# IP/MPLS L3VPN

Rodrigo Ferreira

2025

# Contents

# Glossary

**ASBR** Autonomous System Boundary Router

**BGP** Border Gateway Protocol

**CE** Customer Edge

**CoS** Class of Service

**L2VPN** Layer 2 Virtual Private Network

**L3VPN** Layer 3 Virtual Private Network

**LDP** Label Distribution Protocol

**MP-BGP** Multiprotocol BGP

**MPLS** Multiprotocol Label Switching

**MVPN** Multicast Virtual Private Network

**OSPF** Open Shortest Path First

**P** Provider Router

**PE** Provider Edge

**RD** Route Distinguisher

**RIP** Routing Information Protocol

**RT** Route Target

**SP** Service Provider

**VPN** Virtual Private Network

**VRF** VPN Routing and Forwarding

# 1 Introduction

## 1.1 Requirements

Private network connectivity between geographically dispersed sites for organisations and individual, to access other devices in said network while also being able to be provided with a connection to the Internet is a modern requirement. A VPN provides this connectivity over a shared public or SP network. L3VPNs, also known as BGP/MPLS IP VPNs are defined in RFC 4364 [1] as a solution for this requirement.

Allowing for an SP to provide a scalable, secure, and flexible IP VPN service, meaning a layer 3 (network layer) isolation between different customers/users while leveraging the same physical infrastructure on the layers beneath the network layer, as a L3VPN allows for a logical distribution of resources. This architecture comes with two challenges:

- **Address Space Overlap:** Different VPNs may use the same private IP address spaces
- **Scalability:** The SP's core routers (P routers) should not need to store the routing information for every customer VPN as doing so would create a massive scalability bottleneck.
- **Isolation:** traffic from one VPN must be kept completely separate from another.

By combining a logical partitioning of routing tables on the PE routers with MP-BGP, and an efficient data plane, MPLS an SP should be able to provide L3VPN services to their customers.

## 1.2 L2VPN vs L3VPN

The main difference between L2 and L3 VPNs is in the network layer at which the SP participates in the customer's network functionality. This is an architectural choice between who manages routing control and what L3 protocols to be used, as well as the scalability of the solution. L2VPN are usually setup to connect datacenters while L3VPNs are used for most other type of networks.

- **Operation:**
  - **L2VPNs** operate at the Data Link Layer, with the SP network acting as a transparent tunnel / virtual switch (e.g. VPLS). Forwarding relies on Layer 2 information or virtual circuit identifiers, leveraging technologies like Ethernet and optic cables.
  - **L3VPNs** operate at the Network Layer. SP's PE routers are active participants, using IP addresses and routing and forwarding tables known as VRFs, leveraging MPLS IP.
- **Routing Responsibility:**
  - In **L2VPNs**, routing control is mainly customer manged after informing and obtaining authorizations from the SP. The SP does not participate in the Layer 3 control plane, making the service network protocol agnostic.

  - In **L3VPNs**, routing control is *fully SP managed. The PE router actively exchanges Layer 3 routing information with the CE router and the SP manages the whole systems and protocols.

- **Customer IP Space and Scaling:**
  - **L2VPNs** are suitable for extending a customer's L2 broadcast domain and allow the customer to use the same IP subnet across connected sites., although typically less scalable for complex, full-mesh enterprise WANs.
  - **L3VPNs** are highly scalable and designed for large-scale, complex multipoint-to-multipoint mesh networks. Each customer site is either a separate IP subnet, or an organisation's network with multiple geographically separated systems that the SP handles the inter-subnet routing for.

# 2 Architecture

The L3VPN architecture is built on a few key components and three distinct planes of operation: traffic isolation, control routing and data forwarding.

