

下一代组播默认MDT:配置文件0

目录

[简介](#)

[背景信息](#)

[PIM作为重叠信令](#)

[配置任务](#)

[配置](#)

[验证](#)

[故障排除](#)

简介

本文档介绍在下一代组播中使用多协议标签交换(MPLS)核心时组播数据包如何传输。

背景信息

默认MDT - PIM C — 组播信令

Draft Rosen使用通用路由封装(GRE)作为重叠协议。这意味着所有组播数据包都封装在GRE中。虚拟LAN模拟为VPN中的所有提供商边缘(PE)路由器加入组播组。这称为默认组播分布树(MDT)。默认MDT用于协议独立组播(PIM)呼叫和其他PIM信令，也用于数据流量。如果源设备发送大量流量，则使用默认MDT会效率低下，并且可以创建数据MDT。数据MDT将仅包括具有正在使用的组的接收器的PE。

罗森草案部署相当简单，工作效果良好，但有一些缺点。让我们看一下：

开销 - GRE为数据包增加24字节的开销。与通常添加8或12字节的MPLS相比，每个数据包增加的开销为100%或更多。

核心中的PIM - Draft Rosen要求在核心中启用PIM，因为PE必须加入通过PIM信令完成的默认和或数据MDT。如果核心中使用PIM ASM，则也需要RP。如果PIM SSM在核心中运行，则无需RP。

核心状态 — 由于来自PE的PIM信令，在核心中创建不必要状态。核心应尽可能少地拥有状态。

PIM邻接- PE将成为彼此的PIM邻居。如果是大型VPN和大量PE，将创建大量PIM邻接关系。这会生成大量hello和其他信令，增加路由器的负担。

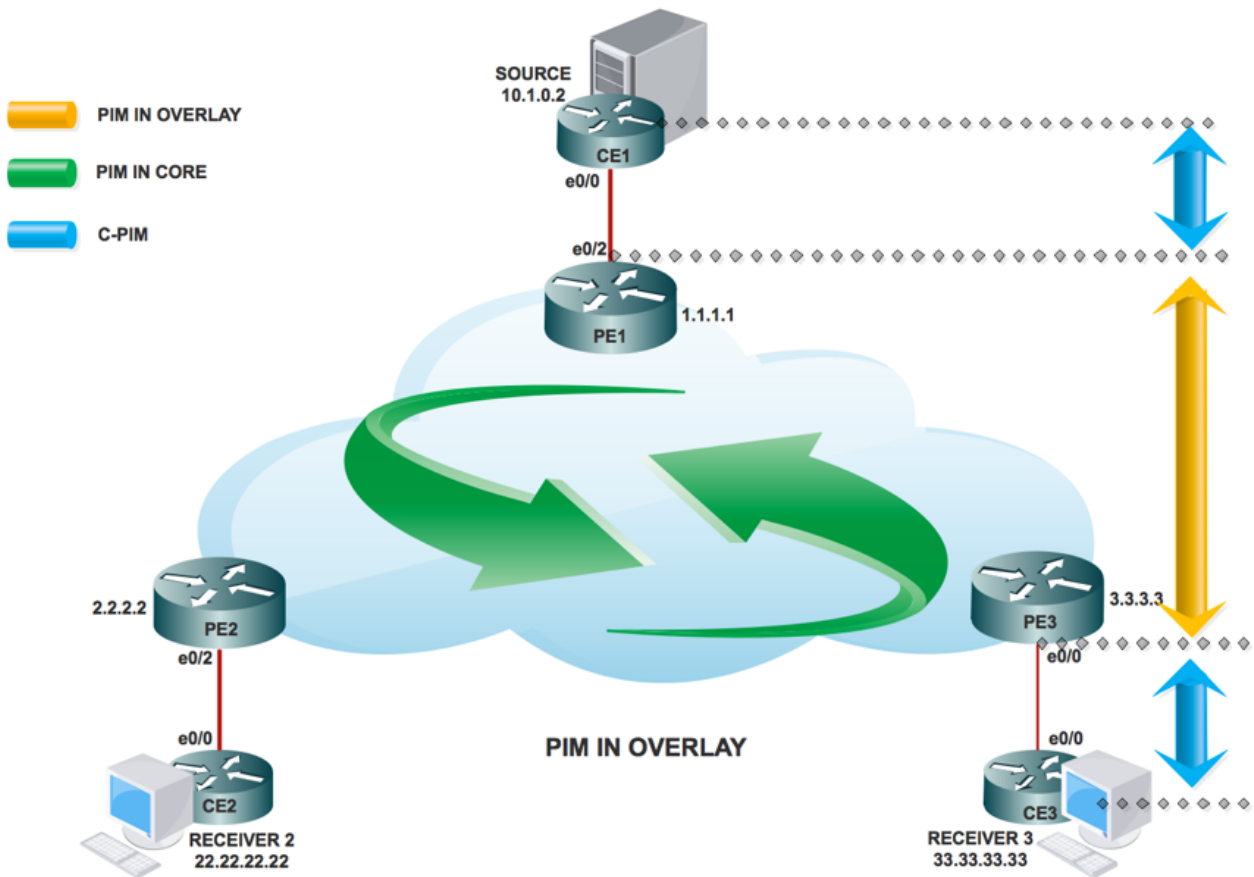
单播与组播 — 单播转发使用MPLS，组播使用GRE。这增加了复杂性，并意味着单播使用的转发机制与组播不同，而组播不是最佳解决方案。

