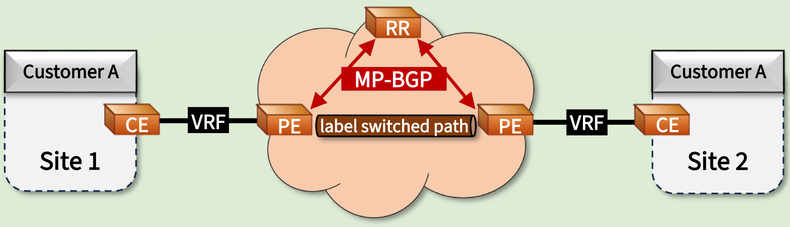
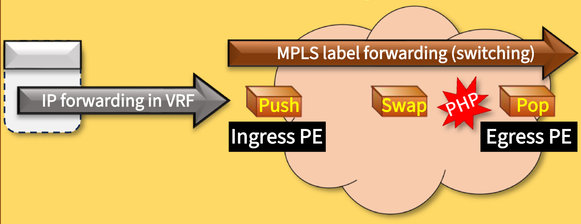
MPLS Layer-3 VPN Introduction

MPLS Layer-3 VPN is a highly scalable service provider technology that enables private network deployment for any number of independent customers. At its foundation, an MPLS L3VPN architecture uses Virtual Routing and Forwarding (VRF) tables as well as Multiprotocol BGP (MP-BGP) VPN routes to separate each customer's network over a shared physical topology. An MPLS L3VPN can leverage a single core network bordered by Provider Edge (PE) routers to create a multi-tenant architecture.



Key points about architecture

- The control plane uses Multiprotocol BGP (MP-BGP) to exchange VPN prefixes between the PE routers, and BGP Route Reflectors (RR) are configured to improve scalability, redundant pairs of RRs are commonly deployed.
- The data plane uses MPLS encapsulation and label switching. Lookups are performed in the Label Forwarding Information Base (LFIB) to determine the label and next-hop. OSPF or IS-IS is used in the MPLS core network. Label Distribution Protocol, MPLS Traffic Engineering tunnels or Segment Routing signals a label switched path.
- An MPLS label stack is a key concept that enables the L3VPN architecture. There are three mechanisms carried out by the Label Switching Routers (LSR):
    1. Label Push – insert label to enable MPLS forwarding (label switching)
    2. Label Swap – replace existing label with a new label at each next-hop
    3. Label Pop – remove a label e.g. due to PHP* or to enable IP forwarding



***Penultimate Hop Pop**
is performed at the last hop before traffic reaches the Egress PE router. At this last hop, the IGP top label is removed (popped), and only the VPN label remains. PHP is signaled by the Egress PE using the implicit NULL label with the reserved value 3.

- The MPLS label stack is pushed by the Ingress PE router and ensures the following two requirements:
    - That the correct Egress PE router is found in the MPLS core network, this is ensured by the top label aka transport label, IGP label, next-hop label, or outer label. These all refer to the same label shown in the packet capture below. At each next-hop within the MPLS core network the top label is swapped.
    - That the correct customer VRF interface is found on the Egress PE, this is ensured by the bottom label aka VPN service label, or inner label. The VPN service label is advertised with MP-BGP from the Egress PE, but inserted (pushed) on the Ingress PE. Each Egress PE router signals its connected VPNs.
- The MPLS label stack is visible in the following packet capture. There are two MPLS headers inserted, 405 is the top label, 609 is the bottom label. As visible, there is no encryption available by default. In order to encrypt label switched traffic, GETVPN can be deployed. GETVPN uses IP header duplication to ensure QoS.

```
Source          Destination     Protocol  Info
192.168.1.1     192.168.2.1     TELNET    Telnet Data ...
‹
› Frame 1: 70 bytes on wire, 70 bytes captured
› Ethernet II, Src: 52:54:00:1f:80:37 (52:54:00:1f:80:37), Dst: 52:54:00:01:44:90 (52:54:00:01:44:90)
› MultiProtocol Label Switching Header, Label: 405, Exp: 6, S: 0, TTL: 252
› MultiProtocol Label Switching Header, Label: 609, Exp: 6, S: 1, TTL: 253
› Internet Protocol Version 4, Src: 192.168.1.1, Dst: 192.168.2.1
› Transmission Control Protocol, Src Port: 35592, Dst Port: 23, Seq: 1, Ack: 1, Len: 8
⌄ Telnet
    Data: P4SSW0RD
```

MPLS Forwarding Equivalence Class (FEC)

Ingress PE router assigns each packet to an MPLS FEC as packets enter the MPLS core network. The assigned MPLS label value identifies the FEC.

On the Ingress PE router, packets forwarded out on the same interface to the same next-hop router and with an identical queuing policy belong to the same FEC.

The FEC is assigned only once and on the Ingress PE router. In other words, the VPN service label advertised by the Egress PE router already tells the Ingress PE router which FEC should be assigned to the incoming packet. Along the way from the Ingress PE to the Egress PE the FEC is not changed.

The FEC and the Label Forwarding Information Base (LFIB) work together. Once the FEC is determined, the LFIB is used to forward packets without a Layer-3 IP lookup. The LFIB contains label values and next-hop information to enable label switching.

Five facts about MPLS L3VPN

- MP-BGP control plane can carry routes for a variety of address families, such as VPNv4 for IPv4, VPNv6 for IPv6, or multicast VPNs. In addition to VPNv4/v6 routes, MP-BGP EVPN (Route Type 5) can also be used as the control plane for an MPLS L3VPN. This is called EVPN over MPLS IRB.
- The access network between the PE – CE router is configured in a VRF and commonly carries routes using eBGP. However, the PE – CE routing protocol can also be EIGRP, OSPF or RIP.
- Quality of Service (QoS) in an MPLS L3VPN can be configured end-to-end for each customer site per-VRF. Between the PE-CE the IP DSCP field is used. The PE performs ToS Reflection, which is to map the IP Precedence (first three DSCP bits) to the MPLS header EXP field. The MPLS EXP is called MPLS Traffic Class (TC) since RFC 5462. The MPLS network uses the EXP/TC field for QoS.
- The MPLS core network operates with label switching. An MPLS header is 4-bytes long. A L3VPN MPLS label stack consists of at least 2 headers, which adds 8-bytes to each packet. MPLS data encapsulation is sometimes compared to tunneling due to a number of similarities.
- With MPLS L3VPN each customer lets the service provider handle its WAN routing table. There is a trust relationship because the customer outsources its enterprise routing table to the SP.