Next Generation Multicast Default MDT: Profile 0

Contents

Introduction

This document describes how Multicast packet traverses with the use of Multiprotocol Label Switching (MPLS) core in Next Generation Multicast.

Background Information

Default MDT - PIM C - mcast signaling

Draft Rosen uses Generic Routing Encapsulation (GRE) as an overlay protocol. That means that all multicast packets are encapsulated inside GRE. A virtual LAN is emulated with all Provider Edge (PE) routers in the VPN join a multicast group. This is known as the default Multicast Distribution Tree (MDT). The default MDT is used for Protocol Independent Multicast (PIM) hello's and other PIM signaling but also for data traffic. If the source sends a lot of traffic it is inefficient to use the default MDT and a data MDT can be created. The data MDT will only include PE's that have receivers for the group in use.

Draft Rosen is fairly simple to deploy and works well but it has a few drawbacks. Let's take a look at these:

Overhead – GRE adds 24 bytes of overhead to the packet. Compared to MPLS which typically adds 8 or 12 bytes there is 100% or more of overhead added to each packet.

PIM in the core – Draft Rosen requires that PIM is enabled in the core because the PE's must join the default and or data MDT which is done through PIM signaling. If PIM ASM is used in the core, an RP is needed as well. If PIM SSM is run in the core, no RP is needed.

Core state – Unnecessary state is created in the core due to the PIM signaling from the PE's. The core should have as little state as possible.

PIM adjacencies – The PE's will become PIM neighbors with each other. If it's a large VPN and a

lot of PE's, a lot of PIM adjacencies will be created. This generates a lot of hello's and other signaling which adds to the burden of the router.

Unicast vs multicast – Unicast forwarding uses MPLS, multicast uses GRE. This adds complexity and means that unicast uses a different forwarding mechanism than multicast, which is not the optimal solution.

Inefficiency – The default MDT sends traffic to all PE's in the VPN regardless if the PE has a receiver in the (*,G) or (S,G) for the group in use.

- Default MDT will be used to connect multicast to all the PE in one VRF.
- Default means it connects all the PE routers.
- By default, it carries all the traffic.
- All the PIM control traffic and the Data plane traffic. Eg (*,G) Traffic and the (S,G) traffic.
- This represent multipoint to multipoint.
- Anybody can send and everybody can receive from the tree.

PIM as Overlay Signaling
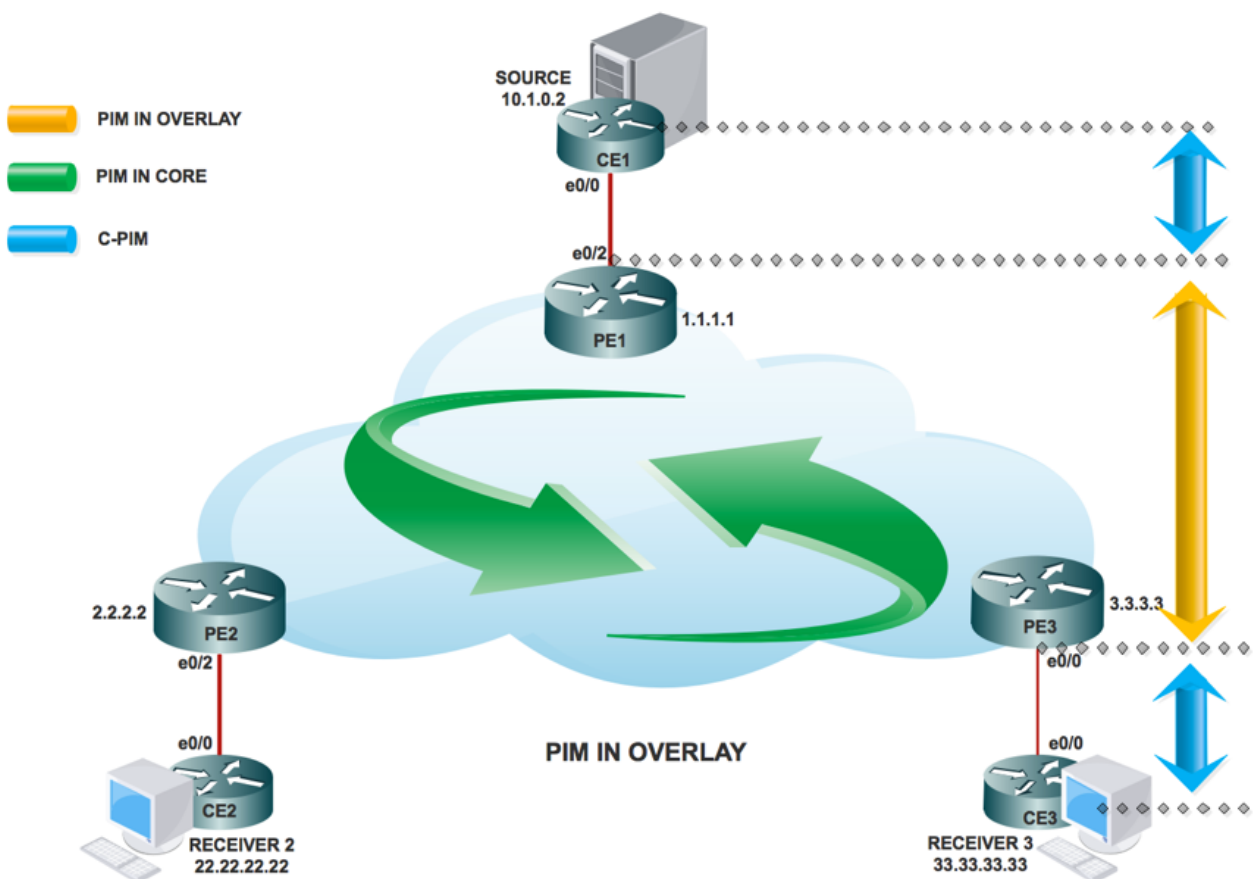
Topology
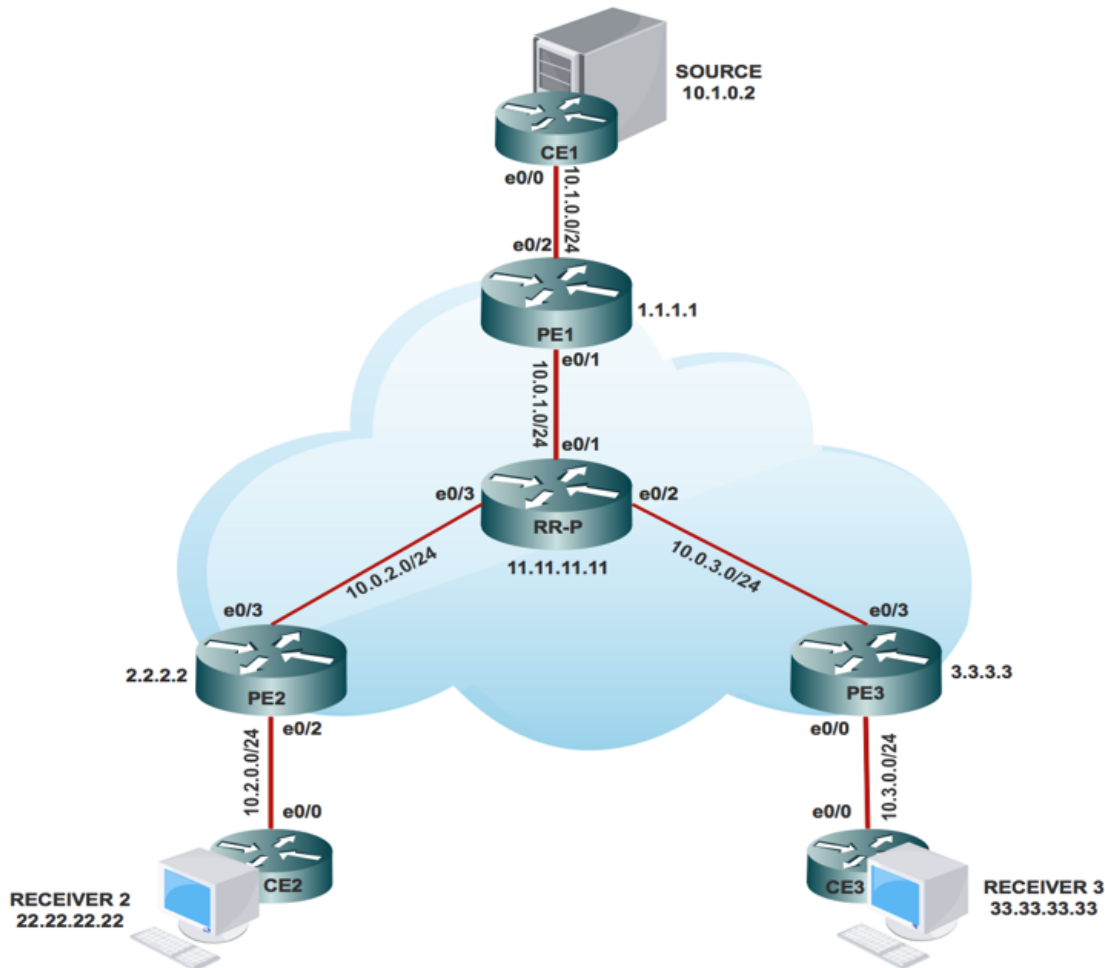
Configuration Tasks

