

How does it built circuit on backbone ATM MPLS IP
and Unified MPLS carry label information example

Massimiliano Sbaraglia

Network Backbone ATM aware

ATM (Asynchronous Transport Mode)

si applica al trasporto, alla moltiplicazione ed alla commutazione di informazioni sotto forma di pacchetti in celle elementari di lunghezza fissa e relativamente brevi:

53 byte di cui 48 byte di informazione e 5 byte di intestazione.

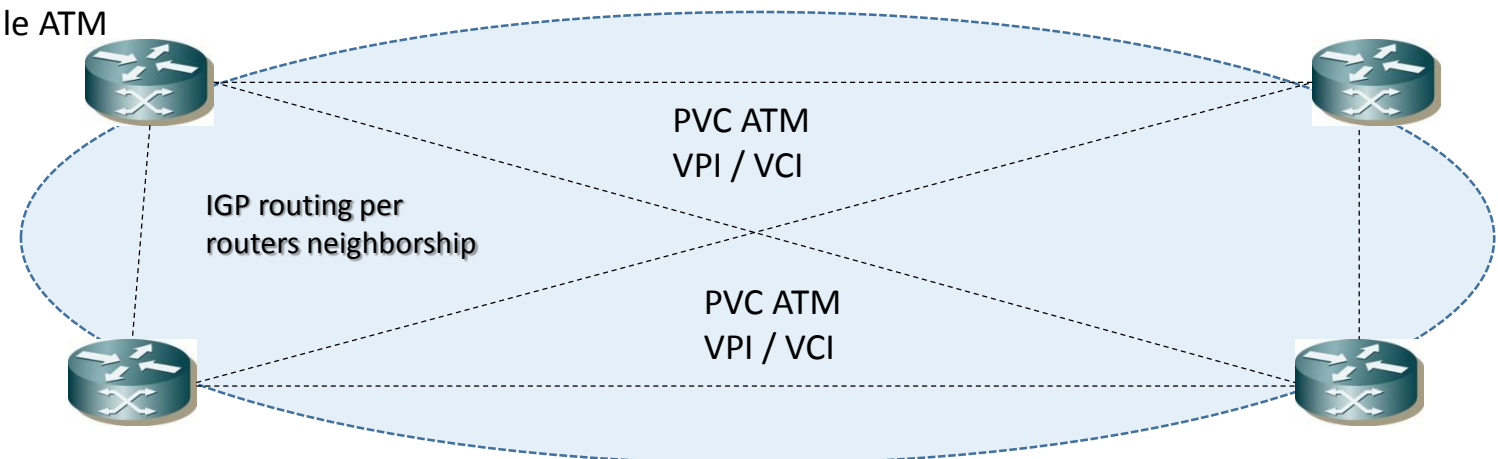
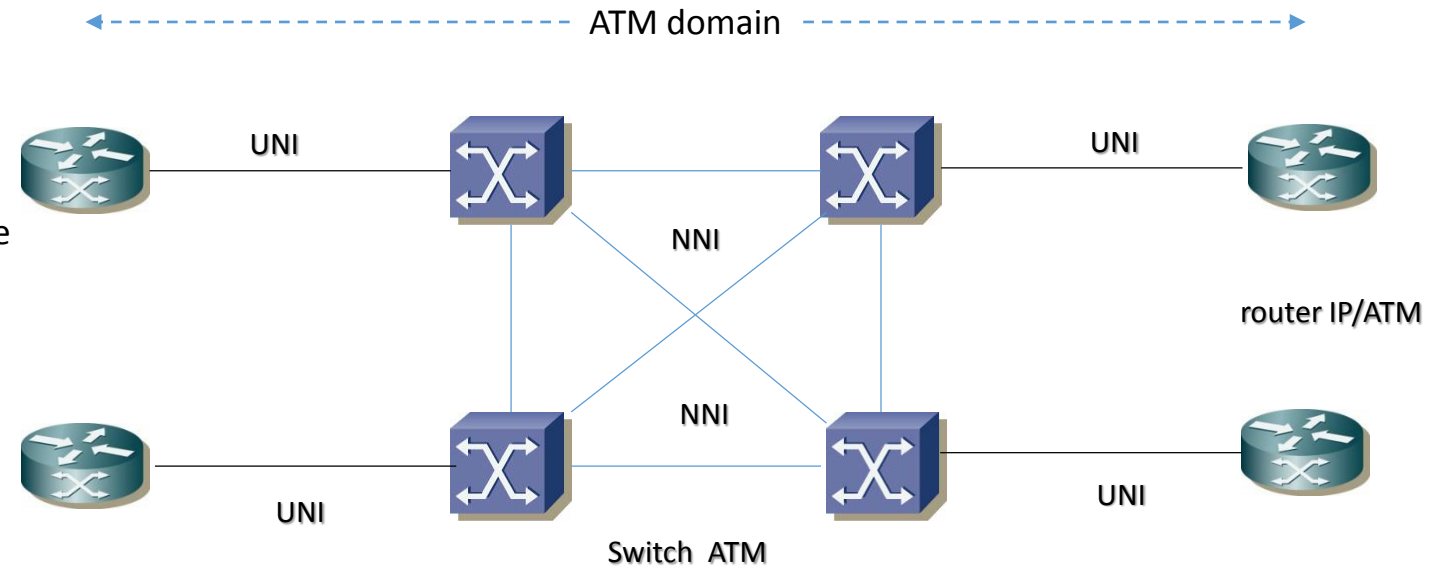
Il principio di funzionamento consiste in una segmentazione di flussi informativi di diverso tipo (voce, dati, immagini) in una sequenza di celle che vengono inserite nella rete di telecomunicazione.

UNI (User to Network):

interfaccia tra il terminale d'utente ed il commutatore locale ATM

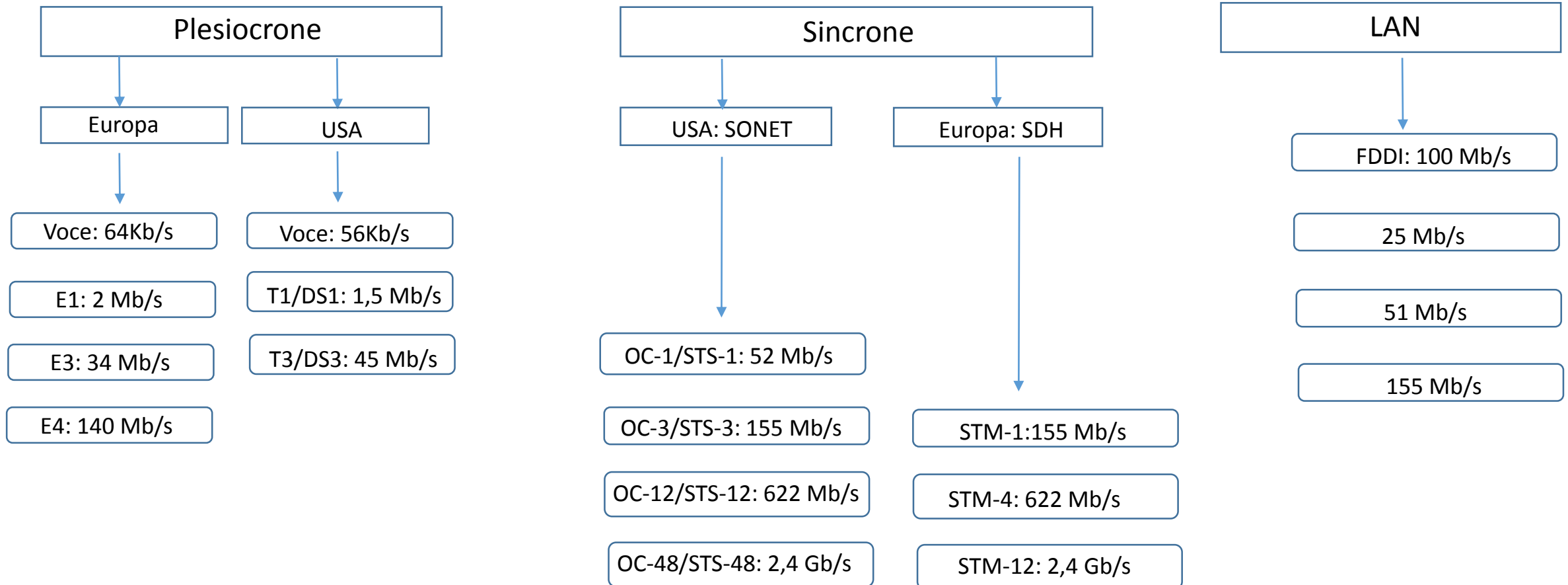
NNI (Network to Network):

interfaccia tra commutatori ATM



ATM aware e gerarchie Plesiocrone e Sincrone

La tecnica ATM può essere utilizzata sia nelle reti TLC locali che pubbliche e di conseguenza l'interfaccia UNI (User to Network) è stata specificata a livello fisico per poter operare con i suddetti standard:



ATM aware UNI e NNI cell packets

GFC (Generic Flow Control): è presente solo in una cella UNI e consente ad un Mux di controllare il flusso della cella da un terminale ATM, ovvero dall'interfaccia locale UNI; fondamentalmente il suo scopo è quello di risolvere le condizioni di sovraccarico a breve termine nella rete. Questo controllo di flusso non è necessario a livello NNI;

VPI (Virtual Path Identifier): occupa diverso spazio in termini di bit a seconda se l'interfaccia è di tipo utente (UNI) oppure network (NNI). La maggior lunghezza del campo VPI in una cella NNI consente di definire più percorsi virtuali all'interfaccia tra gli switch ATM;

VCI (Virtual Circuit Identifier): viene utilizzato per l'instradamento delle celle tra utenti finali.

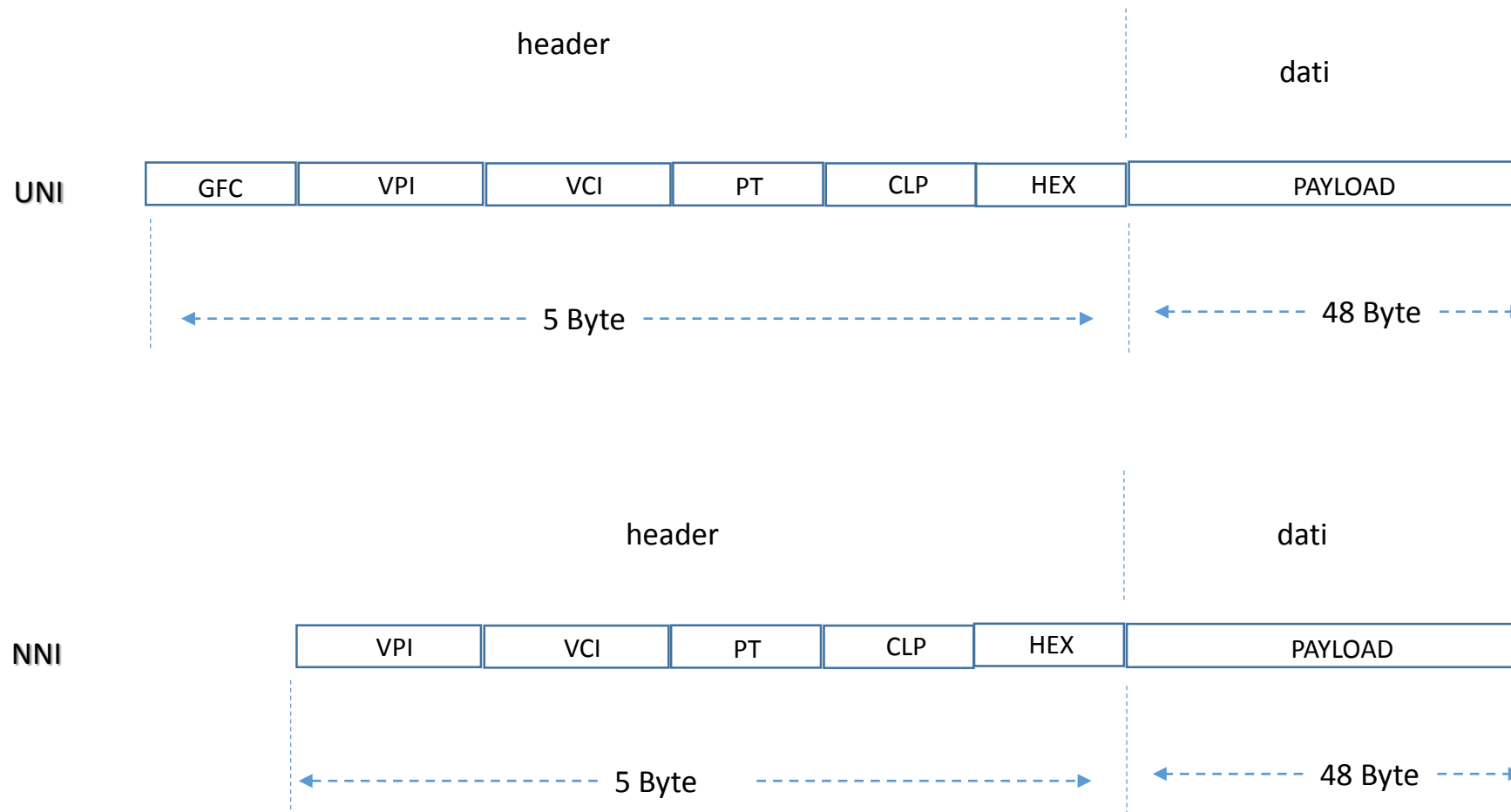
PT (Payload Type): indica se la cella contiene dati di utente, dati di segnalazione o informazioni di manutenzione.

CLP (Cell Loss Priority): indica la priorità relativa della cella e viene utilizzata per il controllo della congestione. Il valore 0 indica priorità alta di una cella che non dovrebbe essere eliminata dalla rete; il valore 1 indica priorità bassa di una cella ed indica che la cella può essere eliminata dalla rete.

HEC (Header Error Check): impiega un codice CRC per la rilevazione e la correzione di errori;

VPI/VCI: l'insieme di questi due campi indica l'etichetta (label) e serve a determinare la connessione a cui una cella appartiene; L'etichetta essendo costituita da un numero limitato di bit, identifica in modo univoco una connessione solo all'interno di un canale e pertanto nel passaggio da un canale in ingresso ad un canale in uscita deve essere opportunamente sostituita;

ATM aware UNI e NNI cell packets



ATM aware e classificazione dei servizi

ATM utilizza per la classificazione dei servizi un livelli di adattamento chiamato AAL, il quale è in grado di offrire una determinata classe di servizi attivabili in base alle esigenze di un cliente; la classificazione tiene conto di:

- riferimento temporale tra sorgente e destinazione
- bit rate
- modalità di connessione

AAL si differisce per:

- **AAL1**: classe A
- **AAL2**: classe B
- **AAL3/4**: classe C e D
- **AAL5**: è una semplificazione del modello AAL3/4 per rendere il protocollo più efficiente e di conseguenza più adatto ad un impiego nelle reti LAN ATM

Tutti i modelli AAL sono suddivisi in:

- **SAR (Segmentation and Reassembly)**: si occupa di segmentare le PDU (unità di informazione) proveniente dai livelli superiori della pila ISO/OSI in celle ATM e poi riassemblele in modo corretto
- **CS (Convergence Sublayer)**: svolge funzioni sulla base delle classi di servizio implementate

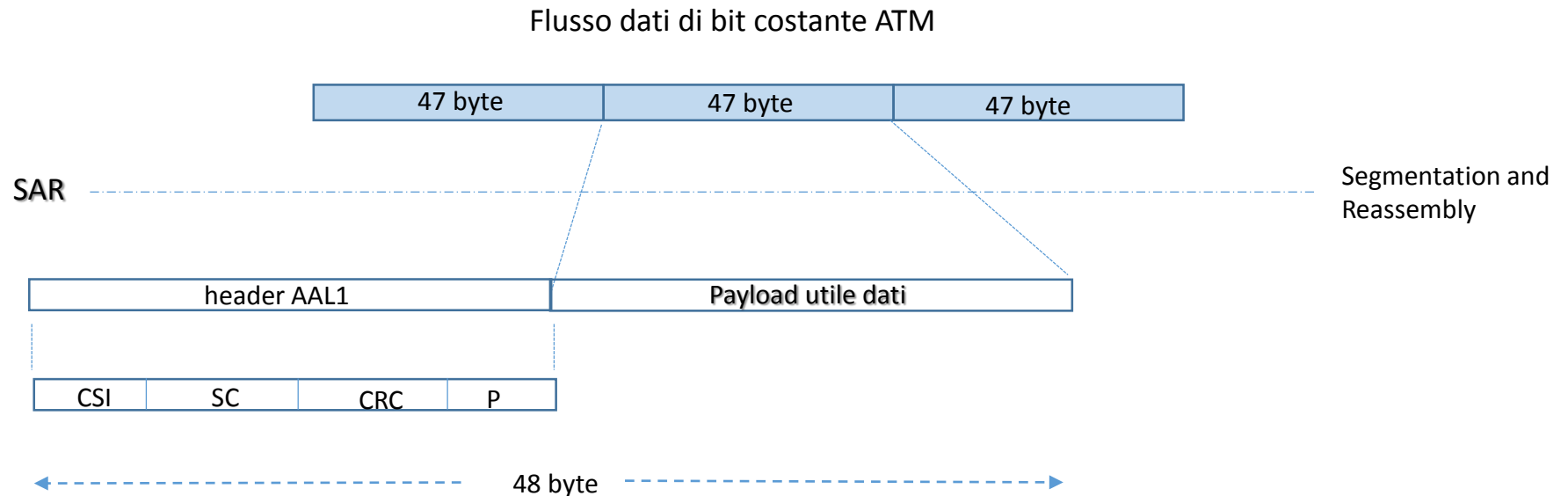
ATM aware e classificazione dei servizi

Classe di Servizio	Parametri QoS	Applicazioni
CBR (Costant Bit Rate)	Velocità di trasferimento a celle costanti	Voce, Video on demand, TV, VideoConferenza
VBR – nrt (Variable Bit Rate – non real time)	Velocità di trasferimento variabile nel tempo dipendente dalla disponibilità delle informazioni da parte utente	Transazioni bancarie, prenotazioni tickets, monitoraggio dei processi
VBR – rt (Variable Bit Rate – real time)	Velocità di trasferimento come sopra ma sensibili ad applicazioni con variazioni di ritardo nel tempo tra le celle	Voce e Video
ABR (Available Bit Rate)	E' possibile definire una velocità minima di trasferimento delle celle che deve essere garantita	Posta elettronica, FTP,
UBR (unspecified Bit Rate)	In questo caso le celle sono trasferite secondo uno schema FIFO e può subire ritardi variabili	FTP, messaging, downloading di testo ed immagini

ATM AAL1 encapsulation

AAL1 specifica il modo di trasferimento di applicazioni CBR (Costant Bit Rate) come ad esempio la voce ed il video in una rete ATM; in particolare descrive l'emulazione di circuito che fa riferimento alle connessioni di servizi telefonici digitali DS1, DS3, E1, ISDN)

- CSI (Convergence Sublayer Identifier): utilizzato per la segnalazione
- SC (Sequence Count): utilizzato per controllo errori punto-punto di un flusso dati attraverso un ordine ed identificazione delle celle ATM
- CRC (Cycles Redundant Check): utilizzato per un controllo di ridondanza ciclica del flusso dati eseguito sui primi quattro bit dell'header ATM
- P (Parity): utilizzato per un controllo di parità che considera i primi sette bit dell'header

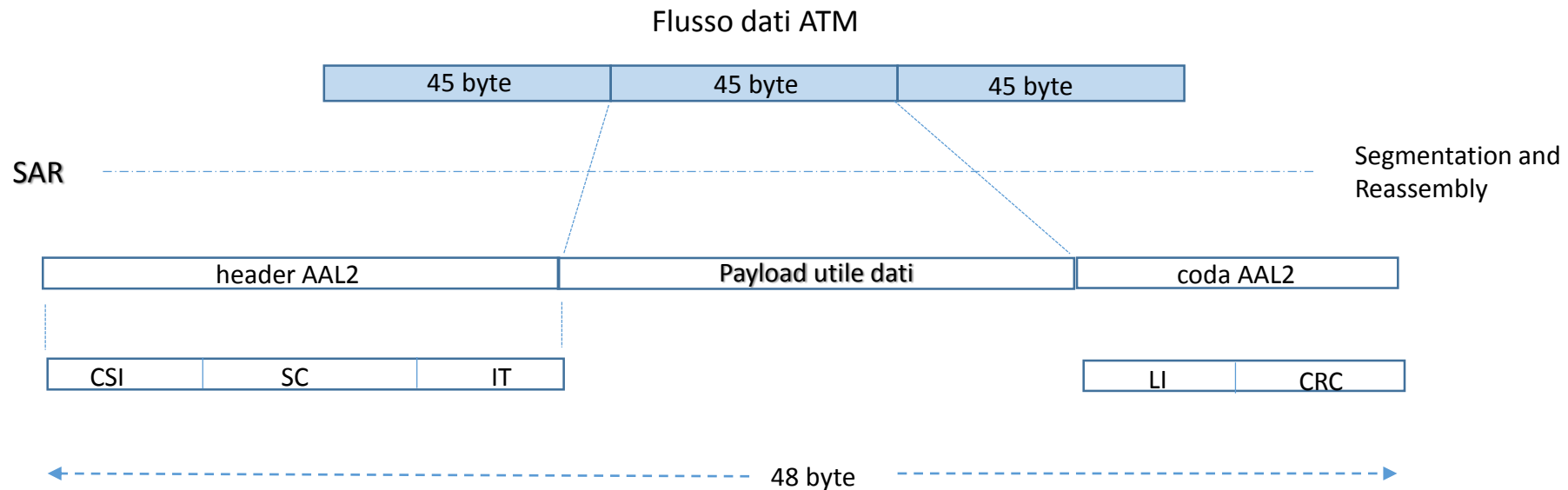


ATM AAL2 encapsulation

AAL2 supporta applicazioni di tipo VBR-rt (Variable Bit Rate – real time) quali voce, dati, telefax e ISDN channel B e D

I circuiti virtuali sono PVC, SVC

- **IT (Information Type)**: campo di 4 bit, identifica che un segmento dati cada all’inizio, in mezzo oppure alla fine di un messaggio
- **LI (Length Indicator)**: quando i bit IT indicano la fine del messaggio, i primi 6 bit della coda indicano quale parte della cella finale è costituita da dati; la parte rimanente è costituita da elementi di riempimento; in pratica, poiché il flusso di bit trasmesso non è divisibile per 45, vengono introdotti dei bit aggiuntivi nell’ultimo segmento, in modo da completare una cella ATM da 48 byte. Il campo LI indica pertanto la posizione di questi bit



ATM AAL3/4 encapsulation

AAL3/4 supporta servizi VBR (Variable Bit Rate) senza connessione come ad esempio i servizi per dati in banda larga CBDS (Connectionless Broadband Data Services) che possono tollerare una variazione nel ritardo.

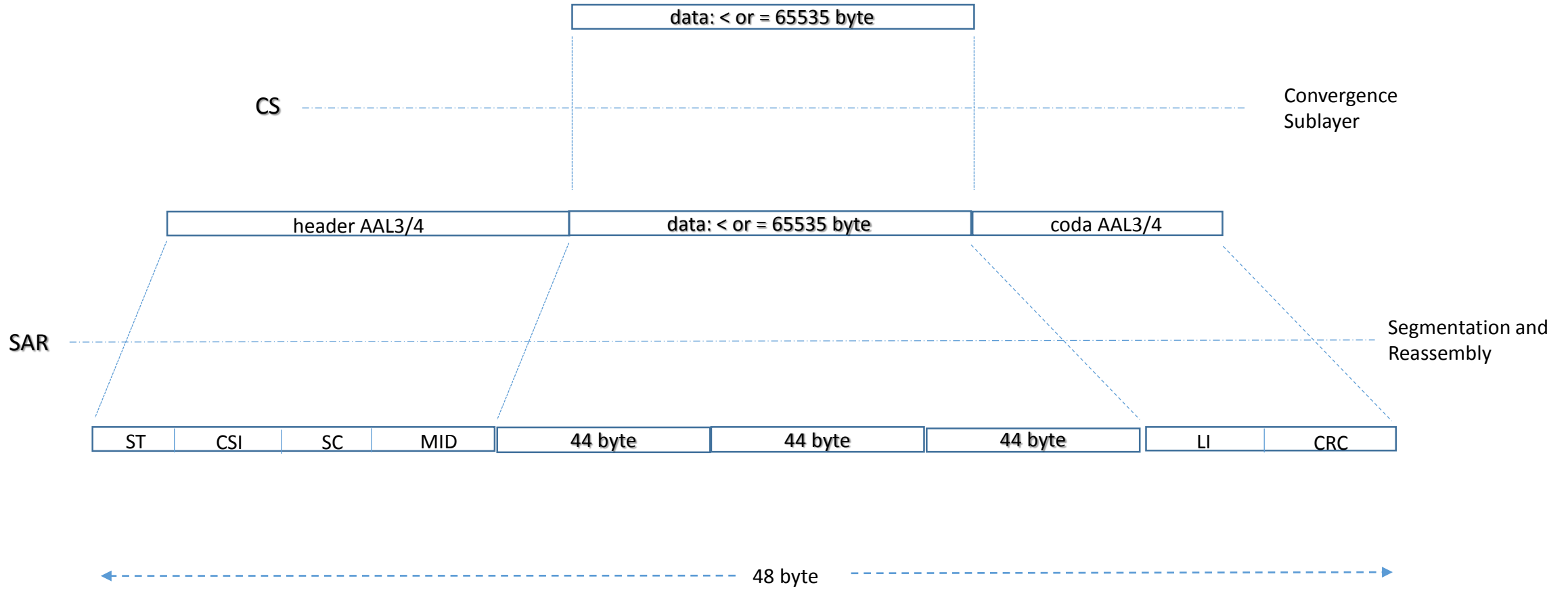
Il sottolivello CS accetta dal livello superiore un pacchetto che può essere costituito da un massimo di 65.535 byte ed aggiunge un'intestazione, una coda e dei bit di riempimento.

