Draft Rosen Multicast

Massimiliano Sbaraglia

MDT Draft Rosen layer 3 Multicast VPN

- Consiste in una VPN multicast con PIM (Protocol Indipendent Multicast) configurato sia all'interno della VPN stessa che a livello di backbone del Services Provider (control plane);
- Definisce per il trasporto dei dati il protocollo GRE oppure IP-in-IP (data plane);
- Ad ogni VRF corrisponde un Multicast Domain (MD); un router PE, le cui interfaccie sono associate alla VRF instances, appartiene al corrispondente Multicast Domain:
- Per ogni MD esiste una Default Multicast Distribution Tree (MDT) attraverso il backbone del Services Provider, al quale sono connessi tutti i PE routers appartenenti al corrispondente MD;
- Ciascun PE routers configurato con un default MDT group address multicast può essere source (sorgente) di un default MDT;
- Draft Rosen conosciuto come rosen 6 L3 Multicast opera con ASM (Any Source Multicast); la versione draft Rosen 7 L3 Multicast, invece, opera con SSM (Source Specific Multicast) con services provider tunnel, permettendo ai P router del backbone SP, di non mantenere informazioni riguardo le specifiche PIM legate alla VPN.

MDT Draft Rosen layer 3 Multicast VPN

MVPN combines multicast with MPLS VPN. PE routers stabilisce virtual PIM neighborships con altri PE routers che sono connessi alla stessa VPN.

The VPN-specific multicast routing and forwarding database è riferita come **MVRF**.

A **MDT** (multicast distribution tree) tunnel interface è una interfaccia che MVRF usa per accedere al multicast domain. MDT tunnels sono "point-to-multipoint".

Multicast packets sono trasmessi dal CE verso l' ingress PE eppoi incapsulati e trasmessi via il core backbone (over the MDT tunnel); l'egress PE ha la funzione di decapsulare I pacchetti eppoi trasmetterli al receiving CE.

Quando vi è una trasmissione di tipo customer VRF traffic, il PEs incapsula il traffico come (**S,G**) state, dove **G** è il MDT group address, ed **S** è il MDT source per il PE. Attraverso il joining via (S,G) MDT (Mutlicast Tree) è parte del PE neighbors, ed è attivo per ricevere il traffico encapsulated multicast traffic per quella VRF.

Tutti iVPN packets passanti via il provider network sono visti come native multicast packets e sono ruotati sulla base delle routing information in core network.

Per supportare MVPN, PE routers hanno solo bisogno di supportare native multicast routing.

RTs dovrebbe essere configurato in modo che il receiver VRF ha una route unicast reachability per il prefixes presente nel source VRF.

MDT Draft Rosen Control Plane Scalability

- Il CE router mantiene una relationship con il PE router; il PE mantiene adiacenze PIM con altri PE del backbone, quali parte di una determinata MVPN; in caso di un numero elevato di PE, il numero di sessioni PIM può essere notevole stressando cosi il piano di controllo (control plane) di ogni PE;

- Draft Rosen non contempla meccanismi di protezioni quali FRR or custom traffic engineering;

MDT Draft Rosen Multicast Tree

- Ci sono due tipi di alberi Multicast:
  - Default Tree:
    - è usato su base Customer PIM (control plane) e low-rate di dati (data Plane);
    - creato di defult, ciaascun router PE che opera per una determinata MVPN è parte del tree;
    - può usare ASM oppure SSM; SSM è preferito perché permette l'assenza di RP (Rendevous Point) ed è più semplice da gestire.
  - Data Tree:
    - creato dinamicamente quando esiste un active multicast source e listeners;
    - solo i PE con active source or listen possono unirsi (join) al tree.
    - MVPN supportano "optimized VPN traffic forwarding for high-bandwidth applications" largamente distribuiti tra receivers;
    - Un dedicato multicast group può essere usato per encapsulare pacchetti da un specifico source ed un optimized MDT può essere creato per trasmettere traffico solo ai PE routers connessi ed interessati a rivecere il traffico;
    - Un unico group per vrf dovrebbe essere usato a livello PEs routers

MDT Draft Rosen PIM/GRE mVPN

- PIM adjacencies between PEs to exchange mVPN routing information;
- unique multicast address per VPN;
- per-VPN PIM adjacencies between PEs and Ces;
- per-VPN MDT (GRE) tunnels between Pes;
- data MDT tunnels for optimization

BGP/MPLS mVPN or NG mVPN

- BGP peerings between PEs to exchange mVPN routing information;
- PIM messages are carried in BGP:
- BGP autodiscovery for inter-PE tunnels;
- MPLS P2MP inclusive tunnels between Pes;
- selective tunnels for optimization.

