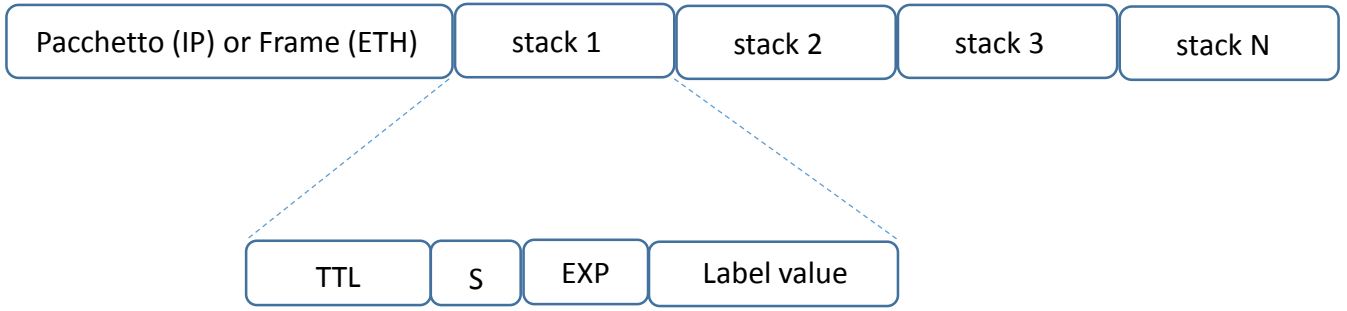
MPLS aware
MPLS L2-VPN VPWS (PW) and VPLS
MPLS L3-VPN unified RFC 3107 carry label information BGPv4
Core Aggregation Access design
configurations example

Massimiliano Sbaraglia

MPLS Header

- L’informazione è contenuta in un pacchetto MPLS costituito da una o più etichette (labels);
- Le etichette possono assumere differenti valori (RFC 3032) di servizio:

  - Label = 0 : IPv4 Explicit Null Label
  - Label = 1 : Router Alert
  - Label = 2 : IPv6 Explicit Null Label
  - Label = 3 : Implicit Null



TTL per Loop Prevention

S = 1 quando l’elemento È ultimo della pila di labels

EXP per differenziare qualità del servizio

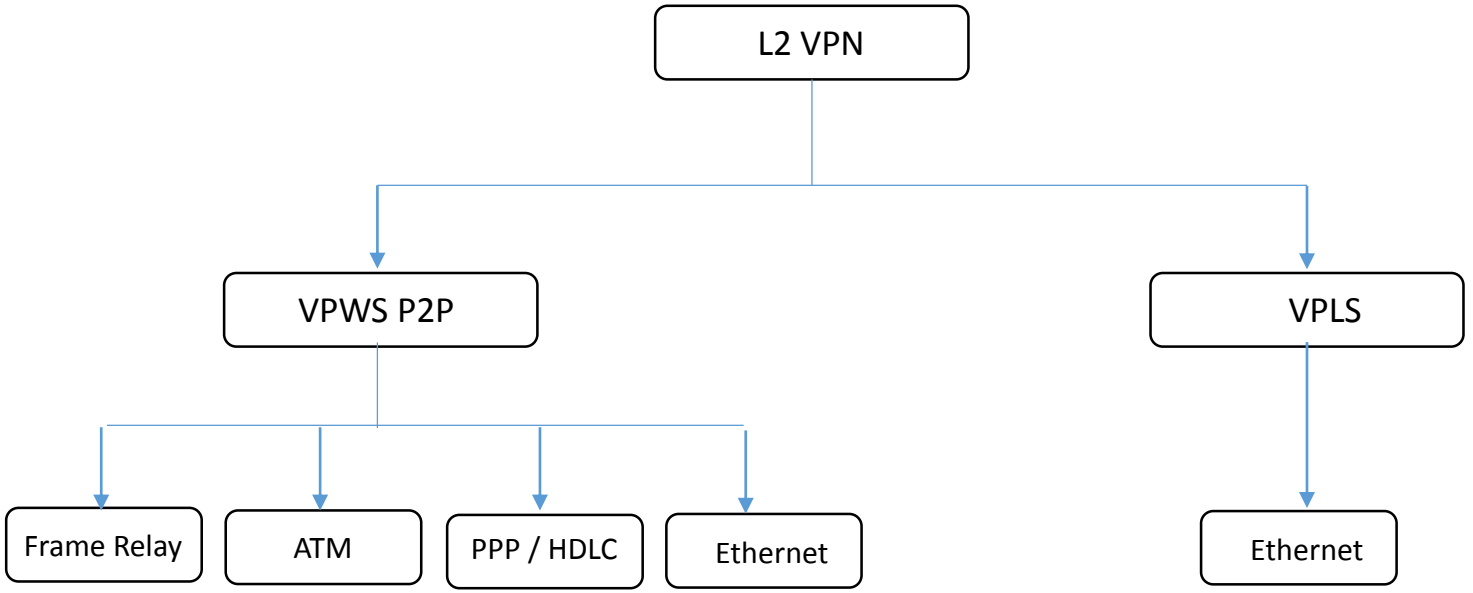
LSP Label Switch Path

- LSP di una rete MPLS può essere considerato analogo ad un PVC (VPI/VCI) di una rete ATM che si instaura tra due end-point routers;
- LSP di una rete MPLS può essere considerato un tunnel in cui transitano pacchetti IP o frame layer 2;
- LSP Hop by Hop è una connessione virtuale creata attraverso un protocollo di routing IGP presente in rete;
- LSP Esplicito è una connessione creata attraverso meccanismi di segnalazione:

    - RSVP: utilizzato per la prenotazione di banda trasmissiva tra due end-point routers
    - CR-LDP: utilizzato nella forma tradizionale MPLS con la distribuzione di labels lungo il percorso LSP

MPLS VPN-L2 application

- **Layer 2 VPN:** sono trasportate frame layer 2 attraverso una rete MPLS aware
    - LSP provvedono ad una connessione end-to-end MPLS rachability tra due end-point routers (transport LSP or tunnel LSP)
    - VPN point-to-point sono chiamate virtual private wire service (VPWS) oppure pseudowire
    - VPN Multi point-to-point chiamate virtual privat lan services (VPLS)



MPLS VPN-L2 application

Sia VPWS che VPLS possono essere create attraverso il metodo Kompella oppure Martini; in entrambi i metodi l'LSP è costruito tra due PE attraverso il protocollo LDP.

I circuiti L2 pseudowire con il metodo Kompella il trasporto delle labels è segnalato via LDP a differenza della VC (virtual circuit) label segnalata via MP-BGP; mentre con il metodo Martini entrambe le labels via LDP.

Il PE routers MPLS, collegato al CE customer devices, analizza la frame ethernet in ingresso e identifica quale egress router PE è usato per il trasporto della frame (primo lookup); il secondo lookup determina l'interfaccia egress presso il router PE egress

I pacchetti MPLS Layer 2 utilizzano due labels:

- Outer Label (topmost) utizzata per raggiungere il devices egress (di uscita)
- Inner Label (VC label) utilizzata per identificare il circuito pseudowire at Egress router PE

VPWS può trasportare qualsiasi payload layer 2; VPLS trasporta solo frame ethernet.

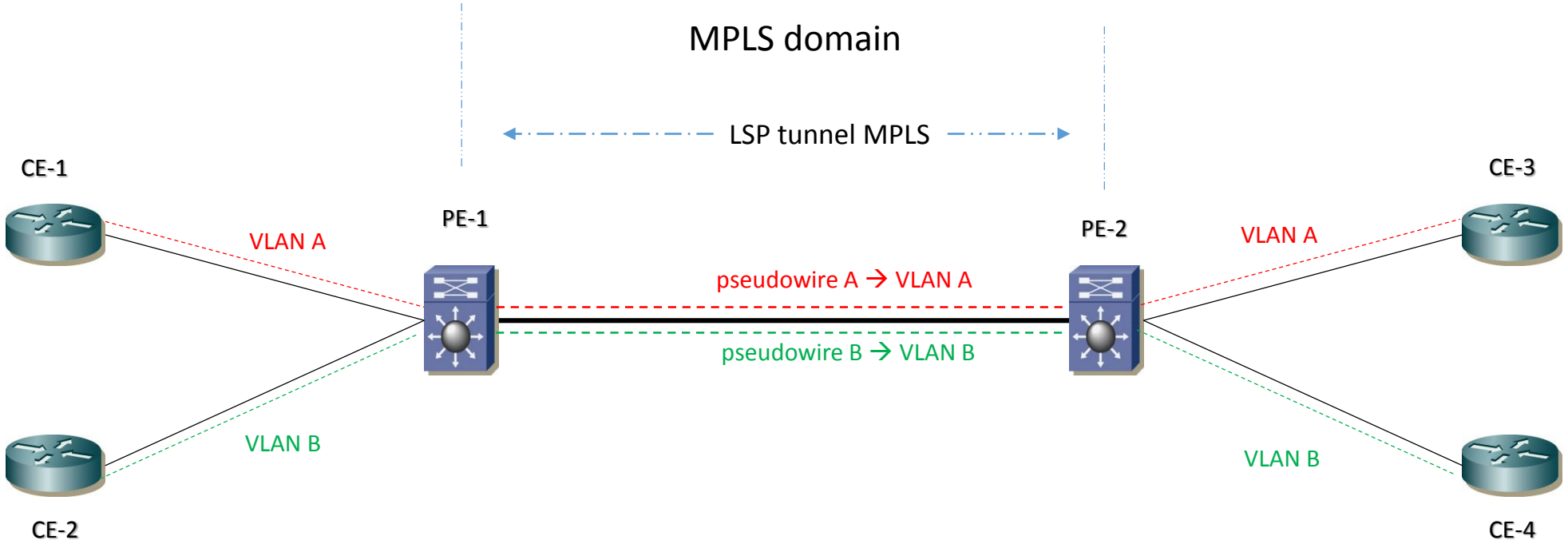
MPLS VPN-L2 application

In VPWS i PE routers imparano solo VLAN se il VC-type è VLAN (se il VC-type è ethernet, i PE routers non apprendono le informazioni VLAN)

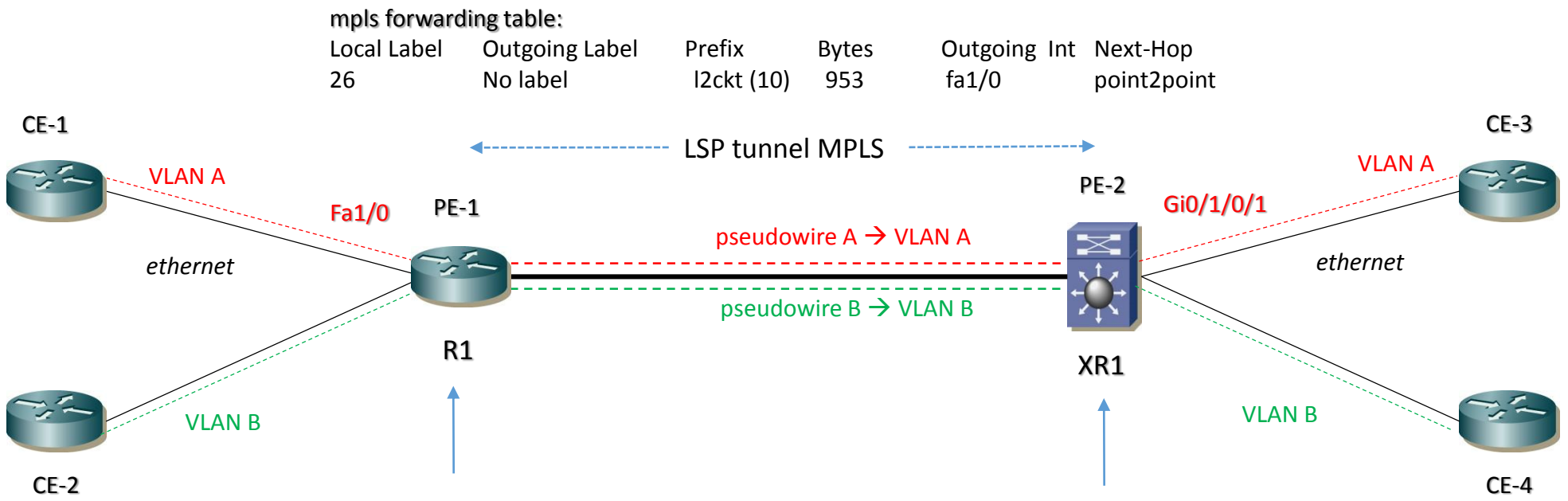
Per VPWS esiste un solo punto di uscita presso il router PE egress, quindi questo significa che i PE MPLS non hanno necessità di mantenere una tabella di MAC addresses per la costruzione del PW.

