

Multiprotocol Label Switching Layer 3 Virtual Private Networks with Open Shortest Path First protocol

PRASAD ATHUKURI

Sree kavitha engineering info technology, kammam

Abstract

This paper aims at implementing Layer 3 Multiprotocol Label Switching Layer 3 Virtual Private Networks with Open Shortest Path First protocol (OSPF) between the Provider Edge and Customer edge routers, to separate the customer routes and to create a Border Gateway Protocol (BGP) free core in the service provider network. Multiprotocol Label Switching (MPLS) uses the Border Gateway Protocol extensions to distinguish customer routes. This enables the customers to connect to the remote sites across the MPLS cloud without leaking their routes to the other customers. The core routers in the service provider network need not run BGP as the packets are label switched. MPLS provides a much more scalable and effective solution to site to site VPN connections across service provider networks as compared to the other WAN solutions like Frame Relay. L3 MPLS VPN services allow businesses to outsource their current network core using a private IP-based service offering from a service provider.

Unlike current overlay networks (such as ATM or Frame Relay service offerings), MPLS VPNs require that the enterprise peer with the SP at the IP L3 level. In this scenario, the SP network is involved in the L3 routing of the IP packets delivered by the enterprise.

This capability is implemented through Virtual Routing/Forwarding (VRF) tables for each customer and MPLS labels to de-multiplex and to tunnel customer traffic through the SP core. Because the SP network participates in the routing of customer traffic, each enterprise must inject its prefixes into the appropriate VRF table in the SP network. The SP is responsible for ensuring that these routes are distributed to the appropriate customer VRF tables.

Routing scenarios can sometimes be complex, such as in a customer hub-and-spoke topology where traffic to and from each spoke is routed through the hub. However, the most common deployment is an any-to-any topology where any customer device can connect directly to the L3 MPLS VPN. Enterprise traffic entering the SP domain is then routed based on the information in the VRF table and encapsulated

with MPLS labels to ensure proper tunneling and de-multiplexing through the core.

INTRODUCTION

Multi protocol label switching:

Multiprotocol Label Switching (MPLS) is a mechanism in high-performance telecommunications networks that directs data from one network node to the next based on short path labels rather than long network addresses, avoiding complex lookups in a routing table. The labels identify virtual links (paths) between distant nodes rather than endpoints. MPLS can encapsulate packets of various network protocols. MPLS supports a range of access technologies, including T1/E1, ATM, Frame Relay and DSL

A virtual private network (VPN): Is a network that uses primarily public telecommunication infrastructure, such as the Internet, to provide remote offices or traveling users an access to a central organizational network.

VPNs typically require remote users of the network to be authenticated, and often secure data with encryption technologies to prevent disclosure of private information to unauthorized parties.

VPNs may serve any network functionality that is found on any network, such as sharing of data and access to network resources, printers, databases, websites, etc. A VPN user typically experiences the central network in a manner that is identical to being connected directly to the central network. VPN technology via the public Internet has replaced the need to requisition and maintain expensive dedicated leased-line telecommunication circuits once typical in wide-area network installations.

Open shortest path first:

OSPF is an interior gateway protocol that routes Internet Protocol (IP) packets solely within a single routing domain (autonomous system). It gathers link state information from available routers and constructs a topology map of the network. The topology determines the routing table presented to the Internet Layer which makes routing decisions based solely on the destination IP address found in IP

packets. OSPF was designed to support variable-length subnet masking (VLSM) or Classless Inter-Domain Routing (CIDR) addressing models.

OSPF detects changes in the topology, such as link failures, very quickly and converges on a new loop-free routing structure within seconds. It computes the shortest path tree for each route using a method based on Dijkstra's algorithm, a shortest path first algorithm.

The link-state information is maintained on each router as a link-state database (LSDB) which is a tree-image of the entire network topology. Identical copies of the LSDB are periodically updated through flooding on all OSPF routers.

The OSPF routing policies to construct a route table are governed by link cost factors (external metrics) associated with each routing interface. Cost factors may be the distance of a router (round-trip time), network throughput of a link, or link availability and reliability, expressed as simple unitless numbers. This provides a dynamic process of traffic load balancing between routes of equal cost.

