# RSVP Signaled Point-to-Multipoint LSPs

## In This Chapter

This section provides information about RSVP signaled point-to-multipoint LSPs.

Topics in this section include:

# Applicability

This feature is applicable to all of the 7750 and 7710 SR series (except SR-1). Tested on release 7.0R5. On all nodes involved with the LSP, at least chassis mode C is required. This means that modular systems should be equipped with IOM-2 linecards or higher. This is supported on 7450 ESS-7 or ESS-12 in mixed-mode since 8.0R1. The 7750 SR-c4 is supported from 8.0R4 and higher.

# Overview

Point-to-MultiPoint (P2MP) MPLS label switched path (LSP) allows the source of multicast traffic to forward packets to one or many multicast receivers over a network without requiring a multicast protocol, such as PIM, to be configured in the network. A P2MP LSP tree is established in the control plane which path consists of a head-end node, one or many branch and BUD nodes, and the leaf nodes. Packets injected by the head-end node are replicated in the data plane at the branching nodes before they are delivered to the leaf nodes.

Similar to point-to-point (P2P) LSPs, also P2MP LSPs are unidirectional, originating on a head-end node (the ingress LER) and terminating on one or more leaf node(s) (the egress LER(s)). Initially, RSVP is used as signaling protocol. A P2MP LSP is modeled as a set of root-to-leaf sub LSPs (S2L). Each S2L is modeled as a point-to-point LSP in the control plane. This means that each S2L has it's own PATH/RESV messages. This is called the de-aggregated method.

The forwarding of multicast packets to the LSP tree is based on static multicast routes initially but will evolve to BGP based VPN routes in the future. Forwarding multicast packets is initially done over P2MP RSVP LSPs in the base router instance but will evolve to VPRNs.

RSVP signalled P2MP LSPs can have fast reroute (FRR) enabled, the facility method (one-to-many) with link protection is supported.
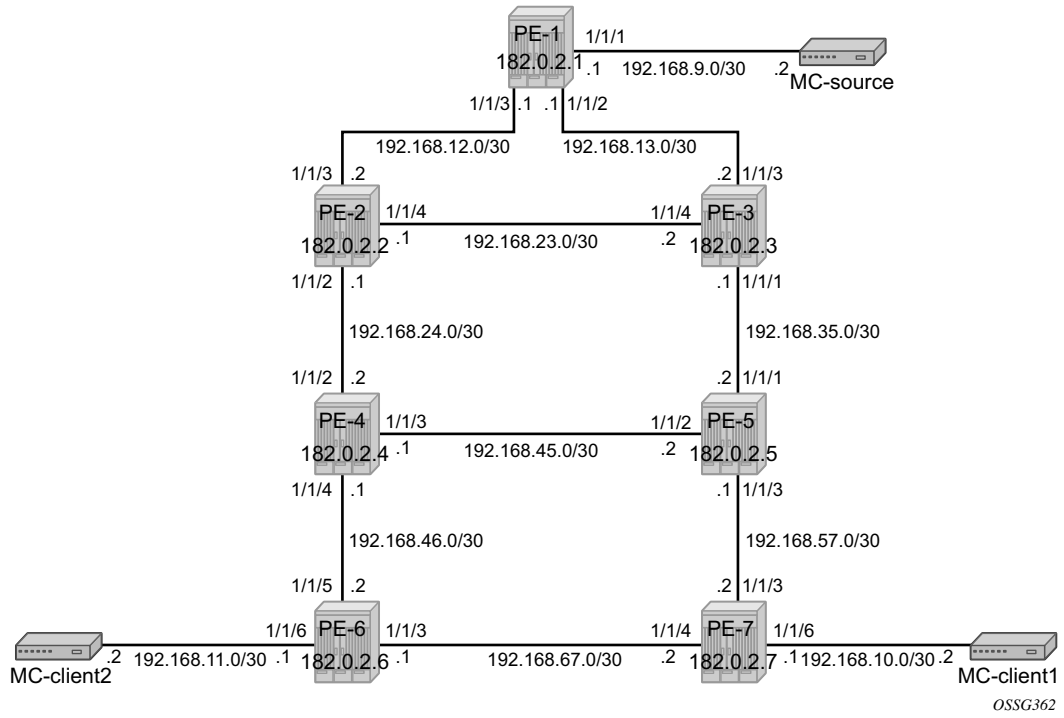
**Figure 169: P2MP Network Topology**

# Configuration

The following sections describe the tasks you must perform to configure RSVP signaled point-to-multipoint LSPs.

## Configuring the IP/MPLS Network

After configuring the cards and MDAs, the correct chassis-mode must be set (C or higher) on all MPLS nodes along the P2MP LSP.

```
A:PE-[1..7]# configure system chassis-mode c

A:PE-[1..7]# show chassis | match chassis mode
    Admin chassis mode         : c
    Oper chassis mode          : c
```

The system addresses and L3 interface addresses are configured according to Figure 169. An interior gateway protocol (IGP) is needed to distribute routing information on all PEs. In our case, the IGP is OSPF using the backbone area (area 0.0.0.0). A configuration example is shown for PE-1. A similar configuration is needed on all 7 PEs

```
A:PE-1# configure router
        interface "int-PE-1-PE-2"
            address 192.168.12.1/30
            port 1/1/3
        exit
        interface "int-PE-1-PE-3"
            address 192.168.13.1/30
            port 1/1/2
        exit
        interface "system"
            address 192.0.2.1/32
        exit

A:PE-1# configure router
        ospf
            traffic-engineering
            area 0.0.0.0
                interface "system"
                exit
                interface "int-PE-1-PE-3"
                    interface-type point-to-point
                exit
                interface "int-PE-1-PE-2"
                    interface-type point-to-point
                exit
            exit
        exit
```

Since fast reroute (FRR) will be enabled on the P2MP LSP, traffic engineering (TE) is needed on the IGP. By doing this, OSPF will generate opaque LSAs which are collected in a traffic engineering database (TED), separate from the traditional OSPF topology database. OSPF interfaces are setup as type *'point-to-point'* to improve convergence, no DR/BDR election process is done.[1]

In our example on PE-1, the interface towards the multicast source is configured as an IES service. This could have been on a router interface instead.

Similar IES services are configured on PE-7 and PE-6 for multicast client 1 and multicast client 2.

```
A:PE-1 # configure service
       …
        ies 1 customer 1 create
           interface "to-MC-source" create
               address 192.168.9.1/30
               sap 1/1/1 create
               exit
           exit
           no shutdown
        exit

A:PE-7# configure service
       …
        ies 1 customer 1 create
           interface "to-MC-client1" create
               address 192.168.10.1/30
               sap 1/1/6 create
               exit
           exit
           no shutdown
        exit

A:PE-6# configure service
       …
        ies 1 customer 1 create
           interface "to-MC-client2" create
               address 192.168.11.1/30
               sap 1/1/6 create
               exit
           exit
           no shutdown
        exit
```

---

1.    Convergence is out of the scope of this configuration note.

To verify that OSPF neighbors are up (state **Full**), **show router ospf neighbor** is performed. To check if L3 interface addresses/subnets are known on all PEs, **show router route-table** or **show router fib** *iom-card-slot* will display the content of the forwarding information base (FIB).

```
A:PE-1# show router ospf neighbor
===============================================================================
OSPF Neighbors
===============================================================================
Interface-Name               Rtr Id        State    Pri  RetxQ   TTL
-------------------------------------------------------------------------------
int-PE-1-PE-3                 192.0.2.3     Full     1    0       36
int-PE-1-PE-2                 192.0.2.2     Full     1    0       32
-------------------------------------------------------------------------------
No. of Neighbors: 2
===============================================================================

A:PE-1# show router route-table
===============================================================================
Route Table (Router: Base)
===============================================================================
Dest Prefix                             Type    Proto   Age        Pref
      Next Hop[Interface Name]                                   Metric
-------------------------------------------------------------------------------
192.0.2.1/32                            Local   Local   15d04h27m  0
      system                                                     0
192.0.2.2/32                            Remote  OSPF    01d22h28m  10
      192.168.12.2                                               10000
192.0.2.3/32                            Remote  OSPF    01d22h27m  10
      192.168.13.2                                               1000
192.0.2.4/32                            Remote  OSPF    01d22h24m  10
      192.168.13.2                                               3000
192.0.2.5/32                            Remote  OSPF    01d22h24m  10
      192.168.13.2                                               2000
192.0.2.6/32                            Remote  OSPF    01d22h23m  10
      192.168.13.2                                               4000
192.0.2.7/32                            Remote  OSPF    01d22h22m  10
      192.168.13.2                                               5000
192.168.12.0/30                         Local   Local   01d22h31m  0
      int-PE-1-PE-2                                              0
192.168.13.0/30                         Local   Local   01d22h31m  0
      int-PE-1-PE-3                                              0
192.168.23.0/30                         Remote  OSPF    01d22h27m  10
      192.168.13.2                                               11000
192.168.24.0/30                         Remote  OSPF    01d22h24m  10
      192.168.13.2                                               13000
192.168.35.0/30                         Remote  OSPF    01d22h26m  10
      192.168.13.2                                               2000
192.168.45.0/30                         Remote  OSPF    01d22h24m  10
      192.168.13.2                                               3000
192.168.46.0/30                         Remote  OSPF    01d22h24m  10
      192.168.13.2                                               4000
192.168.57.0/30                         Remote  OSPF    01d22h24m  10
      192.168.13.2                                               12000
192.168.67.0/30                         Remote  OSPF    01d22h23m  10
      192.168.13.2                                               5000
-------------------------------------------------------------------------------
No. of Routes: 16
===============================================================================
```

```
A:PE-1# show router fib 1
===============================================================================
FIB Display
===============================================================================
Prefix                                           Protocol
    NextHop
-------------------------------------------------------------------------------
192.0.2.1/32                                     LOCAL
    192.0.2.1 (system)
192.0.2.2/32                                     OSPF
    192.168.12.2 (int-PE-1-PE-2)
192.0.2.3/32                                     OSPF
    192.168.13.2 (int-PE-1-PE-3)
192.0.2.4/32                                     OSPF
    192.168.13.2 (int-PE-1-PE-3)
192.0.2.5/32                                     OSPF
    192.168.13.2 (int-PE-1-PE-3)
192.0.2.6/32                                     OSPF
    192.168.13.2 (int-PE-1-PE-3)
192.0.2.7/32                                     OSPF
    192.168.13.2 (int-PE-1-PE-3)
192.168.12.0/30                                  LOCAL
    192.168.12.0 (int-PE-1-PE-2)
192.168.13.0/30                                  LOCAL
    192.168.13.0 (int-PE-1-PE-3)
192.168.23.0/30                                  OSPF
    192.168.13.2 (int-PE-1-PE-3)
192.168.24.0/30                                  OSPF
    192.168.13.2 (int-PE-1-PE-3)
192.168.35.0/30                                  OSPF
    192.168.13.2 (int-PE-1-PE-3)
192.168.45.0/30                                  OSPF
    192.168.13.2 (int-PE-1-PE-3)
192.168.46.0/30                                  OSPF
    192.168.13.2 (int-PE-1-PE-3)
192.168.57.0/30                                  OSPF
    192.168.13.2 (int-PE-1-PE-3)
192.168.67.0/30                                  OSPF
    192.168.13.2 (int-PE-1-PE-3)
-------------------------------------------------------------------------------
Total Entries : 16
-------------------------------------------------------------------------------
===============================================================================
```

