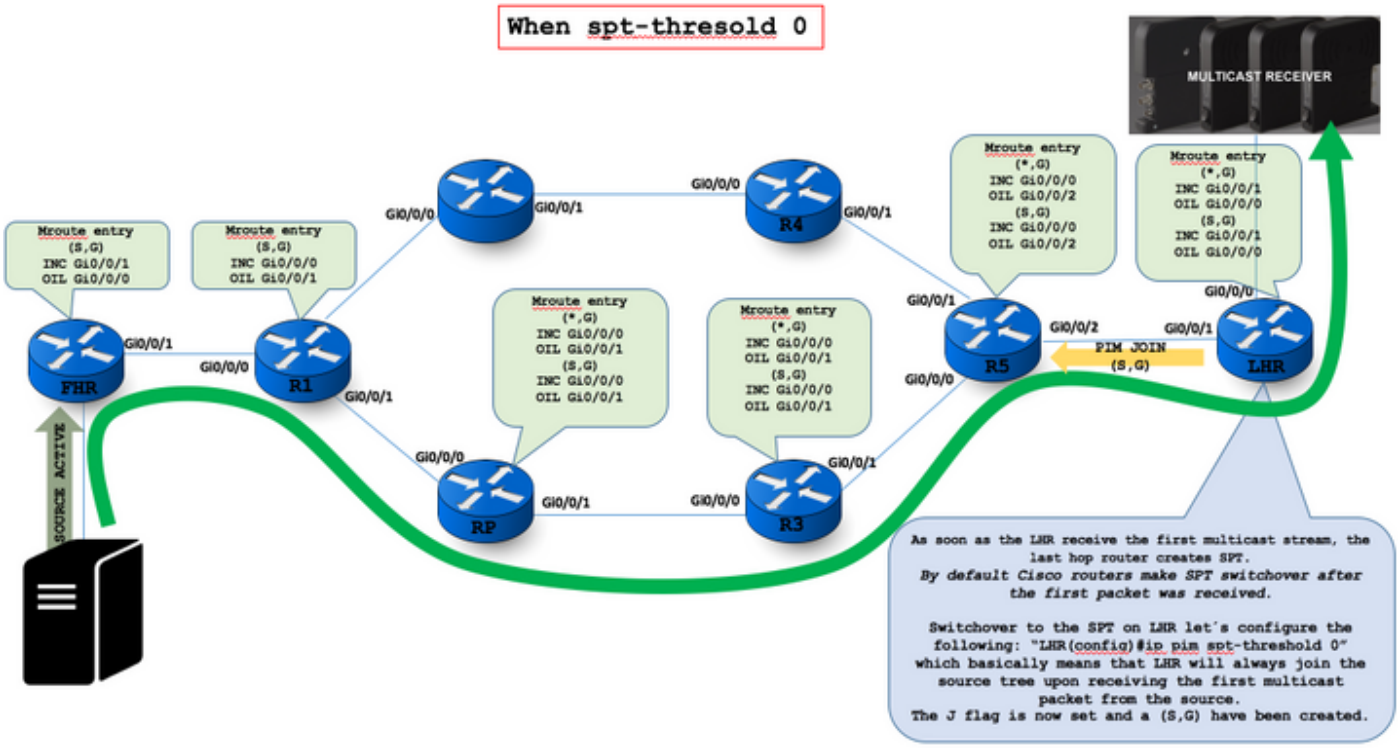
步驟5.組播資料包的第一流，通過共用樹到達接收方

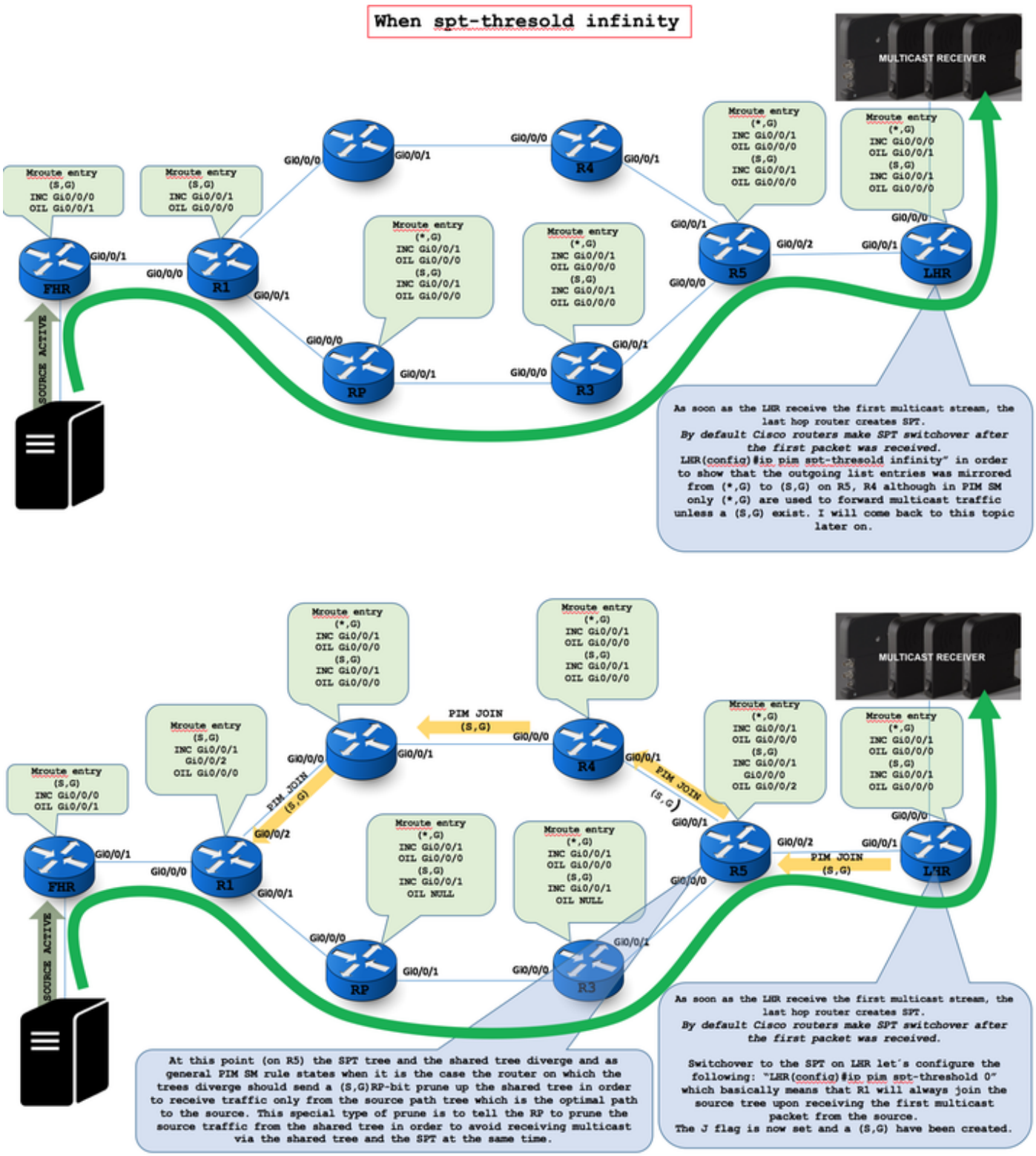
- 僅當在INC /RFP介面上收到組播流時，路由器才會轉發組播流。
- 根據單播RT檢查組播資料包源地址。
- 在傳送加入的源方向確定介面和下一跳組播路由器。
- RP正在加入S的源特定樹，資料包將繼續封裝到RP。當來自S的資料包也開始以本機方式到達RP時，RP將接收每個資料包的兩個副本。
- 此時，RP開始丟棄這些資料包的封裝副本，並向的DR傳送REGISTER STOP消息以防止DR不必要地封裝資料包。