1. Enable Multicast Routing on all nodes.
2. Enable PIM Sparse Mode in all the interface.
3. With existing VRF configure Default MDT.
4. Configure the VRF on the interface Ethernet0/x.
5. Enable Multicast Routing on VRF.
6. Configure PIM SSM Default in all nodes inside the core.
7. Configure BSR RP in CE Node.
8. Pre-Configured:

   - VRF m-GRE
   - mBGP: Address family VPNv4
   - VRF Routing Protocol

Configure

1. Enable Multicast Routing on all nodes.

```
(config)# ip multicast-routing
```
2. Enable PIM Sparse Mode in all the interface.

```
(config)# interface ethernet0/x

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

```
(config)# interface loopback0

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

3. With VRF that already exists, configure Default MDT.

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

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

4. Configure the VRF on the interface Ethernet0/x.

On PE1, PE2 and PE3.

```
(config)# interface ethernet0/x

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

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

5. Enable Multicast Routing on VRF.

On PE1, PE2 and PE3.

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

6. Configure RP for Service Provider core.

On PE1, PE2, PE3 and RR-P Node.

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

7. Configure BSR RP in CE Node (Receiver).

On Receiver2.

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

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

Verify

Use this section in order to confirm that your configuration works properly.

Task 1: Verify Physical Connectivity

- Verify all the connected interface are "UP"

Task 2: Verify Address Family VPNv4 unicast

- Verify that BGP is enabled in all the routers for AF VPNv4 unicast and BGP neighbors are "UP"
- Verify that BGP VPNv4 unicast table has all the Customer prefixes.

Task 3: Verify Multicast Traffic end to end.

- Check PIM neighborship.
- Verify that multicast state is created end to end.
- Verify mRIB entry on PE1, PE2 and PE3
- Verify that (S,G) mFIB entry, packet getting incremented in software forwarding.
- Verify ICMP packets getting reach from CE to 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: 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
```

