

Carrier Supporting Carrier

This document details the concept known as Carrier Supporting Carrier. The design will be explained first followed by an example showing both the control plane advertisements and the data plane traffic flow.

CsC Design

CsC works by having one Service Provider (the Backbone Provider), treat another Provider (the Customer Provider) as a customer. However this isn't a simple Option A setup. In order to maintain a continuous LSP from one part the Customer Providers network to another, prefixes and corresponding labels must be exchanged between the Customer and Backbone Providers. These prefixes correspond to the Customer Providers LSP endpoints which are typically router loopbacks.

There are two ways to do this:

1. Run BGP Labelled Unicast between the Customer and Backbone Providers edge routers.
2. Run IGP and LDP between the Customer and Backbone Providers edge routers.

Both options will be explored in this document. Once these prefixes + labels have been exchanged, they need to be distributed throughout the Customer Providers networks. This can be done either using BGP LU or redistribution into the local IGP, from which LDP will then dynamically allocate labels. Only the later option is shown in the CLI output here, but the downloadable GNS3 lab that accompanies this document uses both.

The Backbone Provider must put the attachment circuits connecting to the Customer Provider into a VRF. This results in VRF aware LDP being run if the IGP + LDP option is used between the Customer and Backbone Providers.

The Customer Provider does not need to put its attachment circuits to the Backbone Provider into a VRF.

The Backbone Provider can then run any internal MPLS setup it wishes. Here we will use standard MPLS VPNv4 between the Backbone Provider PEs. The Customer Provider runs the same type of setup with respect to its customers.

CsC Example

To illustrate these concepts, the sample network shown on the right is used.

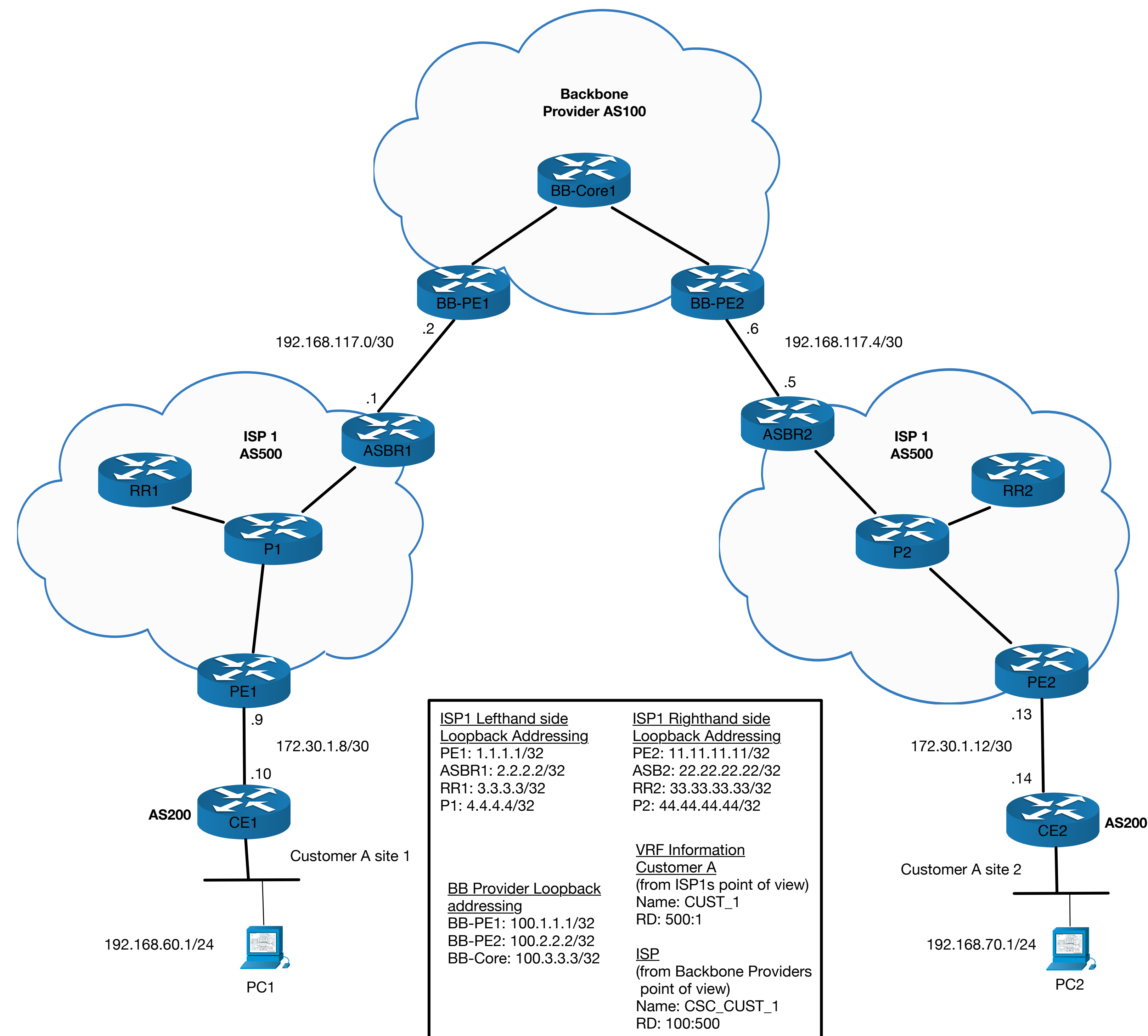
Customer A has two sites connected to ISP1 in a VRF called CUST_1.

