Cisco live!

What You Make Possible

TOMORROW starts here.

CISCO

The CCIE Candidate's Introduction to MPLS L3VPN Networks

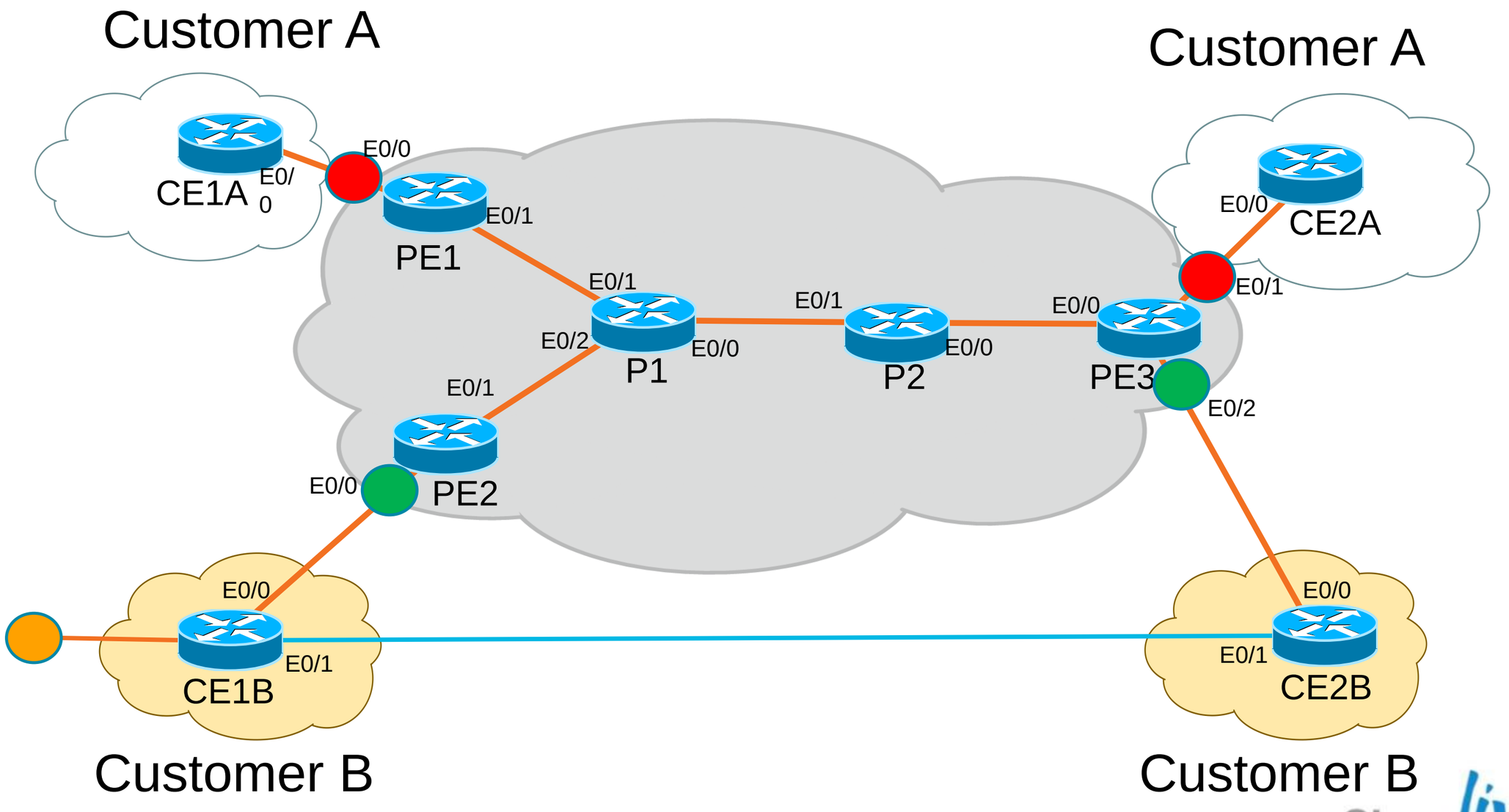
BRKCCIE-3345

Objectives

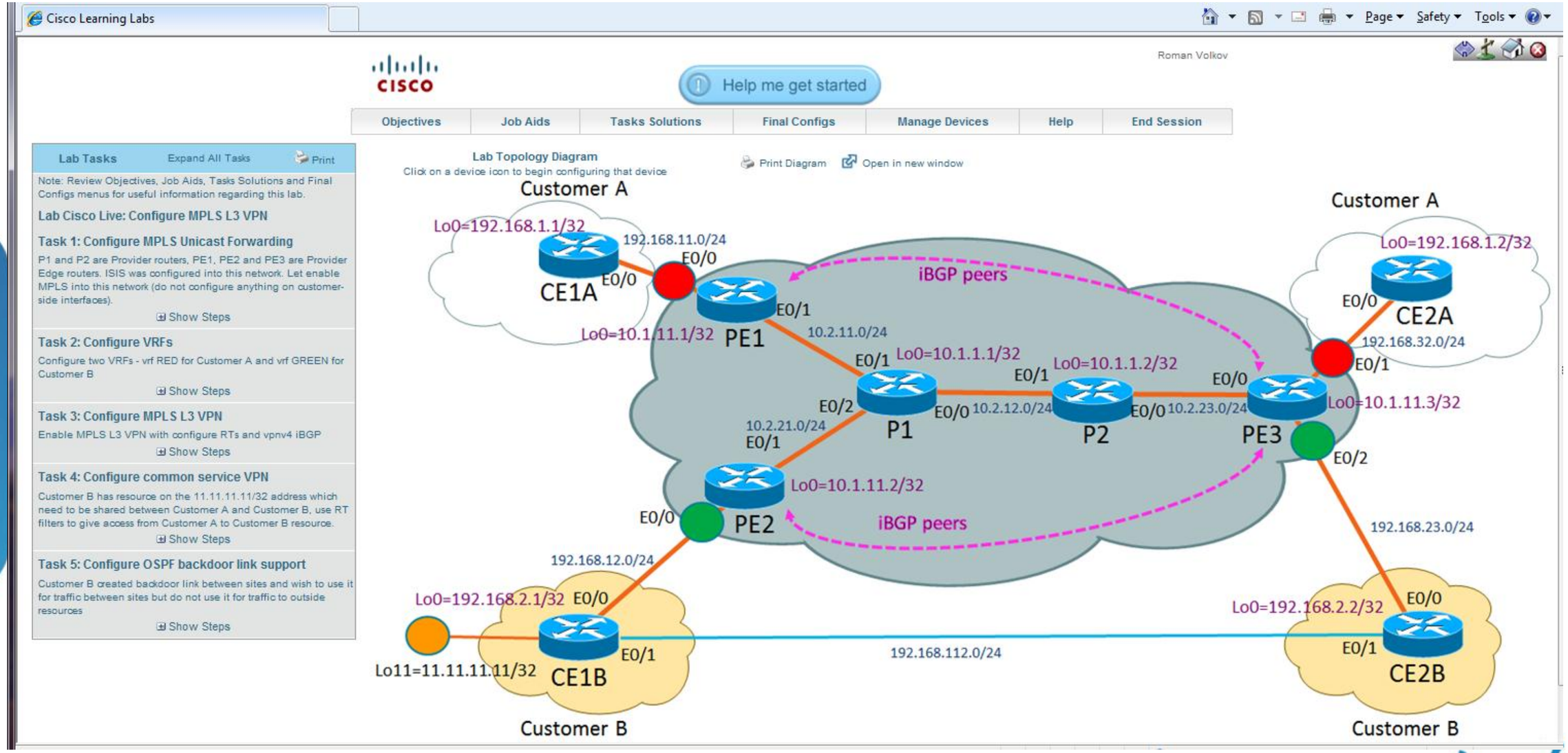
- Define network and tasks
- MPLS IP Unicast Forwarding
- What are VRFs
- MPLS L3 VPNs
- Tasks into CCIE Labs
- MPLS L3 VPN Game

Define Network and Tasks

Topology

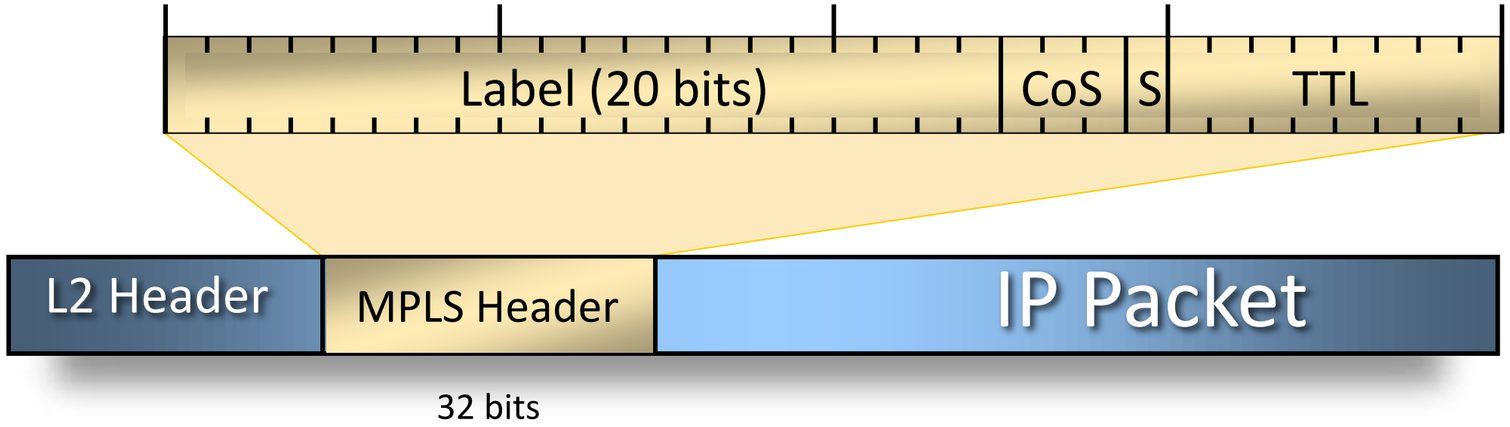


Cisco Learning Lab for this session



MPLS IP Unicast Forwarding

MPLS Header



- MPLS (Layer 2.5) Header Fields:
  - Label, 20 bits
  - Experimental (CoS), 3 bits
  - Stacking bit, 1 bit. This is the bottom-of-stack bit. 1=on=last label.
  - Time to live, 8 bits

Control Plane and Data Plane

- Dynamic protocols build the control plane.
- Packets are forwarded on the data plane.

Who do we Turn to for Lookups?

