BGP Vector Routing

draft-patel-raszuk-bgp-vector-routing-01

Keyur Patel, Robert Raszuk, Burjiz Pithawala, Ali Sajassi, Eric Osborne, Jim Uttaro, Luay Jalil

IETF 88, November 2013, Vancouver, Canada

Motivation

- Network Architectures require additional control over the traffic paths (Inter as well as Intra domain)
  - Need to force the traffic to go through one or more Transit Nodes
  - Transit Nodes could be a TE Node
  - Other examples include Service Nodes like: Firewall, NAT, Load Balancers, etc
- Need a **scalable control plane solution** to advertise "information" so that the traffic gets routed through an ordered set of Transit points before it is forwarded to its destination
  - In context of Transit points as Service Nodes it is known as "Service Chaining". Otherwise it is known as "Traffic Engineering" (TE)

BGP Vector Routing

- BGP based mechanism to create arbitrary forwarding topologies as well as facilitate Service Chaining
  - Does not require changes to the forwarding plane
  - Assumes use of an existing encapsulation/tunneling techniques to forward data
- New BGP attribute known as a Vector Node attribute
- Vector Node attribute consist of one or more TLVs
  - TLVs carry ordered lists of IP Transit Hops that needs to be traversed before the packet is forwarded to its destination
  - TLV information is used to replace the NEXTHOP information when installing the route in RIB/FIB
- Two new TLVs defined as part of this draft
  - Type 1 and Type 2 TLV
- Rules to process and use TLV information of Vector Node Attribute

BGP Vector Routing (Cont'd)

- BGP Vector Node attribute can be applied to any BGP Address Family
- Creation of a BGP Vector Node attribute is outside the scope of the document
  - Assumed to be created using CLI on a router or using an Orchestrated system, or by some automated SDN policy computing engines
- Vector Node attribute is usually inserted at a single point in the network and advertised by BGP to all BGP speakers

BGP Vector Node Attribute TLVs

- TYPE1 TLV consists of a Vector Node address
  - Vector Node address is an address of a transit (services) router and is typically announced in an IGP protocol
- TYPE2 TLV consists of a Vector Node and a Service Node address
  - Vector Node address is an address of a Transit Services router and is typically announced in an IGP protocol
  - Service Node address is an address of a Service Appliance and is directly connected to Vector Node address and is not announced in an IGP. Alternatively Service Node Address could be a Local ID of a Transit Service Router pointing to an Appliance
  - Vector Nodes and Service Nodes may belong to a different Address Families
- Both the TLVs carry AS Number to facilitate Inter-AS announcements

BGP Vector Node Attribute Rules

- 4 Rules defined to process the BGP Vector Node Attribute
- 1st Rule describes Vector Node attribute and AS Number Validation
  - Missing Attribute or a failing AS Number Validation results in use of a BGP address from BGP MP_REACH attribute or from a NEXT_HOP attribute (if BGP MP_REACH Attribute is NOT present) as a NEXTHOP address when adding a route to RIB/FIB
- 2nd Rule describes a case where an AS Number Validation succeeds but a BGP Speaker Address (loopback or connected) is missing in the Vector Node Attribute
  - In such a case BGP Speaker should use the First TLV Vector Node address as a NEXTHOP address when adding a route to RIB/FIB
- 3rd Rule describes a case where an AS Number Validation succeeds but a BGP Speaker Address (loopback or connected) is present in the Vector Node Attribute TLV
  - In such a case BGP Speaker should use the next eligible Vector Node address as a NEXTHOP address when adding a route to RIB/FIB

BGP Vector Node Attribute Rules (Con't)

- 4th Rule describes a case where an AS Number Validation succeeds but a BGP Speaker Address (loopback or connected) is present as the Last Vector Node Attribute TLV address
  - In such a case BGP Speaker should use the BGP address from BGP MP_REACH attribute or from a NEXT_HOP attribute (if BGP MP_REACH Attribute is NOT present) as a NEXTHOP address when adding a route to RIB/FIB

Questions?

Request WG to adopt the draft as a WG document.