Viceversa per VPLS, ricordiamo che il PE emula un dominio di bridge verso il customer CE, la rilegatura tra MAC addresses e circuito pesudowire è necessaria (essendo VPLS multipoint-to-multipoint, la destinazione di una frame può essere qualsiasi PE egress appartenente al domino VPLS)

MPLS VPN-L2 application pseudowire design



MPLS VPN-L2 application pseudowire configuration example




```
psuedowire-class ETH-to-ETH
 encapsulation mpls
 interface fastethernet 1/0
!
int fa1/0
 xconnect 2.2.2.2  10 pw-class ETH-to-ETH
```

VC ID value

```
gigabitethernet0/1/0/1 l2 transport
!
l2vpn
 pw-class ETH-to-ETH
 encapsulation mpls
!
 xconnect group ETH
 p2p GI0/1/0/1
 interface  gi0/1/0/1
  neighbor 1.1.1.1 pw-id 10
  pw-class ETH-to-ETH
```

MPLS VPN-L2 VPLS VCT (Virtual Circuit Table)

Un VCT (Virtual Circuit Table) contiene le seguenti informazioni:

- RD Route Distinguisher: analogo alle VPN-L3 indica un indirizzo univoco proprio della VPN (formato AS:NN)
- VE-ID: un valore intero che identifica in modo univoco il site all'interno della VPN (il numero della site)
- Numero di site membri della VPN
- Label base: è il valore di partenza del blocco di inner-label assegnate dal PE router alla VPN in modo dinamico
- Route Target: analogo alle VPN-L3 è usato per politiche di controllo e filtering di annunci via MP-BGP da parte dei PE routers MPLS

Le VCT sono scambiate tra i PE routers attraverso il protocollo MP-BGP (RFC 2547); attraverso lo scambio della VCT ogni PE router è in grado di realizzare l'auto-discovery degli altri membri

MPLS VPN-L2 VPLS VFT (VPLS Forwarding DataBase)

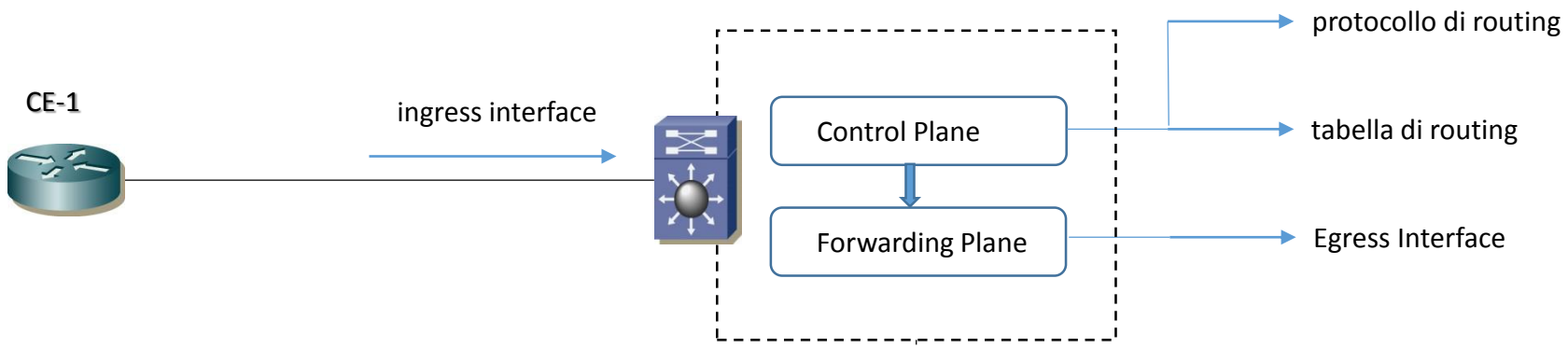
Il FBD (Forwarding DataBase) è contenuto nell'ambito della VFT (VPLS Forwarding DataBase); oltre alle interfacce fisiche come punti di uscita (egress), sono presenti un certo numero di virtual-ports ognuna delle quali emula una porta ethernet di collegamento verso tutti i PE routers membri del dominio VPLS.

Queste porte partecipano a tutti i meccanismi di learning, forwarding e flooding attivi sulle normali porte degli switch ethernet

Le entry relative ad una porta logica contengono ulteriori informazioni rispetto a quelle di una porta fisica ed in particolare:

- VE-ID: è la lista di identificativi dei site remoti
- La inner label da associare alla frame, che nel PE remoto consente a sua volta di associare la frame in ingresso alla specifica VPN-L2 ed al PE router di partenza
    - INNER RX: è il valore della inner label che il PE si aspetta di ricevere per tutte le frame provenienti dai site remoti; questa label è usata per demultiplexare i flussi che arrivano ad un egress PE router; per ogni site remoto è assegnata un valore di label con numero progressivo a partire dal numero base (l'inner RX è propagata via MP-BGP agli altri PE routers)
    - INNER TX: è il valore della inner label che le frame hanno in direzione di altri PE routers (i valori di inner TX per ogni site remoto sono apprese da un PE attraverso il protocollo MP-BGP)
- La outer label che individua il tunnel LSP di egress verso il PE router remoto (appresa via LDP oppure RSVP)

MPLS VPN-L2 VPLS VCT + VFT + FDB design



VCT Circuit Table	Details
RD	AS:NN
VE ID	Site_Name
N° di Site	Site_Member-VPN
Label base	Se stesso
RT	VPN Name

VFT Forwarding	Details
VE ID	Site_Name
Label INNER RX	valore con cui il PE identifica il Site remoto
Label INNER TX	Valore con cui il PE trasmette per indentificare il site di origine
Outer Label	Identifica il tunnel LSP

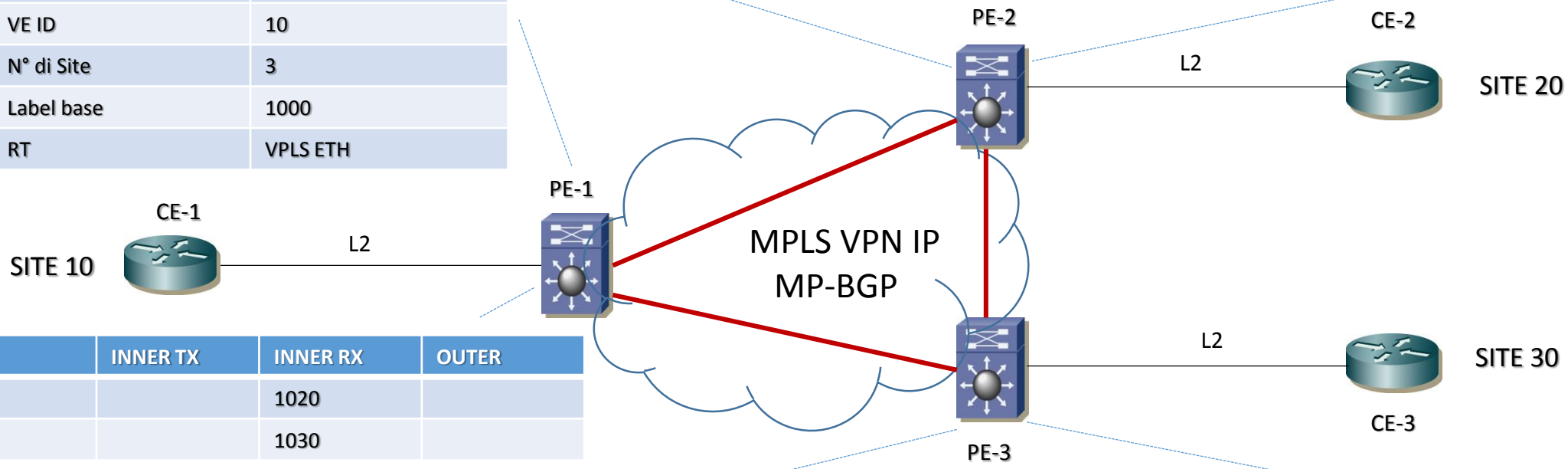
FDB Forward DB	Details
INNER Label	Associazione packet VPN
OUTER Label	LSP di egress verso il PE remoto

MPLS VPN-L2 VPLS fase 0 design

VCT Circuit Table	Details
RD	AS:NN
VE ID	10
N° di Site	3
Label base	1000
RT	VPLS ETH

VCT Circuit Table	Details
RD	AS:NN
VE ID	20
N° di Site	3
Label base	2000
RT	VPLS ETH

VCT Circuit Table	Details
RD	AS:NN
VE ID	30
N° di Site	3
Label base	3000
RT	VPLS ETH



PE1: VFT vpls forwarding Table

VE ID	INNER TX	INNER RX	OUTER
20		1020	
30		1030	

PE2: VFT vpls forwarding Table

VE ID	INNER TX	INNER RX	OUTER
10		2010	
30		2030	

PE3: VFT vpls forwarding Table

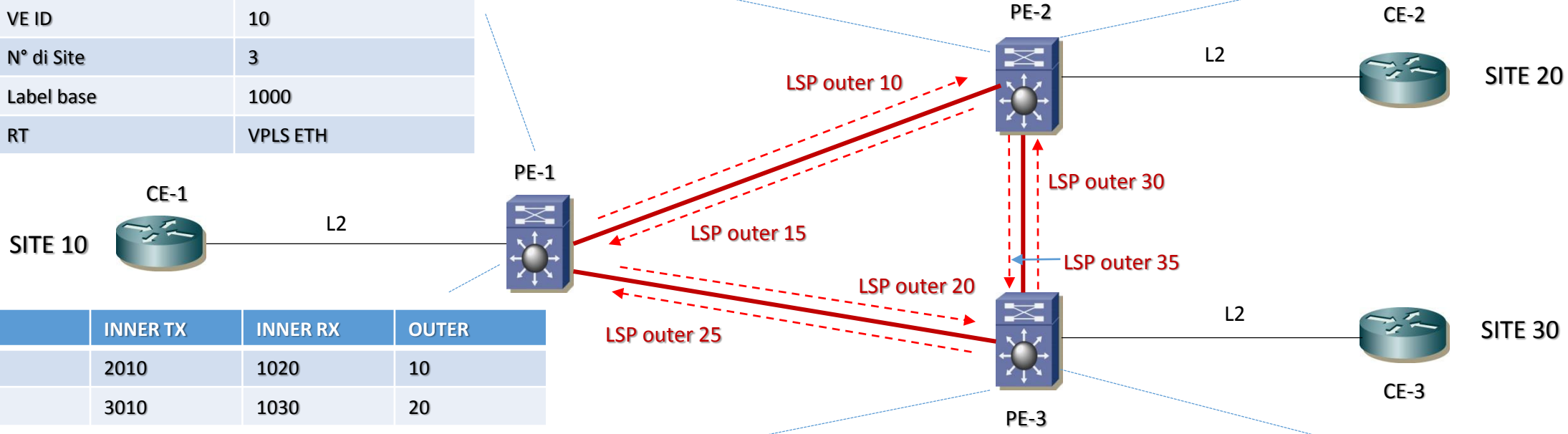
VE ID	INNER TX	INNER RX	OUTER
10		3010	
20		3020	

MPLS VPN-L2 VPLS fase 1 design

VCT Circuit Table	Details
RD	AS:NN
VE ID	10
N° di Site	3
Label base	1000
RT	VPLS ETH

VCT Circuit Table	Details
RD	AS:NN
VE ID	20
N° di Site	3
Label base	2000
RT	VPLS ETH

VCT Circuit Table	Details
RD	AS:NN
VE ID	30
N° di Site	3
Label base	3000
RT	VPLS ETH