Control Plane

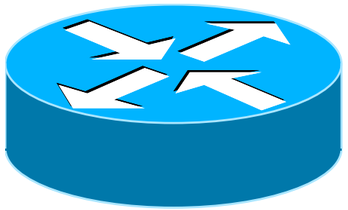
Data Plane

IP Routing protocols

Routing Information Base (RIB)

IP only packets forwarding

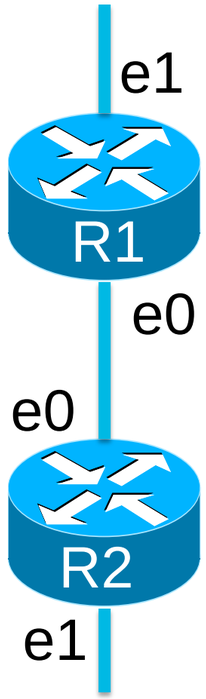
Cisco Express Forwarding (CEF) and its Forwarding Information Base (FIB)



The Birth of a Label

1.1.1.0/24 -> e1
2.2.2.2/32 -> e0

1.1.1.0/24 -> e0
2.2.2.2/32 -> e1

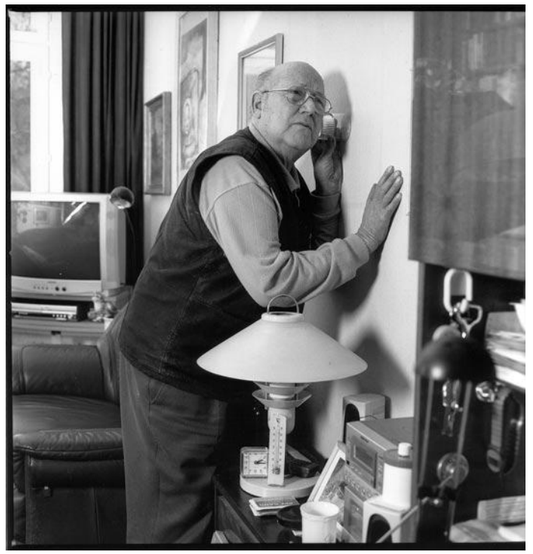


1.1.1.0/24 = label **16** -> e1
2.2.2.2/32 = label 20 -> e0

1.1.1.0/24 = label 35 -> e0
2.2.2.2/32 = label **16** -> e1

- For routes in the routing table, each router assigns a locally significant label for each IP route.

Won't You Be My Neighbor?



- Two step process
  - LDP neighbor discovery

    LDP link hello uses destination UDP port 646 and is sent to 224.0.0.2

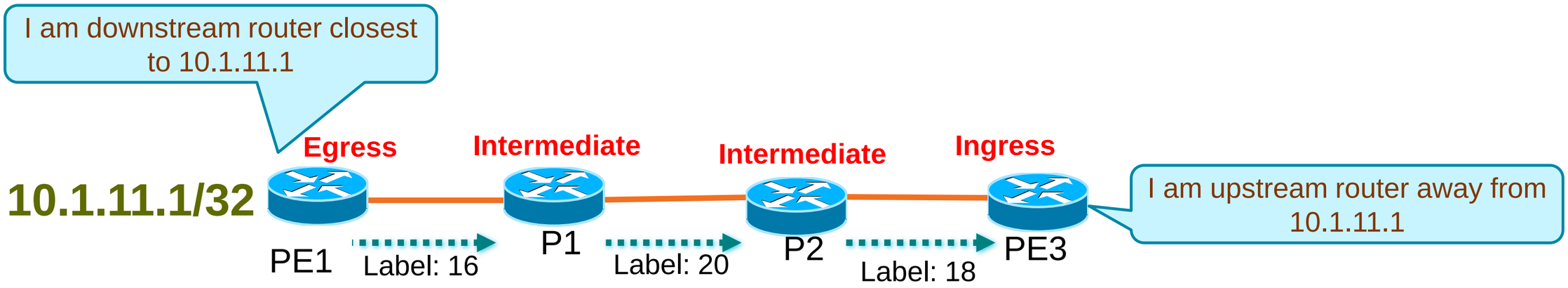
    Hello may include the IP address desired for peering, different than the source IP in the header.
  - Setup LDP session with neighbor who says hello.

    Session is TCP based on destination port 646

    Router with highest LDP router ID will initiate this TCP session (called the active LSR). Keepalives are sent every 60 seconds.

MPLS Label Distribution

- **Labels are created and advertised in Control plane**
- LDP label mapping
  - Each router assigns local label
  - Each router advertises that label
  - Label 3 is a reserved implicit null label for PHP



In	Out
E0/0 16	No label

In	Out
E0/1 20	E0/0 16

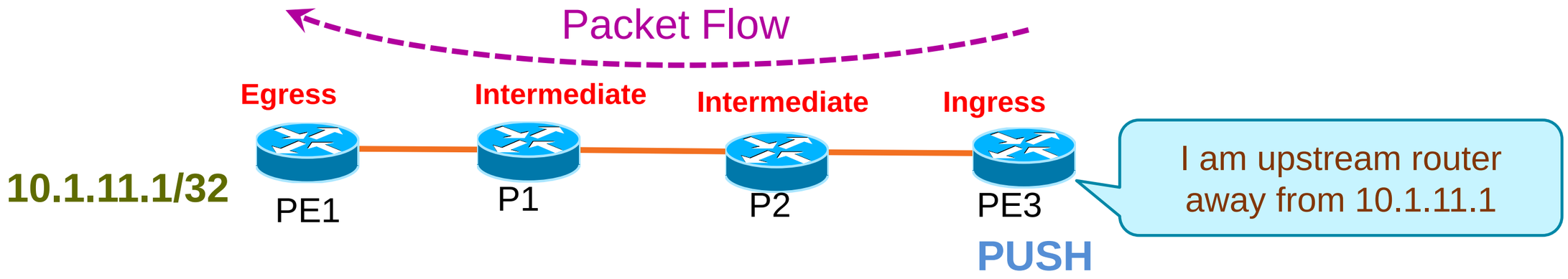
In	Out
E0/0 18	E0/1 20

In	Out
No label	E0/0 18

How Routers Use Labels

Three Major "Operations" Have Been Defined for Data plane

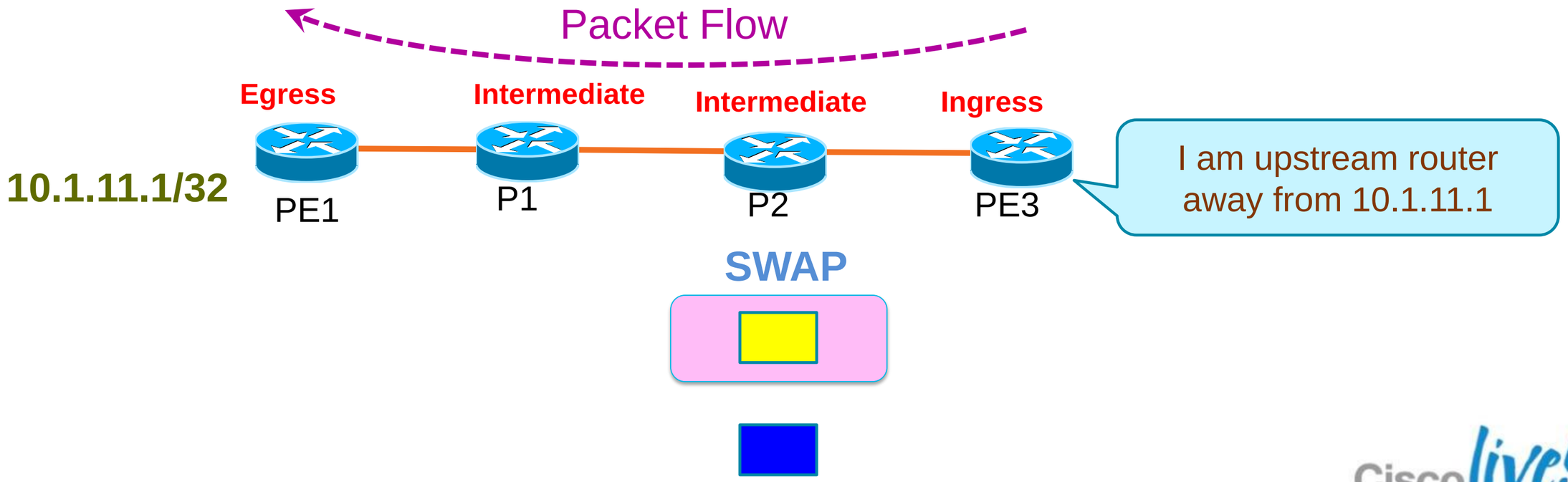
- PUSH – impose label (Ingress Router)
- POP – dispose label (Egress Router)
- SWAP – which is a pop/push combo (Intermediate Router)



How Routers are Use Labels

Three Major "Operations" Have Been Defined for Data plane

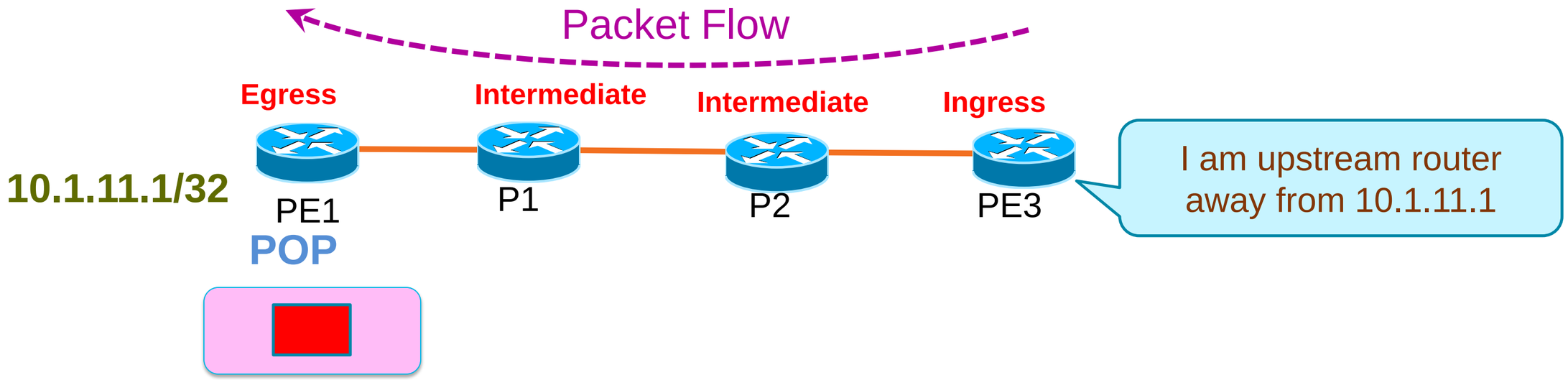
- PUSH – impose label (Ingress Router)
- POP – dispose label (Egress Router)
- SWAP – which is a pop/push combo (Intermediate Router)



