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

# VIRTUAL HUB-AND-SPOKE VPN: A BRIEF REVIEW

Goal

- Reduce the number of routes that V-Spoke1 carries

- V-Spoke1 carries only one IP Default route per VPN
  - Next-hop == V-Hub1

- Traffic from V-Spoke1 traverses V-Hub1

- Traffic to V-Spoke 1 may traverse a more direct route

RR

V-Hub1

Red VPN Site

Red VPN Site

PE1

Red VPN Site (192.0.2.0/24)
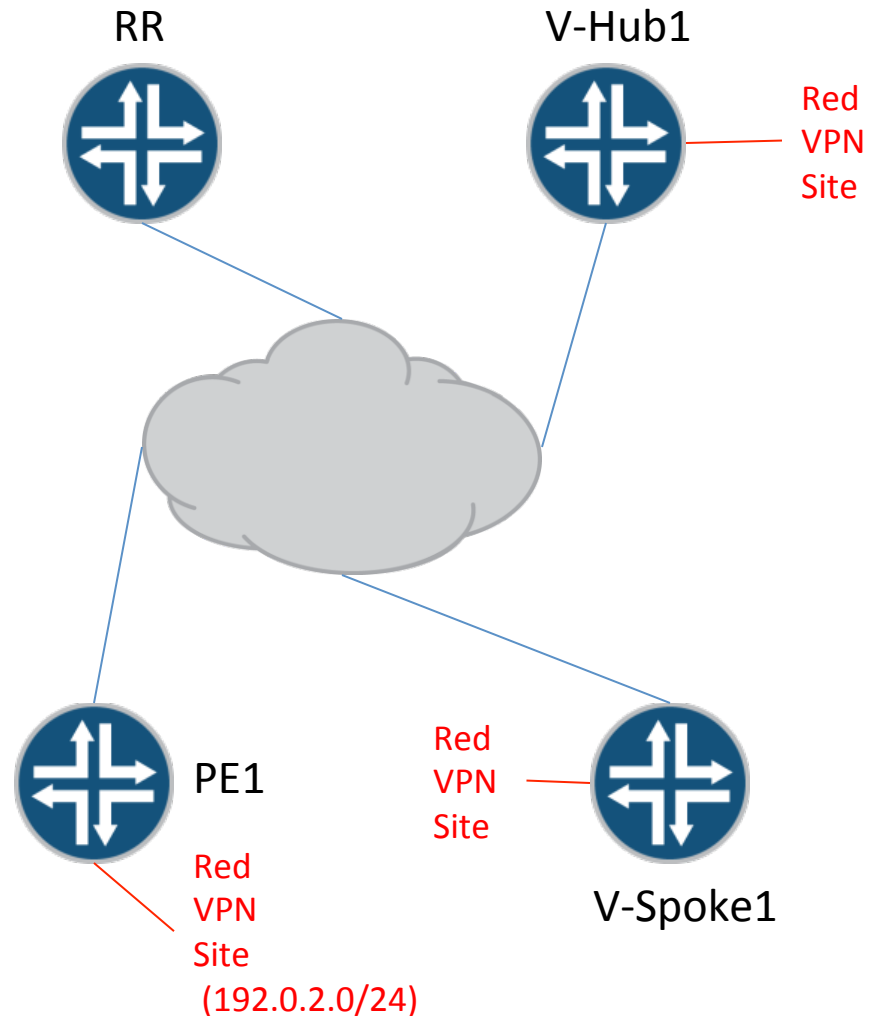
V-Spoke1

BGP Routing Policy

- PE1 and V-Hub1 are clients of a RR
  - V-Spoke1 may be client of RR or V-Hub1
- PE1 and V-Hub1 accept advertisements carrying the RT, RT-RED
- V-Spoke1 accepts advertisements carrying the RT, RT-RED-FROM-HUB1

