TOMORROW starts here.
mVPN Deployment Models

BRKIPM-2011

Luc De Ghein

Technical Leader Services
Agenda

- Introduction
  - mVPN Rosen GRE Recap

- Building Blocks
  - Signalling in the core
    - PIM
    - mLDP
    - P2MP TE
  - Assigning flows to LSPs
  - Customer Signalling

- Deployment Models (Profiles)
  - mVPN – GRE
  - In-band signaling
  - mVPN – mLDP
  - Partitioned MDT
  - mVPN over P2MP TE
  - Ingress Replication

- LSM Conclusion
Before We Start

- **P-** is Provider
- **C-** is Customer
  - E.g. C-(S,G) is an S,G of a customer, so likely in a VRF
- Only IPv4 mentioned
  - All applies equally to IPv6
  - Similar configuration, other address family
Introduction
mVPN Rosen GRE - Recap

- PIM on the edge
- Unicast routing in overlay across MPLS
- Multicast signalling in overlay
- Multicast through core — GRE encapsulation

IPv4 MDT

Default-MDT created in core using PIM

BiDir
SSM
Sparse mode

For Default MDT Group
Configured on PE per VRF
Per VRF: Default MDT
PIM hello’s and other PIM signalling
Default-MDT emulates a virtual LAN

vpn4 or vpn6 needed to have unicast plane
Unicast plane needed to make RPF succeed
mVPN  Rosen GRE - Recap

PIM on the edge
Unicast routing in overlay across MPLS
Mcast signalling in overlay
Mcast through core – GRE encap

Mcast traffic on Default MDT is received by all PE's attached to that VPN
Even if that PE has no receiver in VRF for the (*,G) or (S,G)
mVPN  Rosen GRE - Recap

PIM on the edge
Unicast routing in overlay across MPLS
Mcast signalling in overlay
Mcast through core – GRE encap

I have receiver: I join
I have no receivers: I ignore

Data MDT Join TLV
PIM message carries C-(S,G) & P-group

CE
Leaf PE
Source

Traffic rate exceeds threshold

mcast data
mVPN Rosen GRE - Recap

PIM on the edge
Unicast routing in overlay across MPLS
Mcast signalling in overlay
Mcast through core – GRE encap

PIM Data-MDT Join TLV
C-\((S,G)\)
P-Group

For high rate sources, data-MDT created
Removes traffic from default-MDT to offload PE’s that did not join stream
Why Label Switched Multicast?

- **Past**
  - 1 solution
  - mVPN based on GRE (Rosen-GRE)

- **Now and Future**
  - NG mVPN
  - LSM (Label Switched Multicast) in core
    - Shared control plane with unicast
    - Less protocols to manage in the core
    - Shared forwarding plane with unicast
    - Only MPLS as encapsulation
    - FRR
LSM Solution Space

Service

C-Multicast Signaling (PE-PE)

Core Tree Signaling

Encapsulation/Forwarding

IPv4
Native
IPv6
Native
IPv4
mVPN
IPv6
mVPN
VPLS

PIM
BGP
None

PIM
MLDP
P2MP TE
IR

IP/GRE
LSM
Unicast MPLS
Building Blocks
Signalling in the Core
Protocols for Building Multicast LSPs

- **Multipoint LDP (mLDP)**
  - Extensions to LDP
  - Support both P2MP and MP2MP LSP
  - RFC 6388

- **RSVP-TE P2MP**
  - Extensions to unicast RSVP-TE
  - RFC 4875

- **Unicast MPLS + Ingress Replication (IR)**
  - Use unicast
  - No additional signalling in core
New LSP Types

Point to Multipoint
- Replication of traffic in core
- Allows only the root of the P2MP LSP to inject packets into the tree
- Signaled with MLDP
  - Path based on IP routing
- Signaled with RSVP-TE
  - Constraint-based / explicit routing
  - Admission control

Multipoint to Multipoint
- Replication of traffic in core
- Bidirectional
- All the leafs of the LSP can inject and receive packets from the LSP
- Signaled with multicast LDP (mLDP)
- Path based on IP routing
New LSP Types

Point – Point to Multipoint

- Combination of P2P and P2MP
- P2P from leaf to the root
- P2MP from root to the leafs
- Signaled with mLDP or RSVP-TE
- P2P to root used for control packets
- P2MP used for data & control

- PPMP is an emulated MP2MP LSP
- Control plane is MP2MP, data plane is P2MP
- Does not work with PIM bidir
- Allows PIM to be run over P2MP LSPs, as with RSVP-TE that does not support MP2MP
- Core routers don’t support MP2MP (dual vendor issues)
PIM
PIM Signalling in the Core

- IP Multicast routing in the core
- PIM in the core
- Used by Rosen GRE model
MLDP Overview

- LSPs are build from the leaf to the root
- Supports P2MP and MP2MP LSPs
  - mLDP with MP2MP provides great scalability advantages for “any to any” topologies
    - “any to any” communication applications:
      - mVPN supporting Bidirectional PIM
      - mVPN Rosen model default MDT
      - If a provider does not want tree state per ingress PE source
- Supports Fast Reroute (FRR) via RSVP TE unicast backup path
- No periodic signaling, reliable using TCP
- Control plane is P2MP or MP2MP
- Data plane is P2MP
- Scalable due to receiver driven tree building
- Supports MP2MP
- Does not support traffic engineering
Tree Types
Official Tree Names

Multi-directional Inclusive PMSI

MI-PMSI (default-MDT)

Like E-LAN
Full mesh P2MP or one MP2MP
Good when sources are in every site

Selective PMSI

S-PMSI (data-MDT)

Like NBMA
A single P2MP

Multidirectional Selective PMSI

MS-PMSI (Partitioned-MDT)

Combination E-LAN and NBMA
A single MP2MP
Good when sources in few sites

PMSI = Provider Multicast Service Interface
mLDP FEC and Opaque Values

- Multicast FEC is advertised by mLDP
- Root node address and opaque value identify the P2MP or MP2MP tree
  - Root node address is
    - learned dynamically (BGP next hop address), for P2MP trees
    - configured, for MP2MP trees
  - Opaque value is used to carry multicast stream information, like
    - (S,G) : in-band signalling
    - LSP identifier : Default/Data MDT
    - ...