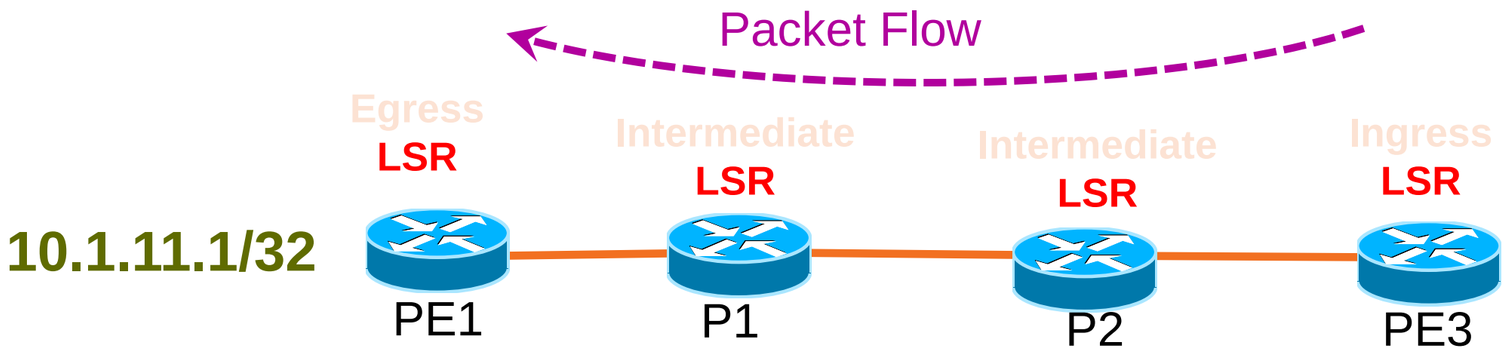
How Routers are Use Labels

Three Major "Operations" Have Been Defined for Data plane

- PUSH – impose label (Ingress Router)
- POP – dispose label (Egress Router)
- SWAP – which is a pop/push combo (Intermediate Router)



Device Roles on the Topology



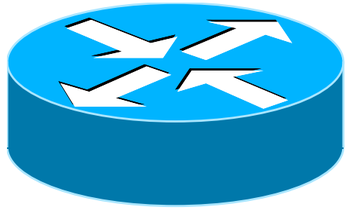
- LSR – Label Switch Router
  - Router that supports MPLS
- Ingress LSR (upstream)
  - Provider Edge (PE) first hop. Takes IP naked transit packet and pushes/imposes a label and forwards.
- Intermediate LSR
  - Provider (P) takes labelled packet and swaps labels and forwards to next LSR
- Egress LSR (downstream)
  - Provider Edge (PE) last hop. Pops/disposes label and forwards naked IP packet

Where Do Routers Turn to for Lookups?

Control Plane

Data Plane

IP Routing protocols



Routing Information Base (RIB)

IP only packets forwarding

Cisco Express Forwarding (CEF) and its Forwarding Information Base (FIB)

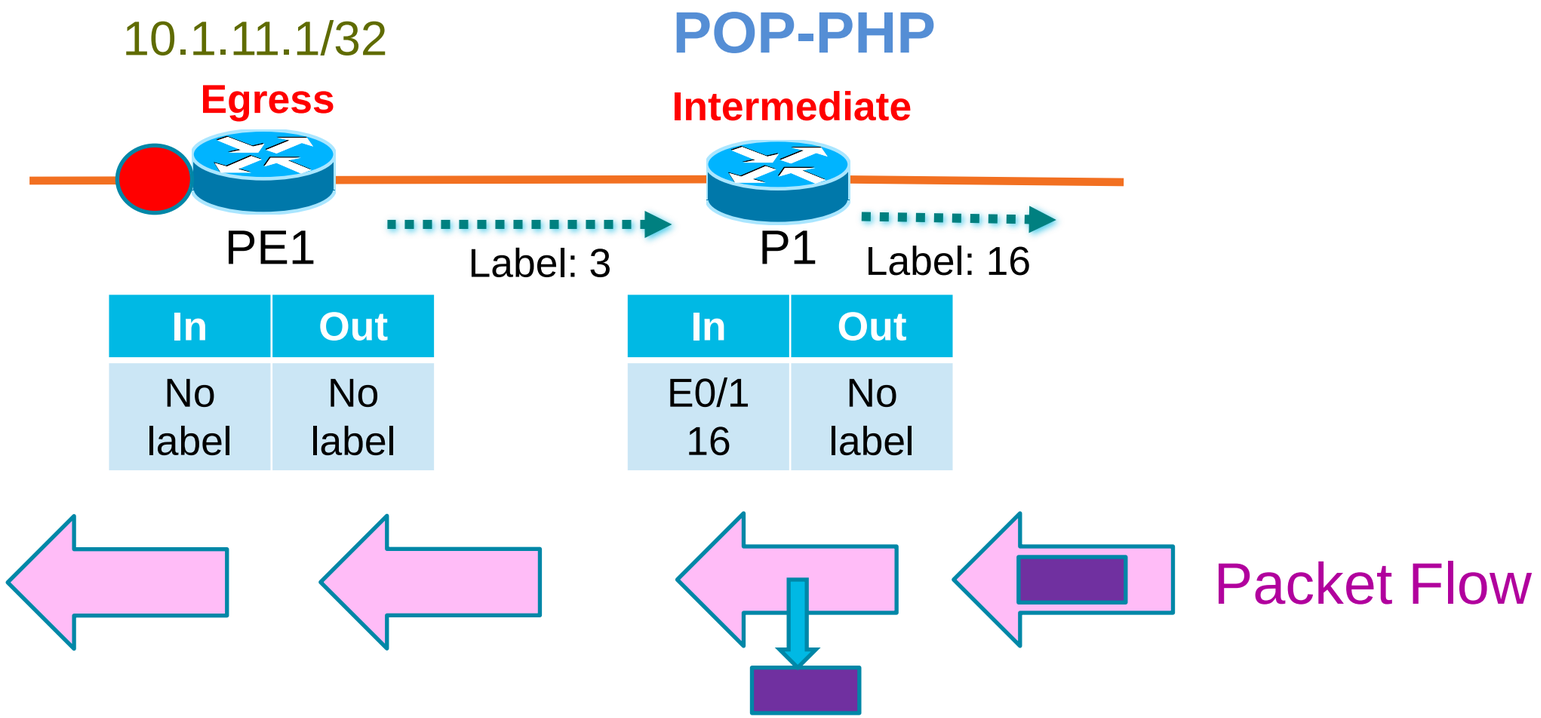
Label Distribution Protocol (LDP)

Label Information Base (LIB)

MPLS labeled packet forwarding

Label Forwarding Information Base (LFIB)

PHP



- PHP – Penultimate Hop Pop
  - Next to last LSR, removes top label, so that egress LSR (PE) doesn't have to
- Penultimate Hop Popping (PHP) saves the egress PE from an extra database lookup.

Follow the Bouncing Ball (Label)

```
PE1#show mpls ldp bindings 10.1.11.1 32
  tib entry: 10.1.11.1/32, rev 4
        local binding:  tag: imp-null
```

```
P1#show mpls ldp bindings 10.1.11.1 32
  tib entry: 10.1.11.1/32, rev 2
        local binding:  tag: 16
        remote binding: tsr: 10.1.1.2:0, tag: 16
        remote binding: tsr: 10.1.11.1:0, tag: imp-null
```

```
P2#show mpls ldp bindings 10.1.11.1 32
  tib entry: 10.1.11.1/32, rev 2
        local binding:  tag: 16
        remote binding: tsr: 10.1.1.1:0, tag: 16
```

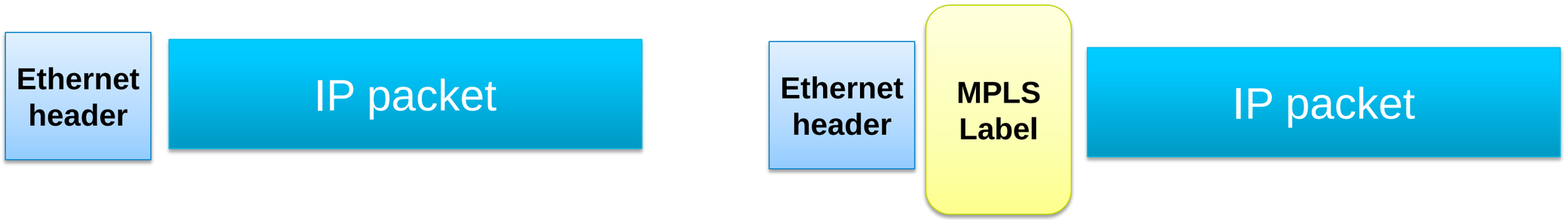
```
PE3#show mpls ldp binding 10.1.11.1 32
  tib entry: 10.1.11.1/32, rev 2
        local binding:  tag: 16
        remote binding: tsr: 10.1.1.2:0, tag: 16
PE3#show mpls forwarding-table 10.1.11.1
Local  Outgoing    Prefix           Bytes tag  Outgoing    Next Hop
tag    tag or VC   or Tunnel Id     switched   interface
16     16          10.1.11.1/32     0          Et0/0       10.2.23.1
```

The Order of Things

- IP IGP routing protocols build the IP tables
- LSRs assign a ***local*** label for each route
- LSRs share their labels with other LSRs using LDP
- LSRs build their own LIB, LFIB and FIBs based on what they have learned from their LDP neighbors

MPLS MTU Problem

- MTU is automatically adjusted
- You can change with **mpls mtu** command
  - mpls mtu 1512 -- would support 3 labels (4 bytes per MPLS header)



- Large packets dropped

  MTU not supported by switches. Multiple labels may be present pushing the MTU to a size not supported by the infrastructure.

