Comprensión del Mecanismo de Afirmación PIM

Contenido

Introducción

Este documento describe el mecanismo de afirmación Protocol Independent Multicast (PIM), se concentra en los criterios del ganador de PIM y profundiza en ciertos casos esquinales.

Prerequisites

Requirements

Cisco recomienda que tenga conocimiento del mecanismo de afirmación PIM.

Componentes Utilizados

La información de este documento se basa en Cisco CSR1000V versión 16.4.1

The information in this document was created from the devices in a specific lab environment. All of the devices used in this document started with a cleared (default) configuration. Si tiene una red en vivo, asegúrese de entender el posible impacto de cualquier comando.

¿Qué es el mecanismo de afirmación PIM?

Cuando hay varios routers habilitados para PIM en un segmento compartido, es posible que estos routers encuentren tráfico multicast duplicado. Este podría ser el caso porque dos o más routers en el mismo segmento compartido podrían tener una entrada válida (S,G) o (*,G) que rellene la interfaz saliente hacia el segmento compartido para el mismo grupo de IP/destino de origen.

El mecanismo de afirmación PIM se utiliza para detectar y eliminar la duplicación del tráfico multicast en un segmento compartido. Es importante tener en cuenta que este mecanismo no evita la duplicación de eventos, sino que utiliza la duplicación del tráfico multicast como disparador para activar este mecanismo que elige un único reenviador para este flujo.

Cuando tiene duplicación de tráfico multidifusión en un segmento compartido, puede suponer que hay varios routers que envían el mismo (S,G) o (*,G) en un segmento compartido. Si elige un router para reenviar este flujo de forma eficaz, elimina la duplicación.

PIM aprovecha los mensajes PIM assert que se activan cuando recibe un paquete multidifusión en la lista de interfaces salientes (OIL). Estos mensajes de afirmación contienen métricas que se utilizan para calcular quién se convertirá en ganador seguro. Los routers descendentes en la LAN también reciben mensajes de afirmación PIM. Estos mensajes son luego utilizados por los dispositivos descendentes para enviar los mensajes de unión/separación adecuados al router ascendente que ganó la elección segura.

Escenario 1. Justificación de LHR

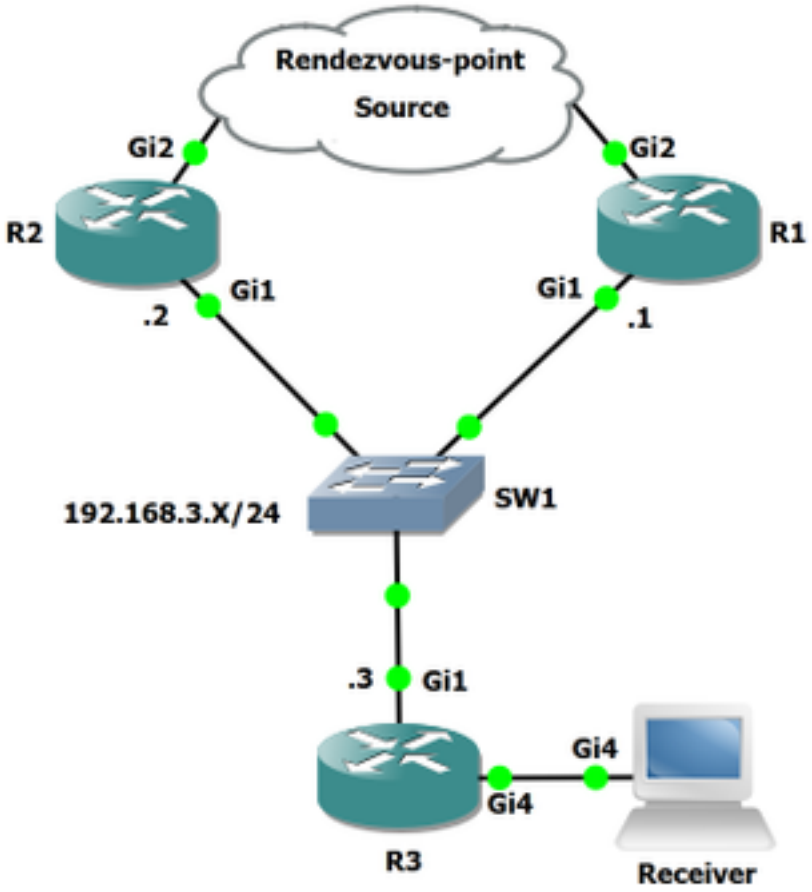


Figura 1.

En el diagrama de red, R3 es el router de último salto (LHR), R3 se conecta a R2 y R1 a través de un segmento compartido.