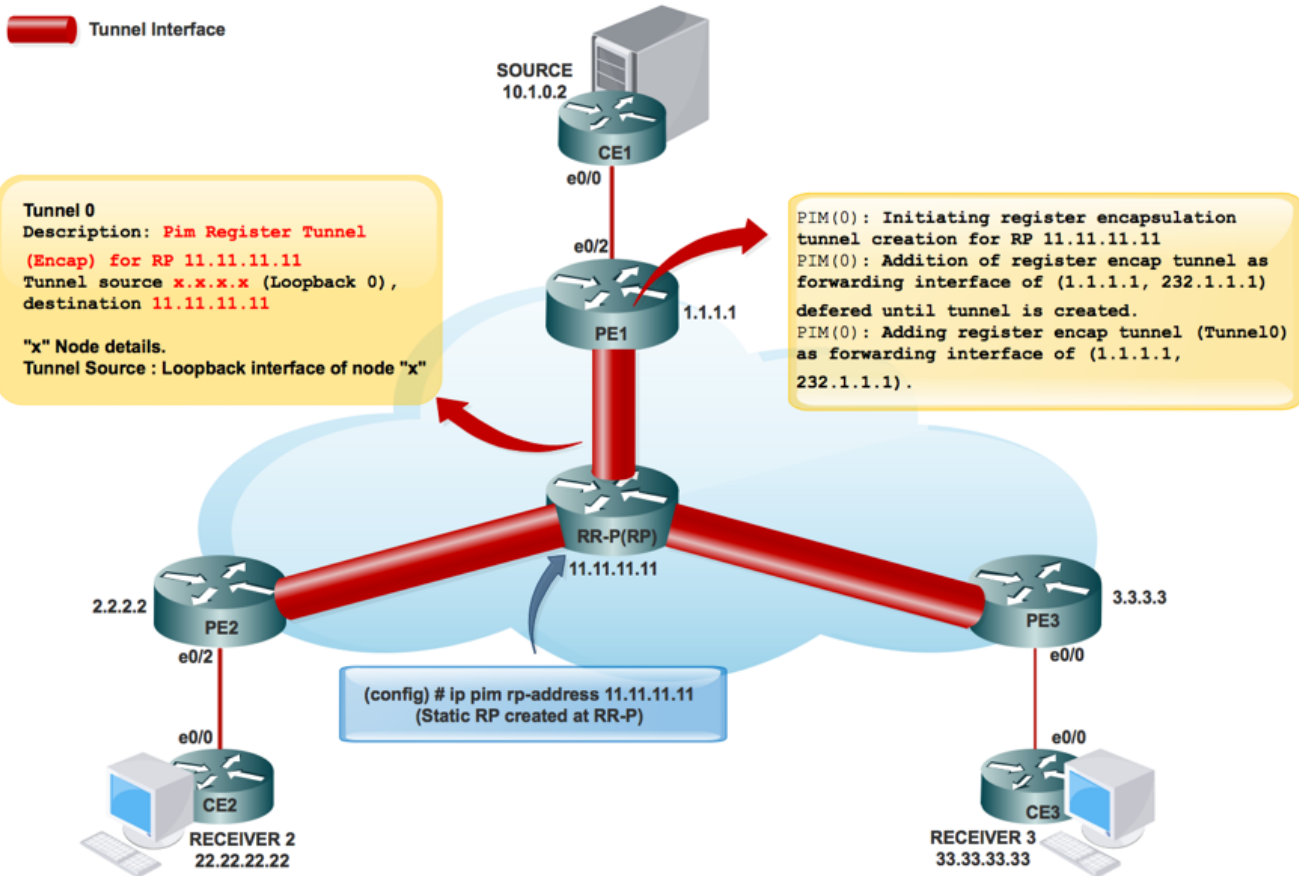**Verify ICMP packets getting reach from CE to CE**

```
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
```

When the Tunnel interfaces are created:

Service Provider RP creation:

Once the RP information flooded in the core. Interface Tunnel 0 is created.

Tunnel Interface

SOURCE
10.1.0.2

CE1
e0/0

Tunnel 0
Description: Pim Register Tunnel
(Encap) for RP 11.11.11.11
Tunnel source x.x.x.x (Loopback 0),
destination 11.11.11.11

"x" Node details.
Tunnel Source : Loopback interface of node "x"

e0/2

PE1   1.1.1.1

PIM(0): Initiating register encapsulation
tunnel creation for RP 11.11.11.11
PIM(0): Addition of register encap tunnel as
forwarding interface of (1.1.1.1, 232.1.1.1)
defered until tunnel is created.
PIM(0): Adding register encap tunnel (Tunnel0)
as forwarding interface of (1.1.1.1,
232.1.1.1).

RR-P(RP)
11.11.11.11

2.2.2.2
PE2
e0/2

e0/0
CE2
RECEIVER 2
22.22.22.22

(config) # ip pim rp-address 11.11.11.11
(Static RP created at RR-P)

3.3.3.3
PE3
e0/0

e0/0
CE3
RECEIVER 3
33.33.33.33

PIM(0): Initiating register encapsulation tunnel creation for RP 11.11.11.11.

PIM(0): Initial register tunnel creation succeeded for RP 11.11.11.11.

PIM(0): Addition of register encap tunnel as forwarding interface of (1.1.1.1, 232.1.1.1) deferred until tunnel is created.

May 9 17:34:56.155: PIM(0): Check RP 11.11.11.11 into the (, 232.1.1.1).