LDP Features

- **(config)# no mpls ldp advertise-labels**
- **(config)# mpls ldp advertise-labels [for (ACL-of-networks)] [to (ACL-peers)]**
- **(config-if)# mpls label range 200 120000**

- Security – Computes MD5 Signatures
  - **(config)# mpls ldp neighbor (ip#) password (pw)**

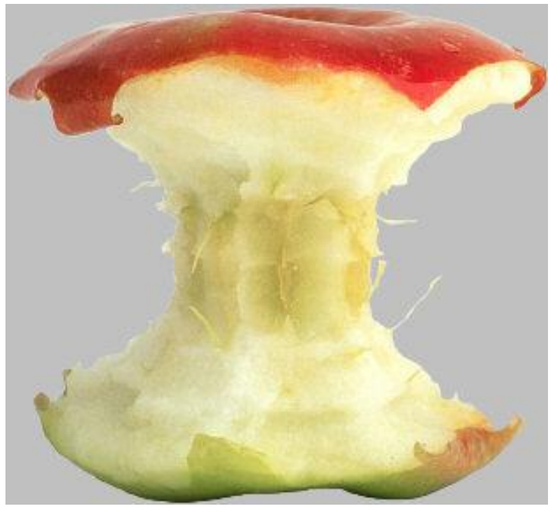
- Label filters – inbound from neighbor
  - **(config)# mpls ldp neighbor (ip#) labels accept (#)**

    (ip#) = IP address of LDP neighbor

    (#) = number of access-list of network prefixes

Hide the MPLS Core from the Client

- Traceroute uses TTL manipulation to trigger feedback.
- Disabling the TTL propagation will not copy the initial IP TTL to the MPLS TTL, and MPLS will start at 255.
- Results: MPLS LSRs become the invisible network to the eyes of traceroute.



No mpls ip propagate-ttl (on All LSRs)

```
PE3(config)#do trace 10.1.11.1

Type escape sequence to abort.
Tracing the route to 10.1.11.1

  1 10.2.23.1 [MPLS: Label 16 Exp 0] 12 msec 0 msec 0 msec
  2 10.2.12.1 [MPLS: Label 16 Exp 0] 0 msec 4 msec 0 msec
  3 10.2.11.1 0 msec *  0 msec
PE3(config)#no mpls ip propagate-ttl
PE3(config)#do trace 10.1.11.1

Type escape sequence to abort.
Tracing the route to 10.1.11.1

  1 10.2.11.1 4 msec *  0 msec
PE3(config)#
```

Troubleshooting MPLS

- LDP neighborship failed

  MPLS not enabled, LDP TCP/646 or TDP TCP/711 ports filtered, no L3 route to LDP neighbor LSR router-id, highest loopback address.

- Labels not assigned

  CEF not enabled

- Labels not shared

  LDP/TDP between neighbors

- Slow convergence

  Get rid of RIP ☺ IGP is biggest factor in convergence delay

- Large packets dropped

  MTU not supported by switches. Multiple labels may be present pushing the MTU to a size not supported by the infrastructure.

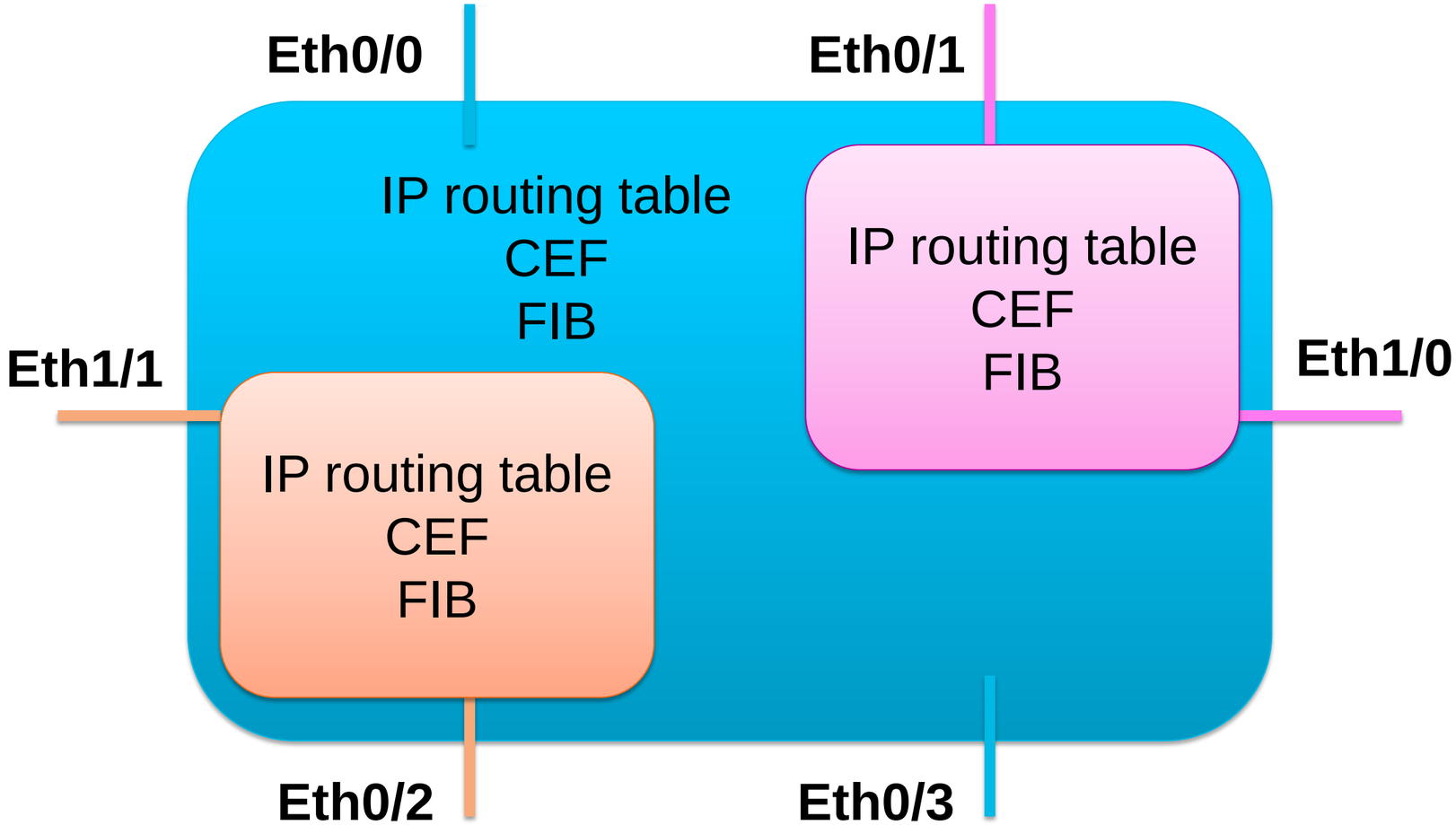
Minimal Configuration to Enable MPLS

- (config)# ip cef
- (config)# mpls ip
- (config)# interface ethernet 0/0
- (config-if)# mpls ip
- (config)# interface ethernet 0/1
- (config-if)# mpls ip

What are VRFs

VRF: The Virtual Routing Table

- VRF: Virtual Routing and Forwarding instance

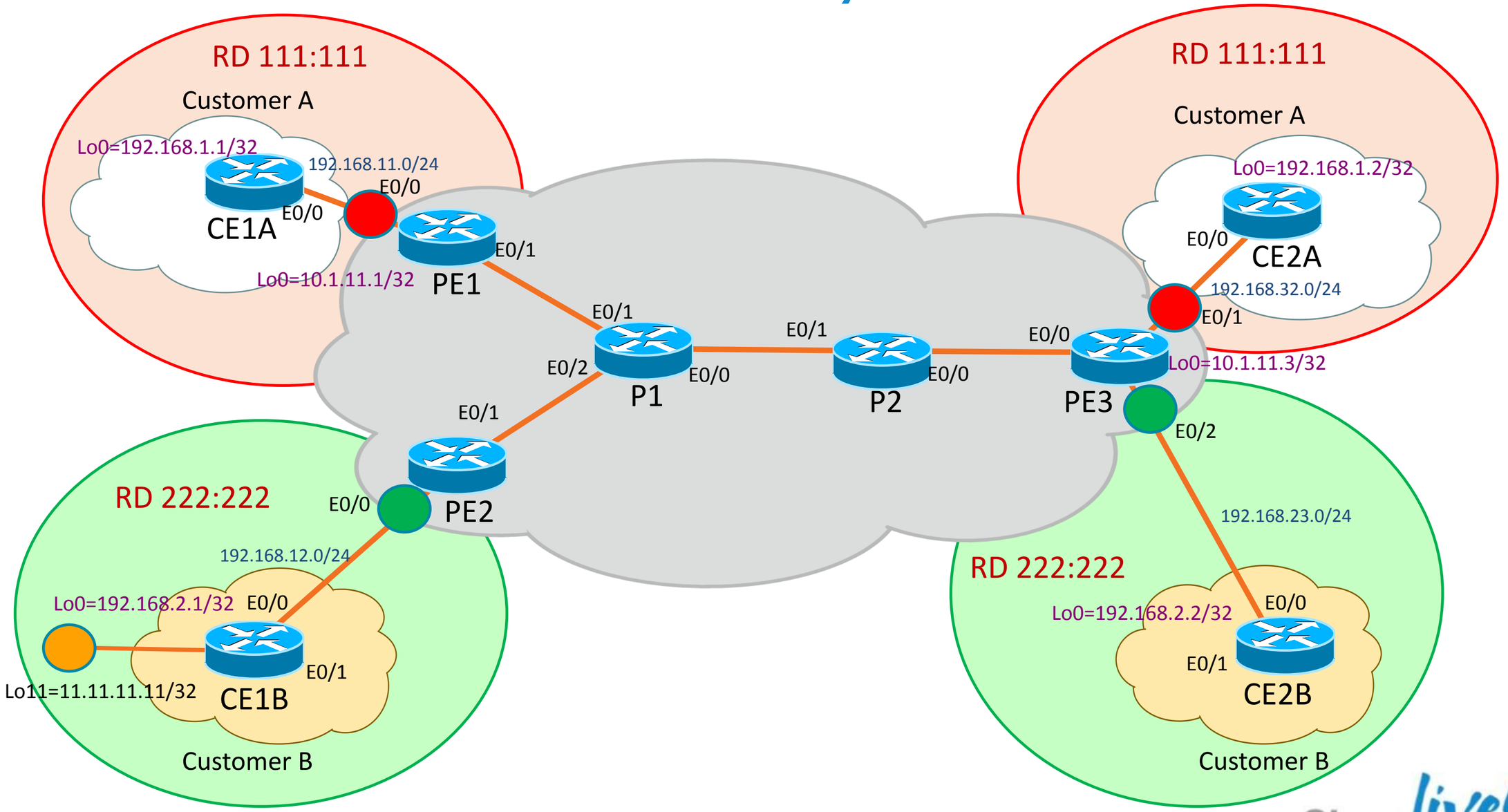


Uniquely Identifying Similar Routes

- What if customer A and customer B both have a 10.0.0.0/8 network, how do we differentiate these?