VE ID	INNER TX	INNER RX	OUTER
20	2010	1020	10
30	3010	1030	20

PE1: VFT vpls forwarding Table

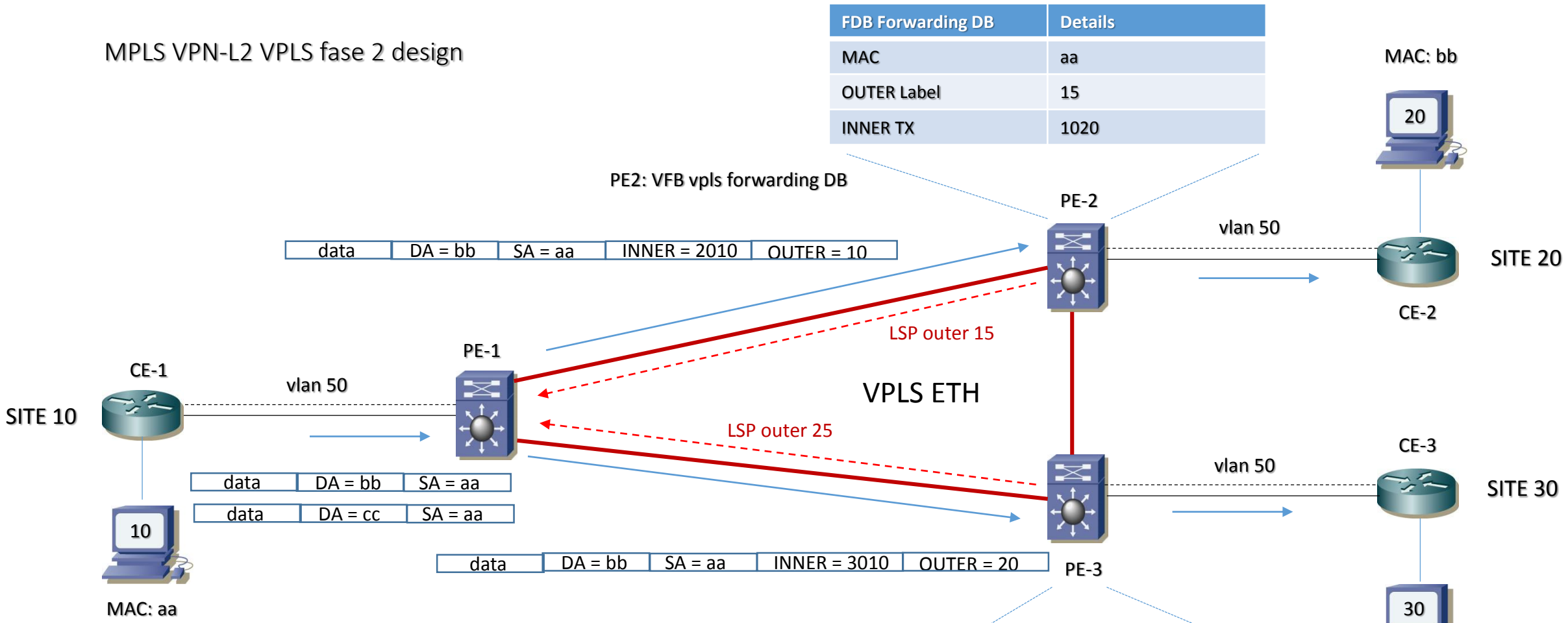
PE2: VFT vpls forwarding Table

VE ID	INNER TX	INNER RX	OUTER
10	1020	2010	15
30	3020	2030	35

VE ID	INNER TX	INNER RX	OUTER
10	1030	3010	25
20	2030	3020	30

PE3: VFT vpls forwarding Table

MPLS VPN-L2 VPLS fase 2 design



FDB Forwarding DB	Details
MAC	aa
OUTER Label	15
INNER TX	1020

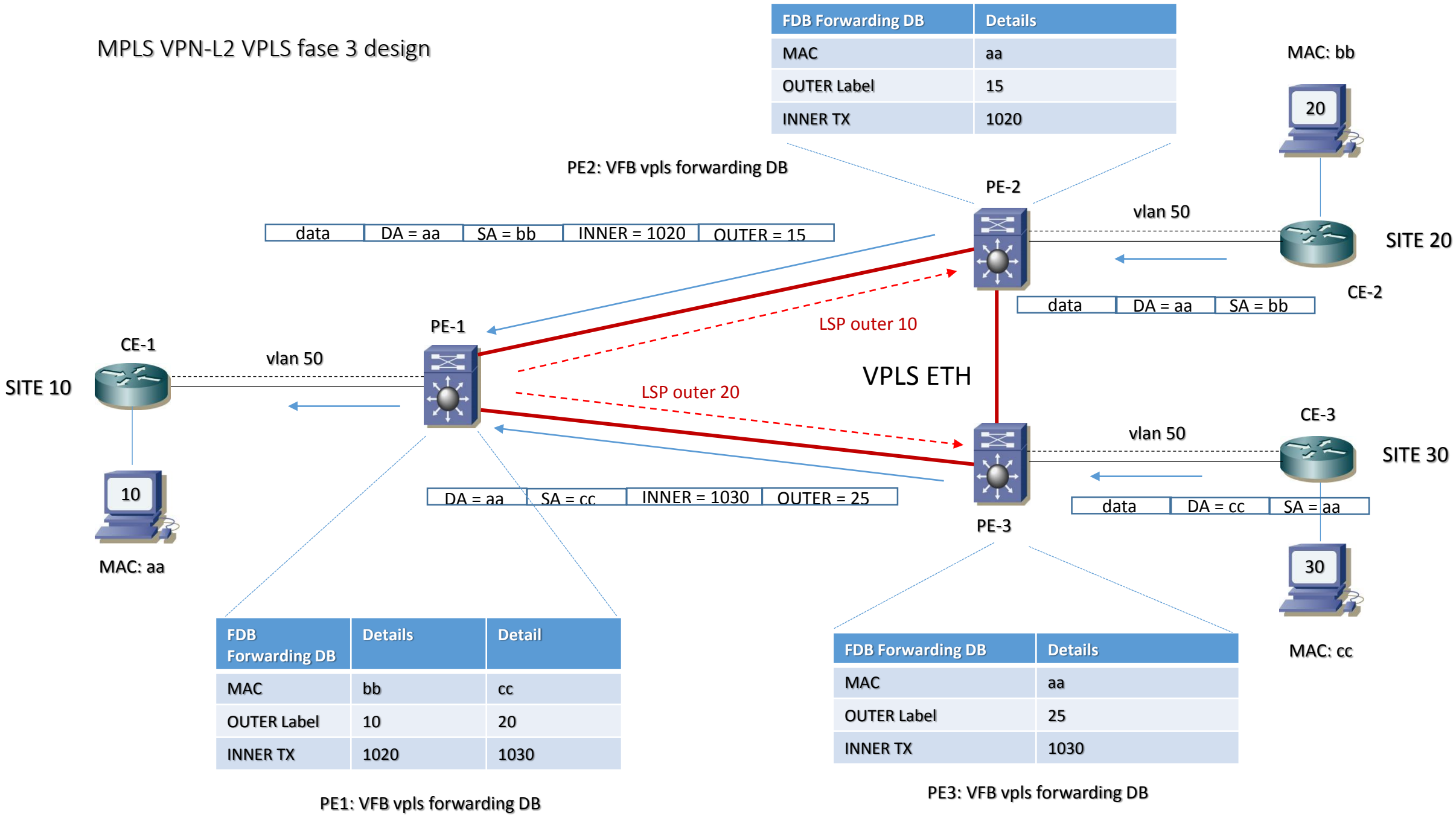
VE ID	INNER TX	INNER RX	OUTER
20	2010	1020	10
30	3010	1030	20

PE1: VFT vpls forwarding Table

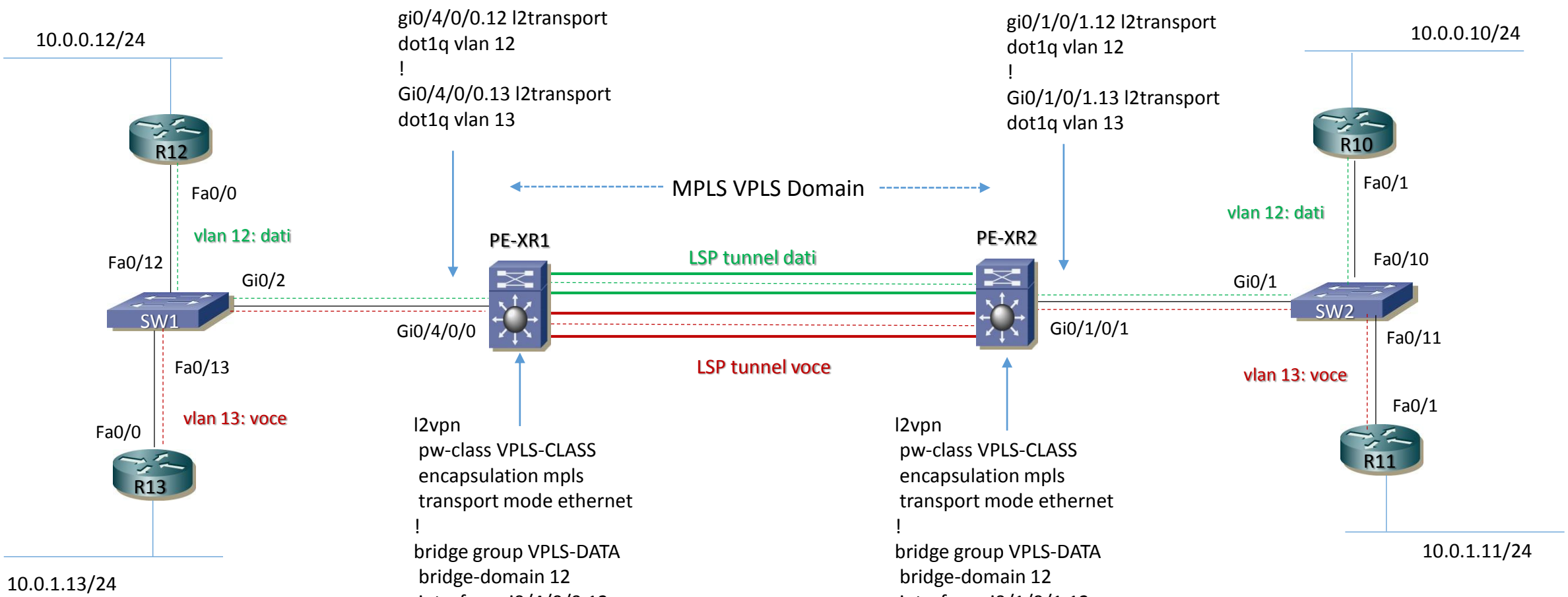
FDB Forwarding DB	Details
MAC	aa
OUTER Label	25
INNER TX	1030

PE3: VFB vpls forwarding DB

MPLS VPN-L2 VPLS fase 3 design



MPLS VPN-L2 VPLS example configuration




```
l2vpn
 pw-class VPLS-CLASS
 encapsulation mpls
 transport mode ethernet
!
bridge group VPLS-DATA
 bridge-domain 12
 interface gi0/4/0/0.12
!
vfi 12
 neighbor 2.2.2.2 pw-id 12
 pw-class VPLS-CLASS
```

The same for vlan 13

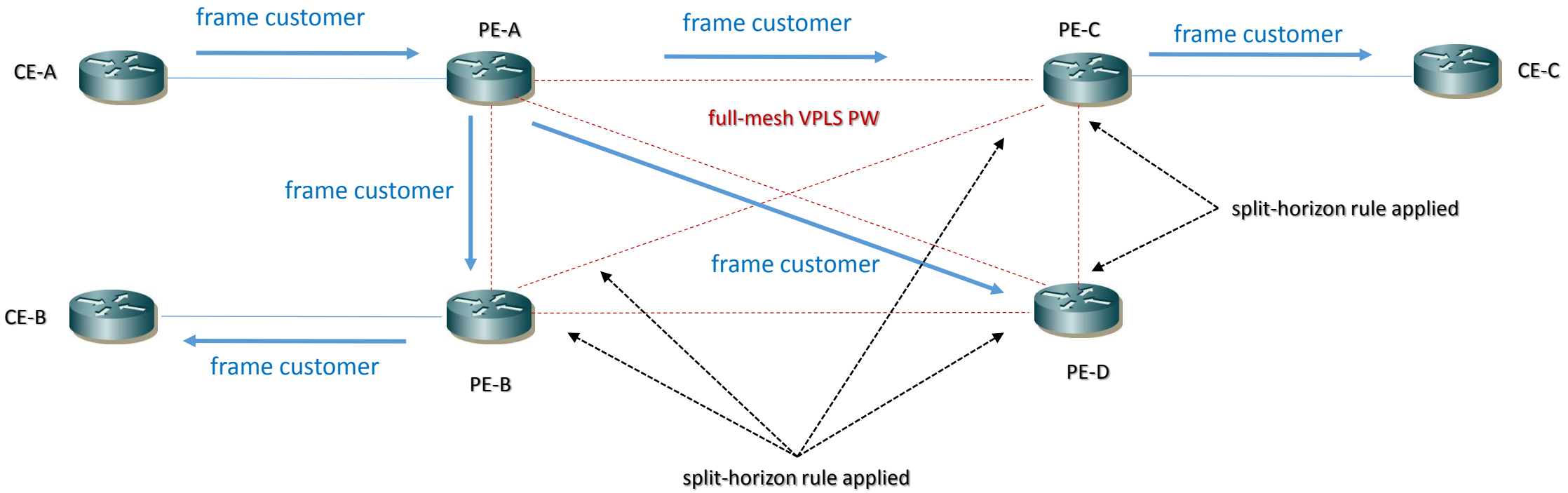
```
l2vpn
 pw-class VPLS-CLASS
 encapsulation mpls
 transport mode ethernet
!
bridge group VPLS-DATA
 bridge-domain 12
 interface gi0/1/0/1.12
!
vfi 12
 neighbor  1.1.1.1 pw-id 12
 pw-class VPLS-CLASS
```

The same for vlan 13

MPLS VPN-L2 VPLS Spanning Tree Avoidance

Non c'è Spanning Tree Protocol all'interno del dominio Core in un Services Provider MPLS per ragioni di loop avoidance in VPLS

La regola split-horizon nel core SP è abilitata di default e nessuna configurazione è necessaria, quindi se una frame customer è ricevuta da un circuito pseudowire, questa non è trasmessa dietro via altro PW