ISP1 has two separate disparate areas. The Backbone Provider implements CsC to connect to the two sides together. ISP1 has one route reflector in each area running VPNv4 to the PE routers, and between each other. IS-IS and LDP is used throughout ISP1.

The connection between ASBR1 and BB-PE1 is running OSPF and LDP. The loopbacks of RR1 and PE1 are redistributed from IS-IS into OSPF and sent to BB-PE1. The loopback for RR1 is needed to provide reachability for the iBGP VPNv4 session between the reflectors. The loopback for PE1 is needed for next-hop reachability when using this PE as an LSP endpoint. ASBR1 also redistributes from OSPF into IS-IS. This is expected to be loopbacks for RR2 and PE2. LDP will then allocate labels for these.

The connection between ASBR2 and BB-PE2 is running BGP-LU. The loopbacks for RR2 and PE2 are advertised to BB-PE2 for the same reason ASBR1 redistributes from IS-IS to OSPF. ASBR2 should be receiving the loopback for RR1 and PE1. Instead of redistributing into the local IGP, ASBR2 runs BGP-LU to RR2 who in turn runs BGP-LU to PE2 into order to communicate the label for PE1.

The examples given in this document will follow the control plane advertisements and traffic flow involved when PC1 sends a packet to PC2. All output shown and referenced is based on Cisco IOS. The local PC is represented by loopbacks on the CE routers.