- Route Distinguisher (RD) is a 64-bit identifier pre-pended to each IP address to make it globally unique into our network
- The resulting 96-bit address is called VPNv4 address
- VPNv4 addresses can only be exchanged via BGP between PE routers
  - BGP supporting other address families than IPv4 addresses is called multi-protocol BGP or MP-BGP

Vrf RED and vrf GREEN, with Interfaces



Routing between Customer and Provider

```
PE2(config)#router ospf 1 vrf GREEN
PE2(config-router)#network 192.168.12.0 0.0.0.255 area 0
PE2(config-router)#
*Nov 19 04:43:14.661: %OSPF-5-ADJCHG: Process 1, Nbr 192.168.2.1 on Ethernet
0/0 from LOADING to FULL, Loading Done
```

To exchange routes between Customer and Provider's VRF we can use OSPF, RIP, EIGRP, BGP or static routing.

Provider side router runs this protocols inside VRFs or uses VRF-aware implementations

Minimal Configuration for VRF

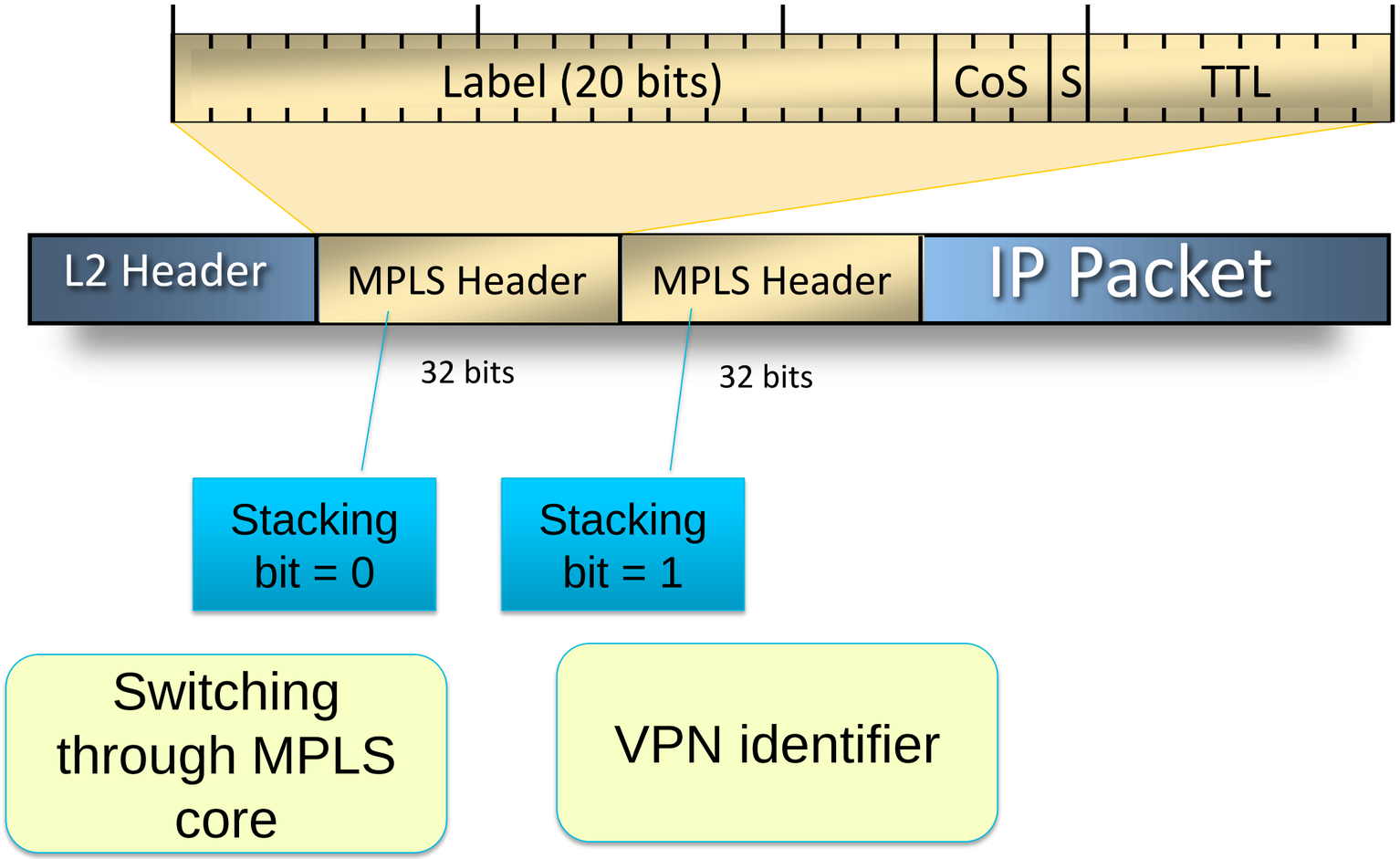
```
ip vrf CompanyC
 rd 300:300

interface e0/0
 ip vrf forwarding CompanyC
 ip address 10.2.22.2 255.255.255.0
```

MPLS L3 VPNs

Ingredient List for MPLS L3 VPNs…

- MPLS
- VRFs
- BGP
- Routers

Stacks



MP-BGP for Exchange Addresses

Only MP-BGP can exchange VPNv4 addresses (if you remember, this is address constructed from RD and IP address).
So we need to configure MP-BGP between PE routers and enable VPNv4 address-family routing exchange

```
PE1(config)#router bgp 65530
PE1(config-router)#address-family vpnv4 unicast
PE1(config-router-af)#neigh 10.1.11.3 activate
```

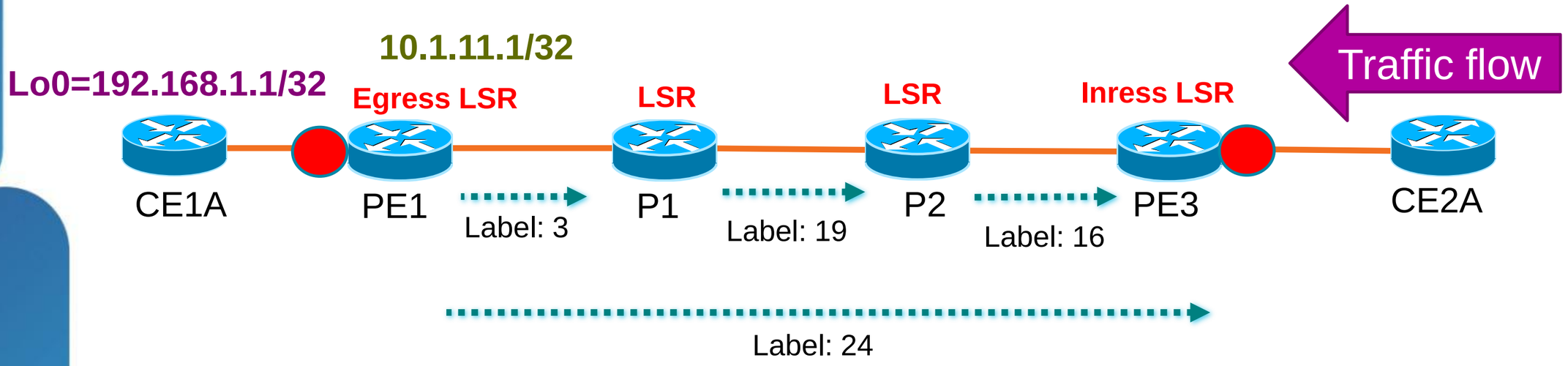
BGP routing instances for VRFs will be created automatically - to be able transfer routes from IGP via VPNv4 (uses routes redistribution)

MP-BGP for Exchange Labels

Labels are delivered into extended communities together with VPNv4 addresses.

```
PE3#show ip bgp vpnv4 vrf RED 192.168.1.1
BGP routing table entry for 111:111:192.168.1.1/32, version 37
Paths: (1 available, best #1, table RED)
  Advertised to update-groups:
        2
  Local, (Received from a RR-client)
    10.1.11.1 (metric 40) from 10.1.11.1 (10.1.11.1)
      Origin incomplete, metric 11, localpref 100, valid, internal, best
      Extended Community: RT:100:100 OSPF DOMAIN ID:0x0005:0x000000020200
        OSPF RT:0.0.0.0:2:0 OSPF ROUTER ID:192.168.11.2:512
      mpls labels in/out nolabel/24
```

MPLS Labels for VPN (Control plane)



LDP

In	Out
No label	No label

In	Out
E0/1 19	No label

In	Out
E0/0 16	E0/1 19

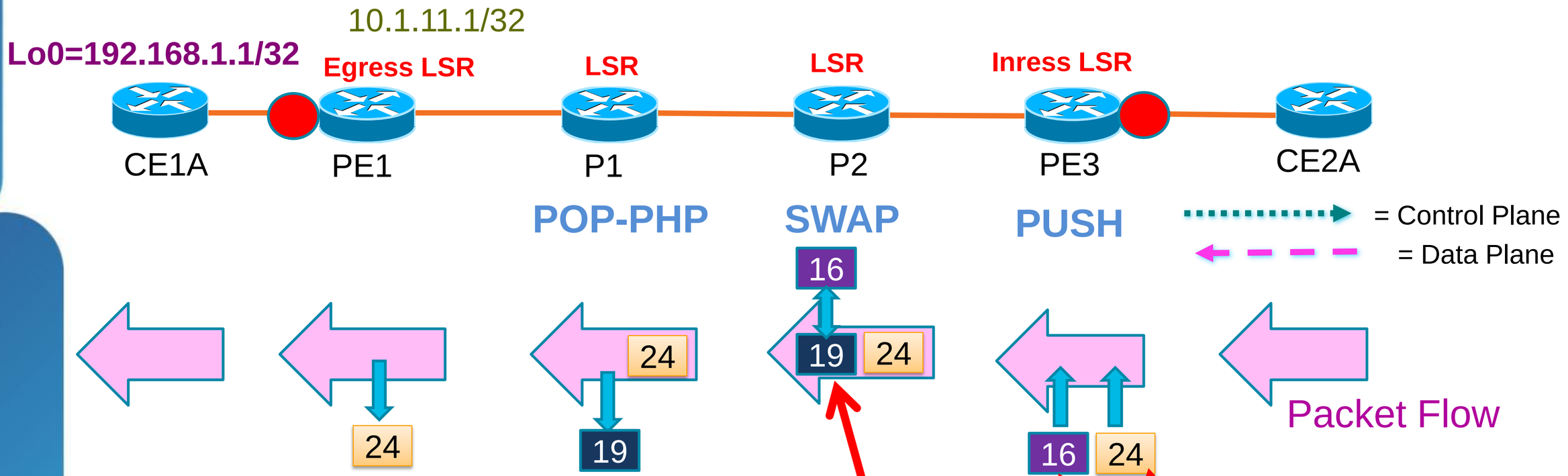
In	Out
No label	E0/0 16

MP-BGP

In	Out
24	No label

In	Out
No label	24

MPLS Labels for VPN (Data plane)