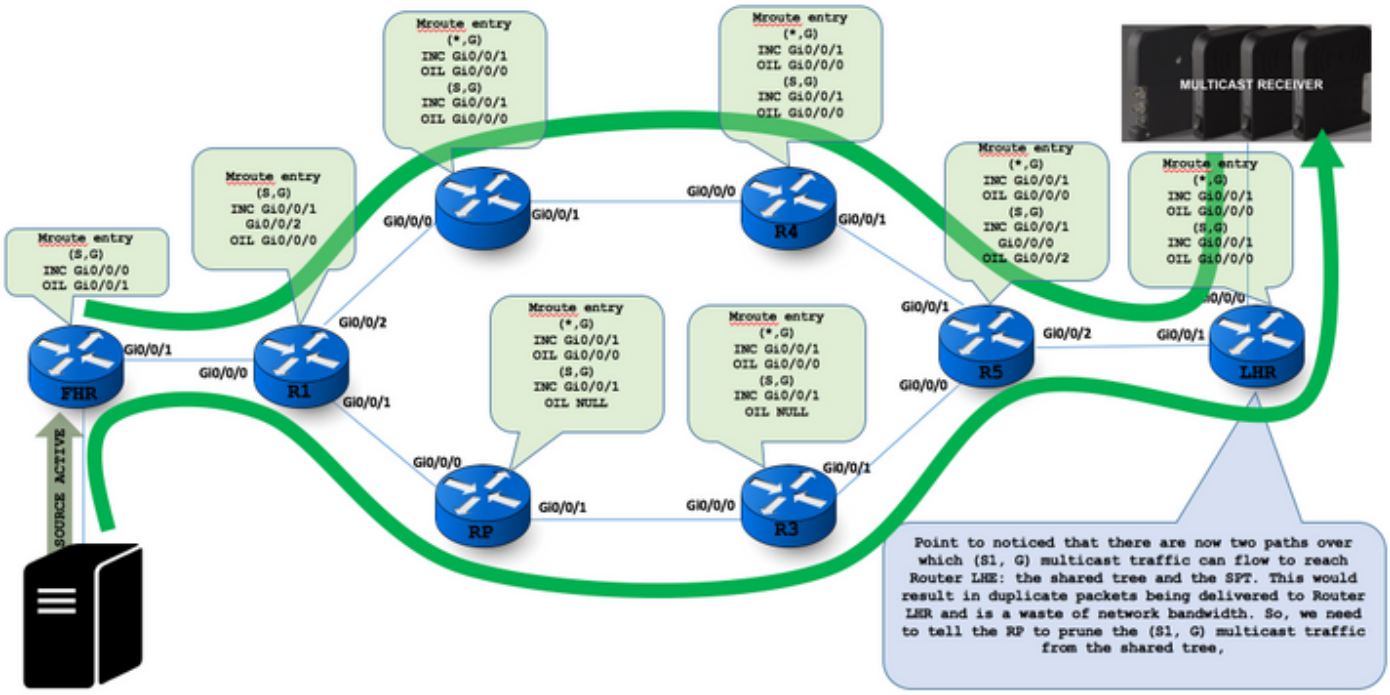


PIM-SM enables a last hop DR (that is, a DR with directly connected hosts that have joined a multicast group) to switch from the shared tree to the SPT for a specific source. This step is usually accomplished by specifying an SPT-Threshold in terms of bandwidth. If this threshold is exceeded, the last-hop DR joins the SPT. (Cisco routers have this threshold set to zero by default, which means that the SPT is joined as soon the first multicast packet from a source has been received via the shared tree.)

- 一旦RP通過源樹接收源流量。建立硬體源路徑。設定「T」標誌，生成樹。
- RP在收到源流量後傳送REGISTER STOP以斷開從FHR到RP的隧道。
- 要停止此註冊過程，RP會傳送PIM註冊停止消息