The next step in the process of setting up a P2MP LSP, is enabling our L3 interfaces in the MPLS and RSVP context on all involved PE nodes (PE-1 <=> PE-7). By default, the system interface is put automatically within the MPLS/RSVP context. When an interface is put in the MPLS context, 7x50 copies it also in the RSVP context. Explicit enabling of MPLS and RSVP context is done by the **no shutdown** command. Below you can find the MPLS/RSVP configuration for PE-1.
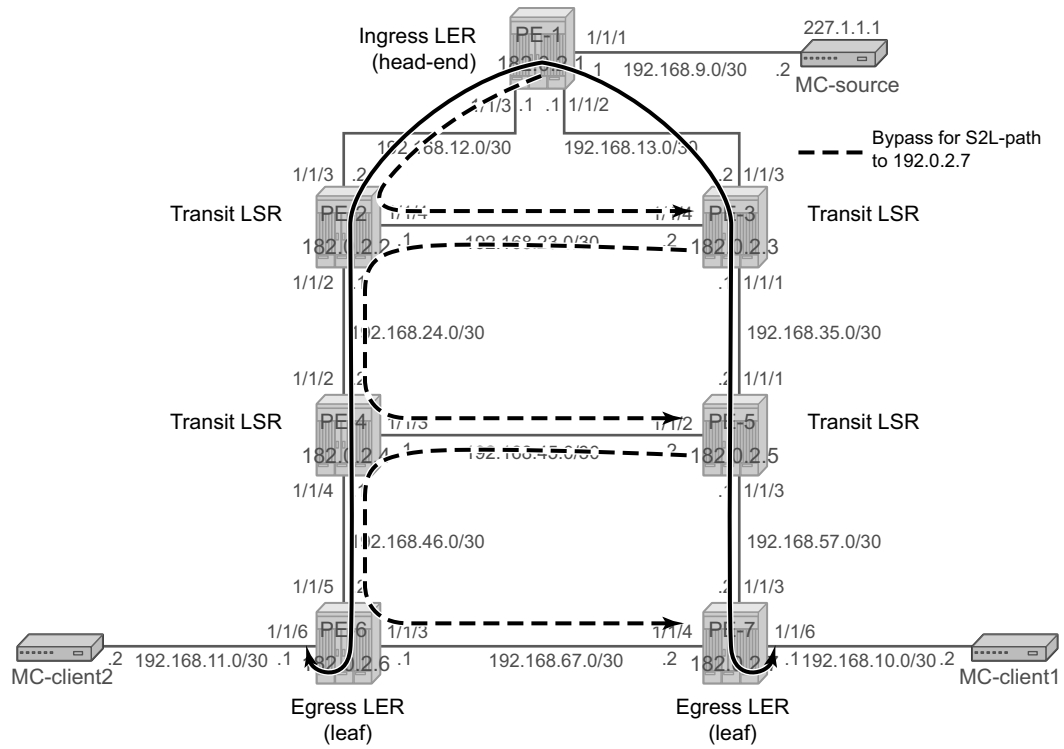
```
A:PE-1# configure router mpls no shutdown

A:PE-1# configure router rsvp no shutdown

A:PE-1# configure router mpls
            interface "system"
            exit
            interface "int-PE-1-PE-3"
            exit
            interface "int-PE-1-PE-2"
            exit
            no shutdown

A:PE-1# configure router rsvp
            interface "system"
            exit
            interface "int-PE-1-PE-3"
            exit
            interface "int-PE-1-PE-2"
            exit
            no shutdown
```

**Figure 170: P2MP LSP LSP-p2mp-1**

# Configuring P2MP RSVP LSP

See for a description of configuring PSMP LSPs.

A P2MP LSP (LSP-p2mp-1) will be setup from PE-1 acting as head-end node and PE-6 and PE-7 acting as leaf nodes. Because FRR is enabled, Constrained Shortest Path First (CSPF) is enabled to do route calculations on the traffic engineering database (TED). FRR method **facility** is used without node protection, **facility** stands for one-to-many meaning that one bypass tunnel can protect a set of primary LSPs with similar backup constraints. When a link failure occurs on one of the active S2L paths, the Point of Local Repair (PLR) node will push an additional MPLS label on the incoming MPLS packet before sending it into the bypass tunnel downstream towards the merge point (MP) node.

In the first example, OSPF (our IGP) will do the path calculation to the two destinations (PE-6 and PE-7). The intermediate hops of the LSP are dynamically assigned by OSPF best route selection, thus S2L paths follows the IGP least cost path. For this, an MPLS path **loose** is configured without specifying any strict/loose hops.

```
A:PE-1# configure router mpls
        …
        path "loose"
            no shutdown
        exit
```

Creation of the P2MP LSP itself is done on the ingress LER or head-end node (PE-1 in our example) and can be seen in following CLI output. P2MP name is **LSP-p2mp-1**. A create time keyword **p2mp-lsp** is added in addition to the P2MP name to make a distinction in configuration between normal point-to-point LSPs and point-to-multipoint LSPs. A primary P2MP instance is initiated using the **primary-p2mp-instance** keyword accompanied with the P2MP instance name p-LSP-p2mp-1. Within this primary P2MP instance, the different S2Ls are defined using the **s2l-path** keyword. Be aware that the same MPLS path name can be used for different S2Ls as long as the destination is different (**to** command).

```
A:PE-1# configure router mpls
        …
        lsp "LSP-p2mp-1" p2mp-lsp
            cspf
            fast-reroute facility
                no node-protect
            exit
            primary-p2mp-instance "p-LSP-p2mp-1"
                s2l-path "loose" to 192.0.2.6
                exit
                s2l-path "loose" to 192.0.2.7
                exit
            exit
            no shutdown
        exit
        no shutdown
```

On the head-end LER node of the P2MP LSP, several show commands can be used. First set of show commands are used to verify the administrative and operational state of the P2MP LSP and his different S2L paths (including FRR bypass information). In our example, 'LSP-p2mp-1' P2MP LSP has two active S2L paths towards leaf node PE-6 and leaf node PE-7

```
A:PE-1# show router mpls p2mp-lsp
===============================================================================
MPLS P2MP LSPs
===============================================================================
LSP Name                        P2mp-ID          Fastfail    Adm   Opr
                                                 Config
-------------------------------------------------------------------------------
LSP-p2mp-1                      0                Yes         Up    Up
-------------------------------------------------------------------------------
LSPs : 1
===============================================================================


A:PE-1# show router mpls p2mp-lsp LSP-p2mp-1 detail
===============================================================================
MPLS P2MP LSPs (Originating) (Detail)
===============================================================================
-------------------------------------------------------------------------------
Type : Originating
-------------------------------------------------------------------------------
LSP Name    : LSP-p2mp-1              LSP Tunnel ID : 11
From        : 192.0.2.1              To            : 0.0.0.0
Adm State   : Up                     Oper State    : Up²
LSP Up Time : 0d 00:08:20            LSP Down Time : 0d 00:00:00
Transitions : 1                      Path Changes  : 1
Retry Limit : 0                      Retry Timer   : 30 sec
Signaling   : RSVP                   Resv. Style   : SE
Hop Limit   : 255                    Negotiated MTU : n/a
Adaptive    : Enabled                ClassType     : 0
FastReroute : Enabled                Oper FR       : Enabled
FR Method   : Facility               FR Hop Limit  : 16
FR Bandwidth: 0 Mbps                 FR Node Protect: Disabled
FR Object   : Enabled
CSPF        : Enabled                ADSPEC        : Disabled
Metric      : Disabled               Use TE metric : Disabled
Include Grps:                        Exclude Grps  :
None                                 None
Type        : P2mpLsp                Least Fill    : Disabled
LdpOverRsvp : Disabled               VprnAutoBind  : Disabled
Oper Metric : Disabled

P2MPInstance: p-LSP-p2mp-1           P2MP-Inst-type : Primary
S2l-Name    : loose                  To            : 192.0.2.6
S2l-Name    : loose                  To            : 192.0.2.7
===============================================================================
```

_____

2.  As long as one S2L path is operationally up (**show router mpls p2mp-lsp** *lsp-name* **p2mp-instance** *instance-name*) , the **Oper State** of the P2MP LSP is **Up**.

```
A:PE-1# show router mpls p2mp-info
===============================================================================
MPLS P2MP Cross Connect Information
===============================================================================
-------------------------------------------------------------------------------
Source IP Address    : 192.0.2.1          Tunnel ID     : 11
P2MP ID              : 0                   Lsp ID        : 27648
S2L Name             : LSP-p2mp-1::loose   To            : 192.0.2.6
Out Interface        : 1/1/2              Out Label      : 131068
Num. of S2ls         : 2
-------------------------------------------------------------------------------
S2L LSP-p2mp-1::loose
-------------------------------------------------------------------------------
Source IP Address    : 192.0.2.1          Tunnel ID     : 11
P2MP ID              : 0                   Lsp ID        : 27648
S2L Name             : LSP-p2mp-1::loose   To            : 192.0.2.7
Out Interface        : 1/1/2              Out Label      : 131068
Num. of S2ls         : 2
-------------------------------------------------------------------------------
P2MP Cross-connect instances : 2
===============================================================================

A:PE-1# show router mpls p2mp-lsp LSP-p2mp-1 p2mp-instance p-LSP-p2mp-1
===============================================================================
MPLS P2MP Instance (Originating)
===============================================================================
-------------------------------------------------------------------------------
Type : Originating
-------------------------------------------------------------------------------
LSP Name    : LSP-p2mp-1              LSP Tunnel ID  : 11
P2MP ID     : 0
Adm State   : Up                      Oper State     : Up

P2MPInstance: p-LSP-p2mp-1            P2MP-Inst-type : Primary
P2MP Inst Id: 3                       P2MP Lsp Id    : 27648
Inst Admin  : Up                      Inst Oper      : Up
Inst Up Time: 0d 00:12:14             Inst Dn Time   : 0d 00:00:00
Hop Limit   : 255                     Adaptive       : Enabled
Record Route: Record                  Record Label   : Record
Include Grps:                         Exclude Grps   :
None                                  None
Bandwidth   : No Reservation          Oper Bw        : 0 Mbps
S2l-Name    : loose                   To             : 192.0.2.6
S2l Admin   : Up                      S2l Oper       : Up
S2l-Name    : loose                   To             : 192.0.2.7
S2l Admin   : Up                      S2l Oper       : Up
-------------------------------------------------------------------------------
P2MP instances : 1
===============================================================================
```