An OSPF network may be structured, or subdivided, into routing areas to simplify administration and optimize traffic and resource utilization. Areas are identified by 32-bit numbers, expressed either simply in decimal, or often in octet-based dot-decimal notation, familiar from IPv4 address notation.

Backdoor connectivity: Either a dynamic or permanent link, outside of the MPLS VPN cloud, over which a routing adjacency is formed to pass routing information that ties two customer domains together. This link is typically used to connect two geographically distinct sites and usually runs the same IGP protocol as the customer site. An example of a backdoor link is illustrated in Figure 2.

Customer router: Customer router that is connected only to other customer devices.

Customer edge: Customer edge router that peers at Layer 3 to the provider edge. The PE-CE interface runs either a dynamic routing protocol (eBGP, RIPv2, EIGRP, or OSPF) or a static routing protocol (Static, Connected).

Global routing/forwarding table: The non-VRF routing and forwarding table used in the SP core for infrastructure addressing reachability.

Label: In this document, this refers to an MPLS frame-based label.

MP-BGP: Multi-Protocol Border Gateway Protocol. In an MPLS VPN context, this protocol is run between PE routers to exchange customer prefixes in a VPNv4 format.

Managed CE service: Some service providers may offer an added service along with the Layer 3 MPLS VPN offering known as a managed CE service. The SP handles the operations, management, and administration of the CE router at one or more sites. There are typically added charges for what is essentially outsourced management of the CE devices.

Provider router: Provider router, which resides in the core of the provider network. In an MPLS VPN context, the P router participates in the control plane for customer prefixes. The P router is sometimes referred to as a label switch router (LSR), in reference to its primary role in the core of the network, performing label switching/swapping of MPLS traffic.

Provider edge router: Provider edge router. The PE router sits at the edge of the MPLS SP cloud. In an MPLS VPN context, separate VRF routing tables are allocated for each user group. Also, the PE still contains a global routing table for routes in the core SP infrastructure. The PE is sometimes referred to as a label edge router (LER) or edge label switch router (ELSR) in reference to its role at the edge of the MPLS cloud, performing label imposition and disposition.

Route distinguisher: Route distinguisher, which is a 64-bit value defined uniquely for each user group. The RD is combined with the customer IPv4 prefix to guarantee that the resulting VPNv4 prefix is unique.

Route target: Route target, which is a 64-bit value used as a BGP extended community attribute. The RT is used to determine the VPNv4 routes that should be installed in the respective VRF tables.

VPNv4: The combination of the RD and customer IPv4 prefix. These VPNv4 prefixes are passed in MP-BGP.

VRF: The virtual routing and forwarding table, which is separate from the global routing table that exists on PE routers. Routes are injected into the VRF from the CE-PE routing protocols for that VRF and any MP-BGP announcements that match the defined VRF route targets (RTs).

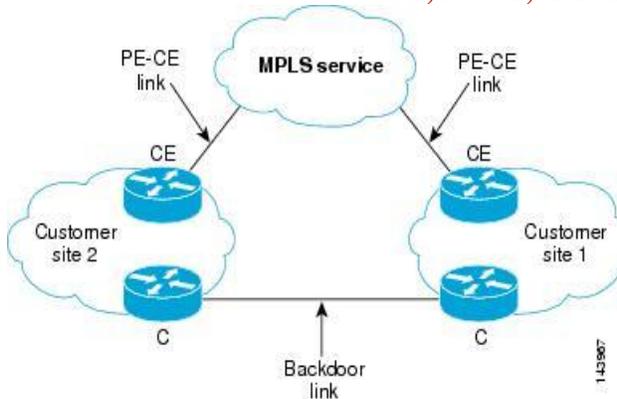


Figure 2 Backdoor Link

Layer 3 MPLS VPN Operation

This section briefly examines the L3 MPLS VPN control and data planes, and includes the following topics:

Layer 3 MPLS VPN Route Distribution Operation

Layer 3 MPLS VPN Forwarding Operations

Layer 3 MPLS VPN Route Distribution Operation

Figure 3 illustrates an example of BGP VPN route distribution using MP-BGP between a VPN that terminates on PE3 and PE7. The customer devices (C1 and CE2 on the left, and CE8 and C9 on the right) participate in the same VPN.