Control Plane

Concept - From CE1 to CE2

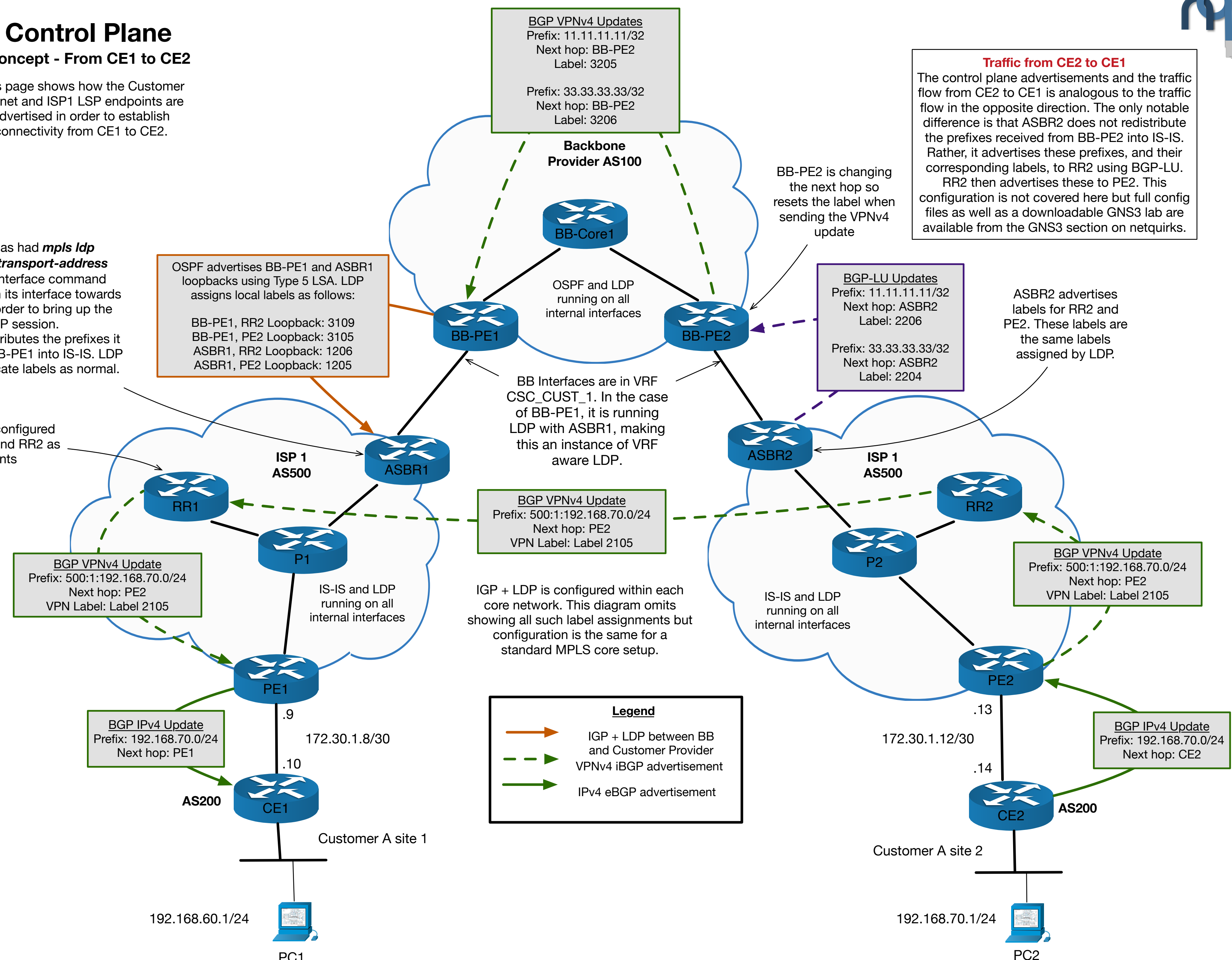
This page shows how the Customer subnet and ISP1 LSP endpoints are advertised in order to establish connectivity from CE1 to CE2.

ASBR1 has had *mpls ldp discovery transport-address interface* command configured on its interface towards BB-PE1 in order to bring up the LDP session. ASBR1 redistributes the prefixes it learns from BB-PE1 into IS-IS. LDP will then allocate labels as normal.

OSPF advertises BB-PE1 and ASBR1 loopbacks using Type 5 LSA. LDP assigns local labels as follows:
BB-PE1, RR2 Loopback: 3109
BB-PE1, PE2 Loopback: 3105
ASBR1, RR2 Loopback: 1206
ASBR1, PE2 Loopback: 1205

RR1 has configured both PE1 and RR2 as clients

Traffic from CE2 to CE1
The control plane advertisements and the traffic flow from CE2 to CE1 is analogous to the traffic flow in the opposite direction. The only notable difference is that ASBR2 does not redistribute the prefixes received from BB-PE2 into IS-IS. Rather, it advertises these prefixes, and their corresponding labels, to RR2 using BGP-LU. RR2 then advertises these to PE2. This configuration is not covered here but full config files as well as a downloadable GNS3 lab are available from the GNS3 section on netquirks.



BGP VPNv4 Update
Prefix: 500:1:192.168.70.0/24
Next hop: PE2
VPN Label: Label 2105

IGP + LDP is configured within each core network. This diagram omits showing all such label assignments but configuration is the same for a standard MPLS core setup.