- The opaque value has meaning to root and leaves
  - Root and leaves are edge devices
  - Opaque value is mapped to PIM state on the edge devices
  - Opaque value is completely transparent to intermediate nodes (P routers)
LSP Types and Forwarding

P2MP tree

<table>
<thead>
<tr>
<th>Local Label</th>
<th>Outgoing Label</th>
<th>Prefix or Tunnel Id</th>
<th>Bytes Label or Tunnel Id</th>
<th>Switched Label or Tunnel Id</th>
<th>Outgoing Switched</th>
<th>Next Hop IP Address</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>23</td>
<td>[ipv4 10.100.1.6]</td>
<td>232.1.1.1</td>
<td>0</td>
<td>Et1/0</td>
<td>10.1.1.1</td>
</tr>
<tr>
<td>18</td>
<td>23</td>
<td>[ipv4 10.100.1.6]</td>
<td>232.1.1.1</td>
<td>0</td>
<td>Et3/0</td>
<td>10.1.3.3</td>
</tr>
</tbody>
</table>

MP2MP tree

<table>
<thead>
<tr>
<th>Local Label</th>
<th>Outgoing Label</th>
<th>Prefix or Tunnel Id</th>
<th>Bytes Label or Tunnel Id</th>
<th>Switched Label or Tunnel Id</th>
<th>Outgoing Switched</th>
<th>Next Hop IP Address</th>
</tr>
</thead>
<tbody>
<tr>
<td>19</td>
<td>21</td>
<td>[mdt 1000:2000]</td>
<td>0</td>
<td>33516</td>
<td>Et2/0</td>
<td>10.1.2.2</td>
</tr>
<tr>
<td>19</td>
<td>18</td>
<td>[mdt 1000:2000]</td>
<td>0</td>
<td>912</td>
<td>Et1/0</td>
<td>10.1.1.1</td>
</tr>
<tr>
<td>20</td>
<td>24</td>
<td>[mdt 1000:2000]</td>
<td>0</td>
<td>1932</td>
<td>Et3/0</td>
<td>10.1.3.3</td>
</tr>
<tr>
<td>23</td>
<td>24</td>
<td>[mdt 1000:2000]</td>
<td>0</td>
<td>1932</td>
<td>Et2/0</td>
<td>10.1.2.2</td>
</tr>
<tr>
<td>23</td>
<td>21</td>
<td>[mdt 1000:2000]</td>
<td>0</td>
<td>33940</td>
<td>Et3/0</td>
<td>10.1.3.3</td>
</tr>
<tr>
<td>23</td>
<td>19</td>
<td>[mdt 1000:2000]</td>
<td>0</td>
<td>912</td>
<td>Et1/0</td>
<td>10.1.1.1</td>
</tr>
</tbody>
</table>

MP2MP LSP is combination of P2MP LSPs

packet replication and forwarding
P2MP TE
RSVP-TE Overview

- LSPs are build from the head-end to the tail-end
- Supports only P2MP LSPs
- Supports traffic engineering
  - Bandwidth reservation
  - Explicit routing
  - Fast ReRoute
- Signaling is periodic
- P2P technology at control plane
  - Inherits P2P scaling limitations
- P2MP at the data plane
  - Packet replication in the core
The Leafs sends a BGP Auto Discovery message to notify the ingress PE
The ingress sends RSVP-TE Path messages to the leaves
The leaves respond with RSVP-TE Resv messages
# Comparisons – Core Protocols

<table>
<thead>
<tr>
<th>PIM</th>
<th>mLDP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mature protocol</td>
<td>New enhancement to existing protocol</td>
</tr>
<tr>
<td>No need for new code on core routers</td>
<td>Need newer code, but base mLDP code (for P routers) has been around for years now</td>
</tr>
<tr>
<td>Soft state (periodic refresh)</td>
<td>Hard state (no periodic updates)</td>
</tr>
<tr>
<td>GRE encap</td>
<td>MPLS label switching</td>
</tr>
<tr>
<td>Customer state present in core with Data MDT</td>
<td>Customer state present in core with Data MDT</td>
</tr>
<tr>
<td>-</td>
<td>Customer state in core with in-band signalling</td>
</tr>
</tbody>
</table>
## Comparisons – Core Protocols

<table>
<thead>
<tr>
<th>mLDP</th>
<th>P2MP TE</th>
</tr>
</thead>
<tbody>
<tr>
<td>New enhancement to existing protocol</td>
<td>New enhancement to existing protocol</td>
</tr>
<tr>
<td>Need newer code, but base mLDP code (for P routers) has been around for years now</td>
<td>Need newer code</td>
</tr>
<tr>
<td>Hard state (no periodic updates)</td>
<td>Soft state (periodic refresh)</td>
</tr>
<tr>
<td>Dynamic tree building</td>
<td>Static P2MP trees</td>
</tr>
<tr>
<td></td>
<td>Dynamic tree building is coming</td>
</tr>
<tr>
<td>FRR possible (P2P TE tunnel)</td>
<td>FRR possible</td>
</tr>
<tr>
<td>No “BW reservation”</td>
<td>“BW reservation”</td>
</tr>
<tr>
<td>Setup driven by tail ends</td>
<td>Setup driven by head ends</td>
</tr>
<tr>
<td>P2MP and MP2MP trees</td>
<td>P2MP trees only</td>
</tr>
<tr>
<td>Suitable for all mcast applications</td>
<td>Mostly suitable for video delivery</td>
</tr>
<tr>
<td>Best for many-to-many</td>
<td>Best for few-to-many</td>
</tr>
<tr>
<td>Inter-as and CsC (GRE + all mLDP profiles)</td>
<td>No CsC</td>
</tr>
</tbody>
</table>
Customer Signalling
Protocols for Assigning Flows to LSPs

- Static
- PIM
  - RFC 6513
- BGP Customer Multicast (C-Mcast)
  - RFC 6514
  - Also describes Auto-Discovery
- mLDP In-band signaling
  - RFC 6826
Flow Mapping
Overview

PIM in Overlay

BGP in Overlay

Static

Inband
Flow Mapping

Static

- Today mostly applicable to RSVP-TE P2MP
- Static configuration of multicast flows per LSP
- Allows aggregation of multiple flows in a single LSP
Flow Mapping

PIM

- Dynamically assigning flows to an LSP by running PIM over the LSP
- Works over MP2MP and PPMP LSP types
- Mostly applicable (but not limited) to mVPN - Rosen
- No changes to PIM in order to support this
- Allows aggregation of multiple flows in a single LSP
Flow Mapping
Auto Discovery

- Auto Discovery (AD)
  - The process of discovering all the PEs with members in a given mVPN
- Used to establish the MDT in the SP core
- Can also be used to discover set of PEs interested in a given C-group (to enable S-PMSI creation)
  - S-PMSI = Data MDT
- Used to advertise address of originating PE and tunnel attribute information (i.e. which kind of tunnel)
Flow Mapping
Auto Discovery Without BGP

- Rosen GRE needs Default MDT
- Rosen GRE did not need BGP AD
- Core is PIM
  - ASM or BiDir
    - RP is configured/learned: the PEs learn of each other through the RP
  - SSM
    - AF IPv4 MDT under BGP is needed to learn the PEs
BGP IPv4/IPv6 MVPN Address Family

- Specified in RFC 4271, using BGP Multiprotocol Extensions [RFC4760] with an AFI of 1 or 2 and an SAFI of MCAST-VPN

- Used for advertisement of AD routes

- Used for advertisement of C-mcast routes (*,G) and (S,G)

- Two new extended communities (used by vpnv4/6 prefixes, so SAFI 128):
  - VRF route import (replacing connector attribute, i.e. storing route originator IP address/used to import mcast routes, similar to RT for unicast routes)
  - Source AS (advertise AS information: used for inter-AS mVPN)

- New BGP attributes
  - PMSI Tunnel Attribute (PTA) (contains information about advertised tunnel)
  - PPMP label attribute (upstream generated label used by the downstream clients to send unicast messages towards the source)