L'intestazione e la coda sono utili per le operazioni di riassetaggio alla destinazione in quanto indicano l'inizio e la fine dell'intero messaggio; i bit di riempimento vengono utilizzati per finire un segmento di 44 byte che verrà passato al sottolivello SAR.

Il sottolivello SAR accetta dal sottolivello CS i segmenti di carico utile di 44 byte e vi aggiunge 2 byte di intestazione e 2 byte di coda per formare un carico utile di 48 byte per la cella ATM.

- **ST (Segment Type):** indicano se il segmento dati cade all'inizio (ST=10), in mezzo (ST=00), oppure alla fine (ST=01) del messaggio o se si tratta di un messaggio costituito da un solo segmento (ST=11).
- **MID (Multiplexing Identification):** campo di 10 bit per 1024 possibili valori, consente di eseguire il multiplexing di più celle provenienti da flussi differenti sulla stessa connessione virtuale.
- **LI (Length Indicator):** quando i bit ST indicano la fine del messaggio o quando ST=11, i n° 6 bit LI indicano quale parte della cella è costituita dai dati e quale parte forma semplicemente la parte di riempimento.

ATM AAL3/4 encapsulation



ATM AAL5 encapsulation

AAL5 è stato definito per servizi orientati alla connessione in modo da ridurre il sovraccarico di elaborazione del protocollo, il sovraccarico di trasmissione e semplificare l'adattamento con altri protocolli di trasporto.

AAL 5 aggiunge solo una coda di 8 byte all'unità PDU del livello superiore; l'unità PDU CS può avere una lunghezza compresa tra 1 e 65535 byte.

Il sottolivello SAR suddivide questa unità CS PDU in segmenti di carico utile della lunghezza di 48 byte per le celle che vengono poi passate al livello ATM.

- **Pad:** bit di riempimento
- **UU (User to User identification):** 1 byte e viene utilizzato per trasferire in modo trasparente le informazioni da un utente all'altro
- **CPI (Common Part Indicator):** interpreta gli altri campi della coda CS PDU. Una funzione del CPI è quella di allineare la coda CS PDU a 64 bit
- **L (Length):** 2 byte, indica la lunghezza del campo di carico utile CS PDU. Indica quanta parte del segmento finale è costituita da dati e quanta da riempimento. Inoltre il ricevitore utilizza il valore del campo L per rilevare la perdita di informazioni.
- **CRC :** $G(x) = x^{32} + x^{26} + x^{23} + x^{22} + x^{16} + x^{12} + x^{11} + x^{10} + x^8 + x^7 + x^5 + x^4 + x^2 + x + 1$

ATM QoS

I parametri di qualità del servizio ATM sono definiti:

- **PCR (Peak Cell Rate):** definisce la velocità massima con la quale l'utente può trasmettere.
- **SCR (Sustained Cell Rate):** è la velocità media delle celle misurata su un lungo periodo di tempo. La velocità di trasmissione effettiva può essere superiore o inferiore e può a volte raggiungere la velocità PCR.

I parametri orientati alla rete definiscono le caratteristiche della rete e sono:

- **CLR (Cell Loss Ratio):** è la percentuale di celle che non vengono consegnate a destinazione. Queste celle possono perdersi o possono essere consegnate talmente in ritardo da essere considerate perse. La perdita di celle può essere dovuta ad una congestione o ad un sovraccarico dei buffer.
- **CER (Cell Error Rate):** rappresenta la frazione di celle che vengono consegnate con errori.
- **CTD (Cell Transfer Delay):** è il ritardo sperimentato da una cella tra il momento in cui è entrata nella rete all'origine ed il momento in cui è uscita alla destinazione.
- **CDV (Cell Delay Variation):** misura la differenza fra il ritardo CTD massimo e minimo.

ATM switching labels

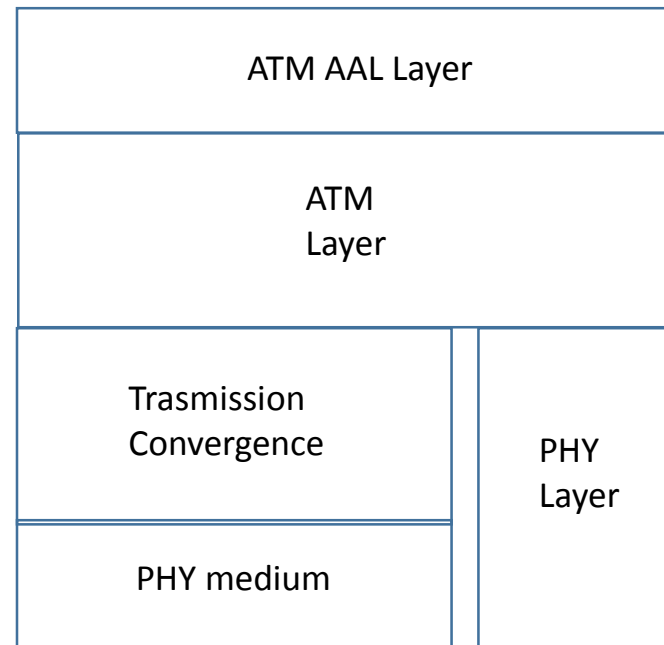
La funzione di traslazione di etichette viene realizzata fisicamente mediante una tabella di lookup associata a ciascuna porta di ingresso del commutatore ATM.

A tale tabella viene riscontrata una variazione di etichetta da quella vecchia a quella nuova, mentre la commutazione di multiplex viene implementata per mezzo di una struttura di commutazione spaziale dotata di buffer, cioè di memorie in cui le celle vengono raccolte in attesa della commutazione per un certo periodo di tempo di durata non deterministica, dipendenti solo dalle condizioni di carico presente in rete.

I buffer sono indispensabili a causa della tecnica asincrona impiegata per il trasferimento di celle ATM e quindi servono a risolvere i problemi di conflitto che si generano tra celle che arrivano su porte di ingresso differenti e sono destinate alla stessa porta di uscita.

CONTROL PLANE
trasporto e trattamento
dell'informazione di segnalazione

CS Convergence Sublayer
SAR Segment and Reassem
Cell header
Cell VPI / VCI
Cell mux / demux
Cell rate decoupling
HEC gener / verific
Cell delineation
Bit Timing
Bit Tx / Rx



USER PLANE
trasporto delle informazioni
utente: voce, video, dati

ATM switching labels

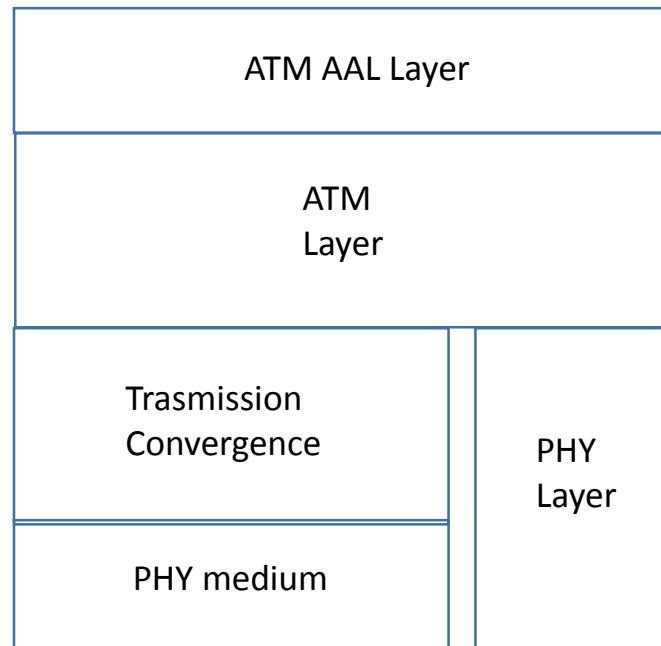
La funzione di traslazione di etichette viene realizzata fisicamente mediante una tabella di lookup associata a ciascuna porta di ingresso del commutatore ATM.

A tale tabella viene riscontrata una variazione di etichetta da quella vecchia a quella nuova, mentre la commutazione di multiplex viene implementata per mezzo di una struttura di commutazione spaziale dotata di buffer, cioè di memorie in cui le celle vengono raccolte in attesa della commutazione per un certo periodo di tempo di durata non deterministica, dipendenti solo dalle condizioni di carico presente in rete.

I buffer sono indispensabili a causa della tecnica asincrona impiegata per il trasferimento di celle ATM e quindi servono a risolvere i problemi di conflitto che si generano tra celle che arrivano su porte di ingresso differenti e sono destinate alla stessa porta di uscita.

CONTROL PLANE
trasporto e trattamento
dell'informazione di segnalazione

CS Convergence Sublayer
SAR Segment and Reassem
Cell header
Cell VPI / VCI
Cell mux / demux
Cell rate decoupling
HEC gener / verific
Cell delineation
Bit Timing
Bit Tx / Rx



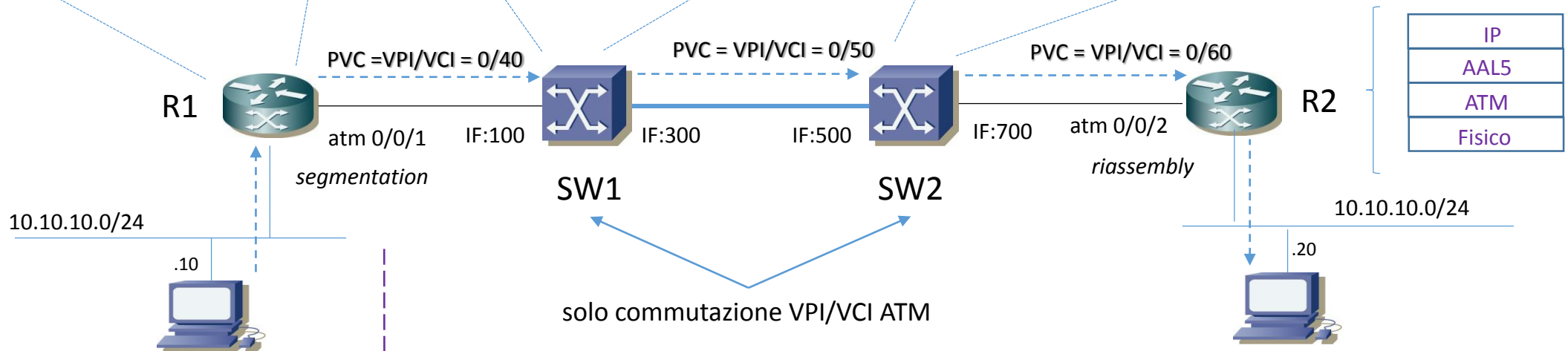
USER PLANE
trasporto delle informazioni
utente: voce, video, dati

ATM packets switching overlay

Network IP	Next-Hop	VPI	VCI
10.10.10.0/24	SW1	0	40

IF IN	VPI/VCI IN	IF OUT	VPI/VCI OUT
100	0/40	300	0/50

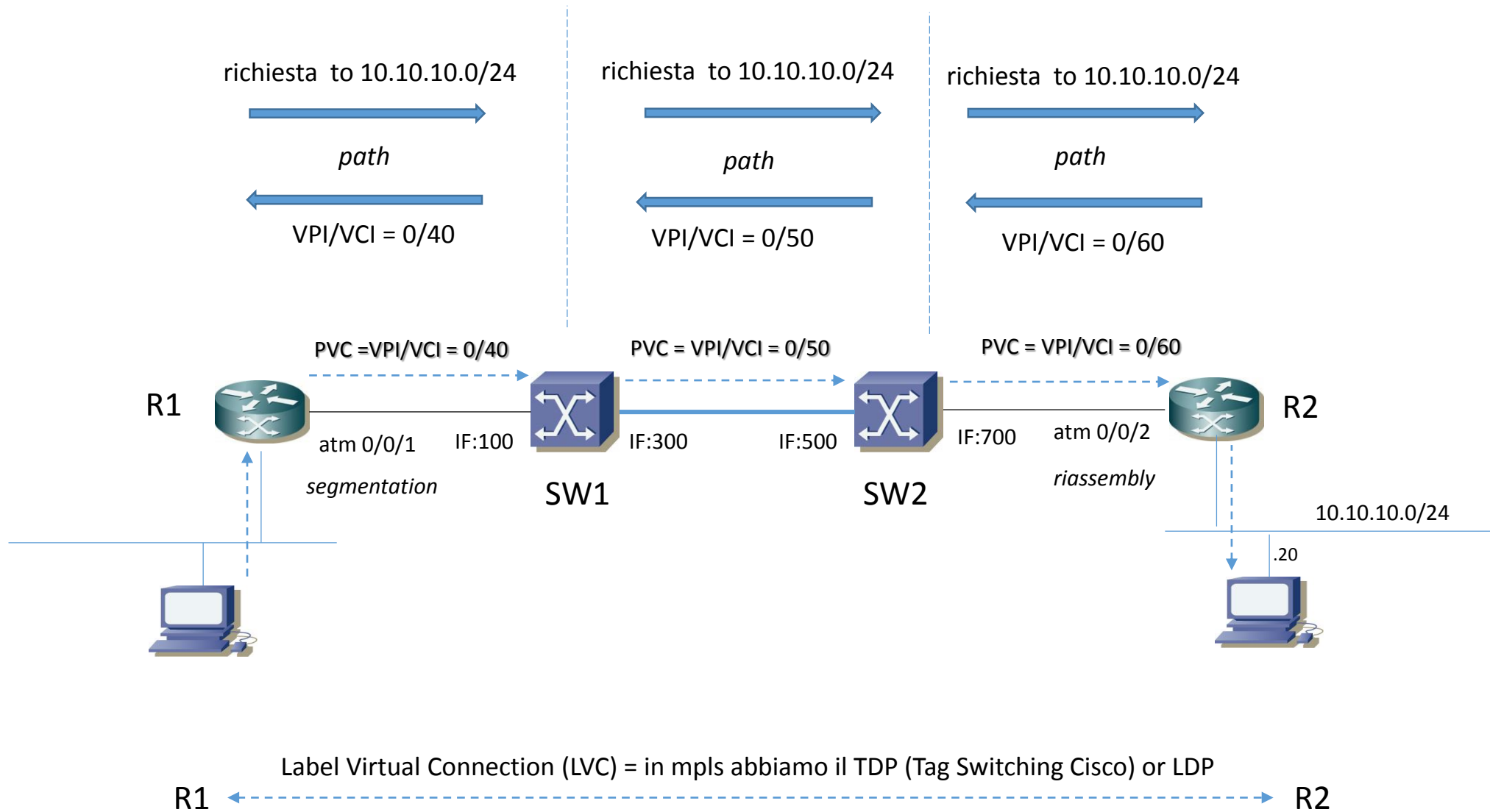
IF IN	VPI/VCI IN	IF OUT	VPI/VCI OUT
500	0/50	700	0/60



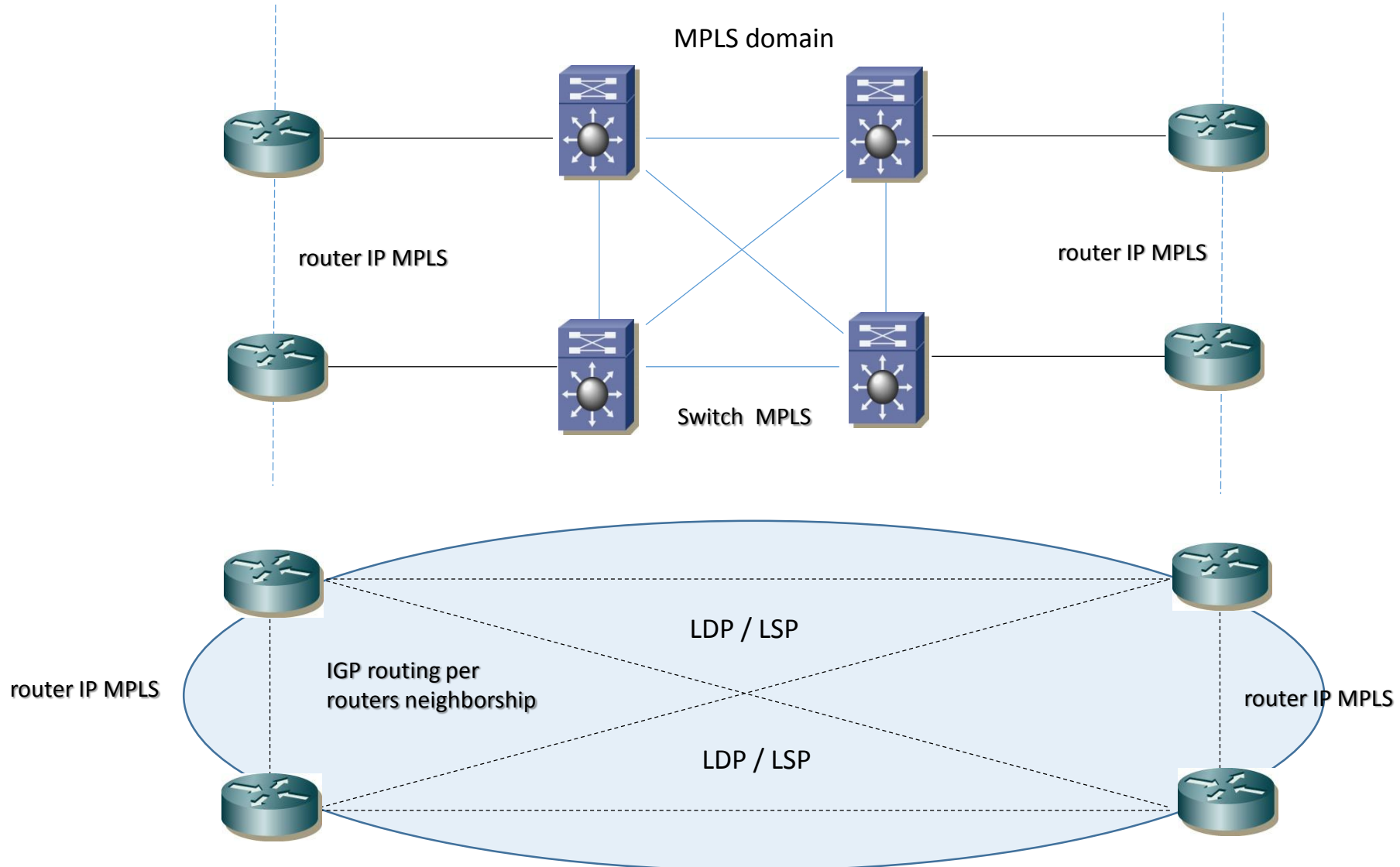
```

int atm 0/0/1
ip address 1.1.1.1 255.255.255.252
pvc 0/40
abr 1000 500 --# PCR = 1 Mbit/s ; MCR = 500 Kbit/s
encapsulation aal5snap
    
```

ATM packets switching overlay associazione VCI path R1-SW1-SW2-R2

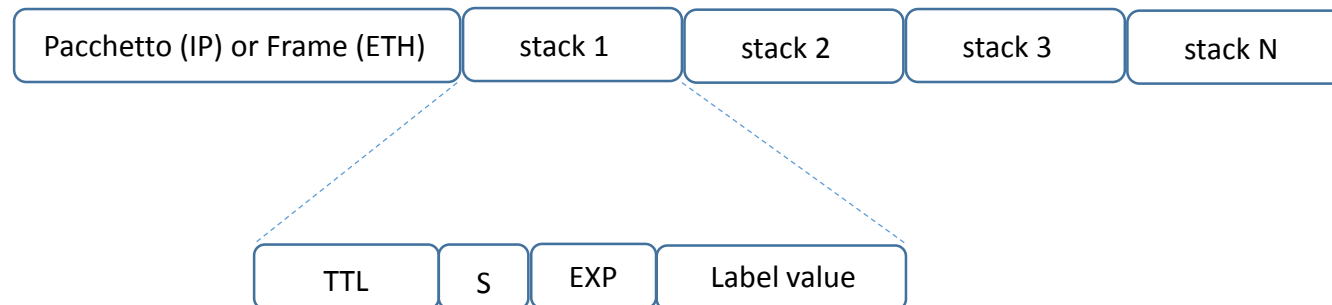


Network Backbone MPLS aware



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: rimozione header MPLS ed annuncio su base del solo prefix IP (S=1)
 - Label = 1 : Router Alert: il pacchetto è soggetto ad un trattamento da parte del nodo MPLS; l'annuncio si basa sul valore di push di una nuova etichetta da inserire nello stack MPLS
 - Label = 2 : IPv6 Explicit Null Label: come sopra per IPv4 ma questa volta per IPv6
 - Label = 3 : Implicit Null: è un valore di etichetta che richiede ad altri nodi MPLS una operazione di pop (rimozione della label su uno stack MPLS)

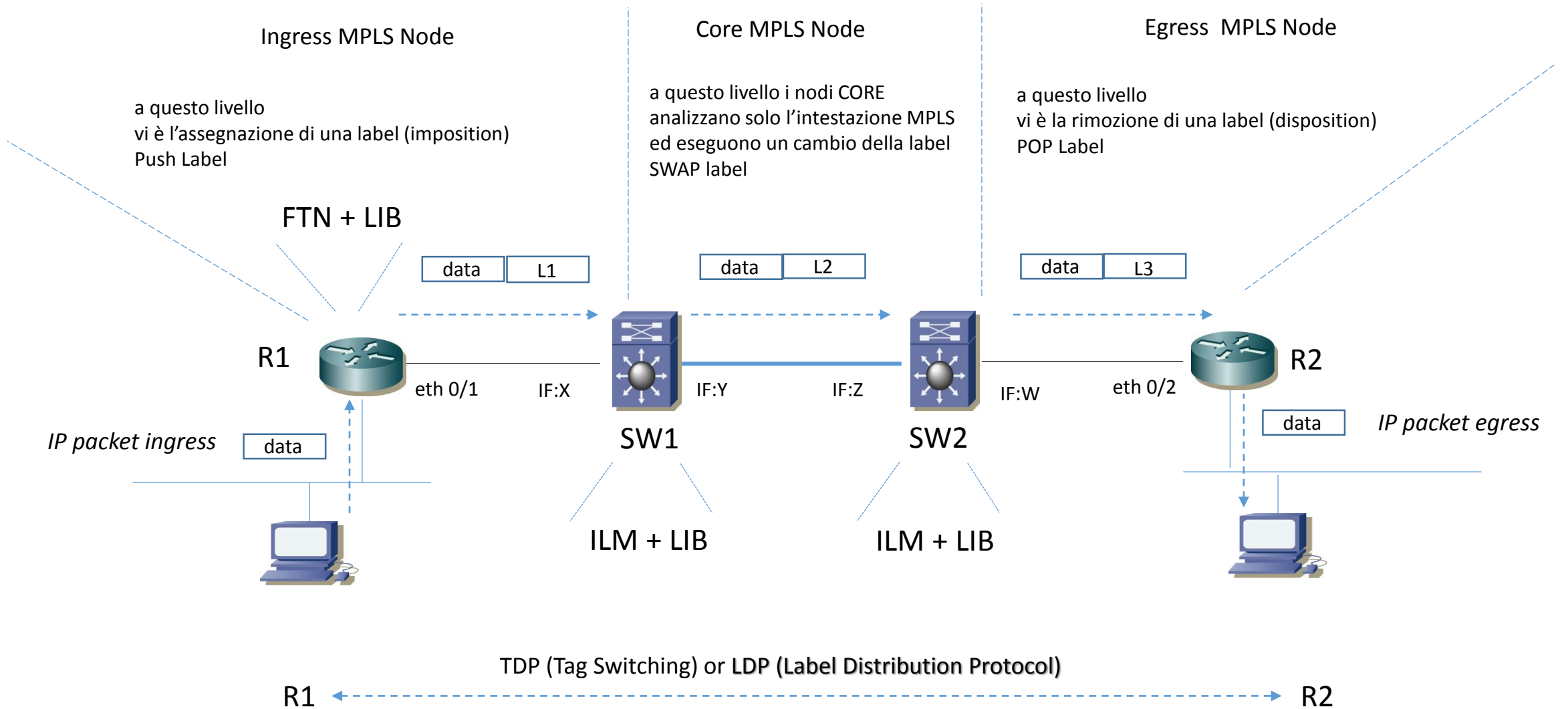


TTL per Loop
Prevention

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

EXP per differenziare qualità del servizio

MPLS packets switching integrato path R1-SW1-SW2-R2



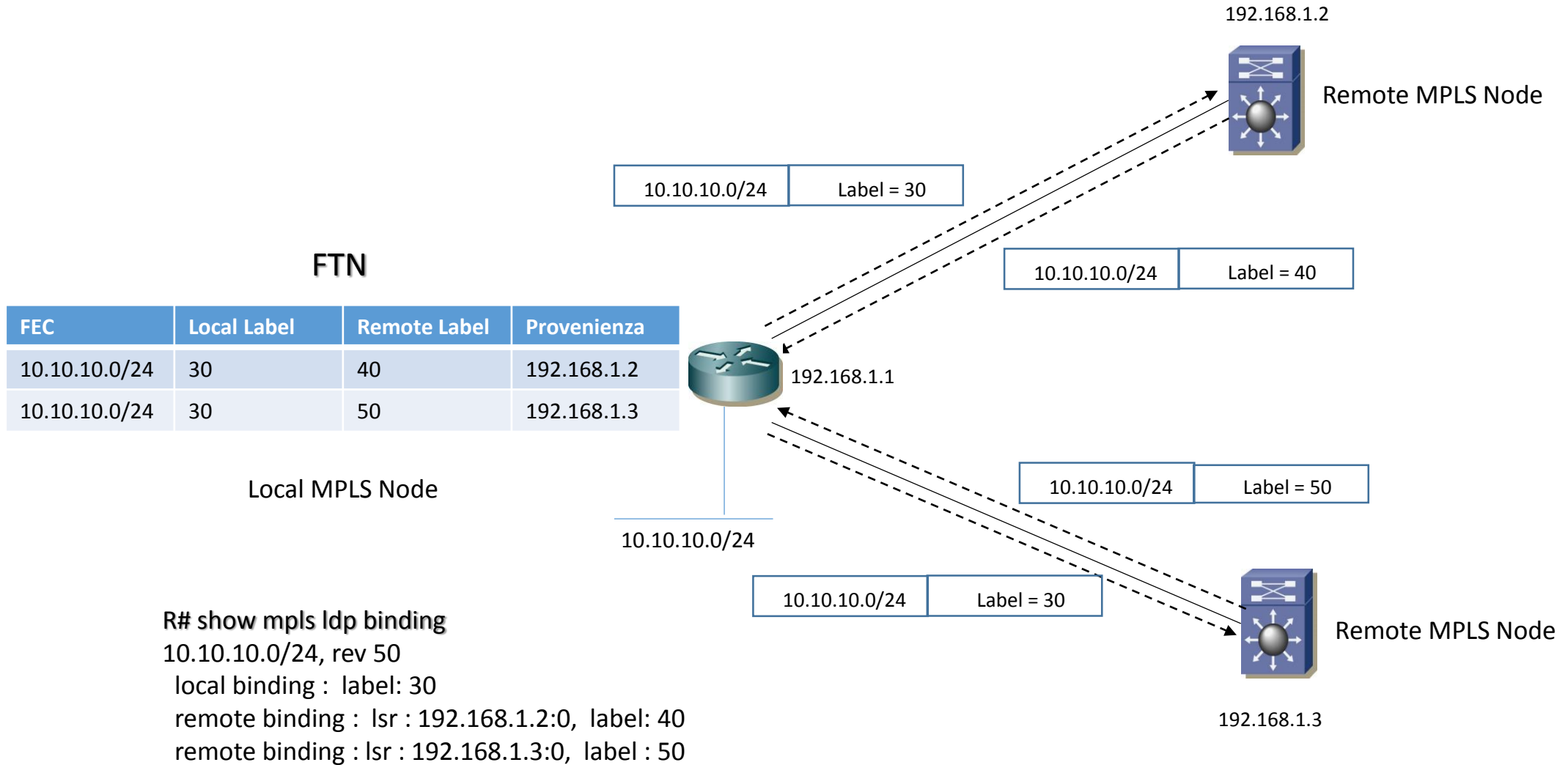
MPLS FTN (FEC To NHLFE)