PIM(0): Adding register encap tunnel (Tunnel0) as forwarding interface of (1.1.1.1, 232.1.1.1).

```
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:

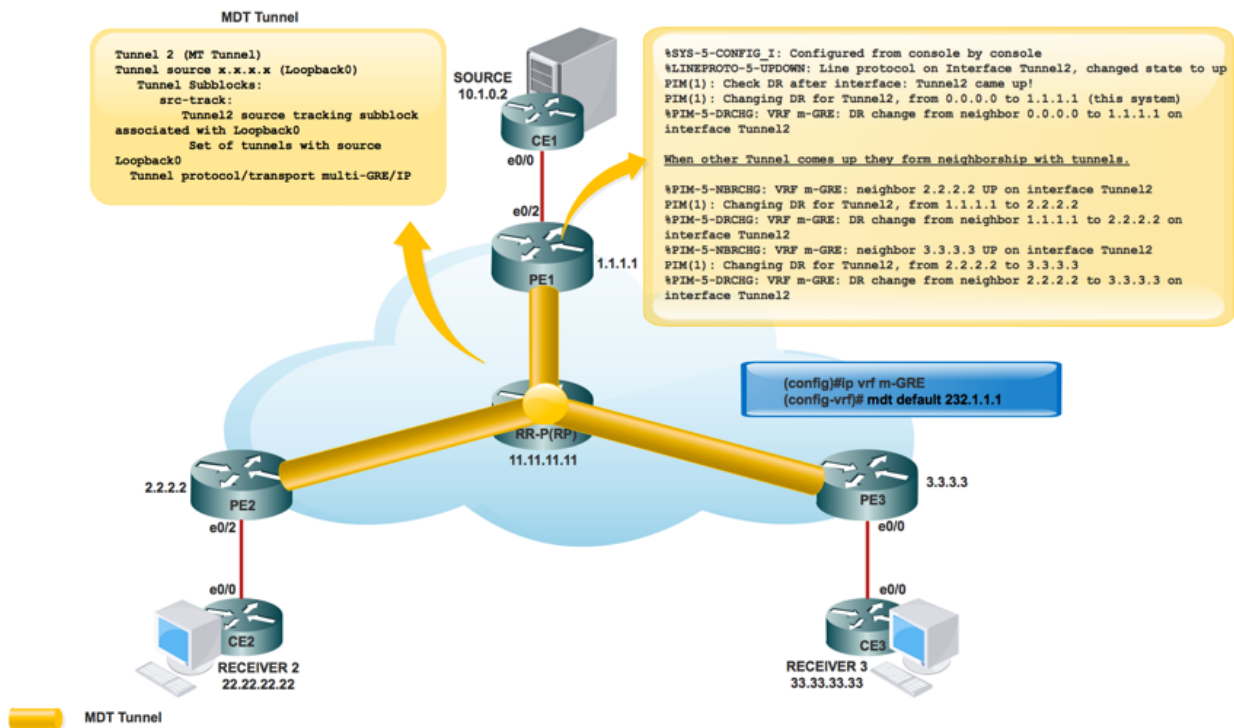   Tunnel0 source tracking subblock associated with Ethernet0/1

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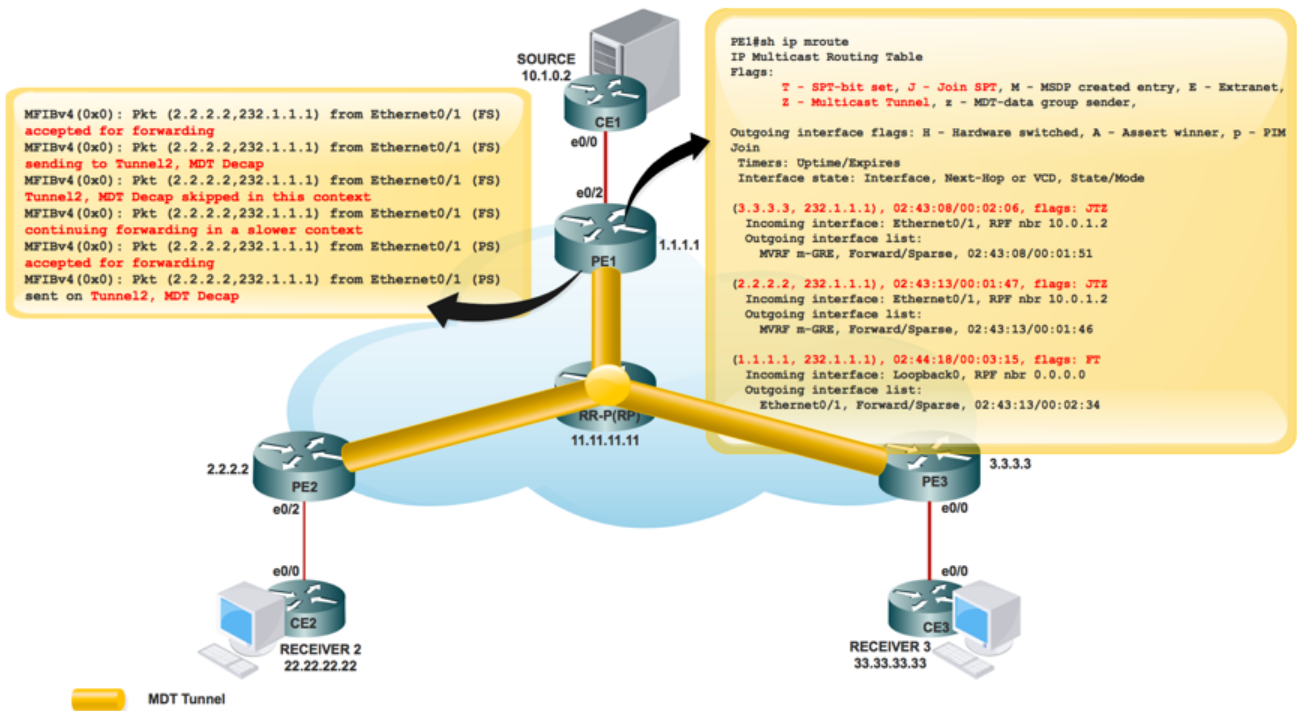<OK>

  Tunnel protocol/transport PIM/IPv4

  Tunnel TOS/Traffic Class 0xC0, Tunnel TTL 255

  Tunnel transport MTU 1472 bytes

MDT Tunnel Creation:

MRIB Creation in the core:

SOURCE
10.1.0.2

CE1

e0/0

e0/2

PE1   1.1.1.1

MFIBv4(0x0): Pkt (2.2.2.2,232.1.1.1) from Ethernet0/1 (FS)
accepted for forwarding
MFIBv4(0x0): Pkt (2.2.2.2,232.1.1.1) from Ethernet0/1 (FS)
sending to Tunnel2, MDT Decap
MFIBv4(0x0): Pkt (2.2.2.2,232.1.1.1) from Ethernet0/1 (FS)
Tunnel2, MDT Decap skipped in this context
MFIBv4(0x0): Pkt (2.2.2.2,232.1.1.1) from Ethernet0/1 (FS)
continuing forwarding list in a slower context
MFIBv4(0x0): Pkt (2.2.2.2,232.1.1.1) from Ethernet0/1 (PS)
accepted for forwarding
MFIBv4(0x0): Pkt (2.2.2.2,232.1.1.1) from Ethernet0/1 (PS)
sent on Tunnel2, MDT Decap

PE1#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,

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

(3.3.3.3, 232.1.1.1), 02:43:08/00:02:06, flags: JTZ
  Incoming interface: Ethernet0/1, RPF nbr 10.0.1.2
  Outgoing interface list:
    MVRF m-GRE, Forward/Sparse, 02:43:08/00:01:51

(2.2.2.2, 232.1.1.1), 02:43:13/00:01:47, flags: JTZ
  Incoming interface: Ethernet0/1, RPF nbr 10.0.1.2
  Outgoing interface list:
    MVRF m-GRE, Forward/Sparse, 02:43:13/00:01:46

(1.1.1.1, 232.1.1.1), 02:44:18/00:03:15, flags: FT
  Incoming interface: Loopback0, RPF nbr 0.0.0.0
  Outgoing interface list:
    Ethernet0/1, Forward/Sparse, 02:43:13/00:02:34

RR-P(RP)
11.11.11.11

2.2.2.2   PE2

e0/2

e0/0

CE2
RECEIVER 2
22.22.22.22

3.3.3.3   PE3

e0/0

e0/0

CE3
RECEIVER 3
33.33.33.33

MDT Tunnel

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:

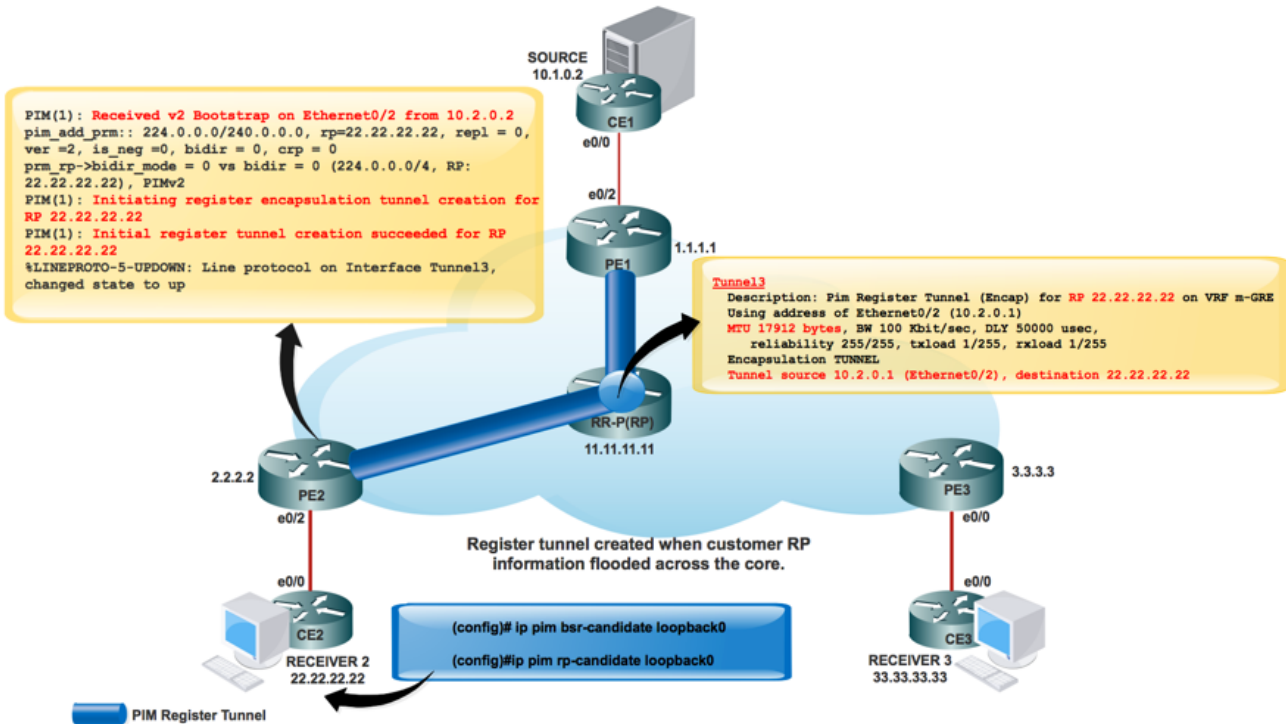   MVRF m-GRE, Forward/Sparse, 00:10:14/00:01:45