Legend

- Orange arrow: IGP + LDP between BB and Customer Provider
- Green dashed arrow: VPNv4 iBGP advertisement
- Green solid arrow: IPv4 eBGP advertisement

BGP-LU Updates
Prefix: 11.11.11.11/32
Next hop: ASBR2
Label: 2206

Prefix: 33.33.33.33/32
Next hop: ASBR2
Label: 2204

BB-PE2 is changing the next hop so resets the label when sending the VPNv4 update

BGP VPNv4 Update
Prefix: 500:1:192.168.70.0/24
Next hop: PE2
VPN Label: Label 2105

BGP IPv4 Update
Prefix: 192.168.70.0/24
Next hop: CE2

BGP VPNv4 Update
Prefix: 500:1:192.168.70.0/24
Next hop: PE2
VPN Label: Label 2105

BGP IPv4 Update
Prefix: 192.168.70.0/24
Next hop: PE1

BB Interfaces are in VRF CSC_CUST_1. In the case of BB-PE1, it is running LDP with ASBR1, making this an instance of VRF aware LDP.

IS-IS and LDP running on all internal interfaces

IS-IS and LDP running on all internal interfaces

192.168.60.1/24

192.168.70.1/24



Control Plane

CLI output - From CE1 to CE2

This page shows relevant CLI output for the CsC setup - demonstrating how label switching is done and what you'd expect to see.

```

BB-PE1#show mpls forwarding-table vrf CSC_CUST_ISP1 11.11.11.11 detail
Local  Outgoing  Prefix          Bytes tag  Outgoing  Next Hop
tag   tag or VC  or Tunnel Id   switched  interface
3105  3205      11.11.11.11/32[V] 18660     Fa0/0     10.30.13.3
      MAC/Encaps=14/22, MRU=1496, Tag Stack(3301 3205)
      C20627A40000C20427A200008847 00CE500000C85000
      VPN route: CSC_CUST_ISP1
      No output feature configured
      Per-packet load-sharing
BB-PE1#sh run | sec router ospf
router ospf 2 vrf CSC_CUST_ISP1
log-adjacency-changes
redistribute bgp 100 subnets
router ospf 1
mpls ldp autoconfig
router-id 100.1.1.1
log-adjacency-changes
passive-interface Loopback0
BB-PE1#sh run interface Fa0/1
Building configuration...

Current configuration : 218 bytes
!
interface FastEthernet0/1
description link to ASBR1
ip vrf forwarding CSC_CUST_ISP1
ip address 192.168.117.2 255.255.255.252
ip ospf network point-to-point
ip ospf 2 area 0
mpls ip
end
BB-PE1#

```

```

BB-PE2#sh bgp vpnv4 unicast vrf CSC_CUST_ISP1 11.11.11.11/32
BGP routing table entry for 100:500:11.11.11.11/32, version 24
Paths: (1 available, best #1, table CSC_CUST_ISP1)
Advertised to update-groups:
 2
 500
 192.168.117.5 from 192.168.117.5 (22.22.22.22)
  Origin IGP, metric 20, localpref 100, valid, external, best
  Extended Community: RT:100:500
  mpls labels in/out 3205/2206
BB-PE2#sh mpls forwarding-table vrf CSC_CUST_ISP1 11.11.11.11 detail
Local  Outgoing  Prefix          Bytes tag  Outgoing  Next Hop
tag   tag or VC  or Tunnel Id   switched  interface
3205  2206      11.11.11.11/32[V] 21564     Fa0/1     192.168.117.5
      MAC/Encaps=14/18, MRU=1500, Tag Stack(2206)
      C20E279E0001C20527A300018847 0089E000
      VPN route: CSC_CUST_ISP1
      No output feature configured
      Per-packet load-sharing
BB-PE2#

```

```

ASBR2#show bgp ipv4 unicast 11.11.11.11
BGP routing table entry for 11.11.11.11/32, version 22
Paths: (1 available, best #1, table Default-IP-Routing-Table)
Advertised to update-groups:
 1
 2
  Local
  10.20.24.4 from 0.0.0.0 (22.22.22.22)
  Origin IGP, metric 20, localpref 100, weight 32768, valid,
  sourced, local, best
  mpls labels in/out 2206(from LDP)/nolabel
ASBR2#show bgp ipv4 unicast 33.33.33.33
BGP routing table entry for 33.33.33.33/32, version 29
Paths: (1 available, best #1, table Default-IP-Routing-Table)
Advertised to update-groups:
 1
 2
  Local
  10.20.24.4 from 0.0.0.0 (22.22.22.22)
  Origin IGP, metric 20, localpref 100, weight 32768, valid,
  sourced, local, best
  mpls labels in/out 2204(from LDP)/nolabel
ASBR2#