低效 — 默认MDT将流量发送到VPN中的所有PE，而不管PE在(*,G)或(S, G)中是否有接收器用于

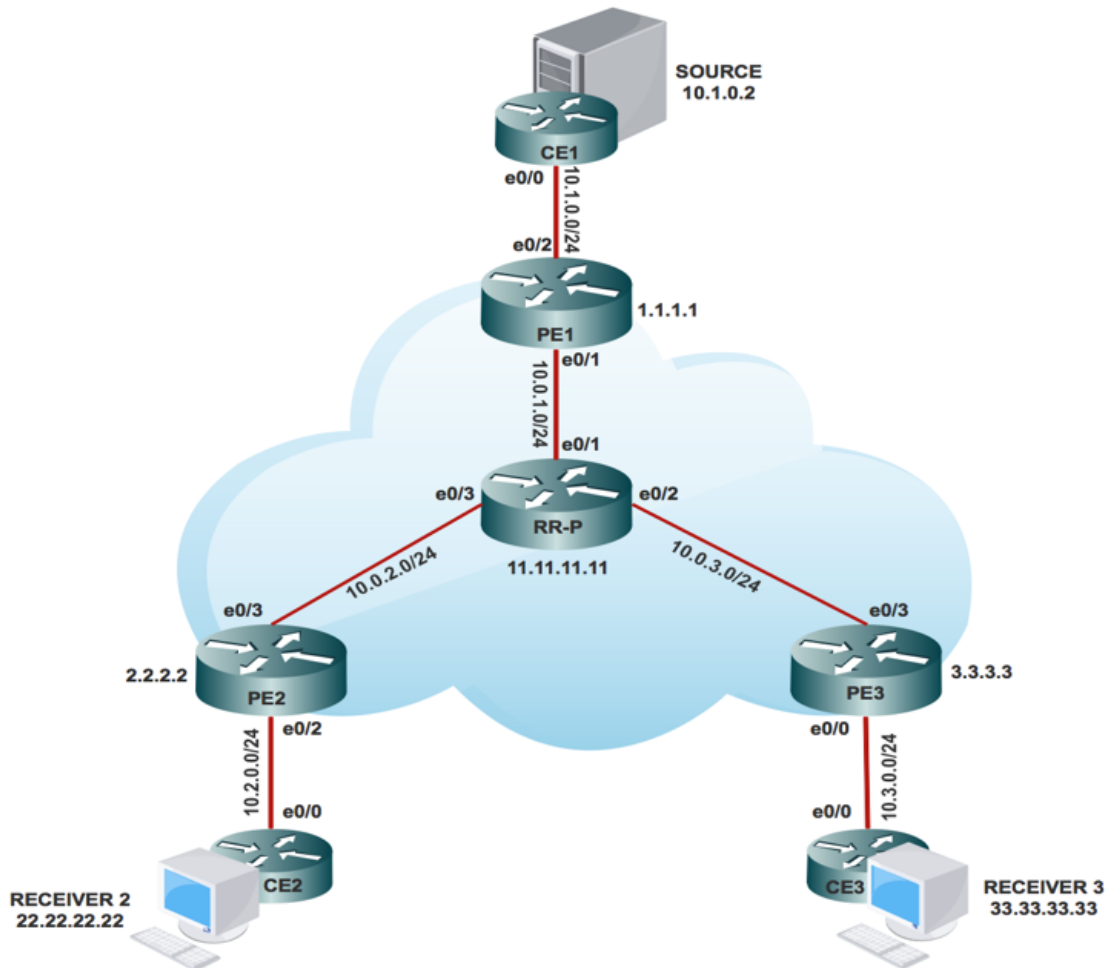
组。

- 默认MDT将用于将组播连接到一个VRF中的所有PE。
- 默认值表示它连接所有PE路由器。
- 默认情况下，它传输所有流量。
- 所有PIM控制流量和数据平面流量。例如(*,G)流量和(S, G)流量。
- 这表示多点对多点。
- 任何人都可以发送，每个人都可以从树上接收。

PIM作为重叠信令



拓扑



配置任务

1. 在所有节点上启用组播路由。
2. 在所有接口中启用PIM稀疏模式。
3. 使用现有VRF配置默认MDT。
4. 在接口Ethernet0/x上配置VRF。
5. 在VRF上启用组播路由。
6. 在核心内的所有节点中配置PIM SSM默认值。
7. 在CE节点中配置BSR RP。
8. 预配置：

- VRF m-GRE
- mBGP:地址系列VPNv4
- VRF路由协议

配置

1. 在所有节点上启用组播路由。

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

2. 在所有接口中启用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.使用已存在的VRF，配置默认MDT。

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

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

4.在接口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上启用组播路由。

在PE1、PE2和PE3上。

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

6.为服务提供商核心配置RP。

在PE1、PE2、PE3和RR-P节点上。

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

7.在CE节点（接收方）中配置BSR RP。

在Receiver2上。

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

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

验证

使用本部分可确认配置能否正常运行。

任务 1：验证物理连接

- 验证所有连接的接口是否都为“UP”

任务 2：验证地址系列VPNv4单播

- 验证AF VPNv4单播的所有路由器中是否启用了BGP，且BGP邻居处于“UP”状态
- 验证BGP VPNv4单播表具有所有客户前缀。

任务 3：检验端到端组播流量。

- 检查PIM邻居关系。
- 检验组播状态是否端到端创建。
- 验证PE1、PE2和PE3上的mRIB条目
- 验证(S, G)mFIB条目，数据包在软件转发时递增。
- 检验从CE到CE的ICMP数据包是否到达。

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: 4500000000000000FF2FD0CA01010101E801010100000800
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