FRR information can be displayed in detail for each S2L path. From this moment onward, we will only focus on the S2L path towards PE-7. As you can see in the show command, link protection is present for links PE-1 <=> PE-3, PE-3 <=> PE-5 and PE-5 <=> PE-7 ( '@'-reference inside show command).

```
A:PE-1# show router mpls p2mp-lsp LSP-p2mp-1 p2mp-instance p-LSP-p2mp-1 s2l loose to
192.0.2.7 detail
===============================================================================
MPLS LSP LSP-p2mp-1 S2L loose (Detail)
===============================================================================
Legend :
    @ - Detour Available                     # - Detour In Use
    b - Bandwidth Protected                  n - Node Protected
    s - Soft Preemption
===============================================================================
-------------------------------------------------------------------------------
LSP LSP-p2mp-1 S2L loose
-------------------------------------------------------------------------------
LSP Name    : LSP-p2mp-1                      S2l LSP ID  : 27652
P2MP ID     : 0                               S2l Grp Id  : 2
Adm State   : Up                              Oper State  : Up
S2l State:  : Active                                      :
S2L Name    : loose                           To          : 192.0.2.7
S2l Admin   : Up                              S2l Oper    : Up
OutInterface: 1/1/2                           Out Label   : 131071
S2L Up Time : 0d 00:26:40                     S2L Dn Time : 0d 00:00:00
RetryAttempt: 0                               NextRetryIn : 0 sec
S2L Trans   : 3                               CSPF Queries: 3
Failure Code: noError                         Failure Node: n/a
ExplicitHops:
    No Hops Specified
Actual Hops :
    192.168.13.1(192.0.2.1) @                 Record Label    : N/A
 -> 192.168.13.2(192.0.2.3) @                 Record Label    : 131071
 -> 192.168.35.2(192.0.2.5) @                 Record Label    : 131071
 -> 192.168.57.2(192.0.2.7)                   Record Label    : 131068
ComputedHops:
    192.168.13.1   -> 192.168.13.2   -> 192.168.35.2   -> 192.168.57.2
LastResignal: n/a
===============================================================================
```

More in detail, **show router mpls bypass-tunnel** can be used. **Actual Hops** gives you the explicit hops of the bypass tunnel used to avoid link PE-1 <=> PE-3. On PE-1 node the MPLS path PE-1 <=> PE-2 <=> PE-3 is followed (see also ).

```
A:PE-1# show router mpls bypass-tunnel protected-lsp p2mp detail
===============================================================================
MPLS Bypass Tunnels (Detail)
===============================================================================
-------------------------------------------------------------------------------
bypass-link192.168.13.2
-------------------------------------------------------------------------------
To             : 192.168.23.2      State          : Up
Out I/F        : 1/1/3             Out Label      : 131064
Up Time        : 1d 23:25:14       Active Time    : n/a
```

```
Reserved BW    : 0 Kbps              Protected LSP Count : 3
Type           : P2mp
SetupPriority  : 7                   Hold Priority       : 0
Class Type     : 0
Actual Hops    :
    192.168.12.1   -> 192.168.12.2   -> 192.168.23.2

Protected LSPs -
LSP Name       : LSP-p2mp-1::loose
From           : 192.0.2.1           To                  : 192.0.2.7
Avoid Node/Hop : 192.168.13.2        Downstream Label    : 131071
Bandwidth      : 0 Kbps
```

On PE-3 node the MPLS path PE-3 <=> PE-2 <=> PE-4 <=> PE-5 is followed (see ) to avoid link PE-3 <=> PE-5.

```
A:PE-3# show router mpls bypass-tunnel protected-lsp p2mp detail
===============================================================================
MPLS Bypass Tunnels (Detail)
===============================================================================
-------------------------------------------------------------------------------
bypass-link192.168.35.2
-------------------------------------------------------------------------------
To             : 192.168.45.2        State               : Up
Out I/F        : 1/1/4               Out Label           : 131071
Up Time        : 1d 23:28:31         Active Time         : n/a
Reserved BW    : 0 Kbps              Protected LSP Count : 3
Type           : P2mp
SetupPriority  : 7                   Hold Priority       : 0
Class Type     : 0
Actual Hops    :
    192.168.23.2   -> 192.168.23.1   -> 192.168.24.2   -> 192.168.45.2

Protected LSPs -
LSP Name       : LSP-p2mp-1::loose
From           : 192.0.2.1           To                  : 192.0.2.7
Avoid Node/Hop : 192.168.35.2        Downstream Label    : 131071
Bandwidth      : 0 Kbps

===============================================================================
```

A similar output can be seen on PE-5 node also. The MPLS path PE-5 <=> PE-4 <=> PE-6 <=> PE-7 is followed (see also ).

On the transit LSRs and egress LER/leaf node (see also ), the **show router mpls p2mp-info** command can be used. Attached is the show command on PE-3 node included for S2L path to 192.0.2.7. Similar outputs are possible for PE-5 and PE-7 node.

```
A:PE-3# show router mpls p2mp-info type transit
===============================================================================
MPLS P2MP LSPs (Transit)
===============================================================================
-------------------------------------------------------------------------------
S2L LSP-p2mp-1::loose
-------------------------------------------------------------------------------
```

```
Source IP Address   : 192.0.2.1            Tunnel ID     : 11
P2MP ID             : 0                     Lsp ID        : 27652
S2L Name            : LSP-p2mp-1::loose     To            : 192.0.2.7
Out Interface       : 1/1/1                 Out Label     : 131071
Num. of S2ls        : 1
-------------------------------------------------------------------------------
P2MP Cross-connect instances : 1
===============================================================================
```

# Mapping Multicast Traffic

To map multicast traffic into the LSP tree from the head-end node until leaf node, PIM and IGMP configurations are needed on the head-end node (PE-1) and leaf nodes (PE-6 and PE-7) of the P2MP RSVP LSP. The intermediate nodes (transit LSR, branch LSR and BUD LSR) do not need any explicit configuration for that.

## Head-end Node (Ingress LER)

PIM must be enabled on the interface towards the MC source and PIM must be enabled on the tunnel interface. A tunnel interface should be seen as an internal representation of a specific P2MP LSP. Creation is done within the PIM context using the **tunnel-interface rsvp-p2mp** command followed by the P2MP LSP name. Translated into configuration commands, this becomes :

```
A:PE-1# configure router pim
        …
        interface "to-MC-source"
        exit
        tunnel-interface rsvp-p2mp LSP-p2mp-1
        …
```

In the data path, when a multicast packet is received on an interface a successful Reverse Path Forwarding (RPF) check must be done for the source address otherwise the packet will be dropped.

Besides enabling PIM on the tunnel interface also IGMP is enabled to do a static <S,G> or <*,G> join of a multicast group address (227.1.1.1 in our example) to the tunnel interface/P2MP LSP. Be aware that there is always a one-to-one mapping between <S,G> or <*,G> and a tunnel interface/ P2MP LSP. In our example a < S,G > will be configured. A <*,G> join scenario is included in .

```
A:PE-1# configure router igmp
      …
      tunnel-interface rsvp-p2mp "LSP-p2mp-1"
          static
              group 227.1.1.1
                  source 192.168.9.2
              exit
          exit
      exit
```

The **show router pim tunnel-interface** command shows you the admin state of the tunnel interface and an association to an internal local ifindex (**73730** in the example).

```
A:PE-1# show router pim tunnel-interface
===============================================================================
PIM Tunnel-Interfaces
===============================================================================
LSP                                      Sender Addr       IfIndex  Oper
-------------------------------------------------------------------------------
LSP-p2mp-1                               None              73730    Up
-------------------------------------------------------------------------------
Interfaces : 2
===============================================================================
```

With **show router igmp group** you see the configured <S,G> entry and outgoing interface (= tunnel interface), represented by mpls-if-73730.

```
A:PE-1# show router igmp group 227.1.1.1
===============================================================================
IGMP Groups
===============================================================================
(192.168.9.2,227.1.1.1)              Up Time : 0d 00:06:27
    Fwd List  : mpls-if-73730
-------------------------------------------------------------------------------
(*,G)/(S,G) Entries : 1
===============================================================================
```

At this moment in time, users can verify if multicast traffic is using P2MP LSP at the head-end node using the **show router pim group** *group-address* **detail** command

```
A:PE-1# show router pim group 227.1.1.1 detail
===============================================================================
PIM Source Group ipv4
===============================================================================
Group Address     : 227.1.1.1
Source Address    : 192.168.9.2
RP Address        : 192.0.2.1
Flags             : spt            Type               : (S,G)
MRIB Next Hop     : 192.168.9.2
MRIB Src Flags    : direct         Keepalive Timer Exp: 0d 00:03:29
Up Time           : 0d 00:08:59    Resolved By        : rtable-u

Up JP State       : Joined         Up JP Expiry       : 0d 00:00:00
Up JP Rpt         : Not Joined StarG  Up JP Rpt Override : 0d 00:00:00

Register State    : Join           Register Stop Exp  : 0d 00:00:00
Reg From Anycast RP: No

Rpf Neighbor      : 192.168.9.2
Incoming Intf     : to-MC-source
Outgoing Intf List : mpls-if-73730

Curr Fwding Rate  : 14.7 kbps
Forwarded Packets : 649            Discarded Packets  : 0
```

```
Forwarded Octets  : 29854              RPF Mismatches    : 0
Spt threshold     : 0 kbps            ECMP opt threshold : 7
Admin bandwidth   : 1 kbps
-------------------------------------------------------------------------------
Groups : 1
===============================================================================
```

# Leaf Node (Egress LER)

In the PIM context, the same tunnel interface must be created as the head-end node with in addition an explicit reference to the head-end system address, translated into the **sender** *systemIP_head-end_node* parameter.

```
A:PE-7# configure router pim
        tunnel-interface rsvp-p2mp LSP-p2mp-1 sender 192.0.2.1
```

The **show router pim tunnel-interface** command shows you the admin state of the tunnel interface and an association to an internal local ifindex (73730 in our example, by coincidence the same ifindex as the one on the head-end node/PE-1).

```
A:PE-7# show router pim tunnel-interface
===============================================================================
PIM Tunnel-Interfaces
===============================================================================
LSP                              Sender Addr       If Index  Oper
-------------------------------------------------------------------------------
LSP-p2mp-1                       192.0.2.1         73730     Up
-------------------------------------------------------------------------------
Interfaces : 1
===============================================================================
```

The main goal on the leaf node(s) is to get traffic off the P2MP LSP/tunnel interface. This is done using a multicast information policy (*multicast-info-policy*). Inside this MC policy, a range of multicast group addresses must be defined under a bundle context *(bundle1)* in order to see traffic *(channel)* . Also inside the bundle context, the P2MP LSP is presented by the tunnel interface *(primary-tunnel-interface)*. Translated into configuration commands, this becomes :

```
A:PE-7# configure mcast-management
        multicast-info-policy "p2mp-pol" create
            bundle "bundle1" create
                primary-tunnel-interface rsvp-p2mp LSP-p2mp-1 sender 192.0.2.1
                channel "227.1.1.1" "227.1.1.1" create
                exit
            exit
            bundle "default" create
            exit
        exit
```

**Note:** The channel command must be seen as a range command with a start-mc-group-address and an end-mc-group-address. In our example, only one MC group address, 227.1.1.1 is seen.