In	Out
No label	No label

In	Out
E0/1 19	No label

In	Out
E0/0 16	E0/1 19

In	Out
No label	E0/0 16

LDP

MP-BGP

In	Out
24	No label

Label: 24

In	Out
No label	24

MP-BGP for Exchange Routes

Check BGP instance for VRF (it should be enabled automatically when you configure VRF and MP-BGP)

```
address-family ipv4 vrf RED
 no synchronization
exit-address-family
```

Configure redistribution

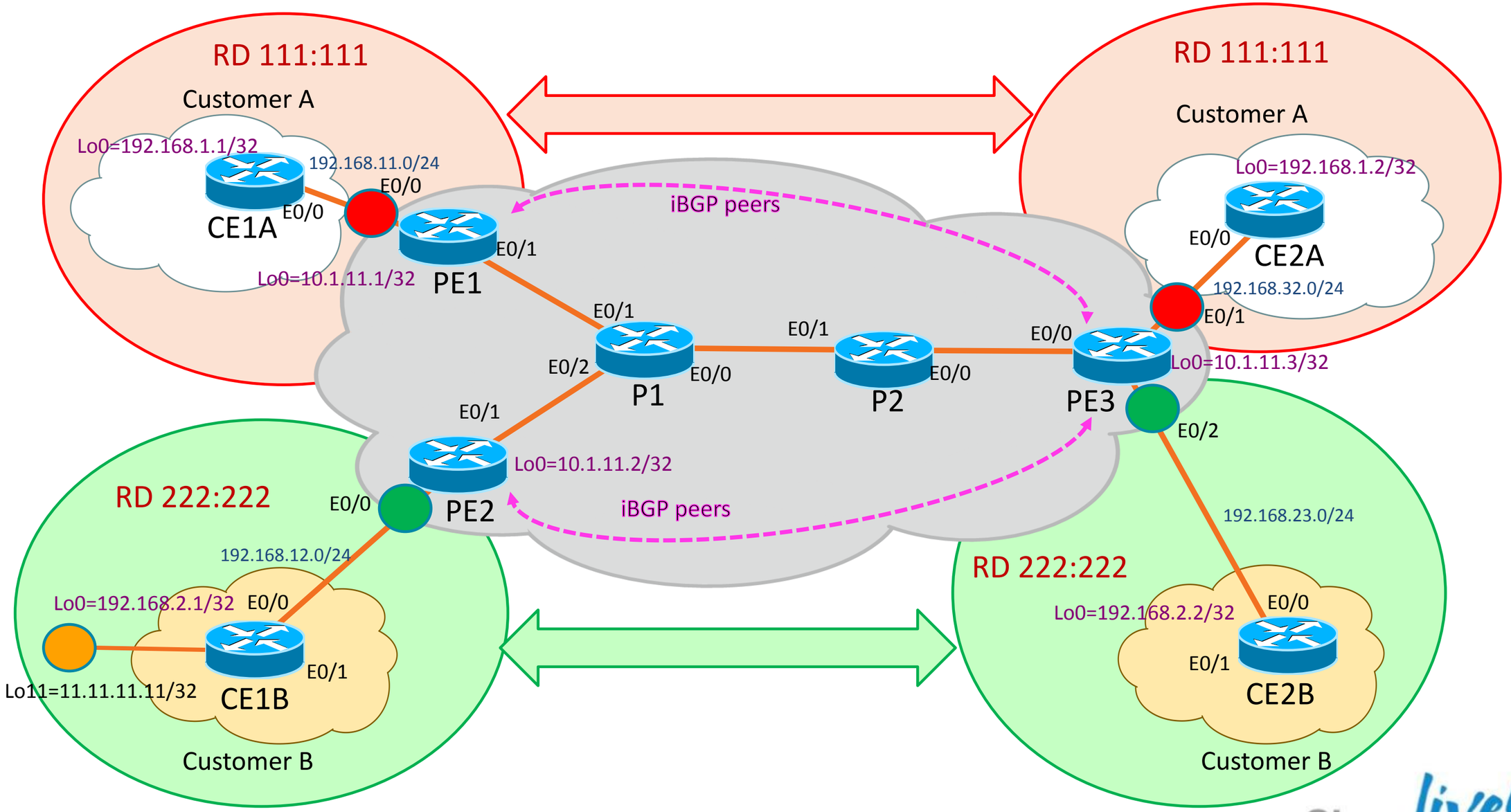
```
PE3(config)#router bgp 65530
PE3(config-router)#address-family ipv4 vrf RED
PE3(config-router-af)#redistribute ospf 2 vrf RED
```

```
PE3(config)#router ospf 2 vrf RED
PE3(config-router)#redistribute bgp 65530 subnets
```

Viewing Routes in MP-BGP by RD

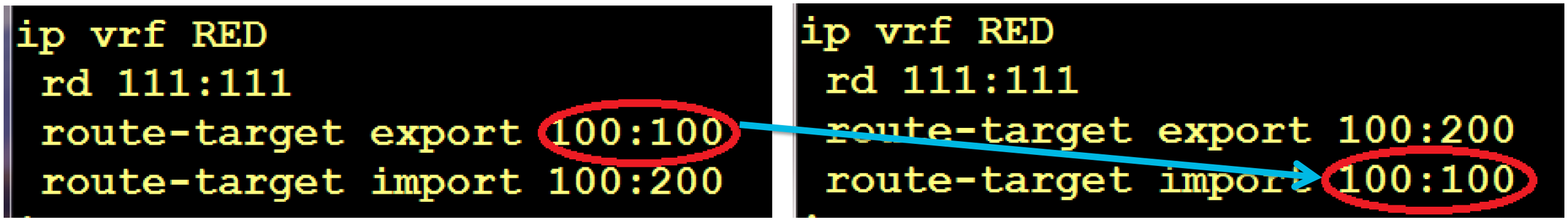
```
PE3#sh ip bgp vpnv4 rd 222:222
BGP table version is 17, local router ID is 10.1.11.3
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
              r RIB-failure, S Stale
Origin codes: i - IGP, e - EGP, ? - incomplete

   Network          Next Hop            Metric LocPrf Weight Path
Route Distinguisher: 222:222 (default for vrf GREEN)
*>i11.11.11.11/32   10.1.11.2                11    100      0 ?
*>i192.168.2.1/32   10.1.11.2                11    100      0 ?
*> 192.168.2.2/32   192.168.23.1             11           32768 ?
*>i192.168.12.0     10.1.11.2                 0    100      0 ?
*> 192.168.23.0     0.0.0.0                   0           32768 ?
*> 192.168.112.0    192.168.23.1             20           32768 ?
```

How do we Deliver the Correct Routes?

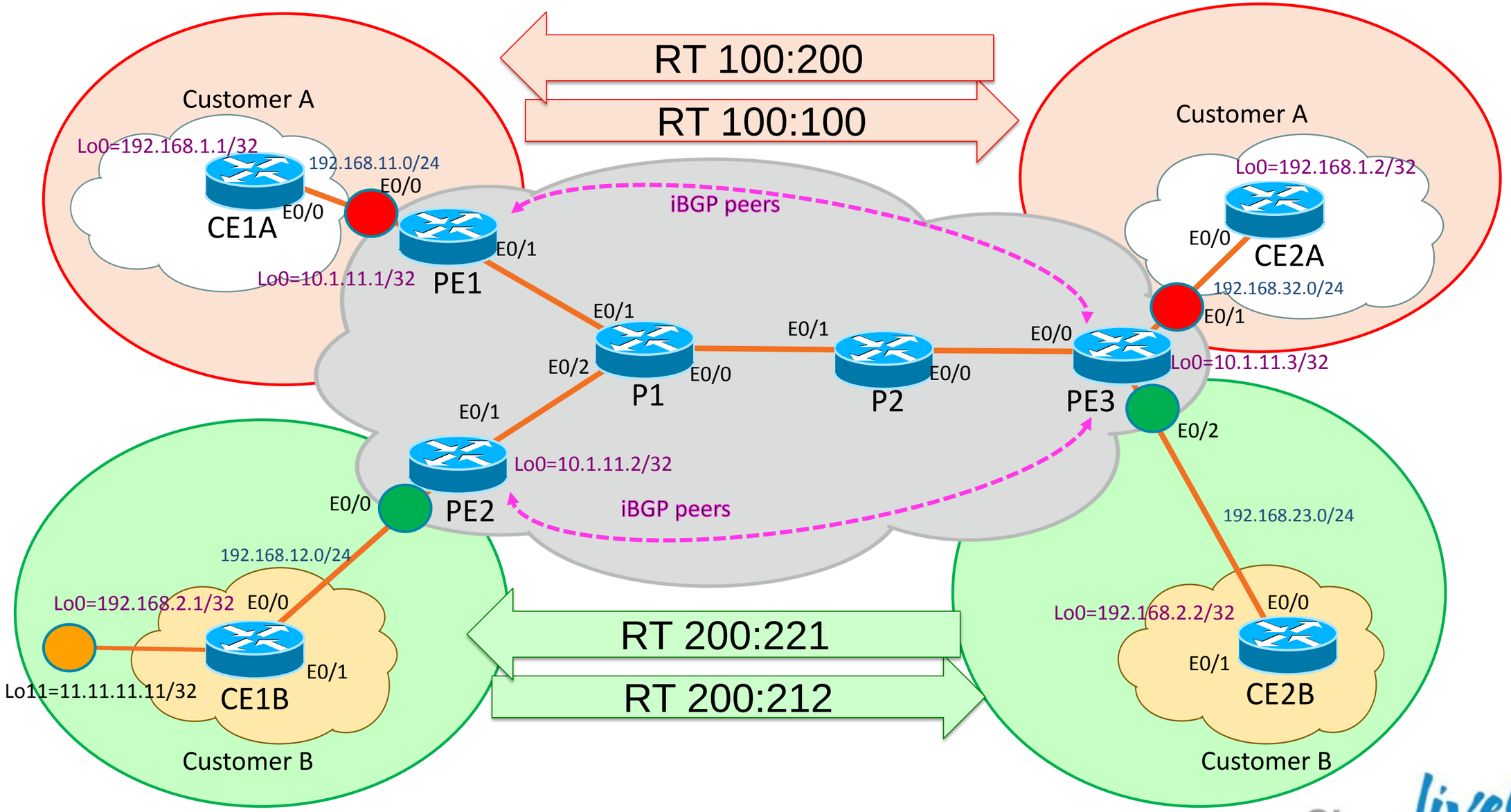


Route Target

- Route Targets are additional attributes attached to VPNv4 BGP routes to indicate VPN membership
- Extended BGP communities are used to encode these attributes
  - Extended communities carry the meaning of the attribute together with its value
- Any number of route targets can be attached to a single route