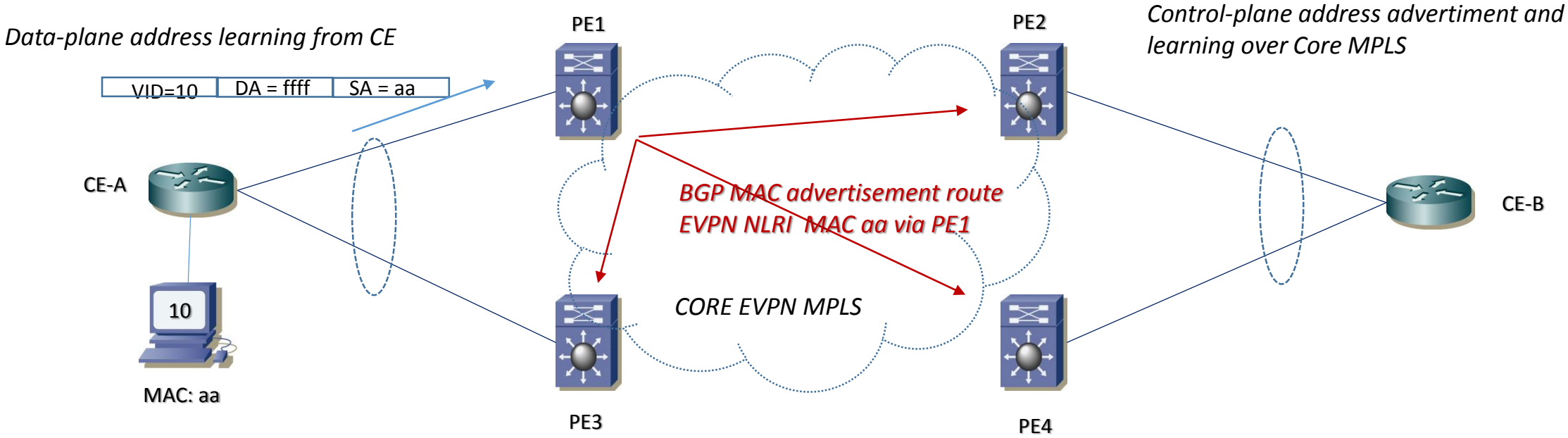


MPLS VPN-L2 VPLS load-balancing EVPN

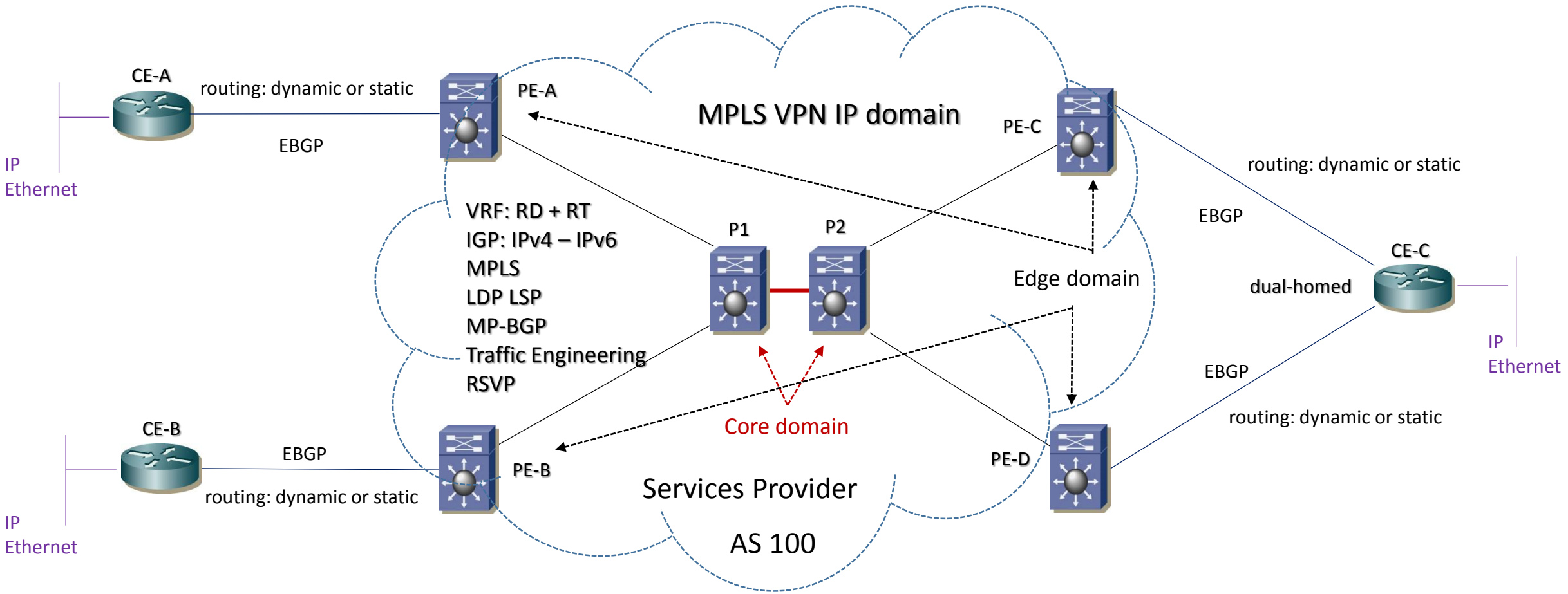
CE customer collegati in dual-homed allo stesso o differente VPLS-PE router del Services Provider, possono utilizzare i due link in modalità active-standby per tutte le vlans oppure utilizzare un vlan-based load-balancing (50% di vlan su un link e l'altro 50% di vlan sull'altro).

EVPN (Ethernet VPN) è la nuova generazione per MPLS ethernet based services (VPLS) può supportare invece active-active flow-based load-balancing e le vlans possono essere utilizzate in entrambi i links attivamente; questo significa anche fast-convergence customer links, PE links ed node failure scenario.

In VPLS la tabella di learning MAC addresses è imparata via data-plane (piano di forwarding) e segnalata attraverso MP-BGP control-plane; in EVPN non esiste un data-plane MAC learning attraverso il Core Network, ma il MAC addresses è appreso direttamente dal circuito layer 2 direttamente connesso ai due end-point routers, sempre via data-plane



MPLS VPN-L3 Traditional Network



UNIFIED MPLS VPN-L3 (RFC 3107) carry label information in BGPv4 via Core and Aggregation

Unified MPLS utilizza il protocollo BGPv4 per scambiare informazioni con labels; quando BGP distribuisce una prefix/route porta con se anche una label MPLS che mappa la prefix/route.

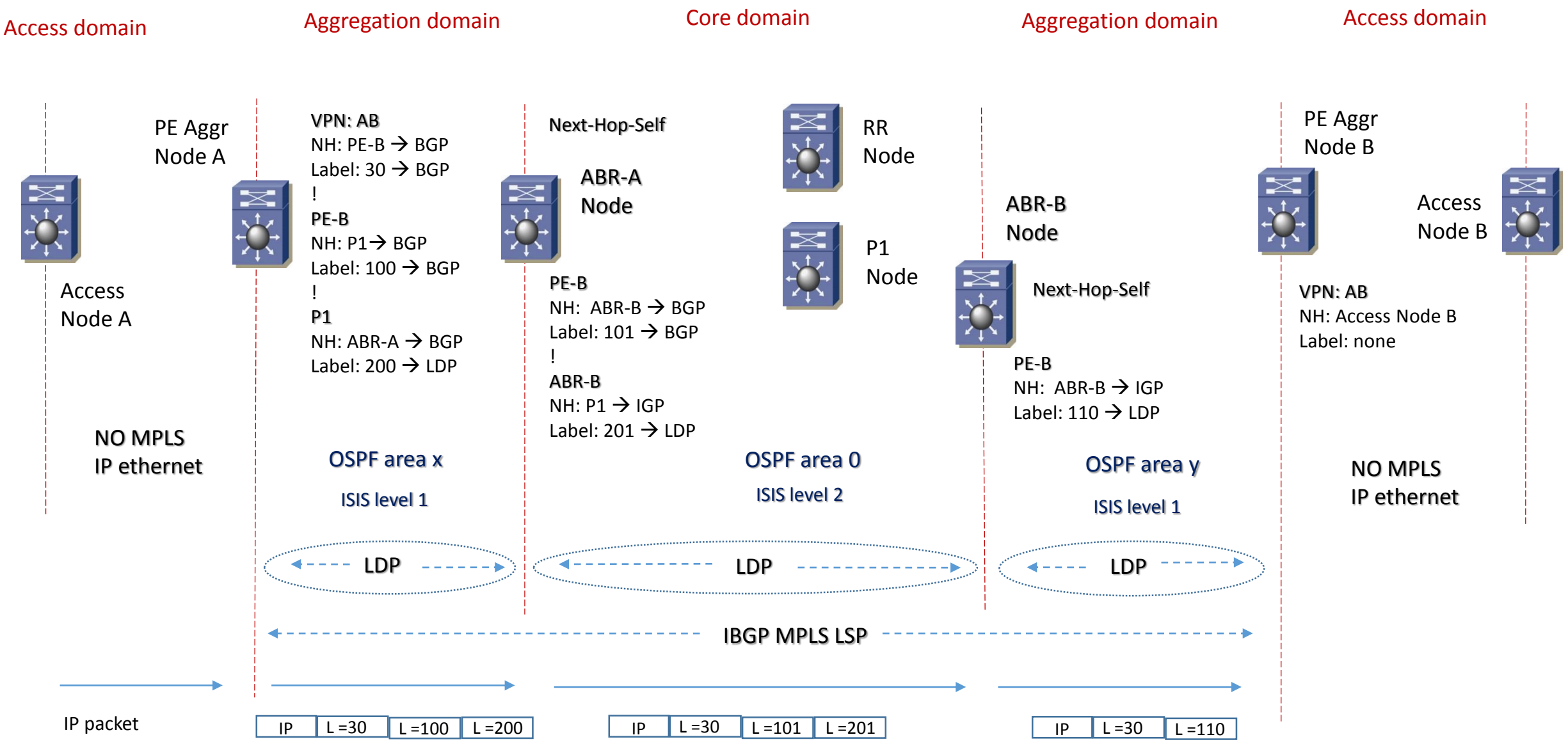
Questo mapping «label-prefix» è trasportato negli update message di BGP che contengono le informazioni della route.

Se il next-hop non cambia la label viene preservata; viceversa la label cambia valore se il next-hop cambia; in unified MPLS il next-hop cambia a livello ABR router

Le caratteristiche di un unified MPLS RFC 3107 enable sono:

- Differenti domini IGP di routing tra il Core Aggregation ed Access Networks; questo si ottiene con la divisione in aree per OSPF ed in differenti level per ISIS (tre domini IGP);
- Tre segmenti LDP/LSP (label distribution protocol / label switch path), ciascuno per i tre domini IGP creati;
- Il layer Access Network non ha LDP/LSP enable;
- Un singolo IBGP domain con path MPLS end-to-end LSP tra i livelli di Aggregation Node Routers;
- Un singolo AS (Autonomous System) IBGP sessions
- I Nodi ABRs che uniscono i segmenti di rete debbono essere inline Route Reflector con Next-Hop Self enable in ordine di trasportare una route IPv4 + label configurata

UNIFIED MPLS VPN-L3 (RFC 3107) carry label information in BGPv4 via Core and Aggregation



UNIFIED MPLS VPN-L3 (RFC 3107) carry label information in BGPv4 via Core, Aggregation and Access

Unified MPLS utilizza il protocollo BGPv4 per scambiare informazioni con labels; quando BGP distribuisce una prefix/route porta con se anche una label MPLS che mappa la prefix/route.

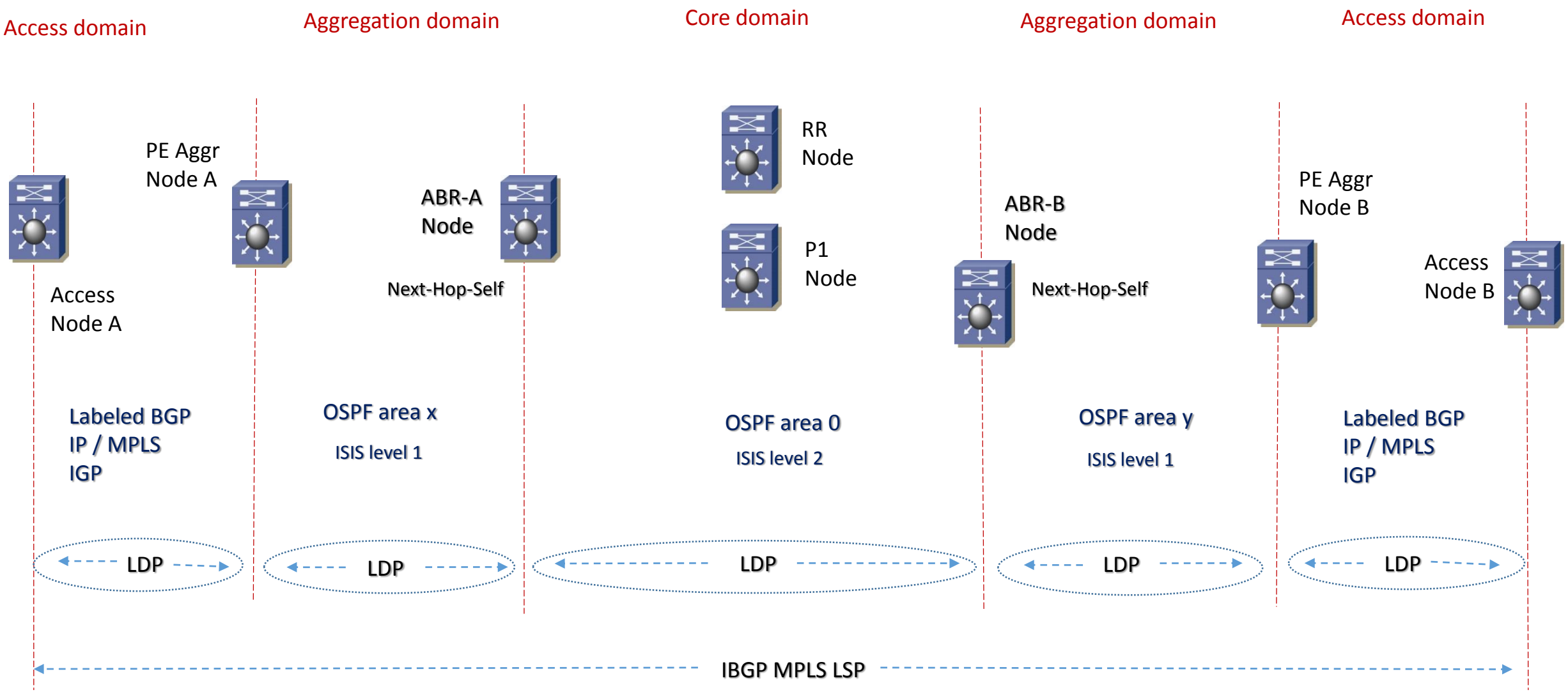
A differenza del modello precedente (labeled Core Aggregation, questo modello prevede MPLS enable anche a livello di accesso, pertanto il BGP labeled è esteso sino all'accesso del dominio

Resta distinto il protocollo IGP in differenti domini di routing tra Core, Aggregation ed Access.