The configured multicast information policy must be applied to the base router instance.

```
A:PE-7# configure router multicast-info-policy p2mp-pol
```

On the leaf node (PE-7/PE-6), MC clients are connected. IGMP is enabled on those MC clients with a static <S,G> join to redirect MC traffic downstream to the MC client. Translated into configuration commands, this becomes :

```
A:PE-7 # configure router igmp
        …
        interface "to-MC-client1"
            static
                group 227.1.1.1
                    source 192.168.9.2
                exit
            exit
        exit
```

With **show router igmp group** you see the configured <S,G> entry and outgoing interface (= to-MC-client1).

```
A:PE-7# show router igmp group 227.1.1.1
===============================================================================
IGMP Groups
===============================================================================
(192.168.9.2,227.1.1.1)              Up Time : 0d 00:00:43
    Fwd List  : to-MC-client1
-------------------------------------------------------------------------------
(*,G)/(S,G) Entries : 1
===============================================================================
```

Now, users can verify if multicast traffic is sent to MC client using the **show router pim group** *group-address* **detail** command

```
A:PE-7# show router pim group 227.1.1.1 detail
===============================================================================
PIM Source Group ipv4
===============================================================================
Group Address     : 227.1.1.1
Source Address    : 192.168.9.2
RP Address        : 0
Flags             :                    Type            : (S,G)
MRIB Next Hop     :
MRIB Src Flags    : remote             Keepalive Timer : Not Running
Up Time           : 0d 00:27:43        Resolved By     : unresolved

Up JP State       : Joined             Up JP Expiry    : 0d 00:00:17
```

```
Up JP Rpt           : Not Joined StarG   Up JP Rpt Override : 0d 00:00:00

Register State      : No Info
Reg From Anycast RP: No

Rpf Neighbor        :
Incoming Intf       : mpls-if-73730
Outgoing Intf List  : to-MC-client1

Curr Fwding Rate    : 14.7 kbps
Forwarded Packets   : 66340            Discarded Packets  : 0
Forwarded Octets    : 3051640          RPF Mismatches     : 0
Spt threshold       : 0 kbps           ECMP opt threshold : 7
Admin bandwidth     : 1 kbps
-------------------------------------------------------------------------------
Groups : 1
===============================================================================
```

# OAM Tool

P2P LSP operation and maintenance (OAM) commands (**oam lsp-ping** and **oam lsp-trace**) are extended for P2MP LSP. The user can instruct the head-end node to generate an P2MP LSP ping or a P2MP LSP trace by entering the command **oam p2mp-lsp-ping** or **oam p2mp-lsp-trace**. The P2MP OAM extensions are defined in **draft-ietf-mpls-p2mp-lsp-ping**.

For P2MP LSP ping, the echo request is sent on the active P2MP instance and is replicated in the data path over all branches of the P2MP LSP instance. By default, all egress LER nodes which are leaves of the P2MP LSP instance will reply. Echo reply messages can be reduced by configuring the **s2l-dest-address** (a maximum of up to five egress nodes in a single run of the OAM command). Replies are sent by IP.

```
A:PE-1# oam p2mp-lsp-ping
 - p2mp-lsp-ping <lsp-name> [p2mp-instance <instance-name> [s2l-dest-address
   <ip-address> [...(upto 5 max)]]] [fc <fc-name> [profile {in|out}]] [size
   <octets>] [ttl <label-ttl>] [timeout <timeout>] [detail]

 <lsp-name>          : [32 chars max]
 <instance-name>     : [32 chars max]
 <ip-address>        : ipv4 address    a.b.c.d - Up to 5 addresses permitted
 <profile>           : in|out - Default: out
 <fc-name>           : be|l2|af|l1|h2|ef|h1|nc - Default: be
 <octets>            : [92|97..9198]- Default: 92
 <label-ttl>         : [1..255] - Default: 255
 <timeout>           : [1..120] seconds - Default: 10
 <detail>            : keyword - displays detailed information

A:PE-1# oam p2mp-lsp-ping LSP-p2mp-1 detail
P2MP LSP LSP-p2mp-1: 92 bytes MPLS payload

===============================================================================
S2L Information
===============================================================================
From            RTT                 Return Code
-------------------------------------------------------------------------------
192.0.2.6       =3.35ms             EgressRtr(3)
192.0.2.7       =6.14ms             EgressRtr(3)
===============================================================================

Total S2L configured/up/responded = 2/2/2,
        round-trip min/avg/max  = 3.35 / 4.74 / 6.14 ms

Responses based on return code:
      EgressRtr(3)=2
```

**Note:** Return codes are based on RFC4379, with value 3 => replying router is an egress for the FEC at stack-depth.

P2MP LSP trace allows the user to trace the path of a single S2L path of a P2MP LSP from head-end node to leaf node. By the use of the downstream mapping TLV, each node along the S2L path

can fill in the appropriate flags : B or E flag. B-flag is set when responding node is a branch LSR, E-flag is set when responding node is an egress LER.

```
A:PE-1# oam p2mp-lsp-trace
 - p2mp-lsp-trace <lsp-name> p2mp-instance <instance-name> s2l-dest-address
   <ip-address> [fc <fc-name> [profile {in|out}]] [size <octets>] [max-fail
   <no-response-count>] [probe-count <probes-per-hop>] [min-ttl
   <min-label-ttl>] [max-ttl <max-label-ttl>] [timeout <timeout>]  [interval
   <interval>] [detail]

<lsp-name>          : [32 chars max]
<instance-name>     : [32 chars max]
<ip-address>        : ipv4 address    a.b.c.d
<fc-name>           : be|l2|af|l1|h2|ef|h1|nc - Default: be
<profile>           : in|out - Default: out
<octets>            : [128..9198] - Default: 128
<no-response-count> : [1..10] - Default: 5
<probes-per-hop>    : [1..10] - Default: 1
<min-label-ttl>     : [1..255] - Default: 1
<max-label-ttl>     : [1..255] - Default: 30
<timeout>           : [1..60] seconds
<detail>            : keyword - displays detailed information

A:PE-1# oam p2mp-lsp-trace LSP-p2mp-1 p2mp-instance p-LSP-p2mp-1 s2l-dest-address
192.0.2.7 detail
P2MP LSP LSP-p2mp-1: 128 bytes MPLS payload
P2MP Instance p-LSP-p2mp-1, S2L Egress 192.0.2.7

  1  192.0.2.3  rtt=4.87 ms rc=8(DSRtrMatchLabel)
     DS 1: IfAddr 192.168.35.2 MRU=1500 label=131064 proto=4(RSVP-TE) B/E flags:0/0
  2  192.0.2.5  rtt=8.02 ms rc=8(DSRtrMatchLabel)
     DS 1: IfAddr 192.168.57.2 MRU=1500 label=131066 proto=4(RSVP-TE) B/E flags:0/0
  3  192.0.2.7  rtt=4.44 ms rc=3(EgressRtr)
```

**Note:** Return codes are based on RFC4379, with value '8' => label switched at stack-depth. Meaning it is an transit LSR, doing MPLS label swapping. No B or E flag is set.

# Additional Topics

## <*,G> IGMP join iso <S,G> IGMP join

In the Head-end Node (Ingress LER) on page 1199 and Leaf Node (Egress LER) on page 1201 steps, a source specific IGMP join (<S,G> join) was used at the head-end node and leaf nodes. Another possibility is to use a source unknown or *starg* IGMP join (<*,G> join). When doing the latter, a rendezvous point (RP) must be defined in the PIM network. The RP allows multicast data flows between sources and receivers to meet at a predefined network location (in our example, loopback address of PE-1 node). It must be seen as an intermediate device to establish a multicast flow.

RP can be defined in a dynamic way (BSR protocol) or a static way. In our example we choose the static way meaning that on all involved PIM nodes, the RP address will be statically configured. The following output shows translated in configuration commands.

```
A:PE-1/PE-6/PE-7# configure router pim
        …
        rp
            static
                address 192.0.2.1
                    group-prefix 227.1.1.1/32
                exit
            exit
            …
        exit
```

The **group-prefix** a mandatory keyword is specified and references a group address or group address range for which this rendez-vous point will be used.

```
A:PE-1/PE-6/PE-7# show router pim rp
===============================================================================
PIM RP Set ipv4
===============================================================================
Group Address                                             Hold Expiry
  RP Address                                    Type    Prio Time Time
-------------------------------------------------------------------------------
227.1.1.1/32
  192.0.2.1                                     Static   1   N/A  N/A
-------------------------------------------------------------------------------
Group Prefixes : 1
===============================================================================
```

As previously mentioned, the configuration of the <*,G> IGMP join is done on the head-end node (PE-1) and leaf nodes (PE-6 and PE-7)

```
A:PE-1# configure router igmp
        tunnel-interface rsvp-p2mp "p-to-mp-1"
            static
```

```
            group 227.1.1.1
                starg
            exit
        exit
    exit

A:PE-6# configure router igmp
    interface "to-MC-client2"
        static
            group 227.1.1.1
                starg
            exit
        exit
    exit

A:PE-7# configure router igmp
    interface "to-MC-client1"
        static
            group 227.1.1.1
                starg
            exit
        exit
    exit
```

The same previous **show** command can be used to verify the multicast traffic on head-end node and leaf nodes, **show router igmp group 227.1.1.1** and **show router pim group 227.1.1.1 detail**.

```
A:PE-7# show router pim group 227.1.1.1 detail
===============================================================================
PIM Source Group ipv4
===============================================================================
Group Address       : 227.1.1.1
Source Address      : *
RP Address          : 192.0.2.1
Flags               :                      Type              : (*,G)
MRIB Next Hop       :
MRIB Src Flags      : remote               Keepalive Timer   : Not Running
Up Time             : 0d 00:05:29          Resolved By       : unresolved

Up JP State         : Joined               Up JP Expiry      : 0d 00:00:30
Up JP Rpt           : Not Joined StarG     Up JP Rpt Override : 0d 00:00:00

Rpf Neighbor        :
Incoming Intf       : mpls-if-73730
Outgoing Intf List  : to-MC-client1

Curr Fwding Rate    : 14.7 kbps
Forwarded Packets   : 12163               Discarded Packets  : 0
Forwarded Octets    : 559498              RPF Mismatches     : 0
Spt threshold       : 0 kbps              ECMP opt threshold : 7
Admin bandwidth     : 1 kbps
-------------------------------------------------------------------------------
Groups : 1
===============================================================================
```

## Influence IGP Metric

Suppose that the IGP metric is increased on all links pointing to/from PE-2 and on the link between PE-5 andPE-7.

```
A:PE-1# configure router ospf area 0 interface int-PE-1-PE-2 metric 10000

A:PE-2# configure router ospf area 0 interface int-PE-2-PE-4 metric 10000
A:PE-2# configure router ospf area 0 interface int-PE-2-PE-3 metric 10000
A:PE-2# configure router ospf area 0 interface int-PE-2-PE-1 metric 10000

A:PE-3# configure router ospf area 0 interface int-PE-3-PE-2 metric 10000

A:PE-4# configure router ospf area 0 interface int-PE-4-PE-2 metric 10000

A:PE-5# configure router ospf area 0 interface int-PE-5-PE-7 metric 10000

A:PE-7# configure router ospf area 0 interface int-PE-7-PE-5 metric 10000
```

The existing P2MP LSP *LSP-p2mp-1* will not take into account these new constraints. The two S2L paths (one *loose* towards PE-6 and another one *loose* towards PE-7) are calculated using the default OSPF metric of 1000. What we can do to trigger MPLS to re-compute the S2L paths, is configure a p2mp-resignal-timer on the head-end node inside the global MPLS context. In this way, each time this timer expires (in our example, every 60 minutes), MPLS will trigger CSPF to re-compute the whole set of S2L paths of all active P2MP instance. MPLS performs a global make-before-break (MBB) and moves each S2L sub-LSP in the instance into its new path using a new P2MP LSP ID if the global MBB is successful. **show router mpls status** gives you an indication when the P2MP resignal timer will expire and which types of LSPs are setup on the node.

```
A:PE-1# configure router mpls p2mp-resignal-timer 60
A:PE-1# show router mpls status
===============================================================================
MPLS Status
===============================================================================
Admin Status      : Up                Oper Status        : Up
Oper Down Reason  : n/a
FR Object         : Enabled           Resignal Timer     : Disabled
Hold Timer        : 1 seconds         Next Resignal      : N/A
Srlg Frr          : Disabled          Srlg Frr Strict    : Disabled
Dynamic Bypass    : Enabled           User Srlg Database : Disabled
Least Fill Min Thd.: 5 percent        LeastFill ReoptiThd: 10 percent

P2mp Resignal Timer: 60 minutes       P2mp Next Resignal : 6 minutes
Sec FastRetryTimer : Disabled

LSP Counts          Originate         Transit           Terminate
-------------------------------------------------------------------------------
…
===============================================================================
```