When spt-thresold infinity
MULTICAST RECEIVER
Mroute entry
(S,G)
INC Gi0/0/0
OIL Gi0/0/1
Mroute entry
(S,G)
INC Gi0/0/1
OIL Gi0/0/0
Mroute entry
(*,G)
INC Gi0/0/1
OIL Gi0/0/0
(S,G)
INC Gi0/0/1
OIL Gi0/0/0
Mroute entry
(*,G)
INC Gi0/0/1
OIL Gi0/0/0
(S,G)
INC Gi0/0/1
OIL Gi0/0/0
Mroute entry
(*,G)
INC Gi0/0/1
OIL Gi0/0/0
(S,G)
INC Gi0/0/1
OIL Gi0/0/0
Mroute entry
(*,G)
INC Gi0/0/0
OIL Gi0/0/1
(S,G)
INC Gi0/0/1
OIL Gi0/0/0
FHR
R1
RP
R3
R4
R5
LHR
Gi0/0/0
Gi0/0/1
Gi0/0/2
SOURCE ACTIVE
As soon as the LHR receive the first multicast stream, the last hop router creates SPT.
By default Cisco routers make SPT switchover after the first packet was received.
LHR(config)#ip pim spt-thresold infinity" in order to show that the outgoing list entries was mirrored from (*,G) to (S,G) on R5, R4 although in PIM SM only (*,G) are used to forward multicast traffic unless a (S,G) exist. I will come back to this topic later on.
MULTICAST RECEIVER
Mroute entry
(S,G)
INC Gi0/0/0
OIL Gi0/0/1
Mroute entry
(S,G)
INC Gi0/0/1
Gi0/0/2
OIL Gi0/0/0
Mroute entry
(*,G)
INC Gi0/0/1
OIL Gi0/0/0
(S,G)
INC Gi0/0/1
OIL Gi0/0/0
Mroute entry
(*,G)
INC Gi0/0/1
OIL Gi0/0/0
(S,G)
INC Gi0/0/1
OIL Gi0/0/0
PIM JOIN
(S,G)
Mroute entry
(*,G)
INC Gi0/0/1
OIL Gi0/0/0
(S,G)
INC Gi0/0/1
Gi0/0/0
OIL Gi0/0/2
Mroute entry
(*,G)
INC Gi0/0/1
OIL Gi0/0/0
(S,G)
INC Gi0/0/1
OIL Gi0/0/0
Mroute entry
(*,G)
INC Gi0/0/1
OIL NULL
Mroute entry
(*,G)
INC Gi0/0/1
OIL Gi0/0/0
(S,G)
INC Gi0/0/1
OIL NULL
FHR
R1
RP
R3
R4
R5
L'HR
PIM JOIN
(S,G)
SOURCE ACTIVE
At this point (on R5) the SPT tree and the shared tree diverge and as general PIM SM rule states when it is the case the router on which the trees diverge should send a (S,G)RP-bit prune up the shared tree in order to receive traffic only from the source path tree which is the optimal path to the source. This special type of prune is to tell the RP to prune the source traffic from the shared tree in order to avoid receiving multicast via the shared tree and the SPT at the same time.
As soon as the LHR receive the first multicast stream, the last hop router creates SPT.
By default Cisco routers make SPT switchover after the first packet was received.
Switchover to the SPT on LHR let´s configure the following: "LHR(config)#ip pim spt-threshold 0" which basically means that R1 will always join the source tree upon receiving the first multicast packet from the source.
The J flag is now set and a (S,G) have been created.



步驟6. LHR從SPT接收流量並向共用樹傳送修剪消息

收到兩個組播流量流後，LHR開始接收來自SPT的流量，並向共用樹傳送修剪消息。

J標誌表示相應的(*,G)狀態是通過枝葉路由器交換SPT。

LHR編號

(10.0.12.1、239.1.1.1)、00:00:38/00:02:21、標志：LJT

傳入介面：FastEthernet0/0,RPF引擎10.0.78.7

傳出介面清單：

GigabitEthernet1/0，轉發/稀疏，00:00:38/00:02:21

「F」標誌通常用於在PIM DR路由器上建立的狀態 — 它指示與向RP註冊的流相對應的轉發狀態。如果「F」標誌持續出現，則您的路由器很可能無法從RP接收PIM註冊停止消息，因此有些源尚未切換到SPT。