- The NLRI field in the MP_REACH_NLRI/MP_UNREACH_NLRI attribute contains the MCAST-VPN NLRI
## Flow Mapping

**BGP Auto Discovery + C-Signalling**

### Inside the BGP update

#### PMSI Tunnel Attribute (PTA)

<table>
<thead>
<tr>
<th>Flags (1 octet)</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-6 – Reserved</td>
<td></td>
</tr>
<tr>
<td>7 – L – Leaf information required</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Tunnel type (1 octet)</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>No tunnel information present</td>
</tr>
<tr>
<td>1</td>
<td>RSVP-TE P2MP LSP</td>
</tr>
<tr>
<td>2</td>
<td>mLDP P2MP LSP</td>
</tr>
<tr>
<td>3</td>
<td>PIM-SSM Tree</td>
</tr>
<tr>
<td>4</td>
<td>PIM-SM Tree</td>
</tr>
<tr>
<td>5</td>
<td>PIM-Bidir Tree</td>
</tr>
<tr>
<td>6</td>
<td>Ingress Replication</td>
</tr>
<tr>
<td>7</td>
<td>mLDP MP2MP LSP</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>MPLS label (3 octects)</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>MPLS label</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Tunnel Identifier (variable)</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>RSVP-TE P2MP LSP - &lt;Extended Tunnel ID, Reserved, Tunnel ID, P2MP ID&gt;</td>
</tr>
<tr>
<td>2</td>
<td>mLDP P2MP LSP - &lt;P2MP FEC Element&gt;</td>
</tr>
<tr>
<td>3</td>
<td>PIM-SSM Tree - &lt;P- Root Node Address, P-Multicast Group&gt;</td>
</tr>
<tr>
<td>4</td>
<td>PIM-SM Tree - &lt;Sender Address, P-Multicast Group&gt;</td>
</tr>
<tr>
<td>5</td>
<td>PIM-Bidir Tree - &lt;Sender Address, P-Multicast Group&gt;</td>
</tr>
<tr>
<td>6</td>
<td>Ingress replication - &lt;unicast tunnel endpoint IP address of the local PE that is to be this PE’s receiving endpoint address for the tunnel&gt;</td>
</tr>
<tr>
<td>7</td>
<td>mLDP MP2MP LSP - &lt;MP2MP FEC Element&gt;</td>
</tr>
</tbody>
</table>

#### mcast vpn NLRI

<table>
<thead>
<tr>
<th>Route type (1 octet)</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Intra-AS I-PMSI A-D route</td>
</tr>
<tr>
<td>2</td>
<td>Inter-AS I-PMSI A-D route</td>
</tr>
<tr>
<td>3</td>
<td>S-PMSI A-D route</td>
</tr>
<tr>
<td>4</td>
<td>Leaf A-D route</td>
</tr>
<tr>
<td>5</td>
<td>Source Active A-D route</td>
</tr>
<tr>
<td>6</td>
<td>Shared Tree Join route</td>
</tr>
<tr>
<td>7</td>
<td>Source Tree Join route</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Length (1 octet)</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Length in octets of the Route Type specific field of MCAST-VPN NLRI</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Route type specific (variable length)</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>One or more of the following:</td>
<td></td>
</tr>
<tr>
<td>RD (8 octets)</td>
<td></td>
</tr>
<tr>
<td>MCAST source length (1 octet)</td>
<td></td>
</tr>
<tr>
<td>MCAST source (variable)</td>
<td></td>
</tr>
<tr>
<td>MCAST group length (1 octet)</td>
<td></td>
</tr>
<tr>
<td>MCAST group (variable)</td>
<td></td>
</tr>
<tr>
<td>Originating router’s IP address</td>
<td></td>
</tr>
</tbody>
</table>
# Flow Mapping – BGP Route Types

<table>
<thead>
<tr>
<th>Route Type</th>
<th>Name</th>
<th>AD or C-mcast</th>
<th>Usage</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Intra-AS I-PMSI A-D route</td>
<td>AD</td>
<td>Advertise member PE</td>
</tr>
<tr>
<td>2</td>
<td>Inter-AS I-PMSI A-D route</td>
<td>AD</td>
<td>Same as type 1, but for inter-AS</td>
</tr>
<tr>
<td>3</td>
<td>S-PMSI A-D route</td>
<td>AD</td>
<td>Advertised by Source PE to signal Data MDT</td>
</tr>
<tr>
<td>4</td>
<td>Leaf A-D route</td>
<td>AD</td>
<td>Advertised by Receiver-PE as a response to type 3 route with leaf info required flag set</td>
</tr>
<tr>
<td>5</td>
<td>Source Active A-D route</td>
<td>AD</td>
<td>Advertised by Source PE with active source to facilitate SPT switchover</td>
</tr>
<tr>
<td>6</td>
<td>Shared Tree Join route</td>
<td>C-mcast</td>
<td>Advertise (*,G) by Receiver PE</td>
</tr>
<tr>
<td>7</td>
<td>Source Tree Join route</td>
<td>C-mcast</td>
<td>Advertise (S,G) by Receiver PE</td>
</tr>
</tbody>
</table>
Flow Mapping
BGP Customer-Multicast

- BGP Customer Multicast (C-Mcast) signalling on overlay
- Tail-end driven updates is not a natural fit to BGP
  - BGP is good in one-2-many, not many-2-one
- In mVPN context, PIM is still the PE-CE protocol
- Easy for SSM
- Complex to understand/troubleshoot for ASM
Flow Mapping
New SAFIs to Support BGP Customer Signalling

- SAFI 2 in VRF and SAFI 129 across the core
  - Allows for different mcast vs unicast topologies across MPLS
  - SAFI 129 = VPN mcast SAFI
  - A PE can select an upstream mcast hop which is different than the unicast next hop (RPF is not the unicast route)
  - e.g.

```
router bgp 1
  address-family vpnv4 unicast
  address-family vpnv6 unicast
  address-family vpnv4 multicast
  address-family vpnv6 multicast

  and

  address-family ipv4 unicast vrf vrf1
  address-family ipv6 unicast vrf vrf1
  address-family ipv4 multicast vrf vrf1
  address-family ipv6 multicast vrf vrf1
```
Flow mapping
In-Band signaling

- Only works with mLDP
- Multicast flow information encoded in the mLDP FEC
- Each customer mcast flow creates state on the core routers
- IPv4 and IPv6 multicast in global or VPN context
- Typical for SSM or PIM Sparse mode sources
- IPTV walled-garden deployment
- draft-ietf-mpls-mldp-in-band-signaling
Flow Mapping
Good to Know

- Rosen (GRE or mLDP) uses Default and Data MDT
  - Default MDT (always on) can carry anything
    - All (*,G) must be on it
    - Some (S,G), low rate
  - Data MDT (on demand)
    - Can only carry (S,G)

- Partitioned MDT
  - Shared
  - MDT carries (*,G) and (S,G) for which no Data MDT is triggered
  - Besides one shared MDT, multiple Data MDTs can be used
## Comparisons – Customer Signalling Protocols