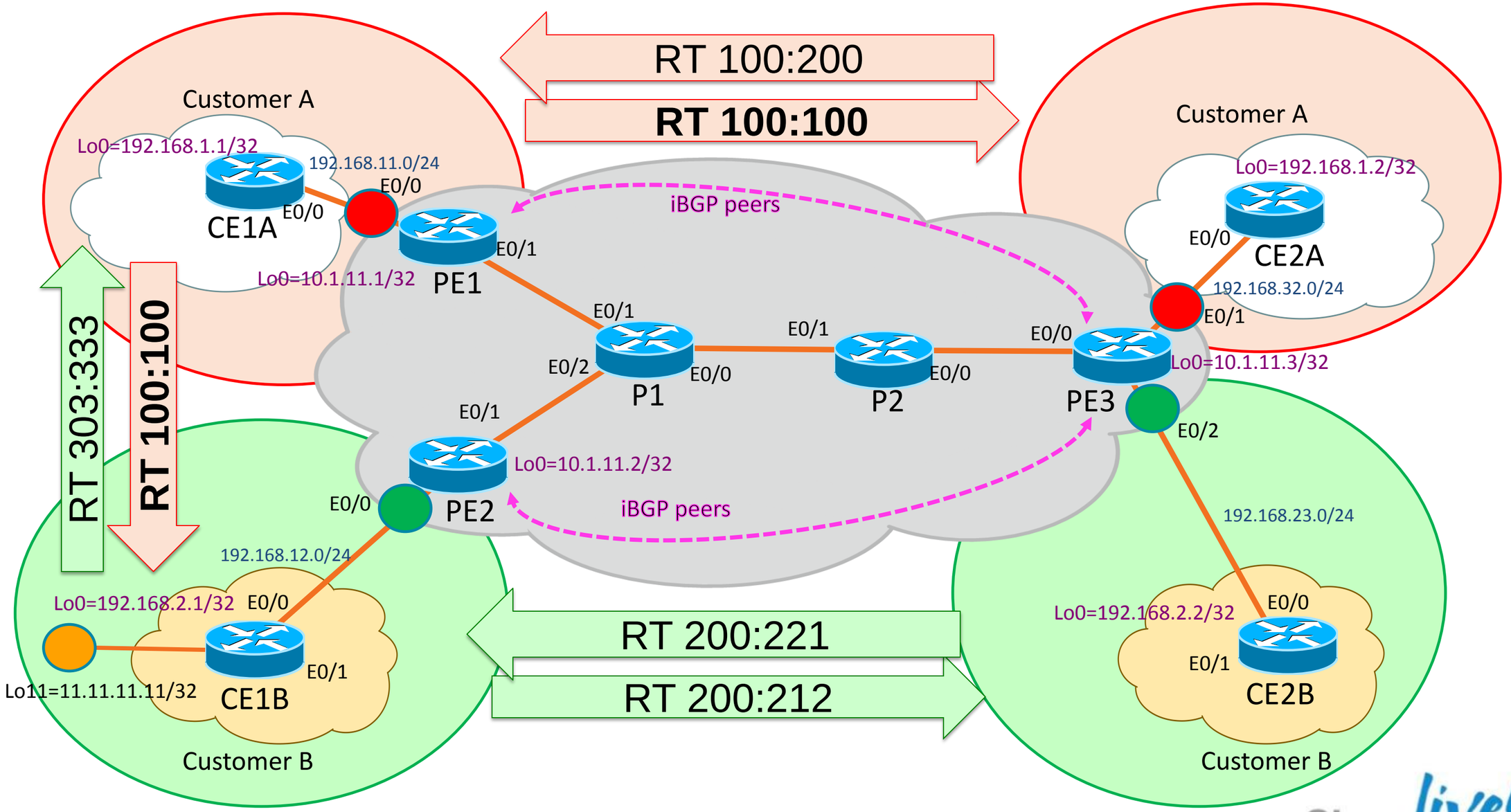
How do we Deliver the Correct Routes?



Viewing the Route Target in MP-BGP

```
PE3#show ip bgp vpnv4 vrf GREEN 192.168.2.1
BGP routing table entry for 222:222:192.168.2.1/32, version 53
Paths: (1 available, best #1, table GREEN)
  Advertised to update-groups:
        2
  Local, (Received from a RR-client)
    10.1.11.2 (metric 40) from 10.1.11.2 (10.1.11.2)
      Origin incomplete, metric 11, localpref 100, valid, internal, best
      Extended Community: RT:200:212 OSPF DOMAIN ID:0x0005:0x000000010200
        OSPF RT:0.0.0.0:2:0 OSPF ROUTER ID:192.168.12.2:512
      mpls labels in/out nolabel/25
```

How do we Deliver the Correct Routes?



How to Control RT?

```
access-list 10 permit 11.11.11.11
!
route-map SHARED permit 10
 match ip address 10
 set extcommunity rt 303:333
```

```
ip vrf GREEN
 rd 222:222
 export map SHARED
 route-target export 200:212
 route-target import 200:221
```

```
ip vrf RED
 rd 111:111
 route-target export 100:100
 route-target import 100:200
 route-target import 303:333
```

We can mark route with the other extcommunity and receive on the other VRF.
Do not forget about REVERSE PATH!!!

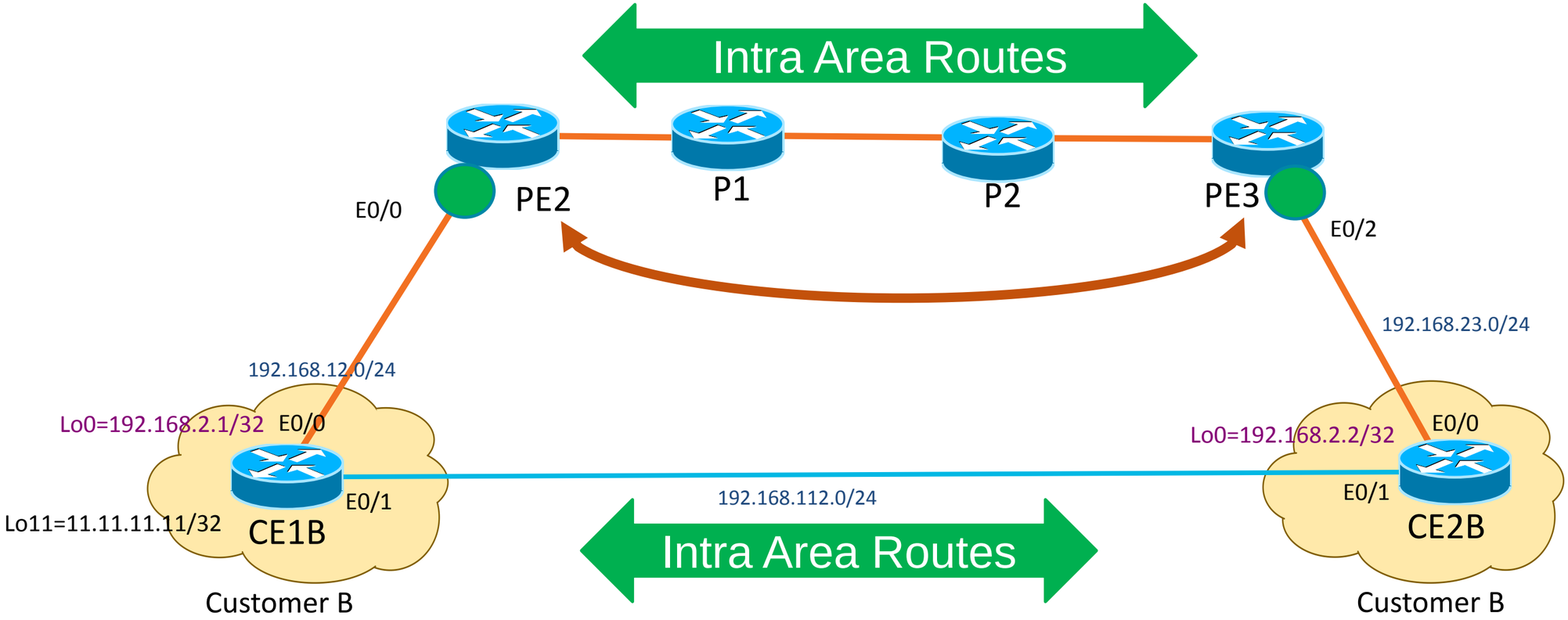
Customer Routing Table (CE1A)

```
CE1A#show ip route
Codes: C - connected, S - static, R - RIP, M - mobile, B - BGP
       D - EIGRP, EX - EIGRP external, O - OSPF, IA - OSPF inter area
       N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
       E1 - OSPF external type 1, E2 - OSPF external type 2
       i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
       ia - IS-IS inter area, * - candidate default, U - per-user static route
       o - ODR, P - periodic downloaded static route

Gateway of last resort is not set

C    192.168.11.0/24 is directly connected, Ethernet0/0
     11.0.0.0/32 is subnetted, 1 subnets
O E2    11.11.11.11 [110/11] via 192.168.11.2, 01:37:16, Ethernet0/0
     192.168.1.0/32 is subnetted, 2 subnets
C       192.168.1.1 is directly connected, Loopback0
O IA    192.168.1.2 [110/21] via 192.168.11.2, 02:57:35, Ethernet0/0
O IA 192.168.32.0/24 [110/11] via 192.168.11.2, 02:57:35, Ethernet0/0
```

MPLS VPN LSP Troubleshooting

- Look for route to BGP next-hop
  - **show ip route**
- Check VRF for route “inside”
  - **show ip route vrf RED**
- Check LFIB for route
  - **show mpls forwarding vrf [name] [address] detail**
- Check MP-BGP VPNv4
  - **show ip bgp vpnv4 vrf [name] [address]**
- Check LFIB for next-hop
  - **show mpls forwarding [address] detail**

Sham-Link to Correct OSPF Backdoor



Configure Sham-Link

```
interface Loopback1
 ip vrf forwarding GREEN
 ip address 22.1.1.2 255.255.255.255
```

```
address-family ipv4 vrf GREEN
 redistribute ospf 1 vrf GREEN
 no synchronization
 network 22.1.1.2 mask 255.255.255.255
```

```
router ospf 1 vrf GREEN
 log-adjacency-changes
 area 0 sham-link 22.1.1.2 22.1.1.3
```

```
interface Loopback1
 ip vrf forwarding GREEN
 ip address 22.1.1.3 255.255.255.255
```

```
address-family ipv4 vrf GREEN
 redistribute ospf 1 vrf GREEN
 no synchronization
 network 22.1.1.3 mask 255.255.255.255
```

```
router ospf 1 vrf GREEN
 log-adjacency-changes
 area 0 sham-link 22.1.1.3 22.1.1.2
```

Configure a separate /32 address on the remote PE for sham-link. This /32 address must meet the following criteria:

- **Belong to a VRF.**
- **Not be advertised by OSPF.**
- **Be advertised by BGP.**

Associate the sham-link with an existing OSPF area.