Un singolo IBGP domain con path MPLS end-to-end LSP tra i livelli Access Node Routers;

La segmentazione tra i domini Access, Aggregation e Core può basarsi su un singolo AS multi-area design oppure su Inter-AS

UNIFIED MPLS VPN-L3 (RFC 3107) carry label information in BGPv4 via Core, Aggregation and Access



UNIFIED MPLS VPN-L3 (RFC 3107) carry label information in BGPv4 via Core and Aggregation with IGP redistribution on Access

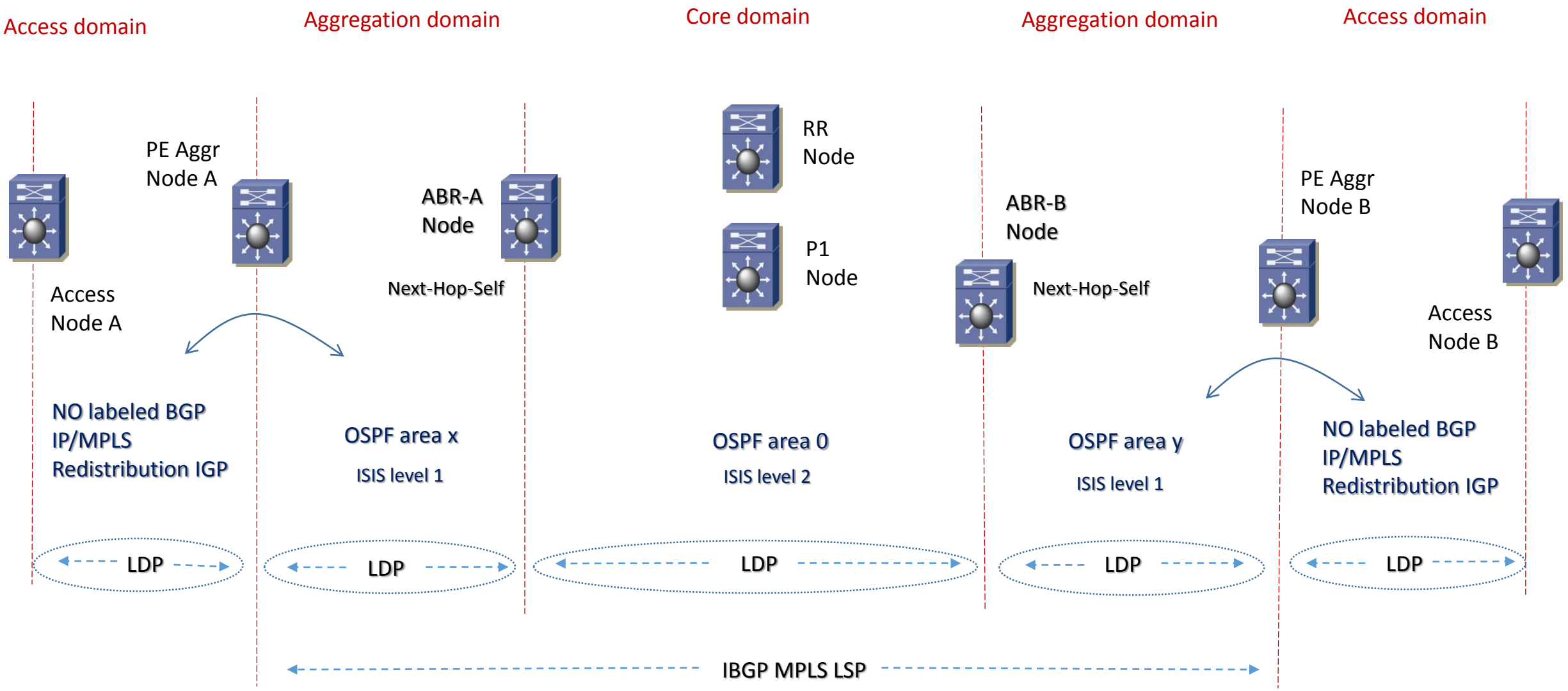
Unified MPLS utilizza il protocollo BGPv4 per scambiare informazioni con labels; quando BGP distribuisce una prefix/route porta con se anche una label MPLS che mappa la prefix/route.

In questo modello il livello Access lavora con IP/MPLS ma non con labeled BGP; un processo di Redistribution è performato a livello di Aggregation Node.

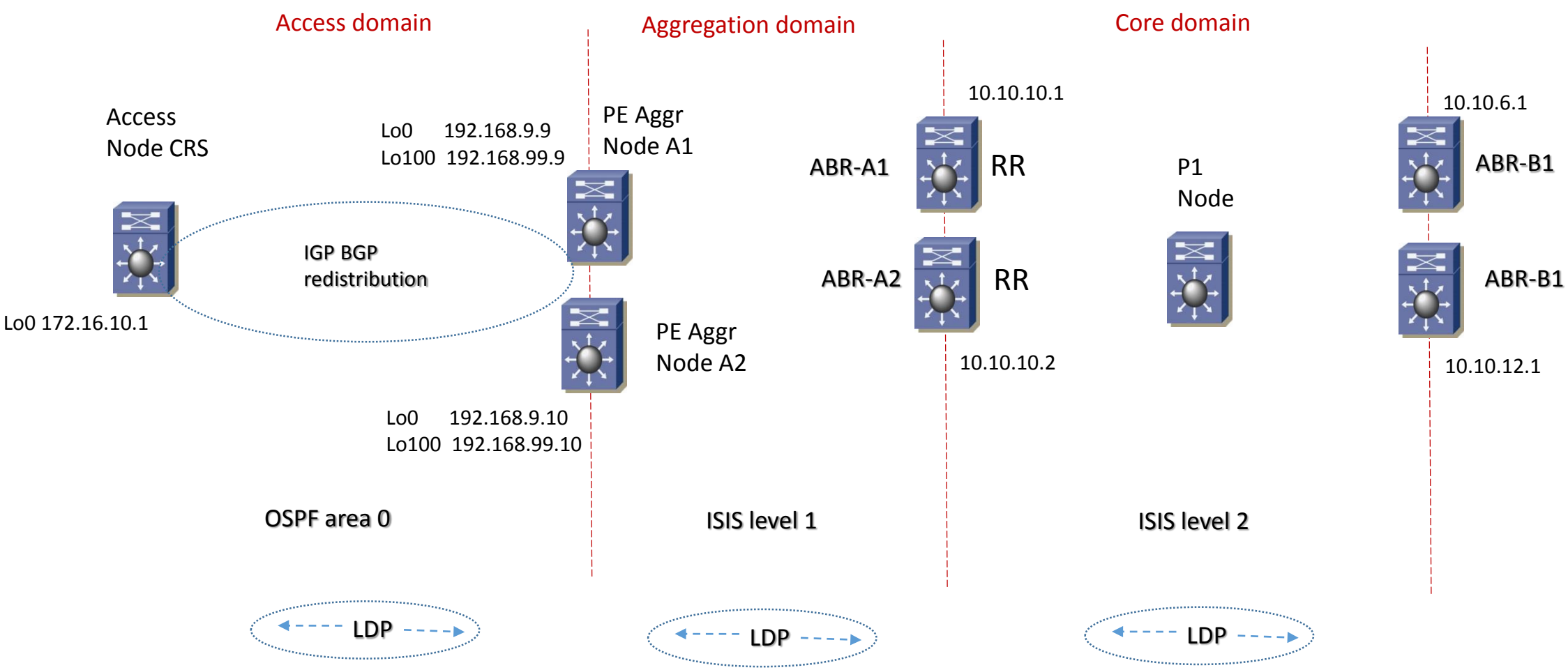
L'indirizzo di loopback del Access Node è redistribuito dentro IBGP domain.

Il dominio Core Aggregation è esteso a livello di Access con la redistribuzione dell'Access IGP dentro il dominio IBGP e la redistribuzione di tutte le necessarie labeled IBGP prefix dentro il dominio Access IGP (via BGP comminity)

UNIFIED MPLS VPN-L3 (RFC 3107) carry label information in BGPv4 via Core and Aggregation with IGP redistribution on Access

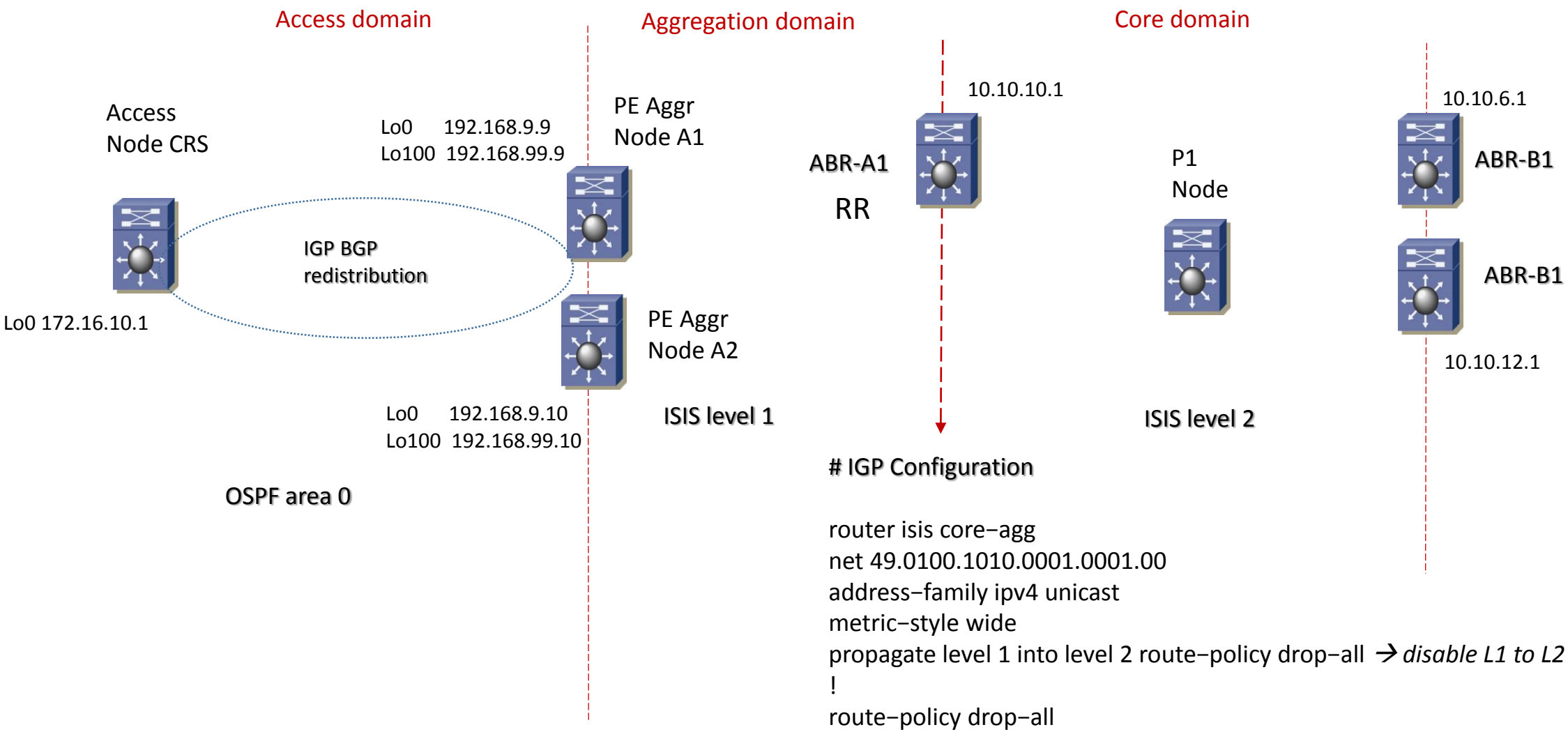


UNIFIED MPLS VPN-L3 (RFC 3107) Core and Aggregation with IGP redistribution on Access Configuration Example