<table>
<thead>
<tr>
<th>PIM</th>
<th>BGP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Older protocol, proven, well known</td>
<td>New enhancement to existing protocol</td>
</tr>
<tr>
<td>No changes needed</td>
<td>New procedures (troubleshooting!) complex for PIM SM</td>
</tr>
<tr>
<td>Soft state (periodic refresh)</td>
<td>Hard state (no periodic updates)</td>
</tr>
<tr>
<td>Info driven to specific PE router</td>
<td>Info driven to all PE routers</td>
</tr>
<tr>
<td>PIM adjacencies to all PE routers</td>
<td>BGP adjacencies to all PE routers but likely only to RRs</td>
</tr>
</tbody>
</table>
**Configuration – BGP-AD & C-mcast Signalling**

### IOS

- `vrf definition one`
- `rd 1:3`
- `route-target export 1:1`
- `route-target import 1:1`
- `address-family ipv4`
  - `mdt overlay use-bgp`
  - `mdt auto-discovery mldp pim-tlv-announce`
- `exit-address-family`

### IOS-XR

- `router pim`
- `vrf one`
- `address-family ipv4`
  - `rpf topology route-policy rpf-for-one`
  - `mdt c-multicast-routing pim`
  - `migration route-policy pim-to-BGP`
- `announce-pim-join-tlv`
- `multicast-routing`
- `address-family ipv4`
  - `bgp auto-discovery mldp | pim | ir | p2mp-te`

**Switch to BGP as C-mcast signalling protocol**

**Specify PIM or BGP**

**Specify to use PIM “Data MDT Join TLV” in dual PIM/mLDP network**

**BGP AD**

**BGP AD**
BGP Signalling

Build Default MDT
- config driven

Relay PIM Join
- PIM driven

Relay PIM Join
- PIM driven

Build Data MDT
- data driven
BGP Signalling

Type 3 with Leaf-Info-Required Flag

- BGP driven

SPT Switchover

- data driven

- This is iBGP signalling
  - All BGP speakers get the BGP mvvpn updates
  - Only targetted PEs store it
    - e.g. Type 6 is only stored by RP-PE
    - e.g. Type 5 is stored by all PEs
BGP Signalling - Routes

DRP/0/3/CPH1:meltdown-drp#show bgp ipv4 mvpn vrf one
Route Distinguisher: 1:2 (default for vrf one)

*>i[1][10.1.100.1]/40 10.1.100.1 100 0 i

*>i[3][0][0.0.0.0][0][0.0.0.0][0][0.1.100.1]/120
10.1.100.1 100 0 i

*>i[4][3][1:2][32][10.2.2.9][32][225.1.1.1][10.1.100.2][10.1.100.5]/224
10.1.100.5 100 0 i

*>i[4][3][1:2][32][10.2.100.9][32][225.1.1.1][10.1.100.2][10.1.100.1]/224
10.1.100.1 100 0 i

*>i[5][32][10.2.2.9][32][225.1.1.1]/88
10.1.100.2 100 0 i

*>i[6][1:5][1][32][10.2.100.12][32][225.1.1.1]/184
10.1.100.1 100 0 i

*>i[7][1:2][1][32][10.2.2.9][32][232.1.1.1]/184
10.1.100.1 100 0 i

Intra-AS I-PMSI AD Route
Partitioned MDT
Data MDT advertised by Source PE
Leaf AD Route
Source Active Route
(*,G) Join Route
(S,G) Join Route

R-PE or Source-PE
BGP C-Signalling

- On Rosen models with C-PIM signalling, data traffic needs to be monitored
  - On Default/Data tree to detect duplicate forwarders over MDT and trigger the assert process
  - On Default MDT to perform SPT switchover (switchover from (*,G) to (S,G))

- On Rosen models with C-BGP signalling
  - There is only one forwarder on MDT
    - There are no asserts
  - The BGP Type 5 routes are used for SPT switchover on PEs
Type 4 Leaf AD Routes

- Not required at this time
- Can be enabled
- Only used for display purposes and to track type 3 S-PMSI (Data MDT) routes
- If the Source PE sets the Leaf-Info-required flag on type 3 routes, the Receiver PE responds with a type 4 route

**IOS-XR**

```
multicast-routing

vrf one

address-family ipv4

bgp auto-discovery ir | mldp | p2mp-te | pim

leaf-info-required
```
router pim
vrf one
  address-family ipv4
  rpf topology route-policy rpf-for-one
  mdt c-multicast-routing pim
  migration route-policy pim-to-BGP
!
interface GigabitEthernet0/1/0/0
  enable

route-policy pim-to-BGP
  if destination in (10.1.100.2/32) then
    set c-multicast-routing bgp
  endif
  pass
end-policy

Migration route policy for C-mcast signalling

If source-PE is 10.1.100.2 then use BGP
else use PIM for C-mcast signalling
**Migrating Core Tree**

- Core tree selection can be based on C-S, C-G, Source PE

```plaintext
route-policy rpf-for-one
  if destination in (225.1.1.0/24) then
    set core-tree pim-default
  endif

  if source in (10.2.2.10) then
    set core-tree mldp-partitioned-p2mp
  endif

  if next-hop in (10.1.100.9) then
    set core-tree mldp-inband
  endif

  if next-hop in (10.1.100.11/32) then
    set core-tree mldp-default
  endif
end-policy
```

- If C-Group is 10.2.2.10 then use Rosen GRE
- If C-Source is 10.2.2.10 then use Partitioned mLDP
- If Source-PE is 10.1.100.9 then use mLDP Inband
- If Source-PE is 10.1.100.11 then use Rosen mLDP
Flow Mapping

Summary

- **Static**
  - Mostly applicable to RSVP-TE

- **PIM**
  - Well known, used since the introduction of mVPN over GRE in 2000

- **BGP A-D**
  - Useful where head-end assigns the flows to the LSP

- **BGP C-mcast**
  - Alternative to PIM in mVPN context
  - Might be required in dual vendor networks

- **mLDP In-band signalling**
  - Method to stitch a PIM tree to a mLDP LSP without any additional signaling
Deployment Models
NG-mVPN - Deployment Models

- In-band signaling
  - Global context
  - VPN context (mVPN)

- Rosen Model mVPN
  - GRE based
  - mLDP based

- Partitioned MDT mVPN over mLDP

- P2MP TE
  - Global context
  - VRF Static over P2MP TE
  - mVPN Rosen over P2MP TE

- Ingress Replication
Configuration

- All unicast must be up and running
- IPv4 MDT is needed for Rosen GRE model
- Enable multicast-routing/PIM on the edge
  - Global or VRF interface to CE
- Enable mLDP/P2MP TE on core
  - IOS
    mLDP enabled by default, if MPLS is enabled
  - IOS-XR
    mpls ldp
    mldp
- Specify the mVPN model
multicast-routing
vrf one
address-family ipv4
  mdt mldp in-band-signaling ipv4
  mdt partitioned mldp ipv4 p2mp (bidir)
  mdt partitioned mldp ipv4 mp2mp (bidir)
  mdt partitioned ingress-replication
  mdt mldp in-band-signaling ipv4
  mdt default mldp ipv4 <root>
  mdt default mldp p2mp (partitioned)(bidir)
  mdt default ingress-replication
  mdt default <ipv4-group>
  mdt default (ipv4) <ipv4-group> partitioned
  mdt data <ipv4-group/length>
  mdt data <max nr of data groups> (threshold)
  mdt static p2mp-te tunnel-te <0-65535>
  mdt static tunnel-mte <0-65535>