Cuando recibe un informe de protocolo de administración de grupos de Internet (IGMP) del receptor, R3 verifica quién es el vecino RPF hacia el RP. En la topología, R1 es el vecino RPF hacia el RP, por lo que R3 envía una unión (*,G) hacia R1. Una vez que R1 baje por la secuencia (suponga que el grupo está activo) R3 envía una unión (S,G) hacia el origen y baja el árbol de origen. R2 es el vecino RPF hacia el árbol de origen, lo que significa que R3 enviará la unión (S,G) hacia R2. R3 tiene la misma interfaz RPF hacia RP y el origen. Aquí puede ver la tabla mroute R3 para el grupo 239.1.1.1.

```
R3#show ip mroute
```

```
IP Multicast Routing Table
Outgoing interface flags: H - Hardware switched, A - Assert winner, p - PIM Join
 Timers: Uptime/Expires
 Interface state: Interface, Next-Hop or VCD, State/Mode
(*, 239.1.1.1), 00:00:55/stopped, RP 192.168.0.100, flags: SJC
  Incoming interface: GigabitEthernet1, RPF nbr 192.168.3.1
  Outgoing interface list:
    GigabitEthernet4, Forward/Sparse, 00:00:55/00:02:04

(10.0.0.2, 239.1.1.1), 00:00:52/00:02:07, flags: JT
  Incoming interface: GigabitEthernet1, RPF nbr 192.168.3.2, Mroute
  Outgoing interface list:
    GigabitEthernet4, Forward/Sparse, 00:00:52/00:02:07

(*, 224.0.1.40), 00:01:22/00:02:09, RP 192.168.0.100, flags: SJPCL
  Incoming interface: GigabitEthernet1, RPF nbr 192.168.3.1
```

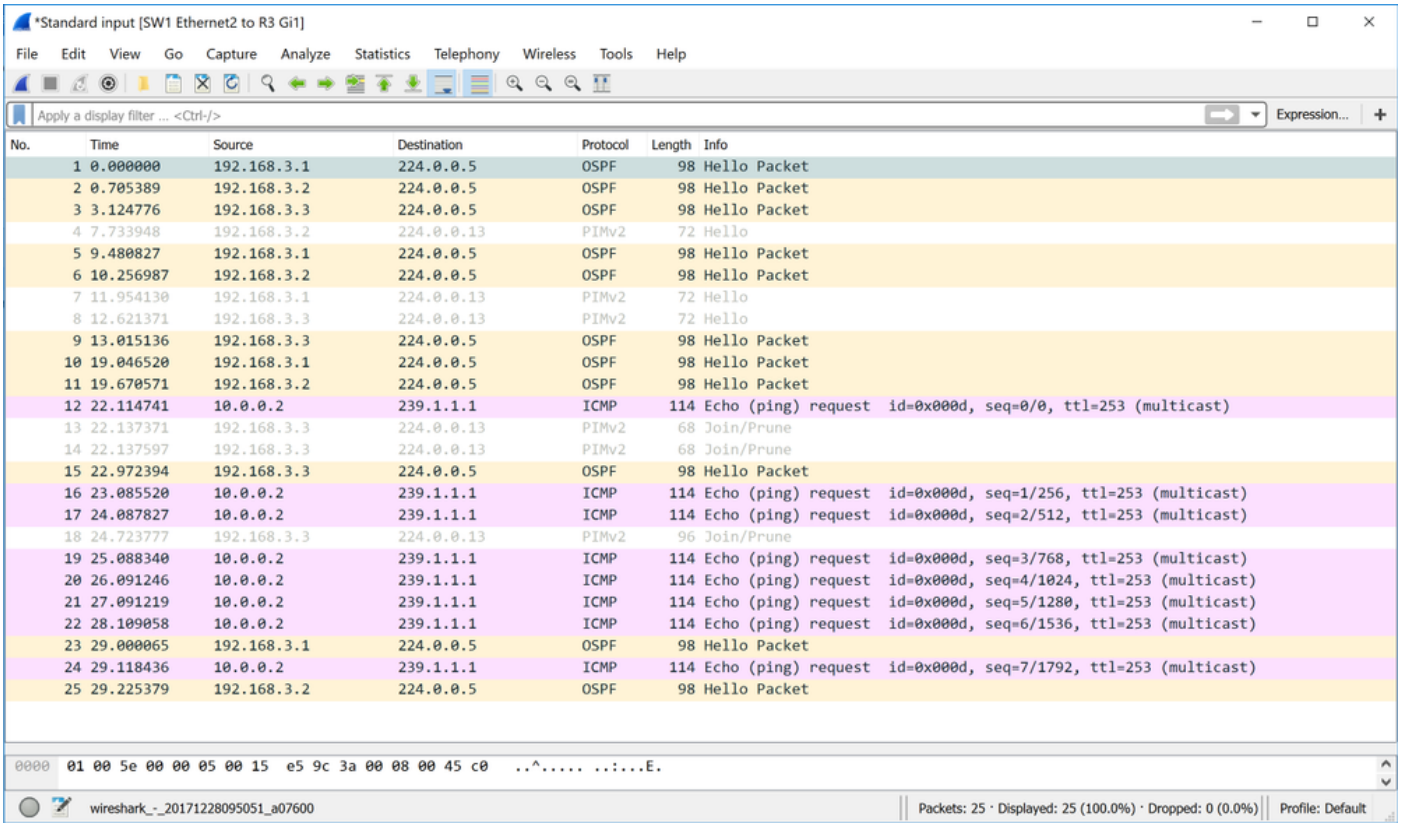
Como puede ver en R3, el vecino (*,G) de RPF es 192.168.3.1 y el vecino de RPF hacia el (S,G) es 192.168.3.2. Ahora, **esto debería resultar en que R1 y R2 tengan un ACEITE válido hacia R3**. Veamos estas entradas:

```
R1#show ip mroute
Outgoing interface flags: H - Hardware switched, A - Assert winner, p - PIM Join
 Timers: Uptime/Expires
 Interface state: Interface, Next-Hop or VCD, State/Mode

(*, 239.1.1.1), 00:15:02/00:02:33, RP 192.168.0.100, flags: S
  Incoming interface: GigabitEthernet2, RPF nbr 192.168.5.2
  Outgoing interface list:
    GigabitEthernet1, Forward/Sparse, 00:15:02/00:02:33

(10.0.0.2, 239.1.1.1), 00:13:24/00:02:33, flags: PR
  Incoming interface: GigabitEthernet2, RPF nbr 192.168.5.2
  Outgoing interface list: Null

(*, 224.0.1.40), 00:29:17/00:02:51, RP 192.168.0.100, flags: SJCL
  Incoming interface: GigabitEthernet2, RPF nbr 192.168.5.2
  Outgoing interface list:
    GigabitEthernet1, Forward/Sparse, 00:16:06/00:02:51
  Outgoing interface list: Null
```

```
R2#show ip mroute
IP Multicast Routing Table
Outgoing interface flags: H - Hardware switched, A - Assert winner, p - PIM Join
 Timers: Uptime/Expires
 Interface state: Interface, Next-Hop or VCD, State/Mode

(*, 239.1.1.1), 00:08:00/stopped, RP 192.168.0.100, flags: SP
  Incoming interface: GigabitEthernet2, RPF nbr 192.168.4.1
  Outgoing interface list: Null

(10.0.0.2, 239.1.1.1), 00:00:03/00:02:56, flags: T
  Incoming interface: GigabitEthernet2, RPF nbr 192.168.4.1
  Outgoing interface list:
    GigabitEthernet1, Forward/Sparse, 00:00:03/00:03:26

(*, 224.0.1.40), 01:37:30/00:02:22, RP 192.168.0.100, flags: SJPL
  Incoming interface: GigabitEthernet2, RPF nbr 192.168.4.1
```

