下一代組播預設MDT:配置檔案0

目錄

簡介

本檔案介紹在下一代多點傳送中使用多重通訊協定標籤交換(MPLS)核心時，多點傳送封包如何通過。

背景資訊

預設MDT - PIM C — 組播信令

Draft Rosen使用通用路由封裝(GRE)作為重疊協定。這表示所有多點傳播封包都封裝在GRE中。虛擬LAN與VPN中的所有提供商邊緣(PE)路由器一起模擬加入組播組。這稱為預設組播分佈樹(MDT)。 預設MDT用於協定無關組播(PIM)Hello和其他PIM信令，也用於資料流量。如果源傳送大量流量，則使用預設MDT效率低下，並且可以建立資料MDT。資料MDT將僅包括具有使用中組的接收器的對等路由器。

Draft Rosen部署相當簡單，工作也很好，但是它也有一些缺點。我們來看一下以下內容：

額外開銷 — GRE為封包增加了24位元組的額外負荷。與通常增加8或12位元組的MPLS相比，每個資料包增加100%或更多的開銷。

核心中的PIM - Draft Rosen要求在核心中啟用PIM，因為PE必須加入通過PIM信令完成的預設和/或資料MDT。如果在核心中使用PIM ASM，則也需要RP。如果PIM SSM在核心上運行，則不需要任何RP。

核心狀態 — 由於來自PE的PIM信令，在核心中建立不必要的狀態。核心應該儘可能少地保持狀態。

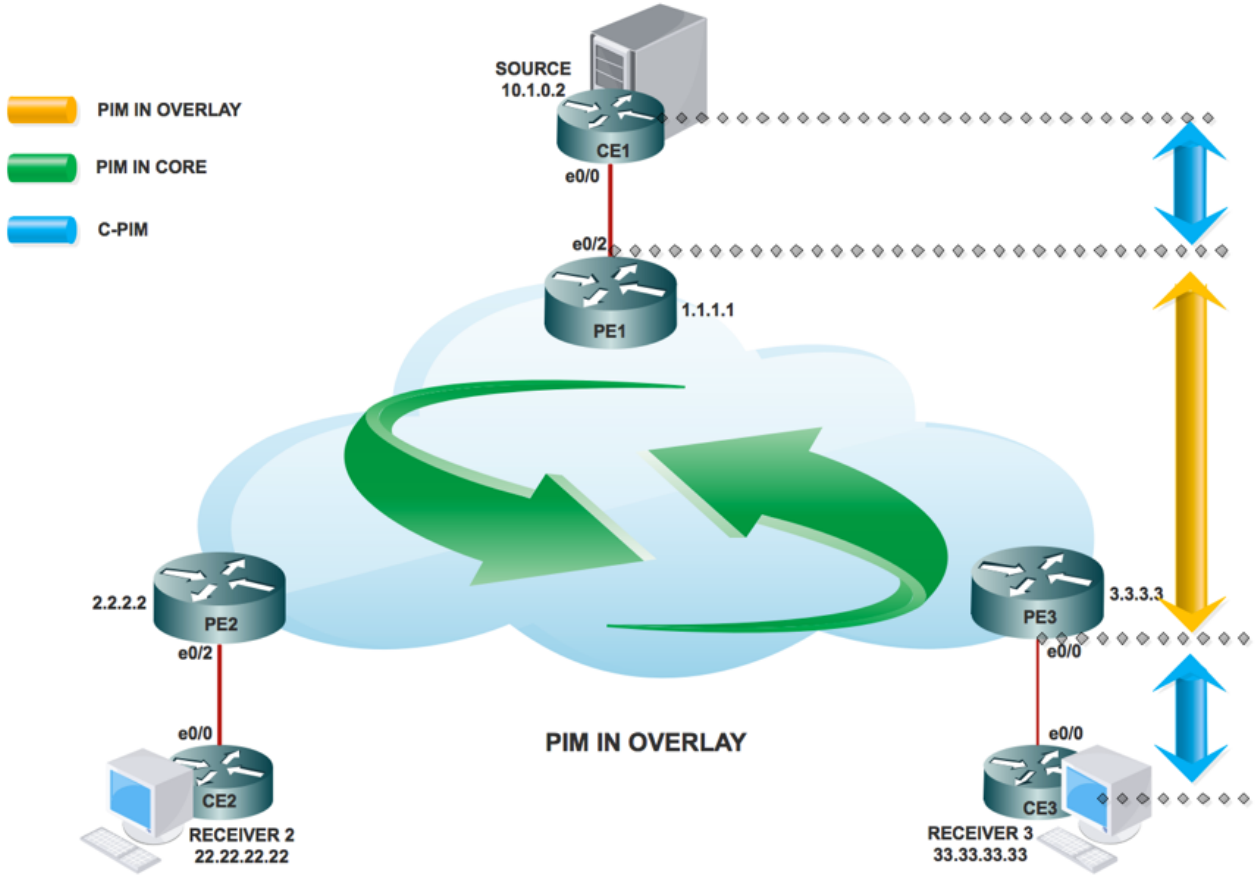
PIM鄰接 — PE將彼此成為PIM鄰居。如果是大型VPN和大量的PE，則會建立許多PIM鄰接關係。這會產生大量hello和其它信令，加重路由器的負擔。

單播與多播 — 單播轉發使用MPLS，組播使用GRE。這增加了複雜性，意味著單播使用的轉發機制與組播不同，後者不是最佳解決方案。

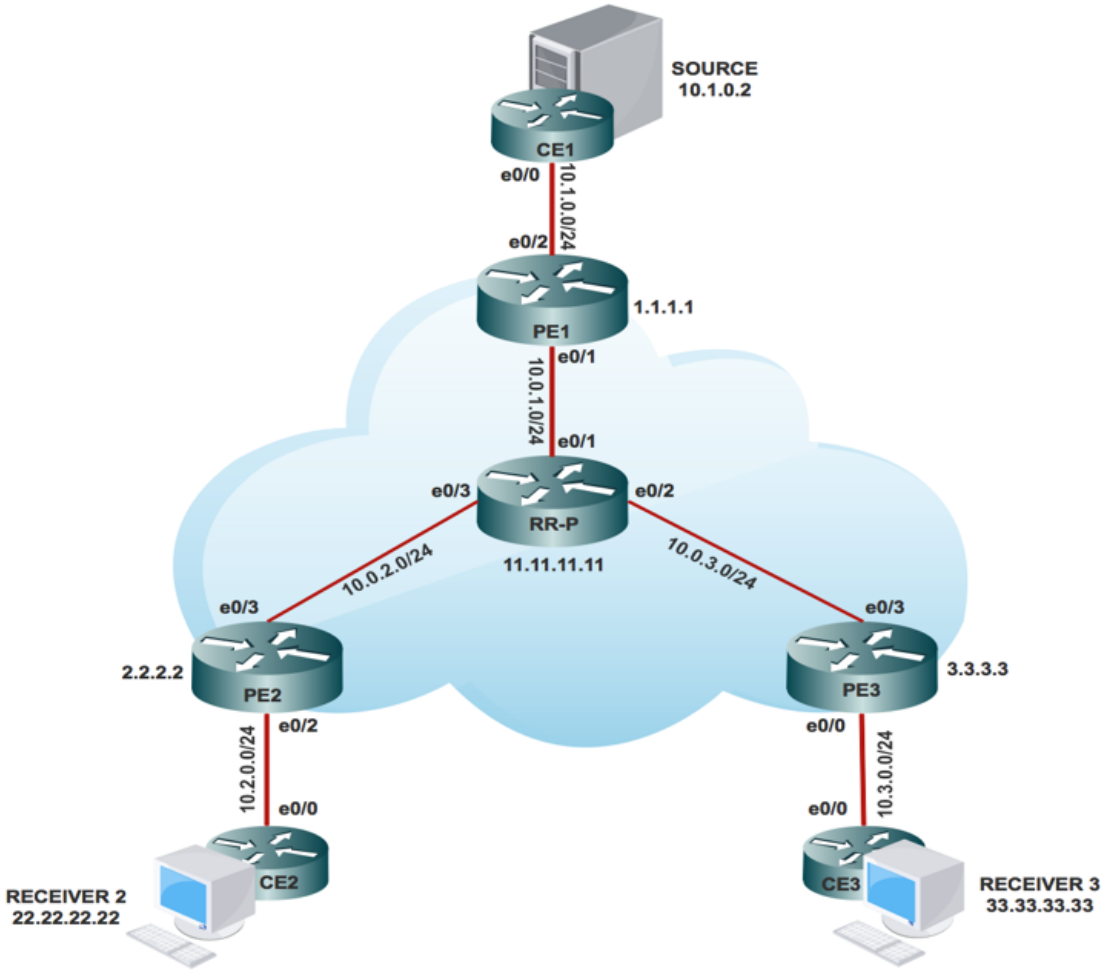
Inefficiency — 預設MDT將流量傳送到VPN中的所有PE，而不管該PE在(*,G)或(S，G)中有接收器用於正在使用的組。

