

Multicast VPNs

Based on IETF **Draft-Rosen**, GRE Tunnels, PIM-SSM, MDT

If PIM-SSM used, P routers don't need VPN-Specific PIM Information

Need a **Default MDT** (multicast distribution trees) and a **Data MDT**

First data is sent over default MDT, all PEs receive but discard it if not needed. Data MDT takes care of that.

PEs discover one another in VRF across the Default MDT through MBGP (MDT-SAFI)

Each PE imports the MDT-SAFI each PE router if the RT matches. Each PE then joins the (S,G) rooted at other PEs

If traffic received crosses a tshold, **Source PE** can move it to a new Data MDT and taken out of Default MDT (after 3sec).

Source PE sends **MDT Join TLV** (UDP) (has (S,G), new Data MDT group address in provider space).

Source PE sends Joins regularly over Default MDT so that everybody hears it.

When PE needs this traffic sends Pim Join directly to source (SSM, he knows the source from the MDT signalling)

If traffic goes down, **Source PE** can move back to Default MDT

Multicast VPNs Advantage: P-routers don't hold state info for all customer groups, scalable (unlike full mesh of GRE tunnels to send multicast over)

Multicast domain: basically a vrf. A default MDT for each multicast domain.

Only pure IP multicast in core (no label).

In **P-network**, only **PIM-SSM** and **PIM-SM** (w/ or w/o **BIDIR**) can be used.

In **C-network**, everything can be used, but only (S,G) will be moved to DataMDTs!!!

so PIMSM can be used if destinations join SPT after going through RP, as per the normal PIMSM rules

DefaultMDT signaled through MPBGP by each PE (sent: MDT Group, RD, PE RID)

DefaultMDT only for control traffic. If more BW needed, move to **DataMDT**, so that traffic does not go everywhere needlessly (some PE routers in the multicast domain may be dormant - no receivers)

After more PEs join a MDT, each one is root and the others are leaves. So each one is root and leaf at the same time.

PEs at source choose to switch to **DataMDT**. Only (S,G) can switch, if Customer has PIM-BIDIR (*,G) no switch to DataMDT!!

PIM **Data-MDT Join** sent to inform PEs to switch (UDP 3232). If less BW needed, can be switched back to Default MDT (less states to be remembered in core-P routers) but minimum 1 minute on DataMDT.

If PE does not have subscribers, caches Data-MDT Join contents for 3 minutes. Source PE refreshes every 1 minute.

Tunnel interfaces used, but not appear in IGP or RIB! RPF-check must be changed. Route to other side CE address from MPBGP is used for RPF .

MVPN without SSM!

Uses a shared tree and rendezvous point (RP) for autodiscovery of the PE routers.

Source PE encapsulates multicast data packets into a PIM register message and sends it unicast to the RP router. The RP then builds a shortest-path tree (SPT) toward the Source PE. A Receiver PE for the MDT multicast group sends (*, G) join messages toward the RP and joins the distribution tree for that group.

o When PIM-SM initializes in a MD, the PE originates a PIM Join message for the MD P-Group address towards the RP in the SP space. This is done for each MD that is configured on the PE.

o Since an MT interface belongs to a VPN, sending a Hello message on this interface does the following:

o The PIM Hello message has the source address of PE's loopback interface in the SP address space and the destination of ALL-PIM-ROUTERS group.

o This PIM Hello gets encapsulated in a GRE header with the source address as the PE's loopback interface and the destination as the MD P-Group address. After the encapsulation, the original PIM-SM Hello travels as the data packet in a PIM-SM Register towards the SP RP.

o RP in the SP network knows about all the receivers (the PEs) because of the earlier PIM Join for the MD P-Group address that it received from all the PEs when they initialized. So, when the RP receives the above PIM-SM register, it decapsulates it and forwards it down to all the PEs. So, all the remote PEs (including the one who sent the packet) receives this data packet which has the source address of the originating PE.

o This PIM Hello packet originated within the VRF travels as the data packet (due to encapsulation) in the SP network towards the RP.

o The above procedure is repeated on all the PEs. Hence, all the PEs receive each other's data packets which contain PIM Hello messages and discover one another. PEs can decide to send the source Join directly to the remote PEs at this point.