Como se ha mencionado, assert se puede activar cuando hay dos routers ascendentes que tienen un OIL válido completado en un segmento compartido. Dado que tanto R1 como R2 tienen un OIL válido, verifique si hay un mecanismo seguro en la captura de paquetes.

Esta captura de paquetes se capturó en la interfaz R3 Gi1 hacia SW1.



En esta captura de paquetes, no se ve ningún paquete asertivo aunque haya todos los requisitos previos para crear duplicación en el segmento compartido entre R1, R2 y R3. ¿Por qué no ve paquetes PIM assert cuando se activó el flujo (S,G)?

Parece que RFC 7761 podría contener la respuesta a estas preguntas.

Resumen de RFC 7761 Sección 4.2.2.

```
4.2.2.  Setting and Clearing the (S,G) SPTbit

   Basically, Update_SPTbit(S,G,iif) will set the SPTbit if we have the
   appropriate (S,G) join state, and if the packet arrived on the
   correct upstream interface for S, and if one or more of the following
   conditions apply:

   1.  The source is directly connected, in which case the switch to the
       SPT is a no-op.

   2.  The RPF interface to S is different from the RPF interface to the
       RP.  The packet arrived on RPF_interface(S), and so the SPT must
       have been completed.

   3.  No one wants the packet on the RP tree.

   4.  RPF'(S,G) == RPF'(*,G).  In this case, the router will never be
       able to tell if the SPT has been completed, so it should just
```

```
      switch immediately.  The RPF'(S,G) != NULL check ensures that the
      SPTbit is set only if the RPF neighbor towards S is valid.
```