```

```

ASBR1#sh mpls forwarding-table
Local  Outgoing  Prefix          Bytes tag  Outgoing  Next Hop
tag   tag or VC  or Tunnel Id   switched  interface
1200  Pop tag   4.4.4.4/32     0         Fa0/0     10.10.24.4
1201  Pop tag   10.10.14.0/24  0         Fa0/0     10.10.24.4
1202  Pop tag   10.10.34.0/24  0         Fa0/0     10.10.24.4
1203  1401     1.1.1.1/32     26328     Fa0/0     10.10.24.4
1204  1400     3.3.3.3/32     17701     Fa0/0     10.10.24.4
1205  3105     11.11.11.11/32 16332     Fa0/1     192.168.117.2
1206  3109     33.33.33.33/32 7877      Fa0/1     192.168.117.2
ASBR1#sh run | sec router ospf
router ospf 1
router-id 2.2.2.2
log-adjacency-changes detail
redistribute isis LAB level-2 subnets route-map LOOPBACKS
match ip address prefix-list LOOPBACKS
ASBR1#sh route-map LOOPBACKS
route-map LOOPBACKS, permit, sequence 10
Match clauses:
  ip address prefix-lists: LOOPBACKS
Set clauses:
  Policy routing matches: 0 packets, 0 bytes
ASBR1#sh ip pref
ASBR1#sh ip prefix-list LOOPBACKS
ip prefix-list LOOPBACKS: 2 entries
 seq 5 permit 3.3.3.3/32
 seq 10 permit 1.1.1.1/32
ASBR1#

```

```

P1#sh mpls forwarding-table
Local  Outgoing  Prefix          Bytes tag  Outgoing  Next Hop
tag   tag or VC  or Tunnel Id   switched  interface
1400  Pop tag   3.3.3.3/32     39241     Fa1/0     10.10.34.3
1401  Pop tag   1.1.1.1/32     44880     Fa0/0     10.10.14.1
1402  Pop tag   2.2.2.2/32     0         Fa0/1     10.10.24.2
1403  Pop tag   192.168.117.0/30 0         Fa0/1     10.10.24.2
1404  1205     11.11.11.11/32 14604     Fa0/1     10.10.24.2
1405  1206     33.33.33.33/32 10487     Fa0/1     10.10.24.2
P1#

```

```

RR1#sh bgp vpnv4 unicast rd 500:1 192.168.70.0
BGP routing table entry for 500:1:192.168.70.0/24, version 33
Paths: (1 available, best #1, no table)
Advertised to update-groups:
 1
 200, (Received from a RR-client)
 11.11.11.11 (metric 20) from 33.33.33.33 (33.33.33.33)
  Origin incomplete, metric 0, localpref 100, valid, internal,
  best
  Extended Community: RT:500:1
  Originator: 11.11.11.11, Cluster list: 33.33.33.33
  mpls labels in/out nolabel/2105
RR1#sh ip route 33.33.33.33
Routing entry for 33.33.33.33/32
  Known via "isis", distance 115, metric 20, type level-2
  Redistributing via isis
  Last update from 10.10.34.4 on FastEthernet0/0, 01:56:36 ago
  Routing Descriptor Blocks:
  * 10.10.34.4, from 2.2.2.2, via FastEthernet0/0
    Route metric is 20, traffic share count is 1