MDT Draft Rosen Multicast Configuration Cisco IOS

```
ip multicast-routing
!
ip pim ssm default
!
interface Loopback0
 ip pim sparse-mode
!
interface X
 ip pim sparse-mode
!
ip multicast-routing vrf VPN
!
vrf definition VPN
 address-family ipv4
  mdt default x.x.x.x
  mdt data x.x.x.x y.y.y.y
 exit-address-family
!
router bgp X
 address-family ipv4 mdt
  neighbor x.x.x.x activate
 exit-address-family
```

MDT Draft Rosen Multicast Configuration Cisco IOS-XR

```
multicast-routing
 address-family ipv4
  interface Loopback0
   enable
  !
  mdt source Loopback0
 !
 vrf VPN
  address-family ipv4
   mdt default ipv4 x.x.x.x
   mdt data y.y.y.y/24
   interface all enable
!
router bgp X
 address-family ipv4 mdt
 !
 neighbor x.x.x.x
  address-family ipv4 mdt
```

"MDT source" is required in IOS-XR (it can be configured under the VRF if it's specific for it).

Sparse mode must be activated on all physical interfaces where multicast will be passing through (global or VRF ones) and **on the loopback interface used for the BGP VPNv4 peerings**.

The RP setup of the CEs must agree with the VRF RP setup on the PEs. In case you manually define the RP (static RP) on the CEs, then this must be done on the PEs too (inside the vrf).

Configuration Cisco IOS PIM inside a VRF Tunnel

```
interface Tunnel1
 ip vrf forwarding VPN-A
 ip address 10.10.10.1 255.255.255.0
 ip pim sparse-mode
 tunnel source 192.168.0.1
 tunnel destination 192.168.0.2
 tunnel vrf VPN-B
!
interface Tunnel1
 ip vrf forwarding VPN-A
 ip address 10.10.10.2 255.255.255.0
 ip pim sparse-mode
 tunnel source 192.168.0.2
 tunnel destination 192.168.0.1
 tunnel vrf VPN-B
```

"ip vrf forwarding" defines the vrf under which the tunnel (10.10.10.0/24) operates; above it's VPN-A.

"tunnel vrf" defines the vrf which is used to build the tunnel (from 192.168.0.1 to 192.168.0.2); above it's VPN-B. If the tunnel source and destination are in the global routing table, then you don't need to define their vrf with the "tunnel vrf X" command.

Configuration Cisco IOS **EXTRANET**

- An extranet site can have either the multicast source or the receivers (otherwise multicast happens intra-as)
- The Source PE has the multicast source behind a directly connected CE through the Source MVRF
- The Receiver PE has one or more receivers behind a directly connected CE through the Receiver MVRF
- In order to achieve multicast connectivity between the Source and Receiver PEs, you must have the same default MDT group in the source and receiver MVRF.

Two solutions:

- Configure the Receiver MVRF on the Source PE router
  - you need each receiver MVRF copied on the Source PE router
- Configure the Source MVRF on the Receiver PE routers
  - you need the Source MVRF copied on all interested Receiver PE routers
- In both cases, the receiver MVRF (wherever placed) must import the source MVRF's RT.

  Only PIM-SM and PIM-SSM are supported.

  The multicast source and the RP must reside in the same site of the MVPN, behind the same PE router.