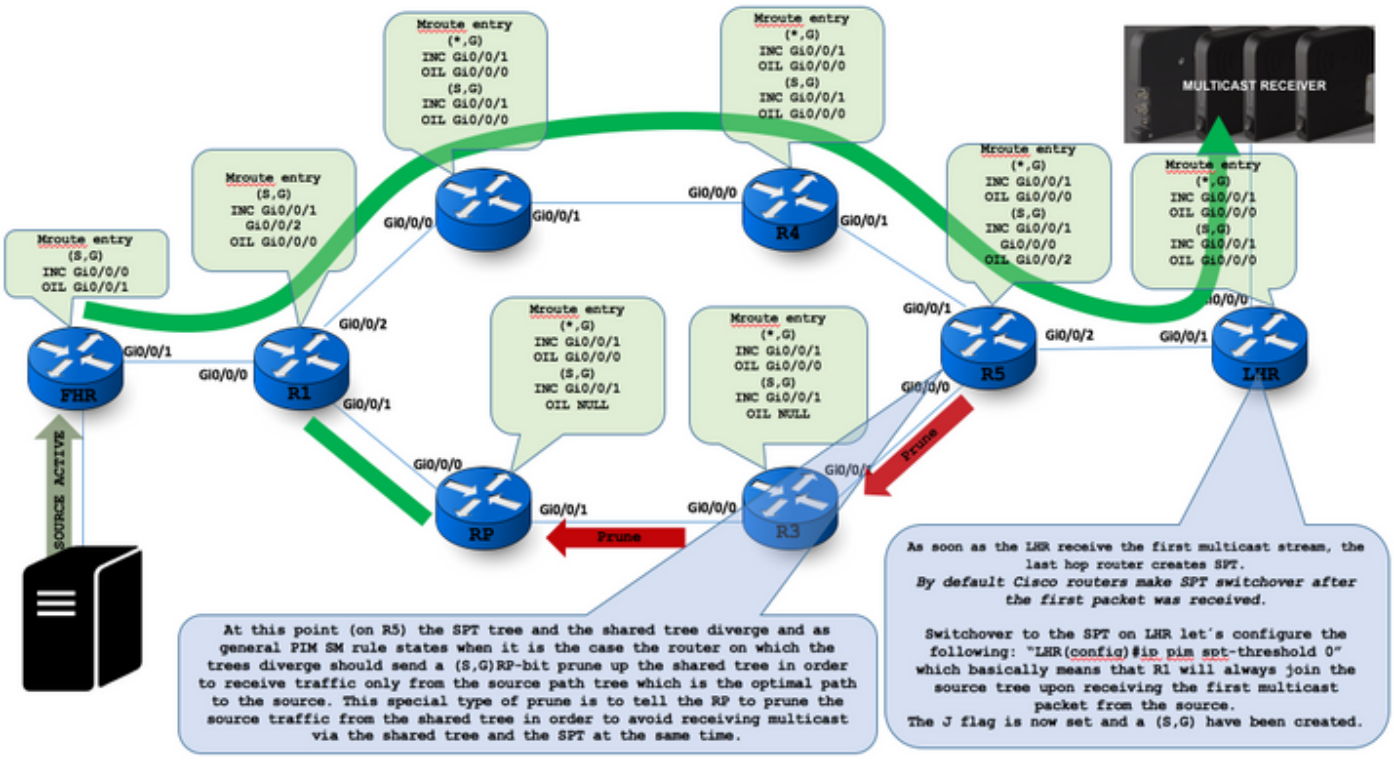
The J flag means the respective (*,G) state is to be switched the SPT by the leaf router.

```
LHR #
(10.0.12.1, 239.1.1.1), 00:00:38/00:02:21, flags: LJT
  Incoming interface: FastEthernet0/0, RPF nbr 10.0.78.7
  Outgoing interface list:
    GigabitEthernet1/0, Forward/Sparse, 00:00:38/00:02:21
```

The "F" flag is typically found for the states created at the PIM DR router - it signalizes the forwarding states that correspond to the flows being registered with the RP. If the "F" flag persists, then your router is most likely not receiving the PIM Register-Stop messages back from the RP, and thus there are sources that has not switched to the SPT tree.

```
FHR #

(*, 239.1.1.1), 00:09:01/stopped, RP 4.4.4.4, flags: SPF
  Incoming interface: Null, RPF nbr 0.0.0.0
  Outgoing interface list: Null

(1.1.1.1, 239.1.1.1), 00:03:02/00:00:15, flags: PFT
  Incoming interface: Loopback0, RPF nbr 0.0.0.0, Registering
  Outgoing interface list: Null
```

There is an (S,G) entry in this table, which has the flag "T" meaning it's a shortest-path and not a shared tree construct. The incoming interface is set to Loopback0 and RPF neighbor to "0.0.0.0" which means the local router is the traffic source.



The receiver (or a router upstream of the receiver) will be receiving two copies of the data: one from the SPT and one from the RPT. When the first traffic starts to arrive from the SPT, the DR or upstream router starts to drop the packets for G from S that arrive via the RP tree. In addition, it sends an (S,G) Prune message towards the RP.
This is known as an (S,G,rpt) Prune. The Prune message travels hop-by-hop, instantiating state along the path towards the RP indicating that traffic from S for G should NOT be forwarded in this direction. The prune is propagated until it reaches the RP or a router that still needs the traffic from S for other receivers.

At this point (on R5) the SPT tree and the shared tree diverge and as general PIM SM rule states when it is the case the router on which the trees diverge should send a (S,G)RP-bit prune up the shared tree in order to receive traffic only from the source path tree which is the optimal path to the source. This special type of prune is to tell the RP to prune the source traffic from the shared tree in order to avoid receiving multicast via the shared tree and the SPT at the same time.

