Inter-AS Option BC for MPLS VPN/EVPN

Jeffrey Zhang, Kireeti Kompella, Bruno Decraene, Luay Jalil

BESS WG, IETF117
VPN Inter-AS Option B

• ASBRs re-advertise routes with nexthop-self and locally allocated service labels, and setup forwarding state to switch traffic based on the service labels
  • Per-service-label forwarding state

• Good for inter-provider scenarios
  • Loopback addresses are not exposed to other providers
  • Inter-ASBR eBGP sessions with rich policy control
  • Label spoof check (LSC) for service labels can be performed at the ASBRs
VPN Inter-AS Option C

• PEs exchange service routes via RRs
  • w/o label/nexthop modification
  • w/o going through the ASBRs

• ASBRs re-advertise labeled transport routes for the service routes’ nexthops
  • Service traffic are label switched based on the transport route labels

• Very good scalability – no per-service-label forwarding state

• Good for intra-provider only
  • A provider may not want to externally expose its internal transport routes
  • A provider may not want to externally exchange routes via RRs
    • Lack of policy control
  • LSC for inter-provider case at the egress ASBR degrades scalability to Option-B level
Data; swap of PE labels at each hop
VPN Inter-AS Option BC

• When an egress ASBR re-advertise service routes:
  • Do change the nexthop to itself
  • If the NLRI has only one label, do not change it but add in the NLRI a locally allocated label that binds to the LSP to the received nexthop
    • Set up forwarding state to label switch traffic with the added label to the LSP
  • If the NLRI has more than one label, change the first label to a locally allocated label that binds to the <received nexthop, received first label>
    • Set up forwarding state to swap the new label to the received first label and tunnel to the nexthop
  • This achieves what the labeled transport route (re-advertised across AS boundary) does in Option C
<100, sprfx1, PE1> where 201 binds to PE1
<201, 100, sprfx1, ASBR1> where 202 binds to ASBR1, 201
<202, 100, sprfx1, ASBR21> where 203 binds to ASBR21, 202
<203, 100, sprfx1, ASBR22> where 204 binds to ASBR22, 203

<100, sprfx2, PE2> where 301 binds to PE2
<301, 100, sprfx2, ASBR1> where 302 binds to ASBR1, 301
<302, 100, sprfx2, ASBR21> where 303 binds to ASBR21, 302
<303, 100, sprfx2, ASBR22> where 304 binds to ASBR22, 303

Blue labels are resolved from the received NH; they get packets to next ASBR/PE
Red labels are the first label in the received multi-label NLRI. On the advertising ASBR, they’re bound to <next hop[, binding] label>. Forwarding is based on red label.
Use Tunnel Encapsulation Attribute

• As an alternative, a new “composite tunnel” type of Tunnel Encapsulation Attribute can be used instead of multi-label NRLI
  • It includes the local address of the re-advertising ASBR, and a locally allocated label that is bound to the received nexthop or the received composite tunnel
  • Because of the new tunnel type (that may not be supported by all PEs), incremental transition procedures are needed
    • Details in the draft
  • This was the initial solution but will likely be removed now that we have the multi-label NLRI solution
VPN Inter-AS Option BC

• A combination of Option B and C
  • Scaling property of Option C
    • traffic is switched based on the PE label, not the service label
  • Control property of Option B
    • Rich policy control across ASBRs
    • No exposure of loopback address

• PE label (advertised separately with transport routes in Option C) is advertised together with service routes
  • W/o tying it to the PE address
  • As the first label in multi-label NLRI, or in a new “Composite Tunnel” in TEA
Can Be Used for SRv6/MPLS Interwork, too

• With SRv6/MPLS service interworking, when re-advertising from SRv6 to MPLS side, a new IWNH is used for the re-advertised routes
  • The IWNH is mapped to the SID value in the received Prefix SID Attribute
    • A transport route for the IWNH is advertised additionally
  • For IPv4 MPLS, an IPv4 address is needed as the IWNH for each SID value
    • This is not desired – there could be many IPv4 addresses needed
  • For IPv6 MPLS, the IWNH can be the SID value itself
    • But the SRv6 operator may not want the SID values exposed to MPLS side

• The Interworking node may re-advertise to MPLS site the service routes with an additional label or with a “composite tunnel” in a TEA instead of using an IWNH
  • The additional label or the composite tunnel maps to the SRv6 SID value
Comments Appreciated!

- The main motivation was to not expose PE addresses outside local AS
- Label Spoof Check was not considered; if it is needed then the egress ASBR’s scaling property degrades to Option-B
  - On other ASBRs, scaling property is still Option-C