Specify the mVPN model

MP2MP core-tree with IPv4 Root address
P2MP core-tree
Configuration – IOS-XR

```conf
router pim
vrf one
  address-family ipv4
    rpf topology route-policy rpf-for-one
    interface GigabitEthernet0/0/0/9
```

**Specify RPF to core-tree**

```
RP/0/RP1/CPU0:MeltDown(config-rpl)#set core-tree ?
  ingress-replication-default Ingress Replication Default MDT core
  ingress-replication-partitioned Ingress Replication Partitioned MDT core
  mldp-default MLDP Default MDT core
  mldp-inband MLDP Inband core
  mldp-partitioned-mp2mp MLDP Partitioned MP2MP MDT core
  mldp-partitioned-p2mp MLDP Partitioned P2MP MDT core
  p2mp-te-default P2MP TE Default MDT core
  pim-default PIM Default MDT core
```
Rosen GRE
Rosen mVPN over GRE – Default MDT

- Default-MDT created in core using single IPv4 mcast group
- PIM used for Customer route signalling over default-MDT
- Default-MDT emulates a virtual LAN
Rosen mVPN over GRE – Data MDT

- For high rate sources data-MDT created using P2MP LSPs
- Removes traffic from default-MDT to offload PE’s that did not join stream
- Creation of data-MDT is signalled dynamically using MDT Join messages or BGP A-D routes
- Only PEs with receivers join Data MDT
vrf definition one
rd 1:3
route-target export 1:1
route-target import 1:1
!
address-family ipv4
  mdt default 232.100.1.1
  mdt data 232.100.100.0 0.0.0.255

ip multicast-routing
ip multicast-routing vrf one

router bgp 1
!
address-family vpnv4
  neighbor 10.1.100.7 activate
  neighbor 10.1.100.7 send-community extended
!
address-family ipv4 mdt
  neighbor 10.1.100.7 activate

**Default and Data IPv4 groups for this VRF in global context**

**Multicast-routing globally enabled (PIM on the interfaces)**

**AF IPv4 MDT needed to auto-discover PEs**
Configuration – IOS-XR

route-policy rpf-for-one
    set core-tree pim-default

router bgp 1
    address-family vpnv4 unicast
        !
    address-family ipv4 mdt
        !
neighbor 10.1.100.5
    address-family vpnv4 unicast
        !
    address-family ipv4 mdt
        !

multicast-routing
    address-family ipv4
        interface Loopback0
            enable
            !
        mdt source Loopback0
        interface all enable

vrf one
    address-family ipv4
        mdt source Loopback0
        mdt data 232.100.100.0/24
        mdt default ipv4 232.100.1.1

router pim
    address-family ipv4
        rpf topology route-policy rpf-for-one
        interface GigabitEthernet0/0/4/0
            enable
            !

specify core-tree type

AF IPv4 MDT needed to auto-discover PEs

Multicast-routing enabled in global context

Default and Data IPv4 groups for this VRF in global context

specify route-policy to select RPF topology

PIM enabled on global interfaces
In-band Signaling VPN Context
- PIM (S,G) VPN tree is mapped to a mLDP P2MP LSP
- Root PE is learned via BGP Next-Hop of the VPNv4 Source address route
- R-PE may use SSM Mapping if Receiver is not SSM aware
- RD of the source VRF is included in the mLDP FEC to allow overlapping (S,G) addresses
In-band Signaling VPN Context

- Multicast flow information encoded in the mLDP FEC
- Not well suited for generic mVPN support
- Scalability concern
  - State in VRF creates state in core (on P routers)
vrf definition one
  rd 1:3
  route-target export 1:1
  route-target import 1:1
!
address-family ipv4
!
address-family ipv6

ip multicast-routing
ip multicast-routing vrf one
  ip multicast vrf one mpls mldp

interface Loopback0
  ip address 10.1.100.3 255.255.255.255
  ip pim sparse-mode

mpls mldp logging notifications

ip pim mpls source Loopback0
  ip pim vrf one mpls source Loopback0

enables mLDP inband signalling for the VRF

specifies source interface for multicast RPF
vrf one
  address-family ipv4 unicast
    import route-target
    1:1
    !
  export route-target
    1:1
    !
  address-family ipv6 unicast
    !
interface GigabitEthernet0/0/0/9
vrf one
  ipv4 address 10.2.1.1 255.255.255.0
route-policy rpf-for-one
  set core-tree mldp-inband
mpls ldp
mldp
  logging notifications
  !
interface GigabitEthernet0/0/0/9
router pim
  vrf one
  address-family ipv4
    rpf topology route-policy rpf-for-one
    interface GigabitEthernet0/0/0/9
  address-family ipv6
    !
multicast-routing
  address-family ipv4
    interface Loopback0
      enable
      !
    mdt source Loopback0
    !
  vrf one
    address-family ipv4
      mdt source Loopback0
      mdt mldp in-band-signaling ipv4
      interface all enable
      !
    router igmp
      vrf one
        interface GigabitEthernet0/0/0/9
        !
      router pim
        vrf one
          address-family ipv4
            rpf topology route-policy rpf-for-one
            interface GigabitEthernet0/0/0/9
Rosen Model mVPN over mLDP
Rosen mVPN over mLDP – Default MDT

- Default-MDT created in core using single MP2MP LSP
  - By configuration
  - Any core router is specified as root
- PIM or BGP used for Customer route signalling over default-MDT
- Default-MDT emulates a virtual LAN
For high rate sources data-MDT created using P2MP LSPs

Removes traffic from default-MDT to offload PE’s that did not join stream

Creation of data-MDT is signalled dynamically using MDT Join messages or BGP A-D routes
Rosen mVPN over mLDP

- Known from Rosen GRE
  - same principles/architecture
    - Default MDT = one MP2MP LSP (opaque type MDT), created as soon as the root is configured on the leaf(s)
    - Data MDT is P2MP LSP (opaque type MDT), signalled by PIM by ingress PE upon reaching threshold rate

- No need to configure IPv4 multicast in core

- Avoiding cumbersome design/management of IPv4 groups for Default MDT or IPv4 group ranges for Data MDTs per customer VRF/PE

- PIM or BGP signalling in overlay
vrf definition one

rd 1:3

vpn id 1:1
route-target export 1:1
route-target import 1:1
!
address-family ipv4

mdt default mpls mldp 10.1.100.7
mdt data mpls mldp 100

ip multicast-routing vrf one

interface Loopback0
ip address 10.1.100.3 255.255.255.255
ip pim sparse-mode
!
ip pim mpls source Loopback0
vrf one

vpn id 1:1

address-family ipv4 unicast

import route-target

1:1
!

export route-target

1:1
!

address-family ipv6 unicast

route-policy rpf-for-one

set core-tree mldp-default

mpls ldp

mldp

multicast-routing

!