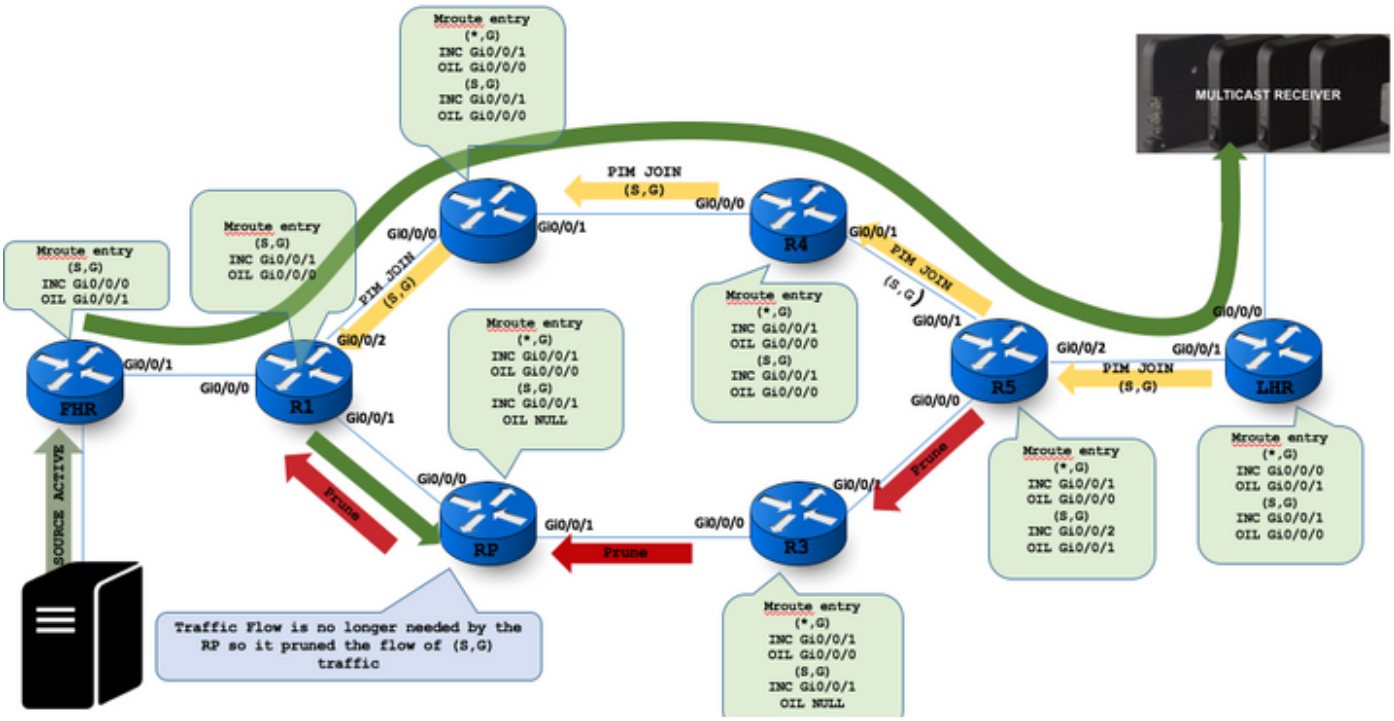
```
RP #
(10.0.12.1, 224.1.1.1), 00:00:10/00:02:53, flags: PTX
  Incoming interface: FastEthernet0/0, RPF nbr 10.0.24.2
  Outgoing interface list: Null
```