200:200 ABC community
300:300 Aggregation community

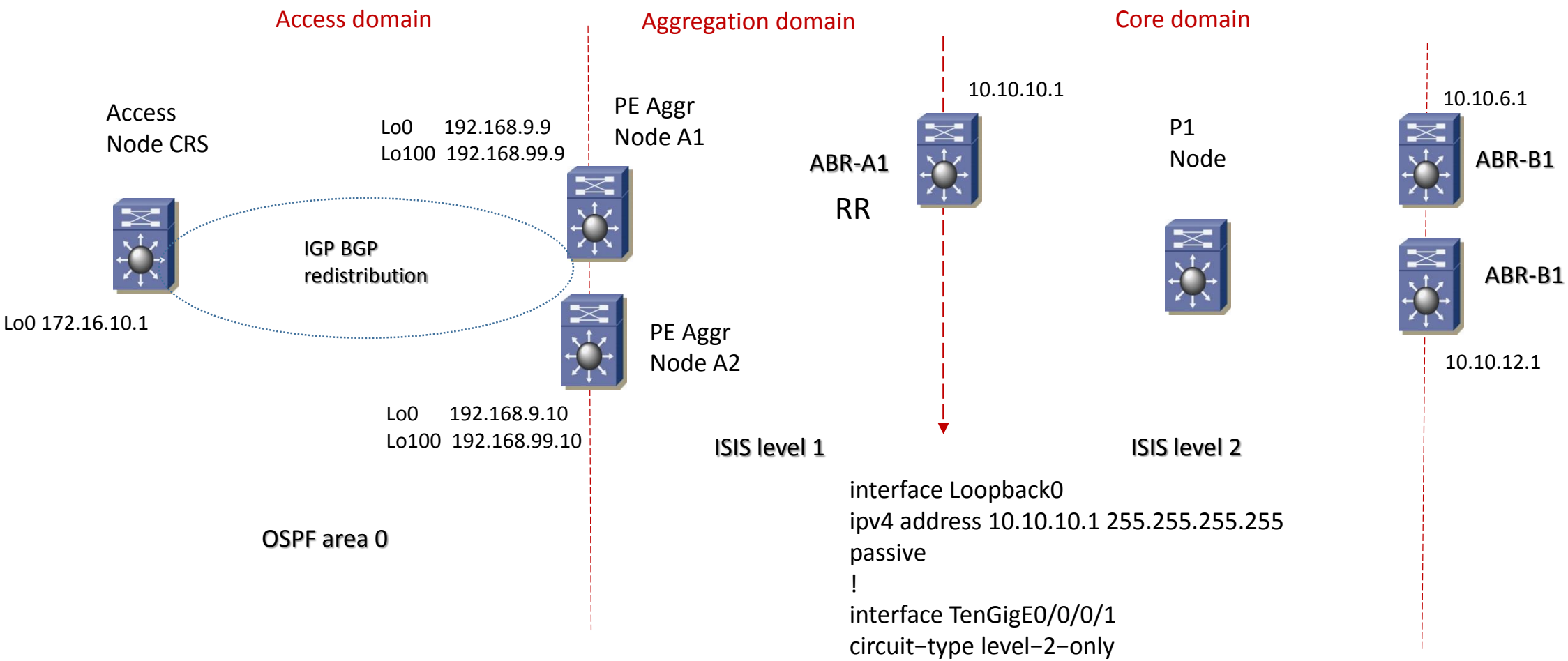
UNIFIED MPLS VPN-L3 (RFC 3107) Core and Aggregation with IGP redistribution on Access Configuration Example



IGP Configuration

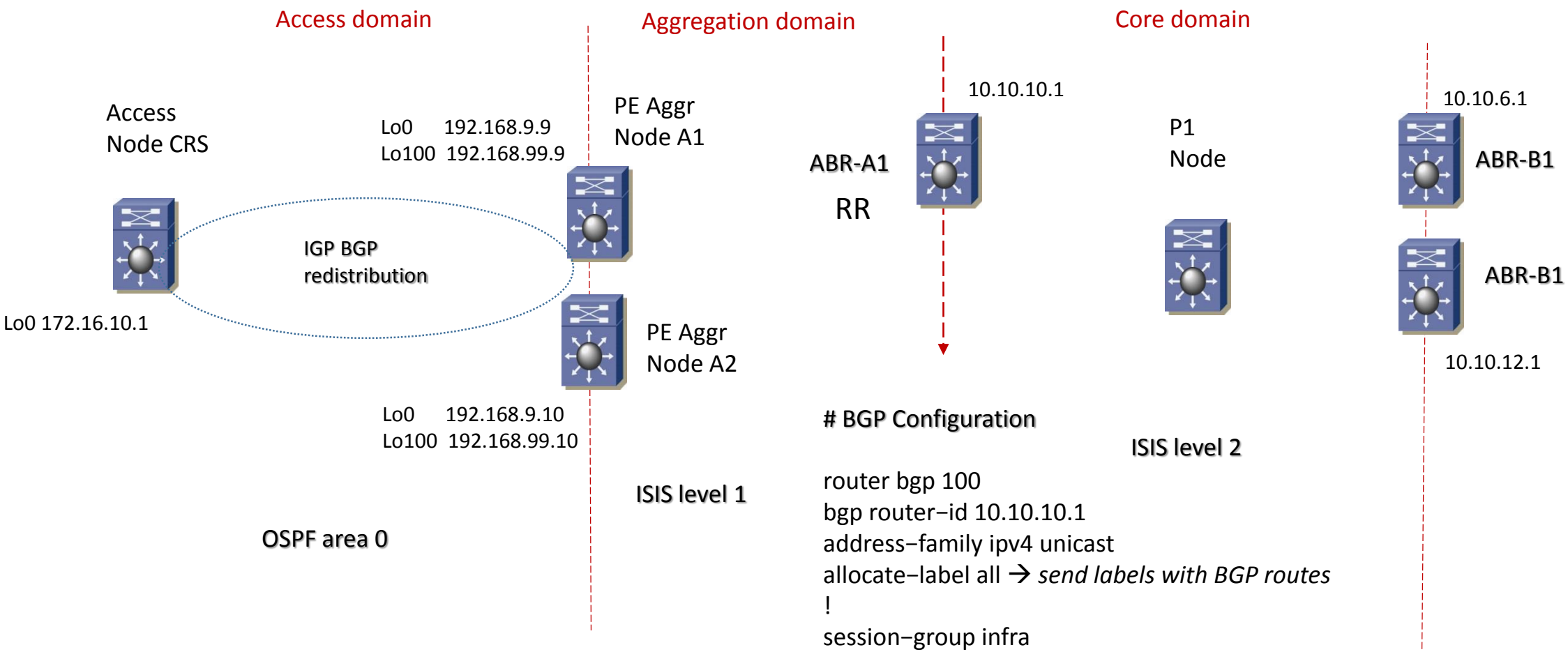
```
router isis core-agg
net 49.0100.1010.0001.0001.00
address-family ipv4 unicast
metric-style wide
propagate level 1 into level 2 route-policy drop-all → disable L1 to L2 redistribution
!
route-policy drop-all
drop
end-policy
```

UNIFIED MPLS VPN-L3 (RFC 3107) Core and Aggregation with IGP redistribution on Access Configuration Example




```
interface Loopback0
ipv4 address 10.10.10.1 255.255.255.255
passive
!
interface TenGigE0/0/0/1
circuit-type level-2-only
!
interface TenGigE0/0/0/2
circuit-type level-1
```

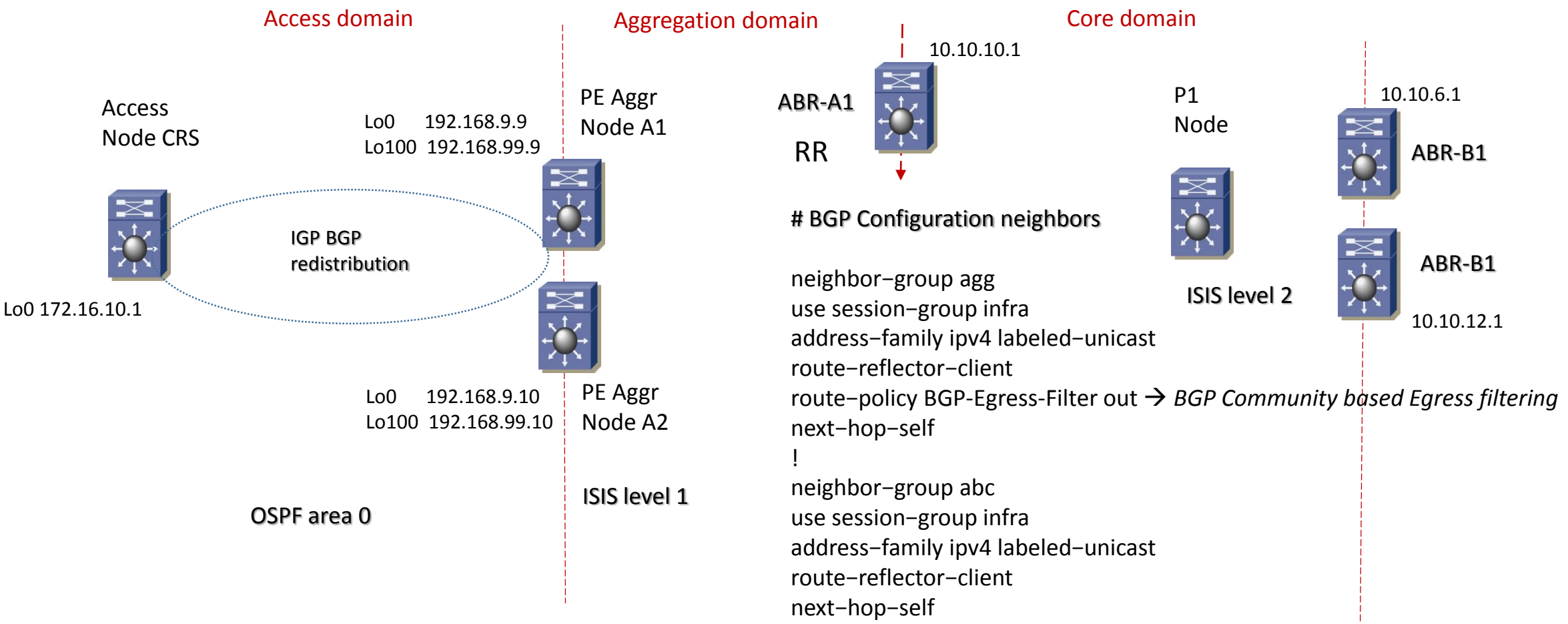
UNIFIED MPLS VPN-L3 (RFC 3107) Core and Aggregation with IGP redistribution on Access Configuration Example



 # BGP Configuration

router bgp 100
bgp router−id 10.10.10.1
address−family ipv4 unicast
allocate−label all → *send labels with BGP routes*
!
session−group infra
remote−as 100
cluster−id 1001
update−source Loopback0

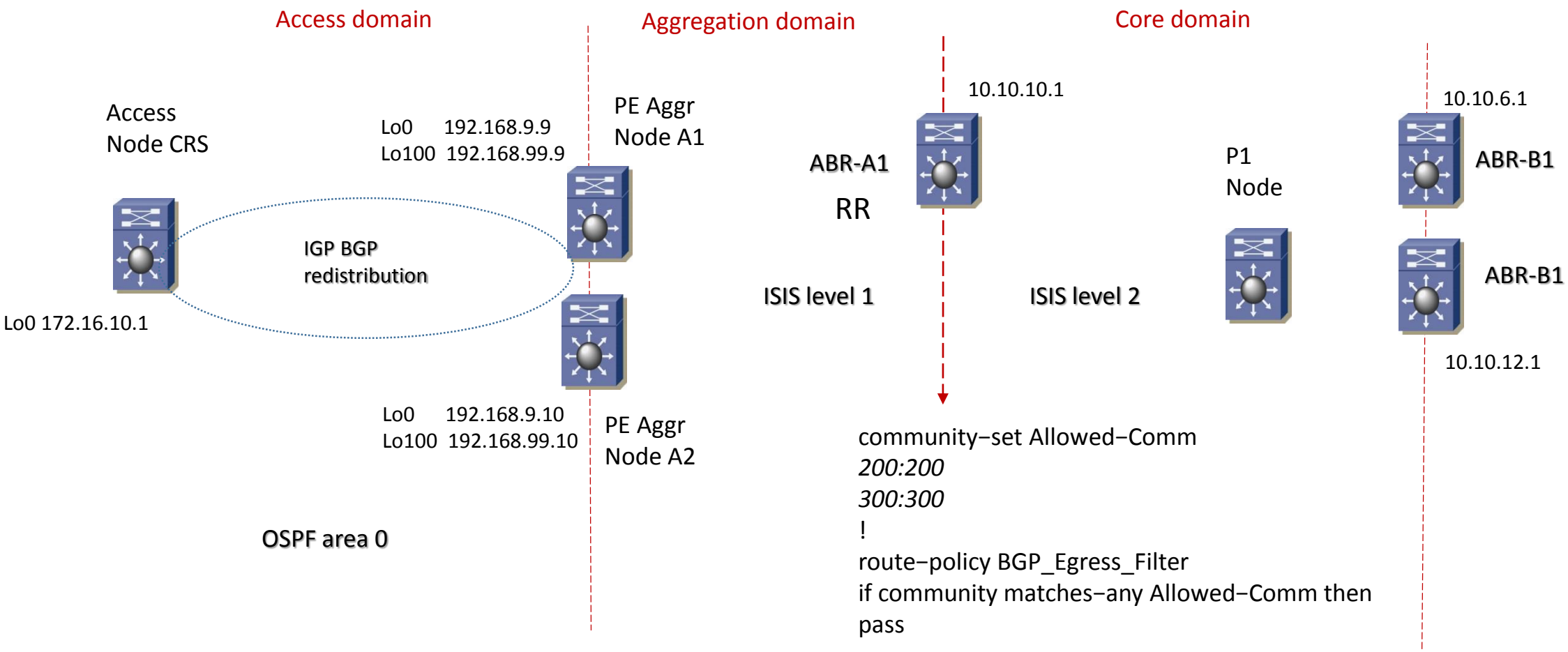
UNIFIED MPLS VPN-L3 (RFC 3107) Core and Aggregation with IGP redistribution on Access Configuration Example