RR

V-Hub1

Red VPN Site

PE1

Red VPN Site

Red VPN Site
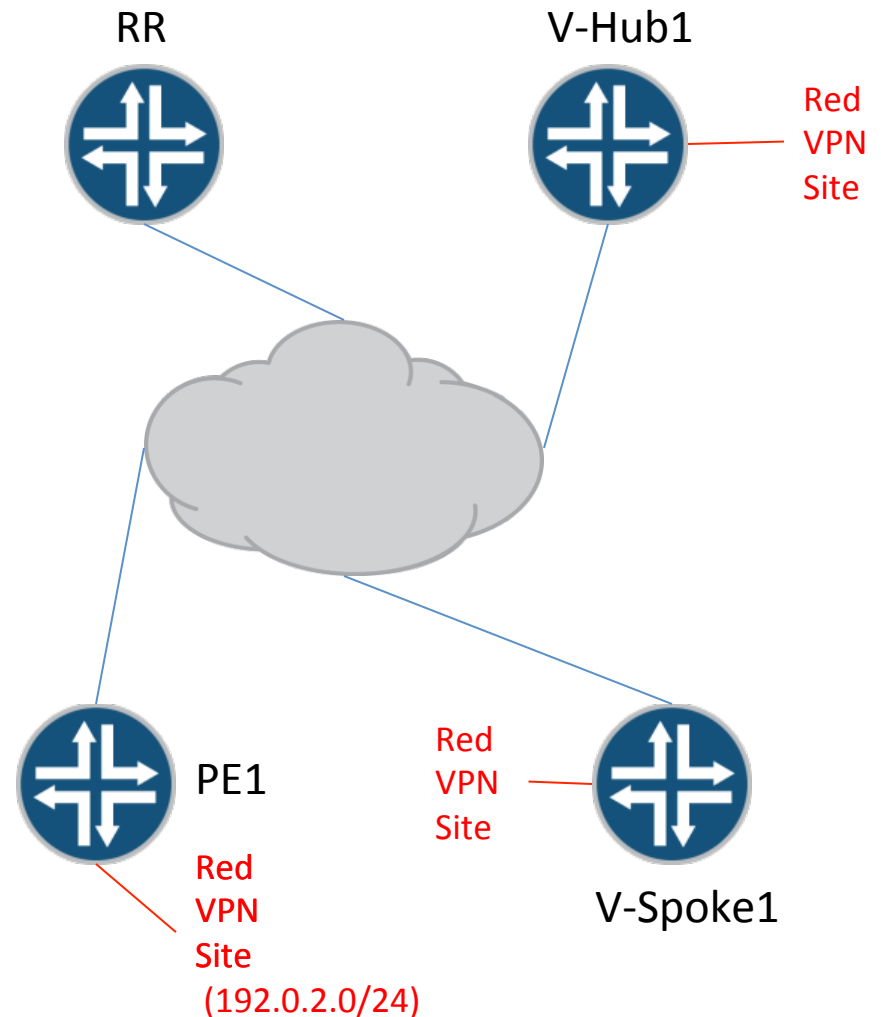
(192.0.2.0/24)

V-Spoke1

BGP Advertisements

- PE1 advertises 192.0.2.0/24 to the RR
  - Next-hop = Self
  - RT = RT-RED
- RR reflects route to V-Hub1
  - V-Hub1 accepts
- RR may also advertise route to V-Spoke1
  - In absence of RT-Constrain
  - If advertised, V-Spoke1 rejects