```
(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
```

Once the RP is created for the customer network:


```
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
```

Tunnel interface is created in order to carry the customer RP information.

PIM(1): Initiating register encapsulation tunnel creation for RP 22.22.22.22.

It is the tunnel created to do Register encapsulation to RP.

For every sparse-mode RP discovered, one Register encapsulation tunnel is created. On the sparse-mode RP itself, there is one de-capsulation tunnel interface created to receive Register packets.

PIM Neighborship:


```
PE1#sh ip mroute

IP Multicast Routing Table

Flags: D - Dense, S - Sparse, B - Bidir Group, s - SSM Group, C - Connected,
```
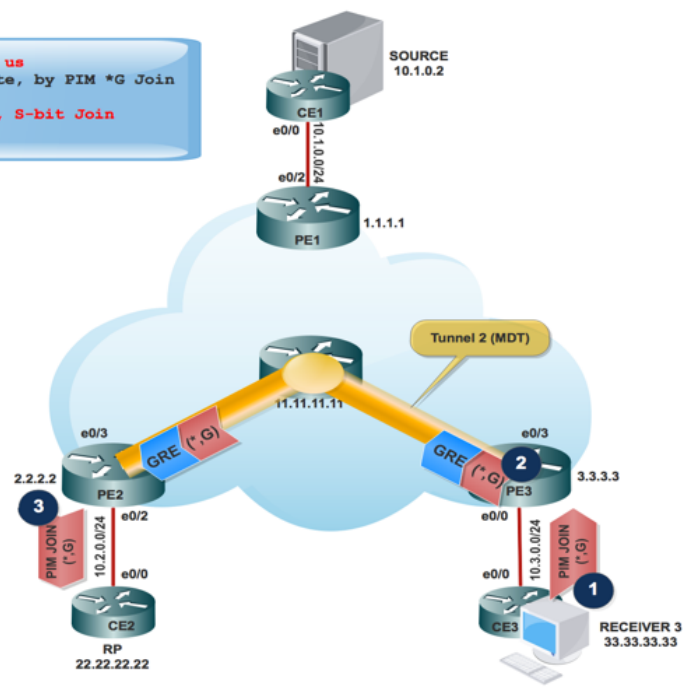
```
        L - Local, P - Pruned, R - RP-bit set, F - Register flag,

        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
```

Packet Flow:

Control plane packet flow divides into two parts.

1. Receiver comes online.
2. Source is active.

When Receiver is active:

1. Receiver comes online, sends PIM JOIN (*,G) towards PE3.

2. PE3 encapsulates the PIM JOIN (*,G) in GRE packet and sends through the tunnel 2 (MDT Tunnel), which is verified from Incoming Interface of **show ip mroute vrf m-GRE**.