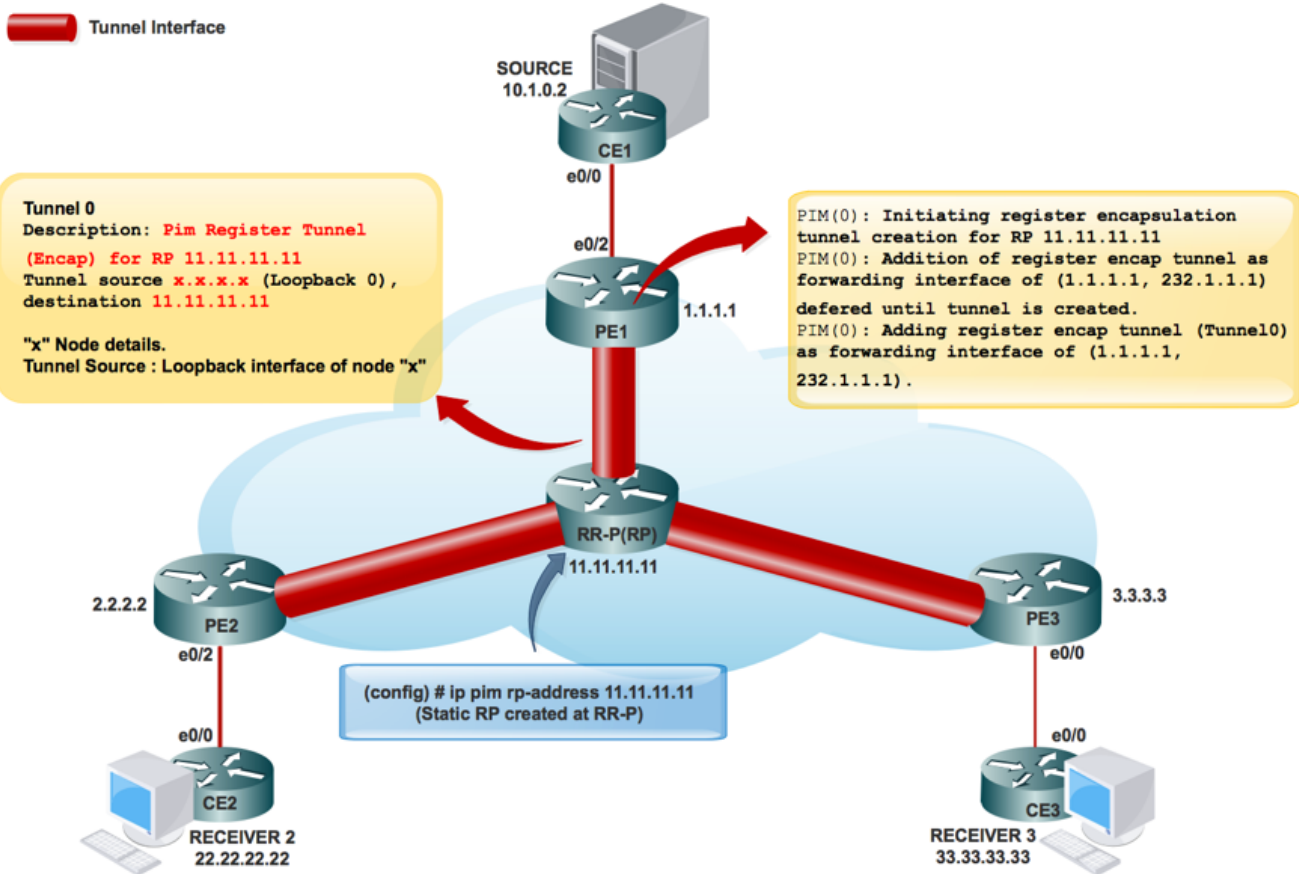
Type escape sequence to abort.
Sending 100, 100-byte ICMP Echos to 230.1.1.1, timeout is 0
seconds:

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
```

创建隧道接口时：

服务提供商RP创建：

一旦RP信息泛洪到核心。接口隧道0已创建。



PIM(0):正在为RP 11.11.11.11启动寄存器封装隧道创建。

PIM(0):RP 11.11.11.11的初始寄存器隧道创建成功。

PIM(0):添加寄存器封装隧道作为(1.1.1.1、232.1.1.1)的转发接口，直到创建隧道。

1955年5月9日17:34:56.155:PIM(0):将RP 11.11.11.11检查到(、232.1.1.1)中。

PIM(0):添加寄存器封装隧道(Tunnel0)作为(1.1.1.1、232.1.1.1)的转发接口。

```
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
```

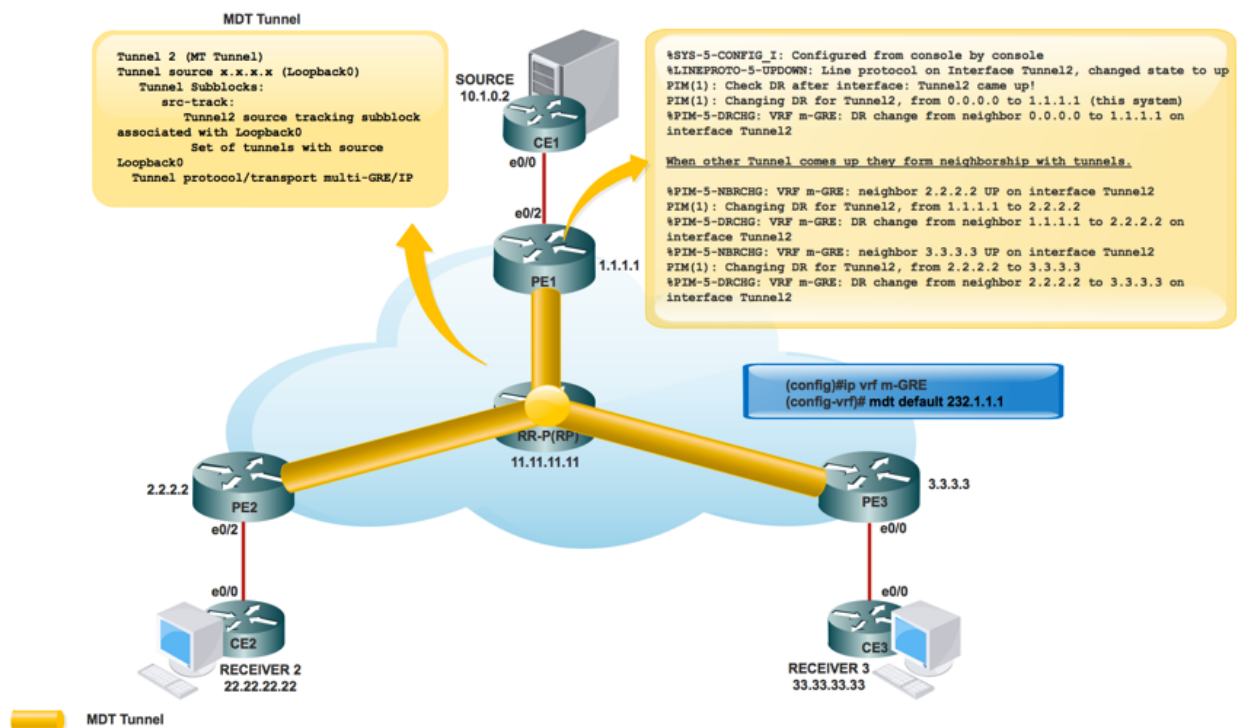
```
<OK> Set of tunnels with source Ethernet0/1, 1 member (includes iterators), on interface
```