```

```

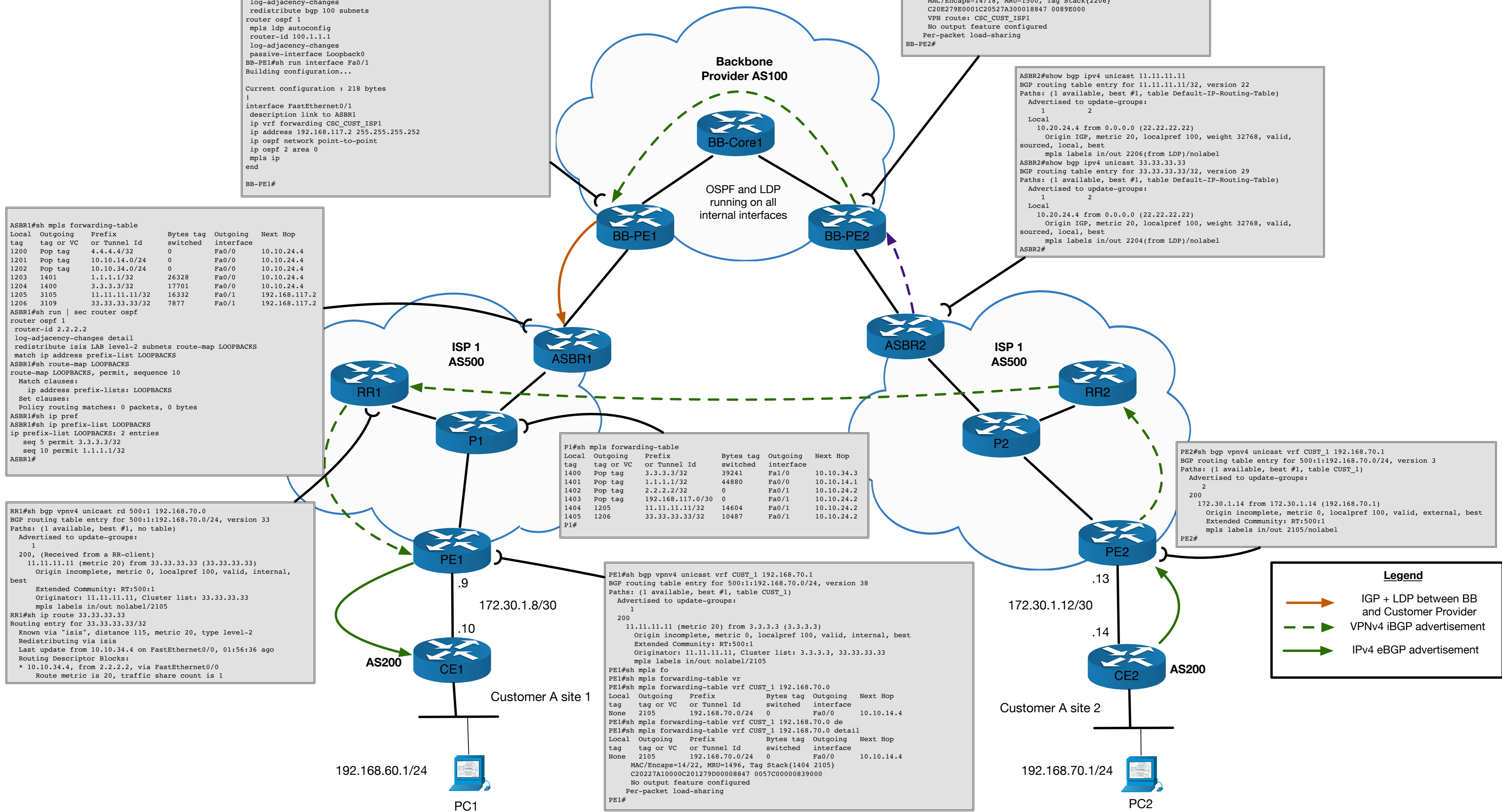
PE1#sh bgp vpnv4 unicast vrf CUST_1 192.168.70.1
BGP routing table entry for 500:1:192.168.70.0/24, version 38
Paths: (1 available, best #1, table CUST_1)
Advertised to update-groups:
 1
 200
 11.11.11.11 (metric 20) from 3.3.3.3 (3.3.3.3)
  Origin incomplete, metric 0, localpref 100, valid, internal, best
  Extended Community: RT:500:1
  Originator: 11.11.11.11, Cluster list: 3.3.3.3, 33.33.33.33
  mpls labels in/out nolabel/2105
PE1#sh mpls fo
PE1#sh mpls forwarding-table vr
PE1#sh mpls forwarding-table vrf CUST_1 192.168.70.0
Local  Outgoing  Prefix          Bytes tag  Outgoing  Next Hop
tag   tag or VC  or Tunnel Id   switched  interface
None  2105      192.168.70.0/24 0         Fa0/0     10.10.14.4
PE1#sh mpls forwarding-table vrf CUST_1 192.168.70.0 de
PE1#sh mpls forwarding-table vrf CUST_1 192.168.70.0 detail
Local  Outgoing  Prefix          Bytes tag  Outgoing  Next Hop
tag   tag or VC  or Tunnel Id   switched  interface
None  2105      192.168.70.0/24 0         Fa0/0     10.10.14.4
      MAC/Encaps=14/22, MRU=1496, Tag Stack(1404 2105)
      C20227A10000C201279D00008847 0057C00000839000
      No output feature configured
      Per-packet load-sharing
PE1#