## 2.1 Components

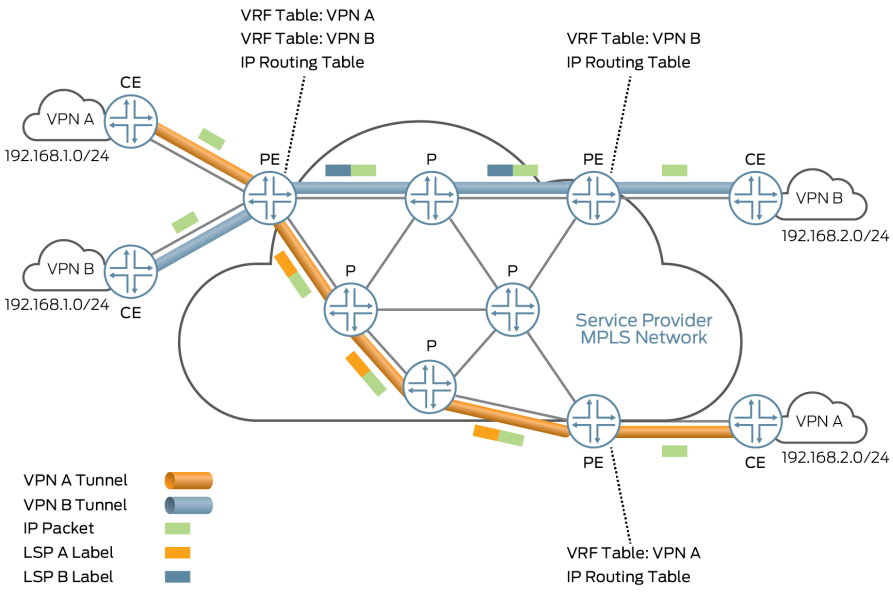


Figure 1: L3VPN Components Overview

- **Customer Edge Router:** A device at the customer's site connected to the SP's network that is a routing peer with the PE.
- **Provider Edge Router:** An SP router at the edge of the provider's backbone responsible for the main routing responsibilities. PEs connect to CEs, maintain separate routing tables for each VPN, and peer with other PEs to exchange VPN routes.

- **Reflector Router:** Used to reflect IP-VPN MPLS labeled packets inside the backbone.
- **Provider Router:** P routers connect PEs to each other and have no knowledge of any customer VPN routes. Their role is to forward MPLS IP packets efficiently towards the PE routers.

## 2.2 VPN Routing and Forwarding

The main mechanism for traffic isolation on a PE router is the VRF table. A PE router maintains multiple independent routing and forwarding tables. The default table for the provider's own routes and the public internet, and one VRF for each customer VPN site it connects to.

Each system from a CE VPN is associated by configuration with a specific VRF table. When a packet arrives to the PE it is processed against the routing table of its associated VRF ensuring that two customers with overlapping addresses are kept isolated from each other as they are logical routers within the PE.

## 2.3 Routing Protocols

- **MP-BGP:** Used to carry routes from multiples addresses families leveraging VPN-IPv4, BGP leverages MPLS to add an extra label to the IP packet received from the CE to route it inside the provider's backbone to the PE on the other side.
- **EBGP:** Used between the PE and CE is a scalable approach
- **IBGP:** Used to distribute eBGP-learned routes inside the provider's backbone.
- **OSPF:** Used between the PE and CE is a fast approach

## 2.4 MP-BGP and VPN-IPv4

To connect customer sites, PEs must exchange routing information, leading to the problem of overlapping addresses, which should be solved using VPN-IPv4 label switching.

- **RD:** When a PE learns an IPv4 route from a CE's VRF, it prepends a unique 8-byte RD.
- **VPN-IPv4 Address:** The resulting 12-byte address is called a VPN-IPv4 address. Even if another customer uses the same prefix its VRF will have a different RD, making the two VPN-IPv4 routes unique.
- **MP-BGP:** PEs use an IBGP session running the VPN-IPv4 address family to advertise and receive these unique VPN-IPv4 routes to and from other PE routers. Some P routers acting as Route Reflectors is required.

## 2.5 Policy Control with RT

While an RD makes a route unique, an RT controls its distribution as BGP extended community attributes

- `vrf-export:` An export policy is applied to the VRF. This policy matches routes within the VRF and attaches one or more RT communities to them as they are advertised to other PEs via MP-BGP.
- `vrf-import:` An import policy is applied to the VRF. This policy examines all VPN-IPv4 routes received from other PEs. If a route contains a matching RT community, the PE removes the RD and imports the standard IPv4 route into the local VRF.