- 將使用預設MDT將組播連線到一個VRF中的所有PE。
- 預設表示連線所有PE路由器。
- 預設情況下，它會傳送所有流量。
- 所有PIM控制流量和資料平面流量。例如(*,G)流量和(S，G)流量。
- 這表示多點到多點。
- 任何人都可以傳送，每個人都可以從樹接收。

作為重疊信令的PIM



拓撲



配置任務

1. 在所有節點上啟用組播路由。
2. 在所有介面中啟用PIM稀疏模式。
3. 使用現有VRF配置預設MDT。
4. 在介面Ethernet0/x上配置VRF。
5. 在VRF上啟用組播路由。
6. 在核心內部的所有節點中配置PIM SSM Default。
7. 在CE節點中配置BSR RP。
8. 預配置：

- VRF m-GRE
- mBGP:地址系列VPNv4
- VRF路由通訊協定

設定

1. 在所有節點上啟用組播路由。

```
(config)# ip multicast-routing
```

2.在所有介面中啟用PIM稀疏模式。

```
(config)# interface ethernet0/x

(config-if)#ip pim sparse-mode


(config)# interface loopback0

(config-if)#ip pim sparse-mode
```

3.對於已經存在的VRF，配置預設MDT。

```
(config)#ip vrf m-GRE

(config-vrf)# mdt default 232.1.1.1
```

4. 在介面Ethernet0/x上配置VRF。

在PE1、PE2和PE3上。

```
(config)# interface ethernet0/x

(config-if)# ip vrf forwarding m-GRE

(config-if)# ip address 10.x.0.1 255.255.255.0
```

5.在VRF上啟用組播路由。

在PE1、PE2和PE3上。

```
(config)# ip multicast-routing vrf m-GRE
```

6.為服務提供商核心配置RP。

在PE1、PE2、PE3和RR-P節點上。

```
(config)# ip pim rp-address 11.11.11.11
```

7.在CE節點（接收器）中配置BSR RP。

在Receiver2上。

```
(config)# ip pim bsr-candidate loopback0

(config)# ip pim rp-candidate loopback0
```

驗證

使用本節內容，確認您的組態是否正常運作。

任務1:檢驗物理連線

- 驗證所有連線的介面是否為「UP」

任務2:驗證地址系列VPNv4單播

- 驗證在所有路由器中為AF VPNv4單播啟用了BGP，且BGP鄰居為「UP」
- 確認BGP VPNv4單播表具有所有客戶字首。

任務3:驗證端到端組播流量。

- 檢查PIM鄰居關係。
- 檢驗是否端到端建立了組播狀態。
- 檢驗PE1、PE2和PE3上的mRIB條目
- 驗證(S，G)mFIB條目，資料包在軟體轉發中遞增。
- 檢驗ICMP資料包是否從CE到達CE。

Verify all the connected interface are "UP"

```
#sh ip interface brief
```

Task 2: Verify Address Family VPNv4 unicast

Address Family VPNv4 unicast and BGP neighbors

```
# show running-config | s r bgp
# show bgp vpnv4 unicast summary all
```

VPNv4 unicast table has all the Customer prefixes

```
PE1#sh bgp vpnv4 unicast all

BGP table version is 31, local router ID is 1.1.1.1

     Network          Next Hop            Metric LocPrf Weight Path
Route Distinguisher: 100:100 (default for vrf m-GRE)

 *>i 22.22.22.22/32   2.2.2.2                  0    100      0 20 i
 *>i 33.33.33.33/32   3.3.3.3                  0    100      0 30 i
 *>  111.111.111.111/32
                      10.1.0.2                 0             0 10 i
```

Check on all the PE nodes (PE1, PE2 and PE3)

Task 3: Verify Multicast Traffic end to end

Verify that multicast state is created end to end

```
#sh ip mroute vrf m-GRE 230.1.1.1 verbose
IP Multicast Routing Table
Flags: T - SPT-bit set,  p - PIM Joins on route

(10.1.0.2, 230.1.1.1), 00:00:35/00:02:24, flags: Tp
  Incoming interface: Ethernet0/2, RPF nbr 10.1.0.2
  Outgoing interface list:
    Tunnel2, GRE MDT: 232.1.1.1 (default), Forward/Sparse,
                              00:00:35/00:02:54, p
```

Check on all the PE nodes (PE1, PE2 and PE3)

Verify that (S,G) mFIB entry, packet getting incremented

```
# sh ip mfib vrf m-GRE 230.1.1.1 verbose
Entry Flags:    NS - Negate Signalling, SP - Signal Present,
                A - Accept, F - Forward, RA - MRIB Accept, RF - MRIB Forward,
                MA - MFIB Accept
Forwarding Counts: Pkt Count/Pkts per second/Avg Pkt Size/Kbits per second
Other counts:      Total/RPF failed/Other drops
I/O Item Counts:   FS Pkt Count/PS Pkt Count
VRF m-GRE

 (10.1.0.2,230.1.1.1) Flags: K DDE
   SW Forwarding: 5/0/100/0, Other: 0/0/0
   Ethernet0/2 Flags: RA A MA
   Tunnel2, MDT/232.1.1.1 Flags: RF F NS
     CEF: Adjacency with MAC: 45000000000000000FF2FD0CA01010101E801010100000800
     Pkts: 4/1
```

Check PIM Neighborship at the core and vrf

```
PE1#sh ip pim neighbor
PIM Neighbor Table
Mode: B - Bidir Capable, DR - Designated Router, N - Default DR Priority,
      P - Proxy Capable, S - State Refresh Capable, G - GenID Capable
Neighbor          Interface                Uptime/Expires    Ver   DR
Address                                                            Prio/Mode
10.0.1.2          Ethernet0/1              01:34:51/00:01:25 v2    1 / DR S P G


PE1#sh ip pim vrf m-GRE neighbor
PIM Neighbor Table
Mode: B - Bidir Capable, DR - Designated Router, N - Default DR Priority,
      P - Proxy Capable, S - State Refresh Capable, G - GenID Capable
Neighbor          Interface                Uptime/Expires    Ver   DR
Address                                                            Prio/Mode
10.1.0.2          Ethernet0/2              01:34:32/00:01:42 v2    1 / DR S P G
3.3.3.3           Tunnel2                  01:32:32/00:01:41 v2    1 / S P G
2.2.2.2           Tunnel2                  01:32:32/00:01:36 v2    1 / S P G
```

Multicast Forwarding Packets

```
#sh ip mfib vrf m-GRE 230.1.1.1 count
Forwarding Counts: Pkt Count/Pkts per second/Avg Pkt Size/
Kilobits per second
Other counts:       Total/RPF failed/Other drops(OIF-null,
rate-limit etc)
VRF m-GRE
 11 routes, 7 (*,G)s, 2 (*,G/m)s
Group: 230.1.1.1
  RP-tree,
   SW Forwarding: 0/0/0/0, Other: 0/0/0
  Source: 10.1.0.2,
   SW Forwarding: 51/0/100/0, Other: 0/0/0
  Totals - Source count: 2, Packet count: 102

Groups: 1, 2.00 average sources per group
```