```

```

PE2#sh bgp vpnv4 unicast vrf CUST_1 192.168.70.1
BGP routing table entry for 500:1:192.168.70.0/24, version 3
Paths: (1 available, best #1, table CUST_1)
Advertised to update-groups:
 2
 200
 172.30.1.14 from 172.30.1.14 (192.168.70.1)
  Origin incomplete, metric 0, localpref 100, valid, external, best
  Extended Community: RT:500:1
  mpls labels in/out 2105/nolabel
PE2#

```



Legend

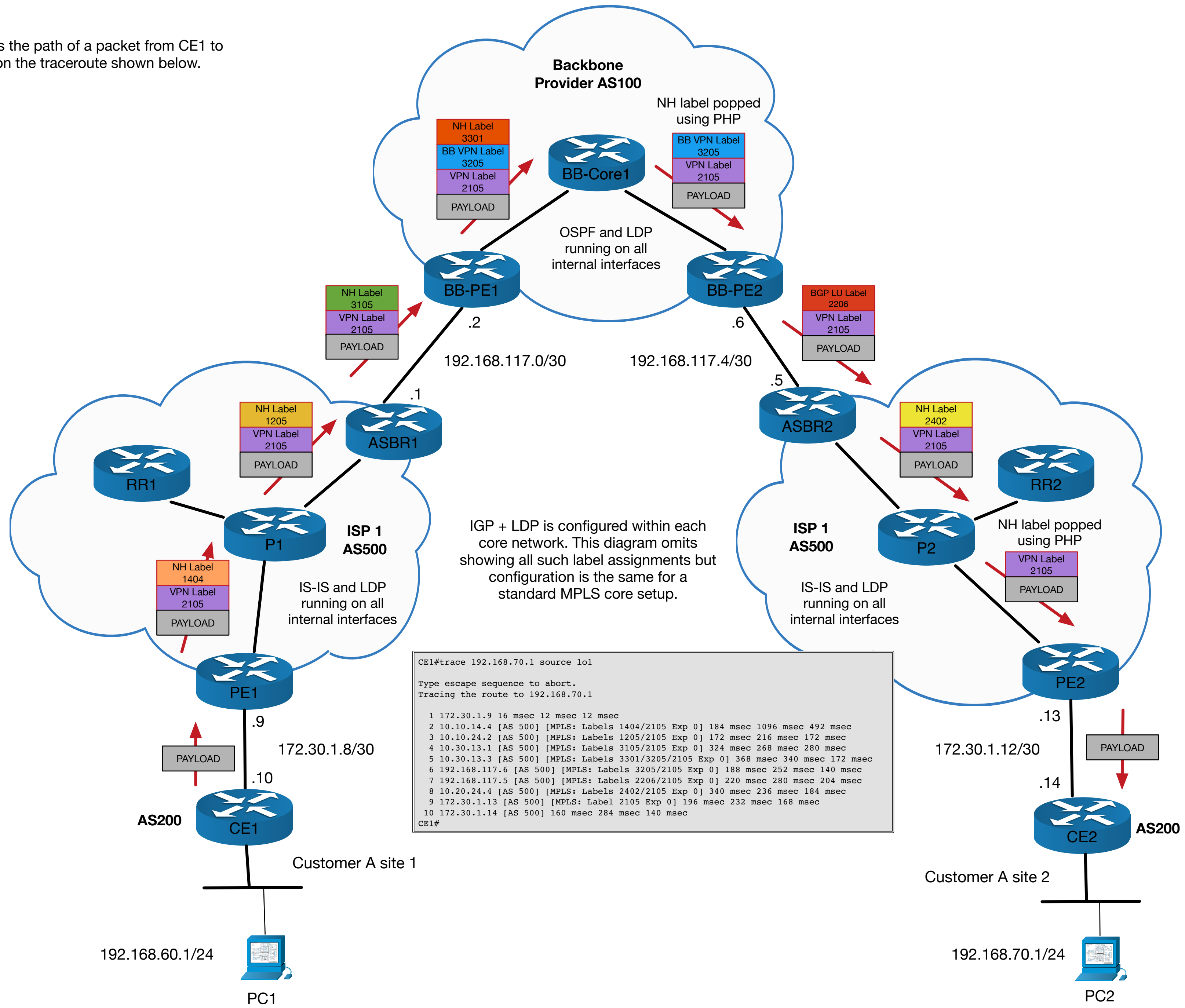
- Orange arrow: IGP + LDP between BB and Customer Provider
- Green dashed arrow: VPNv4 iBGP advertisement
- Green solid arrow: IPv4 eBGP advertisement