```
  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
```

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,

```
 (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
```

1. PE2 received the GRE packet with Source as 3.3.3.3 and Destination 232.1.1.1 and forwards it to MVRF m-GRE based on the OIL.

```
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
```
The GRE Packet gets decapsulated and PIM JOIN sends towards the RP.

> **Note**: RPF Neighbor is 2.2.2.2 because the PIM Join is destined towards the RP address to form the RPT through the core.

> **Note**: WC Bit and RPT Bit:Triggered by the (*,G) state, the DR creates a Join/Prune message with the RP address in its join list and the the wildcard bit (WC-bit) and RP-tree bit (RPT-bit) set to 1. The WC-bit indicates that any source may match and be forwarded according to this entry if there is no longer match; the RPT-bit indicates that this join is sent up the shared RP-tree. The prune list is left empty. When the RPT-bit is set to 1 it indicates that the join is associated with the shared RP-tree and therefore the Join/Prune message is propagated along the RP-tree. When the WC-bit is set to 1 it indicates that the address is an RP and the downstream receivers expect to receivepackets from all sources via this (shared tree) path.

```
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
```

1. GRE encapsulated packet reach at Source PE PE1.

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
```

2. PIM JOIN (S,G) reaches at Source CE.

3. Now the Source got the information of the Interested Receiver and traffic starts send to the Source PE PE1.

4. At Source PE PE1:

```
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
```

At PE2 (RP PE):

```
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
```
PCAP Capture of Multicast Packet from PE1. Tunneled in MDT Default Tunnel. Encapsulated with GRE.

   5. At Receiver PE PE3, packet is received.

```
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
```

> **Note**: RPF Neighbor gets changed once a Multicast Packet from PE1 is received. Earlier it was PE2 as RP hosted behind it. After it receives the first Multicast Packet, it changes the RPF and sets the SPT bit.

Traffic Flow over Default MDT Tunnel:

- Forwarding on the MDT uses GRE, C-packet becomes a P-Packet.
- P-Packet S address = PE's BGP peering address
  G address = MDT-Group address (Default or Data)
- C-Packet IP TOS will be copied to P-Packet.
- MPLS labels are NOT used in core, only native multicast.

Packet Flow:

1. A C-Packet arrives on an VRF configured PE interface, mVRF is implicitly identified. Normal RPF check on C Source.

The C-Packet replicated out interface in the OIL. At this point this would be PE interface in the same VRF.

```
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)
```

If the OIL contains an MTI, then the C packet encapsulates into a P packet. If "y" flag is set on entry destination used is DATA-MDT group otherwise Default MDT group. Source is the PE BGP peer address and the destination is the MDT group address.

2. The P packet is forwarded through the P network as per normal multicast.

Packet arrives at the global interface. Global (S,G) or (*,G) entry for the MDT group referenced. Normal RPF check on P-Source (PE Peer).

3. The P packet is replicated out interface in OIL. At this point this is P/PE in the global mroute table.

4. If the "Z" flag sets the packet is decapsulated to reveal the C Packet. The target mVRF and incoming interface derived from MDT group is the destination of the encapsulated header.

RPF check of the C-Packet in mVRF done, C Packet replicated out OIL in mVRF.

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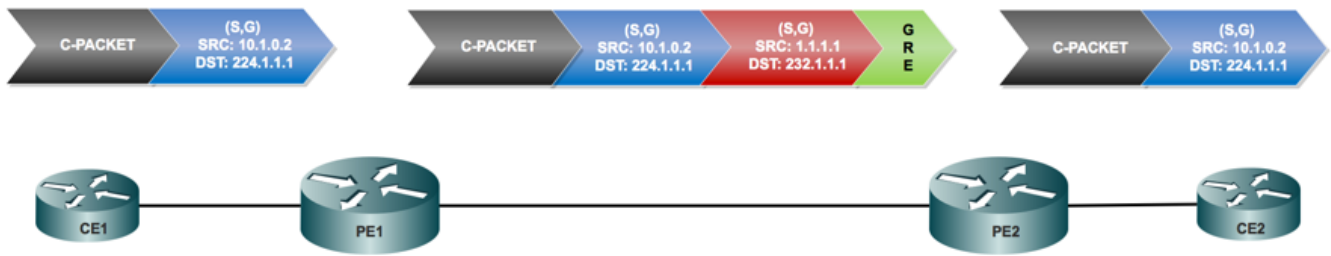
```
(1.1.1.1, 232.1.1.1), 1d01h/00:02:47, flags: JTZ

 Incoming interface: Ethernet0/3, RPF nbr 10.0.3.2
```

```
Outgoing interface list: MVRF m-GRE, Forward/Sparse, 1d01h/stopped
```
   5. Native C Packet reach at the Receiver 3.