Configuration Cisco IOS EXTRANET Configuration

Receiver MVRF on the Source PE

Source PE (IOS)

```
ip vrf VPN1-S-MVRF
 rd 100:1
 route-target export 100:1
 route-target import 100:1
 mdt default 232.1.1.1
!
ip vrf VPN2-R-MVRF
 rd 100:2
 route-target export 100:2
 route-target import 100:2
 route-target import 100:1
 mdt default 232.2.2.2
!
ip multicast-routing
ip multicast-routing vrf VPN1-S-MVRF
ip multicast-routing vrf VPN2-R-MVRF
```

Receiver PE (IOS)

```
ip vrf VPN2-R-MVRF
 rd 100:2
 route-target export 100:2
 route-target import 100:2
 route-target import 100:1
 mdt default 232.2.2.2
!
ip multicast-routing
ip multicast-routing vrf VPN2-R-MVRF
```

Configuration Cisco IOS EXTRANET Configuration

Source MVRF on the Receiver PE

Source PE (IOS)

```
ip vrf VPN1-S-MVRF
 rd 100:1
 route-target export 100:1
 route-target import 100:1
 mdt default 232.1.1.1
!
ip multicast-routing
ip multicast-routing vrf VPN1-S-MVRF
```

Receiver PE (IOS)

```
ip vrf VPN1-S-MVRF
 rd 100:1
 route-target export 100:1
 route-target import 100:1
 mdt default 232.1.1.1
!
ip vrf VPN2-R-MVRF
 rd 100:2
 route-target export 100:2
 route-target import 100:2
 route-target import 100:1
 mdt default 232.2.2.2
!
ip multicast-routing
ip multicast-routing vrf VPN1-S-MVRF
ip multicast-routing vrf VPN2-R-MVRF
```

Configuration Cisco IOS **Fixing RPF**

There are two options:

static mroute between VRFs:

Receiver PE:

```
ip mroute vrf VPN2-R-MVRF 192.168.1.1 255.255.255.255 fallback-lookup vrf VPN1-S-MVRF
```

group-based VRF selection:

Receiver PE:

```
ip multicast vrf VPN2-R-MVRF rpf select vrf VPN1-S-MVRF group-list 1
ip multicast vrf VPN2-R-MVRF rpf select vrf VPN3-S-MVRF group-list 3
!
access-list 1 permit 231.0.0.0 0.255.255.255
access-list 3 permit 233.0.0.0 0.255.255.255
```

Configuration Cisco IOS **Inter-AS configuration steps:**

To establish a Multicast VPN between two ASes, a MDT-default tunnel must be setup between the involved PE routers. The appropriate MDT-default group is configured on the PE router and is unique for each VPN.

All three (A,B,C) inter-as options are supported. For option A nothing extra is required since every AS is completely isolated from the others.

In order to solve the various RPF issues imposed by the limited visibility of PEs between different ASes, each VPNv4 route carries a new transitive attribute (the **BGP connector** attribute) that defines the route's originator.

Inside a common AS, the BGP connector attribute is the same as the next hop. Between ASes the BGP connector attribute stores (in case of ipv4 mdt) the ip address of the PE router that originated the VPNv4 prefix and is preserved even after the next hop attribute is rewritten by ASBRs.

The BGP connector attribute also helps ASBRs and receiver PEs insert the RPF vector needed to build the inter-AS MDT for source PEs in remote ASes.

The **RPF proxy vector** is a PIM TLV that contains the ip address of the router that will be used as proxy for RPF checks (helping in the forwarding of PIM Joins between ASes).

A new PIM hello option has also been introduced along with the PIM RPF Vector extension to determine if the upstream router is capable of parsing the new TLV. An RPF Vector is included in PIM messages only when all PIM neighbors on an RPF interface support it.

The RPF proxy (usually the ASBR) removes the vector for the PIM Join message when it sees itself in it.

Configuration Cisco IOS **Inter-AS configuration steps:**

- Configuration Steps

Option A
  - no MDT sessions between ASes is required
  - intra-as MDT sessions are configured as usual

- Option B
  - intra-as MDT sessions between PEs, ASBRs and RRs
  - inter-as MDT session between ASBRs
  - RPF proxy vector on all PEs for their VRFs
  - RPF proxy vector on all Ps and ASBRs
  - next-hop-self on the MDT ASBRs

- Option C
  - intra-as MDT sessions between PEs and RRs
  - inter-as MDT sessions between RRs
  - RPF proxy vector on all PEs for their VRFs
  - RPF proxy vector on all Ps and ASBRs
  - next-hop-unchanged on the MDT RRs

MSDP will be required if using an RP on both ASes. Prefer to use SSM in the core of both ASes.