mdt source Loopback0

vrf one

address-family ipv4

mtd source Loopback0

mtd default mldp ipv4 10.1.100.7

mtd data 100

interface all enable

router pim

vrf one

address-family ipv4

rpf topology route-policy rpf-for-one

interface GigabitEthernet0/0/0/9

enable

vpn id needs to be configured

interface used to set MDT source address

root address of mp2mp tree

data MDTs

specify core-tree type

specify route-policy to select RPF topology

enables mLDP
Rosen mVPN over mLDP – BGP AD

- No need for BGP AD to autodiscover the PEs
  - Root is known by configuration
- BGP AD signaling can be used instead of PIM Join TLV signaling
  - When Data MDT threshold is crossed
  - AF IPv4 mvpn is used to advertise (C-S, C-G) to S-PMSI tunnel binding

**IOS**

```bash
vrf definition one
address-family ipv4

mtt auto-discovery mldp pim-tlv-announce
mtt default mpls mldp 10.1.100.7
mtt data mpls mldp 100

router bgp 1
address-family ipv4 mvpn
neighbor 10.1.100.7 activate
neighbor 10.1.100.7 send-community extended
```

**IOS-XR**

```bash
multicast-routing
vrf one

address-family ipv4

bgp auto-discovery mldp

router bgp 1
address-family ipv4 mvpn
vrf one

address-family ipv4 mvpn
```
Rosen mVPN over mLDP - Data MDT  BGP AD

BGP IPv4 mvpn update
PMSI type
FEC Element
FEC type
adress family
Root
vpn-id
PMSI ID

traffic rate exceeds threshold

Data-MDT

Source
CE
Leaf PE
Leaf PE
Leaf PE
Leaf PE
Leaf PE
Leaf PE
Leaf PE
CE
CE
CE
CE

traffic rate exceeds threshold
Rosen Co-Existence/Migration

- Rosen GRE and Rosen mLDP can co-exist, even in one VRF
  - GRE is preferred if both are present

- Migration is possible from Rosen GRE to Rosen mLDP

- Steps
  1. Enable mLDP in core
  2. Migrate per PE, per VRF at a time
     - Preference command to prefer mLDP over GRE
  3. Remove Rosen GRE when all VRFs/PEs are in Rosen mLDP
### Rosen Co-Existence/Migration

#### IOS

```bash
vrf definition one
rd 1:3
vpn id 1:1
route-target export 1:1
route-target import 1:1
!
address-family ipv4

mdt preference mldp

mdt default mpls mldp 10.1.100.7
mdt data mpls mldp 100
mdt default 232.100.1.1
mdt data 232.100.100.0 0.0.0.255
```

#### iOS-XR

```bash
route-policy rpf-for-one
set core-tree mldp-default

multicast-routing
vrf one
address-family ipv4
mdt source Loopback0
mdt data 232.100.100.0/24
mdt default ipv4 232.100.1.1
mdt default mldp ipv4 10.1.100.7
mdt data 100
interface all enable
```

**RPF topology mLDP used**

**Prefer mldp over GRE**

**Both Rosen mLDP and Rosen GRE configured (and signalled)**
Rosen mVPN over mLDP

Summary

- Default-MDT created using MP2MP LSP
- Functionality as you know, but instead of GRE, MPLS encapsulation and mLDP replaces PIM/multicast in core
- MP2MP is more scalable than PIM SM/SSM since no per PE state created in provider core
- BGP A-D support for Data-MDTs
Rosen Model mVPN over mLDP – Full Mesh P2MP
Rosen mVPN over mLDP – Default MDT is Full Mesh P2MP

- Default-MDT created in core using full mesh of P2MP LSPs
- Each PE is root of one P2MP mLDP LSP
- BGP AD must be enabled!
- Might be used when interoperating
- P2MP might be easier understood than MP2MP
Configuration – IOS-XR

```
router bgp 1
  address-family vpn4 unicast
    address-family ipv4 mvpn
  !
  neighbor 10.1.100.5
    address-family vpn4 unicast
    address-family ipv4 mvpn
  !
  vrf one
    rd 1:1
    address-family ipv4 unicast
    redistribute connected
  !
  address-family ipv4 mvpn
  !
  route-policy rpf-for-one
    set core-tree mldp-default
```

```
  multicast-routing
    mdt source Loopback0
    vrf one
      address-family ipv4
        mdt source Loopback0
        mdt default mldp p2mp
        mdt data 100
        interface all enable
        bgp auto-discovery mldp
  router pim
    vrf one
      address-family ipv4
      rpf topology route-policy rpf-for-one
      interface GigabitEthernet0/0/0/9
          enable
```

**AF ipv4 mvpn is needed**

**Default MDT is full mesh of P2MP mLDP LSPs**

**BGP AD is needed**

**specify route-policy to select RPF topology**

**specify core-tree type**
Partitioned MDT mVPN over mLDP
Partitioned MDT mVPN over mLDP

Introduction

- Dynamic version of Rosen model
- Key difference
  - MDT built only when customer traffic needs to be transported across core
- Address issues in Rosen model
  - Optimizes deployments where customer sources are mostly co-located in few sites
  - Supports Anycast sources
    - With Rosen -> Default MDT -> Asserts
  - Default MDT is MP2MP or P2MP
  - Data MDT is always P2MP
  - Reduces the number of PIM neighbors
    - PIM neighborhood is unidirectional: ingress PE sees egress PEs as PIM neighbors
Initially there is no (default) MDT

Candidate PE’s advertise their LSP identifier as [*,*] wildcard S-PMSI via BGP AD, we call this the MS-PMSI

In this example PE1 and PE2 are candidates

Note, using BGP AD MVPN SAFI is optional, Cisco also supports it without
Partitioned MDT mVPN over mLDP

Setting up the MDT

- PE3 determines that S1 is reachable via PE1 by doing a RIB lookup
- PE3 joins the [*,*] mLDP LSP as advertised by BGP AD for PE1
- When the mLDP LSP is ready, PE3 sends PIM join. Tree is either MP2MP or P2MP
- PE5 joins same (S1,G) it follows same procedures as PE3
- PE4 does not see P-MDT traffic and signalling for (S1,G)
### Setting up the MDT

- PE4 determines that S2 is reachable via PE2 by doing a RIB lookup.
- PE4 joins the [*,*] mLDP LSP as advertised by BGP AD for PE2.
- When the mLDP LSP is ready, PE4 sends PIM join. Tree is either MP2MP or P2MP.
- PE2 joins (S2,G) in customer site and forwards packet down LSP.
- PE5 now joins (S2,G) and follows same procedures as PE4.
Partitioned MDT mVPN over mLDP

PPMP Usage

- PPMP is needed when P2MP is used and PIM as overlay signalling protocol
- Root advertises BGP MVPN prefix with PPMP label
- Leafs use the PPMP label to encapsulate PIM Joins/Prunes to root, which turns around the packet and sends it out mcast on the P2MP tree to all egress PEs
- A P2P LSP is not set up explicitly, its an existing P2P LSP that is used to reach the root
- Why does the PIM Join/Prune need to be received by all egress PE routers?
  - Because of the way PIM Sparse mode works (e.g. PIM router needs to see Joins/Prunes from other PIM routers)
Control Tree

- C-PIM is SM or BiDir
- How do the PE & C-routers learn RP?