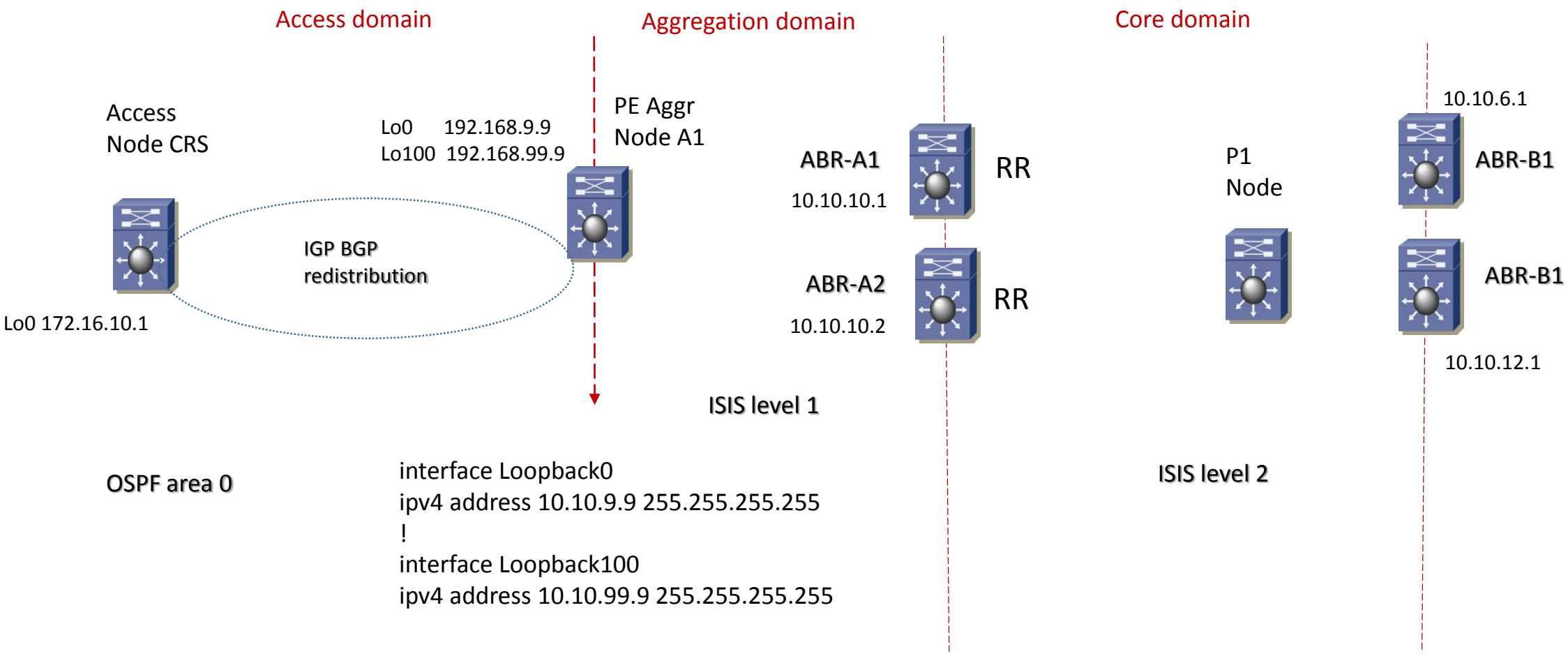

```
# BGP Configuration neighbors

neighbor-group agg
use session-group infra
address-family ipv4 labeled-unicast
route-reflector-client
route-policy BGP-Egress-Filter out → BGP Community based Egress filtering
next-hop-self
!
neighbor-group abc
use session-group infra
address-family ipv4 labeled-unicast
route-reflector-client
next-hop-self
!
neighbor-group core
use session-group infra
address-family ipv4 labeled-unicast
next-hop-self
```

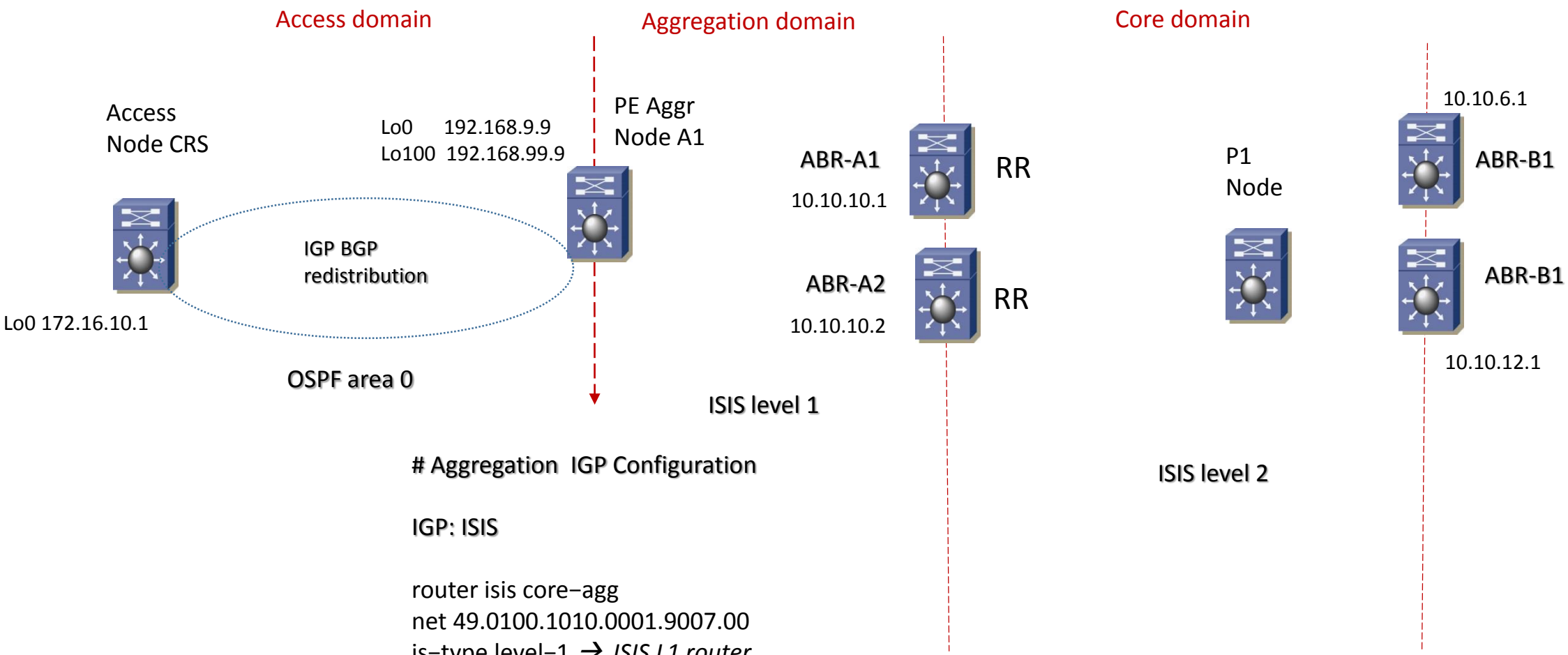
UNIFIED MPLS VPN-L3 (RFC 3107) Core and Aggregation with IGP redistribution on Access Configuration Example



UNIFIED MPLS VPN-L3 (RFC 3107) Core and Aggregation with IGP redistribution on Access Configuration Example



UNIFIED MPLS VPN-L3 (RFC 3107) Core and Aggregation with IGP redistribution on Access Configuration Example

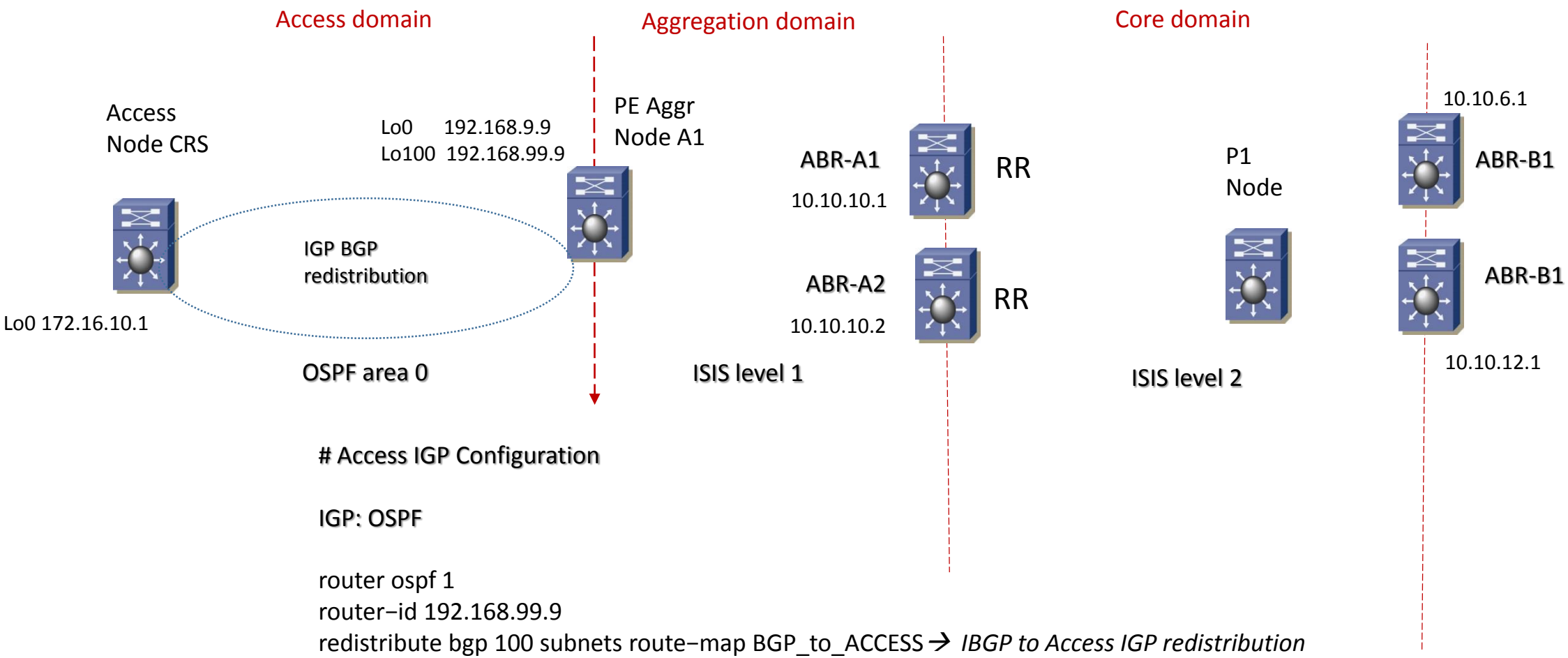


Aggregation IGP Configuration

IGP: ISIS

router isis core−agg
net 49.0100.1010.0001.9007.00
is−type level−1 → *ISIS L1 router*
metric−style wide
passive−interface Loopback0

UNIFIED MPLS VPN-L3 (RFC 3107) Core and Aggregation with IGP redistribution on Access Configuration Example