MDT Draft Rosen Multicast Configuration Cisco IOS Verification

Verification

- There should be (S,G) entries for each BGP neighbor, where S=BGP loopback and G=MDT default address;
- There should be a bidirectional PIM adjacency across a tunnel between the PEs, but inside each PE's VRF;
- If an RP is used on a CE, then each remote CE should know this RP;
- Sources/Receivers from any site should be viewable on the RP;
- There should be an MDT data (S,G) entry for each pair of customer (S,G) entries.

MDT Draft Rosen Multicast Configuration Cisco IOS Verification

Verification (using only a default mdt)

MDT default (S,G) entries

```
PE1#sh ip mroute sum

IP Multicast Routing Table
Flags: D - Dense, S - Sparse, B - Bidir Group, s - SSM Group, C - Connected,
       L - Local, P - Pruned, R - RP-bit set, F - Register flag,
       T - SPT-bit set, J - Join SPT, M - MSDP created entry, E - Extranet,
       X - Proxy Join Timer Running, A - Candidate for MSDP Advertisement,
       U - URD, I - Received Source Specific Host Report,
       Z - Multicast Tunnel, z - MDT-data group sender,
       Y - Joined MDT-data group, y - Sending to MDT-data group,
       V - RD & Vector, v - Vector

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

(*, 239.255.255.1), 00:34:36/stopped, RP 192.168.0.1, OIF count: 1, flags: SJCFZ
  (192.168.0.10, 239.255.255.1), 00:24:11/00:02:18, OIF count: 1, flags: JTZ
  (192.168.0.11, 239.255.255.1), 00:34:35/00:02:54, OIF count: 1, flags: FT
```

MDT Draft Rosen Multicast Configuration Cisco IOS Verification

Verification (using only a default mdt)

MDT default (S,G) entries

```
PE1#sh ip mroute 239.255.255.1

IP Multicast Routing Table
Flags: D - Dense, S - Sparse, B - Bidir Group, s - SSM Group, C - Connected,
       L - Local, P - Pruned, R - RP-bit set, F - Register flag,
       T - SPT-bit set, J - Join SPT, M - MSDP created entry, E - Extranet,
       X - Proxy Join Timer Running, A - Candidate for MSDP Advertisement,
       U - URD, I - Received Source Specific Host Report,
       Z - Multicast Tunnel, z - MDT-data group sender,
       Y - Joined MDT-data group, y - Sending to MDT-data group,
       V - RD & Vector, v - Vector

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

(*, 239.255.255.1), 00:46:12/stopped, RP 192.168.0.1, flags: SJCFZ
  Incoming interface: FastEthernet0/0.15, RPF nbr 192.168.5.1
  Outgoing interface list:
    MVRF VPN, Forward/Sparse, 00:46:12/00:01:46

(192.168.0.11, 239.255.255.1), 00:35:47/00:02:28, flags: JTZ
  Incoming interface: FastEthernet0/0.57, RPF nbr 192.168.7.7
  Outgoing interface list:
    MVRF VPN, Forward/Sparse, 00:35:47/00:01:46

(192.168.0.10, 239.255.255.1), 00:46:12/00:03:19, flags: FT
  Incoming interface: Loopback0, RPF nbr 0.0.0.0
  Outgoing interface list:
    FastEthernet0/0.57, Forward/Sparse, 00:35:46/00:03:11
```

MDT Draft Rosen Multicast Configuration Cisco IOS Verification

Verification (using only a default mdt)

```
PE1#sh bgp ipv4 mdt all 192.168.0.11/32

BGP routing table entry for 100:1:192.168.0.11/32        version 2
Paths: (1 available, best #1, table IPv4-MDT-BGP-Table)
  Not advertised to any peer
  Local
    192.168.0.11 from 192.168.0.1 (192.168.0.1)
      Origin incomplete, metric 0, localpref 100, valid, internal, best
      Originator: 10.0.0.6, Cluster list: 10.0.0.1, 10.0.0.20,
      MDT group address: 239.255.255.1
```