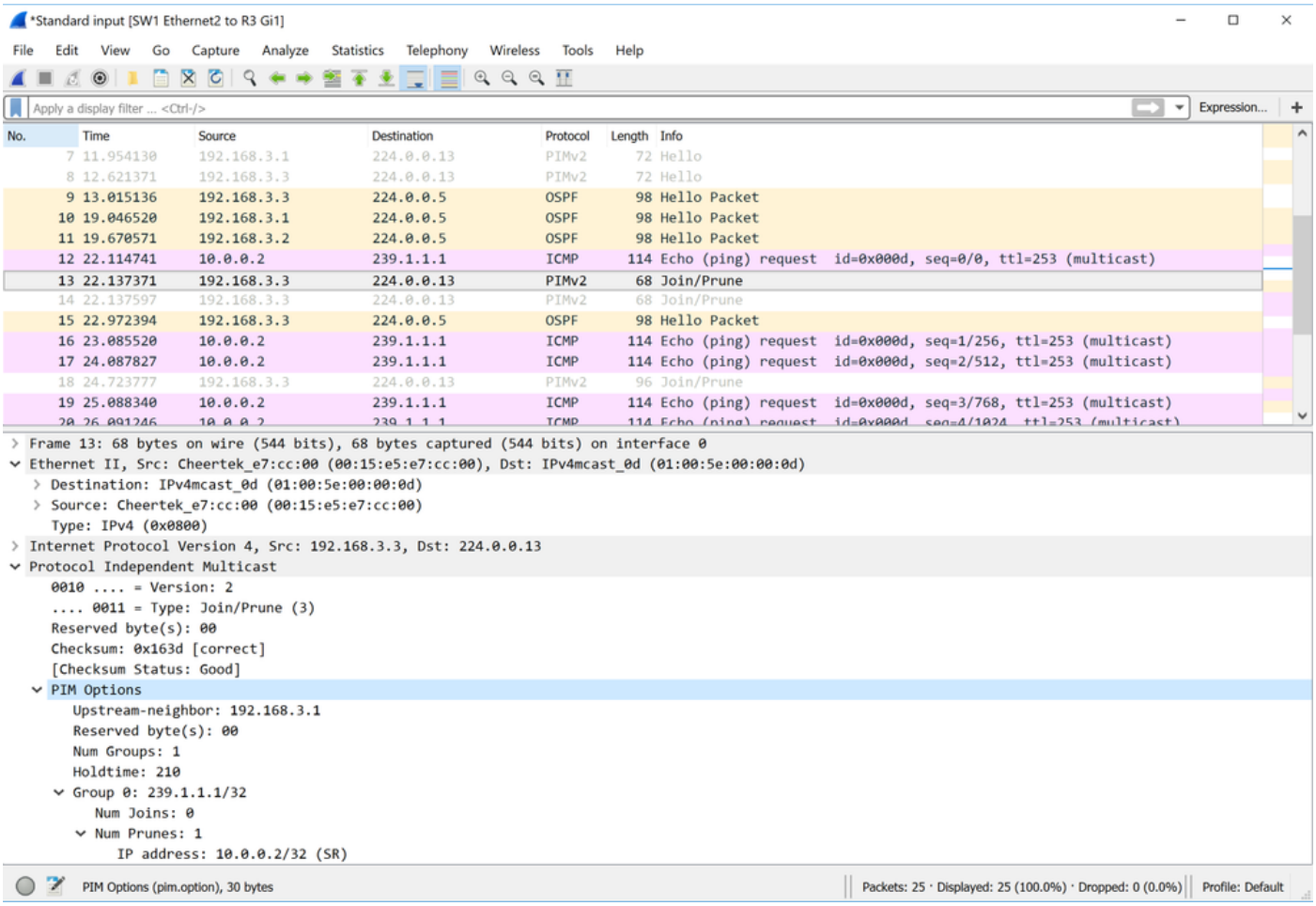
In the case where the RPF interface is the same for the RP and for S, but RPF'(S,G) and RPF'(*,G) differ, we wait for an Assert(S,G), which indicates that the upstream router with (S,G) state believes the SPT has been completed.

El SPTbit (S,G) se utiliza para distinguir si se reenvía en el estado (*,G) o en el estado (S,G). Cuando cambia del árbol RP al árbol de origen, hay un período de transición cuando llegan datos debido al estado ascendente (*,G) mientras se establece el estado ascendente (S,G), en ese momento el router debe continuar reenviando sólo en el estado (*,G). Esto evita los agujeros negros temporales que serían causados por el envío de una(S,G,rpt) antes de que el estado ascendente (S,G) haya terminado de establecerse.

Aunque parece que el escenario puede correlacionarse con el último punto mencionado anteriormente. En el caso de que la interfaz RPF sea la misma para el RP y para S, pero los RPF'(S,G) y RPF'(*,G) difieren, esperamos un Assert(S,G), que indica que el router ascendente con el estado (S,G) cree que el SPT se ha completado.

Para que se active la afirmación, el router debe recibir un paquete duplicado en su OIL ya poblado para el mismo grupo de IP/destino de origen en el segmento. R3 es también un LHR, lo que significa que está designado para cambiar de (*,G) a SPT (S,G) cuando se recibe un paquete de (*,G).

En la captura de paquetes observamos que no se activan aserciones. Aunque sí vemos una poda enviada inmediatamente después de recibir el primer eco ICMP.

