Covering Prefixes Outbound Route Filter for BGP-4

draft-bonica-l3vpn-orf-covering-prefixes-01

H. Jeng, l. Jalil, R. Bonica, Y. Rekhter, K. Patel, L. Yong

Overview

- Define a new ORF-type, called the "Covering Prefixes ORF (CP-ORF)"

- Realizes a "route pull" model in BGP
  - BGP speaker, on demand, pull certain routes from peer

- Applicability
  - Ethernet MPLS/BGP Virtual Private Networks (EVPN)
  - Virtual Hub-and-Spoke VPN's (RFC 7024)

EVPN WITH DEFAULT MAC GATEWAY: A BRIEF REVIEW

Goal

- Reduce the number of routes that PE2 carries
- PE2 carries only one Default MAC Route (DMR)
  - Next-hop == DMG1
- Traffic from PE2 traverses DMG1
- Traffic to PE2 may traverse a more direct route
- DMG1 is a logical function
  - Can be implemented on a router the performs as an NVE, on a separate router or on the RR

RR   DMG1

Red EVI

PE1

Red EVI

Red EVI (MAC-Address-A)

PE2
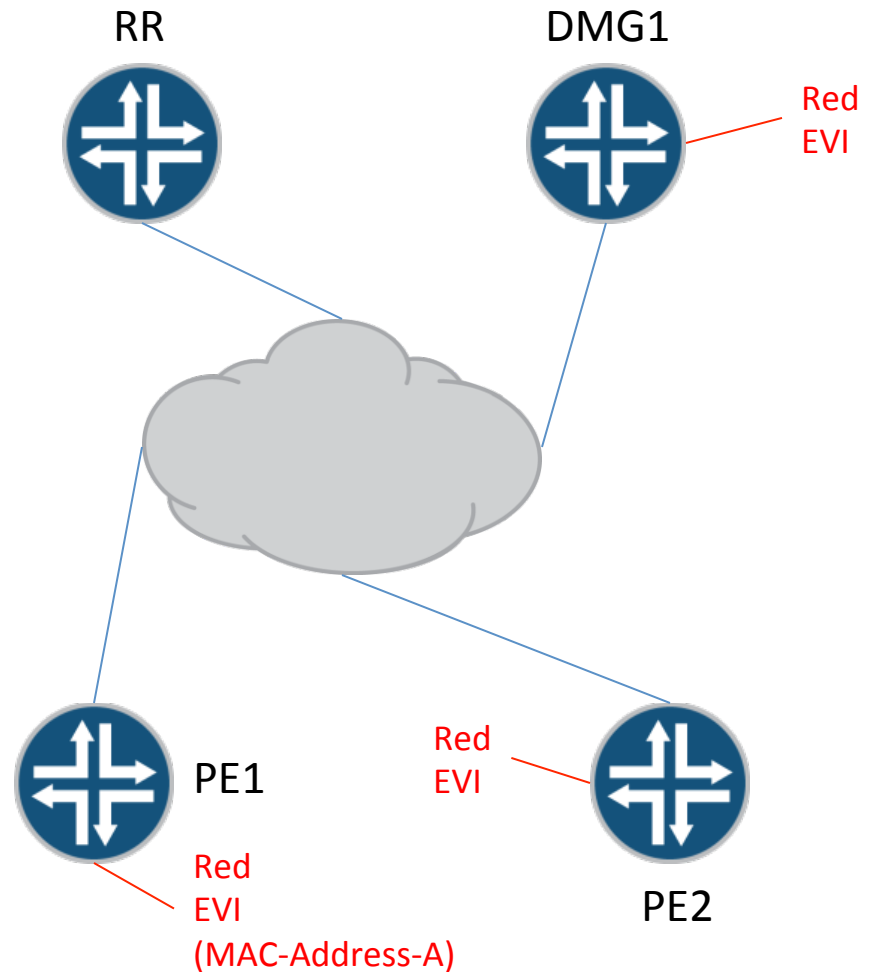
Assumption

- NVEs support Default MAC Route (DMR)
- See Section 10.1 of draft-ietf-l2vpn-evpn

BGP Routing Policy

- PE1 and DMG1 are clients of a RR
  - PE2 may be client of RR or DMG1
- PE1 and DMG1 accept advertisements carrying the RT, RT-RED
- PE2 accepts advertisements carrying the RT, RT-RED-FROM-DMG1

RR

DMG1

Red EVI

PE1

Red EVI

Red EVI (MAC-Address-A)