As an alternative the user can also perform a manual resignal of a P2MP instance on the head-end node using a tools command.

```
A:PE-1# tools perform router mpls resignal p2mp-lsp LSP-p2mp-1 p2mp-instance p-p-to-mp-1

A:PE-1# tools perform router mpls resignal p2mp-delay 0
```
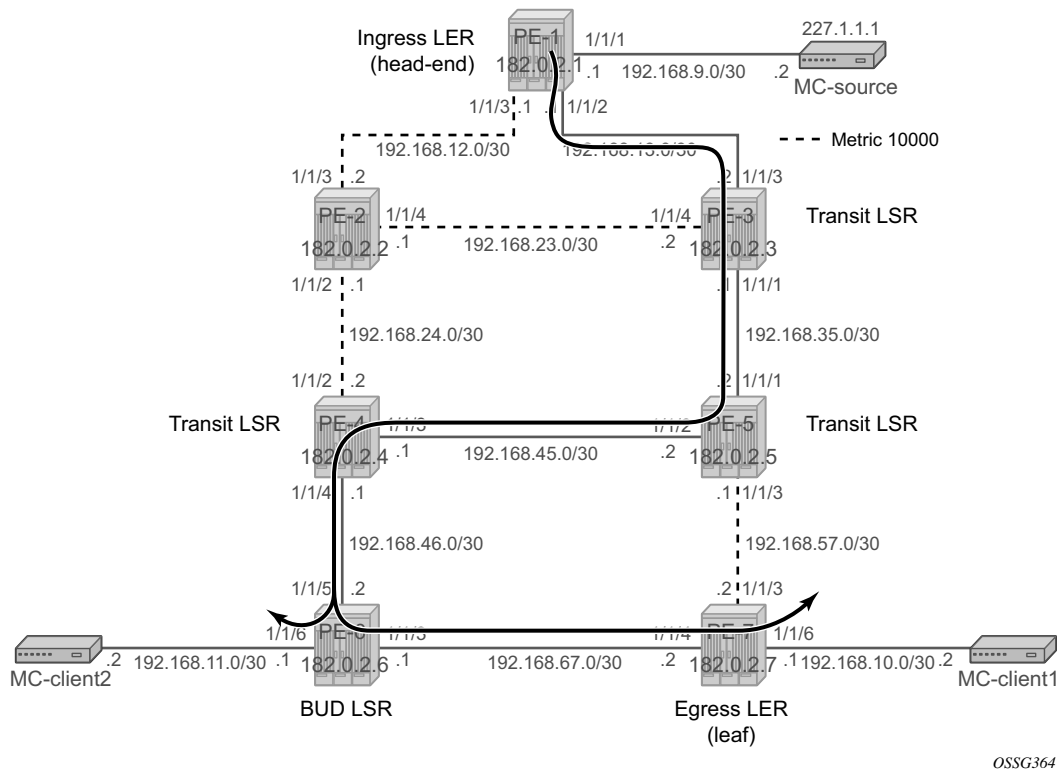


**Figure 171: P2MP LSP p-to-mp-1 with Metric Change**

After an instantaneous **tools resignal** command with executed with no delay (**p2mp-delay 0**), the S2L paths can be verified will be according to Figure 171. PE-6 node is acting now as BUD LSR node (iso egress LER before)

```
A:PE-1# show router mpls p2mp-lsp LSP-p2mp-1 p2mp-instance p-LSP-p2mp-1 s2l loose to
192.0.2.7 detail
…
LSP Name    : LSP-p2mp-1                    S2l LSP ID  : 27658
P2MP ID     : 0                             S2l Grp Id  : 2
Adm State   : Up                            Oper State  : Up
S2l State:  : Active                                    :
S2L Name    : loose                         To          : 192.0.2.7
S2l Admin   : Up                            S2l Oper    : Up
OutInterface: 1/1/2                          Out Label   : 131065
S2L Up Time : 0d 02:14:31                   S2L Dn Time : 0d 00:00:00
```

```
RetryAttempt: 0                                  NextRetryIn : 0 sec
S2L Trans   : 6                                  CSPF Queries: 6
Failure Code: noError                            Failure Node: n/a
ExplicitHops:
   No Hops Specified
Actual Hops :
    192.168.13.1(192.0.2.1) @                    Record Label   : N/A
 -> 192.168.13.2(192.0.2.3) @                    Record Label   : 131065
 -> 192.168.35.2(192.0.2.5) @                    Record Label   : 131062
 -> 192.168.45.1(192.0.2.4) @                    Record Label   : 131061
 -> 192.168.46.2(192.0.2.6) @                    Record Label   : 131065
 -> 192.168.67.2(192.0.2.7)                      Record Label   : 131063
ComputedHops:
    192.168.13.1   -> 192.168.13.2   -> 192.168.35.2    -> 192.168.45.1
 -> 192.168.46.2    -> 192.168.67.2
LastResignal: n/a
===================================================================
A:PE-1#

A:PE-1# show router mpls p2mp-lsp LSP-p2mp-1 p2mp-instance p-LSP-p2mp-1 s2l loose to
192.0.2.6 detail
…
LSP Name    : LSP-p2mp-1                         S2l LSP ID  : 27658
P2MP ID     : 0                                  S2l Grp Id  : 1
Adm State   : Up                                 Oper State  : Up
S2l State:  : Active                                          :
S2L Name    : loose                              To          : 192.0.2.6
S2l Admin   : Up                                 S2l Oper    : Up
OutInterface: 1/1/2                              Out Label   : 131065
S2L Up Time : 0d 02:22:00                        S2L Dn Time : 0d 00:00:00
RetryAttempt: 0                                  NextRetryIn : 0 sec
S2L Trans   : 6                                  CSPF Queries: 6
Failure Code: noError                            Failure Node: n/a
ExplicitHops:
   No Hops Specified
Actual Hops :
    192.168.13.1(192.0.2.1) @                    Record Label   : N/A
 -> 192.168.13.2(192.0.2.3) @                    Record Label   : 131065
 -> 192.168.35.2(192.0.2.5) @                    Record Label   : 131062
 -> 192.168.45.1(192.0.2.4) @                    Record Label   : 131061
 -> 192.168.46.2(192.0.2.6)                      Record Label   : 131065
ComputedHops:
    192.168.13.1   -> 192.168.13.2   -> 192.168.35.2    -> 192.168.45.1
 -> 192.168.46.2
LastResignal: n/a
===================================================================
A:PE-1#
```

An **oam p2mp-lsp-trace** command toward PE-7 will now set the E flag on PE-6 since that PE acts also as an egress LER node.

```
A:PE-1# oam p2mp-lsp-trace LSP-p2mp-1 p2mp-instance p-LSP-p2mp-1 s2l-dest-address
192.0.2.7 detail
P2MP LSP LSP-p2mp-1: 128 bytes MPLS payload
P2MP Instance p-LSP-p2mp-1, S2L Egress 192.0.2.7
```

```
1  192.0.2.3  rtt=3.08 ms rc=8(DSRtrMatchLabel)
   DS 1: IfAddr 192.168.35.2 MRU=1500 label=131062 proto=4(RSVP-TE) B/E flags:0/0
2  192.0.2.5  rtt=3.56 ms rc=8(DSRtrMatchLabel)
   DS 1: IfAddr 192.168.45.1 MRU=1500 label=131061 proto=4(RSVP-TE) B/E flags:0/0
3  192.0.2.4  rtt=11.6 ms rc=8(DSRtrMatchLabel)
   DS 1: IfAddr 192.168.46.2 MRU=1500 label=131065 proto=4(RSVP-TE) B/E flags:0/0
4  192.0.2.6  rtt=4.84 ms rc=8(DSRtrMatchLabel)
   DS 1: IfAddr 192.168.67.2 MRU=1500 label=131063 proto=4(RSVP-TE) B/E flags:0/1
5  192.0.2.7  rtt=6.05 ms rc=3(EgressRtr)
```

As a next step, the S2L path towards PE-7 is changed from **loose** to a **strict** direct MPLS path (**strict-to-PE-7***)*. In that way, OSPF is not calculating anymore the shortest path to the leaf node.

```
A:PE-1# configure router mpls
        …
        path "strict-to-PE-7"
            hop 10 192.168.13.2 strict
            hop 20 192.168.35.2 strict
            hop 30 192.168.57.2 strict
            no shutdown
        exit
```

Before applying this new S2L path to the existing P2MP LSP (*LSP-p2mp-1*), the existing S2L path towards PE-7 must be removed.

```
A:PE-1# configure router mpls lsp "LSP-p2mp-1" primary-p2mp-instance "p-LSP-p2mp-1" s2l-
path "loose" to 192.0.2.7 shutdown

A:PE-1# configure router mpls lsp "LSP-p2mp-1" primary-p2mp-instance "p-LSP-p2mp-1" no
s2l-path "loose" to 192.0.2.7

A:PE-1# configure router mpls lsp "LSP-p2mp-1" primary-p2mp-instance "p-LSP-p2mp-1" s2l-
path "strict-to-PE-7" to 192.0.2.7
```
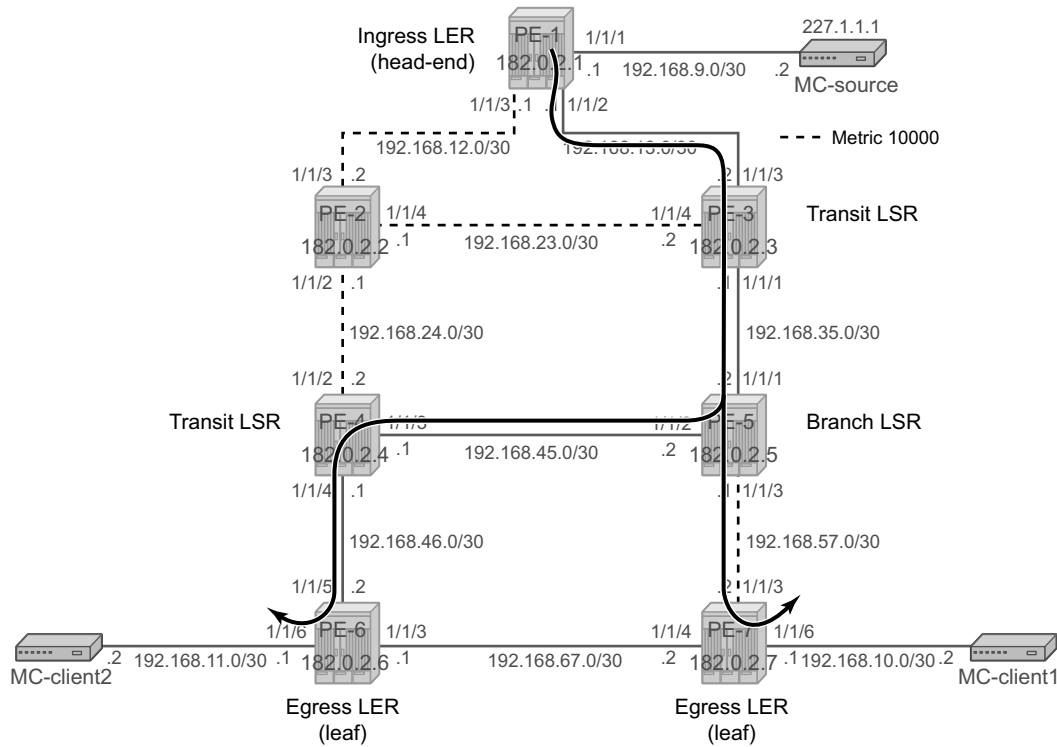
*OSSG365*