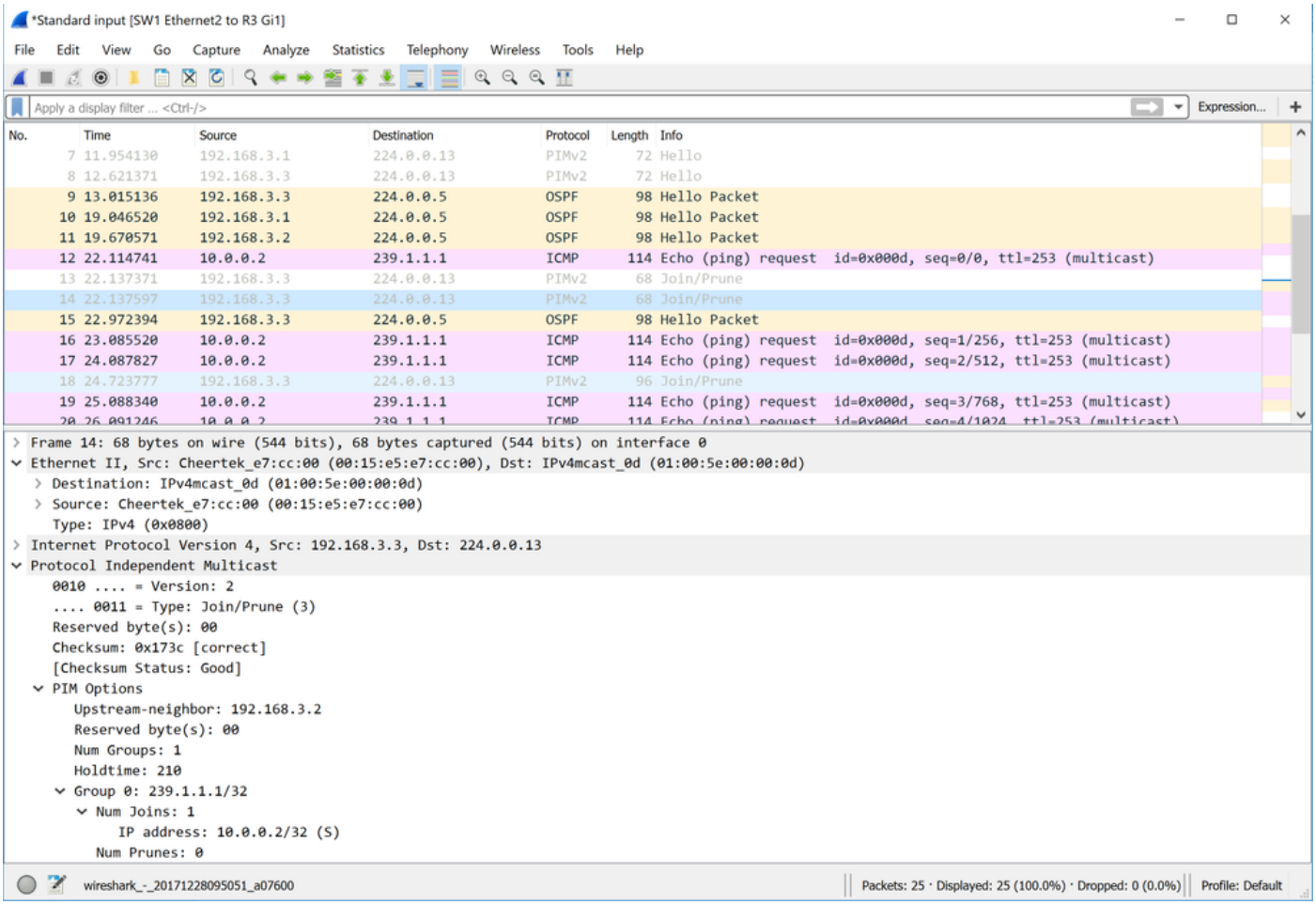


Como puede ver, una vez que se recibe el primer paquete de solicitud del protocolo de mensajes de control de Internet (ICMP) en la interfaz R3 G1, se envía una poda de bits SR (*,G) hacia el vecino ascendente 192.168.3.1. Este recorte (*,G) para el origen específico definido.

Puede ver estos indicadores configurados también: (SR):

```
The S flag: indicates that the multicast group is a sparse mode group.
The R flag: The R flag is the RP-bit flag and indicates that the information in the (S, G) entry
is applicable to the shared tree.
```

En el segundo paquete PIM Nº 14, puede ver que R3 intenta unirse al árbol (S,G).



Se observa que una vez que se recibe el primer plano de datos, el paquete R3 recorta el (*,G) y genera el (S,G). Esta es la razón por la que no ve los paquetes PIM assert. Este escenario determinado está en vigor cuando tiene un LHR que tiene la misma interfaz RPF para (S,G) y (*,G). Aunque este comportamiento puede diferir ligeramente de RFC 7761, no debería causar ningún problema.

Ahora continuemos con la situación 2. El diagrama de esta situación se puede ver aquí:

Situación hipotética 2. Selección de ruta de aserción

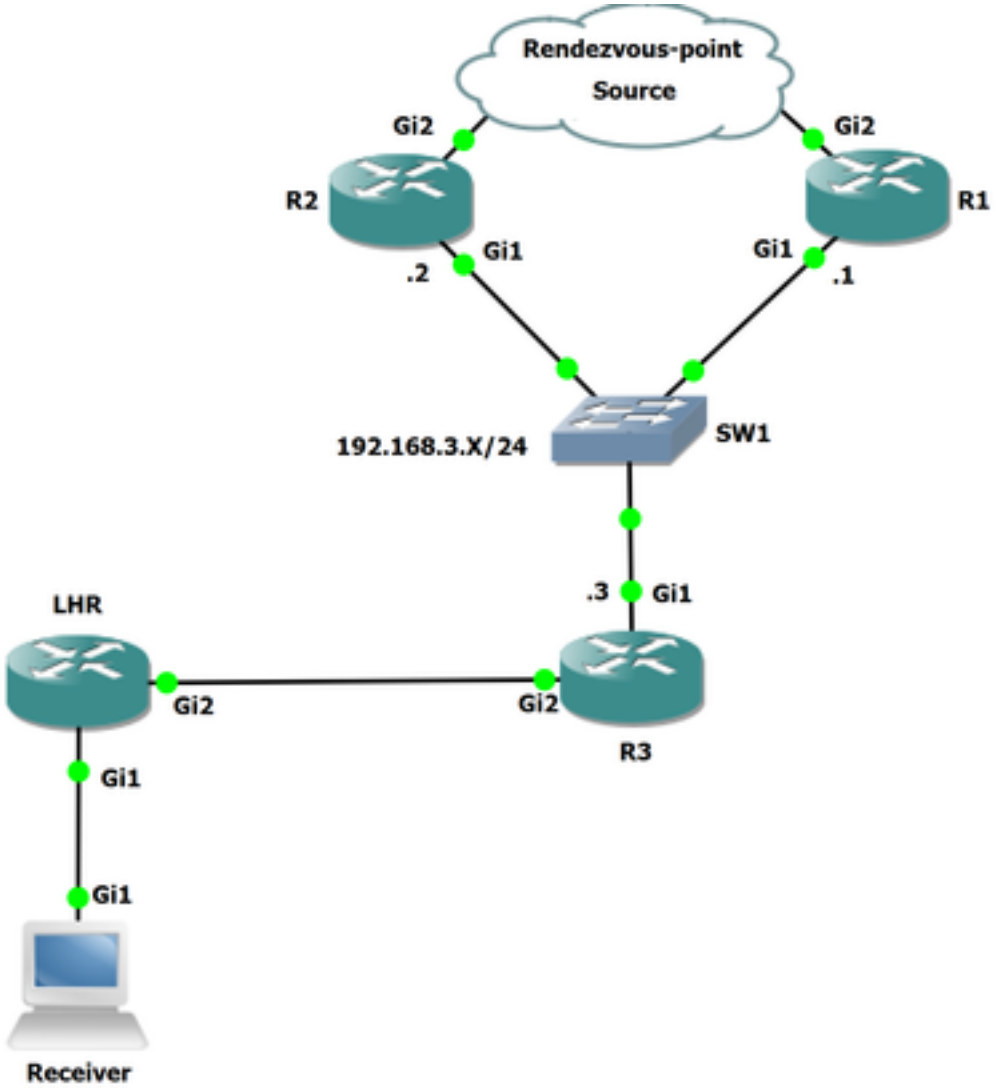


Figura 2

En esta topología, hay otro router conectado en R3 que es el LHR. El LHR se conecta directamente al receptor. El origen y el RP están por encima de R2 y R1. El vecino RPF en R3 hacia el RP es R1 y el vecino RPF hacia el origen es R2.

Verifiquemos el vecino RPF tanto para el origen como para el RP.

Aquí puede ver el vecino RPF hacia el RP: 192.168.0.100 es 192.168.3.1.

```
R3#show ip rpf 192.168.0.100
RPF information for ? (192.168.0.100)
  RPF interface: GigabitEthernet1
  RPF neighbor: ? (192.168.3.1)
  RPF route/mask: 192.168.0.100/32
  RPF type: unicast (ospf 1)
  Doing distance-preferred lookups across tables
  RPF topology: ipv4 multicast base, originated from ipv4 unicast base
```

Aquí puede ver el vecino RPF hacia el origen: 10.0.0.2 es 192.168.3.2.

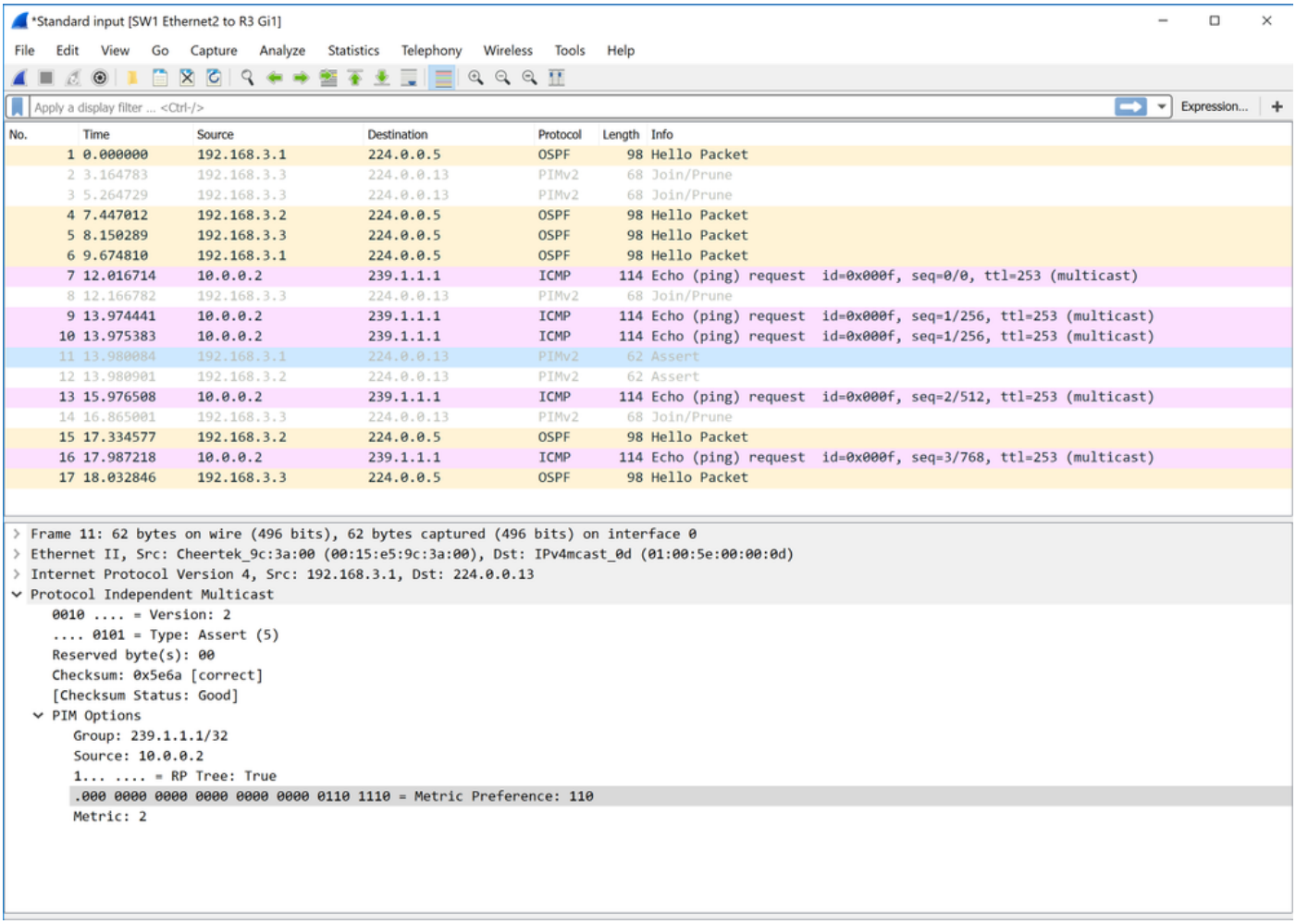
```
R3#show ip rpf 10.0.0.2
RPF information for ? (10.0.0.2)
  RPF interface: GigabitEthernet1
```

```
  RPF neighbor: ? (192.168.3.2)
  RPF route/mask: 10.0.0.0/24
  RPF type: unicast (ospf 1)
  Doing distance-preferred lookups across tables
  RPF topology: ipv4 multicast base, originated from ipv4 unicast base
```