MDT Draft Rosen Multicast Configuration Cisco IOS Verification

Verification (using only a default mdt)

```
PE1#sh ip pim mdt

  * implies mdt is the default MDT
  MDT Group/Num   Interface   Source                VRF
* 239.255.255.1   Tunnel1     Loopback0             VPN


PE1#sh ip pim mdt bgp

MDT (Route Distinguisher + IPv4)              Router ID          Next Hop
  MDT group 239.255.255.1
   100:1:192.168.0.11                         192.168.0.1        192.168.0.11
```

MDT Draft Rosen Multicast Configuration Cisco IOS Verification

Verification (using default and data mdt)

MDT default (S,G) entries
MDT data (S,G) entries

```
PE2#sh ip mroute

IP Multicast Routing Table
Flags: D - Dense, S - Sparse, B - Bidir Group, s - SSM Group, C - Connected,
       L - Local, P - Pruned, R - RP-bit set, F - Register flag,
       T - SPT-bit set, J - Join SPT, M - MSDP created entry, E - Extranet,
       X - Proxy Join Timer Running, A - Candidate for MSDP Advertisement,
       U - URD, I - Received Source Specific Host Report,
       Z - Multicast Tunnel, z - MDT-data group sender,
       Y - Joined MDT-data group, y - Sending to MDT-data group,
       V - RD & Vector, v - Vector

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

(1.1.1.1, 232.0.0.1), 00:08:53/00:03:27, flags: sT
  Incoming interface: Loopback0, RPF nbr 0.0.0.0
  Outgoing interface list:
    FastEthernet0/0.24, Forward/Sparse, 00:08:53/00:03:27

(10.10.10.10, 232.0.0.1), 00:50:48/stopped, flags: sTIZ
  Incoming interface: FastEthernet0/0.24, RPF nbr 20.2.4.4
  Outgoing interface list:
    MVRF VPN, Forward/Sparse, 00:50:48/00:00:11

(10.10.10.10, 232.0.1.0), 00:08:23/00:00:12, flags: sTIZ
  Incoming interface: FastEthernet0/0.24, RPF nbr 20.2.4.4
  Outgoing interface list:
    MVRF VPN, Forward/Sparse, 00:02:47/00:00:12

(10.10.10.10, 232.0.1.1), 00:01:59/00:01:00, flags: sTIZ
  Incoming interface: FastEthernet0/0.24, RPF nbr 20.2.4.4
  Outgoing interface list:
    MVRF VPN, Forward/Sparse, 00:01:59/00:01:00
```

MDT Draft Rosen Multicast Configuration Cisco IOS Verification

Verification (using default and data mdt)

MDT default (S,G) entries
MDT data (S,G) entries

PE2#sh ip pim mdt

```
  * implies mdt is the default MDT
  MDT Group/Num   Interface   Source                VRF
* 232.0.0.1       Tunnel0     Loopback0             VPN
  232.0.1.0       Tunnel0     Loopback0             VPN
  232.0.1.1       Tunnel0     Loopback0             VPN
```

PE2#sh ip pim mdt bgp

```
MDT (Route Distinguisher + IPv4)              Router ID         Next Hop
  MDT group 232.0.0.1
   100:1:10.10.10.10                          10.10.10.10       10.10.10.10
```

MDT Draft Rosen Multicast Configuration Cisco IOS Verification

Verification (using default and data mdt)

In both scenarios, you can also verify the mGRE tunnels by looking at the tunnel interface itself.

```
PE1#sh int tun1 | i protocol/transport

  Tunnel protocol/transport multi-GRE/IP
```

When all PIM adjacencies come up, as PIM neighbors in a VRF you should see all the other MDT PEs though a tunnel and all the local connected CEs through a physical interface.

```
PE1#sh ip pim vrf VPN nei

PIM Neighbor Table
Mode: B - Bidir Capable, DR - Designated Router, N - Default DR Priority,
      P - Proxy Capable, S - State Refresh Capable, G - GenID Capable
Neighbor          Interface                Uptime/Expires    Ver   DR
Address                                                            Prio/Mode
192.168.59.9      FastEthernet0/0.59       00:00:22/00:01:22 v2    1 / DR S G
192.168.0.11          Tunnel1              00:25:52/00:01:27 v2    1 / DR S P G
```