RR

V-Hub1

Red VPN Site

PE1

Red VPN Site

Red VPN Site

(192.0.2.0/24)
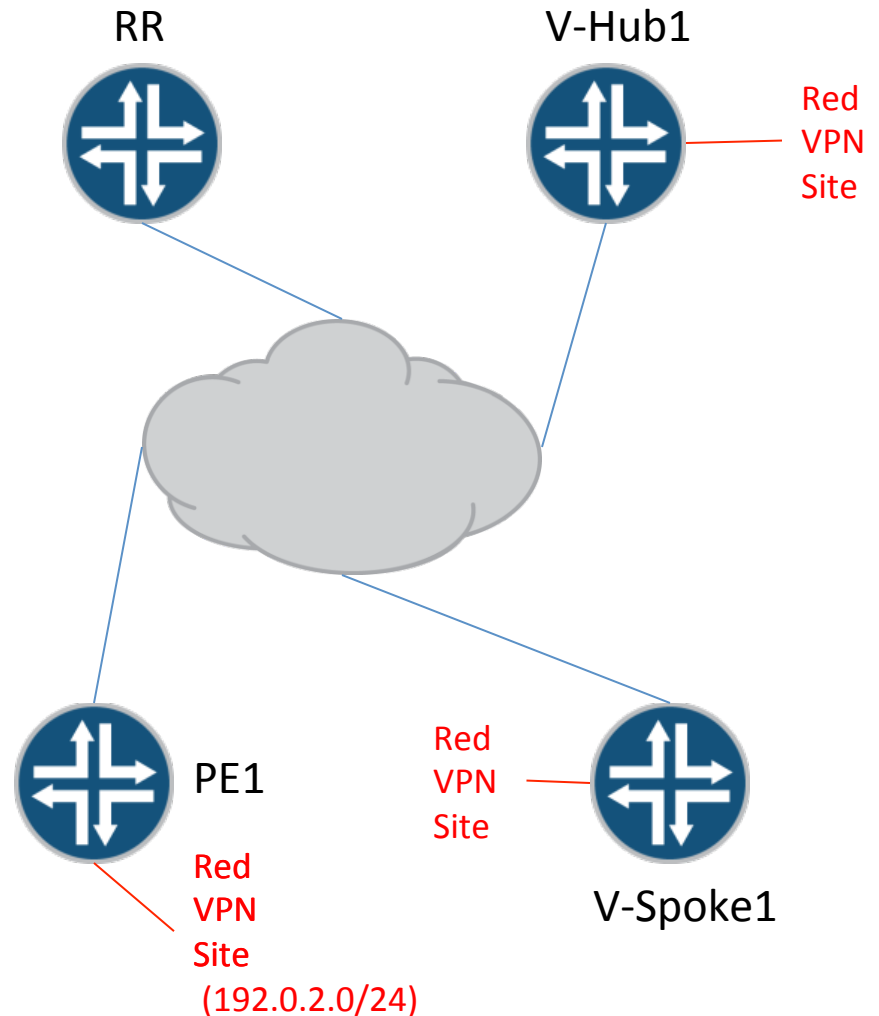
V-Spoke1

BGP Advertisements (continued)

- V-Hub1 advertises VPN-IP default route to the RR
  - Next-hop = Self
  - RT = RT-RED-FROM-HUB1
- RR reflects route to V-Spoke1
  - V-Spoke1 accepts