PE2
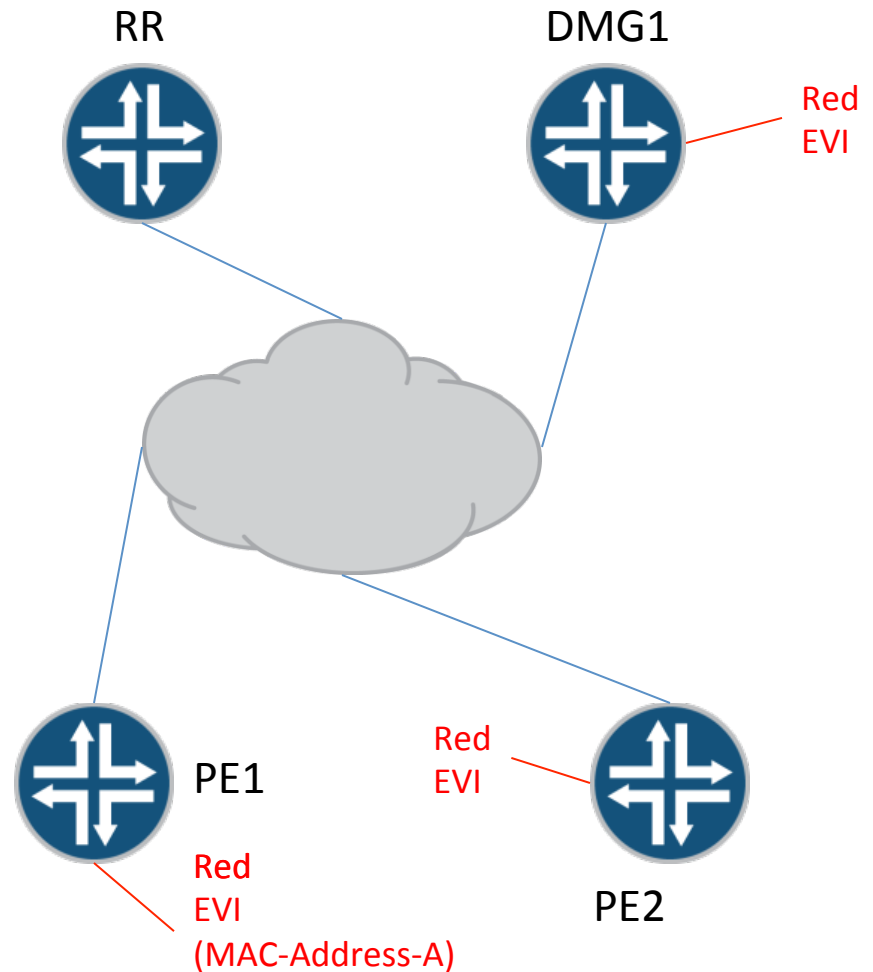
BGP Advertisements

- PE1 advertises MAC-Address-A to the RR
  - Next-hop = Self
  - RT = RT-RED
- RR reflects route to DMG1
  - DMG1 accepts
- RR may also advertise route to PE2
  - In absence of RT-Constrain
  - If advertised, PE2 rejects

RR

DMG1

Red EVI

PE1

Red EVI

PE2

Red EVI (MAC-Address-A)
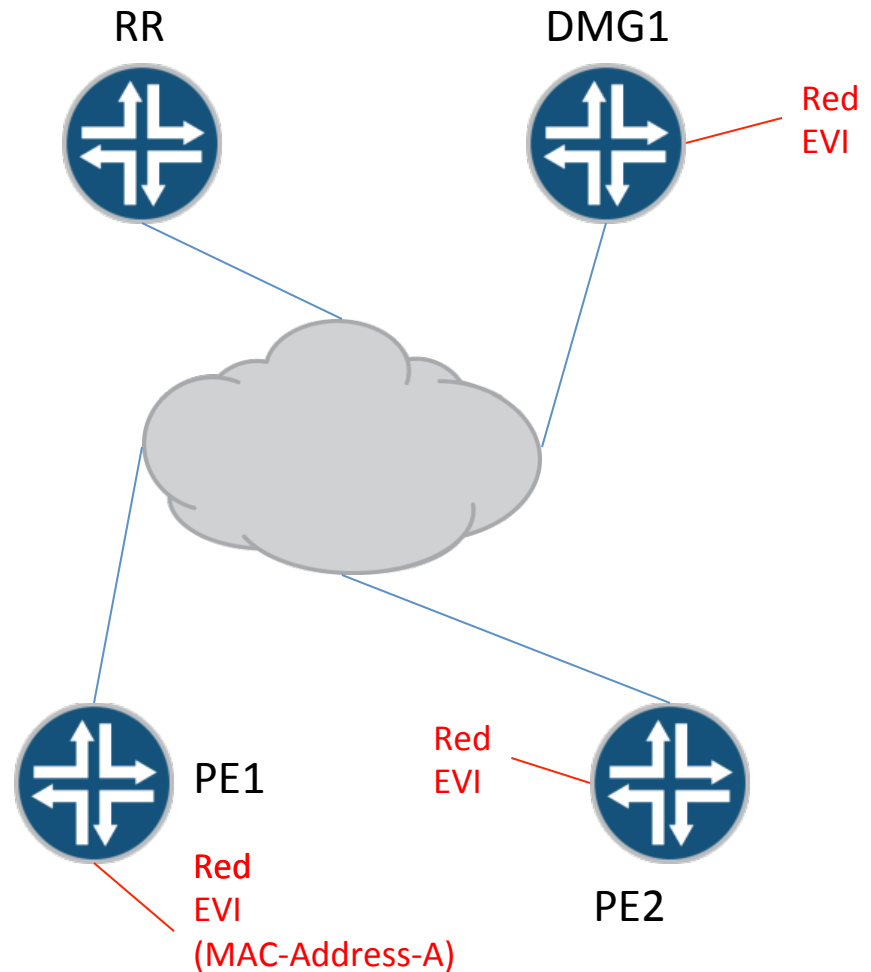
BGP Advertisements (continued)