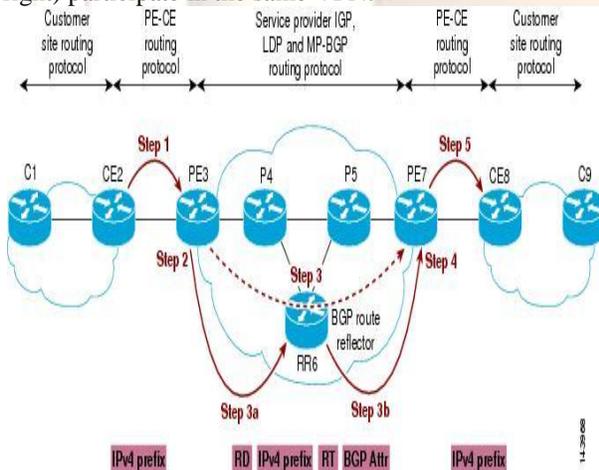


Figure 3 BGP VPN Route Distribution

The distribution steps are as follows:

1. Customer routes are injected into the VRF table at PE3 using static, RIPv2, OSPF, or BGP routing protocol between the PE and the CE. The customer routes are passed as IPv4 prefixes (shown in the red shaded box under Step 1).
2. At PE3, the routes in the customer VRF are exported into MP-BGP as VPNv4 prefixes. To ensure VPNv4 route uniqueness, the customer IPv4 routes are prepended with a uniquely defined RD to create a distinct VPNv4 prefix. Every VRF configuration

requires an RD to be defined. Its uniqueness guarantees customer VPNv4 uniqueness.

3. The exported routes are sent across the MPLS backbone between the BGP peers in PE3 and PE7.

This process repeats for any other BGP peers that have members in the same VPN. Note that this step shows a logical connection between the two BGP peers. There can be a series of BGP route reflectors in between performing the VPN distribution as shown in Steps 3a and 3b.

The VPNv4 prefix (shown in red shaded boxes under Step 3) is composed of the RD and the customer IPv4 prefix. Because this VPNv4 prefix is a BGP route, multiple mandatory and optional BGP attributes are carried along with the prefix. One of these attributes is the route target (RT), which is an extended community BGP attribute.

4. The routes are imported into the correct VRF at PE7. Every VRF configuration contains VRF import and export definitions. The export definitions define which RTs are attached to the BGP VPNv4 prefix, as described in Step 3. The export definitions define the RTs that are carried along with the VPNv4 prefix on export. The import definitions define the RT tagged prefixes that are imported into the VRF. Only VPNv4 prefixes with a matching RT tag to the VRF import RT definitions are imported into that VRF.
5. The routes are accessible from a VPN at each site.

Layer 3 MPLS VPN Forwarding Operations

Figure 4 illustrates the process of packet forwarding for a packet originating from the customer cloud containing C1 and CE2 to the far-end customer cloud containing CE8 and C9.

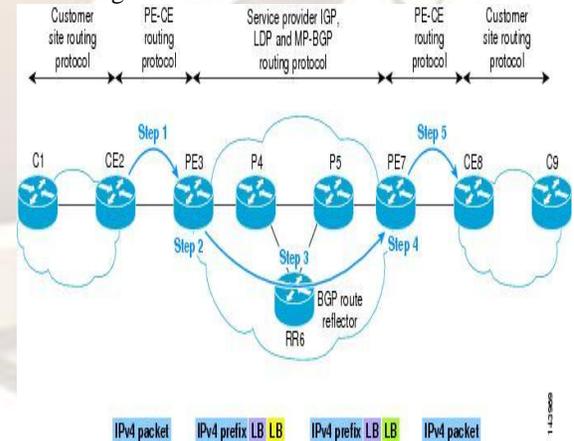


Figure 4 MPLS Data Forwarding Example

1. The customer cloud composed of C1 and CE2 originates an IPv4 packet destined to an address at the far end (CE8 and C9). The routing entry on CE2 for the destination prefix forwards the packet to the PE3 device.