- **FTN (FEC To NHLFE)**: è una tabella presente sui nodi MPLS di Edge MPLS; associa ad ogni pacchetto IP entrante inserito in una determinata FEC (Forwarding Equivalent Class) un insieme di istruzioni (to NHLFE = verso il Next Hop Label Forwarding Entry) necessarie per l'inoltro del pacchetto al nodo successivo
- **ILM (Incoming Label Map)**: è una tabella presente negli switch MPLS intermedi (No Edge Routers); associa ad ogni label entrante un insieme di istruzione (NHLFE) necessarie per l'inoltro del pacchetto al nodo successivo

La costruzione di queste due tabelle, richiede una terza tabella chiamata LIB (Label Information Base) che contiene le seguenti informazioni:

- associazione FEC – Label locale (eseguita dal nodo locale = se stesso)
 - local label binding
- associazione FEC – label remota (eseguita dai nodi adiacenti e non = peer node)
 - Remote label binding
- Provenienza: identificativo de nodo remoto che ha comunicato l'associazione FEC-label

MPLS FTN (FEC To NHLFE)

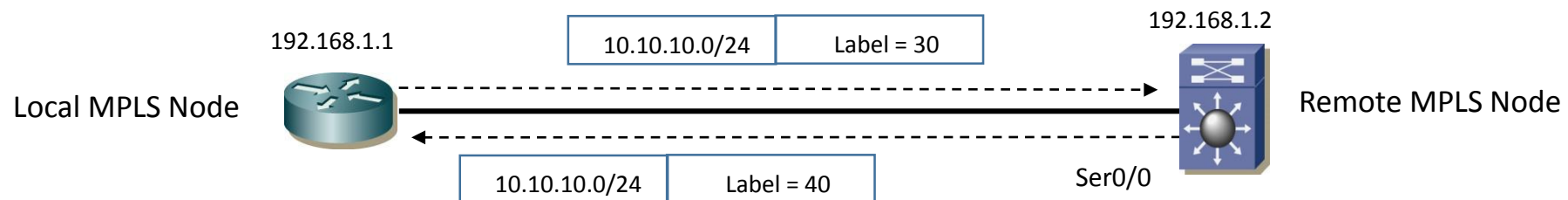


MPLS ILM (Incoming Label Mapping)

ILM (Incoming Label Map): è una tabella presente negli switch MPLS intermedi (No Edge Routers); associa ad ogni label entrante un insieme di istruzioni (NHLFE) necessarie per l'inoltro del pacchetto al nodo successivo

La ILM è costruita attraverso due modalità:

- **Dinamica**: su base della routing table e dalla LIB presente nel nodo MPLS; in questo caso il Next-Hop viene ricavato da un protocollo di routing OSPF, ISIS
- **Eplicita**: su base segnalazione per mezzo di un protocollo quale RSVP, LDP; in questo caso il Next-Hop viene determinato esplicitamente in modo manuale oppure automatico



FTN

FEC	Local Label	Remote Label	Provenienza
10.10.10.0/24	30	40	192.168.1.2
10.10.10.0/24	30	50	192.168.1.3



LIB

ILM

Label Entry	NHLFE Label	Interface	Next-Hop
30	40	Ser 0/0	192.168.1.1

MPLS ILM (Incoming Label Mapping): example output

R# show mpls forwarding-table

Local tag	Outgoing tag or VC	Prefix or Tunnel-ID	Bytes tag switched	Outgoing interface	Next-Hop
101	Untagged	10.10.10.0/24	0	Ser 0/0	point2point
102	Pop tag	192.168.0.50/32	0	fas 1/0	192.168.1.1
103	Untaggeg	172.16.5.0/24	0	fas 1/0	192.168.1.2
104	204	10.12.13.0/24	0	fas 1/0	192.168.1.2

ingress-label	egress-label	FEC	Egress Interface	Next-Hop
---------------	--------------	-----	------------------	----------

Untagged = significa che alla FEC non è stata associata nessuna label;

POP Tag = indica una operazione di rimozione della label in testa allo stack; in genere significa che il nodo è il penultimo del path verso queste destinazioni

MPLS distribuzione labels LDP protocols

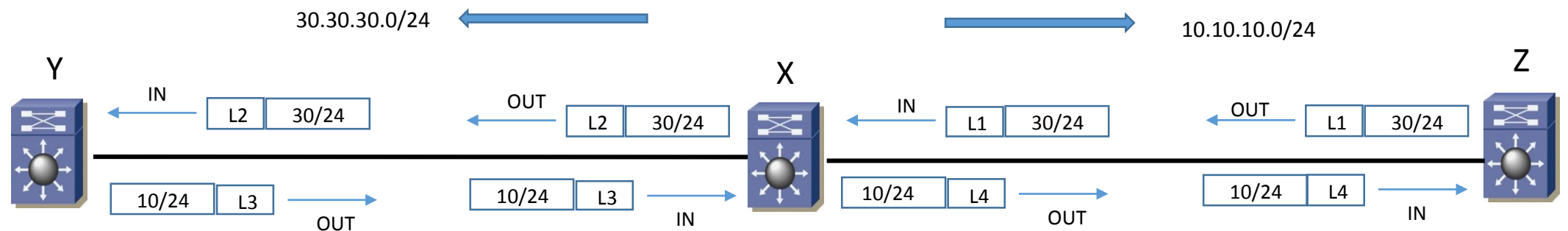
La distribuzione associazione FEC-label avvengono per:

- direzione della distribuzione: direzione del traffico oppure in direzione contraria ?
- a quali nodi inviare le associazioni: solo a chi ne fa richiesta oppure a tutti i nodi MPLS ?
- quando inviare le associazioni FEC-label: appena creata l'associazione oppure dopo aver ricevuto l'associazione dal next-hop ?

Ci sono due distinzioni da fare per identificare label-entrante e label-uscente dal punto di vista di un nodo MPLS

- Label entrante è quella definita locale ed è in direzione downstream rispetto alla direzione del traffico
- Label uscente è quella definita remota ed è in direzione upstream rispetto alla direzione del traffico

Un nodo si definisce in posizione upstream rispetto ad un altro nodo lungo il path se nell'ordine di percorrenza di trova prima rispetto alla direzione del traffico.



Dowstream rispetto ad X per 30.30.30.0/24
Upstream rispetto ad X per 10.10.10.0/24

Upstream rispetto ad Z per 10.10.10.0/24
Upstream rispetto a Y per 30.30.30.0/24

Dowstream rispetto ad X per 10.10.10.0/24
Upstream rispetto ad X per 30.30.30.0/24

MPLS LSP Label Switch Path

LSP è un path end-to-end tra due nodi edge MPLS; i nodi intermedi sono trasparenti al contenuto del pacchetto MPLS (sia per contenuti di livello 3 IP che di livello 2 ethernet), lavorando quindi solo per la commutazione dei pacchetti basandosi solo sul valore della label (la label più esterna in caso di uno stack di labels).

Gli LSP sono monodirezionali, pertanto il trasporto del traffico è in una sola direzione tra i due nodi MPLS; il traffico di ritorno transita su un LSP differente.

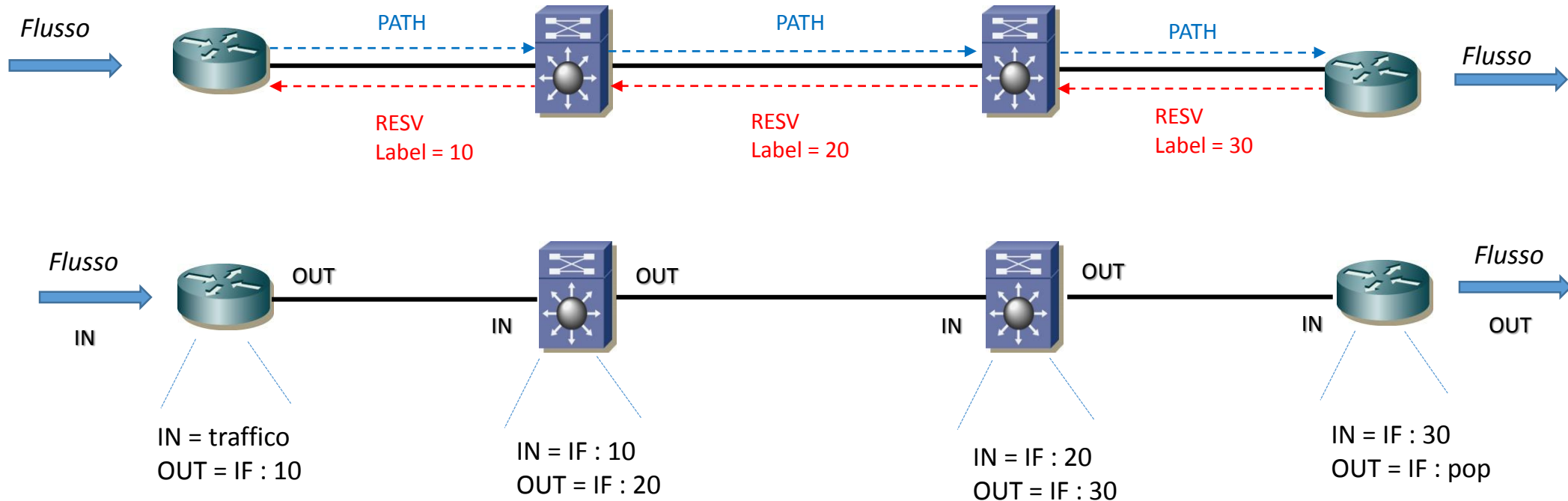
Il percorso di un LSP può essere determinato da un protocollo di routing IGP (LSP di tipo hop-by-hop) oppure in modo esplicito attraverso un protocollo di segnalazione (LSP explicitly routed molto utilizzati nelle problematiche di ingegnerizzazione del traffico)

Il protocollo di segnalazione si basa su:

- **RSVP-TE:** è una estensione del protocollo RSVP per la prenotazione di banda trasmissiva
- **CR-LDP:** è una estensione del protocollo LDP per la distribuzione delle labels

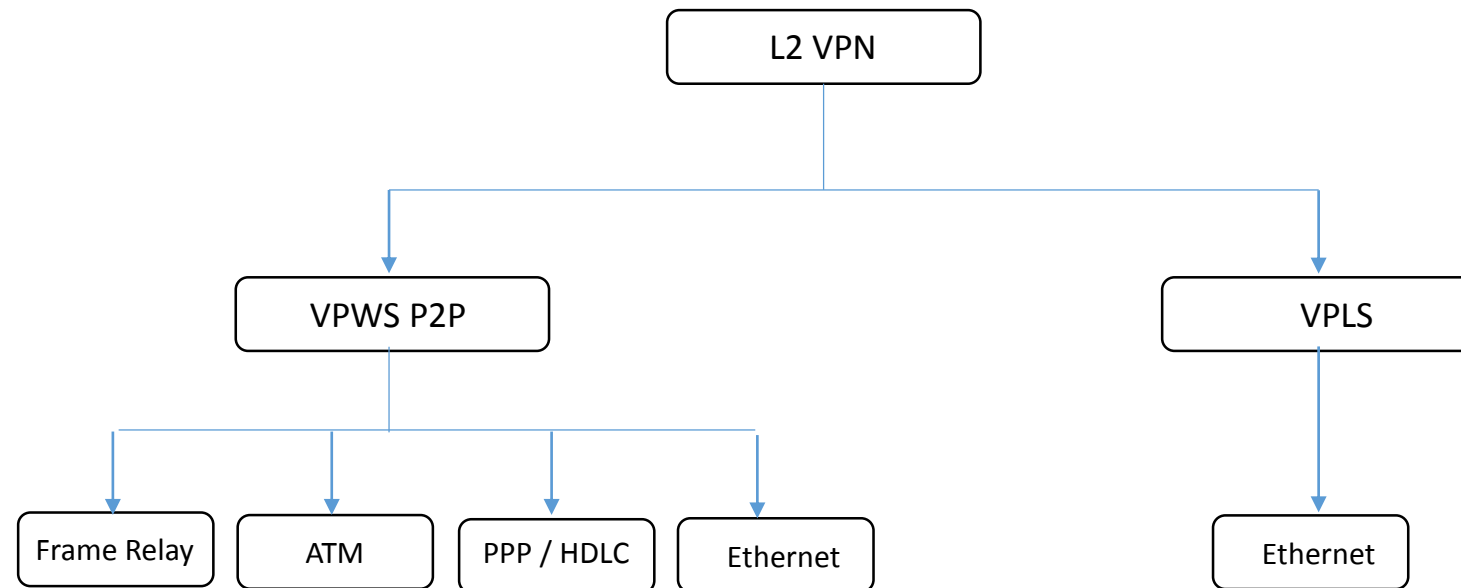
MPLS RSVP signaling

- prenotazione di banda unidirezionale tra sorgente e destinazione per un determinato flusso di traffico
- soft state signaling: continui rinfreschi su stati di prenotazione per un certo periodo di tempo, scaduto il quale viene cancellato
- la prenotazione di banda avviene attraverso messaggi di segnalazione Path e Reservation State, di cui:
 1. PATH STATE: relativo al processo di instradamento
 2. RESV STATE: relativo al processo di prenotazione della banda



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 reachability 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) utilizzata 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.

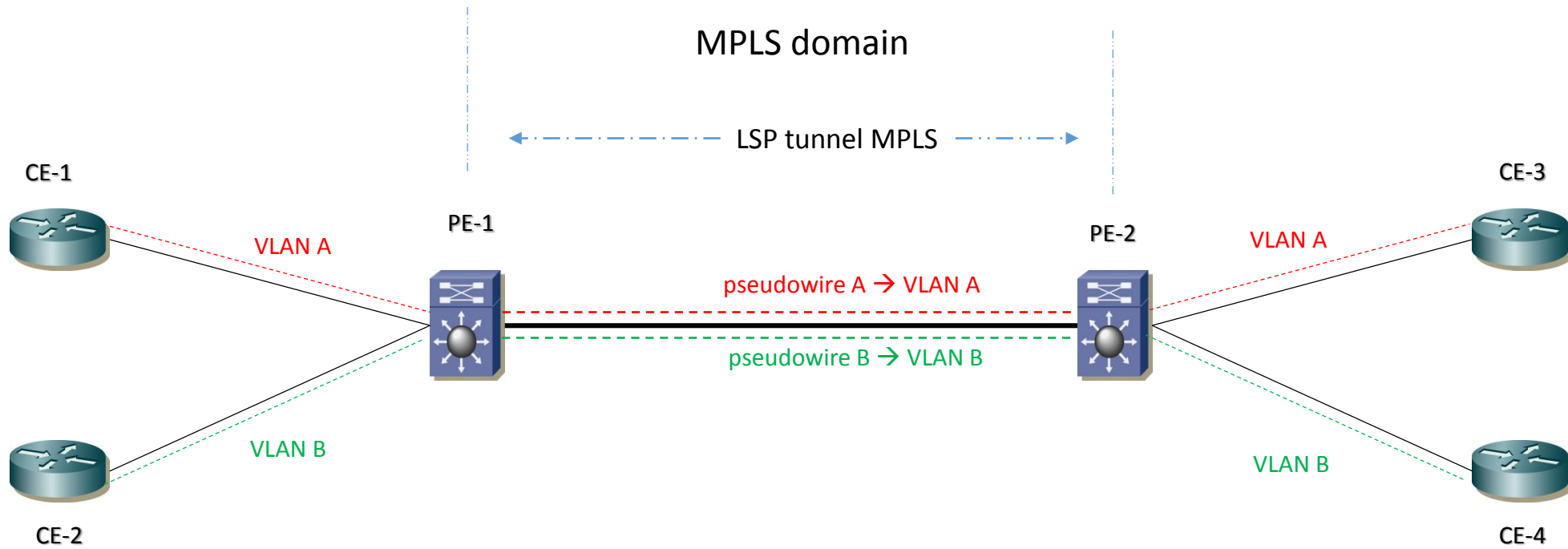
MPLS VPN-L2 application

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)

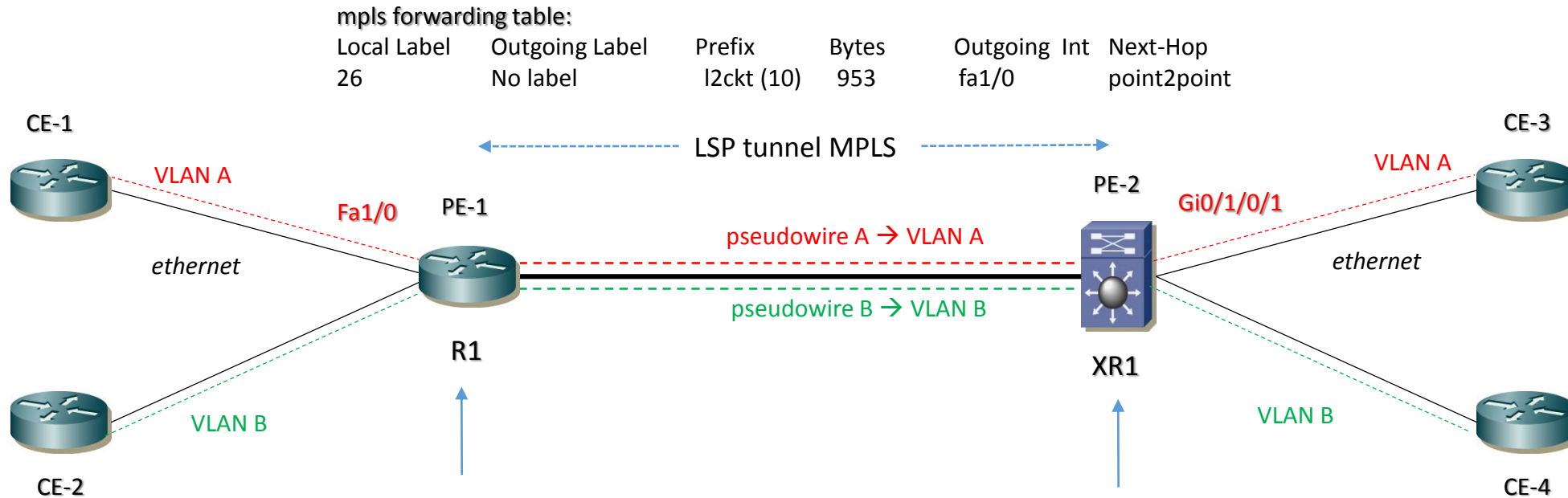
Per VPWS esiste un solo punto di uscita presso il router PE egress, quindi questo significa che i PE MPLS non hanno necessità di mantenere una tabella di MAC addresses per la costruzione del PW.

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

MPLS VPN-L2 application pseudowire design



MPLS VPN-L2 application pseudowire configuration example



mpls forwarding table:

Local Label	Outgoing Label	Prefix	Bytes	Outgoing Int	Next-Hop
26	No label	I2ckt (10)	953	fa1/0	point2point

```

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 I2 transport
 !
 I2vpn
 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)

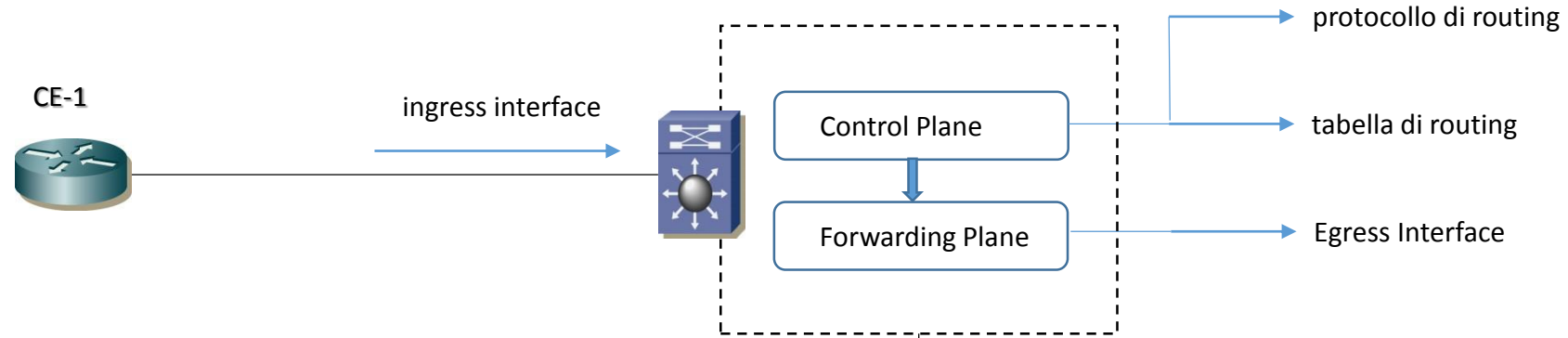
Il FBD (Forwarding DataBase) è contenuto nell'ambito della VFT (VPLS Forwarding DataBase); oltre alle interfacce fisiche come punti di uscita (egress), sono presenti un certo numero di virtual-ports ognuna delle quali emula una porta ethernet di collegamento verso tutti i PE routers membri del dominio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

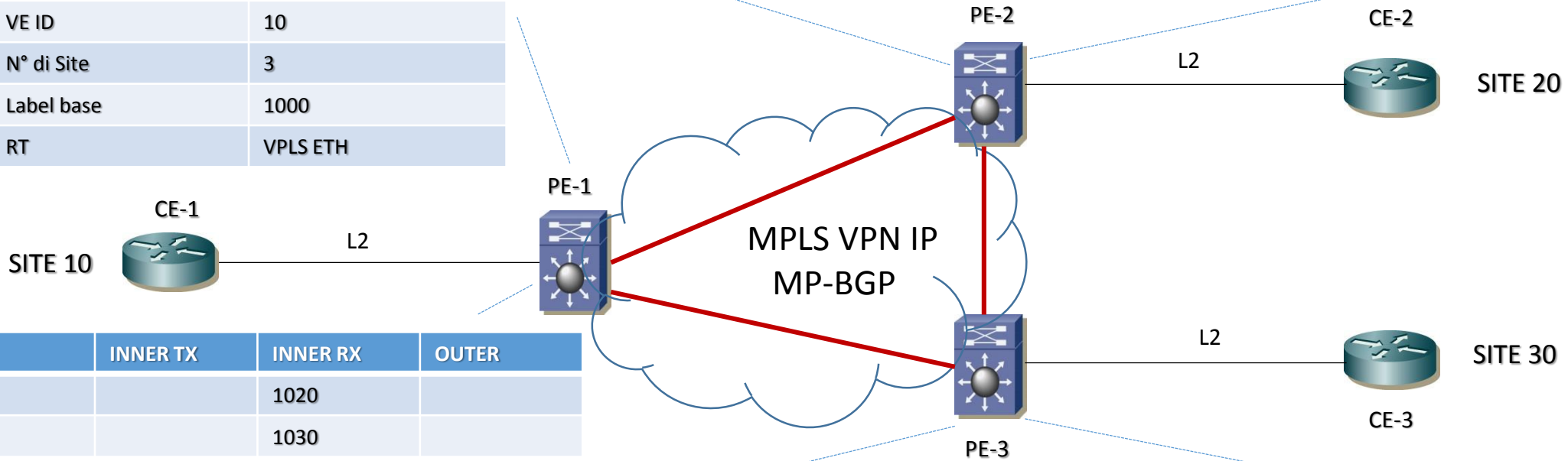
VCT Circuit Table	Details
RD	AS:NN
VE ID	20
N° di Site	3
Label base	2000
RT	VPLS ETH

VE ID	INNER TX	INNER RX	OUTER
10		2010	
30		2030	

VE ID	INNER TX	INNER RX	OUTER
20		1020	
30		1030	

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
10		3010	
20		3020	



PE1: VFT vpls forwarding Table

PE3: VFT vpls forwarding Table

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

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

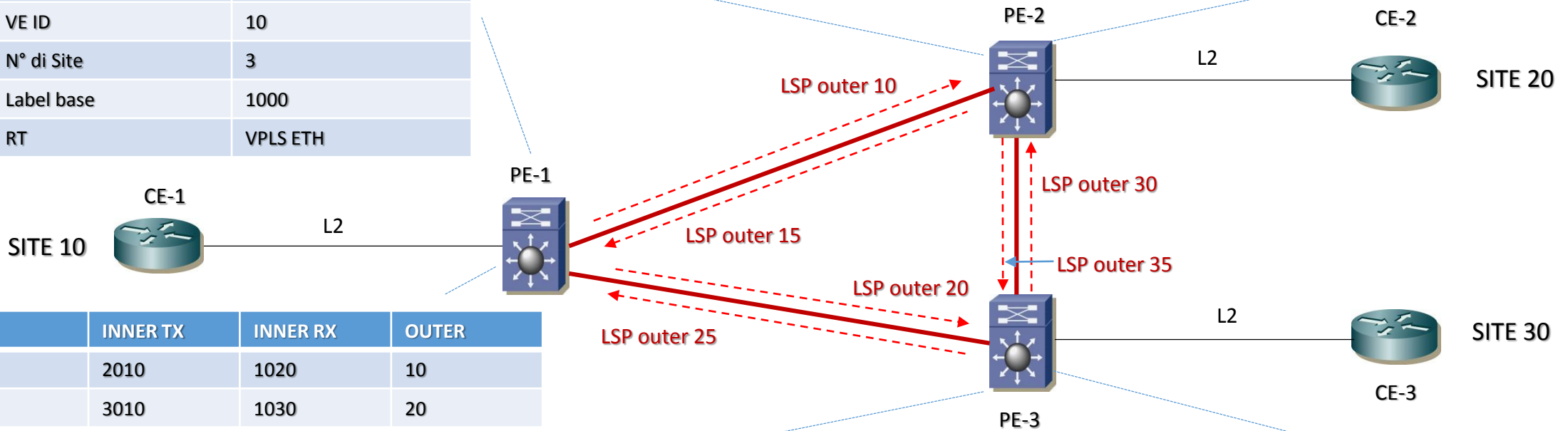
VE ID	INNER TX	INNER RX	OUTER
20	2010	1020	10
30	3010	1030	20

PE1: VFT vpls forwarding Table

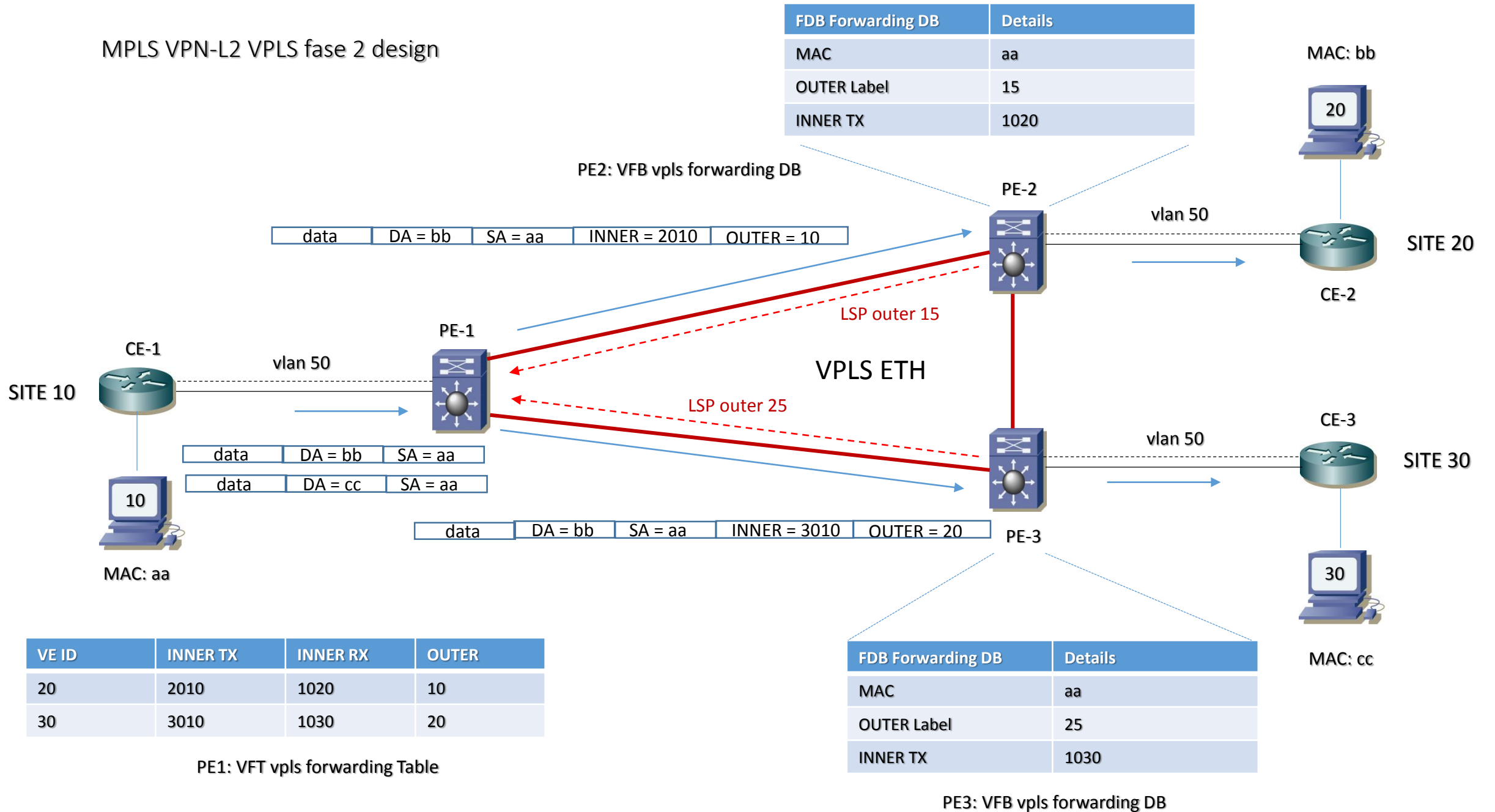
VCT Circuit Table	Details
RD	AS:NN
VE ID	30
N° di Site	3
Label base	3000
RT	VPLS ETH

VE ID	INNER TX	INNER RX	OUTER
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

PE2: VFB vpls forwarding DB

data	DA = bb	SA = aa	INNER = 2010	OUTER = 10
------	---------	---------	--------------	------------

data	DA = bb	SA = aa
data	DA = cc	SA = aa

data	DA = bb	SA = aa	INNER = 3010	OUTER = 20
------	---------	---------	--------------	------------

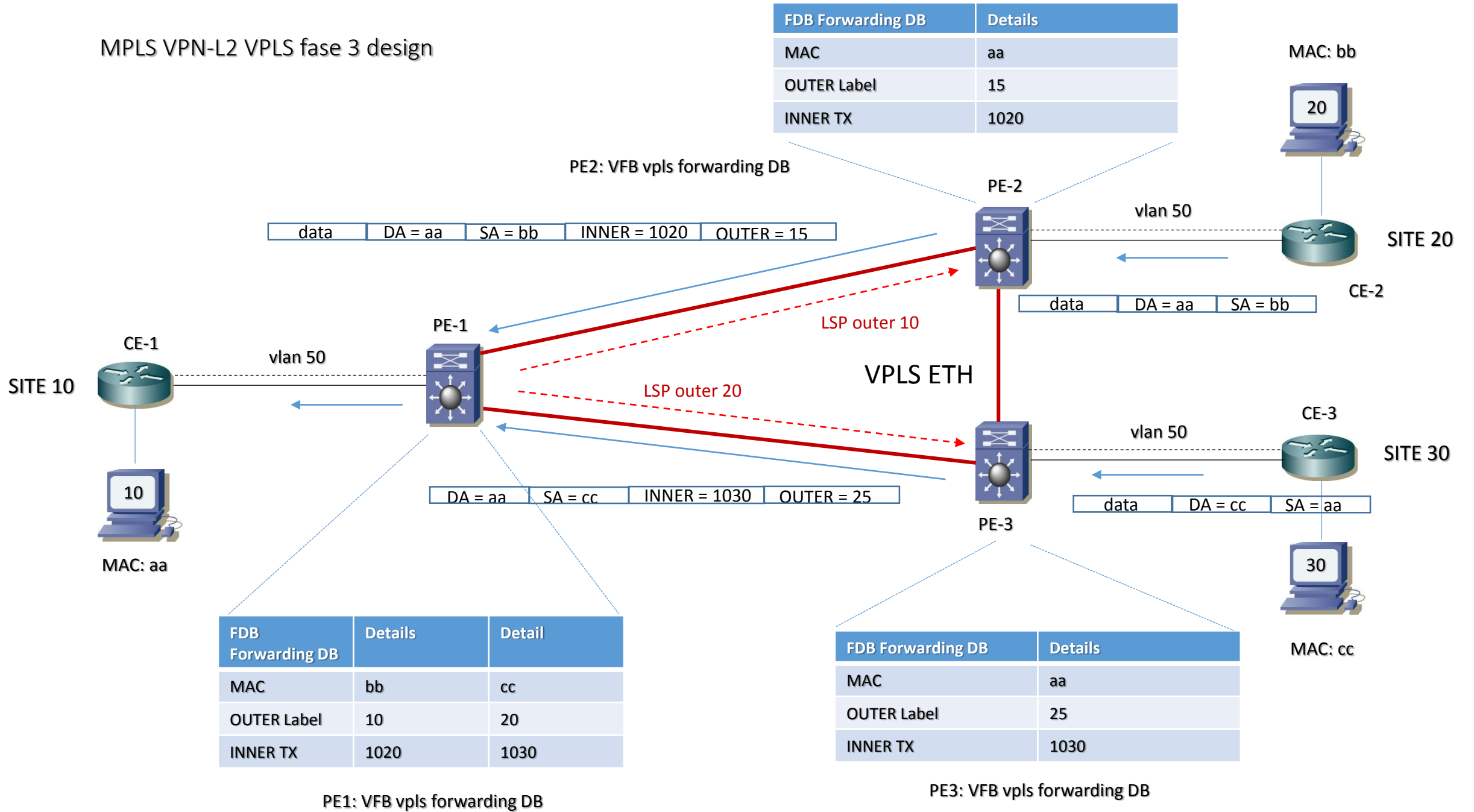
VE ID	INNER TX	INNER RX	OUTER
20	2010	1020	10
30	3010	1030	20

PE1: VFT vpls forwarding Table

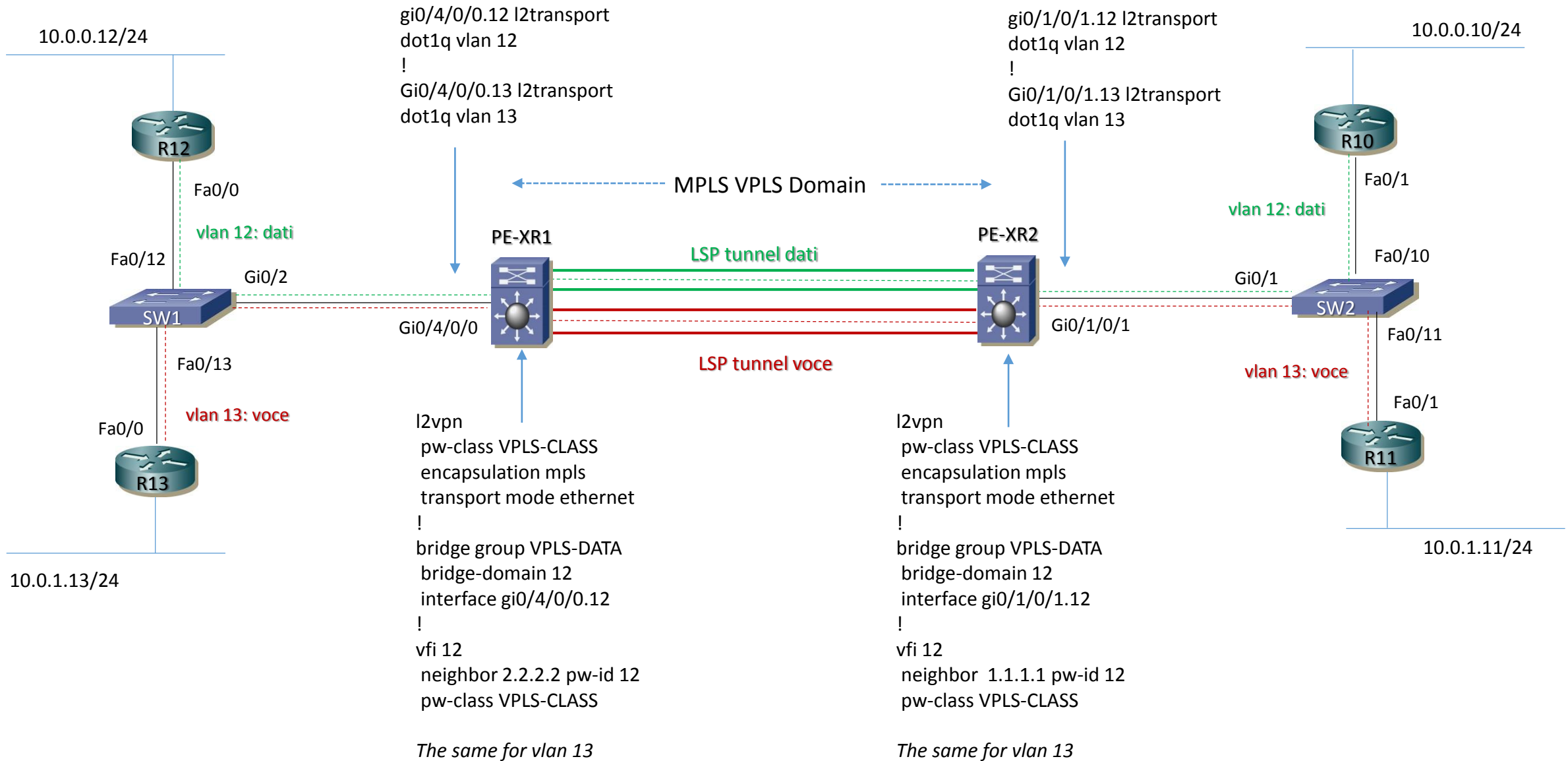
FDB Forwarding DB	Details
MAC	aa
OUTER Label	25
INNER TX	1030

PE3: VFB vpls forwarding DB

MPLS VPN-L2 VPLS fase 3 design



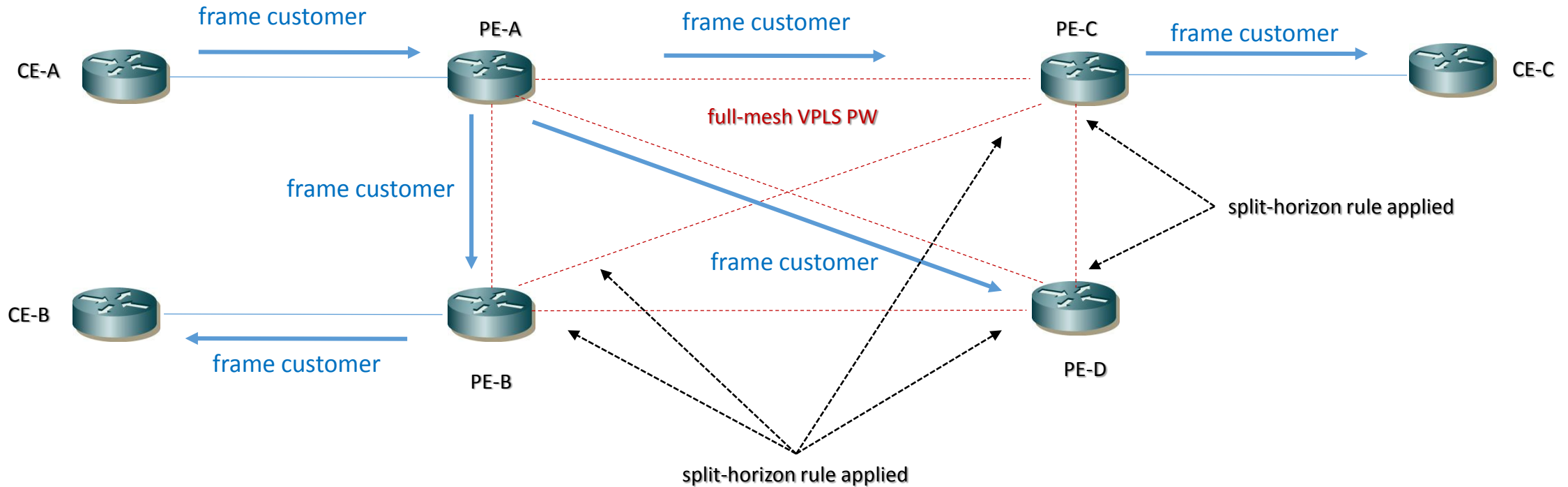
MPLS VPN-L2 VPLS example configuration



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

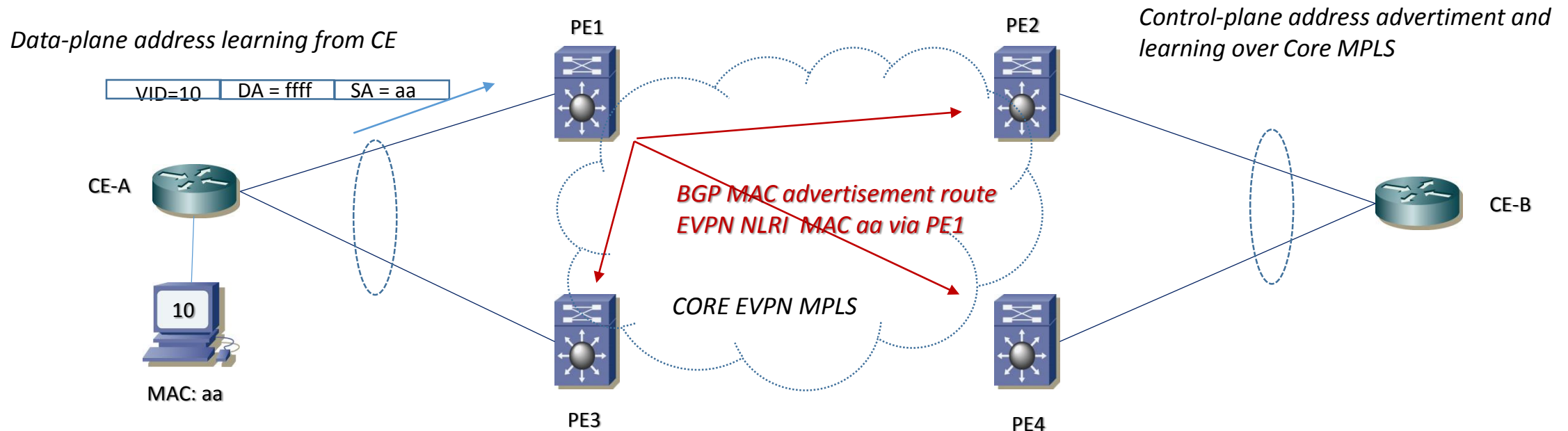


MPLS VPN-L2 VPLS load-balancing EVPN

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 distribuzione labels LDP protocols via BGP (RFC3107 carry label information)

La distribuzione associazione FEC-label può essere eseguita via BGP opportunamente esteso.

L'associazione FEC-label viene trasportata nel campo NLRI (Network Layer Reachability Information) parte dell'attributo MP-BGP «MP-REACH-NLRI», utilizzato per gli annunci delle destinazioni, i cui campi sono:

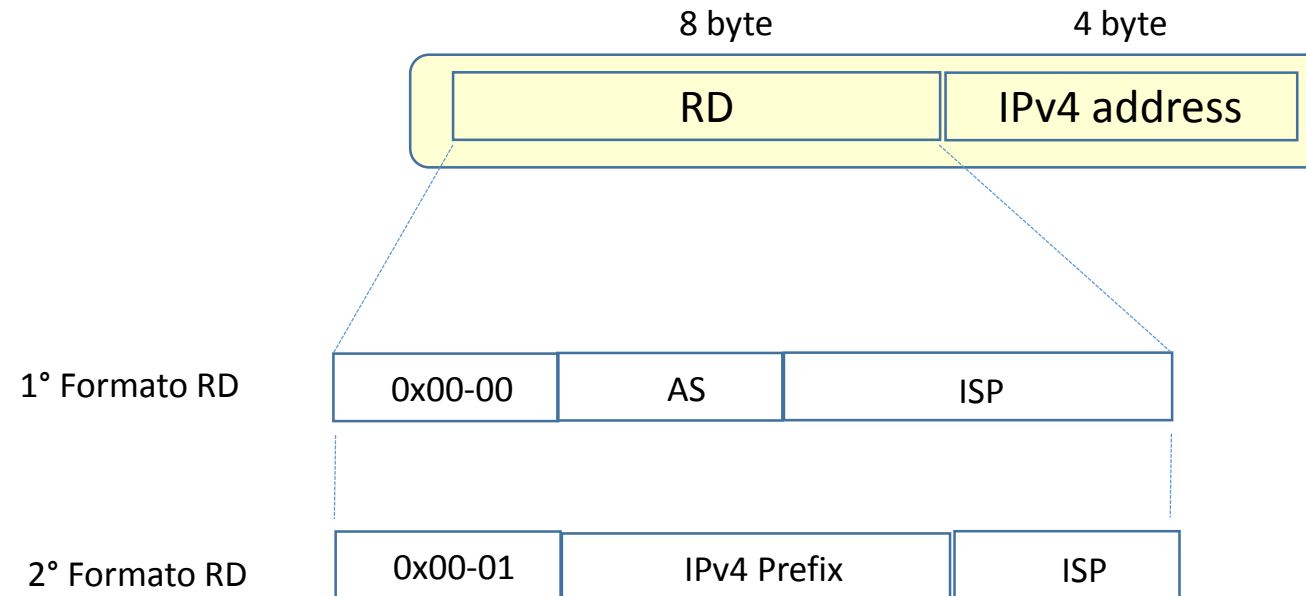
- **AFI (Address Family Identifier):** indica il tipo di indirizzo trasportato (per IP, AFI = 1)
- **Subsequence AFI (SAFI):** assume nel caso di trasporto di labels il valore 4 (SAFI = 4)
- **Network Address of Next-Hop:** identifica il nodo che ha generato l'annuncio
- **NLRI (Network Layer Reachability Information):** contiene i seguenti valori:
 - **Length:** lunghezza in bit del campo « prefix + label »
 - **Label:** contiene uno o più elementi MPLS senza il campo TTL
 - **Prefix:** l'indirizzo IP associato all'annuncio (formato RD + prefix IP)

NOTA: l'utilizzo di BGP per la distribuzione di associazione FEC-labels è utile in un contesto multi-AS, ossia per destinazioni esterne all'AS (Autonomous System) e/o destinazioni di clienti/users esterne al dominio AS.

MPLS BGP VPNv4 address