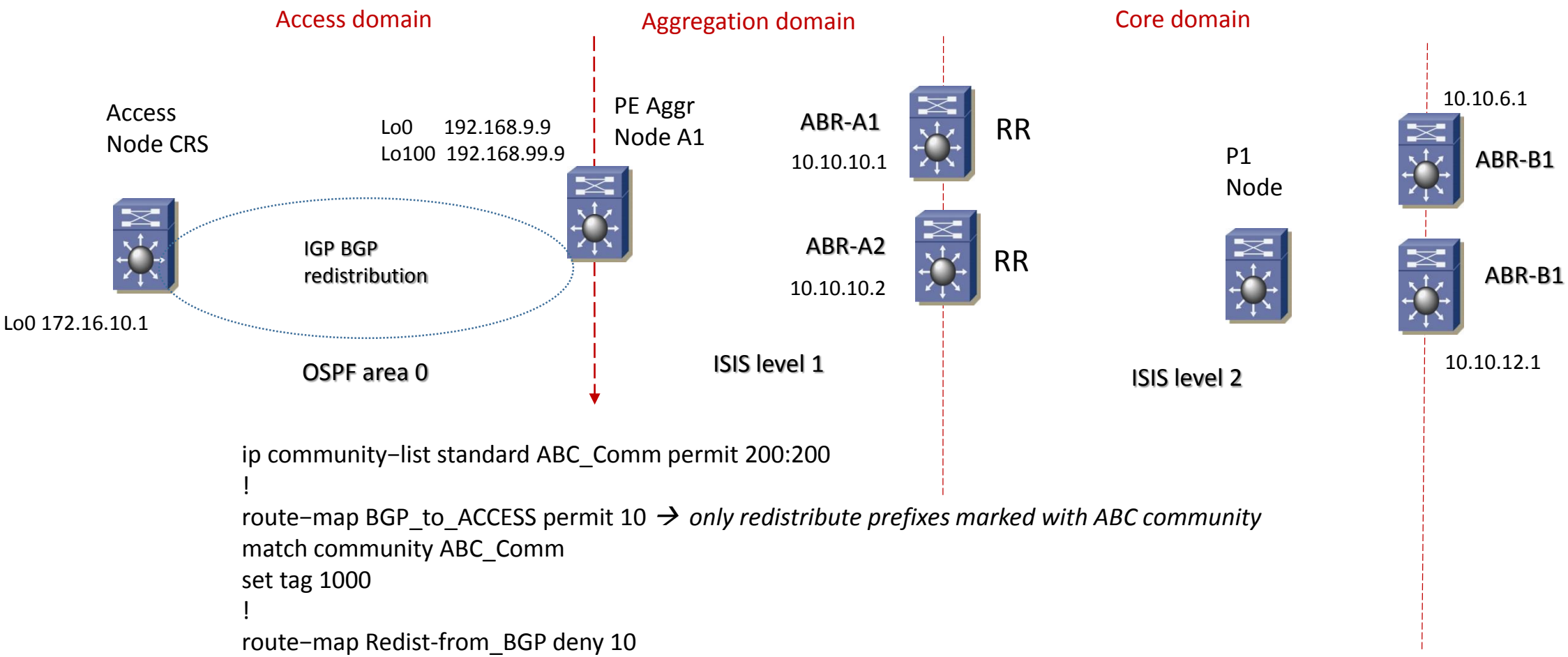


 # Access IGP Configuration

IGP: OSPF

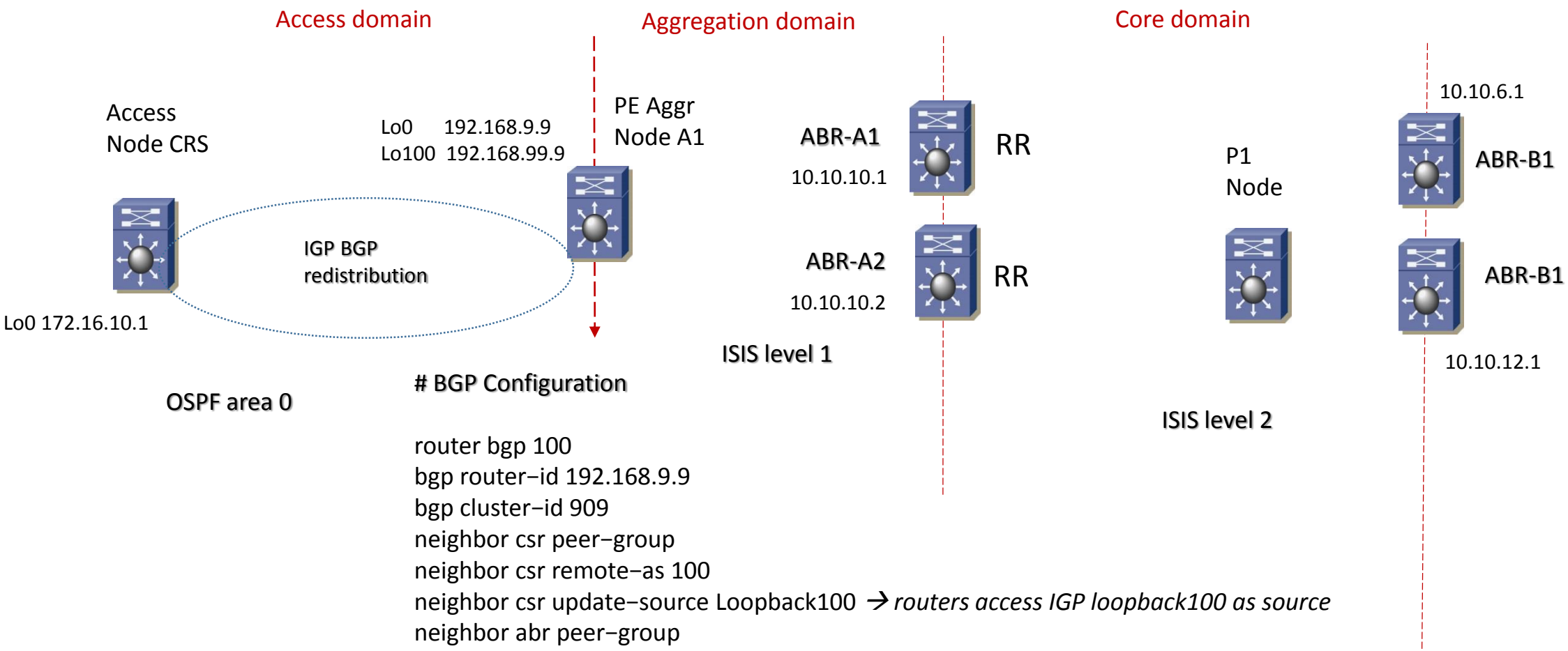
router ospf 1
router−id 192.168.99.9
redistribute bgp 100 subnets route−map BGP_to_ACCESS → *IBGP to Access IGP redistribution*
network 192.168.9.9 0.0.0.1 area 0
network 192.168.99.9 0.0.0.0 area 0
network 10.9.10.0 0.0.0.1 area 0
distribute−list route−map Redist-from_BGP in → *inbound filtering to prefer labeled BGP learnt prefixes*

UNIFIED MPLS VPN-L3 (RFC 3107) Core and Aggregation with IGP redistribution on Access Configuration Example




```
ip community-list standard ABC_Comm permit 200:200
!
route-map BGP_to_ACCESS permit 10 → only redistribute prefixes marked with ABC community
match community ABC_Comm
set tag 1000
!
route-map Redist-from_BGP deny 10
match tag 1000
!
route-map Redist-from_BGP permit 20
```

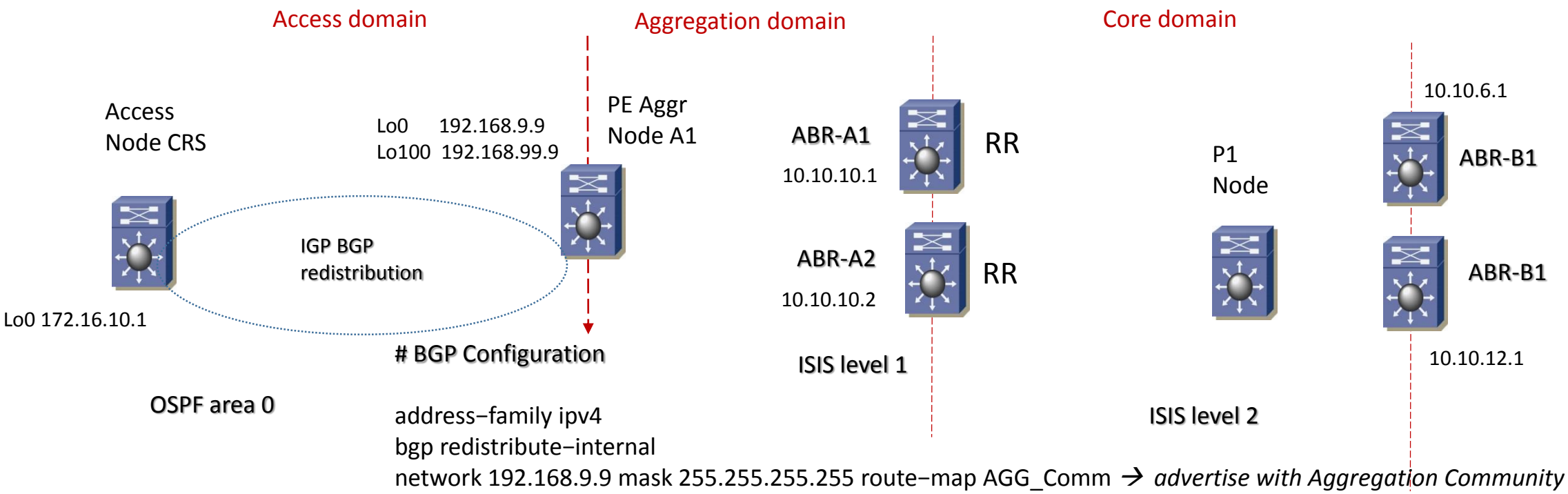
UNIFIED MPLS VPN-L3 (RFC 3107) Core and Aggregation with IGP redistribution on Access Configuration Example



BGP Configuration

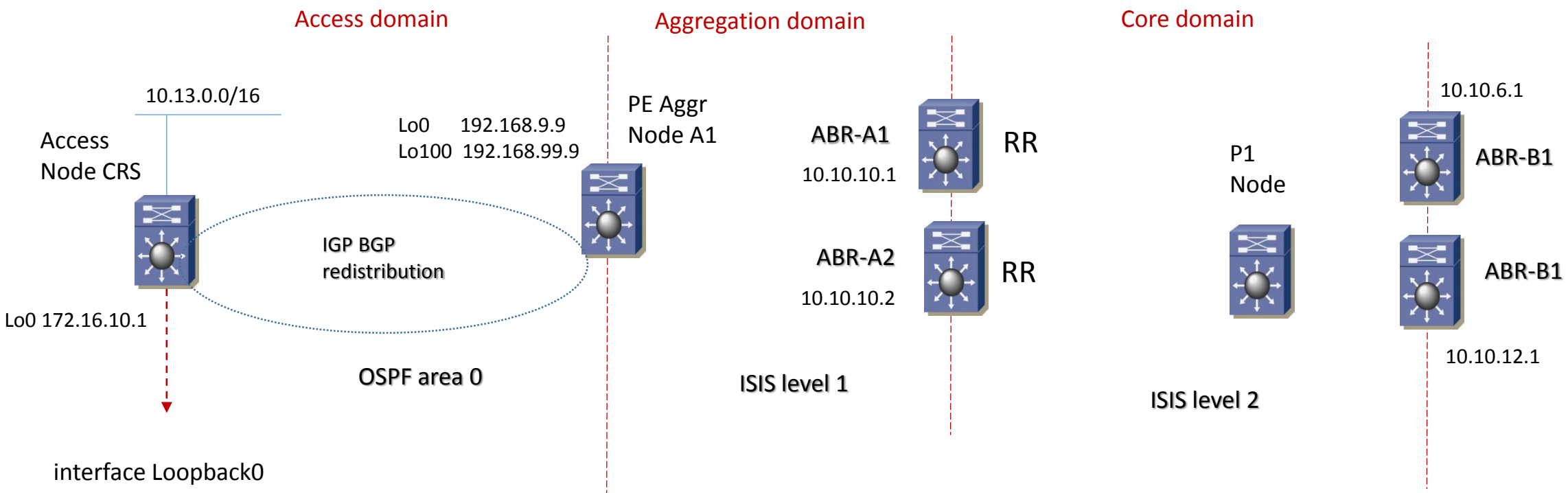
router bgp 100
bgp router−id 192.168.9.9
bgp cluster−id 909
neighbor csr peer−group
neighbor csr remote−as 100
neighbor csr update−source Loopback100 → *routers access IGP loopback100 as source*
neighbor abr peer−group
neighbor abr remote−as 100
neighbor abr update−source Loopback0 → *Core POP ABRs – core−agg IGP loopback0 as source*
neighbor 10.10.10.1 peer−group abr
neighbor 10.10.10.2 peer−group abr
neighbor 172.16.10.1 peer-group crs

UNIFIED MPLS VPN-L3 (RFC 3107) Core and Aggregation with IGP redistribution on Access Configuration Example



address−family ipv4
bgp redistribute−internal
network 192.168.9.9 mask 255.255.255.255 route−map AGG_Comm → *advertise with Aggregation Community*
redistribute ospf 1 → *redistribute Access IGP prefixes*
neighbor abr send−community
neighbor abr next−hop−self
neighbor abr send-label → *send labels with BGP routes*
neighbor 10.10.10.1 activate
neighbor 10.10.10.2 activate
exit−address−family
!
route−map AGG_Comm permit 10
set community 300:300

UNIFIED MPLS VPN-L3 (RFC 3107) Core and Aggregation with IGP redistribution on Access Configuration Example



interface Loopback0
ip address 172.16.10.1 255.255.255.255

IGP Configuration

router ospf 1
router−id 172.16.10.1
network 10.9.10.0 0.0.0.1 area 0
network 10.13.0.0 0.0.255.255 area 0