Packet Encapsulation**:**
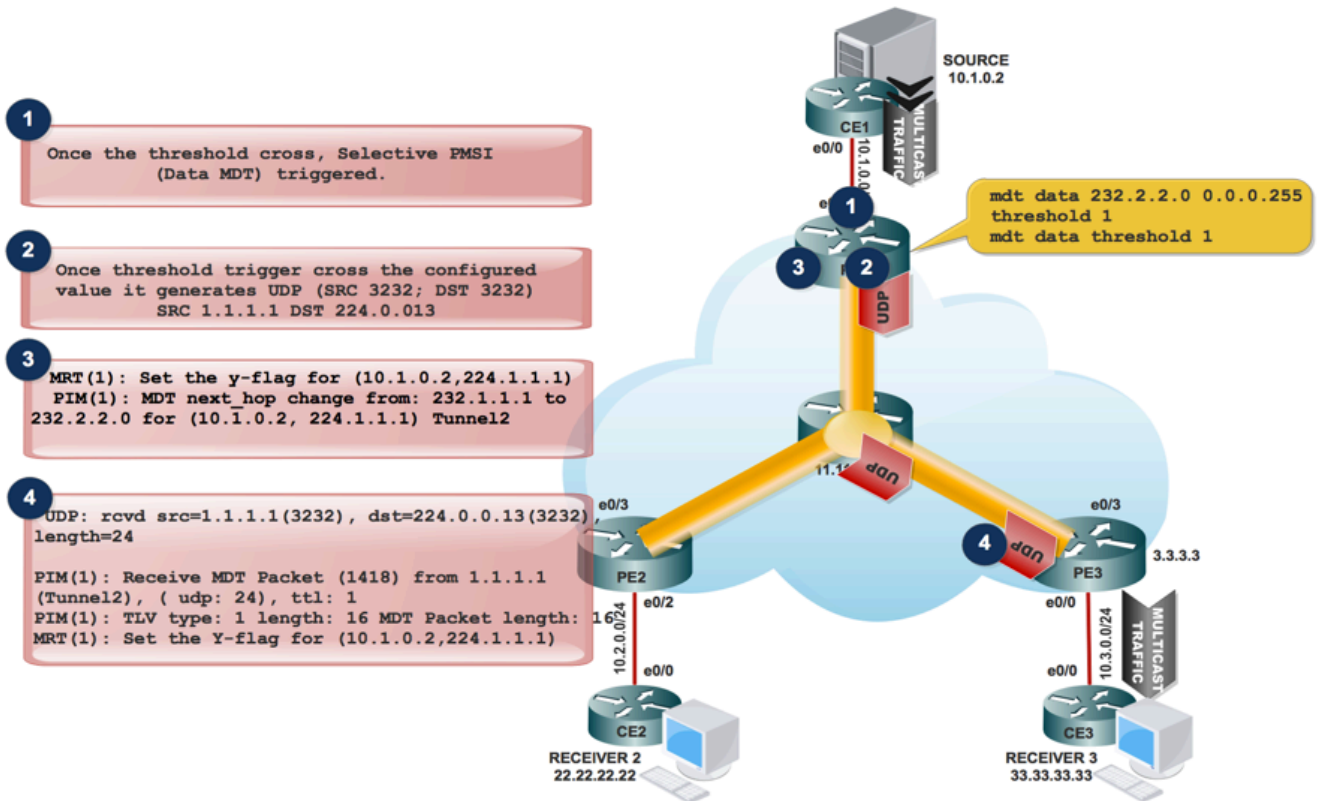

Data MDT:

What is Data MDT?

It is optional. It is created in demand, it carries specific (S,G) traffic. In the latest IOS® release, the threshold configured is as "0" and "infinite". Whenever a first packet hits the VRF, the Data MDT initializes, and if infinity then the Data MDT will never be created, and the traffic moves forward in the default MDT. The Data MDT is always the receiving tree, they never send any traffic. Data MDT is only for the (S,G) traffic.

Selective PMSI:
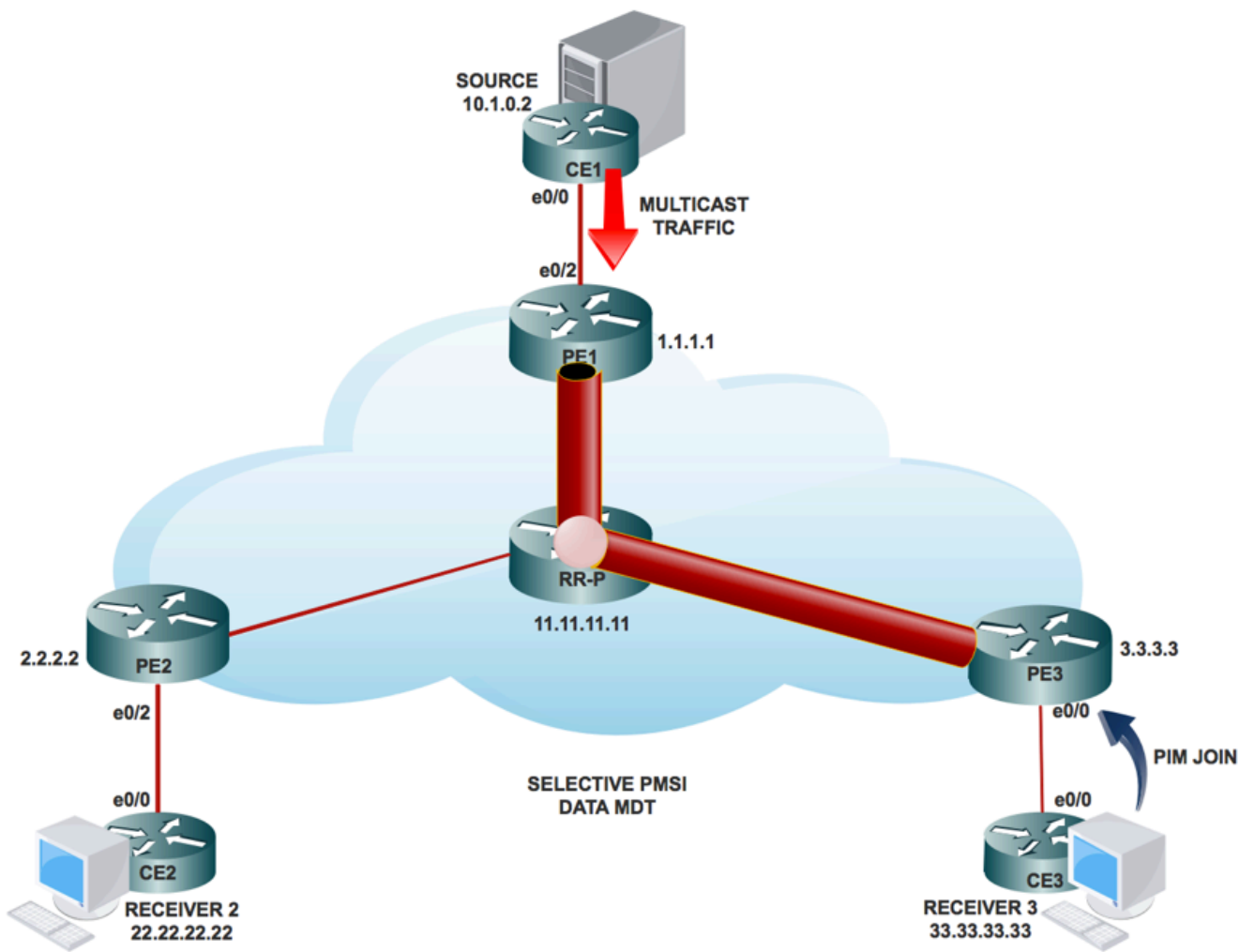
- It's optional. It is created in demand, it carries specific (S, G) traffic.
- Whenever a first packet hits the VRF, the Data MDT initializes, and if infinity then the Data MDT is never created, and the traffic moves forward in the default MDT.
- The Data MDT is always be the receiving tree, they never send any traffic. Data MDT is only for the (S, G) traffic.
- PIM message carries C- (S, G) & P-Group.