- PE3 receives the customer packet and does a routing lookup according to the VRF table that is bound to that interface. In this case, the route resolves to a BGP prefix originated from PE7. PE3 imposes two labels on the IPv4 packet. The first label, referred to in this document as the VPN label, (shown in the purple "LB" shaded box) is the label that is used to uniquely identify a customer VPN prefix. The second label, referred to in this document as the forwarding label (shown by the yellow "LB" shaded box) is the label used to tunnel the packet through the P core to the far-end PE7 device.
- The labeled packet is now forwarded at each hop through the SP core. Each P router makes a forwarding decision based on the top level label, and this top level label is swapped with a new label. This is shown by the yellow "LB" shaded box, and the outgoing packet is shown with a green "LB" shaded box. The underlying packet and inner label are left undisturbed during this process.
- Eventually, PE7 recognizes the labeled packet and recognizes the inner VPN label (purple "LB") as a VPN label for that specific customer prefix. The VPN label is stripped and a forwarding decision for the IPv4 packet is made based on the VPN label. P5 may remove the top level label, leaving only the inner label when forwarding to PE7. This concept is known as penultimate hop popping (PHP), where the penultimate hop removes the top level label. The relevance to the enterprise is that in a PHP scenario, the SP-marked EXP value may not be copied down to the inner label. This depends on the MPLS QoS mode chosen. This is relevant only if the traffic from the PE to the CE (for example, PE7 to CE8 in figure 4) must be queued based on the SP EXP marking
- The original IPv4 packet is forwarded by the switch to the appropriate customer VRF interface. The MPLS label is a 32-bit shim that is inserted between the L2 data link header and the underlying payload (in this case an IPv4 packet). Figure 5 illustrates the format of the 32-bit label.

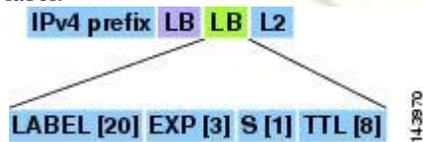


Figure 5 MPLS Label Detail

Table 3 MPLS Label Field Descriptions

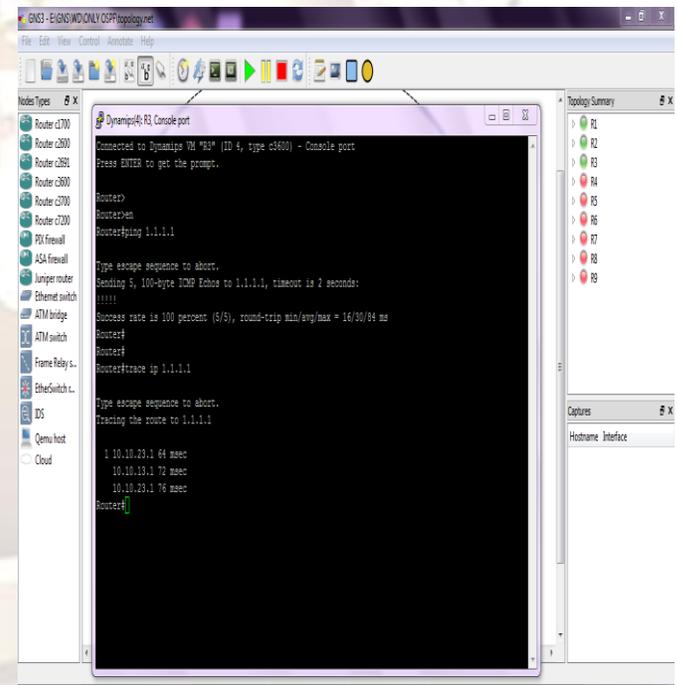
Field ID	Length	Purpose
LABEL	20 bits	Allocated for the actual label value.
	3 bits	MPLS experimental bits. A Cisco

EXP	convention is to use these experimental bits as a means of representing the class of service (CoS) of the MPLS frame.
-----	---

Table 3 describes each field in this label: MPLS VPNs, unlike other VPN types such as IPsec, perform no encryption. Despite this, however, a Layer 3 MPLS VPN service offers equivalent security to that of an ATM/Frame Relay service offering through the use of distinct routing tables and label spoofing mechanisms.