L'unione di un Route Distinguisher ed il prefisso IP costituisce un indirizzo vpnv4, utilizzato per discriminare indirizzi IP di stessa classe all'interno di un backbone services provider multi-VRF (Virtual Routing Forwarding).

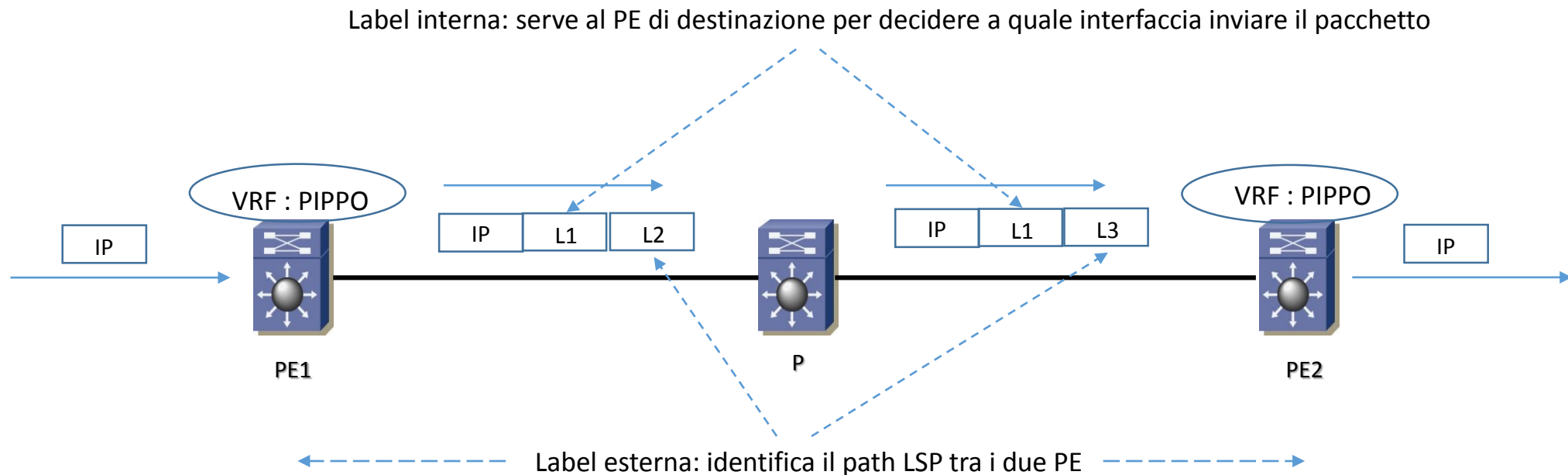
Il controllo ed il filtering di annunci tra multi-VRF è permesso attraverso una community BGP chiamata Route-Target che assume lo stesso formato del RD ma con diverso type (0x0002 per AS:ISP e 0x0102 per IPv4:ISP) attraverso politiche di import - export



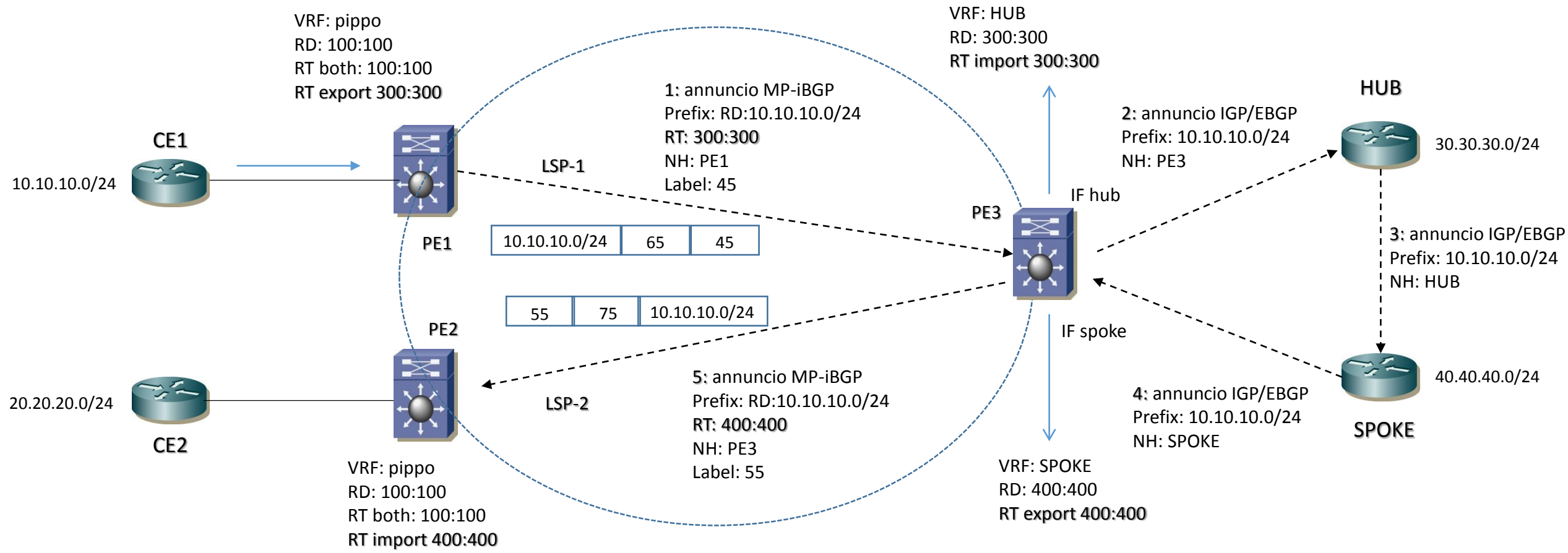
MPLS inoltro packets via BGP (RFC3107 carry label information)

L'inoltro dei pacchetti è possibile attraverso l'uso di due labels:

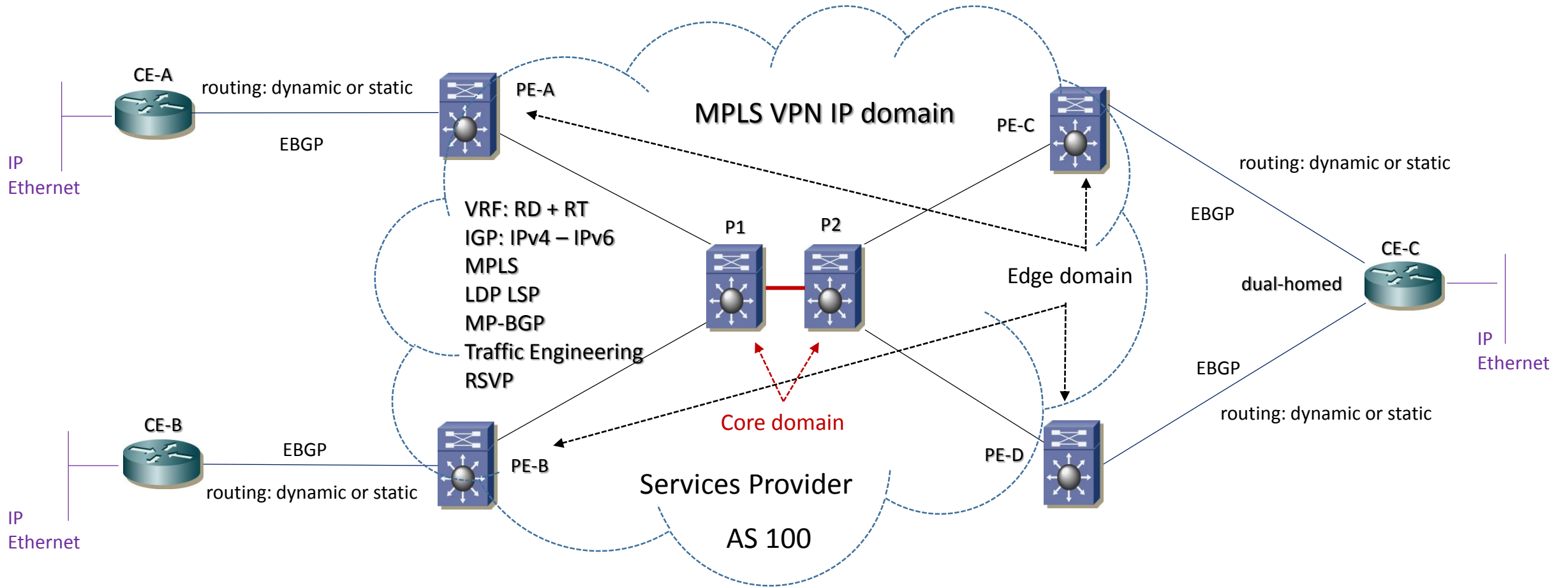
- La labels più esterna dello stack MPLS indica il path (LSP) da utilizzare per raggiungere il PE dove è attestato la rete di destinazione del pacchetto
- La labels più interna indica quale VRF consultare per il corretto inoltro del pacchetto (l'interfaccia di uscita del PE destinazione dove inviare il pacchetto ; questa viene distribuita incapsulata negli annunci MP-BGP.



MPLS intro packets via BGP (RFC3107 carry label information) example design



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.

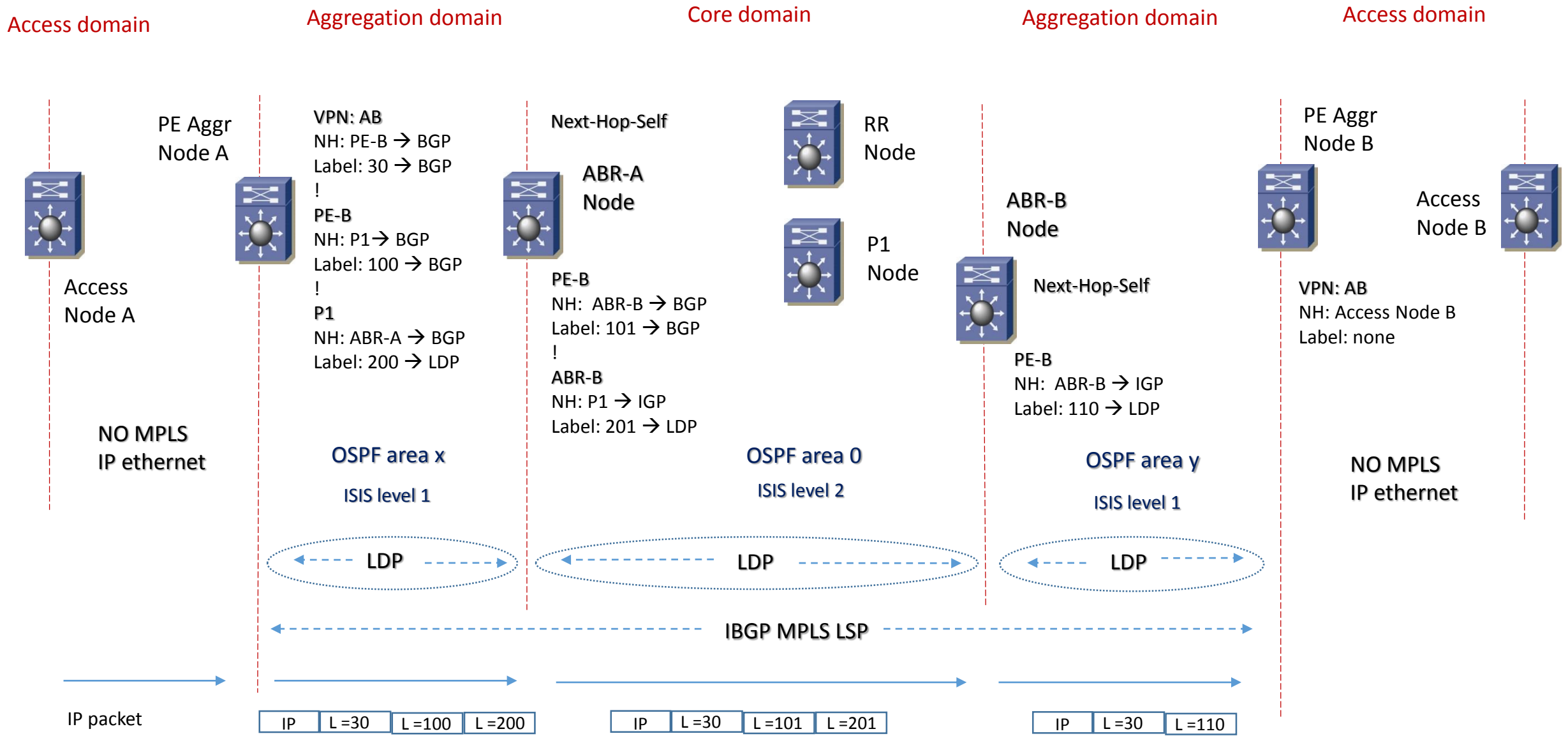
Questo mapping «label-prefix» è trasportato negli update message di BGP che contengono le informazioni della route.

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 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.

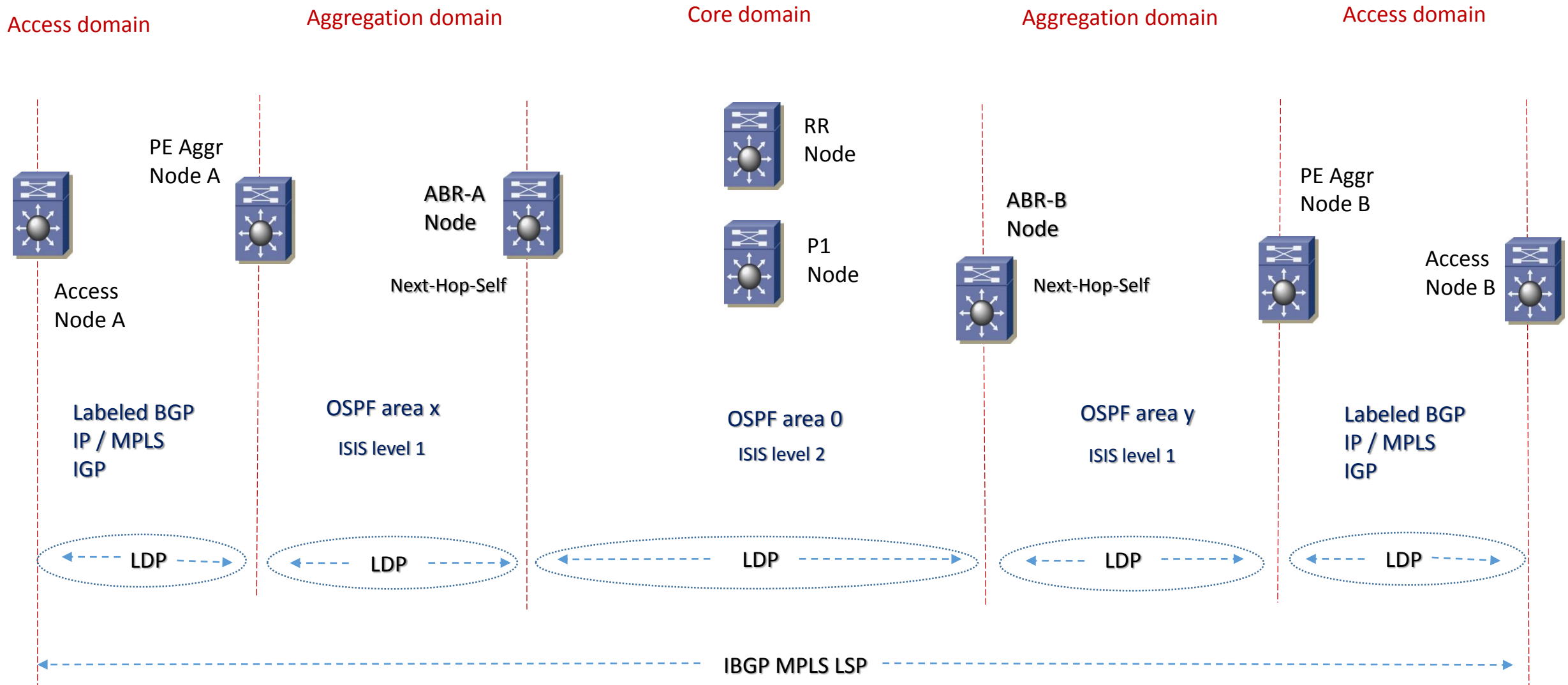
A differenza del modello precedente (labeled Core Aggregation, questo modello prevede MPLS enable anche a livello di accesso, pertanto il BGP labeled è esteso sino all'accesso del dominio

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

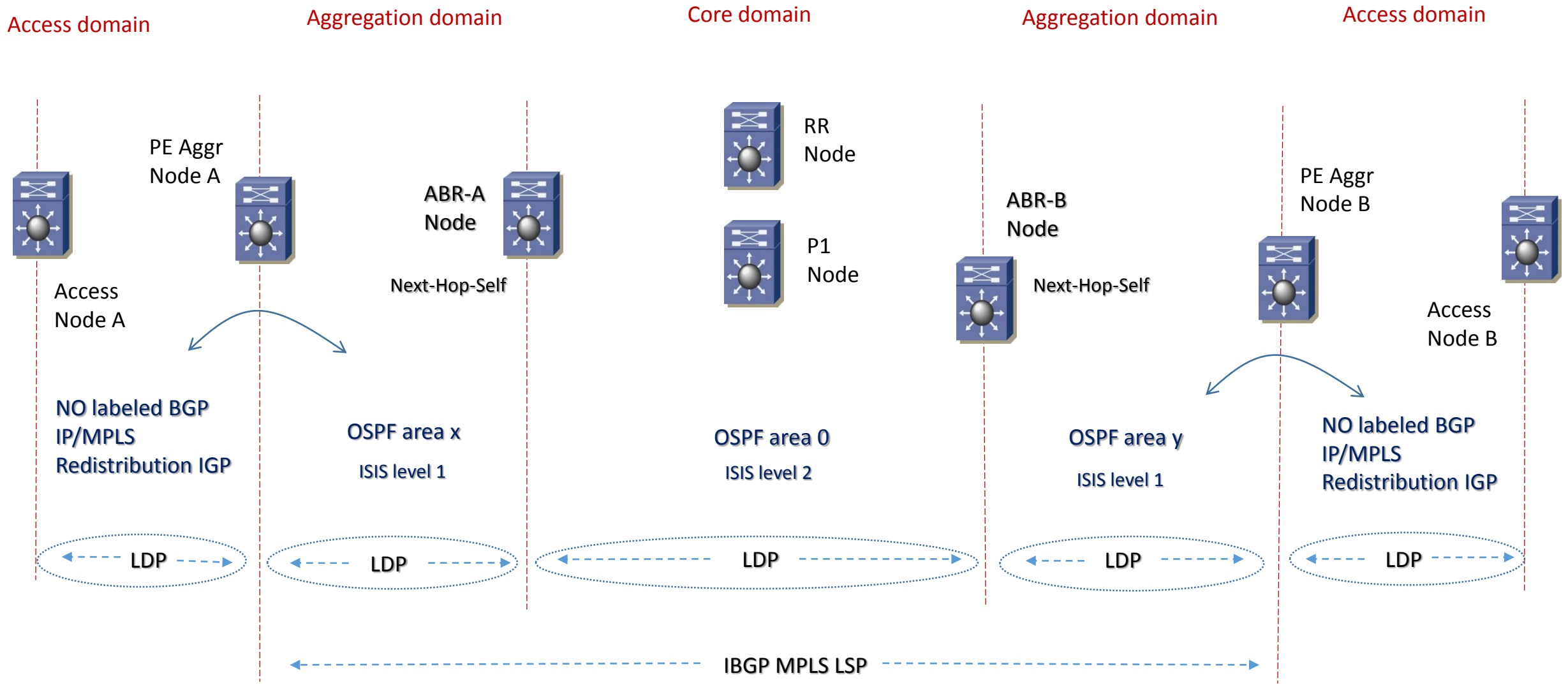
Unified MPLS utilizza il protocollo BGPv4 per scambiare informazioni con labels; quando BGP distribuisce una prefix/route porta con se anche una label MPLS che mappa la prefix/route.

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.

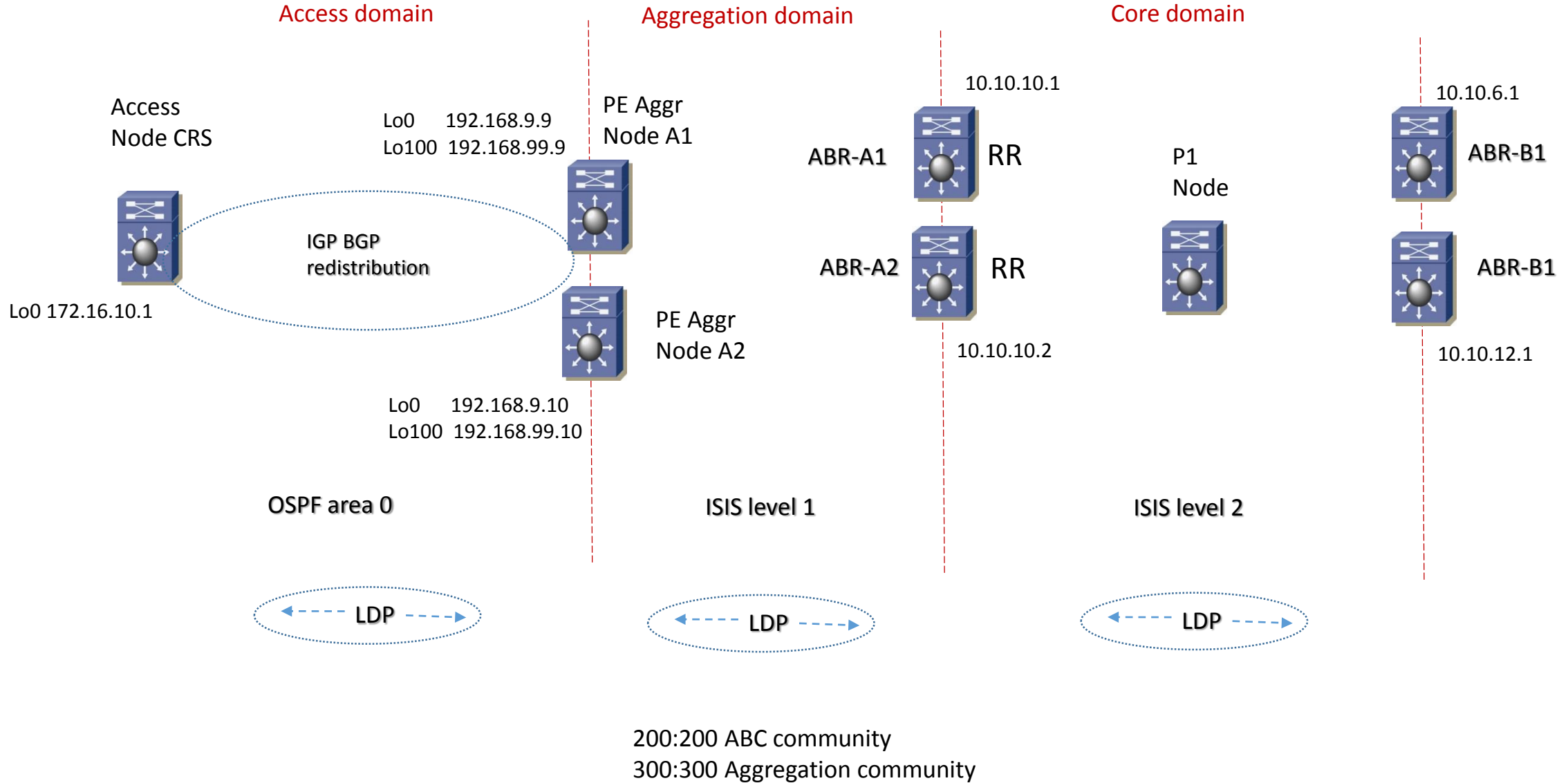
L'indirizzo di loopback del Access Node è redistribuito dentro IBGP domain.

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 community)

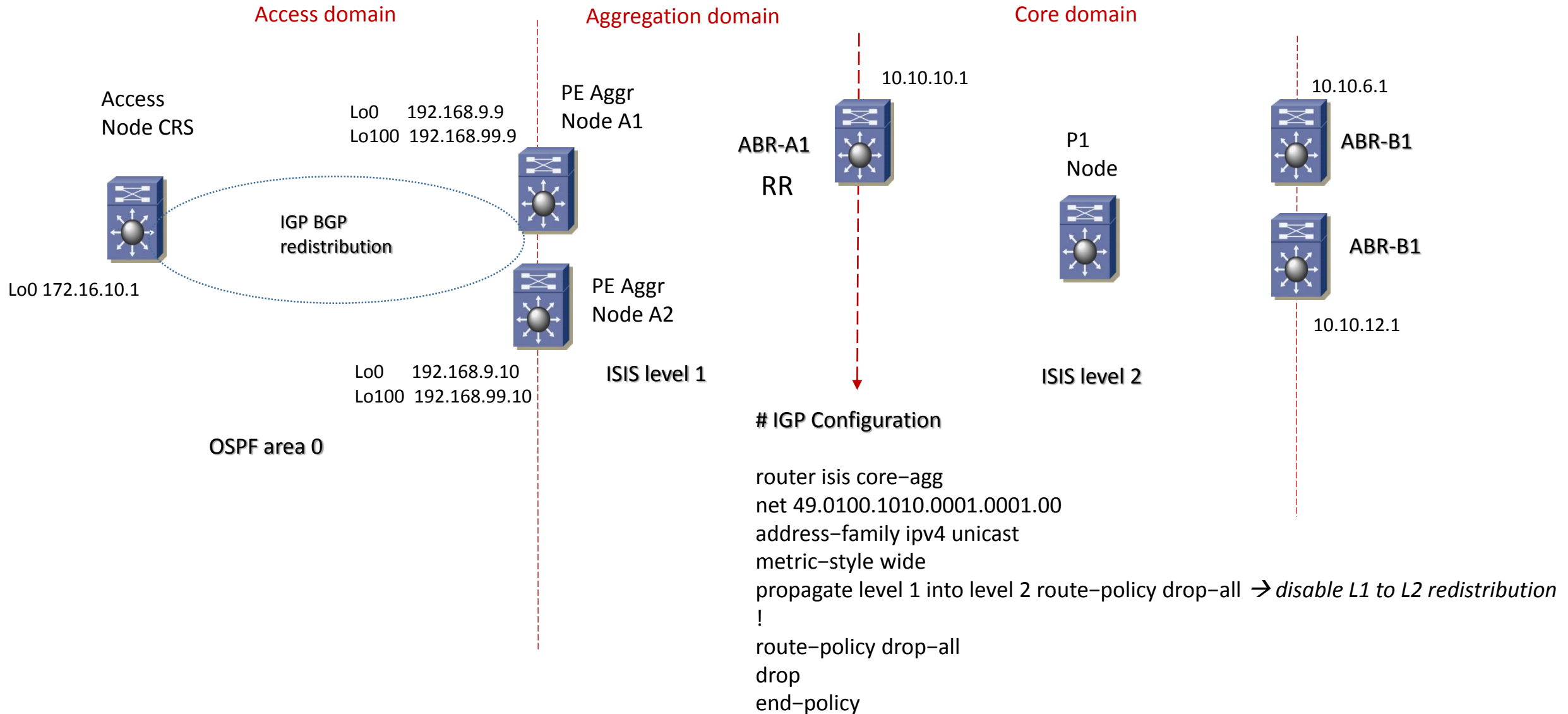
UNIFIED MPLS VPN-L3 (RFC 3107) carry label information in BGPv4 via Core and Aggregation with IGP redistribution on Access



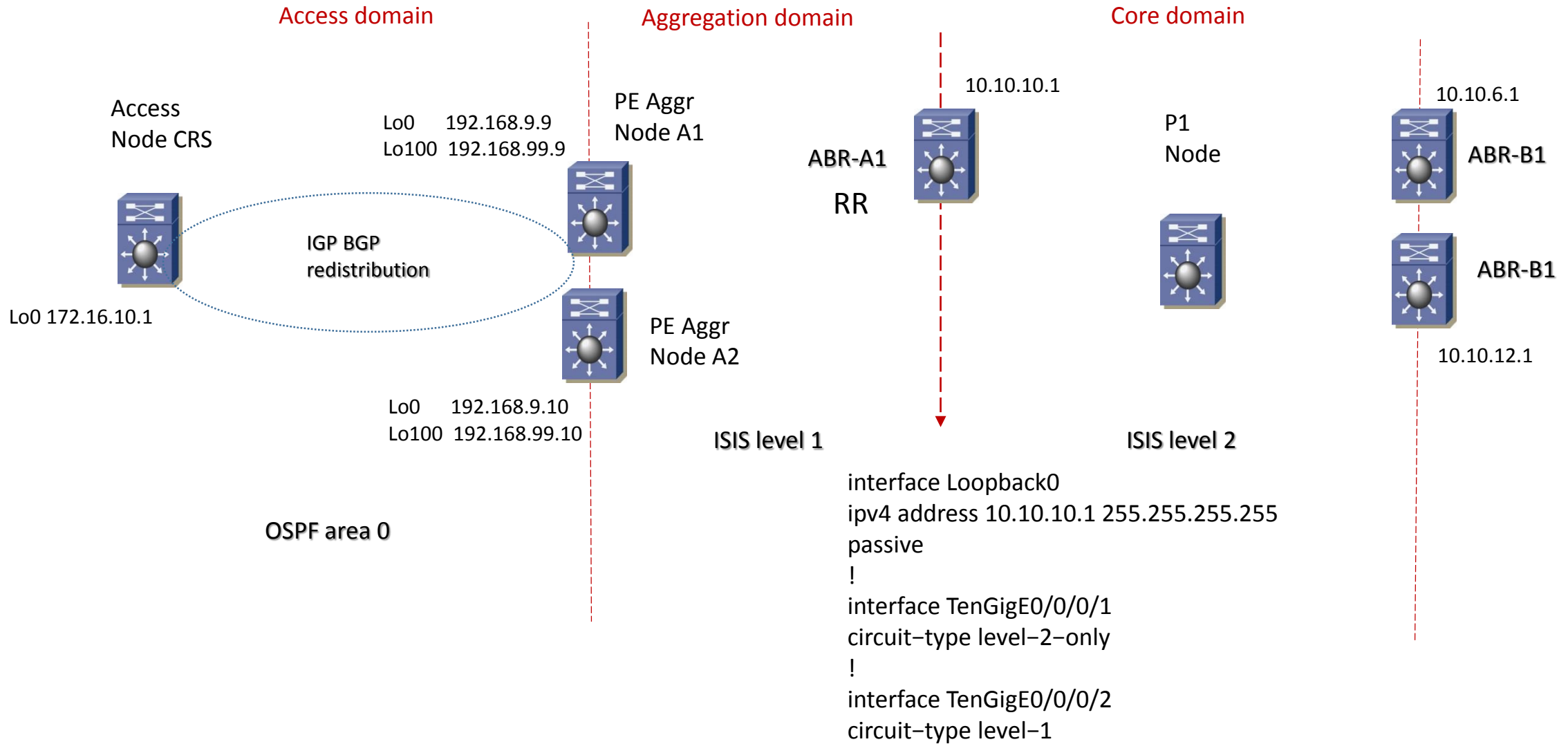
UNIFIED MPLS VPN-L3 (RFC 3107) Core and Aggregation with IGP redistribution on Access Configuration Example



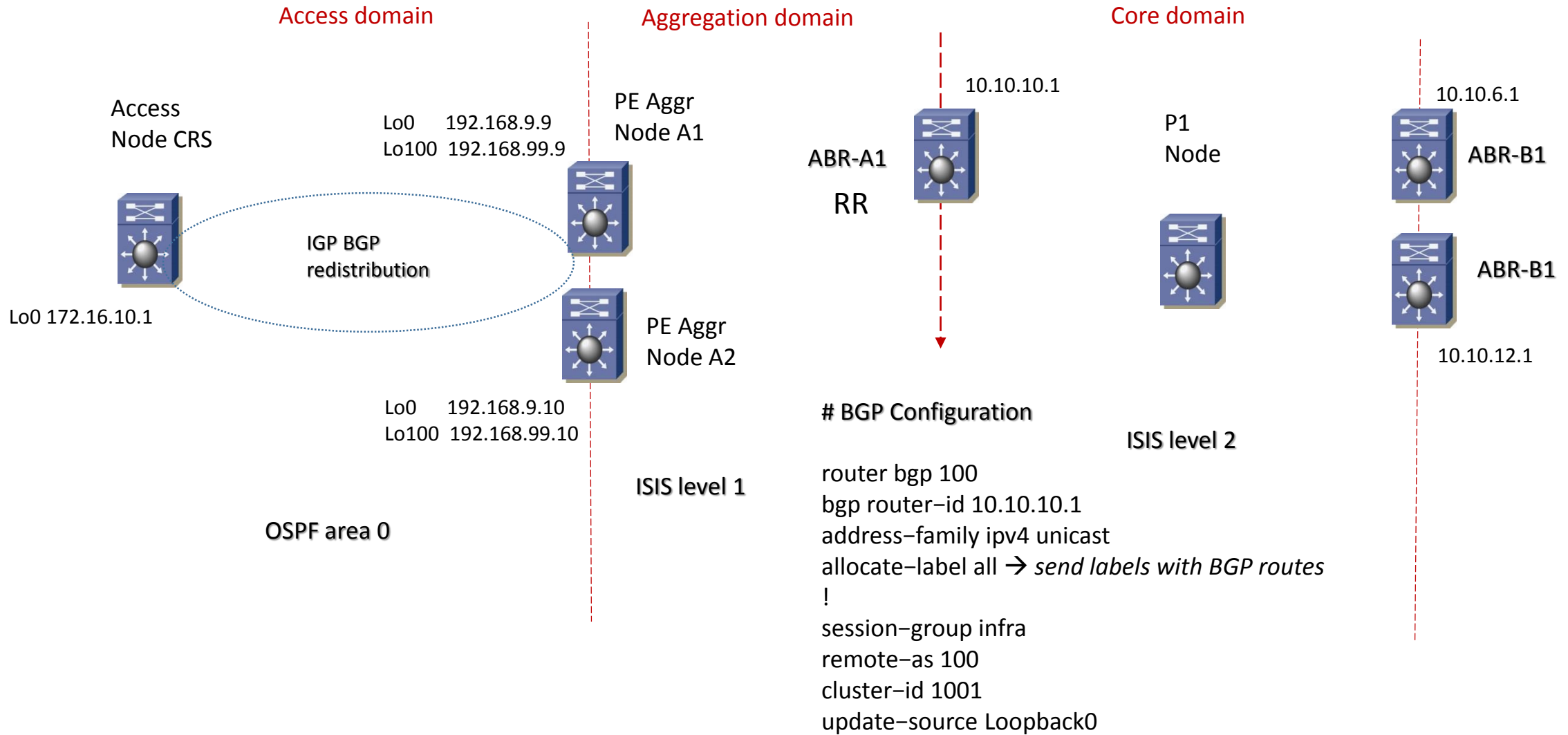
UNIFIED MPLS VPN-L3 (RFC 3107) Core and Aggregation with IGP redistribution on Access Configuration Example



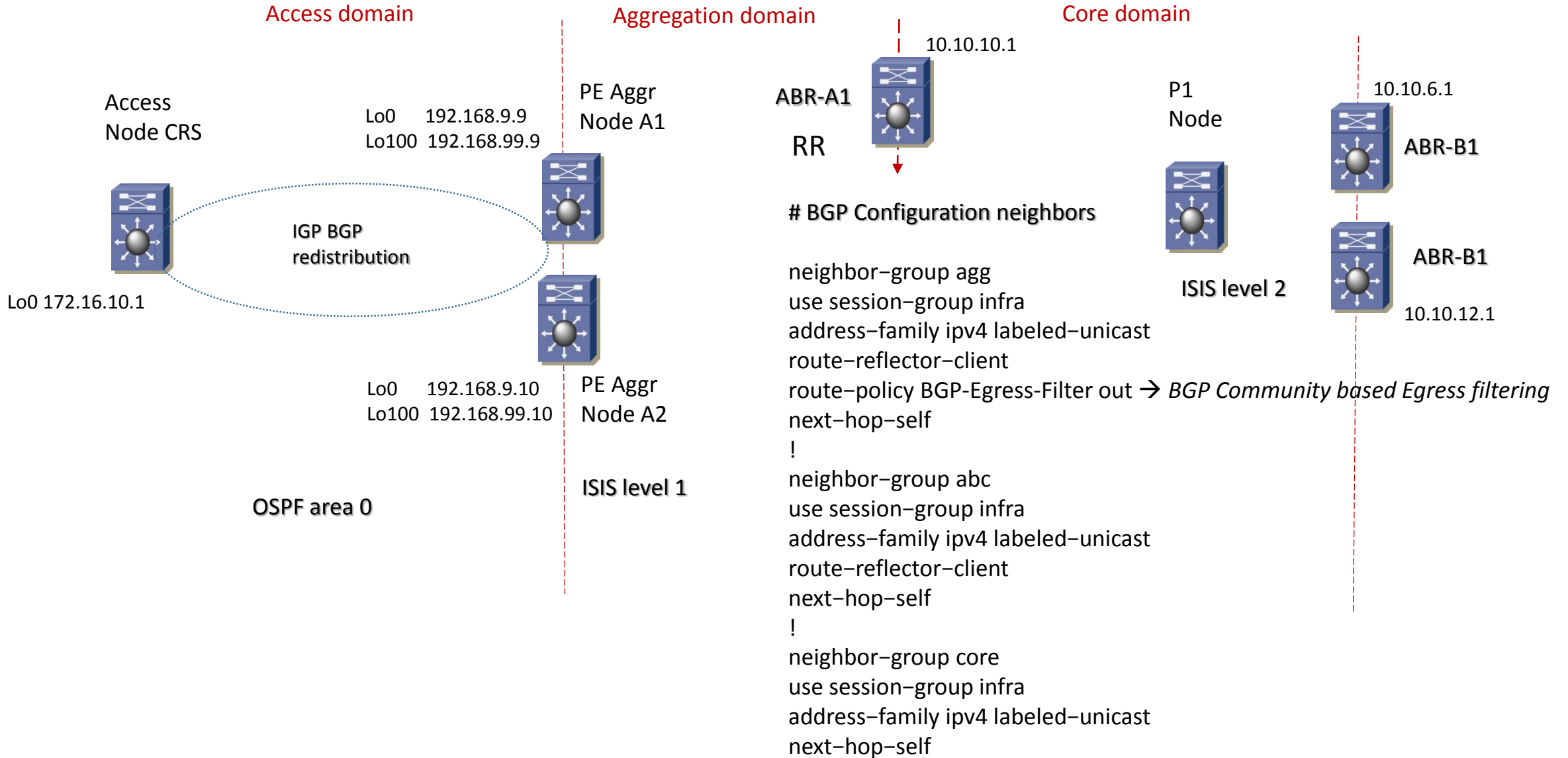
UNIFIED MPLS VPN-L3 (RFC 3107) Core and Aggregation with IGP redistribution on Access Configuration Example



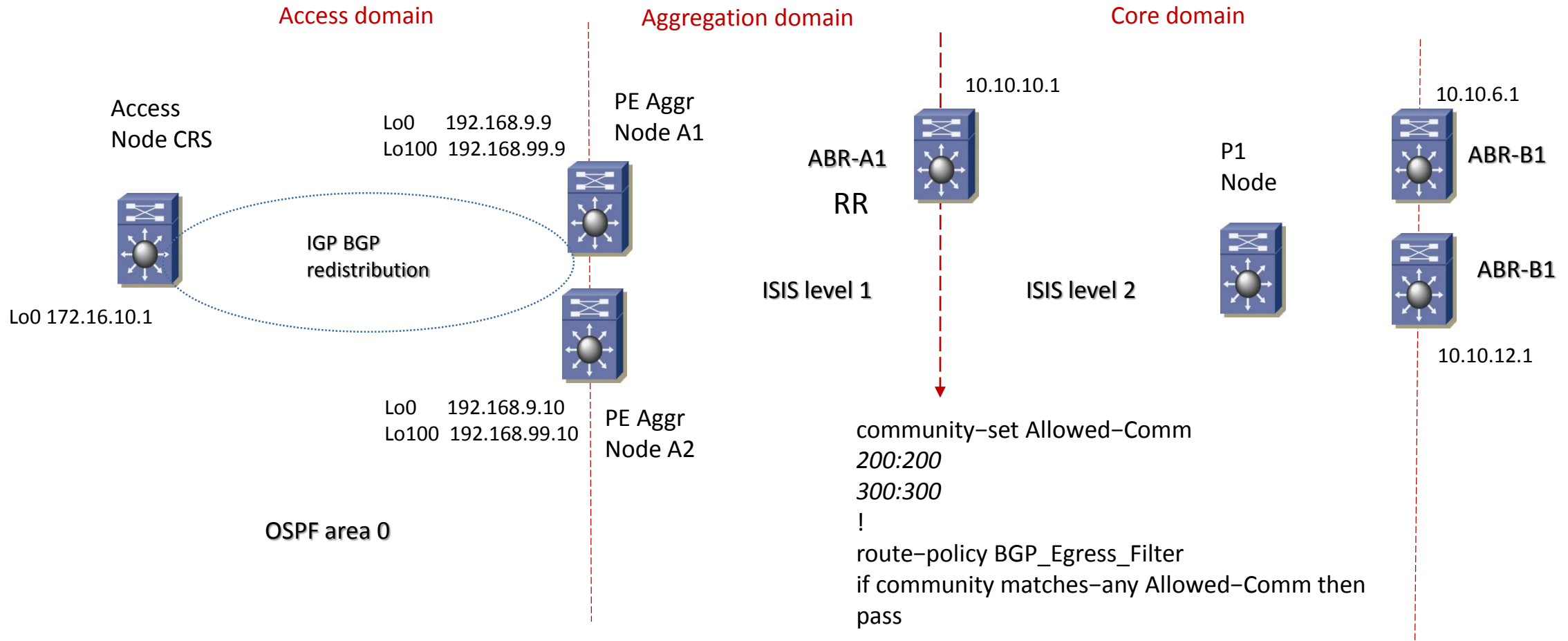
UNIFIED MPLS VPN-L3 (RFC 3107) Core and Aggregation with IGP redistribution on Access Configuration Example



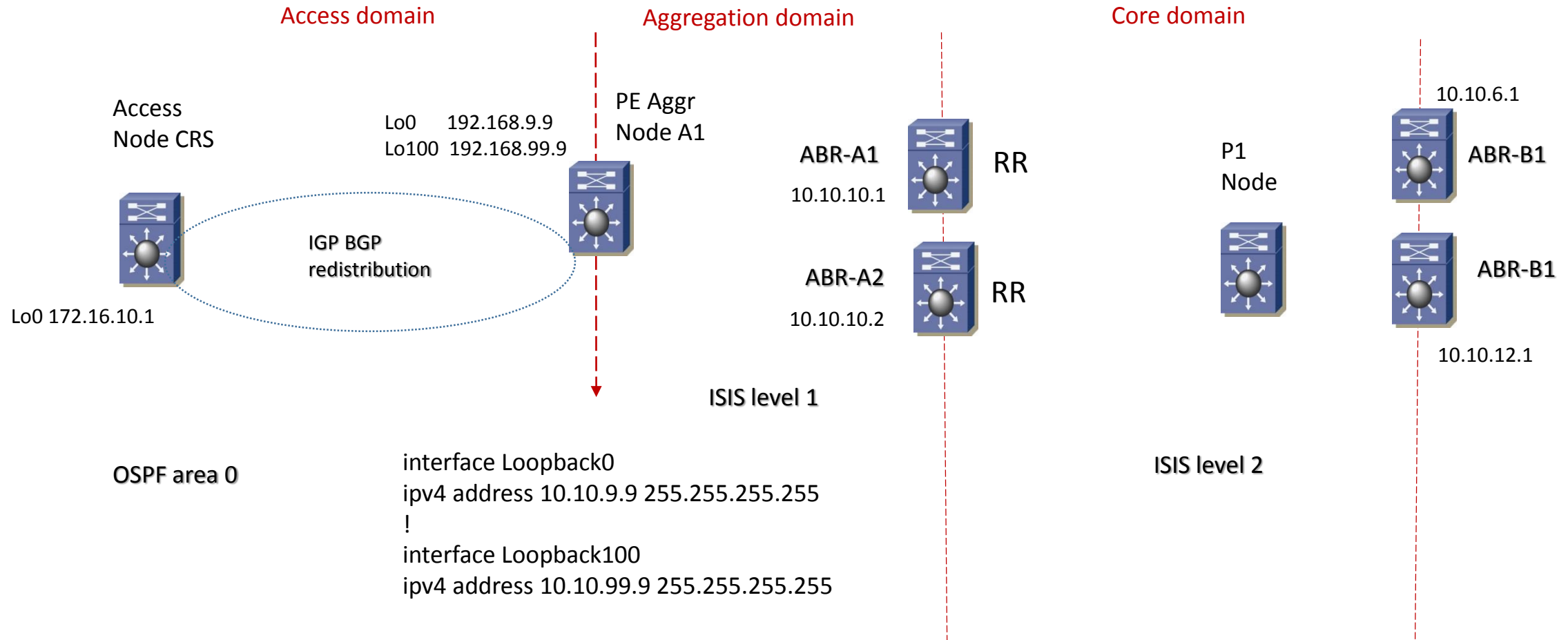
UNIFIED MPLS VPN-L3 (RFC 3107) Core and Aggregation with IGP redistribution on Access Configuration Example



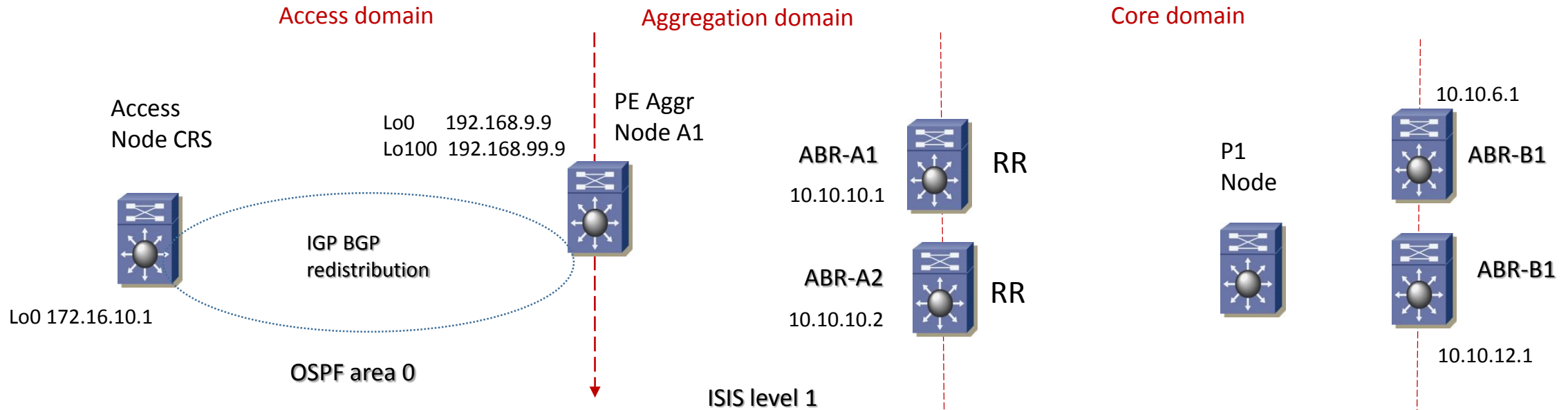
UNIFIED MPLS VPN-L3 (RFC 3107) Core and Aggregation with IGP redistribution on Access Configuration Example



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



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

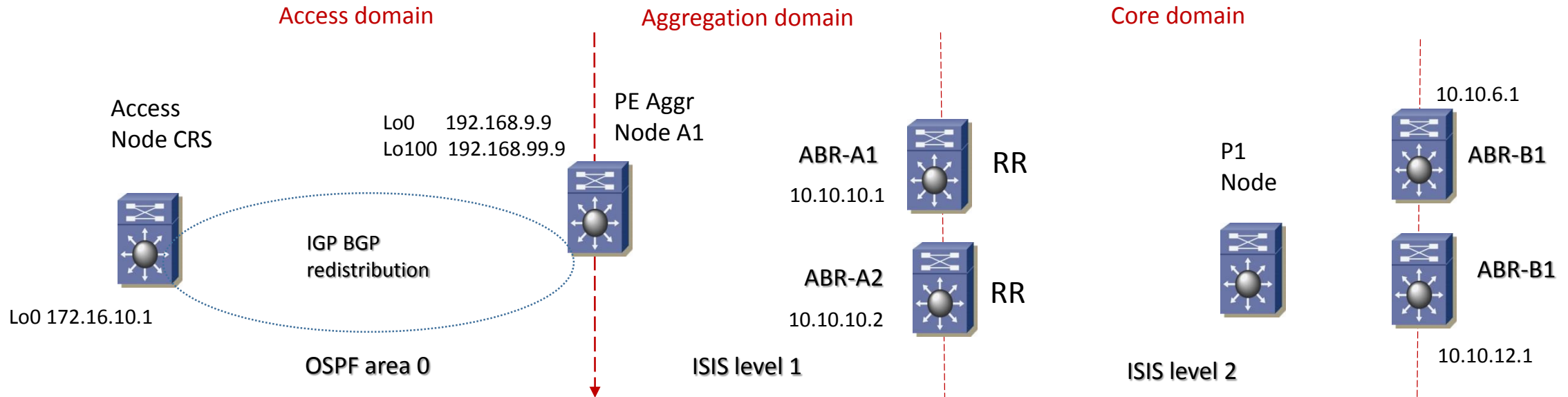


Aggregation IGP Configuration

IGP: ISIS