RR

V-Hub1

Red VPN Site

PE1

Red VPN Site

Red VPN Site
(192.0.2.0/24)
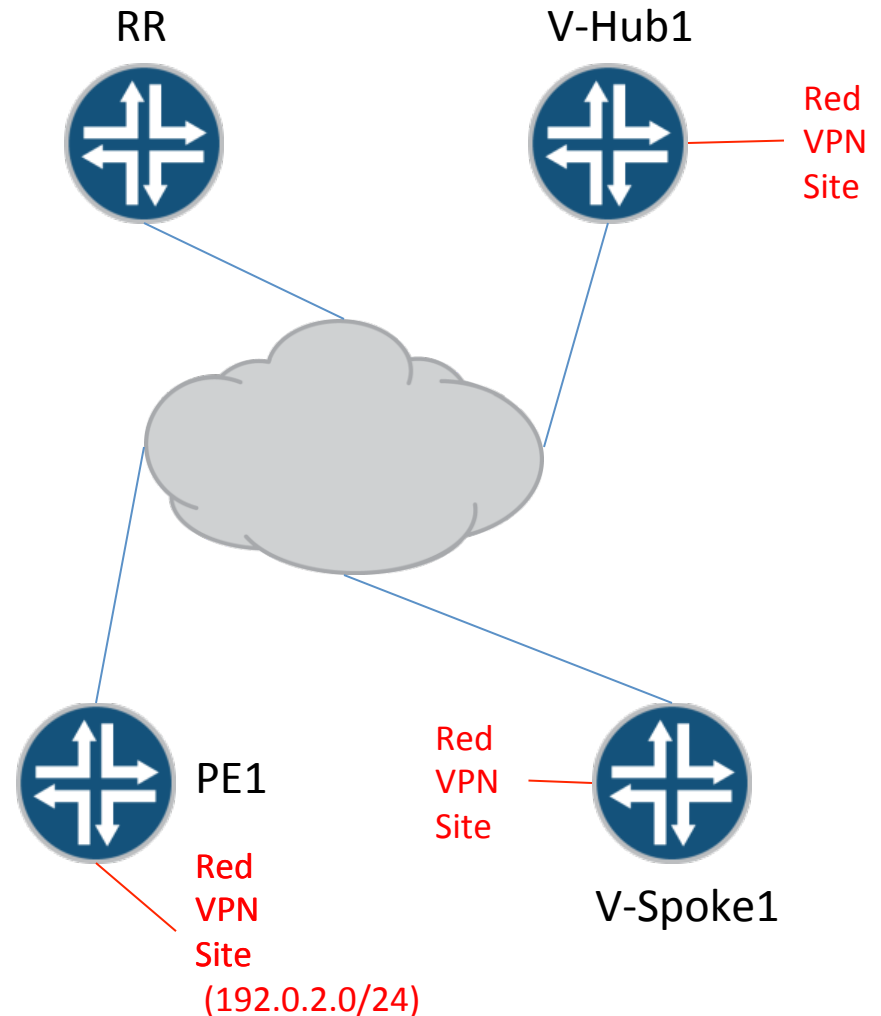
V-Spoke1

COVERING PREFIX ORF

Problem to Be Solved

- The VPN site served by V-Spoke1 originates an "exceptional" flow to 192.0.2.1
  - Large, latency sensitive, etc.
- Flow traverses V-Hub1
- Flow might benefit from a more direct route to 192.0.2.1
  - If such a route exists
- The criteria determining that a flow might benefit from a more direct route are strictly local to V-Spoke1

RR

V-Hub1

Red VPN Site

PE1

Red VPN Site

Red VPN Site
(192.0.2.0/24)
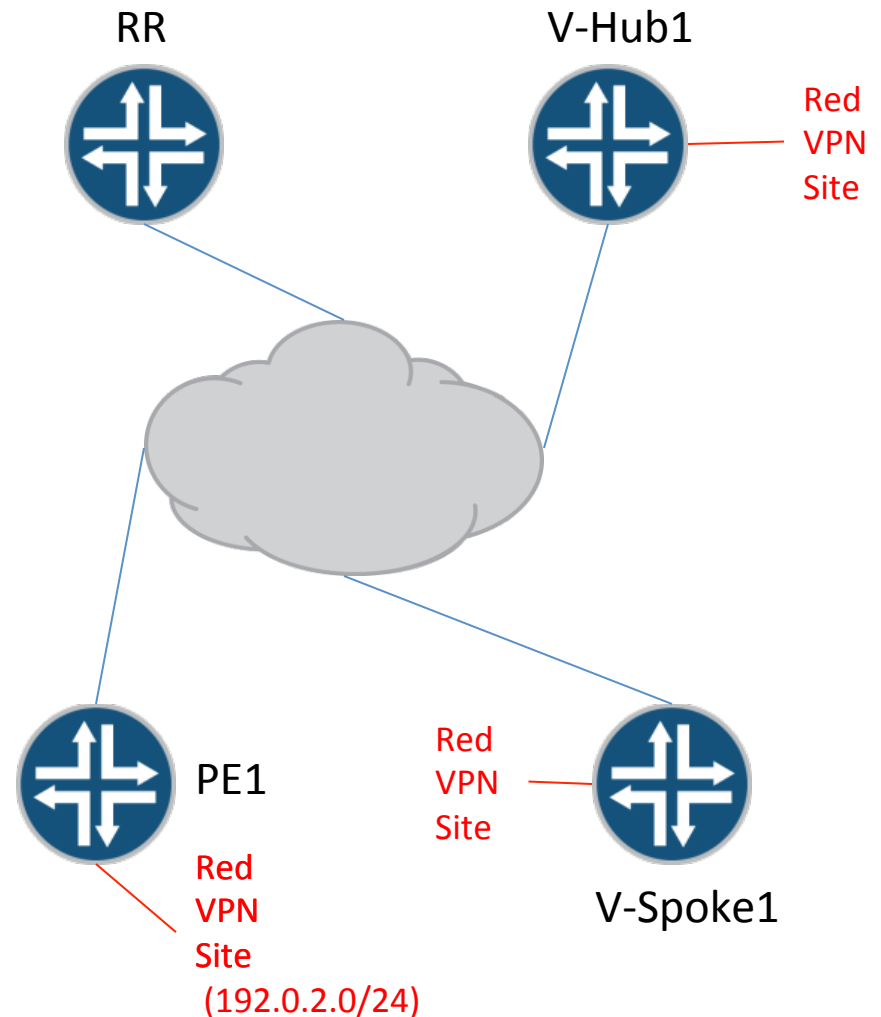
V-Spoke1

Solution

- V-Spoke1 requests the most specific route covering 192.0.2.1 from the RR

  – Carrying additional RT, RT-RED-FROM-HUB1

- Pull versus push

RR

V-Hub1

Red VPN Site

PE1

Red VPN Site

Red VPN Site
(192.0.2.0/24)