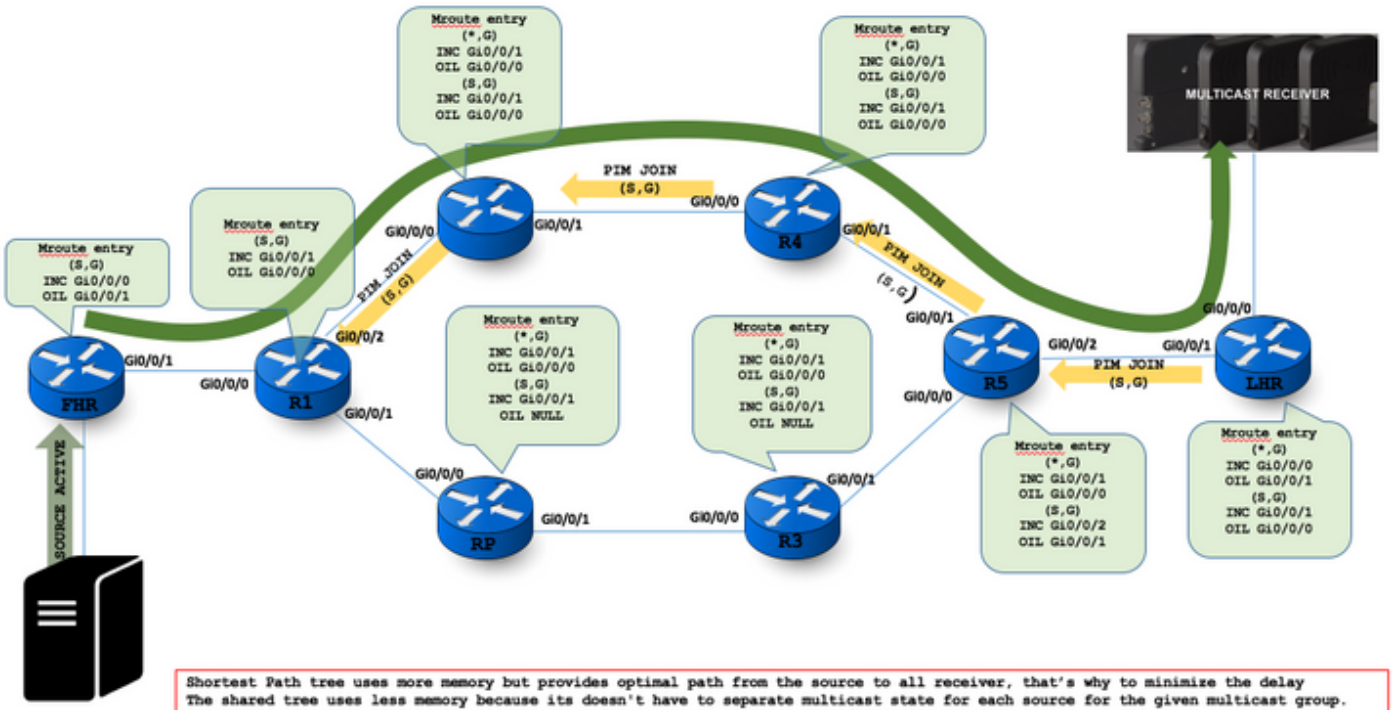
P Bit (Prune Flag) received from the diverge point.

```
LHR #

(10.0.12.1, 224.1.1.1), 00:01:59/00:01:00, flags: LJT
  Incoming interface: FastEthernet0/0, RPF nbr 10.0.78.7
  Outgoing interface list:
    GigabitEthernet1/0, Forward/Sparse, 00:01:59/00:02:57
```

J Flag Join the SPT// T Flag Tree formed

"PIM Join/Prune Messages" the RP flag (also referred to as the RP-bit) indicates that this message is applicable to the shared tree and should be forwarded up the shared tree toward the RP. Setting this flag/bit in an (S1, G) Prune and sending it up the shared tree tells the routers along the shared tree to prune Source S1 multicast traffic from the shared tree.





Shortest Path tree uses more memory but provides optimal path from the source to all receiver, that's why to minimize the delay
The shared tree uses less memory because its doesn't have to separate multicast state for each source for the given multicast group.
But may create a suboptimal routing for some receiver. Shared tree also introduced extra delay.

"Incoming interface" is set to Null, which means there is no incoming traffic for this group. If any physical interface the traffic is their.

"C" means there is a group-member directly connected

"L" means the router itself joined the group.

```
R5#sh ip mroute

(*, 239.1.1.1), 00:27:32/00:02:08, RP 4.4.4.4, flags: SJCL
  Incoming interface: FastEthernet0/0, RPF nbr 10.0.78.7
  Outgoing interface list:
    GigabitEthernet1/0, Forward/Sparse, 00:27:32/00:02:08
```

possibly the next-hop router

Expire times (How soon the group will expired if no refreshed

Uptime (How long this state has been created)

Incoming interface: Null, RPF nbr 155.29.0.5

If the incoming interface is null and the RPF neighbor is IP address, then there is a RPF failure. Mtrace will confirm the issue.