```
Tunnel protocol/transport PIM/IPv4
```

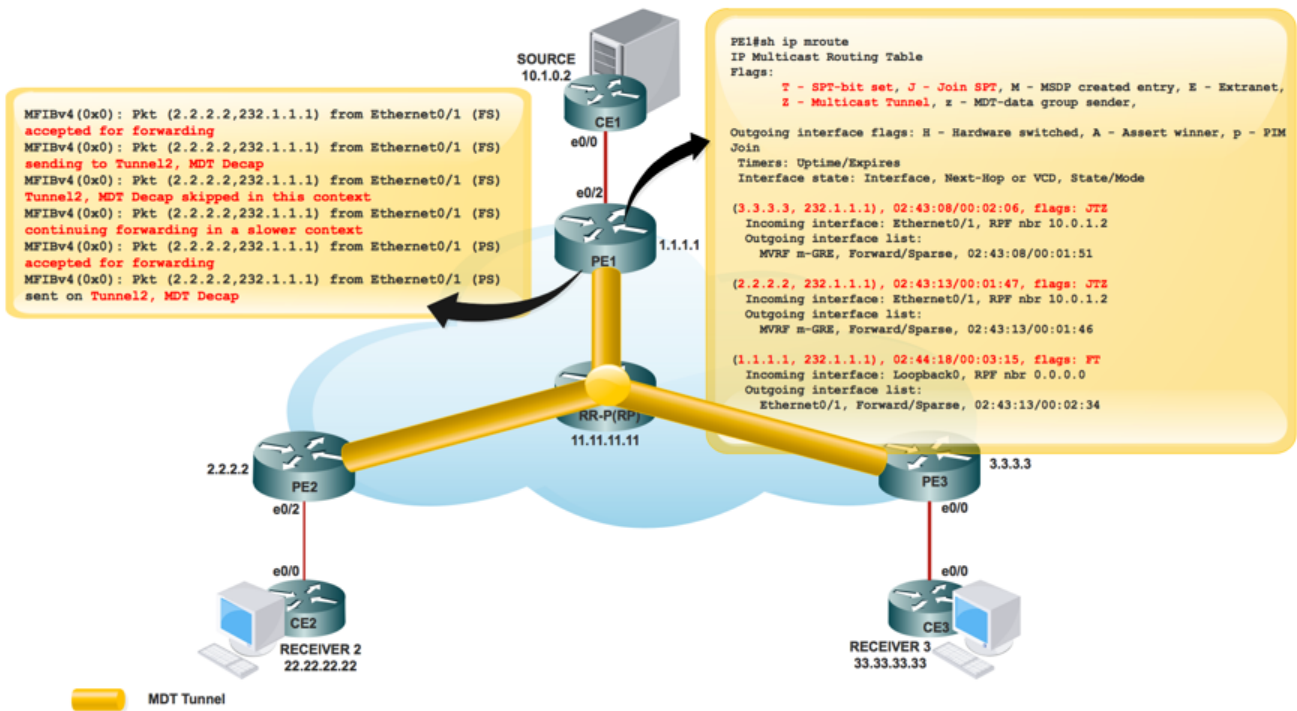
```
Tunnel TOS/Traffic Class 0xC0, Tunnel TTL 255
```