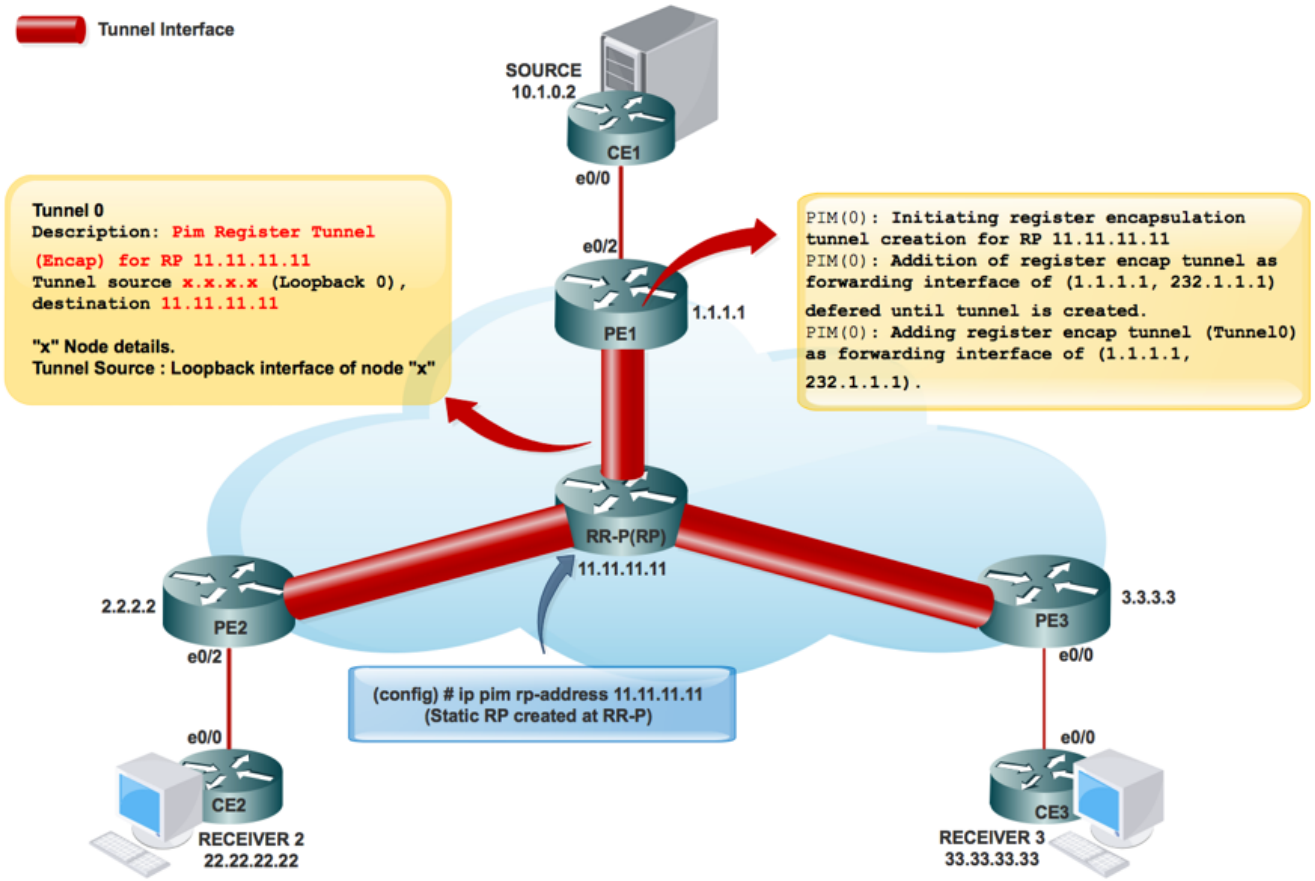
Verify ICMP packets getting reach from CE to CE

```
SOURCE1#ping  230.1.1.1 repeat 100 timeout 0

Type escape sequence to abort.
Sending 100, 100-byte ICMP Echos to 230.1.1.1, timeout is 0
seconds:

Reply to request 0 from 10.3.0.2, 4 ms
Reply to request 1 from 10.3.0.2, 4 ms
Reply to request 2 from 10.3.0.2, 4 ms
Reply to request 3 from 10.3.0.2, 4 ms
```

建立隧道介面時：

建立服務提供商RP:

RP資訊在核心中泛洪後。已建立介面隧道0。



PIM(0):正在為RP 11.11.11.11啟動暫存器封裝隧道建立。

PIM(0):RP 11.11.11.11的初始註冊隧道建立成功。

PIM(0):新增暫存器封裝通道作為(1.1.1.1、232.1.1.1)的轉送介面，此新增功能會延遲到建立通道為止。

5月9日17:34:56.155:PIM(0):將RP 11.11.11.11簽入(, 232.1.1.1)。

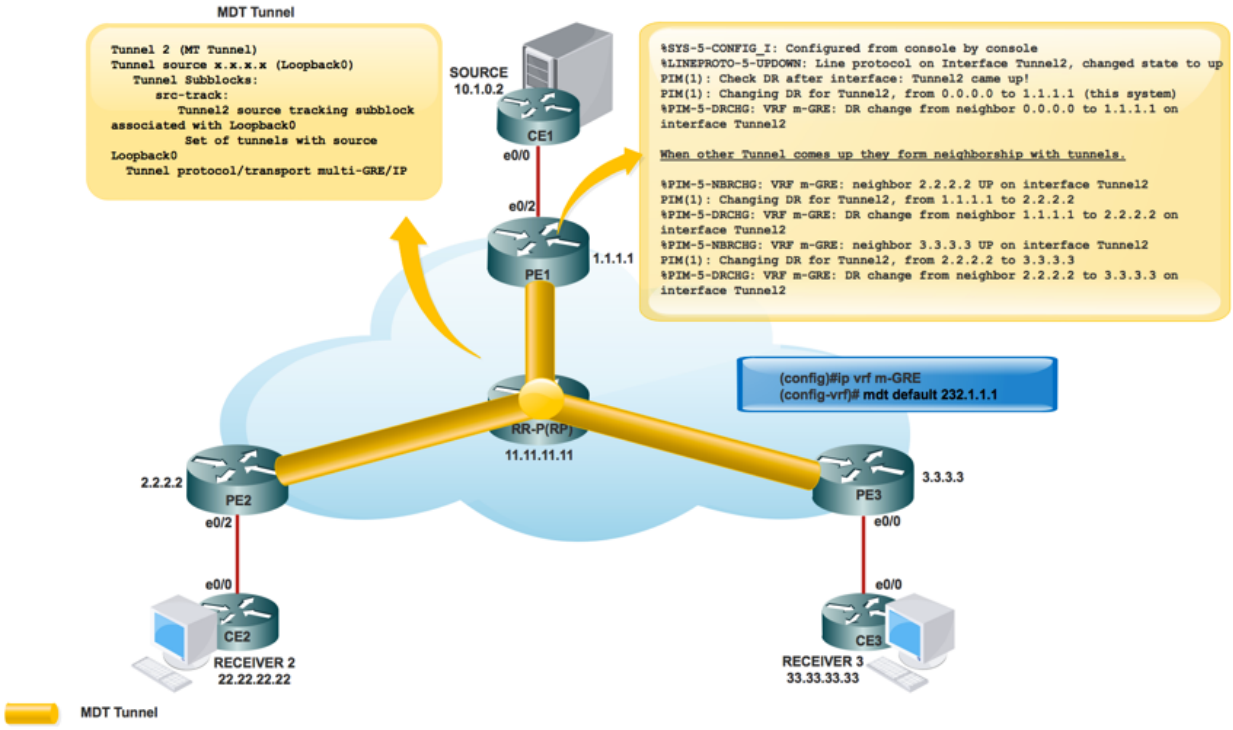
PIM(0):新增暫存器封裝隧道(Tunnel0)作為(1.1.1.1、232.1.1.1)的轉發介面。

```
PE1#sh int tunnel 0

Tunnel0 is up, line protocol is up

  Hardware is Tunnel

  Description: Pim Register Tunnel (Encap) for RP 11.11.11.11

  Interface is unnumbered. Using address of Ethernet0/1 (10.0.1.1)

  MTU 17912 bytes, BW 100 Kbit/sec, DLY 50000 usec,

     reliability 255/255, txload 1/255, rxload 1/255

  Encapsulation TUNNEL, loopback not set

  Keepalive not set
```

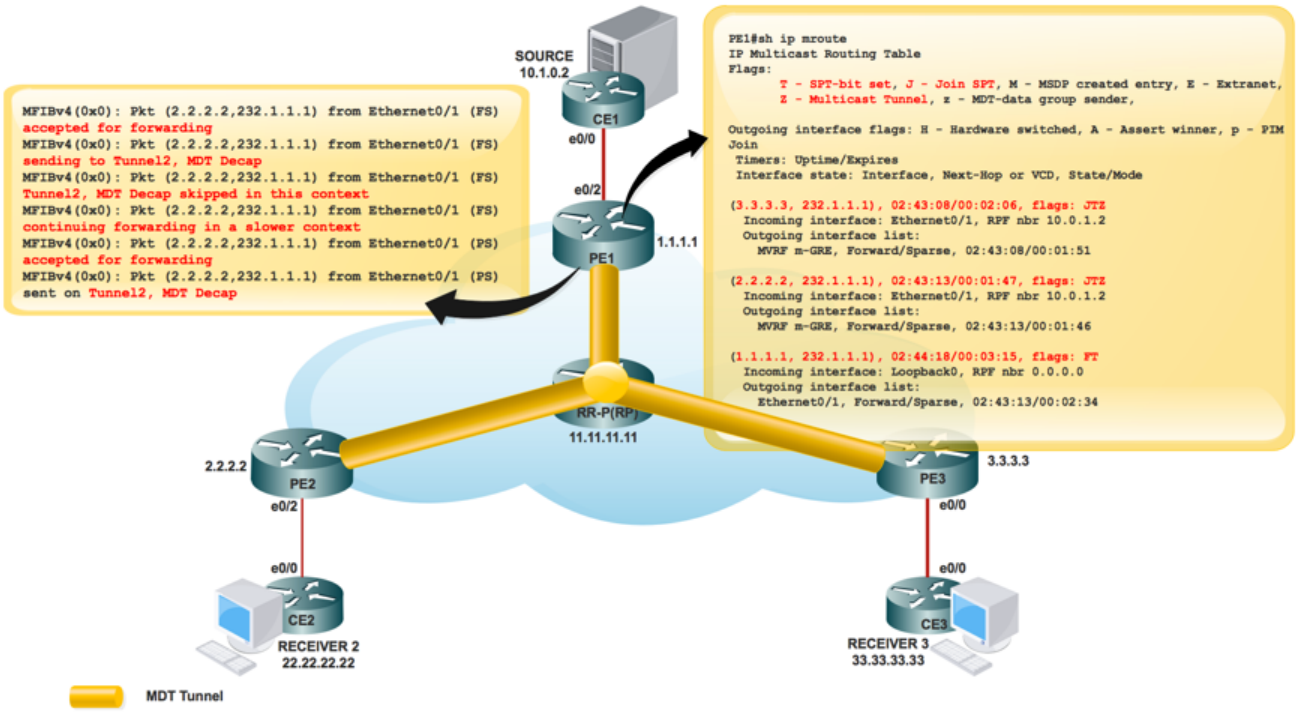
Tunnel source 10.0.1.1 (Ethernet0/1), destination 11.11.11.11 >>>>>>>>>> Tunnel Source and destination

```
   Tunnel Subblocks:

      src-track:

         Tunnel0 source tracking subblock associated with Ethernet0/1

          Set of tunnels with source Ethernet0/1, 1 member (includes iterators), on interface
<OK>

  Tunnel protocol/transport PIM/IPv4

  Tunnel TOS/Traffic Class 0xC0,  Tunnel TTL 255

  Tunnel transport MTU 1472 bytes
```