RESULTS

GNS 3 is used to emulate our proposed protocol



For example if two same loop back interfaces are created on two customer networks and are accessed from the customer network on the other area through service provider if we implement OSPF between them the provider edge router can't distinguish between the two it chooses the path with less cost if both the path have equal costs it distributes the packet between them as show above. to loop back interference of customer1 via 10.10.13.1 with MPLS label 16/22 .similarly figure2 shows the

trace of route to loop back interference of customer2 via 10.10.23.1 with MPLS label 16/20. In this way the provider edge router by assigning different labels are able to distinguish between the two routes with same IP address.

one customer are not leaked to other customer as shown above the figure one shows the trace of route

REFERENCES:

1. wikipedia.org/wiki/Multiprotocol_LabelSwitching
2. MPLS Virtual Private Networks
http://www.cisco.com/univercd/cc/td/doc/product/software/ios122/122cgcr/fiprrp_r/1rfospf.htm
3. Configuring OSPF
<http://www.cisco.com/univercd/cc/td/doc/product/software/ios120/120newft/120t/120t5/vpn.htm>
4. Configuring BGP
http://www.cisco.com/univercd/cc/td/doc/product/software/ios122/122cgcr/fipr_c/ipcprt2/1cfospf.htm
5. RFC 1163, A Border Gateway Protocol
http://www.cisco.com/univercd/cc/td/doc/product/software/ios122/122cgcr/fipr_c/ipcprt2/1cfbgp.htm
6. information on the OSPF configuration procedures,
http://www.cisco.com/univercd/cc/td/doc/product/software/ios122/122cgcr/fiprrp_r/1rfospf.htm

```

Dynamips(20): R9, Console port
changed state to down
*Mar 1 00:01:52.927: %LINEPROTO-5-UPDOWN: Line protocol on Interface Serial1/0,
*Mar 1 00:02:31.667: %OSPF-5-ADJCHG: Process 1, Nbr 10.10.79.1 on Serial1/0 fro
Router>
Router>ping 1.1.1.1

Type escape sequence to abort.
Sending 5, 100-byte ICMP Echos to 1.1.1.1, timeout is 2 seconds:
!!!!
Success rate is 100 percent (5/5), round-trip min/avg/max = 116/152/240 ms
Router>
Router>
Router>trace ip 1.1.1.1

Type escape sequence to abort.
Tracing the route to 1.1.1.1

  1 10.10.79.1 40 msec 36 msec 32 msec
  2 10.10.67.1 [MPLS: Labels 16/22 Exp 0] 164 msec 140 msec 136 msec
  3 10.10.56.1 [MPLS: Labels 16/22 Exp 0] 132 msec 172 msec 116 msec
  4 10.10.45.1 [MPLS: Labels 16/22 Exp 0] 92 msec 136 msec 104 msec
  5 10.10.13.2 [MPLS: Label 22 Exp 0] 68 msec 116 msec 84 msec
  6 10.10.13.1 136 msec 156 msec *
Router>
    
```

Figure1

```

Dynamips(28): R8, Console port
*Mar 1 00:01:02.195: %CRYPTO-6-ISAKMP ON OFF: ISAKMP is OFF
*Mar 1 00:01:02.199: %CRYPTO-6-GDOI ON OFF: GDOI is OFF
*Mar 1 00:01:02.199: %CRYPTO-6-ISAKMP ON OFF: ISAKMP is OFF
*Mar 1 00:01:04.771: %OSPF-5-ADJCHG: Process 1, Nbr 10.10.78.1 on Serial1/0 fro
Router>
Router>ping 1.1.1.1

Type escape sequence to abort.
Sending 5, 100-byte ICMP Echos to 1.1.1.1, timeout is 2 seconds:
!!!!
Success rate is 100 percent (5/5), round-trip min/avg/max = 120/143/180 ms
Router>
Router>
Router>trace ip 1.1.1.1

Type escape sequence to abort.
Tracing the route to 1.1.1.1

  1 10.10.78.1 60 msec 24 msec 52 msec
  2 10.10.67.1 [MPLS: Labels 16/20 Exp 0] 176 msec 124 msec 148 msec
  3 10.10.56.1 [MPLS: Labels 16/20 Exp 0] 156 msec 324 msec 196 msec
  4 10.10.45.1 [MPLS: Labels 16/20 Exp 0] 148 msec 240 msec 148 msec
  5 10.10.23.2 [MPLS: Label 20 Exp 0] 96 msec 228 msec 112 msec
  6 10.10.23.1 120 msec 160 msec *
Router>
    
```

Figure2

If we implement MPLS VPN stated earlier between them the provider edge router is able to distinguish between the two loop back interference. The routes of