**Figure 172: P2MP LSP LSP-p2mp-1 with Strict S2L Path Towards PE-7**

As a consequence of this, only the **S2l Grp Id** has changed while **S2L LSP ID** remains the same
as before. Now, S2L paths can be verified according to Figure 172. PE-5 is acting now as branch
LSR node (iso transit LSR before).

```
A:PE-1# show router mpls p2mp-lsp LSP-p2mp-1 p2mp-instance p-LSP-p2mp-1 s2l strict-to-PE-7
to 192.0.2.7 detail
===============================================================================
MPLS LSP LSP-p2mp-1 S2L strict-to-PE-7 (Detail)
===============================================================================
Legend :
    @ - Detour Available                      # - Detour In Use
    b - Bandwidth Protected                   n - Node Protected
    s - Soft Preemption
===============================================================================
-------------------------------------------------------------------------------
LSP LSP-p2mp-1 S2L strict-to-PE-7
-------------------------------------------------------------------------------
LSP Name    : LSP-p2mp-1                       S2l LSP ID : 27658
P2MP ID     : 0                                S2l Grp Id : 3
Adm State   : Up                               Oper State : Up
S2l State:  : Active                                      :
S2L Name    : strict-to-PE-7                   To         : 192.0.2.7
S2l Admin   : Up                               S2l Oper   : Up
OutInterface: 1/1/2                            Out Label  : 131065
S2L Up Time : 0d 00:01:04                      S2L Dn Time : 0d 00:00:00
```

```
RetryAttempt: 0                                    NextRetryIn : 0 sec
S2L Trans   : 1                                    CSPF Queries: 1
Failure Code: noError                              Failure Node: n/a
ExplicitHops:
    192.168.13.2   -> 192.168.35.2   -> 192.168.57.2
Actual Hops :
    192.168.13.1(192.0.2.1) @                   Record Label    : N/A
 -> 192.168.13.2(192.0.2.3) @                   Record Label    : 131065
 -> 192.168.35.2(192.0.2.5) @                   Record Label    : 131062
 -> 192.168.57.2(192.0.2.7)                     Record Label    : 131068
ComputedHops:
    192.168.13.1   -> 192.168.13.2   -> 192.168.35.2   -> 192.168.57.2
LastResignal: n/a
===============================================================================
```

An **oam p2mp-lsp-trace** command towards PE-7 will now set the B flag on PE-5 since that
became a branch LSR now.

```
A:PE-1# oam p2mp-lsp-trace LSP-p2mp-1 p2mp-instance p-LSP-p2mp-1 s2l-dest-address
192.0.2.7 detail
P2MP LSP LSP-p2mp-1: 128 bytes MPLS payload
P2MP Instance p-LSP-p2mp-1, S2L Egress 192.0.2.7

  1  192.0.2.3  rtt=2.67 ms rc=8(DSRtrMatchLabel)
     DS 1: IfAddr 192.168.35.2 MRU=1500 label=131062 proto=4(RSVP-TE) B/E flags:0/0
  2  192.0.2.5  rtt=3.35 ms rc=8(DSRtrMatchLabel)
     DS 1: IfAddr 192.168.57.2 MRU=1500 label=131068 proto=4(RSVP-TE) B/E flags:1/0
  3  192.0.2.7  rtt=5.35 ms rc=3(EgressRtr)
```

# Intelligent Remerge

Intelligent remerge protects users from receiving duplicate multicast during convergence. It also protects against duplicate traffic in case of badly designed S2L paths. Initially, three cases exist for which intelligent remerge is implemented.

**Case 1**

When the paths of two different S2Ls of the same P2MP LSP instance have Ingress Label Maps (ILMs) on different ports but go out on the same Next-hop Label Forwarding Entry (NHLFE).
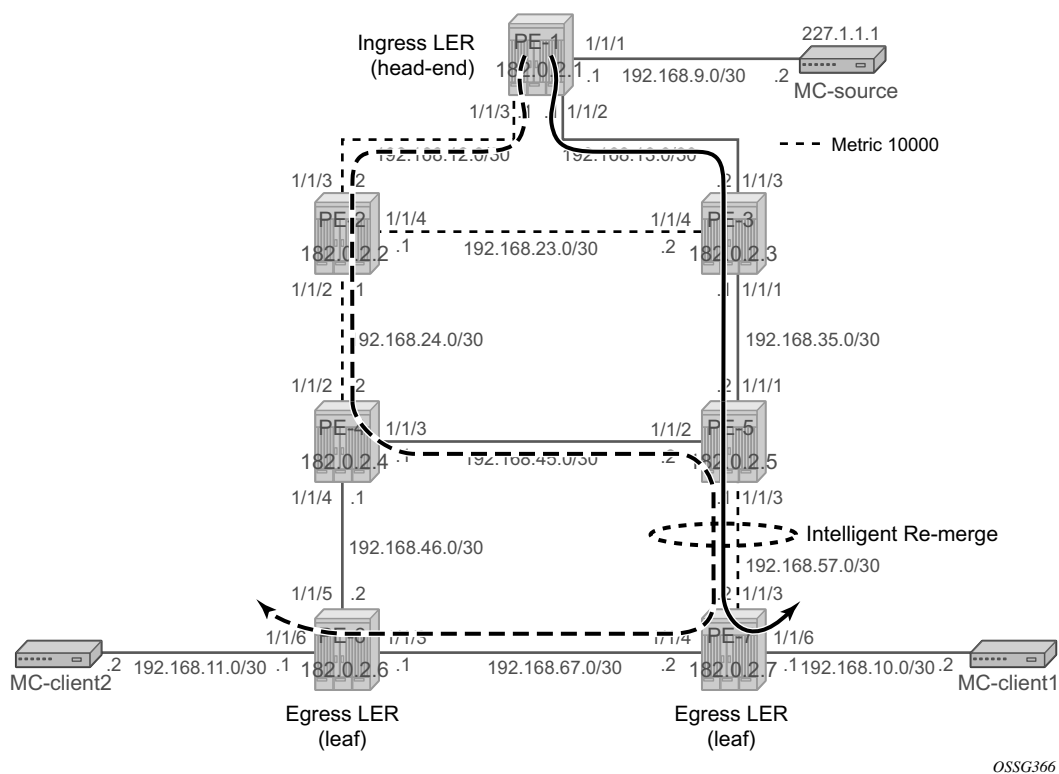


**Figure 173: Intelligent Remerge, Case 1**

On the head-end node (PE-1), a new P2MP LSP *('LSP-p2mp-2')* will be created with two strict direct MPLS paths *('strict-to-PE-7'* and *'strict-to-PE-6'*). See Figure 173 for detailed address information. Intelligent re-merge is performed at node PE-5.

```
A:PE-1# configure router mpls
        …
        path "strict-to-PE-7"
            hop 10 192.168.13.2 strict
            hop 20 192.168.35.2 strict
            hop 30 192.168.57.2 strict
            no shutdown
        exit
        path "strict-to-PE-6"
            hop 10 192.168.12.2 strict
            hop 20 192.168.24.2 strict
            hop 30 192.168.45.2 strict
            hop 40 192.168.57.2 strict
            hop 50 192.168.67.1 strict
            no shutdown
        exit
        …
        lsp "LSP-p2mp-2" p2mp-lsp
            primary-p2mp-instance "p-LSP-p2mp-2"
                s2l-path "strict-to-PE-7" to 192.0.2.7
                exit
                s2l-path "strict-to-PE-6" to 192.0.2.6
                exit
            exit
            no shutdown
        exit
        no shutdown

A:PE-1# show router mpls p2mp-lsp LSP-p2mp-2 p2mp-instance p-LSP-p2mp-2
===============================================================================
MPLS P2MP Instance (Originating)
===============================================================================
-------------------------------------------------------------------------------
Type : Originating
-------------------------------------------------------------------------------
LSP Name    : LSP-p2mp-2                    LSP Tunnel ID  : 13
P2MP ID     : 0
Adm State   : Up                            Oper State     : Up

P2MPInstance: p-LSP-p2mp-2                   P2MP-Inst-type : Primary
P2MP Inst Id: 4                             P2MP Lsp Id    : 22016
Inst Admin  : Up                            Inst Oper      : Up
Inst Up Time: 0d 00:01:56                   Inst Dn Time   : 0d 00:00:00
Hop Limit   : 255                           Adaptive       : Enabled
Record Route: Record                        Record Label   : Record
Include Grps:                               Exclude Grps   :
None                                        None
Bandwidth   : No Reservation                Oper Bw        : 0 Mbps
S2l-Name    : strict-to-PE-7                To             : 192.0.2.7
S2l Admin   : Up                            S2l Oper       : Up
S2l-Name    : strict-to-PE-6                To             : 192.0.2.6
S2l Admin   : Up                            S2l Oper       : Up
-------------------------------------------------------------------------------
P2MP instances : 1
===============================================================================
```

To verify that PE-5 node is not sending duplicate multicast traffic downstream towards PE-7 while it receives two incoming multicast streams, a new tunnel interface and a new static <S,G> IGMP join will be configured on head-end node (PE-1) and leaf nodes (PE-6 and PE-7). Also on the leaf nodes, an extension to the existing multicast information policy is needed. Translated into configuration commands, this becomes

```
A:PE-1# configure router pim tunnel-interface rsvp-p2mp LSP-p2mp-2

A:PE-1# configure router igmp
        …
        tunnel-interface rsvp-p2mp "LSP-p2mp-2"
            static
                group 227.2.2.2
                    source 192.168.9.2


A:PE-6# configure router pim tunnel-interface rsvp-p2mp LSP-p2mp-2 sender 192.0.2.1

A:PE-6# configure router igmp interface to-MC-client2 static group 227.2.2.2 source
192.168.9.2

A:PE-6# configure mcast-management
        multicast-info-policy "p2mp-pol" create
            bundle "bundle2" create
                primary-tunnel-interface rsvp-p2mp LSP-p2mp-2 sender 192.0.2.1
                channel "227.2.2.2" "227.2.2.2" create
                exit
            exit

A:PE-7# configure router pim tunnel-interface rsvp-p2mp LSP-p2mp-2 sender 192.0.2.1

A:PE-7# configure router igmp interface to-MC-client1 static group 227.2.2.2 source
192.168.9.2

A:PE-7# configure mcast-management
        multicast-info-policy "p2mp-pol" create
            bundle "bundle2" create
                primary-tunnel-interface rsvp-p2mp LSP-p2mp-2 sender 192.0.2.1
                channel "227.2.2.2" "227.2.2.2" create
                exit
            exit
```

For verification of incoming/outgoing multicast traffic at PE-5 node, the **monitor** command is used.

```
A:PE-5# monitor port 1/1/1 1/1/2 1/1/3 rate interval 3 repeat 100
===============================================================================
Monitor statistics for Ports
===============================================================================
                                                 Input              Output
-------------------------------------------------------------------------------
…
-------------------------------------------------------------------------------
At time t = 12 sec (Mode: Rate)
-------------------------------------------------------------------------------
Port 1/1/1
-------------------------------------------------------------------------------
Octets                                            2714                  50
Packets                                             40                   0
Errors                                               0                   0
Utilization (% of port capacity)                  0.02               ~0.00

Port 1/1/2
-------------------------------------------------------------------------------
Octets                                            2769                  50
Packets                                             40                   0
Errors                                               0                   0
Utilization (% of port capacity)                  0.02               ~0.00

Port 1/1/3
-------------------------------------------------------------------------------
Octets                                              21                2714
Packets                                              0                  40
Errors                                               0                   0
Utilization (% of port capacity)                 ~0.00               0.02
```

As a conclusion we can say that two incoming multicast streams are seen at PE-5 node (**port 1/1/2** and **port 1/1/1**) and only one outgoing multicast stream (**port 1/1/3**) is sent. No traffic duplication is seen.