- DMG1 advertises DMR to the RR
  - Next-hop = Self
  - RT = RT-RED-FROM-DMG1

- RR reflects route to PE2
  - PE2 accepts

RR

DMG1

Red EVI

Red EVI

PE1

Red EVI (MAC-Address-A)

PE2
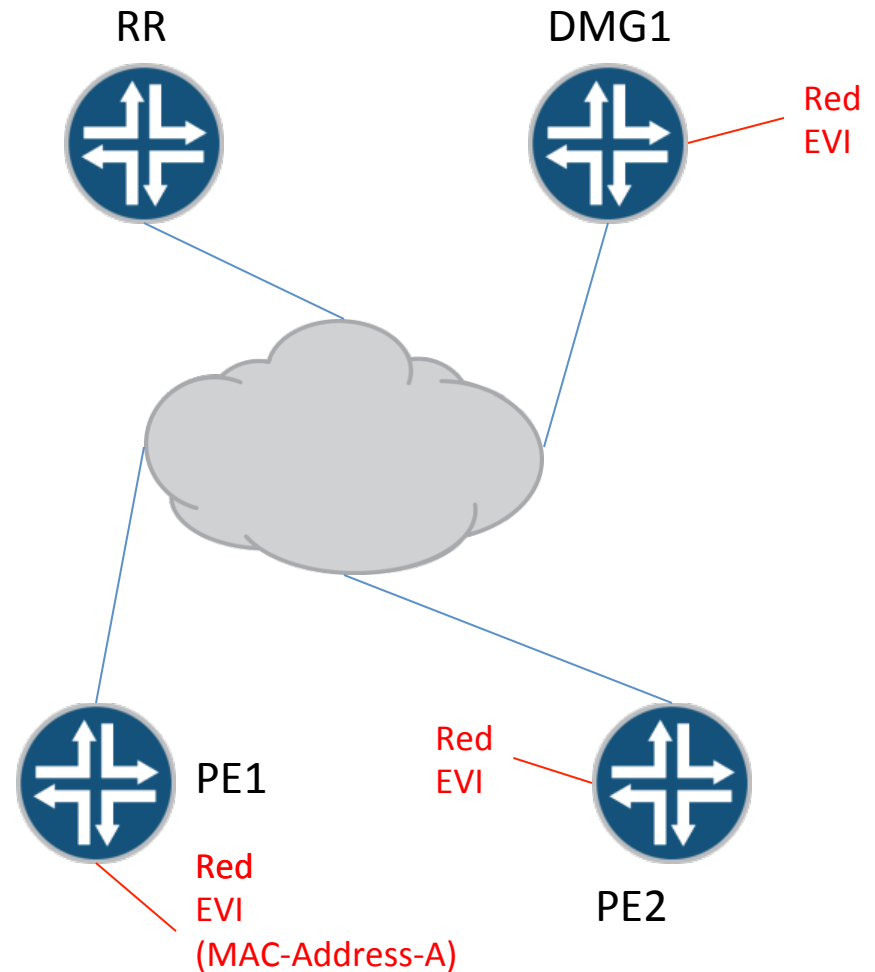
COVERING PREFIX ORF

Problem to Be Solved

- A host served by PE2 originates an "exceptional" flow to MAC-Address-A
  - Large, latency sensitive, etc.
- Flow traverses DMG1