```
router isis core-agg
net 49.0100.1010.0001.9007.00
isis-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



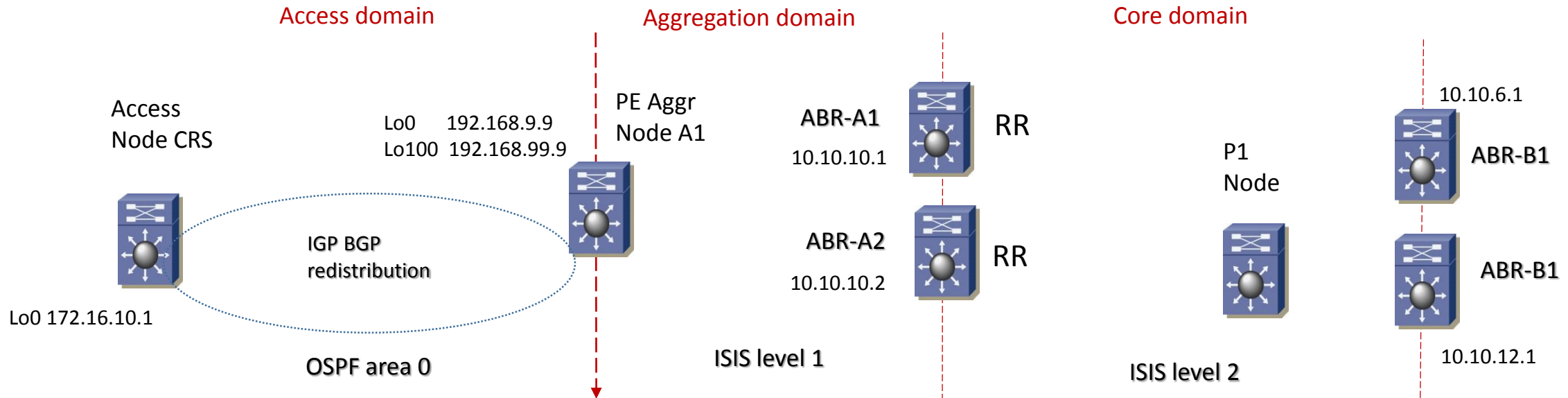
Access IGP Configuration

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
    
```

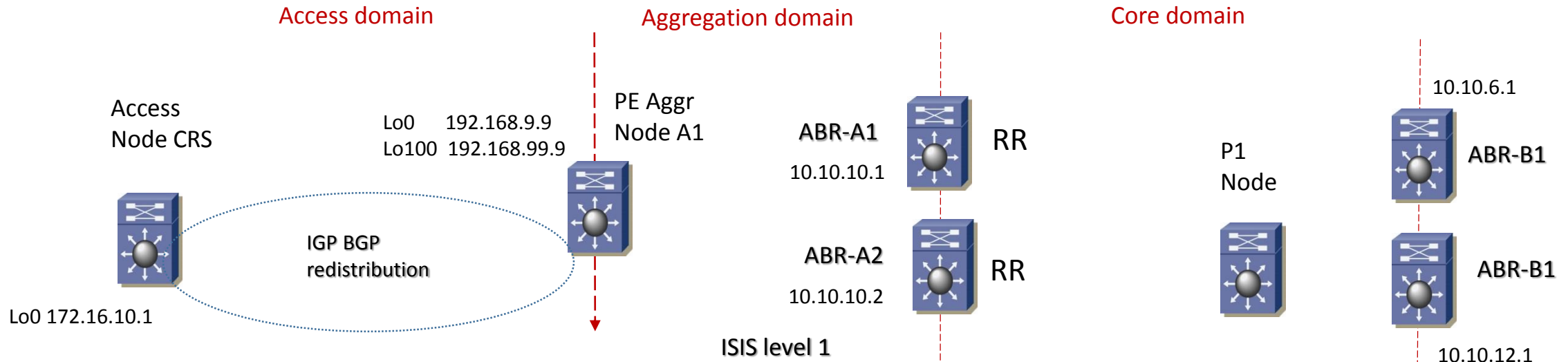

UNIFIED MPLS VPN-L3 (RFC 3107) Core and Aggregation with IGP redistribution on Access Configuration Example