**Case 2**

When two paths of the same S2L have ILMs on different incoming ports and go out on the same NHLFE. This is the case when we perform make-before-break (MBB) on an S2L path due to graceful shutdown or global revertive. This is only a temporary situation since the original path will be torn down.
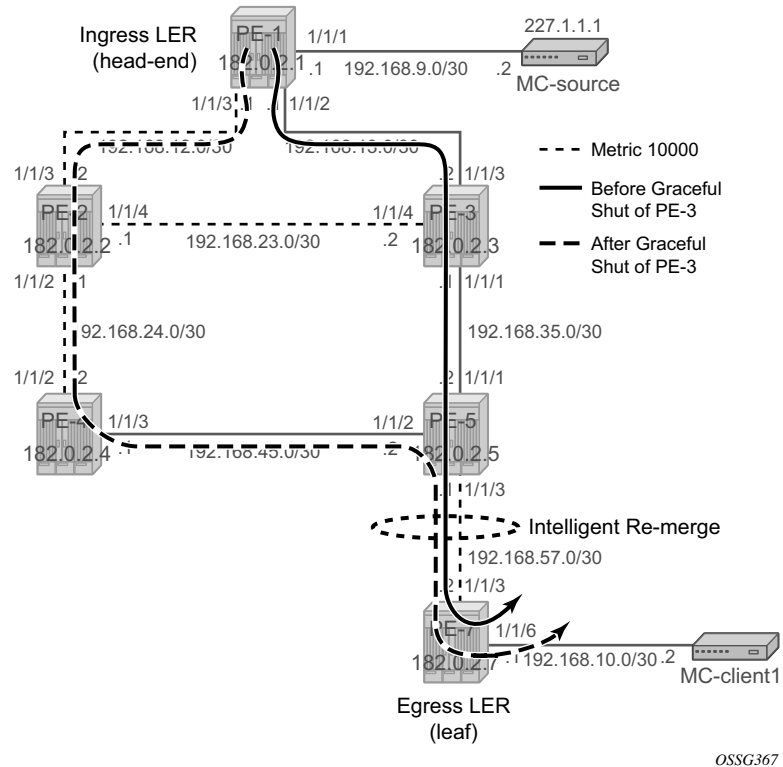


**Figure 174: Intelligent Remerge, Case 2**

For this test, only one MC-client will be looked at (the one connected to head-end node PE-7). On PE-4 and PE-7 nodes, port 1/1/4 will be shutdown to isolate PE-6. On the head-end node (PE-1), a new P2MP LSP (LSP-p2mp-3) will be created with one loose MPLS path (*loose*). See Figure 173 for detailed address information. Also in this case, intelligent re-merge is performed at node PE-5.

```
A:PE-1# configure router mpls
        …
        path "loose"
            no shutdown
        exit
        …
        lsp "LSP-p2mp-3" p2mp-lsp
            primary-p2mp-instance "p-LSP-p2mp-3"
                s2l-path "loose" to 192.0.2.7
```

```
                    exit
                exit
                no shutdown
            exit
            no shutdown

A:PE-1# show router mpls p2mp-lsp LSP-p2mp-3 p2mp-instance p-LSP-p2mp-3 s2l loose to
192.0.2.7 detail
===============================================================================
MPLS LSP LSP-p2mp-3 S2L loose (Detail)
===============================================================================
Legend :
    @ - Detour Available                    # - Detour In Use
    b - Bandwidth Protected                 n - Node Protected
    s - Soft Preemption
===============================================================================
-------------------------------------------------------------------------------
LSP LSP-p2mp-3 S2L loose
-------------------------------------------------------------------------------
LSP Name    : LSP-p2mp-3                  S2l LSP ID  : 17410
P2MP ID     : 0                           S2l Grp Id  : 5
Adm State   : Up                          Oper State  : Up
S2l State:  : Active                                  :
S2L Name    : loose                       To          : 192.0.2.7
S2l Admin   : Up                          S2l Oper    : Up
OutInterface: 1/1/2                        Out Label   : 131069
S2L Up Time : 0d 00:00:27                  S2L Dn Time : 0d 00:00:00
RetryAttempt: 0                            NextRetryIn : 0 sec
S2L Trans   : 3                            CSPF Queries: 0
Failure Code: noError                      Failure Node: n/a
ExplicitHops:
    No Hops Specified
Actual Hops :
    192.168.13.1(192.0.2.1)               Record Label    : N/A
 -> 192.168.13.2(192.0.2.3)               Record Label    : 131069
 -> 192.168.35.2(192.0.2.5)               Record Label    : 131068
 -> 192.168.57.2(192.0.2.7)               Record Label    : 131071
LastResignal: n/a
===============================================================================
```

In a normal situation, the P2MP LSP would follow the nodes PE-1, PE-3, PE-5 and PE-7. This can be verified with MC traffic. Therefore, a new tunnel interface and a new static <S,G> IGMP join will be configured on head-end node (PE-1) and leaf node (PE-7). On the leaf node, an extension to the existing multicast information policy is needed. Translated into configuration commands, this becomes

```
 A:PE-1# configure router pim tunnel-interface rsvp-p2mp LSP-p2mp-3

A:PE-1# configure router igmp
        …
        tunnel-interface rsvp-p2mp "LSP-p2mp-3"
            static
                group 227.3.3.3
                    source 192.168.9.2

A:PE-7# configure router pim tunnel-interface rsvp-p2mp LSP-p2mp-3 sender 192.0.2.1

A:PE-7# configure router igmp interface to-MC-client1 static group 227.3.3.3 source
192.168.9.2

A:PE-7# configure mcast-management
            multicast-info-policy "p2mp-pol" create
                bundle "bundle3" create
                    primary-tunnel-interface rsvp-p2mp LSP-p2mp-3 sender 192.0.2.1
                    channel "227.3.3.3" "227.3.3.3" create
                    exit
                exit

A:PE-5# monitor port 1/1/1 1/1/2 1/1/3 rate interval 3 repeat 999
===============================================================================
Monitor statistics for Ports
===============================================================================
                                              Input              Output
-------------------------------------------------------------------------------
…
-------------------------------------------------------------------------------
At time t = 231 sec (Mode: Rate)
-------------------------------------------------------------------------------
Port 1/1/1
-------------------------------------------------------------------------------
Octets                                         2714                 108
Packets                                          40                   1
Errors                                            0                   0
Utilization (% of port capacity)               0.02              ~0.00


-------------------------------------------------------------------------------
Port 1/1/2
-------------------------------------------------------------------------------
Octets                                           50                  50
Packets                                           0                   0
Errors                                            0                   0
Utilization (% of port capacity)             ~0.00              ~0.00

Port 1/1/3
-------------------------------------------------------------------------------
Octets                                           81                2801
```

```
Packets                                                         1              40
Errors                                                          0               0
Utilization (% of port capacity)                          ~0.00            0.02
```

Now, perform an RSVP graceful shut on PE-3 node. Global revertive is triggered on head-end node (PE-1). A new MPLS path will be calculated (see dotted line in Figure 173). For a few seconds, the old path and new path are active (two incoming MC streams on PE-5 node). PE-5 node is doing intelligent re-merge, not sending duplicate multicast traffic downstream towards PE-7 :

```
A:PE-3# configure router rsvp graceful-shutdown
A:PE-5# monitor port 1/1/1 1/1/2 1/1/3 rate interval 3 repeat 999
===============================================================================
Monitor statistics for Ports
===============================================================================
                                                     Input             Output
-------------------------------------------------------------------------------
…
-------------------------------------------------------------------------------
At time t = 31 sec (Mode: Rate)
-------------------------------------------------------------------------------
Port 1/1/1
-------------------------------------------------------------------------------
Octets                                                2714                108
Packets                                                 40                  1
Errors                                                   0                  0
Utilization (% of port capacity)                      0.02              ~0.00


-------------------------------------------------------------------------------
Port 1/1/2
-------------------------------------------------------------------------------
Octets                                                2714                108
Packets                                                 40                  1
Errors                                                   0                  0
Utilization (% of port capacity)                     ~0.00              ~0.00

Port 1/1/3
-------------------------------------------------------------------------------
Octets                                                  81               2801
Packets                                                  1                 40
Errors                                                   0                  0
Utilization (% of port capacity)                     ~0.00              0.02
```

No traffic duplication is seen.

**Case 3**

When a bypass is active on the S2L path and the new global revertive path of the same S2L arrives on the same incoming interface as the original path (interface flapped) at the FRR merge point node. The implementation recognizes this specific case and will signal a different label from the original S2L path coming on that same interface.
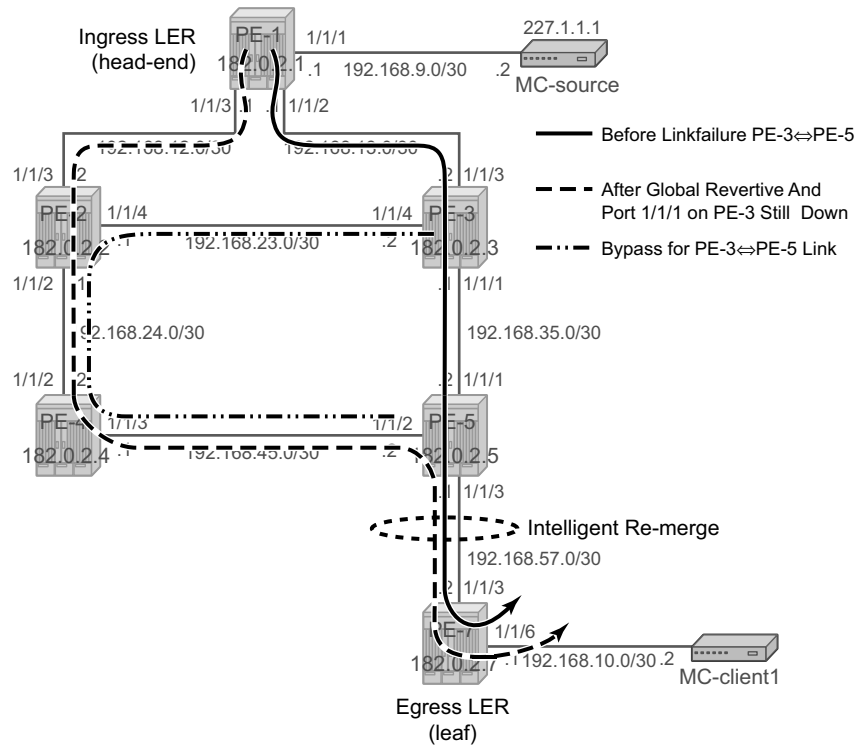