MDT隧道建立：



核心中的MRIB建立：



PE1#sh ip mroute

IP Multicast Routing Table

Flags: D - Dense, S - Sparse, B - Bidir Group, s - SSM Group, C - Connected,

L - Local, P - Pruned, R - RP-bit set, F - Register flag,

T - SPT-bit set, J - Join SPT, M - MSDP created entry, E - Extranet,

X - Proxy Join Timer Running, A - Candidate for MSDP Advertisement,

U - URD, I - Received Source Specific Host Report,

Z - Multicast Tunnel, z - MDT-data group sender,

(3.3.3.3, 232.1.1.1), 00:10:13/00:01:01, flags: **JTZ**

Incoming interface: Ethernet0/1, RPF nbr 10.0.1.2

Outgoing interface list:

MVRF m-GRE, Forward/Sparse, 00:10:13/00:01:46

(2.2.2.2, 232.1.1.1), 00:10:14/00:00:57, flags: **JTZ**

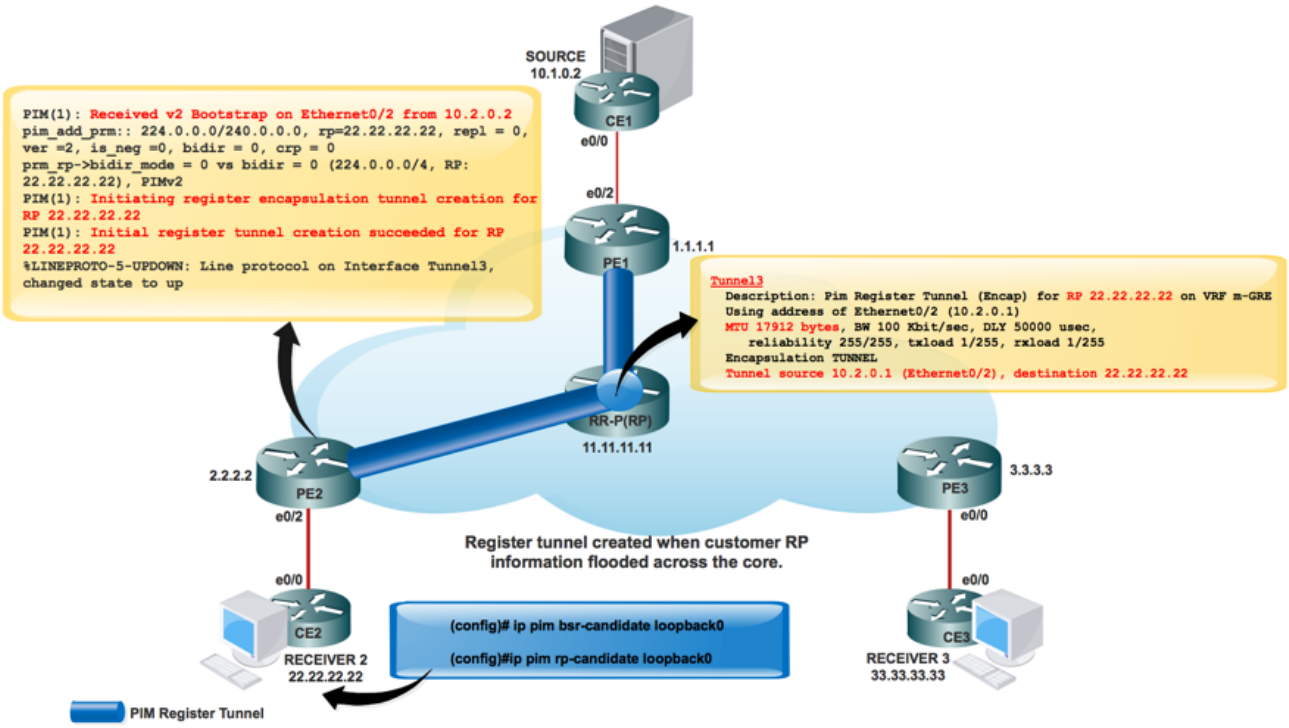
Incoming interface: Ethernet0/1, RPF nbr 10.0.1.2

Outgoing interface list:

MVRF m-GRE, Forward/Sparse, 00:10:14/00:01:45

```
(1.1.1.1, 232.1.1.1), 00:10:15/00:03:20, flags: FT

  Incoming interface: Loopback0, RPF nbr 0.0.0.0

  Outgoing interface list:

    Ethernet0/1, Forward/Sparse, 00:10:15/00:03:04
```

為客戶網路建立RP之後：




```
*May  9 18:54:42.170:  prm_rp->bidir_mode = 0 vs bidir = 0 (224.0.0.0/4, RP:33.33.33.33), PIMv2

*May  9 18:54:42.170: PIM(1): Initiating register encapsulation tunnel creation for RP
33.33.33.33

*May  9 18:54:42.170: PIM(1): Initial register tunnel creation succeeded for RP 33.33.33.33

*May  9 18:54:43.173: %LINEPROTO-5-UPDOWN: Line protocol on Interface Tunnel2, changed state to
up
```