```

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



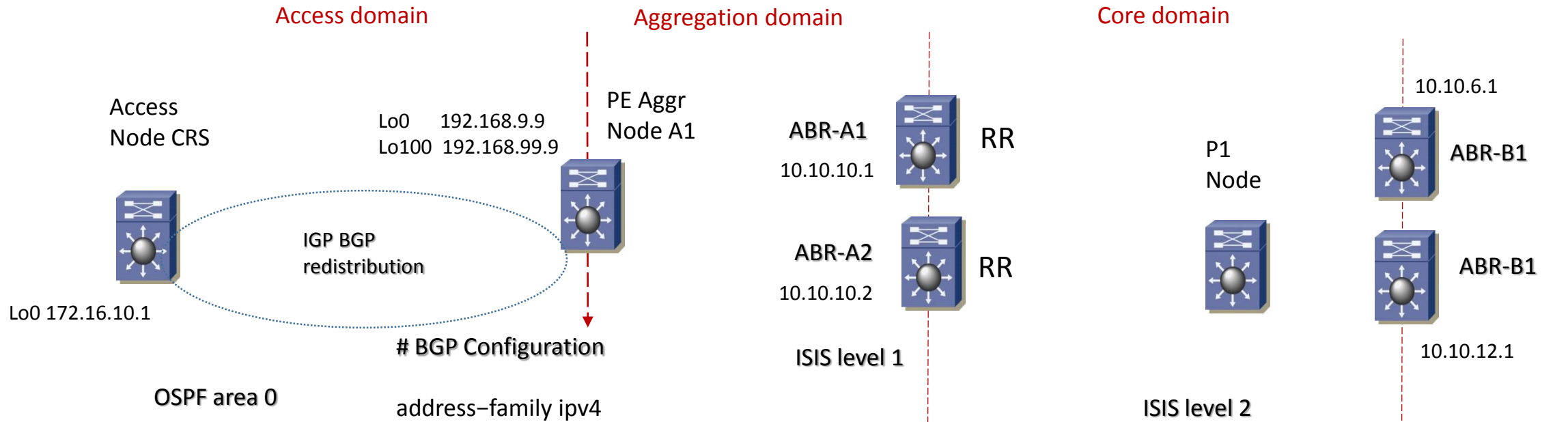
OSPF area 0

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



BGP Configuration

```

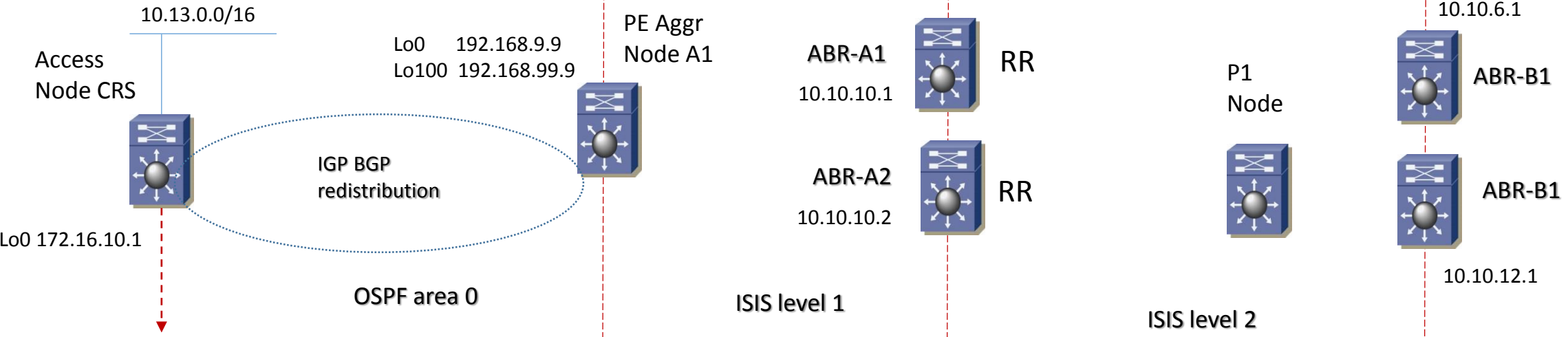
address-family ipv4
  bgp redistribute-internal
  network 192.168.9.9 mask 255.255.255.255 route-map AGG_Comm → advertise with Aggregation Community
  redistribute ospf 1 → redistribute Access IGP prefixes
  neighbor abr send-community
  neighbor abr next-hop-self
  neighbor abr send-label → send labels 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

Access domain

Aggregation domain

Core domain

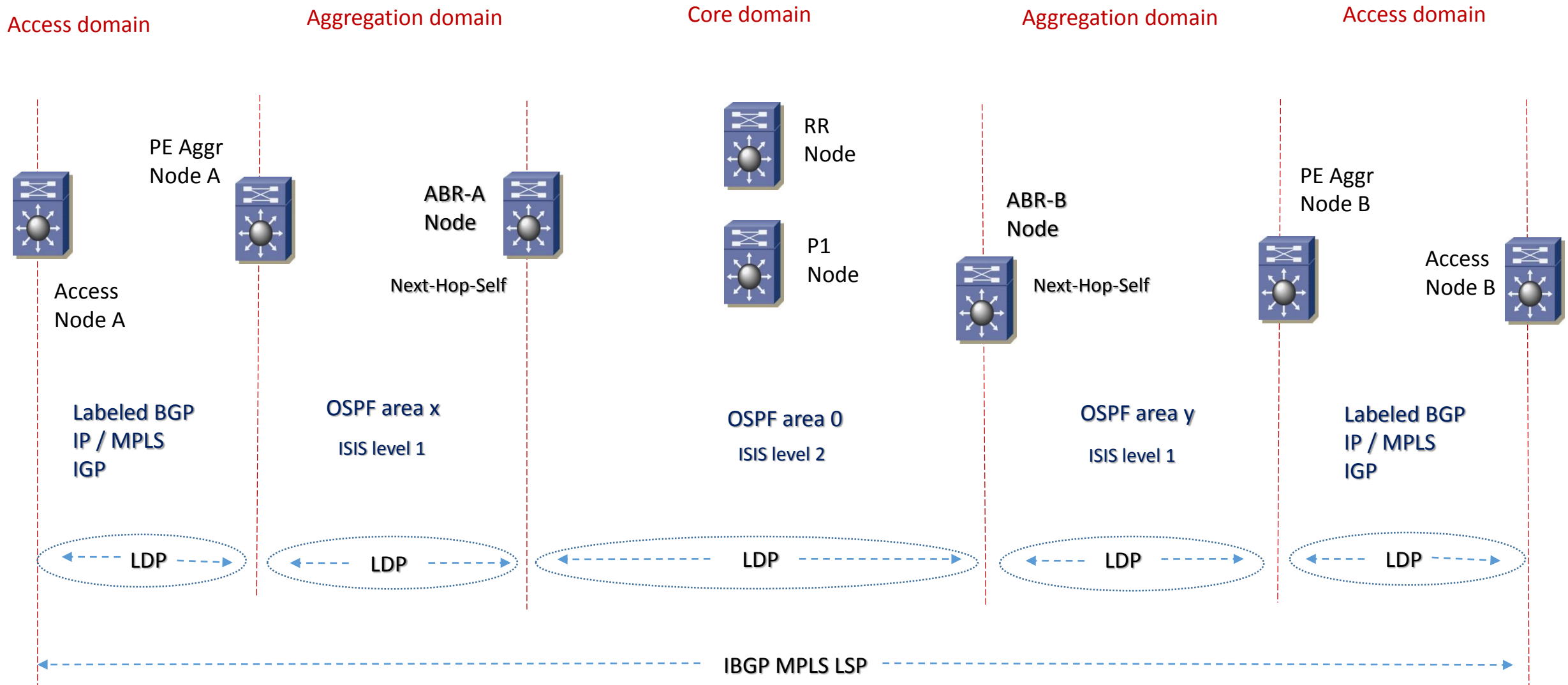


```
interface Loopback0
ip address 172.16.10.1 255.255.255.255
```

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
```

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

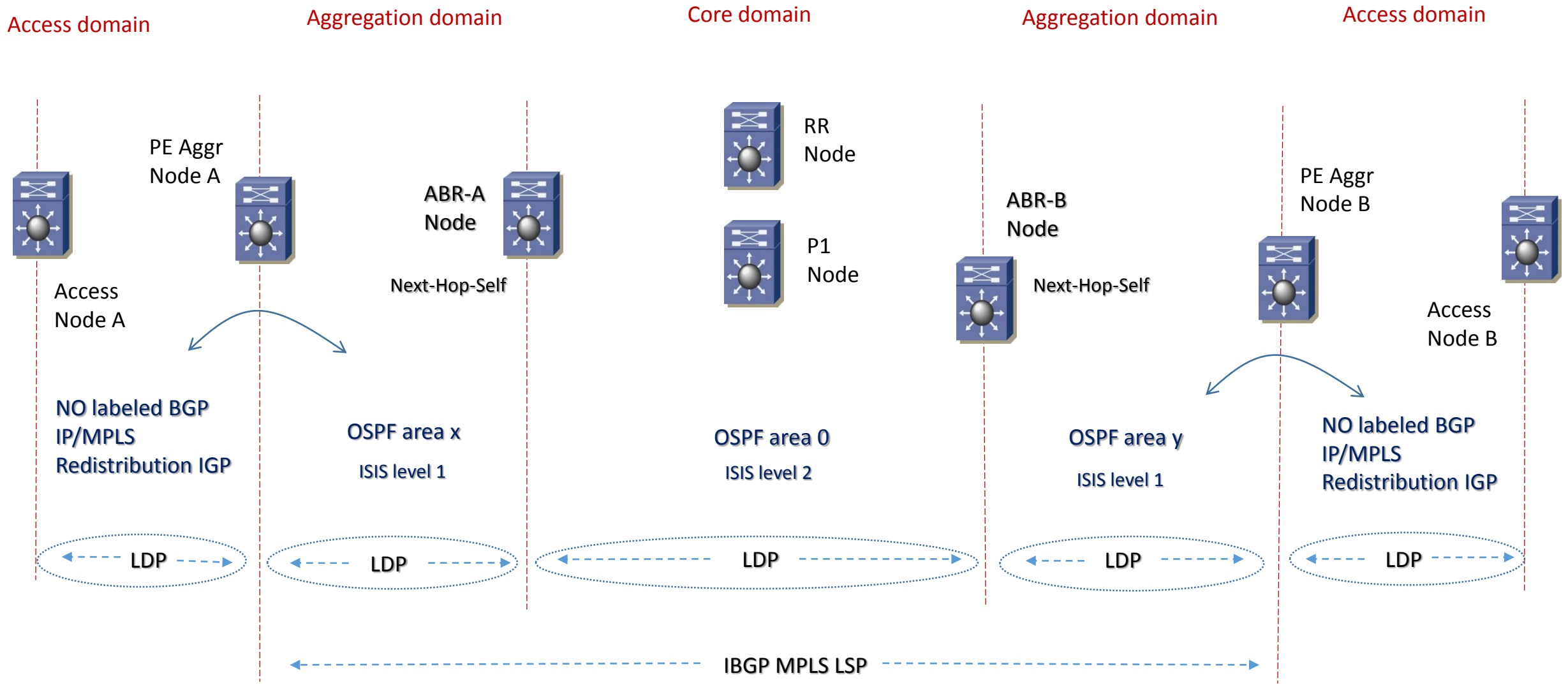
Unified MPLS utilizza il protocollo BGPv4 per scambiare informazioni con labels; quando BGP distribuisce una prefix/route porta con se anche una label MPLS che mappa la prefix/route.

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.

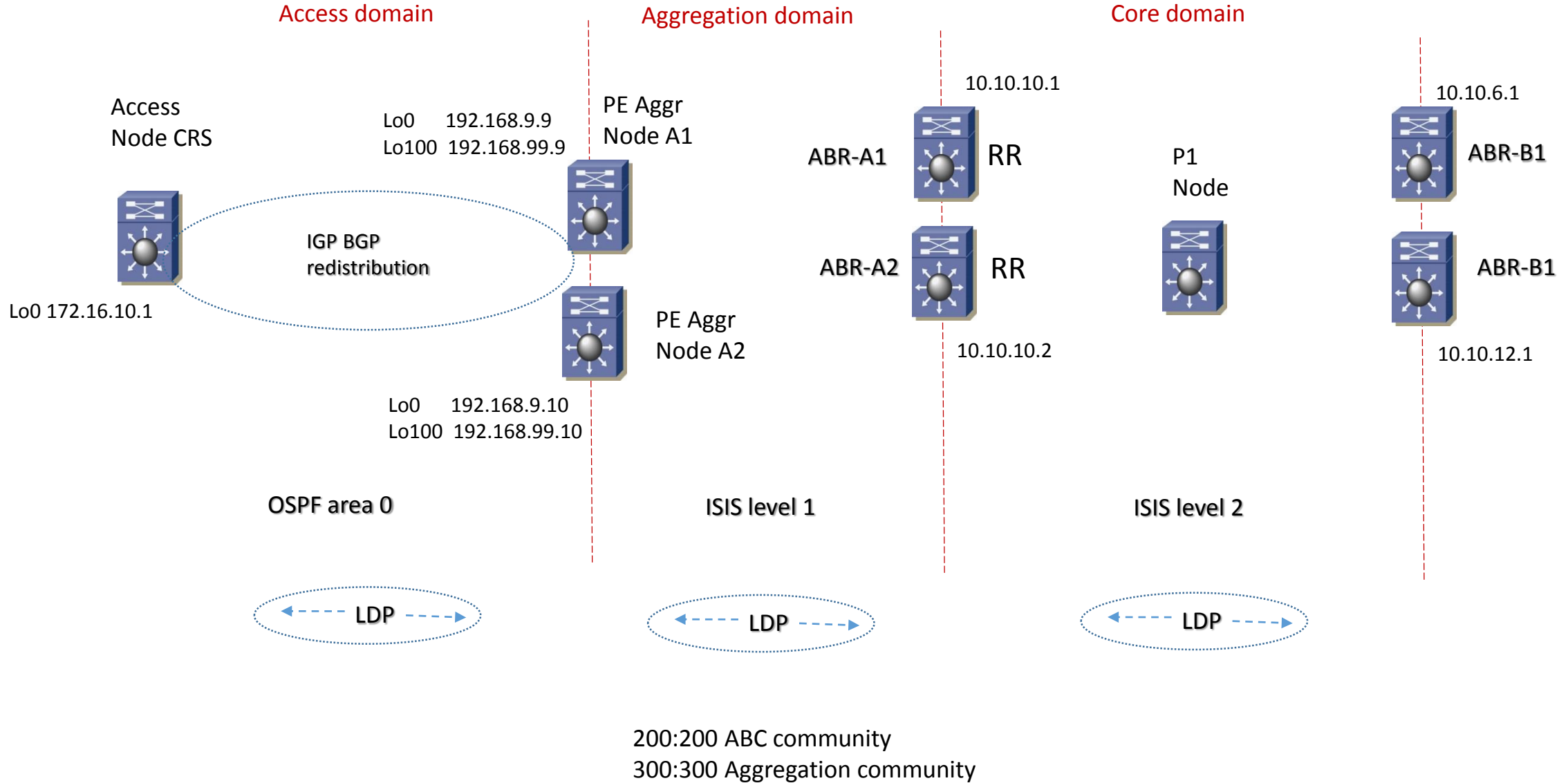
L'indirizzo di loopback del Access Node è redistribuito dentro IBGP domain.

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 community)

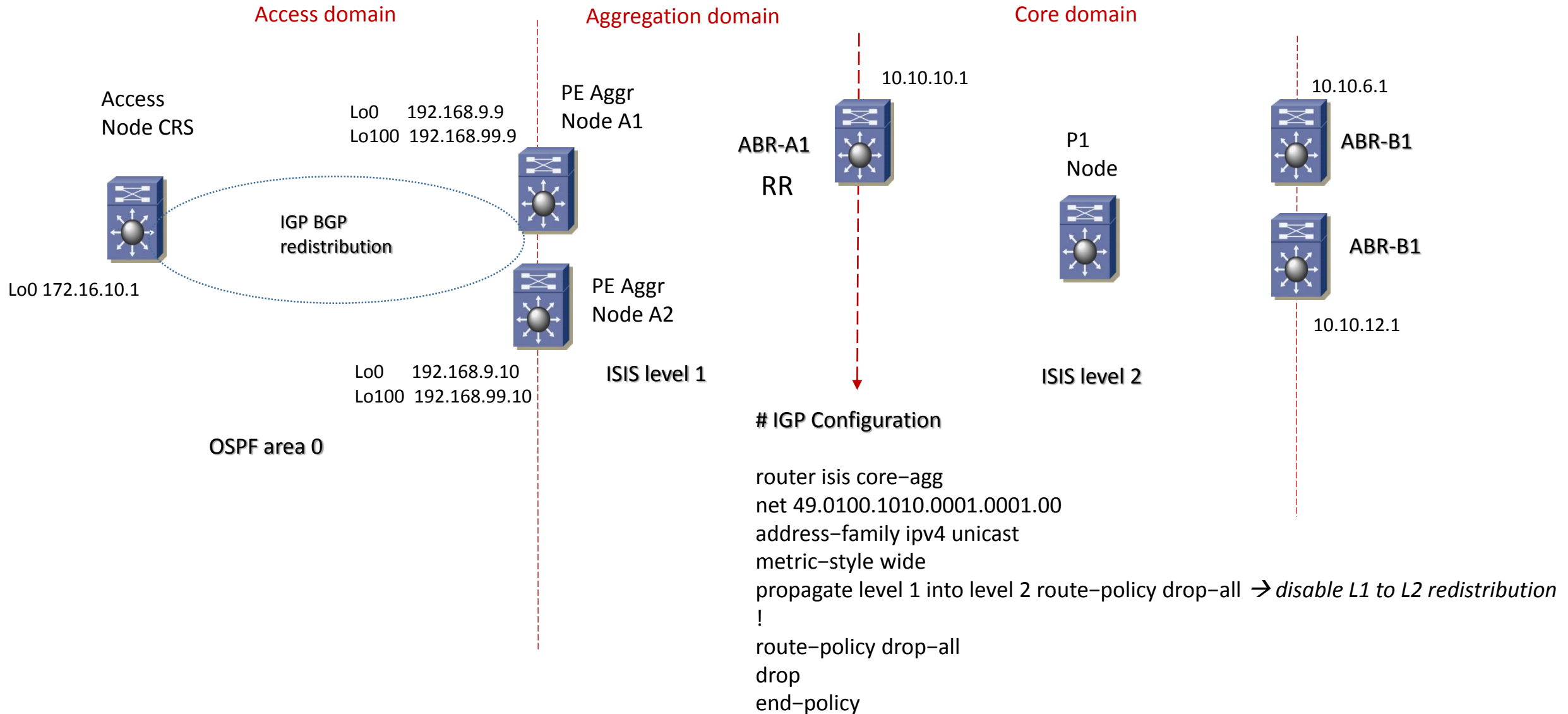
UNIFIED MPLS VPN-L3 (RFC 3107) carry label information in BGPv4 via Core and Aggregation with IGP redistribution on Access



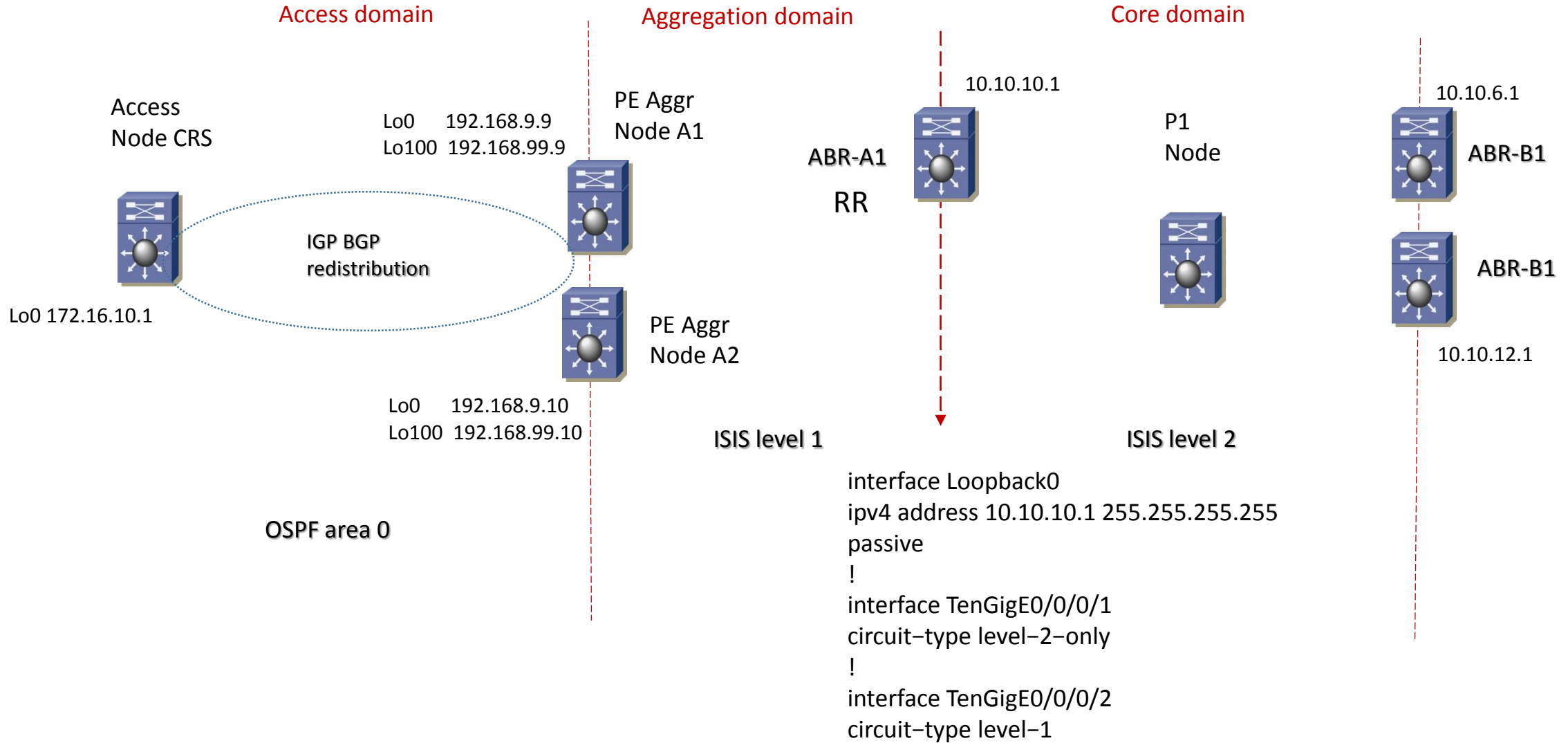
UNIFIED MPLS VPN-L3 (RFC 3107) Core and Aggregation with IGP redistribution on Access Configuration Example