Data Plane

Life of a packet from CE1 to CE2

This page traces the path of a packet from CE1 to CE2 based on the traceroute shown below.



Traceroute Step	Forwarding Process
1	CE1 has a route for 192.168.70.0/24 via BGP with a next-hop of 192.30.1.9. It forwards the packet to PE1.
2	PE1 has a BGP VPNv4 route for 192.168.70.0/24 with a VPN label of 2105 and next-hop of 11.11.11.11/32 (PE2). PE1 recursively looks for a route to PE2. It finds an IS-IS route with a next-hop of P1, out of a directly connected interface running LDP. P1's local label for PE2 is 1404. So, PE1 imposes a VPN label of 2105 and then a transport (or next-hop) label of 1404.
3	P1 receives the labelled packet, sees the top label (1404) and matches its local label for PE2. P1's next hop for PE2 is ASBR1. P1 swaps 1404 label for 1205 - ASBR1's local label for PE2
4	ASBR1 receives the labelled packet and sees the top label matches its local label for PE2. This label was assigned by the LDP session running between ASBR1 and BB-PE1. ASBR2 swaps 1205 with 3105 - BB-PE1's local label for PE2.
5	BB-PE1 receives the labelled packet and sees the top label matches its local label for the 11.11.11.11/32 route in its VRF CSC_CUST_ISP1. BB-PE1 has a BGP VPNv4 route for 11.11.11.11/32 with a VPN label of 3205 and next-hop of 100.1.1.1 (BB-PE2). BB-PE1 recursively looks for a route to BB-PE2. It finds an OSPF route with a next-hop of BB-PE3, out of a directly connected interface running LDP. BB-PE3's local label for BB-PE2 is 3301. BB-PE1 swaps the top label 3105 with the VPN label 3205. It then imposes the transport label of 3301. At this point there are 3 labels on the stack (from top to bottom): 3301/3205/2105
6	BB-PE3 receives the labelled packet and sees the top label matches its local label for BB-PE2. BB-PE3's next hop for BB-PE2 is via a directly connected interface running LDP. BB-PE3 is the penultimate hop to BB-PE2 so it simply pops the top transport label and forwards the packet to BB-PE2 (normal PHP behaviour).
7	BB-PE2 receives the labelled packet and sees the top label matches its VPN label for 11.11.11.11/32 (PE2) in VRF CSC_CUST_ISP1. BB-PE2's VRF route to PE2 is known via BGP-LU with a next-hop of 192.168.117.5. BB-PE2 has also received a label value of 2206 for this prefix over this BGP-LU session so its LFIB has a swap entry. BB-PE2 swaps the top label 3205 with the BGP-LU learned label 2206 and forwards it to ASBR2.
8	ASBR2 receives the labelled packet and sees the top label matches its local LDP assigned label for 11.11.11.11/32 (PE2). ASBR2's next hop for PE2 is P2. ASBR2 swaps 2206 label for 2402 - P2's local label for PE2.
9	P2 receives the labelled packet and sees the top label matches its local label for PE2. P2's next hop for PE2 is via a directly connected interface running LDP. P2 is the penultimate hop to PE2 so it simply pops the top transport label and forwards the packet to PE2 (normal PHP behaviour).
10	PE2 receives the labelled packet and sees the top (and only) label matches its VPN label for 192.168.70.0/24 (CE2's loopback 1) in VRF CUST_1. PE2 VRF next-hop for 192.168.70.0/24 is known via IPv4 BGP out of a local attachment circuit in VRF CUST_1. PE2 removes the 2105 label and forwards the ICMP packet unlabelled.