- Flow might benefit from a more direct route to MAC-Address-A
  - If such a route exists
- The criteria determining that a flow might benefit from a more direct route are strictly local to PE2
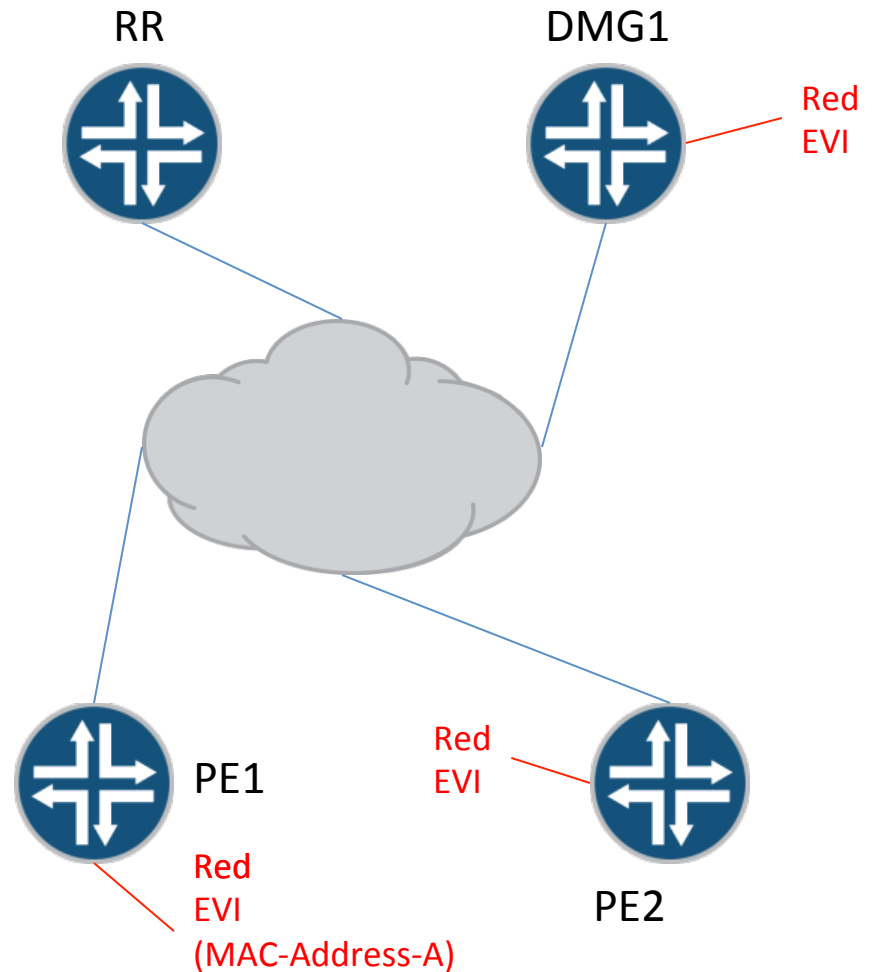
Solution

- PE2 requests the most specific route to MAC-Address-A from the RR
  - Carrying additional RT, RT-RED-FROM-DMG1
- Pull versus push