Antes de activar el origen, echemos un vistazo a la tabla mroute en R3, ya que puede ver que ya hay (*,G) para el grupo 239.1.1.1. Esto se debe a que el receptor conectado a LHR ya ha solicitado el grupo especificado.

```
R3#show ip mroute
IP Multicast Routing Table
Outgoing interface flags: H - Hardware switched, A - Assert winner, p - PIM Join
 Timers: Uptime/Expires
 Interface state: Interface, Next-Hop or VCD, State/Mode

(*, 239.1.1.1), 00:00:57/00:02:32, RP 192.168.0.100, flags: S
  Incoming interface: GigabitEthernet1, RPF nbr 192.168.3.1
  Outgoing interface list:
    GigabitEthernet2, Forward/Sparse, 00:00:57/00:02:32

(*, 224.0.1.40), 00:11:24/00:02:41, RP 192.168.0.100, flags: SJCL
  Incoming interface: GigabitEthernet1, RPF nbr 192.168.3.1
  Outgoing interface list:
    GigabitEthernet2, Forward/Sparse, 00:02:02/00:02:41
```

Ahora, active el origen y capture paquetes en la interfaz R3 Gi1.



Como puede ver en esta captura de paquetes, PIM afirma que los paquetes ya están presentes.

Frame 11:

```
> Frame 11: 62 bytes on wire (496 bits), 62 bytes captured (496 bits) on interface 0
> Ethernet II, Src: Cheertek_9c:3a:00 (00:15:e5:9c:3a:00), Dst: IPv4mcast_0d (01:00:5e:00:00:0d)
> Internet Protocol Version 4, Src: 192.168.3.1, Dst: 224.0.0.13
v Protocol Independent Multicast
    0010 .... = Version: 2
    .... 0101 = Type: Assert (5)
    Reserved byte(s): 00
    Checksum: 0x5e6a [correct]
    [Checksum Status: Good]
  v PIM Options
      Group: 239.1.1.1/32
      Source: 10.0.0.2
      1... .... = RP Tree: True
      .000 0000 0000 0000 0000 0000 0110 1110 = Metric Preference: 110
      Metric: 2
```

Frame 12:

```
> Frame 12: 62 bytes on wire (496 bits), 62 bytes captured (496 bits) on interface 0
> Ethernet II, Src: Cheertek_8b:3e:00 (00:15:e5:8b:3e:00), Dst: IPv4mcast_0d (01:00:5e:00:00:0d)
> Internet Protocol Version 4, Src: 192.168.3.2, Dst: 224.0.0.13
v Protocol Independent Multicast
    0010 .... = Version: 2
    .... 0101 = Type: Assert (5)
    Reserved byte(s): 00
    Checksum: 0xde6a [correct]
    [Checksum Status: Good]
  v PIM Options
      Group: 239.1.1.1/32
      Source: 10.0.0.2
      0... .... = RP Tree: False
      .000 0000 0000 0000 0000 0000 0110 1110 = Metric Preference: 110
      Metric: 2
```

Cuando observa estos paquetes, debe poder determinar quién es el ganador seguro. Ahora echemos un vistazo a la selección del reenviador de afirmación PIM.

La preferencia de métrica es la distancia administrativa (AD). Esto se refiere a la distancia administrativa del protocolo de ruteo que instala la ruta en la tabla de ruteo, que se utiliza para buscar la dirección IP de origen y la métrica es el costo de la ruta.

También hay otros atributos que se utilizan para determinar quién es el ganador de afirmación. Puede ver estos detalles en RFC 7761.

Resumen de RFC 7761 Sección 4.6.3.

```
4.6.3.  Assert Metrics

   Assert metrics are defined as:

     struct assert_metric {
       rpt_bit_flag;
       metric_preference;
```

```
      route_metric;
      ip_address;
    };
  When comparing assert_metrics, the rpt_bit_flag, metric_preference,
  and route_metric fields are compared in order, where the first lower
  value wins.  If all fields are equal, the primary IP address of the
  router that sourced the Assert message is used as a tie-breaker, with
  the highest IP address winning.
```

Con el uso de estos campos definidos y la selección de rutas, puede determinar quién será el ganador de aserción en este escenario. Si vuelve a echar un vistazo a los paquetes asertivos, puede ver que la preferencia métrica no se compara ya que la decisión se toma en el primer criterio de selección que es rpt_bit_flag.