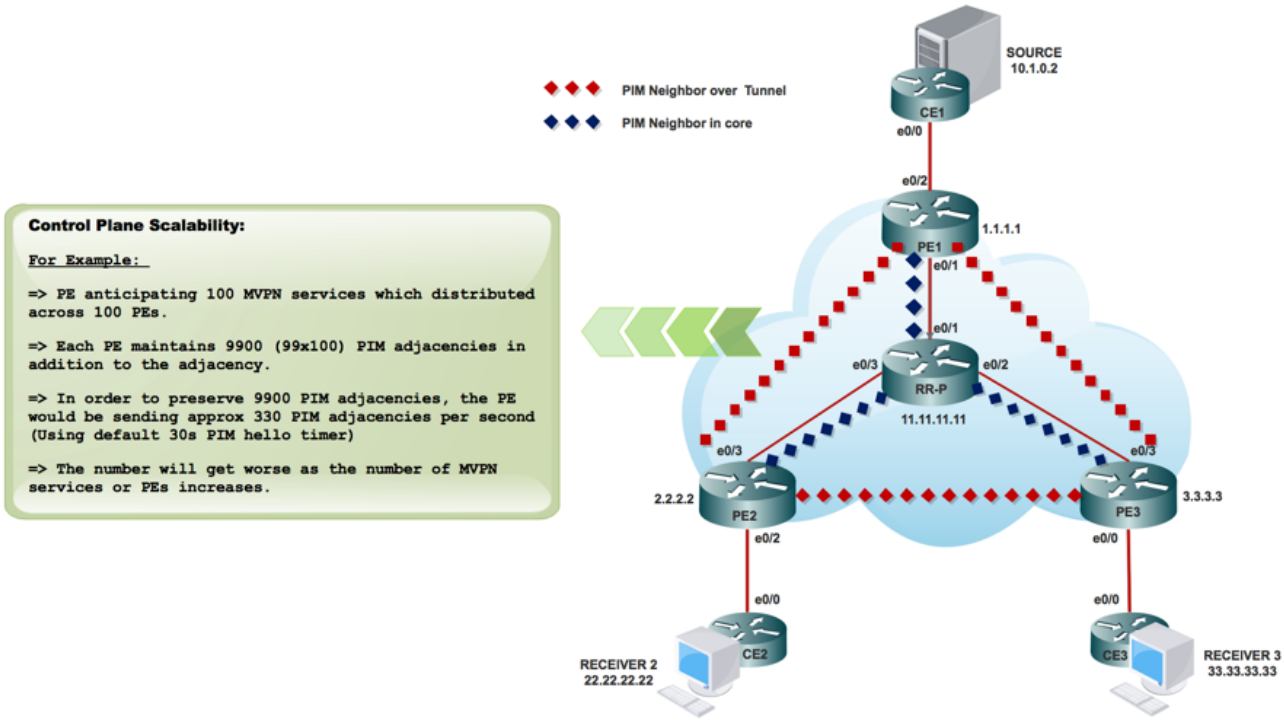
建立通道介面是為了傳輸客戶RP資訊。

PIM(1):正在為RP 22.22.22.22啟動暫存器封裝隧道建立。

它是為向RP註冊封裝而建立的隧道。

對於每個發現的稀疏模式RP，會建立一個註冊封裝隧道。在稀疏模式RP本身上，會建立一個解封裝隧道介面來接收註冊資料包。

PIM鄰居關係：




```
PE1#sh ip pim interface

Address          Interface                Ver/   Nbr    Query  DR     DR
                                          Mode   Count  Intvl  Prior
1.1.1.1          Loopback0                v2/S   0      30     1      1.1.1.1
10.0.1.1         Ethernet0/1              v2/S   1      30     1      10.0.1.2


PE1#sh ip pim vrf  m-GRE neighbor
PIM Neighbor Table
Mode: B - Bidir Capable, DR - Designated Router, N - Default DR Priority,
      P - Proxy Capable, S - State Refresh Capable, G - GenID Capable
Neighbor          Interface                Uptime/Expires    Ver   DR
Address                                                            Prio/Mode
10.1.0.2          Ethernet0/2              03:08:34/00:01:43 v2    1 / DR S P G
3.3.3.3           Tunnel1                  01:44:24/00:01:41 v2    1 / DR S P G
2.2.2.2           Tunnel1                  01:44:24/00:01:38 v2    1 / S P G
```

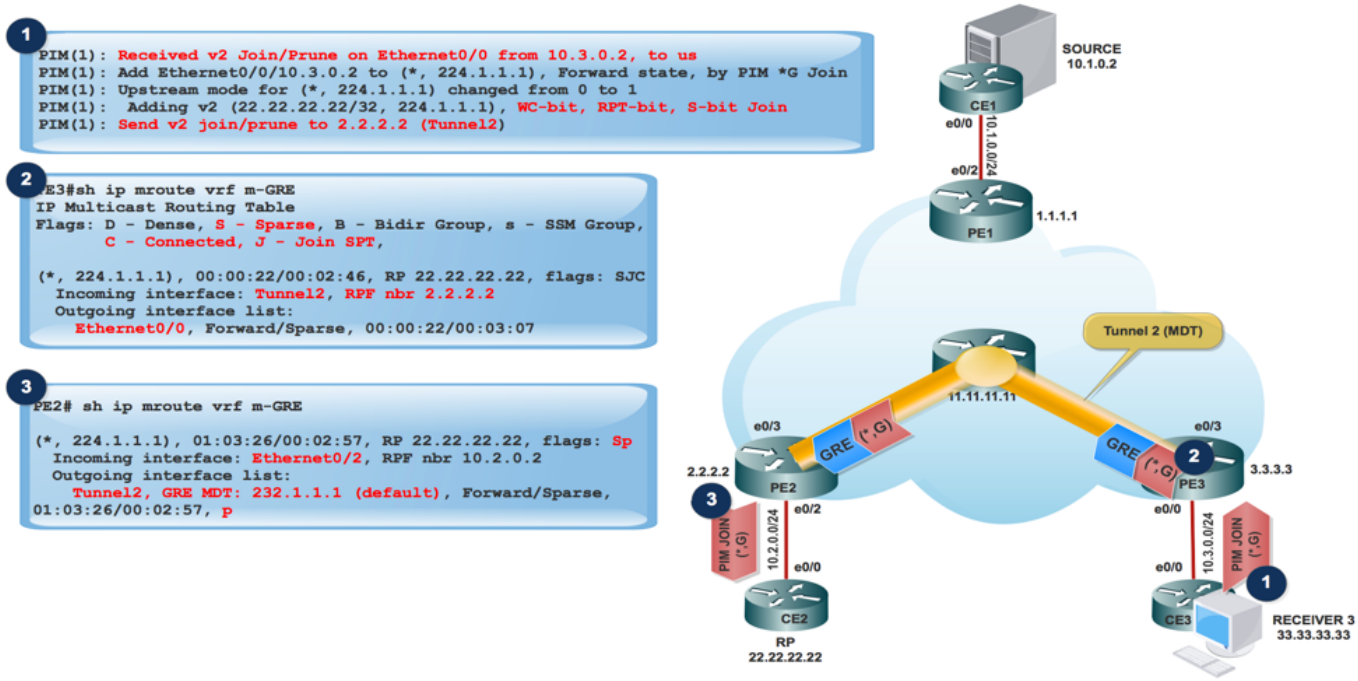
封包流量:

控制平面資料包流分為兩部分。

1. 接收者已聯機。

2. 源處於活動狀態。

當接收器處於活動狀態時：



1. 接收方聯機，向PE3傳送PIM加入(*,G)。

2. PE3將PIM加入(*,G)封裝在GRE資料包中，並通過隧道2（MDT隧道）傳送，該隧道從**show ip mroute vrf m-GRE的傳入介面驗證。**

```
    42 26.584402     3.3.3.3               224.0.0.13            PIMv2      92 Join/Prune
▶ Frame 42: 92 bytes on wire (736 bits), 92 bytes captured (736 bits) on interface 0
▶ Ethernet II, Src: aa:bb:cc:00:50:30 (aa:bb:cc:00:50:30), Dst: IPv4mcast_01:01:01 (01:00:5e:01:01:01)
▶ Internet Protocol Version 4, Src: 3.3.3.3, Dst: 232.1.1.1
▶ Generic Routing Encapsulation (IP)
▶ Internet Protocol Version 4, Src: 3.3.3.3, Dst: 224.0.0.13
▼ Protocol Independent Multicast
    0010 .... = Version: 2
    .... 0011 = Type: Join/Prune (3)
    Reserved byte(s): 00
    Checksum: 0xc0b8 [correct]
  ▼ PIM Options
      Upstream-neighbor: 2.2.2.2
      Reserved byte(s): 00
      Num Groups: 1
      Holdtime: 210
    ▼ Group 0: 224.1.1.1/32
      ▼ Num Joins: 1
          IP address: 22.22.22.22/32 (SWR)
        Num Prunes: 0
```

```
PE3#sh ip mroute

IP Multicast Routing Table

Flags: D - Dense, S - Sparse, B - Bidir Group, s - SSM Group, C - Connected,

       L - Local, P - Pruned, R - RP-bit set, F - Register flag,

       T - SPT-bit set, J - Join SPT, M - MSDP created entry, E – Extranet
```

```
(3.3.3.3, 232.1.1.1), 10:20:04/00:02:56, flags: FT

 Incoming interface: Loopback0, RPF nbr 0.0.0.0

 Outgoing interface list:

   Ethernet0/3, Forward/Sparse, 10:20:04/00:02:40
```

1. PE2收到源地址為3.3.3.3且目的地址為232.1.1.1的GRE資料包，然後根據OIL將其轉發到MVRF m-GRE。

```
PE2#sh ip mroute

IP Multicast Routing Table

Flags:

      T - SPT-bit set, J - Join SPT, M - MSDP created entry, E - Extranet,

      Z - Multicast Tunnel, z - MDT-data group sender,


(3.3.3.3, 232.1.1.1), 11:47:30/00:01:01, flags: JTZ

 Incoming interface: Ethernet0/3, RPF nbr 10.0.2.2

 Outgoing interface list:

   MVRF m-GRE, Forward/Sparse, 11:47:30/00:00:29
```

將GRE封包解除封裝，並向RP傳送PIM加入。

附註：RPF Neighbor是2.2.2.2，因為PIM加入的目的地是RP地址，通過核心形成RPT。

附註：WC位和RPT位：由(*,G)狀態觸發，DR會建立連線/修整消息，其連線清單中的RP地址以及萬用字元位（WC位）和RP樹位（RPT位）設定為1。WC位表示任何源可以匹配，如果不再匹配，則根據此條目轉發；rpt-bit表示此連線通過共用RP樹向上傳送。修剪清單為空。當RPT位設定為1時，它表示連線與共用RP樹相關聯，因此會沿RP樹傳播連線/修剪消息。當WC位元設定為1時，表示位址是RP，且下游接收者預期會透過此（共用樹狀目錄）路徑接收來自所有來源的封包。

```
PE2#sh ip mroute verbose

IP Multicast Routing Table

Flags: D - Dense, S - Sparse, B - Bidir Group, s - SSM Group, C - Connected,

 L - Local, P - Pruned, R - RP-bit set, F - Register flag,

 T - SPT-bit set, J - Join SPT, M - MSDP created entry, E -

 V - RD & Vector, v - Vector, p - PIM Joins on route
```

```
(2.2.2.2, 232.1.1.1), 22:48:12/00:02:04, flags: FTp

Incoming interface: Loopback0, RPF nbr 0.0.0.0

Outgoing interface list:Ethernet0/3, Forward/Sparse, 22:48:12/00:03:12, p
```