```
Tunnel transport MTU 1472 bytes
```

MDT隧道创建：



核心中的MRIB创建：



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:

MVRFP 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:

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

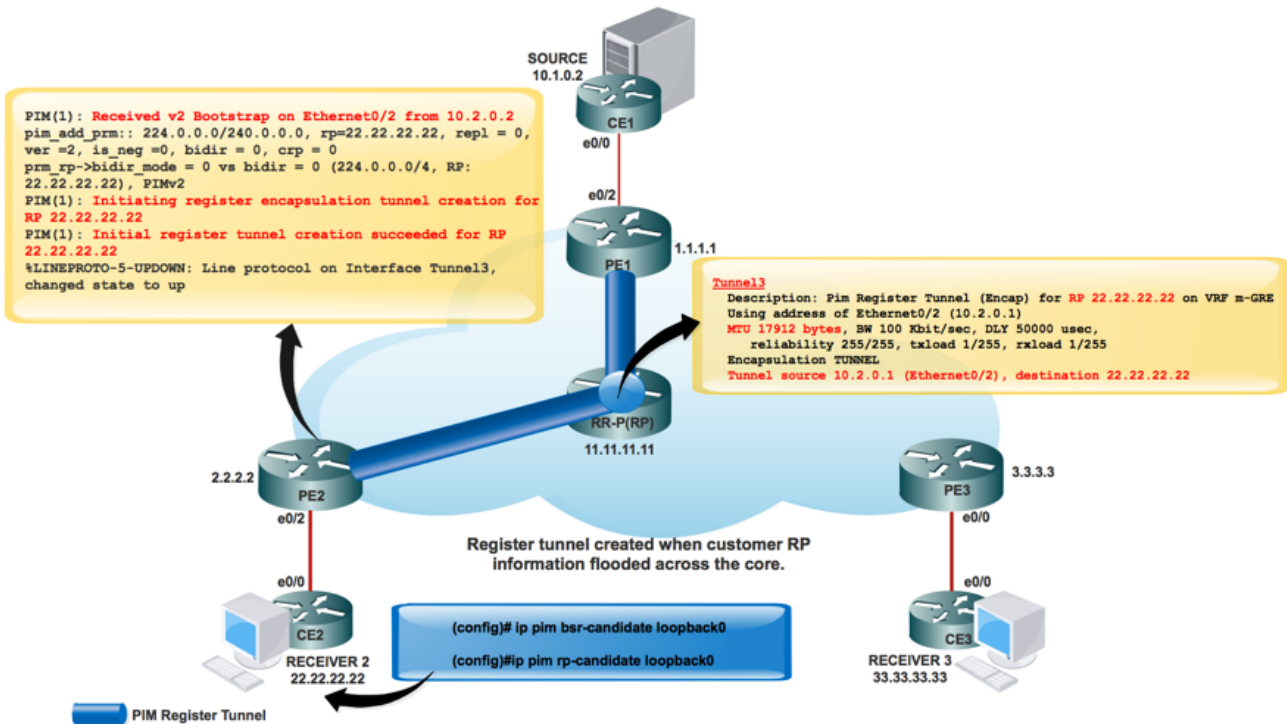
(1.1.1.1, 232.1.1.1), 00:10:15/00:03:20, flags: FT

Incoming interface: Loopback0, RPF nbr 0.0.0.0

Outgoing interface list:

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

为客户网络创建RP后：



*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

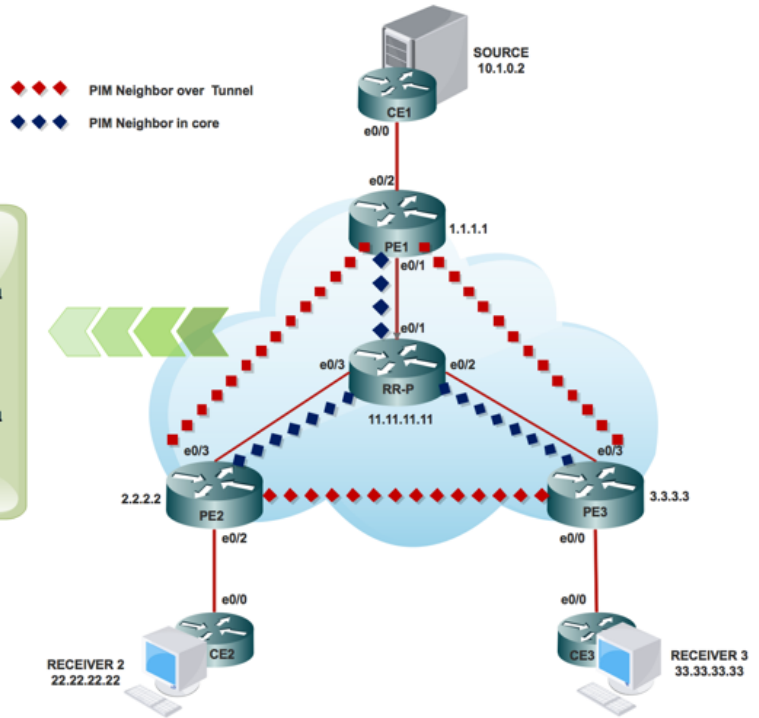
创建隧道接口以传送客户RP信息。

PIM(1):正在为RP 22.22.22.22启动寄存器封装隧道创建。

它是为向RP注册封装而创建的隧道。

对于发现的每个稀疏模式RP，会创建一个寄存器封装隧道。在稀疏模式RP本身上，创建了一个解封隧道接口来接收寄存器数据包。

PIM邻居关系：



Control Plane Scalability:

For Example:

- => PE anticipating 100 MVPN services which distributed across 100 PEs.
- => Each PE maintains 9900 (99x100) PIM adjacencies in addition to the adjacency.
- => In order to preserve 9900 PIM adjacencies, the PE would be sending approx 330 PIM adjacencies per second (Using default 30s PIM hello timer)
- => The number will get worse as the number of MVPN services or PEs increases.

```
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	Tunnell	01:44:24/00:01:41	v2	1 / DR S P G
2.2.2.2	Tunnell	01:44:24/00:01:38	v2	1 / S P G

数据包流:

控制平面数据包流分为两部分。

1. 接收器上线。

2. 源处于活动状态。

接收方处于活动状态时：

```

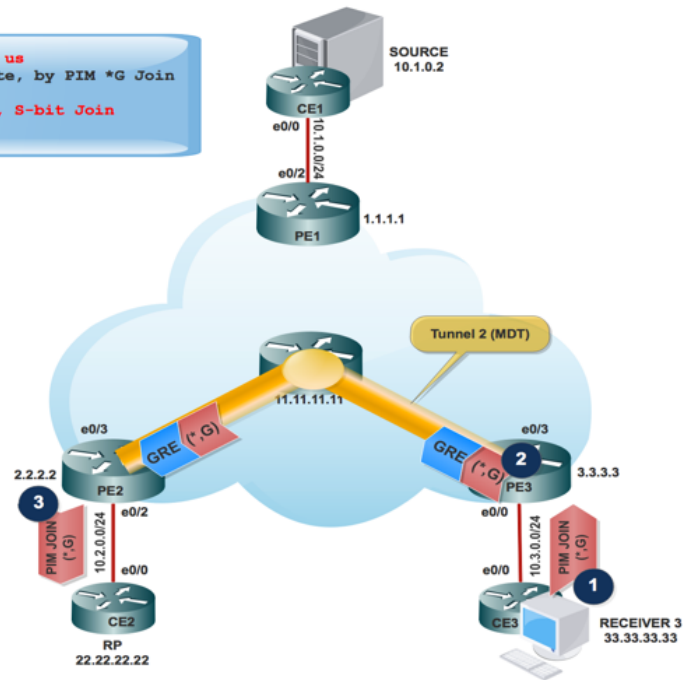
1
PIM(1): Received v2 Join/Prune on Ethernet0/0 from 10.3.0.2, to us
PIM(1): Add Ethernet0/0/10.3.0.2 to (*, 224.1.1.1), Forward state, by PIM *G Join
PIM(1): Upstream mode for (*, 224.1.1.1) changed from 0 to 1
PIM(1): Adding v2 (22.22.22.22/32, 224.1.1.1), WC-bit, RPT-bit, S-bit Join
PIM(1): Send v2 join/prune to 2.2.2.2 (Tunnel2)
    
```

```

2
PE3#sh ip mroute vrf m-GRE
IP Multicast Routing Table
Flags: D - Dense, S - Sparse, B - Bidir Group, s - SSM Group,
       C - Connected, J - Join SPT,
(*, 224.1.1.1), 00:00:22/00:02:46, RP 22.22.22.22, flags: SJC
Incoming interface: Tunnel2, RPF nbr 2.2.2.2
Outgoing interface list:
Ethernet0/0, Forward/Sparse, 00:00:22/00:03:07
    
```

```

3
PE2# sh ip mroute vrf m-GRE
(*, 224.1.1.1), 01:03:26/00:02:57, RP 22.22.22.22, flags: Sp
Incoming interface: Ethernet0/2, RPF nbr 10.2.0.2
Outgoing interface list:
Tunnel2, GRE MDT: 232.1.1.1 (default), Forward/Sparse,
01:03:26/00:02:57, p
    
```