RR

DMG1

Red EVI

PE1

Red EVI

Red EVI MAC-Address-A

Red EVI

PE2
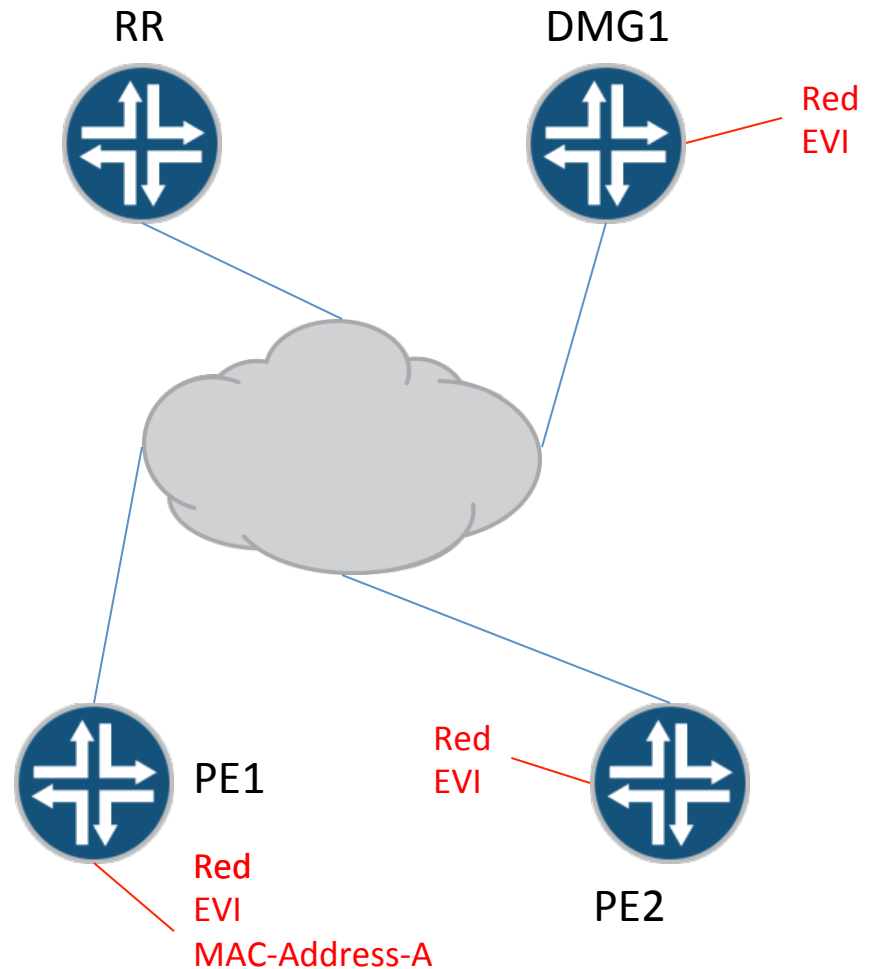
Route Refresh Message With CP-ORF

- AFI = L2VPN
- SAFI = BGP EVPN
- When-to-refresh = IMMEDIATE
- ORF Type = CP-ORF (value TBD)
- ORF entry
  - Action = ADD
  - Match = PERMIT
  - Type Specific Information

CP-ORF Type Specific Information

```
+----------------------------------+
| Sequence (32 bits)               |
+----------------------------------+
| Minlen (8 bits)                  |
+----------------------------------+
| Maxlen (8 bits)                  |
+----------------------------------+
| VPN Route Target (64 bits)       |
+----------------------------------+
| Import Route Target (64 bits)    |
+----------------------------------+
| Host Address (48 bits)           |
| ....                             |
+----------------------------------+
```

Solution In Detail

- At startup, PE2 establishes BGP session with RR
  - Negotiates CP-ORF Capability
  - Negotiates Multiprotocol Extensions Capability
- PE2 sends RR a Route Refresh message containing no ORF entries
  - RR sends PE2 the DMR
    - Next-hop = DMG1
    - RT = RT-RED-FROM-DMG1
- Later, PE2 detects an "exceptional" flow to MAC-Address-A
- PE2 sends RR a Route Refresh message containing CP-ORF entry
  - RR refreshes advertisements to PE2, sending route to MAC-Address-A
- PE2 periodically withdraws ORFs that are no longer required

Solution In Detail: RR Perspective

- RR validates ROUTE REFRESH
  - Ignore entire message if invalid
- If the Action is ADD, RR adds the CP-ORF entry to the Outbound Filter associated with the peer
- If the Action is REMOVE, RR removes the CP-ORF entry from the Outbound Filter associated with the peer
- If the Action is ADD, RR check routes in Loc-RIB for CP-ORF match condition:
  - Route length >= minlen + 64 (for L2VPN, minlen is 48)
  - Route length <= maxlen + 64 (for L2VPN, maxlen is 48)
  - the route carries RT whose value is the same as the CP-ORF VPN Route Target
  - the route matches the CP-ORF Host Address
- Place matching routes into Adj-RIB-Out associated with the peer
- Add CP-ORF Import Route Target to the matching routes that are in Adj-RIB-Out
- Send newly added routes to the peer

Benefit of Route Refresh Semantic

- A BGP speaker can respond to a ROUTE REFRESH message containing ORFs by refreshing all routes or by refreshing only those routes affected by the ORFs [RFC 5291]. The CP-ORF draft recommends that the BGP speaker refreshes only those routes that are affected by the ORFs.

- Because ORFs are carried by ROUTE REFRESH messages, they are not propagated. In this application, propagation is neither required nor desirable. The application requires PE routers to pull routes from the route reflector. It does not require the pull to be propagated

- ORF is easily extensible

Next Step

- Draft currently describes distribution of EVPN MAC/IP Advertisement Routes
- Does not describe distribution of other EVPN route types
  - Will be covered in next draft version
    - Shortly after IETF 89
- Cross WG effort
  - L3VPN/L2VPN
  - Asking for adoption in L3VPN with L2VPN cooperation