**Figure 175: Intelligent Remerge, Case 3**

For this test, all the non-default OSPF metrics are removed from the interfaces. Only one MC-client will be looked at (the one connected to head-end node PE-7). On PE-4 and PE-7 node, port 1/1/4 will be shutdown to isolate PE-6. On the head-end node (PE-1), a new P2MP LSP (LSP-p2mp-4) will be created with one loose MPLS path (loose) and FRR enabled. See Figure 175 for detailed address information. Also in this case, intelligent re-merge is performed at node PE-5.

```
A:PE-1# configure router mpls
        …
        path "loose"
            no shutdown
        exit
        …
        lsp "LSP-p2mp-4" p2mp-lsp
```

```
              cspf
              fast-reroute facility
                  no node-protect
              exit
              primary-p2mp-instance "p-LSP-p2mp-4"
                  s2l-path "loose" to 192.0.2.7
                  exit
              exit
              no shutdown
          exit
          no shutdown

A:PE-1# show router mpls p2mp-lsp LSP-p2mp-4 p2mp-instance p-LSP-p2mp-4 s2l loose to
192.0.2.7 detail
===============================================================================
MPLS LSP LSP-p2mp-4 S2L loose (Detail)
===============================================================================
Legend :
    @ - Detour Available                  # - Detour In Use
    b - Bandwidth Protected               n - Node Protected
    s - Soft Preemption
===============================================================================
-------------------------------------------------------------------------------
LSP LSP-p2mp-4 S2L loose
-------------------------------------------------------------------------------
LSP Name    : LSP-p2mp-4                   S2l LSP ID : 39940
P2MP ID     : 0                            S2l Grp Id : 4
Adm State   : Up                           Oper State : Up
S2l State:  : Active                                  :
S2L Name    : loose                        To         : 192.0.2.7
S2l Admin   : Up                           S2l Oper   : Up
OutInterface: 1/1/2                         Out Label  : 131069
S2L Up Time : 0d 00:00:31                   S2L Dn Time : 0d 00:00:00
RetryAttempt: 0                            NextRetryIn : 0 sec
S2L Trans   : 1                            CSPF Queries: 1
Failure Code: noError                      Failure Node: n/a
ExplicitHops:
    No Hops Specified
Actual Hops :
    192.168.13.1(192.0.2.1) @              Record Label    : N/A
 -> 192.168.13.2(192.0.2.3) @              Record Label    : 131069
 -> 192.168.35.2(192.0.2.5)               Record Label    : 131069
 -> 192.168.57.2(192.0.2.7)               Record Label    : 131070
ComputedHops:
    192.168.13.1   -> 192.168.13.2   -> 192.168.35.2   -> 192.168.57.2
LastResignal: n/a
===============================================================================
```

In the normal situation, the P2MP LSP follows the nodes PE-1, PE-3, PE-5 and PE-7. This can be verified with MC traffic. Therefore, a new tunnel interface and a new static <S,G> IGMP join will be configured on head-end node (PE-1) and leaf node (PE-7). On the leaf node, an extension to the existing multicast information policy is needed. Translated into configuration commands, this becomes:

```
A:PE-1# configure router pim tunnel-interface rsvp-p2mp LSP-p2mp-4

A:PE-1# configure router igmp
        …
        tunnel-interface rsvp-p2mp "LSP-p2mp-4"
            static
                group 227.4.4.4
                    source 192.168.9.2

A:PE-7# configure router pim tunnel-interface rsvp-p2mp LSP-p2mp-4 sender 192.0.2.1

A:PE-7# configure router igmp interface to-MC-client1 static group 227.4.4.4 source
192.168.9.2

A:PE-7# configure mcast-management
            multicast-info-policy "p2mp-pol" create
                bundle "bundle4" create
                    primary-tunnel-interface rsvp-p2mp LSP-p2mp-4 sender 192.0.2.1
                    channel "227.4.4.4" "227.4.4.4" create
                    exit
                exit

A:PE-5# monitor port 1/1/1 1/1/2 1/1/3 rate interval 3 repeat 999
===============================================================================
Monitor statistics for Ports
===============================================================================
                                                  Input               Output
-------------------------------------------------------------------------------
…
-------------------------------------------------------------------------------
At time t = 3 sec (Mode: Rate)
-------------------------------------------------------------------------------
Port 1/1/1
-------------------------------------------------------------------------------
Octets                                             2737                    21
Packets                                              40                     0
Errors                                                0                     0
Utilization (% of port capacity)                   0.02                ~0.00

Port 1/1/2
-------------------------------------------------------------------------------
Octets                                               21                    21
Packets                                               0                     0
Errors                                                0                     0
Utilization (% of port capacity)                  ~0.00                ~0.00

Port 1/1/3
-------------------------------------------------------------------------------
Octets                                               21                  2714
Packets                                               0                    40
```

```
Errors                                                        0                    0
Utilization (% of port capacity)                         ~0.00                 0.02
```

Now we do an link failure on PE-3 <=> PE-5 interface.

```
A:PE-3# configure port 1/1/1 shutdown
```

As a consequence of this, traffic will be flowing over the bypass link (see also Figure 175 + '#' symbol in next show command).

```
A:PE-1# show router mpls p2mp-lsp LSP-p2mp-4 p2mp-instance p-LSP-p2mp-4 s2l loose to
192.0.2.7 detail
===============================================================================
MPLS LSP LSP-p2mp-4 S2L loose (Detail)
===============================================================================
Legend :
    @ - Detour Available                    # - Detour In Use
    b - Bandwidth Protected                 n - Node Protected
    s - Soft Preemption
===============================================================================
-------------------------------------------------------------------------------
LSP LSP-p2mp-4 S2L loose
-------------------------------------------------------------------------------
LSP Name    : LSP-p2mp-4                    S2l LSP ID  : 39940
P2MP ID     : 0                             S2l Grp Id  : 4
Adm State   : Up                            Oper State  : Up
S2l State:  : Active                                    :
S2L Name    : loose                         To          : 192.0.2.7
S2l Admin   : Up                            S2l Oper     : Up
OutInterface: 1/1/2                         Out Label   : 131069
S2L Up Time : 0d 00:04:21                   S2L Dn Time : 0d 00:00:00
RetryAttempt: 0                             NextRetryIn : 0 sec
S2L Trans   : 1                             CSPF Queries: 1
Failure Code: tunnelLocallyRepaired         Failure Node: 192.0.2.3
ExplicitHops:
    No Hops Specified
Actual Hops :
    192.168.13.1(192.0.2.1) @               Record Label    : N/A
 -> 192.168.13.2(192.0.2.3) @ #             Record Label    : 131069
 -> 192.168.35.2(192.0.2.5)                 Record Label    : 131069
 -> 192.168.57.2(192.0.2.7)                 Record Label    : 131070
ComputedHops:
    192.168.13.1   -> 192.168.13.2   -> 192.168.35.2   -> 192.168.57.2
LastResignal: n/a
In Prog MBB :
 MBB Type   : GlobalRevert                  NextRetryIn : 1 sec
 Started At : 11/13/2009 14:46:50           RetryAttempt: 0
 FailureCode: noError                       Failure Node: n/a
===============================================================================
```

In the meantime, PE-3 will trigger a global revertive action (sending PATHerr message) towards the head-end node (PE-1).

```
A:PE-1# show router mpls p2mp-lsp LSP-p2mp-4 p2mp-instance p-LSP-p2mp-4 s2l loose to
192.0.2.7 detail
===============================================================================
```

```
MPLS LSP LSP-p2mp-4 S2L loose (Detail)
===============================================================================
Legend :
    @ - Detour Available                     # - Detour In Use
    b - Bandwidth Protected                  n - Node Protected
    s - Soft Preemption
===============================================================================
-------------------------------------------------------------------------------
LSP LSP-p2mp-4 S2L loose
-------------------------------------------------------------------------------
LSP Name    : LSP-p2mp-4              S2l LSP ID  : 39940
P2MP ID     : 0                       S2l Grp Id  : 5
Adm State   : Up                      Oper State  : Up
S2l State:  : Active                                :
S2L Name    : loose                   To          : 192.0.2.7
S2l Admin   : Up                      S2l Oper    : Up
OutInterface: 1/1/3                   Out Label   : 131071
S2L Up Time : 0d 00:06:44             S2L Dn Time : 0d 00:00:00
RetryAttempt: 0                       NextRetryIn : 0 sec
S2L Trans   : 2                       CSPF Queries: 2
Failure Code: noError                 Failure Node: n/a
ExplicitHops:
    No Hops Specified
Actual Hops :
    192.168.12.1(192.0.2.1) @        Record Label   : N/A
 -> 192.168.12.2(192.0.2.2)          Record Label   : 131071
 -> 192.168.24.2(192.0.2.4)          Record Label   : 131071
 -> 192.168.45.2(192.0.2.5)          Record Label   : 131071
 -> 192.168.57.2(192.0.2.7)          Record Label   : 131070
ComputedHops:
    192.168.12.1   -> 192.168.12.2   -> 192.168.24.2   -> 192.168.45.2
 -> 192.168.57.2
LastResignal: n/a
Last MBB    :
 MBB Type   : GlobalRevert            MBB State   : Success
 Ended At   : 11/13/2009 14:47:23
===============================================================================
```

For some time, PE-5 will receive two incoming MC streams (both arriving on port 1/1/2). One from 'bypass' path (PE-3 => PE-2 => PE-4 => PE-5) and one from new MPLS path ( PE-1 => PE-2 => PE-4 => PE-5 => PE-7). Port 1/1/3 on PE-5 performs intelligent remerge, only one MC stream is sent downstream towards leaf node PE-7.

```
A:PE-5# monitor port 1/1/1 1/1/2 1/1/3 rate interval 3 repeat 999
===============================================================================
Monitor statistics for Ports
===============================================================================
                                                 Input              Output
-------------------------------------------------------------------------------
…
-------------------------------------------------------------------------------
At time t = 39 sec (Mode: Rate)
-------------------------------------------------------------------------------
Port 1/1/1
-------------------------------------------------------------------------------
Octets                                               0                 261
Packets                                              0                   1
Errors                                               0                   0
Utilization (% of port capacity)                  0.00              ~0.00

Port 1/1/2
-------------------------------------------------------------------------------
Octets                                            5540                 393
Packets                                             80                   2
Errors                                               0                   0
Utilization (% of port capacity)                  0.04              ~0.00

Port 1/1/3
-------------------------------------------------------------------------------
Octets                                             128                3041
Packets                                              1                  40
Errors                                               0                   0
Utilization (% of port capacity)                 ~0.00               0.02

-------------------------------------------------------------------------------
```

# Conclusion

Looking to the configuration point of view, a P2MP LSP is only configured on the head-end node of that P2MP LSP, no explicit configuration is needed on the transit LSRs, branch LSRs, BUD LSRs and egress LERs/leaf nodes.

Since PIM protocol is only needed on the head-end node and leaf node(s), we can work in a PIM-free core network. Although convergence is not covered in this configuration note, failures in the core will be resolved by MPLS (in case of FRR, traffic loss < 50ms is expected). This is a major improvement compared to PIM convergence.