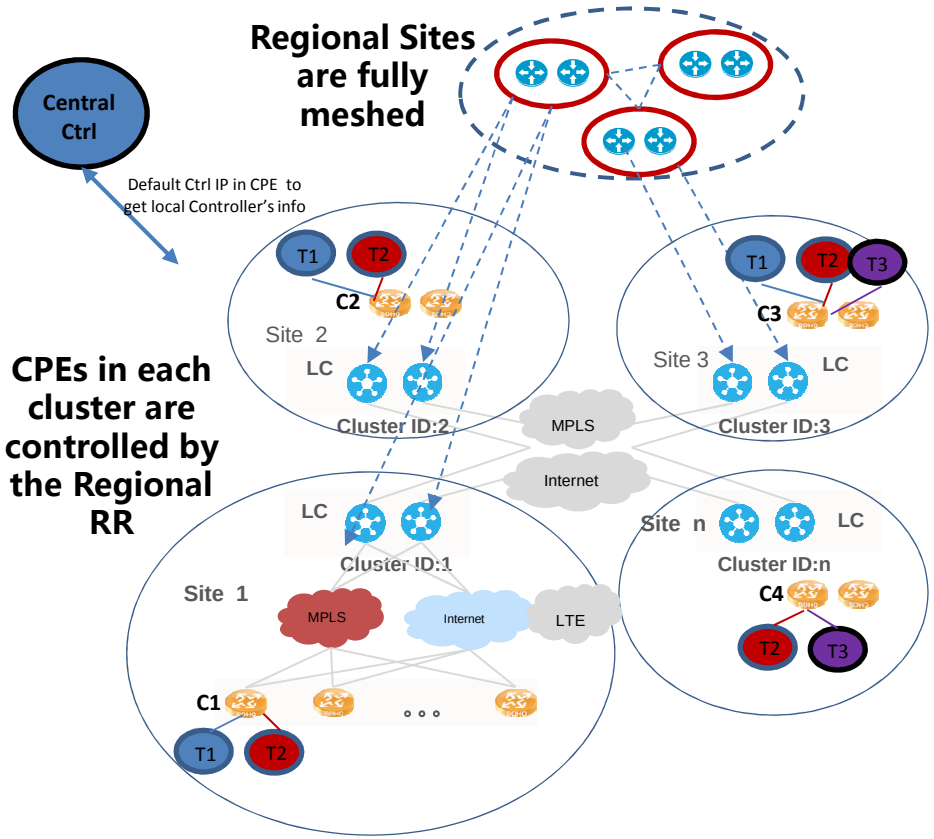
BGP Extension for SDWAN Overlay Networks

draft-dunbar-idr-bgp-sdwan-overlay-ext-01

Defines a new BGP SAFI with a new NLRI in order to advertise a SD-WAN edge node's capabilities in establishing SD-WAN overlay tunnels with other SD-WAN nodes through third party untrusted networks.