1. 接收方联机，向PE3发送PIM JOIN(*,G)。

2. PE3将PIM JOIN(*,G)封装到GRE数据包中，并通过隧道2 (MDT隧道) 发送，该隧道从show ip mroute vrf m-GRE的传入接口进行验证。

```

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收到源为3.3.3.3且目的为232.1.1.1的GRE数据包，并根据OIL将其转发到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
```

GRE数据包被解封，PIM加入向RP发送。

注意：RPF邻居为2.2.2.2，因为PIM加入的目的地是RP地址，以通过核心形成RPT。

注意：WC位和RPT位：由(*,G)状态触发，DR会创建加入/修剪消息，其中RP地址在其加入列表中，通配符位（WC位）和RP树位（RPT位）设置为1。WC位表示任何源可能匹配并根据此条目转发不再匹配的；RPT位表示此加入是向共享RP树发送的。修剪列表留空。当RPT位设置为1时，它表示连接与共享RP树关联，因此，连接/修剪消息沿RP树传播。当WC位设置为1时，它表示地址是RP，而下游接收方期望通过此（共享树）路径从所有源接收数据包。

```
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. 源PE PE1处的GRE封装数据包到达。

```
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)到达源CE。

3. 现在 , 源设备获取了感兴趣的接收方的信息 , 流量开始发送到源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
```

从PE1捕获组播数据包的PCAP。在MDT默认隧道中隧道传输。使用GRE封装。

5. 在接收方PE PE3处，接收数据包。

```
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：收到来自PE1的组播数据包后，RPF邻居将更改。之前它是PE2，作为RP托管在后面。收到第一个组播数据包后，它会更改RPF并设置SPT位。

通过默认MDT隧道的流量：

- MDT上的转发使用GRE，C数据包成为P数据包。
- P-Packet S地址= PE的BGP对等地址
G地址= MDT组地址（默认或数据）
- C-Packet IP TOS将复制到P-Packet。
- MPLS标签不用于核心，而仅用于本地组播。

数据包流:

1. C-Packet到达VRF配置的PE接口，mVRF被隐式标识。C源上的正常RPF检查。C数据包从OIL中的接口复制出去。此时，这将是同一VRF中的PE接口。

```
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，则C数据包封装成P数据包。如果在条目目标上设置“y”标志，则使用DATA-MDT组，否则使用默认MDT组。源是PE BGP对等体地址，目标是MDT组地址。

2. P数据包按照正常组播通过P网络转发。

数据包到达全局接口。引用的MDT组的全局(S, G)或(*,G)条目。P-Source (PE对等体)上的正常RPF检查。

3. P数据包在OIL中从接口复制出去。此时，全局mroute表中为P/PE。

4. 如果“Z”标志设置了数据包，则解封数据包以揭示C数据包。从MDT组派生的目标mVRF和传入接口是封装报头的目标。

在mVRF中对C-Packet进行RPF检查，在mVRF中对C-Packet进行OIL复制。

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,

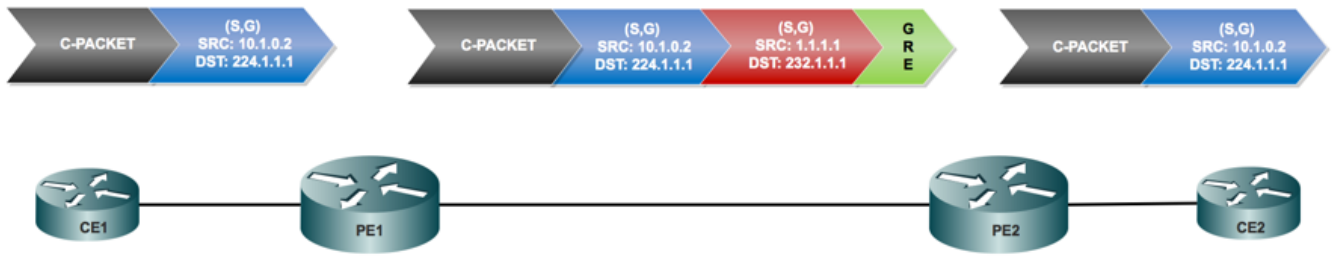
(1.1.1.1, 232.1.1.1), 1d01h/00:02:47, flags: **JTZ**

Incoming interface: Ethernet0/3, RPF nbr 10.0.3.2

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

5. 接收方3上的本地C数据包到达。

数据包封装:

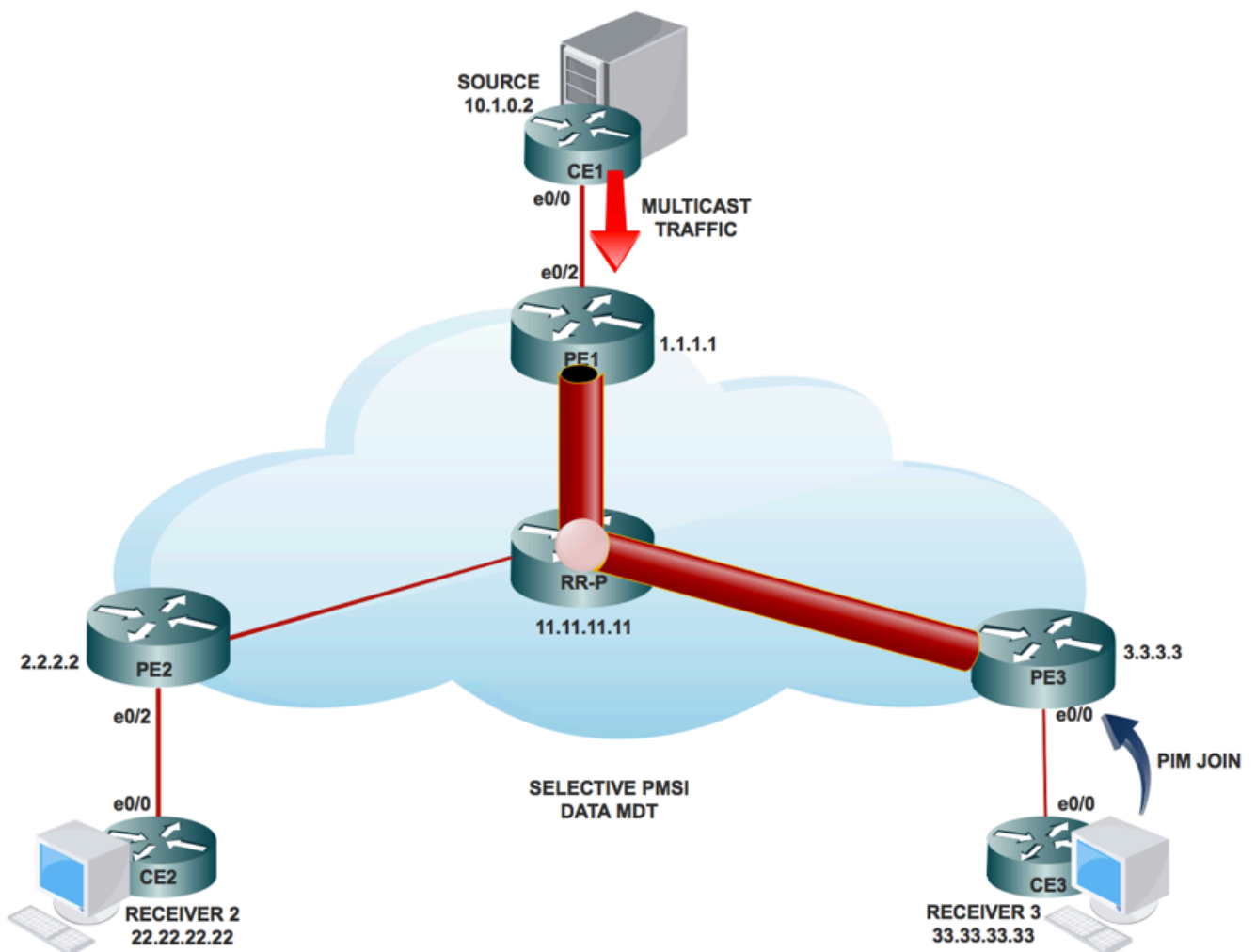


数据MDT:

什么是数据MDT?

它是可选的。它是按需创建的，它传输特定(S, G)流量。在最新IOS®版本中，配置的阈值为“0”和“infinite”。当第一个数据包到达VRF时，数据MDT会初始化，如果无穷大，则永远不会创建数据MDT，并且流量在默认MDT中向前移动。数据MDT始终是接收树，它们从不发送任何流量。数据MDT仅用于(S, G)流量。

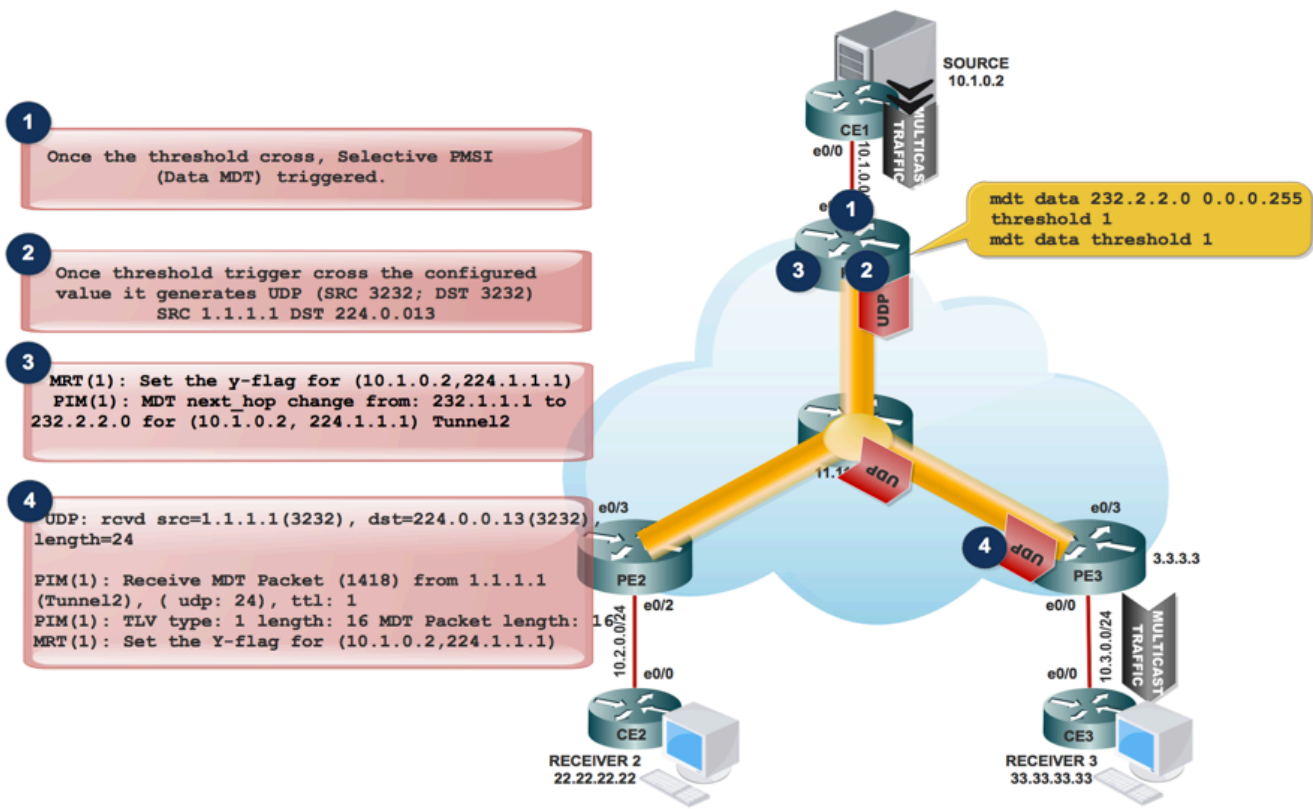
选择性PMSI:



- 是可选的。它是按需创建的，它传输特定(S, G)流量。
- 每当第一个数据包到达VRF时，数据MDT将初始化，如果无限，则从不创建数据MDT，并且流量在默认MDT中前进。

- 数据MDT始终是接收树，它们从不发送任何流量。数据MDT仅用于(S, G)流量。
- PIM消息传送C-(S, G)和P-Group。