1. GRE封裝的資料包到達源PE PE1。

```
PE1#sh ip mroute verbose

IP Multicast Routing Table

Flags: D - Dense, S - Sparse, B - Bidir Group, s - SSM Group, C - Connected,

      L - Local, P - Pruned, R - RP-bit set, F - Register flag,

      T - SPT-bit set, J - Join SPT, M - MSDP created entry, E - Extranet,

      X - Proxy Join Timer Running, A - Candidate for MSDP Advertisement,

      U - URD, I - Received Source Specific Host Report,

      Z - Multicast Tunnel, z - MDT-data group sender,


 (2.2.2.2, 232.1.1.1), 22:55:50/00:02:45, flags: JTZ

 Incoming interface: Ethernet0/1, RPF nbr 10.0.1.2

 Outgoing interface list:MVRF m-GRE, Forward/Sparse, 22:55:50/00:01:09


PIM(1): Received v2 Join/Prune on Tunnel2 from 2.2.2.2, to us

PIM(1): Join-list: (10.1.0.2/32, 224.1.1.1), S-bit set
```

2. PIM JOIN(S，G)到達源CE。

3. 現在，源裝置獲得了感興趣的接收裝置的資訊，流量開始傳送到源PE PE1。

4. 在源PE PE1:

```
PIM(1): Add Tunnel2/2.2.2.2 to (10.1.0.2, 224.1.1.1), Forward state, by PIM SG Join


MFIBv4(0x1): Pkt (10.1.0.2,224.1.1.1) from Ethernet0/2 (PS) accepted for forwarding

MFIBv4(0x1): Pkt (10.1.0.2,224.1.1.1) from Ethernet0/2 (PS) sending to Tunnel2, MDT/232.1.1.1

MFIBv4(0x1): Pkt (10.1.0.2,224.1.1.1) from Ethernet0/2 (PS) sent on Tunnel2, MDT/232.1.1.1
```

在PE2(RP PE):

```
PIM(1): Prune-list: (10.1.0.2/32, 224.1.1.1) RPT-bit set

PIM(1): Cancel sending Join for (10.1.0.2/32, 224.1.1.1) on Tunnel2
```

```
PE2#sh ip mroute vrf m-GRE
IP Multicast Routing Table
Flags: L - Local, P - Pruned, R - RP-bit set, F - Register flag,

(10.1.0.2, 224.1.1.1), 00:03:52/00:01:29, flags: R
 Incoming interface: Ethernet0/2, RPF nbr 10.2.0.2
 Outgoing interface list:
   Tunnel2, Forward/Sparse, 00:00:52/00:02:58
```

PCAP捕獲來自PE1的組播資料包。在MDT預設隧道中隧道。採用GRE封裝。

5. 在接收器PE PE3，收到資料包。

```
PE3#sh ip mroute verbose
IP Multicast Routing Table
Flags: D - Dense, S - Sparse, B - Bidir Group, s - SSM Group, C - Connected,
       L - Local, P - Pruned, R - RP-bit set, F - Register flag,
       T - SPT-bit set, J - Join SPT, M - MSDP created entry, E - Extranet,
       Z - Multicast Tunnel, z - MDT-data group sender,

(1.1.1.1, 232.1.1.1), 23:12:51/00:02:50, flags: JTZ
 Incoming interface: Ethernet0/3, RPF nbr 10.0.3.2
 Outgoing interface list:
   MVRF m-GRE, Forward/Sparse, 23:12:51/stopped
 PIM(1):  Building Join/Prune packet for nbr 2.2.2.2
 PIM(1):  Adding v2 (10.1.0.2/32, 224.1.1.1), RPT-bit, S-bit Prune
 PIM(1):  Send v2 join/prune to 2.2.2.2 (Tunnel2)
 PIM(1):  Building Join/Prune packet for nbr 1.1.1.1
 MFIBv4(0x1): Pkt (10.1.0.2,224.1.1.1) from Tunnel2, MDT/232.1.1.1 (PS) accepted for forwarding
 MFIBv4(0x1): Pkt (10.1.0.2,224.1.1.1) from Tunnel2, MDT/232.1.1.1 (PS) sent on Ethernet0/0
 MFIBv4(0x1): Pkt (10.1.0.2,224.1.1.1) from Tunnel2, MDT/232.1.1.1 (PS) accepted for forwarding
 MFIBv4(0x1): Pkt (10.1.0.2,224.1.1.1) from Tunnel2, MDT/232.1.1.1 (PS) sent on Ethernet0/0
*Jun  2 20:09:11.817: PIM(1): Received v2 Join/Prune on Ethernet0/0 from 10.3.0.2, to us
```

```
PE3#sh ip mroute  vrf m-GRE  verbose

IP Multicast Routing Table

Flags: D - Dense, S - Sparse, B - Bidir Group, s - SSM Group, C - Connected,

      L - Local, P - Pruned, R - RP-bit set, F - Register flag,

      T - SPT-bit set, J - Join SPT, M - MSDP created entry, E - Extranet,

      V - RD & Vector, v - Vector, p - PIM Joins on route


(10.1.0.2, 224.1.1.1), 00:00:07/00:02:52, flags: Tp

 Incoming interface: Tunnel2, RPF nbr 1.1.1.1

 Outgoing interface list:

   Ethernet0/0, Forward/Sparse, 00:00:07/00:03:22, p


RPF Change at PE3 (Receiver PE)


MRT(1): (10.1.0.2,224.1.1.1), RPF change from  /2.2.2.2 to Tunnel1/1.1.1.1

MRT(1): Create (10.1.0.2 ,224.1.1.1), RPF (Tunnel2, 1.1.1.1, 200/0)

MRT(1): Set the T-flag for (10.1.0.2, 224.1.1.1)

MRT(1): WAVL Insert interface: Tunnel1 in (10.1.0.2,224.1.1.1) Successful

MRT(1): set min mtu for (10.1.0.2, 224.1.1.1) 18010->1500
```

注意：收到來自PE1的組播資料包後，RPF鄰居會更改。更早的版本是PE2作為RP託管。收到第一個組播資料包後，它會更改RPF並設定SPT位。

通過預設MDT隧道的流量：

- MDT上的轉發使用GRE，C資料包將成為P資料包。
- P-Packet S address = PE的BGP對等地址
  G address = MDT-Group address（預設或資料）
- C-Packet IP TOS將複製到P-Packet。
- MPLS標籤不在核心中使用，僅用於本地組播。

封包流量:

1. C資料包到達配置了VRF的PE介面，mVRF被隱式識別。對C源執行常規RPF檢查。

C-Packet從OIL中的介面複製出去。此時，這是同一VRF中的PE介面。

```
PE1#sh ip mroute vrf  m-GRE verbose

IP Multicast Routing Table

Flags: D - Dense, S - Sparse, B - Bidir Group, s - SSM Group, C - Connected,
```

```
      L - Local, P - Pruned, R - RP-bit set, F - Register flag,

      T - SPT-bit set, V - Vector, p - PIM Joins on route

Outgoing interface flags: H - Hardware switched, A - Assert winner, p - PIM Join

 Timers: Uptime/Expires

 Interface state: Interface, Next-Hop or VCD, State/Mode




(10.1.0.2, 224.1.1.1), 00:00:03/00:02:56, flags: Tp

 Incoming interface: Ethernet0/2, RPF nbr 10.1.0.2

 Outgoing interface list:

   Tunnel2, GRE MDT: 232.1.1.1 (default), Forward/Sparse, 00:00:03/00:03:26, p (Small "p" indicates downstream PIM join)
```

如果OIL包含MTI，則C資料包封裝成P資料包。如果在使用的條目目標上設定「y」標誌，則為DATA-MDT組，否則為Default MDT組。源是PE BGP對等體地址，目標是MDT組地址。

2. P資料包按照正常組播通過P網路轉發。

資料包到達全域性介面。所引用的MDT組的全域性(S，G)或(*,G)條目。對P源（PE對等點）執行常規RPF檢查。

3. 將P資料包從OIL中的介面複製。此時，這是全域性mroute表中的P/PE。

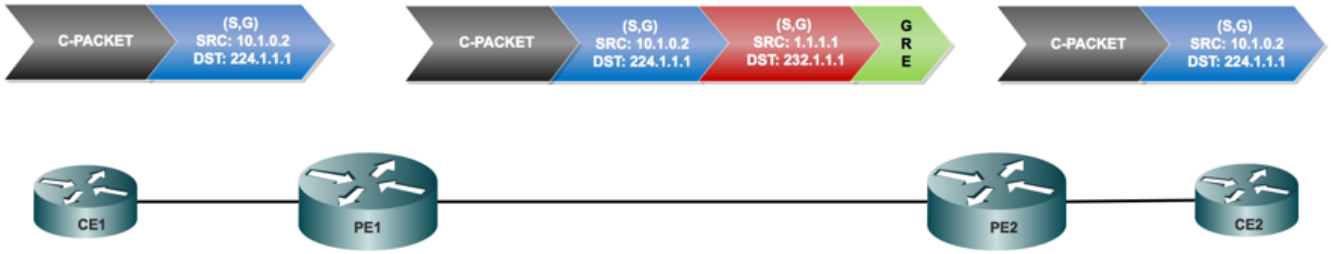
4. 如果「Z」標誌設定，資料包將被解除封裝以顯示C資料包。從MDT組派生的目標mVRF和傳入介面是封裝報頭的目標。

mVRF中的C-Packet的RPF檢查完成，mVRF中的C Packet複製出OIL。

```
PE3#sh ip mroute verbose

IP Multicast Routing Table

Flags: D - Dense, S - Sparse, B - Bidir Group, s - SSM Group, C - Connected,

      L - Local, P - Pruned, R - RP-bit set, F - Register flag,

      T - SPT-bit set, J - Join SPT, M - MSDP created entry, E - Extranet,

      Z - Multicast Tunnel, z - MDT-data group sender,



(1.1.1.1, 232.1.1.1), 1d01h/00:02:47, flags: JTZ

 Incoming interface: Ethernet0/3, RPF nbr 10.0.3.2

 Outgoing interface list: MVRF m-GRE, Forward/Sparse, 1d01h/stopped
```