Minimal MPLS L3 VPN configuration

- ip vrf CompanyA
- route-target both 300:300

- interface Ethernet 0/0
- ip vrf forwarding CompanyA
- ip address 192.168.10.1 255.255.255.0

- router bgp 65530
- address-family vpnv4 unicast
- neighbor 10.10.10.1 activate
- address-family ipv4 vrf CompanyA
- redistribute connected

- ip vrf CompanyA
- route-target both 300:300

- interface Ethernet 0/0
- ip vrf forwarding CompanyA
- ip address 192.168.20.1 255.255.255.0

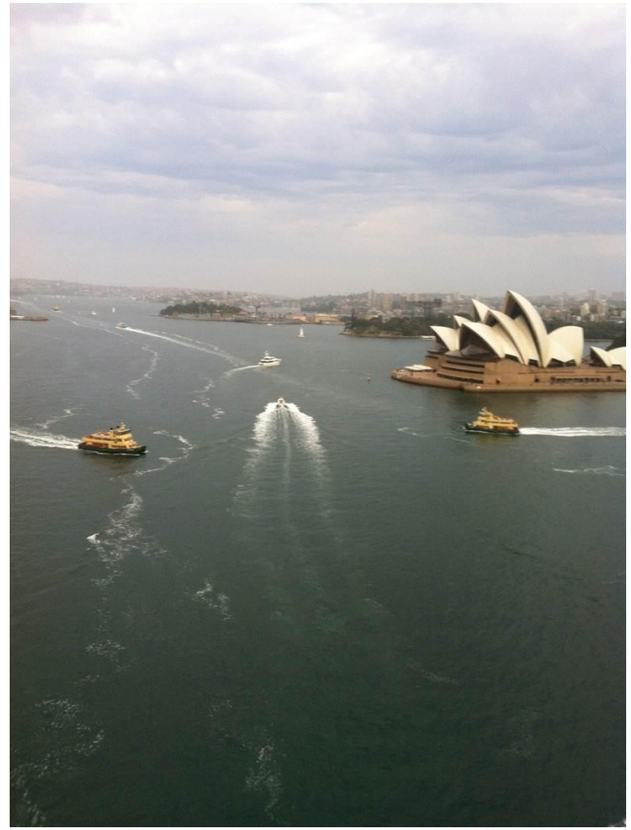
- router bgp 65530
- address-family vpnv4 unicast
- neighbor 10.10.20.1 activate
- address-family ipv4 vrf CompanyA
- redistribute connected

Tasks into CCIE Labs

Into CCIE Labs and into Real Life

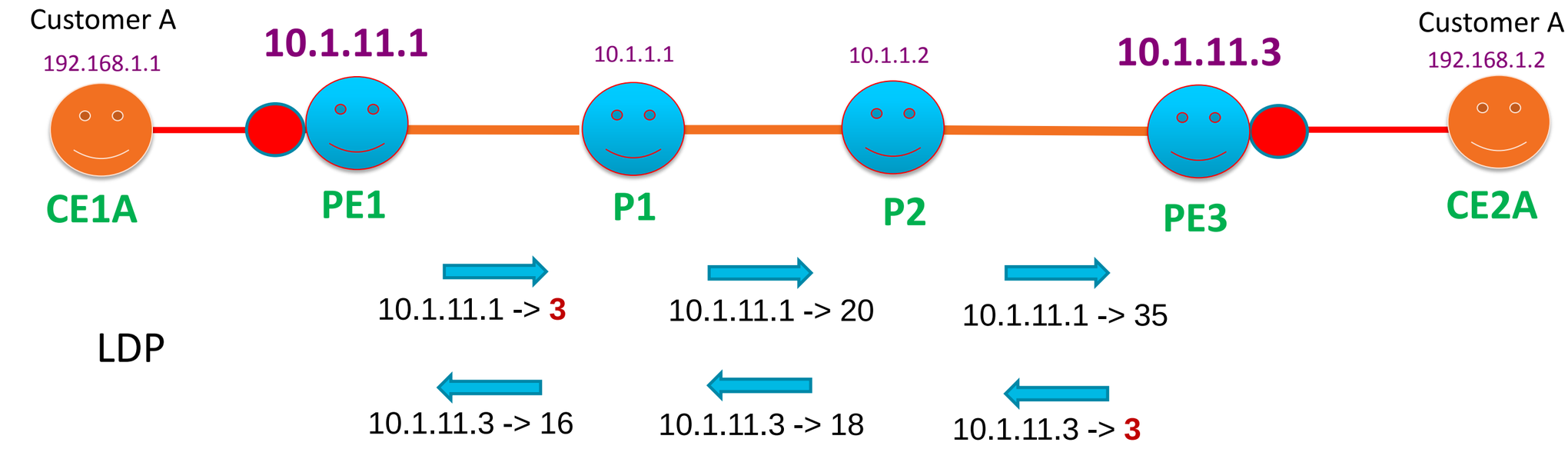
- CCIE R&S TS (Troubleshooting) part – problems with MPLS, LDP, BGP
- CCIE R&S CFG (Configuration) part – MPLS L3VPN configuration with features

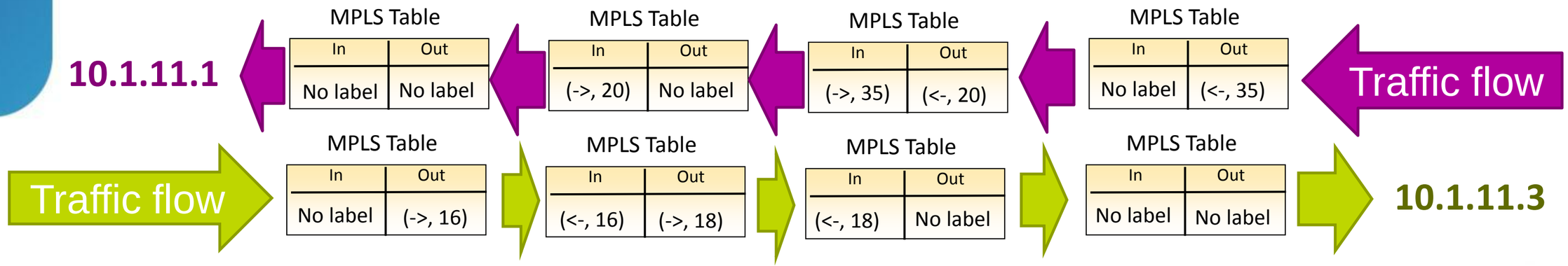
- CCIE SP – ☺
- CCIE Security
- CCIE Wireless
- CCIE Voice
- CCIE DataCentre – VRF, MPLS



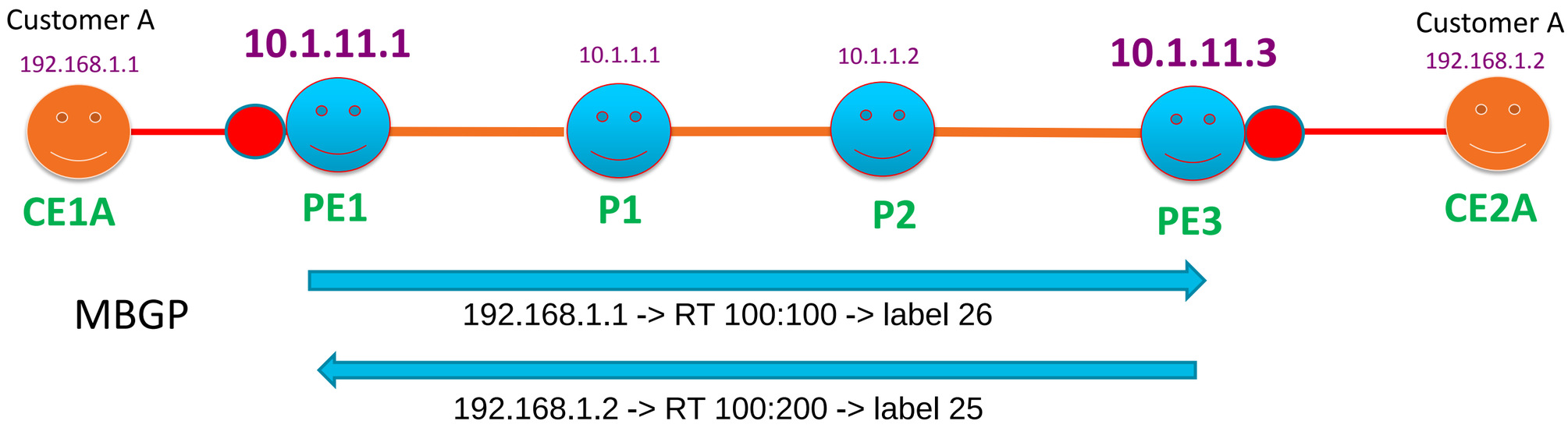
MPLS L3 VPN Game

Label Assignment for MPLS Core





Label Assignment for VPN



Q & A

Complete Your Online Session Evaluation

Give us your feedback and receive a Cisco Live 2013 Polo Shirt!

Complete your Overall Event Survey and 5 Session Evaluations.

- Directly from your mobile device on the Cisco Live Mobile App
- By visiting the Cisco Live Mobile Site www.ciscoliveaustralia.com/mobile
- Visit any Cisco Live Internet Station located throughout the venue

Polo Shirts can be collected in the World of Solutions on Friday 8 March 12:00pm-2:00pm

Cisco live! 365

Don’t forget to activate your Cisco Live 365 account for access to all session material, communities, and on-demand and live activities throughout the year. Log into your Cisco Live portal and click the "Enter Cisco Live 365" button.
www.ciscoliveaustralia.com/portal/login.ww

Cisco live!