MP2MP or P2MP LSP

- Additional P2MP tree
- Only for RP discovery traffic (AutoRP and BSR)
- Not needed for any other model than the partitioned model
- BGP-AD must be enabled
- When an egress PE receives this AD route, it will join the P2MP tree
vrf definition one
rd 1:3
vpn id 1:1
route-target export 1:1
route-target import 1:1
!
address-family ipv4

mdt partitioned mldp p2mp
mdt data mpls mldp 100
exit-address-family

router bgp 1
address-family ipv4 mvpn
neighbor 10.1.100.7 activate
neighbor 10.1.100.7 send-community extended

ip multicast-routing vrf one
vrf one
vpn id 1:1
address-family ipv4 unicast
import route-target
1:1
export route-target
1:1

route-policy rpf-for-one
  set core-tree mldp-partitioned-p2mp

mpls lsp idp
mldp
router bgp 1
  address-family ipv4 mvpn
  neighbor 10.1.100.7
  address-family ipv4 mvpn

multicast-routing
multicast-routing
  !
  mdt source Loopback0
vrf one
  address-family ipv4
  mdt source Loopback0
  mdt partitioned mldp ipv4 p2mp
  mdt data 100
  interface all enable
  interface GigabitEthernet0/0/0/9
router pim
vrf one
  address-family ipv4
  rpf topology route-policy rpf-for-one
  interface GigabitEthernet0/0/0/9
  enable
Partitioned MDT mVPN over mLDP

Summary

- Only PIM join sent towards root
- Only root (ingress PE) is seen as PIM adjacency - unidirectionally
- Core tree is either MP2MP or P2MP
  - If PIM bidir needs to be supported, MP2MP is required
- BGP A-D used to signal core tree [*,*]
- BGP A-D used/needed to signal data-MDT
  - BGP is needed because Partitioned MDT does not lead to full mesh of PIM across core tree (Default MDT)
- PIM or BGP as overlay signalling protocol
- No core state if no customer traffic
- Optimised when sources are co-located in few sites
- Smaller PIM broadcast domain than Rosen mLDP
  - Fewer unnecessary PIM Join/Prune messages
  - No asserts (only one root per P-MDT)
VRF Static over P2MP TE
VRF Static over P2MP TE

- Static mapping of C-(S,G) onto MTE tunnels, on PE headend
- Egress PEs (leaves of P2MP TE tunnel) is statically configured destination list
- There is no C-mcast signalling in overlay

S-PMSI update is used to map C-(S,G) in VRF to tailend of P2MP TE tunnel (MPLS label is incoming label)

static over P2MP TE model also exists in global context
interface tunnel-mtel
ipv4 unnumbered Loopback0

destination 10.1.100.1
  path-option 1 dynamic
! 

destination 10.1.100.3
  path-option 1 dynamic
! 

destination 10.1.100.5
  path-option 1 dynamic
! 

Tunnel type is MTE

Path-option can be explicit or dynamic

Destination list is static

Destination learning (dynamic egress PE set) through BGP AD is future
Configuration – IOS-XR – Headend

vrf one

vpn id 1:1
address-family ipv4 unicast
import route-target 1:1
export route-target 1:1
router pim
vrf one
address-family ipv4

rpf topology route-policy rpf-for-one
interface tunnel-mte1
enable
interface GigabitEthernet0/1/0/0
enable

route-policy rpf-for-one
set core-tree p2mp-te-default

multicast-routing
vrf one
address-family ipv4
mdt source Loopback0

core-tree-protocol rsvp-te
mdt static p2mp-te tunnel-mte1
interface all enable

bgp auto-discovery p2mp-te

router igmp
vrf default
!
vrf one

interface tunnel-mte1
static-group 232.1.1.1 10.2.2.9

specify core-tree protocol rsvp-te
Bind P2MP TE tunnel to VRF
BGP AD needed for P2MP TE
specify route-policy to select RPF topology
C-(S,G) mapping to P2MP TE tunnel
vrf one
vpn id 1:1
address-family ipv4 unicast
  import route-target
    1:1
  export route-target
    1:1

router pim
vrf one
address-family ipv4
  rpf topology route-policy rpf-for-one
  interface GigabitEthernet0/0/0/9
    enable

route-policy rpf-for-one
  set core-tree p2mp-te-default

multicast-routing
  address-family ipv4
  interface Loopback0
    enable
    mdt source Loopback0
    !
  vrf one
  address-family ipv4
    mdt source Loopback0
    core-tree-protocol rsvp-te
  interface all enable
  bgp auto-discovery p2mp-te

specify core-tree protocol rsvp-te
specify route-policy to select RPF topology
BGP AD needed for P2MP TE
mVPN Rosen over P2MP TE
mVPN Rosen over P2MP TE

- Default MDT is full mesh of Static P2MP TE tunnels - Data MDT is future
- BGP AD is required
- No type 3 BGP route is needed
- PIM or BGP C-signalling in overlay
vrf one

vpn id 1:1

import route-target 1:1

export route-target 1:1

router pim

vrf one

address-family ipv4

rpf topology route-policy rpf-for-one

interface GigabitEthernet0/1/0/0

enable

route-policy rpf-for-one

set core-tree p2mp-te-default

multicast-routing

vrf one

address-family ipv4

mtd source Loopback0

core-tree-protocol rsvp-te

mdt default p2mp-te static tunnel-mtel

interface all enable

bgp auto-discovery p2mp-te

router igmp

vrf default

!

vrf one

interface tunnel-mtel

static-group 232.1.1.1 10.2.2.9

specify core-tree protocol rsvp-te

Bind Default P2MP TE tunnel to VRF

BGP AD needed for P2MP TE

specify route-policy to select RPF topology

C-(S,G) mapping to P2MP TE tunnel
vr

address-family ipv4 unicast
import route-target
1:1
export route-target
1:1
router pim
vr
address-family ipv4
rpf topology route-policy rpf-for-one
interface GigabitEthernet0/0/0/9
enable
route-policy rpf-for-one
set core-tree p2mp-te-default

multicast-routing
address-family ipv4
interface Loopback0
enable
mdt source Loopback0
!
vr
address-family ipv4
mdt source Loopback0
core-tree-protocol rsvp-te
mdt default p2mp-te static tunnel-mte1
interface all enable
bgp auto-discovery p2mp-te

specify core-tree protocol rsvp-te
Bind Default P2MP TE tunnel to VRF
BGP AD needed for P2MP TE
specify route-policy to select RPF topology
Ingress Replication (IR)
Ingress Replication (IR)

- Existing unicast LSPs are used (LDP or TE)
- BGP AD is needed
- Default-MDT (Rosen-like) and Partitioned MDT is supported (both with Data-MDT)
- PPMP LSP is needed for Partitioned IR for egress PE sending PIM Join to ingress PE
- PIM and BGP signalling in overlay is supported
Configuration – IOS-XR

vrf one

vpn id 1:1

address-family ipv4 unicast

import route-target

1:1

!

export route-target

1:1

router pim

vrf one

address-family ipv4

mdt c-multicast-routing bgp

!

interface GigabitEthernet0/0/0/9

enable

multicast-routing

!