Linda Dunbar
Wang Haibo
Hao WeiGuo
Oct 2018

Use Case: Managed Overlay WAN Services: 100's or 1000's CPEs



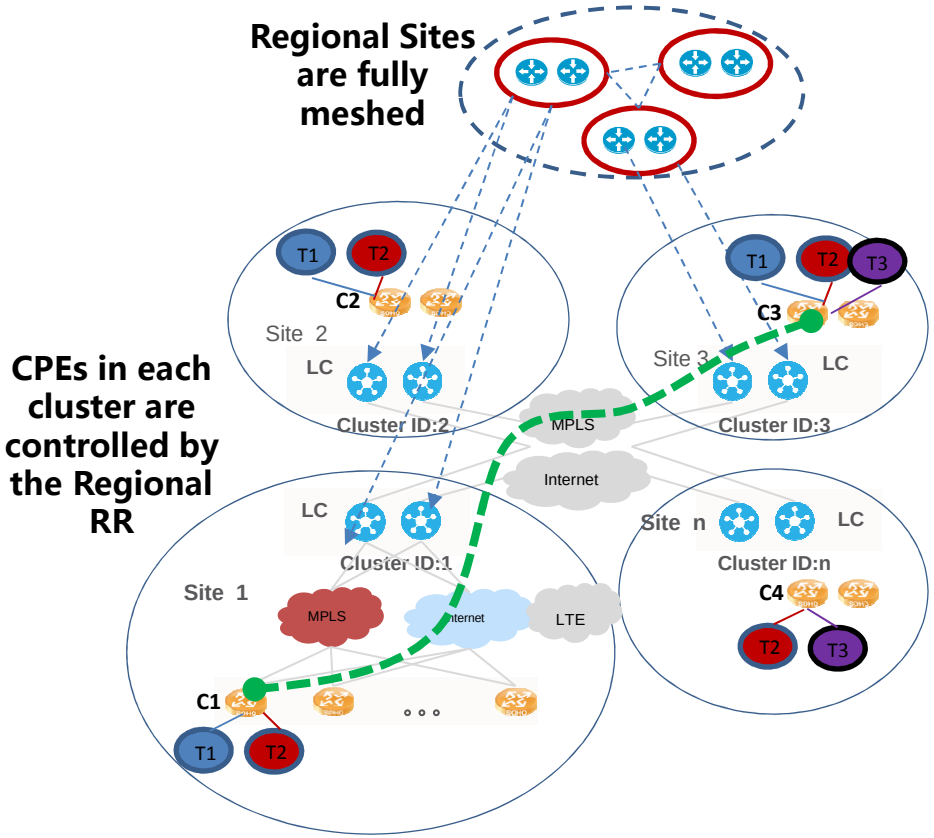
Characteristics of SD-WAN Deployment

- **Zero Touch Provisioning for CPEs**
  - Upon powered up, CPE sends request to its factory default Central Controller address to retrieve its local RR address.
- **The IP address of ports to a SD-WAN node can be dynamic (e.g. assigned by DHCP); therefore, there is no fixed IP address that can be used to uniquely to represent a SD-WAN tunnel end-point.**
- **Apps/Hosts attached to each CPE can belong to different tenants.**
  C1 node alone has to establish following SD-WAN tunnels:
  - two SD-WAN tunnels to C2: one for Tenant 1, another one for Tenant 2,
  - One SD-WAN tunnel to C3 for Tenant 1, and
  - two SD-WAN tunnels to C4: one for Tenant 1, another one for Tenant 2,

Hierarchical management

- **Partition CPEs into Sites, each Site is a logical entity for remote sites**
- **Why:**
  - Enable detour based on sites, instead of CPEs,
  - Avoid complexity of managing full mesh of all CPEs.
  - Hide CPE identity from others (some deployment needs this feature)

Goal: for End points to exchange information across untrusted underlays



a. to advertise the identifiers of ports that support establishing SD-WAN overlay tunnels to other peers,
b. to advertise ports private addresses (or dynamically assigned IP addresses),
c. to advertise its supported IPsec capability, such as the supported encryption algorithms,
d. etc.