5. 本機C資料包在接收器3到達。

封包封裝:

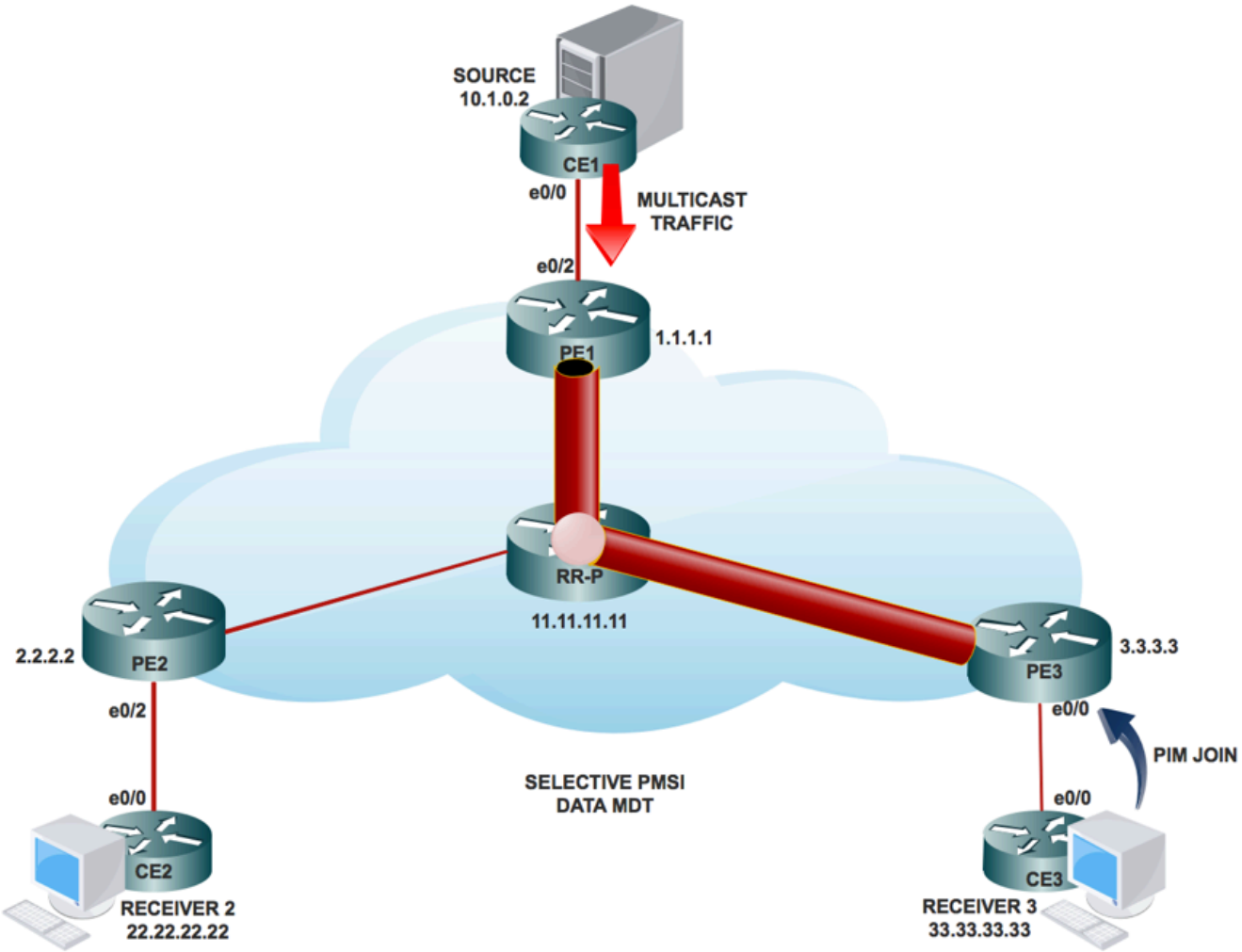


資料MDT:

什麼是Data MDT?

它是可選的。它是按需建立的，它承載特定(S，G)流量。在最新版本的IOS®中，設定的閾值為「0」和「infinite」。每當第一個資料包到達VRF時，資料MDT就會初始化，如果無窮大，則永遠不會建立資料MDT，流量將以預設MDT向前移動。Data MDT始終是接收樹，它們從不傳送任何流量。資料MDT僅用於(S，G)流量。

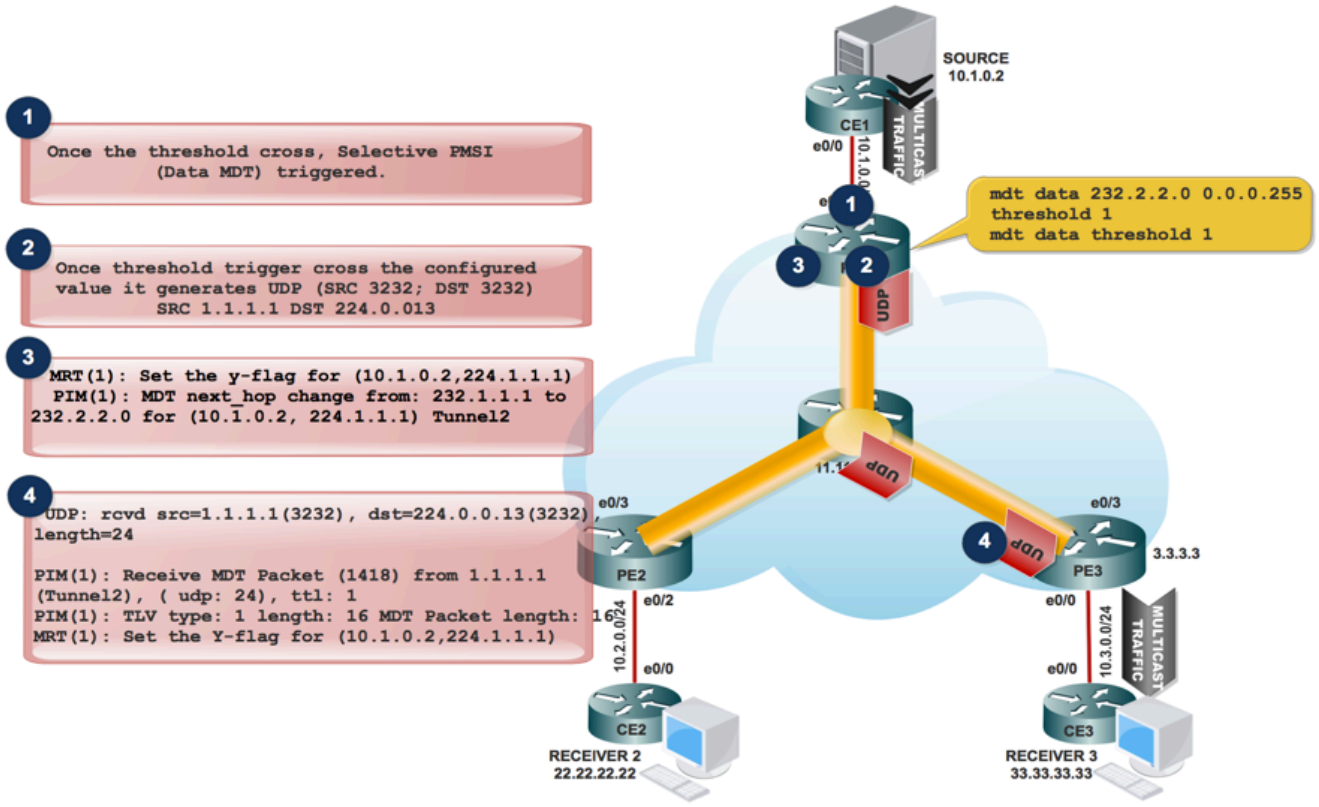
選擇性PMSI:



- 它是可選的。它是按需建立的，可承載特定(S、G)流量。
- 每當第一個資料包到達VRF時，資料MDT就會初始化，如果無窮大，則從不建立資料MDT，流量將以預設MDT向前移動。