These import and export mechanisms are what builds the VPN. For a standard full-mesh VPN, all VRFs for a given customer will be configured to export the same RT and import routes that are labelled with that same RT.

## 2.6 MPLS Label Forwarding

Packet forwarding is achieved using a two label MPLS stack, which is what allows the P routers to remain oblivious of the VPN routes.

1. A packet from CE1 arrives at the ingress PE1.
2. PE1 performs a lookup in the VRF associated with CE1. The route points to the egress PE2 as the BGP next hop.
3. PE1 pushes a stack of two MPLS labels onto the packet:
    - **VPN Label:** This label was advertised by PE2 via MP-BGP along with the VPN-IPv4 route. This label tells PE2 which VRF to use to forward the packet to CE2.
    - **Tunnel Label:** This label is learned via an internal protocol like LDP. It tells the P routers how to reach the BGP next hop. The provider's IGP is used to establish reachability between PEs, and LDP distributes the labels for these routes.
4. The packet is forwarded into the SP backbone core. P routers only inspect the tunnel label. They swap this label at each hop, forwarding the packet toward PE2 without ever looking at the inner label or the IP header.
5. The packet arrives at PE2
6. PE2 inspects the VPN label. This label tells PE2 to perform a lookup in the VRF for CE2.
7. PE2 removes the final label and forwards the original IP packet to CE2.

## 2.7 Topologies

A single customer may have sites connected to different service providers, or a global SP may use regional SPs for access, Interprovider VPNs are used for this purpose.

- **VRF-to-VRF:** The ASBRs of the two providers are connected, where each ASBR is configured with a VRF for each shared VPN. The ASBRs peer with each other and exchange standard, unlabeled IPv4 routes as if they were CE routers to each other. This method is simple to configure but scales poorly as it scales inversely to the number of VPNs.
- **MP-EBGP between ASBRs:** The ASBRs of the two providers peer with each other using MP-BGP session and exchange labeled VPN-IPv4 routes. This is far more scalable as the ASBRs do not need to maintain VRFs for each VPN, but simply pass along the VPN routes.
- **MP-EBGP between PEs:** The PEs in different ASs peer directly with each other via MP-EBGP (or via route reflectors). The ASBRs in this case do not participate in the VPN-IPv4 labeling at all and simply forward labeled packets. This is highly scalable and is a common model for large and complex backbones.

## 2.8 Features

- **Hub-and-Spoke VPNs:** By manipulating RT import/export policies, spoke sites can only communicate with the central hub site, and not directly with other spokes.
- **Overlapping VPNs:** Allows for complex scenarios where a site may belong to more than one VPN e.g. intranet extranet communications.
- **MVPNs:** Provides a framework for transporting IP multicast traffic for customers across the SP's MPLS backbone.
- **Class of Service:** L3VPNs can support CoS, allowing the SP to honor a customer's QoS markings and provide differentiated services across the backbone.

# 3 Conclusion

The MP-BGP L3VPN architecture provides a powerful, scalable, and secure solution for service providers to offer L3VPN services. By combining separate VRFs for isolation, and VPN-IPv4 to handle overlapping addresses, MP-BGP for route distribution, and high-performance stateless switching with MPLS, it became standard for building scalable private IP networks. Different implementations can be done such as Juniper's, that provide practical import and export policies to build complex advanced network topologies for interconnected L3VPN services.

# References

[1] E. Rosen (Cisco Systems Inc.), Y. Rekhter (Juniper Networks), "BGP/MPLS IP Virtual Private Networks (VPNs)," RFC 4364, Feb. 2006. Available: https://www.rfc-editor.org/rfc/rfc4364

[2] HPE Juniper Networking, *Junos OS Documentation: Layer 3 VPNs User Guide*, 2025. Available: https://www.juniper.net/documentation/us/en/software/junos/vpn-l3/