How DATA MDT is created:

1 Once the threshold cross, Selective PMSI (Data MDT) triggered.

2 Once threshold trigger cross the configured value it generates UDP (SRC 3232; DST 3232) SRC 1.1.1.1 DST 224.0.013

3 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

4 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), (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)

mdt data 232.2.2.0 0.0.0.255 threshold 1
mdt data threshold 1

1. When multicast traffic enters in the VRF and when the traffic rate reach the threshold. It Generates a MDT Packet.

2. The MDT packet get encapsulated in UDP with Source and Destination 3232. And send it to interested receiver.

```
     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. After it sends the UDP packet to interested receiver, it sets "y" flag and change the MDT next_hop to new MDT Group address.

At Source 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)
```

Note: The OIL Next Hop Changes to 232.2.2.0.

2. AT PE3, when it receives the MDT Packet encapsulated in UDP SRC port 3232 and DST
   port 3232.

```
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)
```
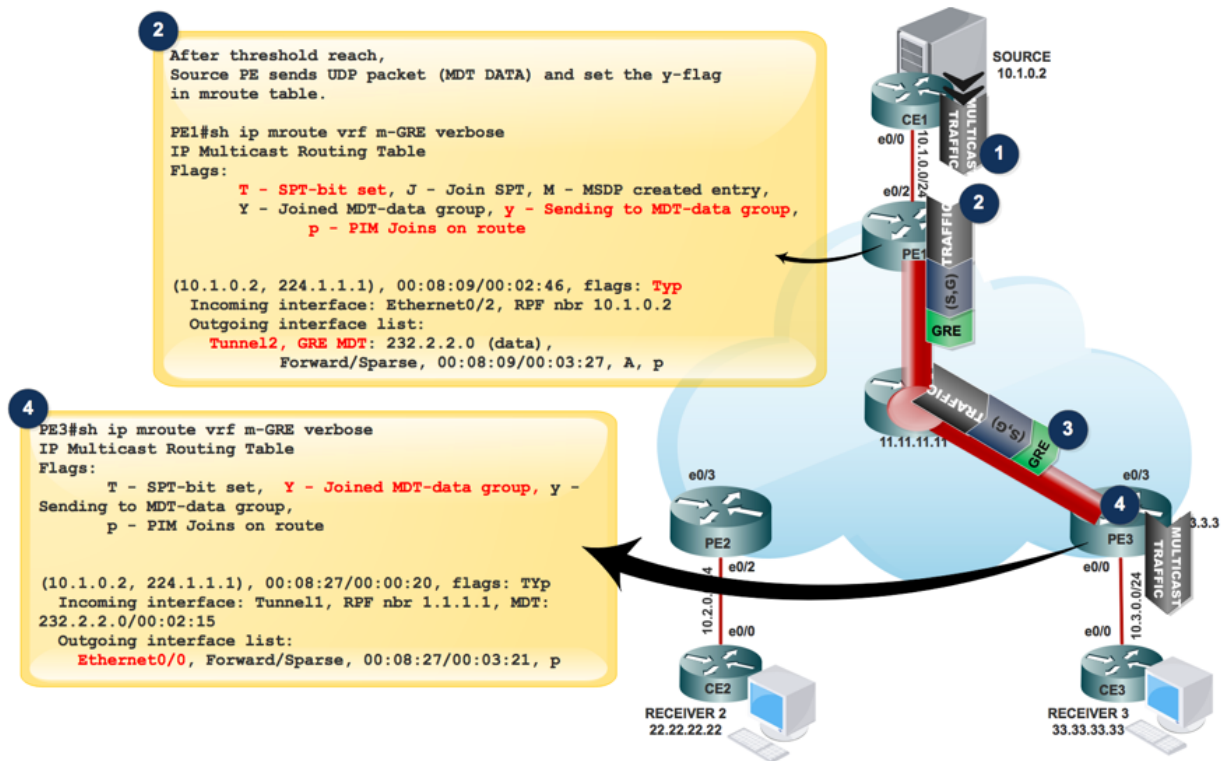The S-PMSI Join message is a UDP-encapsulated message whose destination address is ALL-
PIM-ROUTERS (224.0.0.13) and whose destination port is 3232.

The S-PMSI Join message contains this information: An identifier for the particular multicast
stream that is to be bound to the P-tunnel. This can be represented as an (S,G) pair. An identifier

for the particular P-tunnel to which the stream is to be bound. This identifier is a structured field that includes this information:

Multicast Traffic flow in MDT DATA Tunnel:

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)

- If the OIL contains a tunnel interface, then the packet is encapsulated with the use of GRE, with the source being the BGP peering address of the local PE router and the destination being the MDT Group address.

- The decision the Data-MDT group is selected depends on whether the y flag is set on the (S, G) entry in the mVRF.

- If the (S, G) or (*, G) entry has the Z flag set, then this is a Default- or Data-MDT with an associated mVRF.

- The P-packet must be de-encapsulated to reveal the C-packet.

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

- Because only a single MTI exists in the mVRF per multicast domain, both the Data-MDT and the Default-MDT use the same tunnel interface for customer traffic.

- The Y/y flags are necessary to distinguish Default-MDT traffic from Data-MDT traffic and ensure that customer multicast routing entries use the correct MDT-Data group and refer to an internal table that holds the (S, G, Data-MDT) mappings.

Troubleshoot

There is currently no specific troubleshooting information available for this configuration.