V-Spoke1
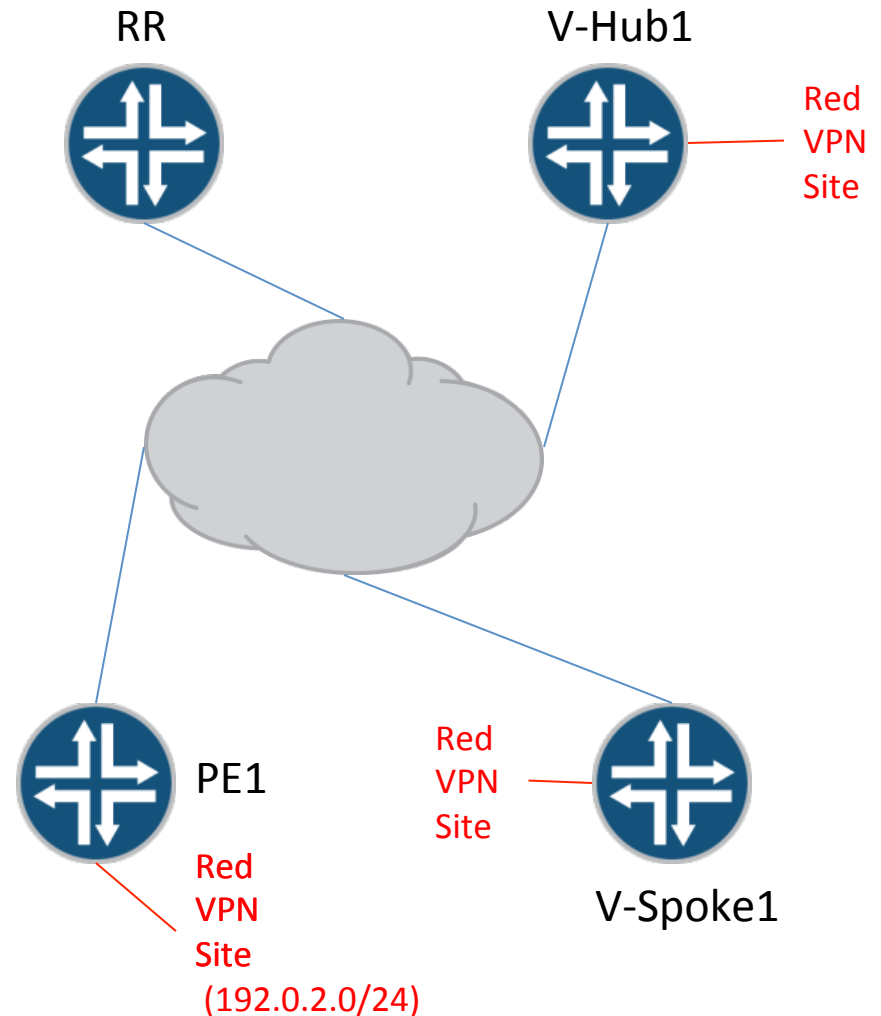
Route Refresh Message With CP-ORF

- AFI = IPv4 or IPv6
- SAFI = MPLS-Labeled-VPN-Address
- When-to-refresh = IMMEDIATE
- ORF Type = CP-ORF (value TBD)
- ORF entry
  - Action = ADD or REMOVE
  - 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 (32 or 128 bits)          |
| ....                                   |
+----------------------------------------+
```

Solution In Detail

- At startup, V-Spoke1 establishes BGP session with RR
  - Negotiates CP-ORF Capability
  - Negotiates Multiprotocol Extensions Capability
- V-Spoke1 sends RR a Route Refresh message containing no ORF entries
  - RR sends V-Spoke1 IP VPN default route
    - Next-hop = V-Hub1
    - RT = RT-RED-FROM-HUB1
- Later, V-Spoke1 detects an "exceptional" flow to 192.0.2.1
- V-Spoke1 sends RR a Route Refresh message containing CP-ORF entry
  - RR refreshes advertisements to V-Spoke1, sending route(s) covering 192.0.2.1 (i.e., 192.0.2.0/24)
- V-Spoke1 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 prefix length >= minlen + 64
  - Route prefix length <= maxlen + 64
  - the route carries RT whose value is the same as the CP-ORF VPN Route Target
  - the route covers 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 spokes to pull routes from the route reflector. It does not require the pull to be propagated

- ORF is easily extensible

Conclusion

- Adopt as WG draft