Gap: RFC5512 & draft-ietf-idr-tunnel-encaps-10

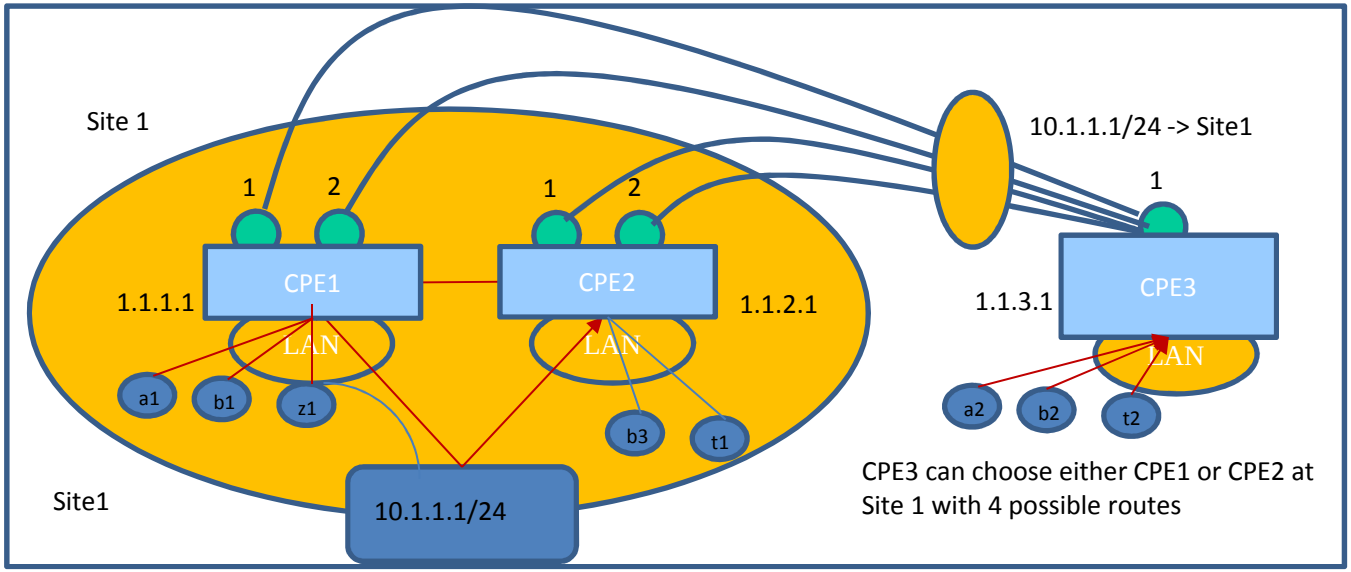
- Tunnel-Encap removed SAFI =7 (RFC5512) for distributing encapsulation tunnel information. Tunnel-Encap requires Tunnels being associated with routes.
- It was suggested to a “Fake Route” for a SD-WAN node to use [Tunnel-Encap] to advertise its SD-WAN tunnel end-points properties:
  - using “Fake Route” can create deployment complexity for large SD-WAN networks with many tunnels. E.g. ,
    - for a SD-WAN network with hundreds of nodes, with each node having many ports & many end-points to establish SD-WAN tunnels to their corresponding peers, the node would need many “fake addresses”. For large SD-WAN networks (such as has more than 10000 nodes), each node might need 10’s thousands of “fake addresses”, which is very difficult to manage and needs lots of configuration to get the nodes provisioned.

Gap of draft-rosen-bess-secure-l3vpn-01

- The use case is specific about a remote CPE node to be integrated with the L3VPN network. With IPsec key pre-provisioned.
- It assumes that C-PE and RR are connected by IPsec tunnel.
  - With zero touch provisioning, we need an automatic way to synchronize the IPSec SA between C-PE and RR. The draft assumes:
  - A C-PE must also be provisioned with whatever additional information is needed in order to set up an IPsec SA with each of the red RRs
- No periodic refreshment of the keys.
- IPsec usually only send configuration parameters to two end points and let the two end points to negotiate the KEY. Now we assume that RR is responsible for creating the KEY for all end points. When one end point is compromised, all other connections are impacted.