- Data MDT始終是接收樹，它們從不傳送任何流量。資料MDT僅用於(S， G)流量。
- PIM消息攜帶C-(S， G)和P-Group。

如何建立DATA MDT:



1. 當組播流量進入VRF以及流量速率達到閾值時。生成MDT資料包。

2.將MDT資料包封裝到源和目標3232的UDP中。並傳送給感興趣的接收者。

```
    252 23.108432     1.1.1.1              224.0.0.13           UDP         82 3232 → 3232  Len=16
▸ Frame 252: 82 bytes on wire (656 bits), 82 bytes captured (656 bits) on interface 0
▸ Ethernet II, Src: aa:bb:cc:00:30:20 (aa:bb:cc:00:30:20), Dst: IPv4mcast_01:01:01 (01:00:5e:01:01:01)
▸ Internet Protocol Version 4, Src: 1.1.1.1, Dst: 232.1.1.1
▸ Generic Routing Encapsulation (IP)
▸ Internet Protocol Version 4, Src: 1.1.1.1, Dst: 224.0.0.13
▸ User Datagram Protocol, Src Port: 3232 (3232), Dst Port: 3232 (3232)
▸ Data (16 bytes)
```

3.將UDP資料包傳送給感興趣的接收方後，設定「y」標誌並將MDT next_hop更改為新的MDT組地址。

在源PE PE1:

```
MRT(1): Set the y-flag for (10.1.0.2,224.1.1.1)

PIM(1): MDT next_hop change from: 232.1.1.1 to 232.2.2.0 for (10.1.0.2, 224.1.1.1) Tunnel2


PE1#sh ip mroute vrf m-GRE verbose

IP Multicast Routing Table
```

```
Flags:

      T - SPT-bit set, J - Join SPT, M - MSDP created entry, E - Extranet,

      Y - Joined MDT-data group, y - Sending to MDT-data group,

        p - PIM Joins on route




(10.1.0.2, 224.1.1.1), 00:08:09/00:02:46, flags: Typ

 Incoming interface: Ethernet0/2, RPF nbr 10.1.0.2

 Outgoing interface list:

   Tunnel2, GRE MDT: 232.2.2.0 (data), Forward/Sparse, 00:08:09/00:03:27, A, p (Small "p"
indicates downstream PIM join)
```

附註：OIL下一跳更改為232.2.2.0。

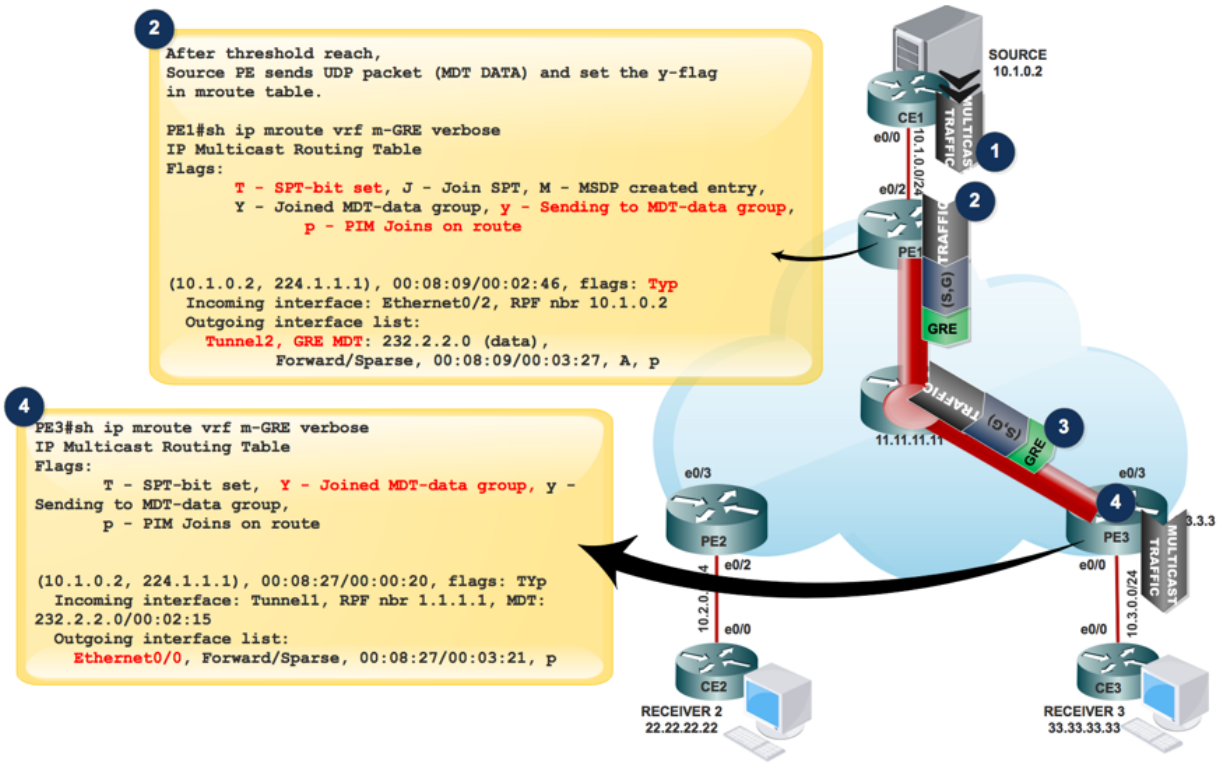
2. 在PE3，當它收到封裝在UDP SRC埠3232和DST埠3232中的MDT資料包時。

```
UDP: rcvd src=1.1.1.1(3232), dst=224.0.0.13(3232), length=24

PIM(1): Receive MDT Packet (1418) from 1.1.1.1 (Tunnel2), length (ip: 44, udp: 24), ttl: 1

PIM(1): TLV type: 1 length: 16 MDT Packet length: 16

MRT(1): Set the Y-flag for (10.1.0.2,224.1.1.1)


PE3#sh ip mroute vrf m-GRE verbose

IP Multicast Routing Table

Flags:

      T - SPT-bit set, Y - Joined MDT-data group, y - Sending to MDT-data

      p - PIM Joins on route


(10.1.0.2, 224.1.1.1), 00:08:27/00:00:20, flags: TYp

 Incoming interface: Tunnel1, RPF nbr 1.1.1.1, MDT:232.2.2.0/00:02:15

 Outgoing interface list:

   Ethernet0/0, Forward/Sparse, 00:08:27/00:03:21, p
```

S-PMSI加入消息是目標地址為ALL-PIM-ROUTERS(224.0.0.13)且目標埠為3232的UDP封裝消息。

S-PMSI加入消息包含以下資訊：要繫結到P隧道的特定組播流的識別符號。這可以表示為(S，G)對。流要繫結到的特定P隧道的識別符號。此識別符號是包含以下資訊的結構化欄位：

MDT資料隧道中的組播流量：




```
PE1#sh ip pim mdt send


MDT-data send list for VRF: m-GRE

 (source, group)                      MDT-data group/num   ref_count

 (10.1.0.2, 224.1.1.1)                232.2.2.0            1
```

```
PE3#sh ip pim mdt receive


Joined MDT-data [group/mdt number : source]  uptime/expires for VRF: m-GRE

 [232.2.2.0 : 1.1.1.1]  00:00:41/00:02:18
```

- 如果OIL包含隧道介面，則使用GRE封裝資料包，源是本地PE路由器的BGP對等地址，目標是MDT組地址。
- 選擇Data-MDT組的決定取決於mVRF中的(S， G)條目上是否設定了y標誌。
- 如果(S， G)或(*, G)條目設定了Z標誌，則這是具有相關mVRF的Default — 或Data-MDT。
- 必須解封P資料包以顯示C資料包。

Flag	Name	Description
Z	Multicast Tunnel	Indicates that this entry is an IP multicast group that belongs to the Default or Data MDT tunnel. All packets received for this IP multicast state are sent to the MDT tunnel for decapsulation . Set on receiving PE. Global mulitcast routing table
Y	Joined MDT-data group	Indicates that the traffic was received through a Data MDT tunnel that was set up specifically for this source and group. MVRF multicast routing table
z	MDT-data group sender	Set on sending PE. Global mulitcast routing table
y	Sending to MDT-data group	Indicates that the traffic was sent through a Data MDT tunnel that was set up specifically for this source and group. MVRF multicast routing table
V	RD & Vector	
v	Vecor	
E	Extranet source mroute entry	Indicates that a (*, G) or (S, G) entry in the VRF routing table is a source Multicast VRF (MVRF) entry and has extranet receiver MVRF entries linked to it

- 由於每個組播域的mVRF中僅存在一個MTI，因此Data-MDT和Default-MDT都會對客戶流量使用相同的隧道介面。
- Y/y標誌是區分預設MDT流量和資料MDT流量所必需的，並確保客戶組播路由條目使用正確的MDT-Data組，並引用包含(S、G、Data-MDT)對映的內部表。

疑難排解

目前尚無適用於此組態的具體疑難排解資訊。