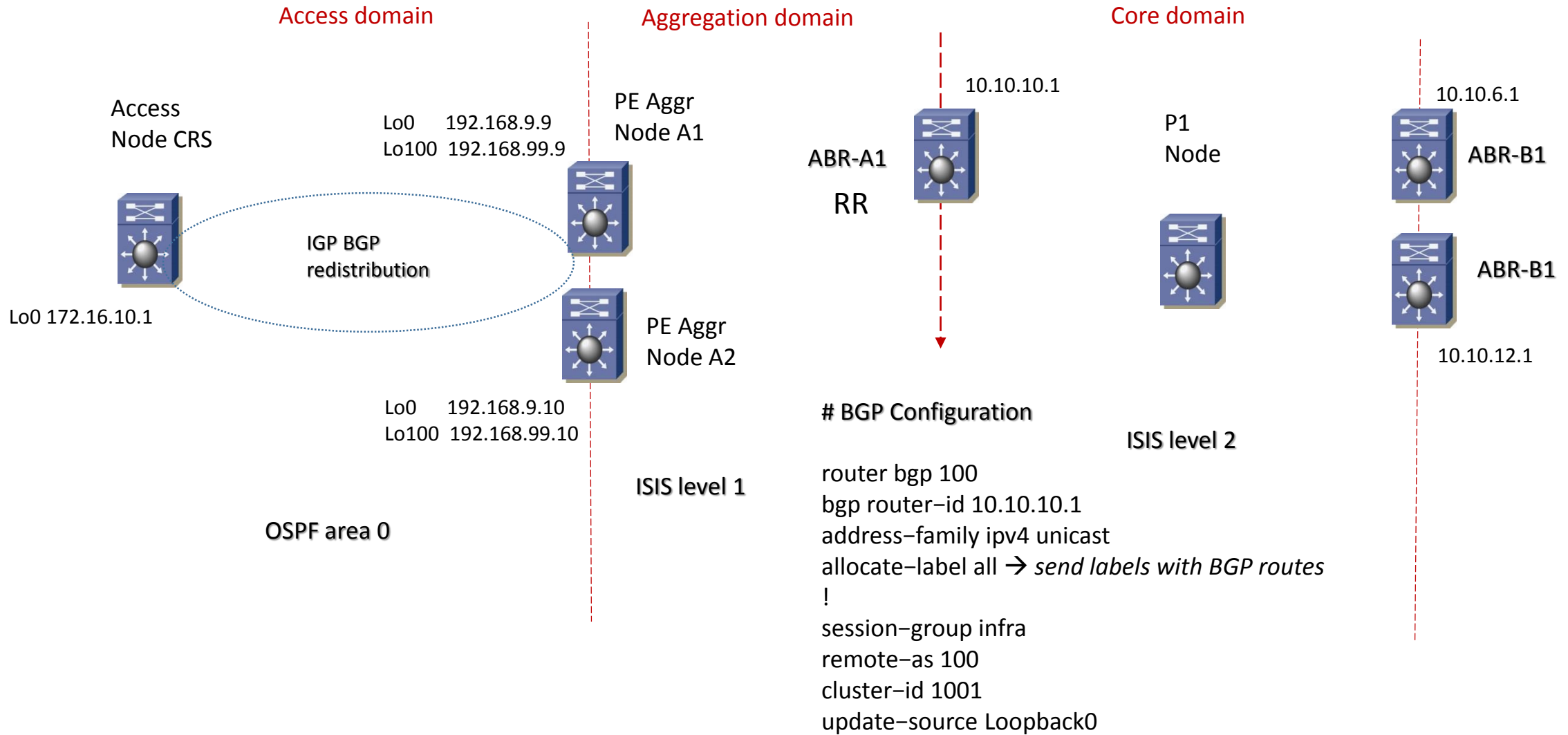
UNIFIED MPLS VPN-L3 (RFC 3107) Core and Aggregation with IGP redistribution on Access Configuration Example



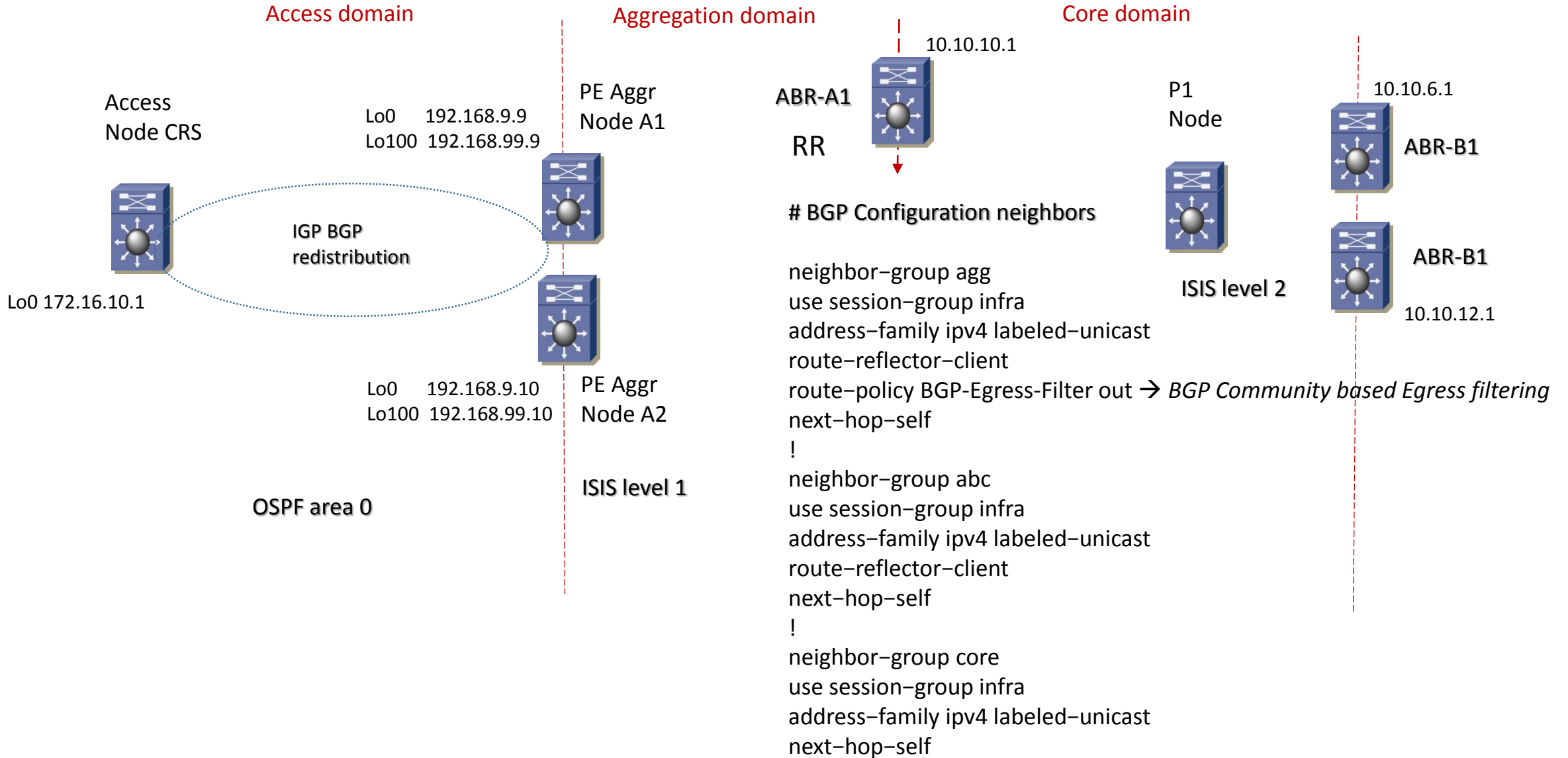
UNIFIED MPLS VPN-L3 (RFC 3107) Core and Aggregation with IGP redistribution on Access Configuration Example



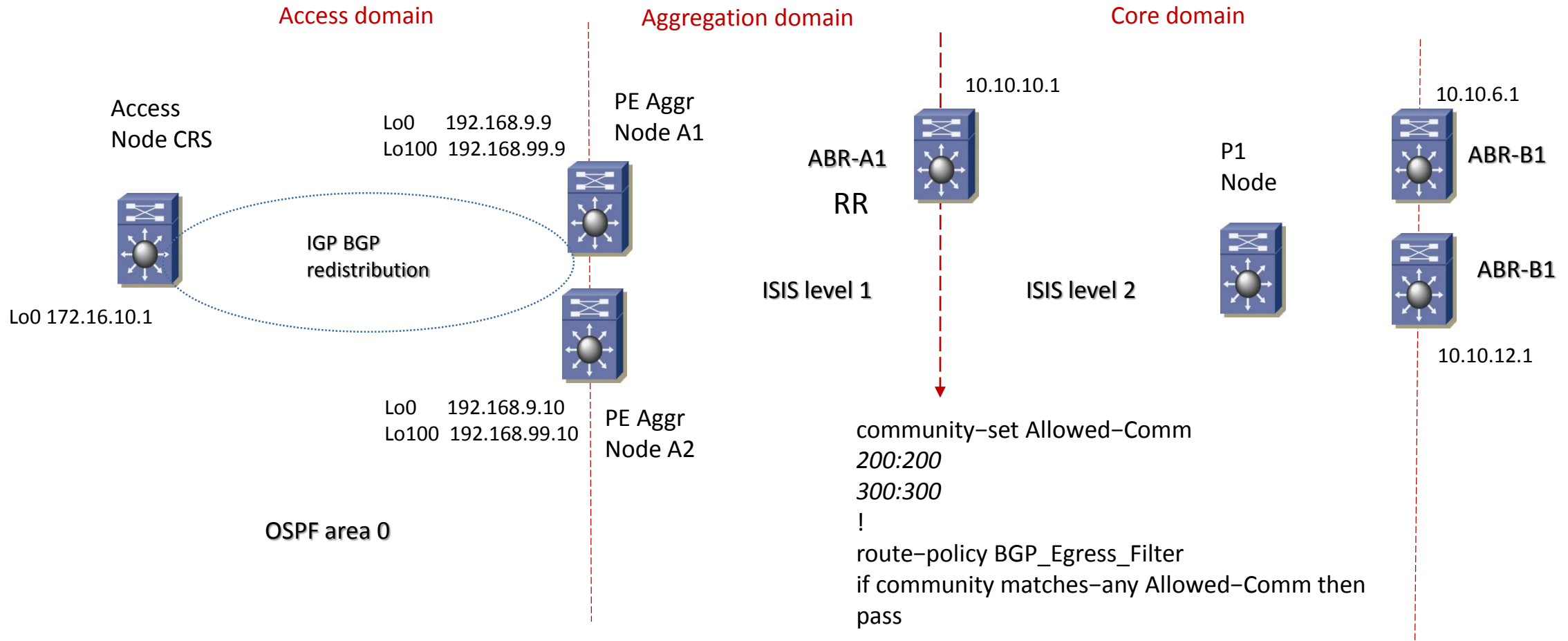
UNIFIED MPLS VPN-L3 (RFC 3107) Core and Aggregation with IGP redistribution on Access Configuration Example



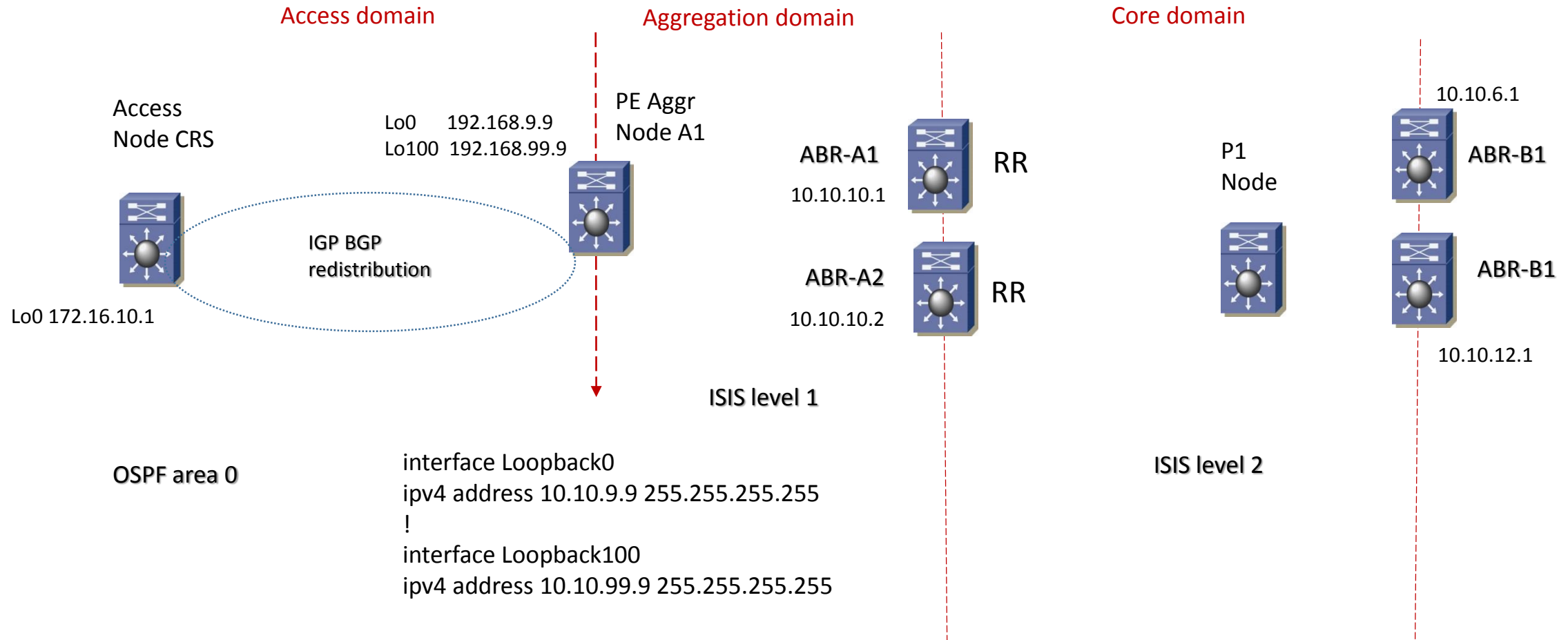
UNIFIED MPLS VPN-L3 (RFC 3107) Core and Aggregation with IGP redistribution on Access Configuration Example



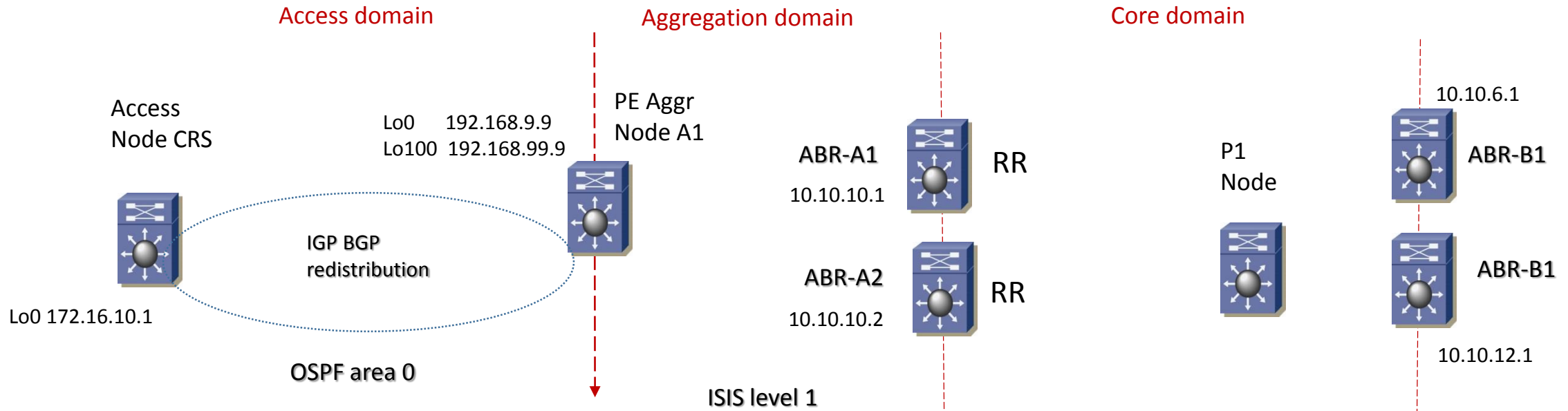
UNIFIED MPLS VPN-L3 (RFC 3107) Core and Aggregation with IGP redistribution on Access Configuration Example



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



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

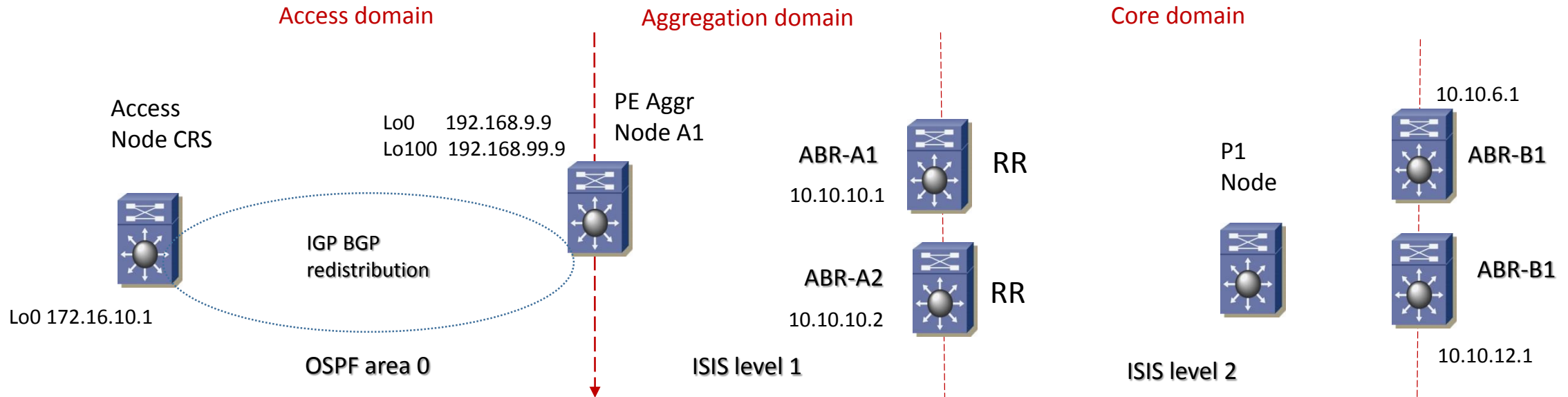


Aggregation IGP Configuration

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



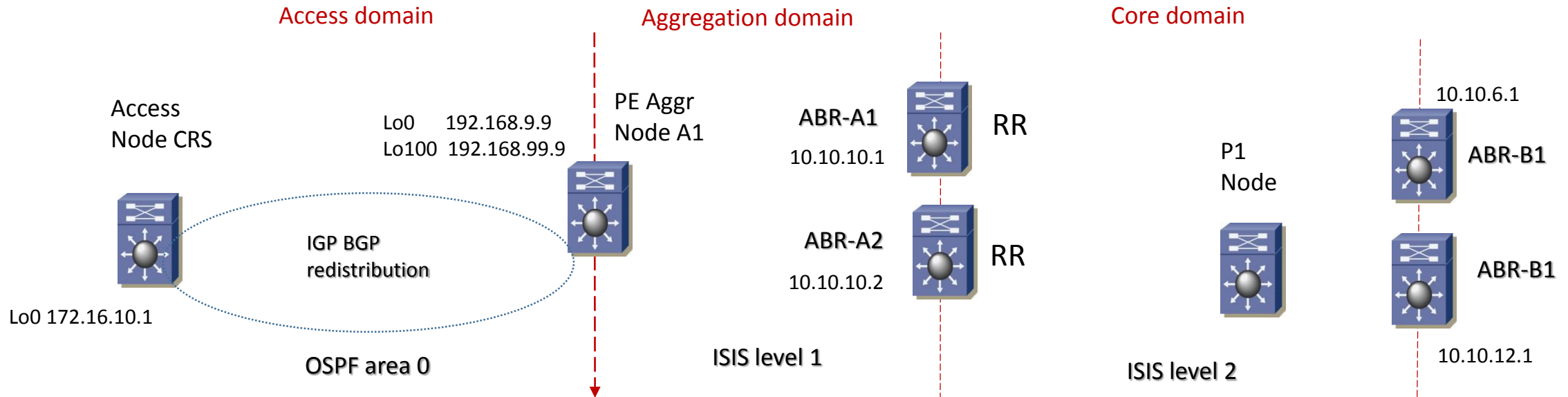
Access IGP Configuration

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
    
```

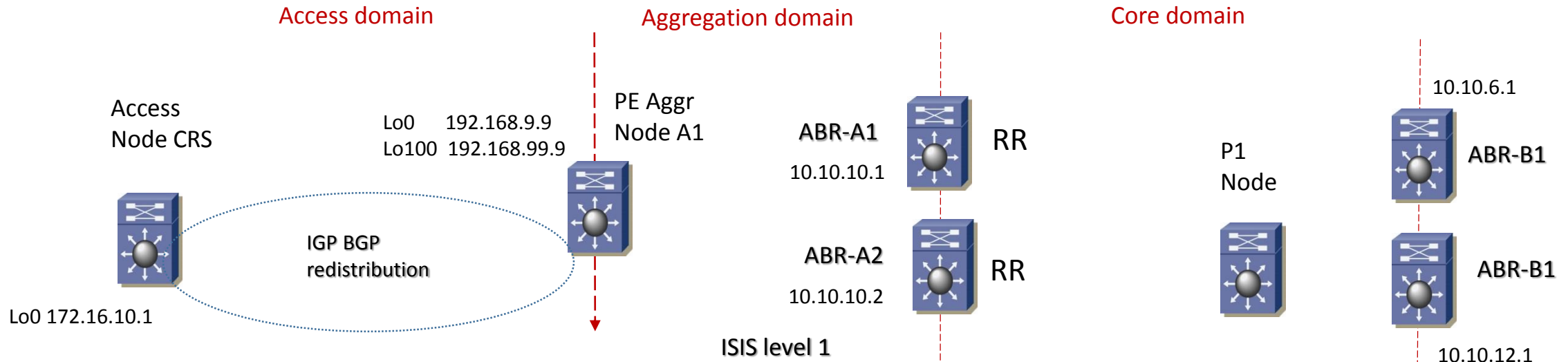

UNIFIED MPLS VPN-L3 (RFC 3107) Core and Aggregation with IGP redistribution on Access Configuration Example



```

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



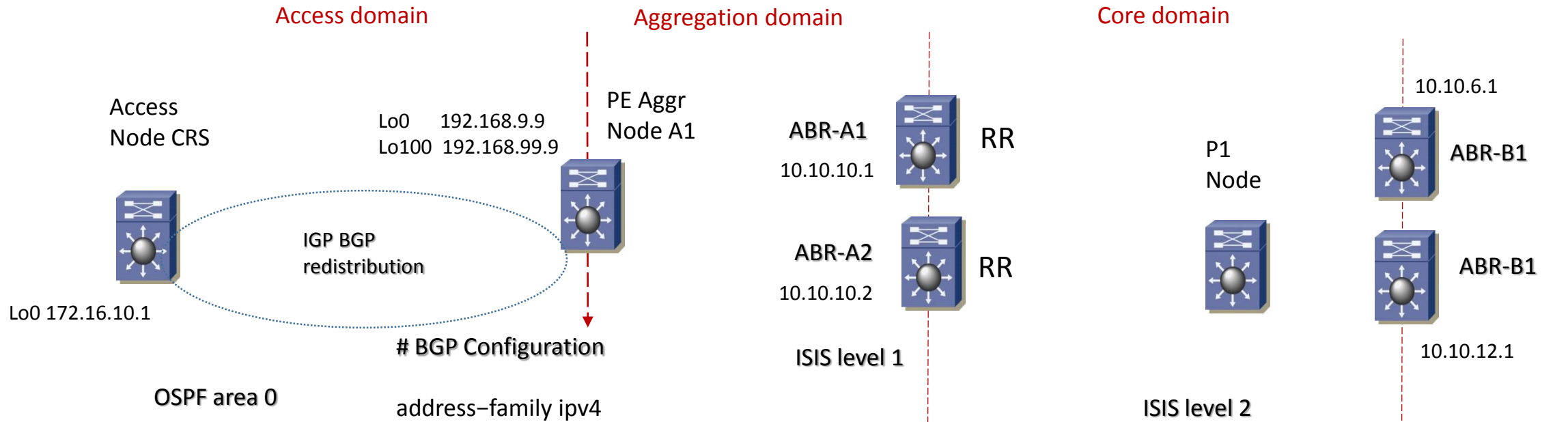
OSPF area 0

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



BGP Configuration

```

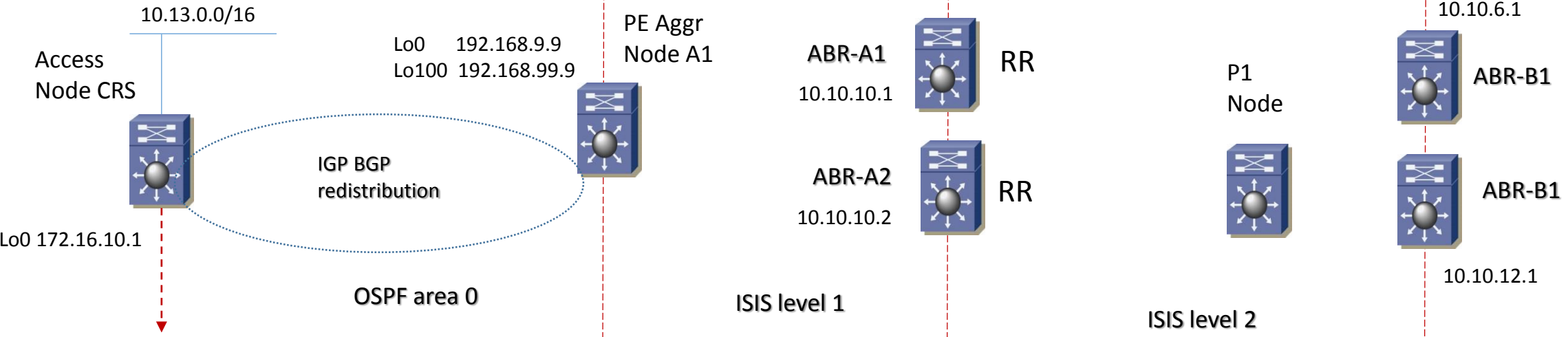
address-family ipv4
  bgp redistribute-internal
  network 192.168.9.9 mask 255.255.255.255 route-map AGG_Comm → advertise with Aggregation Community
  redistribute ospf 1 → redistribute Access IGP prefixes
  neighbor abr send-community
  neighbor abr next-hop-self
  neighbor abr send-label → send labels 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

Access domain

Aggregation domain

Core domain



```
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
```