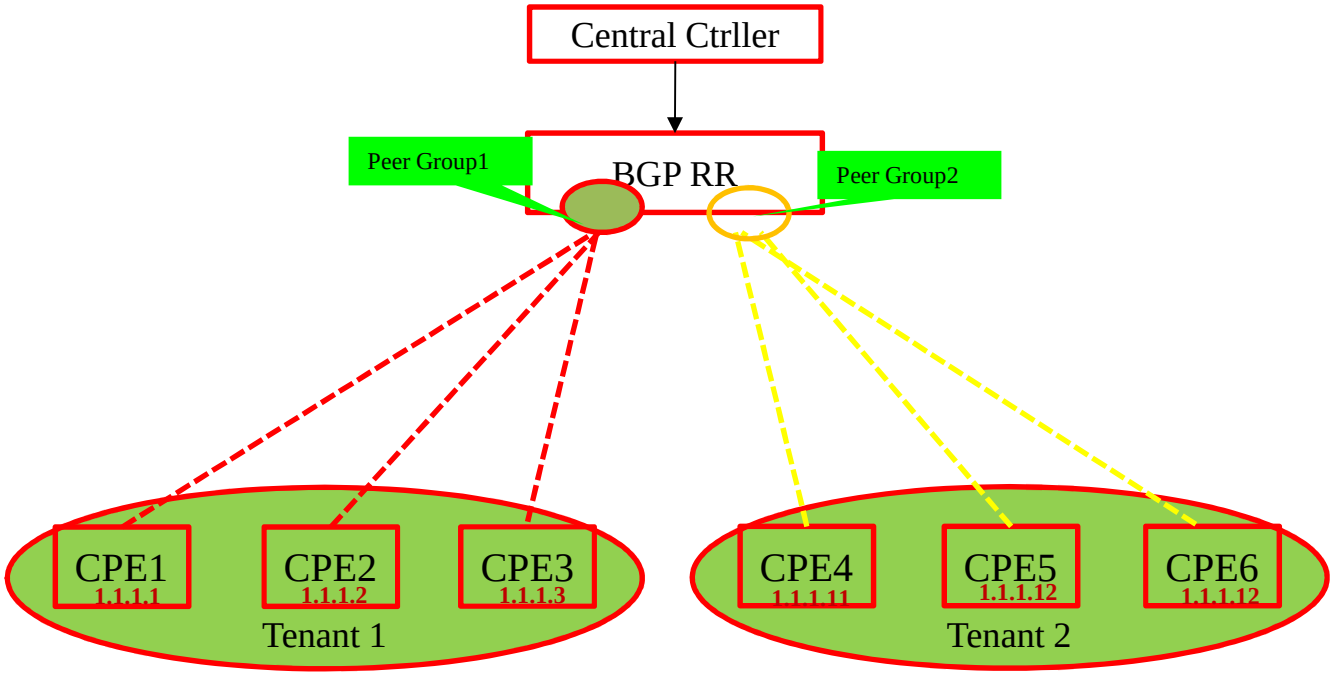
BGP Solution Briefing

- Unlike NHRP/DSVPN, every node use BGP to distribute its Tunnel end point properties
- Defines a new BGP SAFI with a new NLRI in order to advertise a SD-WAN edge node's capabilities in establishing SD-WAN overlay tunnels with other SD-WAN nodes through third party networks.
  - The goal is for SD-WAN network to scale, enabling SD-WAN overlay tunnels among large number of SD-WAN nodes to be established with few provisioning needed



Tunnel Information Advertisement Method

- Tenant Separation Method：



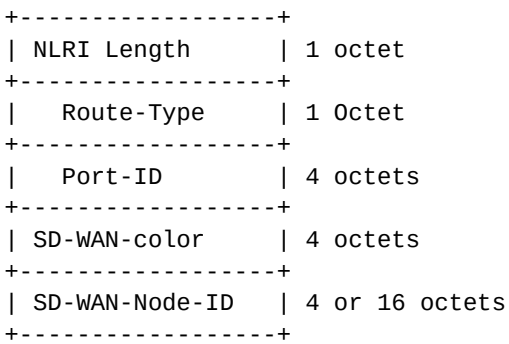
CPE1：
- Receiving SD-WAN IPSEC infor、report WAN ports information to Controller via SD-WAN SAFI
- RR send to CPE2、CPE3（using Policy Filtering to only send to Peers belong to same tenant)
- CPE2、CPE3 upon receiving SD-WAN IPSEC config information, start to negotiate with peers on IPsec tunnel establishing and establish the key.
- SD-WAN IPSEC tunnel is added to the Service Tunnel (IDR-Tunnel-encap) to be used for Route Advertisement

For Tenant Separation: CPEs belonging to same Tenant are added to a Peer Group

```
peer group1 route-policy tenant1-in import
peer group1 route-policy tenant1-out export
route-policy tenant1-in permit node 10
  apply community 100:1 addtitive
route-policy tenant1-out permit node 10
  if-match community-filter 1
ip community-filter 1 permit 100:1
```

Others are configured in similar way

New NLRI for SD-WAN Overlay



- Route-Type: to define the encoding of the rest of the SD-WAN Overlay NLRI.
- Port ID: one (SD-WAN) node can have multiple ports, and each port can support multiple SD-WAN tunnels to different peers. The Port ID is used to identify the port, a.k.a. link identifier.
- SD-WAN-color: used to identify a common property shared by a set of SD-WAN nodes, such as the property of a specific geographic location.
- SD-WAN Node ID: the SD-WAN NLRI advertisement is sent out by the SD-WAN node to indicate all the available ports supporting SD-WAN tunnels. The SD-WAN Node ID can be the node's system ID, such as the loopback address of the SD-WAN node.

BGP Extension Details

- Utilize the Tunnel Encapsulation Attribute specified in draft-ietf-idr-tunnel-encaps-10
  - Tunnel Type: SD-WAN-Tunnel
  - EncapExt SubTLV (value is 128-255)
    - for describing additional information about the SD-WAN tunnel end-points, such as NAT property.
  - IPsec-SA Attribute SubTLV
    - for establishing IPsec SA with other peers.