En este escenario, se compara la comparación de R1 y R2. Ambos routers envían mensajes asertivos que se vieron anteriormente y una vez que ambos dispositivos ven los mensajes asertivos de cada uno, pueden comparar métricas entre sí para determinar quién es el ganador.

Dado que R2 envía un mensaje de aserción con el árbol RP: False que tiene un valor de 0, es ciertamente menor que el R1 enviado con un árbol RP: True que tiene un valor de 1. El bit de árbol RP está configurado en 0 o 1.

El bit de árbol RP cuando se establece en 1 significa que actualmente está en el árbol compartido; el bit RPT borrado indica que el remitente de la aserción tenía el estado de reenvío (S,G) en una interfaz.

Como afirma (S,G), las afirmaciones tienen prioridad sobre (*,G), R2 debe ser el ganador seguro. Transición al estado "Soy el ganador de la declaración". Como se mencionó en la sentencia anterior en RFC 7761, se prefiere el valor más bajo.

Echemos un vistazo a R1 y R2 para ver quién es el ganador seguro.

```
R2#show ip mroute
IP Multicast Routing Table
Outgoing interface flags: H - Hardware switched, A - Assert winner, p - PIM Join
 Timers: Uptime/Expires
 Interface state: Interface, Next-Hop or VCD, State/Mode

(*, 239.1.1.1), 00:42:52/stopped, RP 192.168.0.100, flags: SP
  Incoming interface: GigabitEthernet2, RPF nbr 192.168.4.1
  Outgoing interface list: Null

(10.0.0.2, 239.1.1.1), 00:42:52/00:01:40, flags: T
  Incoming interface: GigabitEthernet2, RPF nbr 192.168.4.1
  Outgoing interface list:
    GigabitEthernet1, Forward/Sparse, 00:42:52/00:03:07, A

(*, 224.0.1.40), 00:43:23/00:02:25, RP 192.168.0.100, flags: SJPL
  Incoming interface: GigabitEthernet2, RPF nbr 192.168.4.1
  Outgoing interface list: Null
```

En este resultado, puede ver que el (S,G) en R2 tiene el indicador A configurado en el PETRÓLEO, lo que indica que es el ganador de la aserción. Aquí en R1, usted no tiene un PETRÓLEO en el (S,G) y el indicador P está configurado, lo que significa que el particular (S,G) ha sido recortado en este caso: no es el ganador seguro.

> **Nota:** Cuando assert está presente en un segmento compartido, los vecinos descendentes envían mensajes periódicos Join(*,G) y Join(S,G) al vecino RPF apropiado, es decir, al vecino RPF modificado por el proceso assert. No siempre se envían al vecino RPF como indica el MRIB.

```
R1#show ip mroute
IP Multicast Routing Table
Outgoing interface flags: H - Hardware switched, A - Assert winner, p - PIM Join
 Timers: Uptime/Expires
 Interface state: Interface, Next-Hop or VCD, State/Mode

(*, 239.1.1.1), 00:44:32/00:03:09, RP 192.168.0.100, flags: S
  Incoming interface: GigabitEthernet2, RPF nbr 192.168.5.2
  Outgoing interface list:
    GigabitEthernet1, Forward/Sparse, 00:44:32/00:03:09, A

(10.0.0.2, 239.1.1.1), 00:44:19/00:03:09, flags: PR
  Incoming interface: GigabitEthernet2, RPF nbr 192.168.5.2
  Outgoing interface list: Null

(*, 224.0.1.40), 00:44:50/00:02:53, RP 192.168.0.100, flags: SJCL
  Incoming interface: GigabitEthernet2, RPF nbr 192.168.5.2
  Outgoing interface list:
    GigabitEthernet1, Forward/Sparse, 00:43:56/00:02:53
```

Si es el caso de que R1 y R2 tengan el bit de árbol RP configurado en 1. luego puede considerar el router con el AD más bajo; si es igual, consulte la métrica. Si el bit de árbol RP es verdadero en ambos routers, la métrica se compara con la dirección IP RP. Si el bit de árbol RP es 0, la métrica se compara con el origen del flujo multicast.

Si todos estos valores son iguales, el mensaje de afirmación de la dirección IP de abastecimiento más alto es el ganador.

Summary

En la situación uno, no observó paquetes asertivos, sin embargo, según RFC, se deberían haber disparado. Como se mencionó, esto se debió a que R3 estaba recortando (*,G) antes de que se construyera el plano de control para (S,G).

Mientras que en el escenario dos, verá asertar paquetes. Cuando se recibió el primer paquete en LHR, enviaba una unión/separación (S,G) hacia R3 para extraer el origen/grupo. R3 entonces enviará un paquete de unión/separación hacia R2 para el mismo origen/grupo. Esto haría que R1 y R2 tuvieran OILs válidos rellenados. Ahora R3 sólo recorta (S,G) con el bit RP configurado cuando el indicador T se rellena en el estado R3s (S,G). Para que esto suceda, necesita recibir otro paquete de plano de datos del segmento compartido. Debido a que el plano de control ya se ha creado para (S,G), esto conduce a la duplicación en el segmento compartido que activa los mensajes de aserción.