DATA MDT的创建方式：



1. 当组播流量进入VRF和流量速率达到阈值时。生成MDT数据包。
2. MDT数据包封装在UDP中，源地址和目的地址为3232。然后发送给感兴趣的接收方。

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.在将UDP数据包发送到感兴趣的接收方后，它会设置“y”标志并将MDT next_hop更改为新的MDT组地址。

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)

注意：OIL下一跳更改为232.2.2.0。

2. 在PE3，当它收到封装在UDP SRC端口3232和DST端口3232中的MDT数据包时。

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:

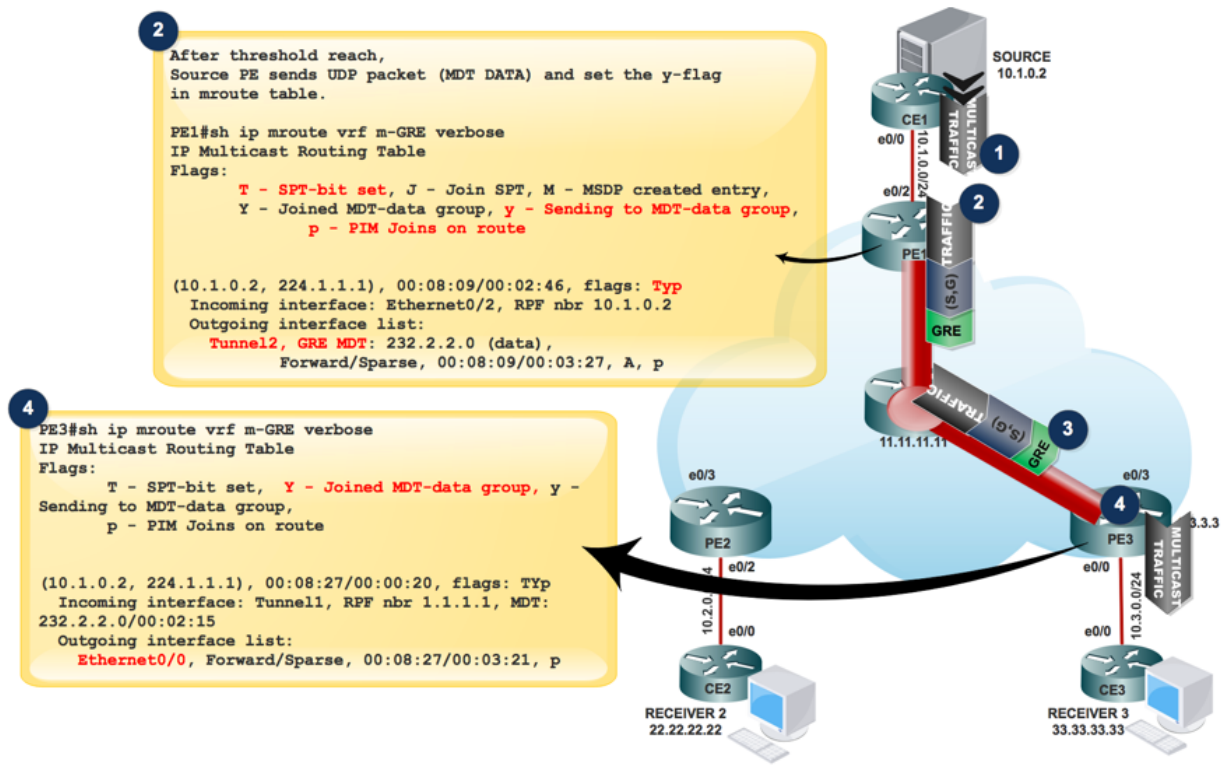
Ethernet0/0, Forward/Sparse, 00:08:27/00:03:21, p

S-PMSI加入消息是UDP封装的消息，其目的地址为ALL-PIM-ROUTERS(224.0.0.13)，目的端口为3232。

S-PMSI加入消息包含以下信息：要绑定到P隧道的特定组播流的标识符。这可以表示为(S, G)对。

流量绑定到的特定P隧道的标识符。此标识符是包含以下信息的结构化字段：

MDT数据隧道中的组播流量：



```
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包含隧道接口，则使用GRE封装数据包，源是本地PE路由器的BGP对等地址，目的地是MDT组地址。
- 选择Data-MDT组的决定取决于是否在mVRF中的(S, G)条目上设置y标志。
- 如果(S, G)或(*, G)条目设置了Z标志，则这是带有相关mVRF的Default — 或Data-MDT。
- 必须解封P数据包，才能显示C数据包。

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 <u>receiving</u> 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

- 由于每个组播域的mVRF中只存在一个MTI，因此Data-MDT和Default-MDT对客户流量使用相同的隧道接口。
- Y/y标志是区分Default-MDT流量与Data-MDT流量的必需标志，并确保客户组播路由条目使用正确的MDT-Data组，并引用包含(S、G、Data-MDT)映射的内部表。

故障排除

目前没有针对此配置的故障排除信息。