mdt source Loopback0

core-tree-protocol rsvp-te

static-rpf 10.2.2.9 32 mpls 10.1.100.2

interface all enable

!

vrf one

address-family ipv4

mdt source Loopback0

mdt partitioned ingress-replication

interface all enable

bgp auto-discovery ir

route-policy rpf-for-one

set core-tree ingress-replication-partitioned

specify MDT type

BGP AD for IR

specify core-tree type

specify route-policy to select RPF topology
Ingress Replication (IR)

Summary

- Using a P2MP LSP is not efficient
- Replication on the ingress PE
- Can be used for interop issues
- Transport mcast across MPLS on a part/segment of the network
- Can be done if amount of mcast traffic is really low
## Comparison – Deployment Models

<table>
<thead>
<tr>
<th>Rosen GRE</th>
<th>Rosen mLDP</th>
</tr>
</thead>
<tbody>
<tr>
<td>PIM/mcast in core needed</td>
<td>mLDP in core</td>
</tr>
<tr>
<td>Encapsulation in core is mGRE</td>
<td>Encapsulation in core is MPLS</td>
</tr>
<tr>
<td>Huge number of deployments</td>
<td>Deployments ramping up</td>
</tr>
<tr>
<td>C-PIM in overlay, but BGP is possible</td>
<td>C-PIM in overlay, but BGP is possible</td>
</tr>
<tr>
<td>Control traffic over Default MDT</td>
<td>Control traffic over Default MDT</td>
</tr>
<tr>
<td>Data MDT is signalled by PIM (or BGP)</td>
<td>Data MDT is signalled by PIM (or BGP)</td>
</tr>
<tr>
<td>in overlay</td>
<td>in overlay</td>
</tr>
<tr>
<td>Data MDT is P2MP mcast IP tree</td>
<td>Data MDT is P2MP FEC</td>
</tr>
</tbody>
</table>
## Comparison – Deployment Models

<table>
<thead>
<tr>
<th>VRF mLDP In-band Signalling</th>
<th>mLDP Partitioned MDT</th>
</tr>
</thead>
<tbody>
<tr>
<td>mLDP in core</td>
<td>mLDP in core</td>
</tr>
<tr>
<td>Encapsulation in core is MPLS</td>
<td>Encapsulation in core is MPLS</td>
</tr>
<tr>
<td>C-(S,G) state is present on P routers (in mLDP)</td>
<td>C-(S,G) state is not present on P routers There is one P-MDT per ingress PE per VRF</td>
</tr>
<tr>
<td>No overlay signalling</td>
<td>Overlay signalling</td>
</tr>
<tr>
<td></td>
<td>Control traffic over control tree</td>
</tr>
<tr>
<td></td>
<td>Data MDT (new P2MP FEC) can only be signalled by BGP</td>
</tr>
<tr>
<td></td>
<td>PIM hello’s are one way only</td>
</tr>
</tbody>
</table>
Inter-AS & CsC

- **Inter-AS**
  - GRE: Option C
  - mLDP: Option B & C
  - P2MP TE: Options B & C

- **CsC**
  - Recursive FEC needed

make sure ASBRs have VPNv4/v6 and IPv4/v6 MVPN routes in case no VRF is configured

- **IOS**
  - router bgp 1
  - address-family ipv4 mvpn
  - no bgp default route-target filter

- **IOS-XR**
  - router bgp 1
  - address-family ipv4 mvpn
  - retain route-target all|route-policy

---

**IOS-XR**

```
multicast-routing
vrf one
  address-family ipv4
  bgp auto-discovery mldp
inter-as
```

removes the no-export community

needed for:
- seamless MPLS
- Inter-AS Option B
- BGP-free core
LSM Conclusion
Multicast over MPLS Profiles

<table>
<thead>
<tr>
<th>PIM</th>
<th>P2MP TE</th>
<th>mLDP</th>
<th>mLDP</th>
<th>mLDP</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Inband Signalling</td>
<td>Rosen mLDP</td>
<td>Partitioned MDT</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Static Global P2MP TE (static (S,G))</td>
<td>Global Inband mLDP</td>
<td>Rosen mLDP MP2MP</td>
<td>Partitioned MDT MP2MP</td>
</tr>
<tr>
<td></td>
<td>VRF Inband mLDP</td>
<td>Rosen mLDP MP2MP with BGP-AD</td>
<td>Partitioned MDT MP2MP with BGP-AD</td>
<td></td>
</tr>
<tr>
<td>Rosen GRE</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rosen GRE with BGP-AD</td>
<td>VRF Static over P2MP TE with BGP-AD (static (S,G))</td>
<td>Rosenberg</td>
<td>Rosenberg</td>
<td>Rosenberg</td>
</tr>
<tr>
<td>Rosenberg with BGP-AD and BGP c-mcast signalling</td>
<td>Rosenberg with BGP-AD and BGP c-mcast signalling</td>
<td>Rosenberg with BGP-AD and BGP c-mcast signalling</td>
<td>Rosenberg with BGP-AD and BGP c-mcast signalling</td>
<td>Rosenberg with BGP-AD and BGP c-mcast signalling</td>
</tr>
</tbody>
</table>

*static = static destination list in the core

*no Data MDT

BGP AD

BGP C-mcast signalling

core tree
Multicast over MPLS Profiles bis

<table>
<thead>
<tr>
<th>Rosen IR</th>
<th>Partitioned IR</th>
<th>Rosen P2MP TE</th>
<th>Partitioned P2MP-TE</th>
</tr>
</thead>
<tbody>
<tr>
<td>IR</td>
<td>IR</td>
<td>P2MP</td>
<td>P2MP</td>
</tr>
</tbody>
</table>

- Rosen IR with BGP-AD and PIM c-mcast signalling
- Partitioned IR with BGP-AD and PIM c-mcast signalling
- Rosen P2MP with BGP-AD and PIM c-mcast signalling
- Partitioned P2MP-TE with BGP-AD and PIM c-mcast signalling
- Partitioned IR BGP c-mcast signalling
- Partitioned IR BGP c-mcast signalling
- Rosen P2MP-TE with BGP c-mcast signalling
- Partitioned P2MP-TE BGP c-mcast signalling

Core tree
- Global
- VRF

BRKIPM-2011

© 2014 Cisco and/or its affiliates. All rights reserved.
Cisco Public
Conclusion

- LSM provides unified unicast/multicast forwarding
- mLDP and RSVP are both useful tree building protocols for transporting multicast over MPLS
- It depends on the application and the scalability/feature requirements which protocol is preferred
- Aggregation is useful to limit the number of LSPs that are created
  - Too much aggregation causes flooding
- There are different options to assign multicast flows to LSP’s, Static, PIM, BGP, and mLDP in-band signaling
- For general purpose mVPN we recommend mLDP for tree building and PIM for assigning flows to the LSP
- With NG mVPN, you can choose any model (per VPN/customer)!
  - Even per-source/per-group/next-hop with flexible route-policy (RPL)
Questions?
Complete Your Online Session Evaluation

- Complete your online session evaluation
- Complete four session evaluations and the overall conference evaluation to receive your Cisco Live T-shirt