EVPN 101

- Mufaddal

Disclaimer

- We will only cover the service provider configuration for EVPN and not cover the data-center based EVPN deployments.
- This session is based on my learnings from some soon to be deployment experience.
- This session only covers the configuration and output from IOS-XR based devices.
- Unfortunately no further LOTR references.

Important Message:

Missed the Cr anyone has a spare ticket t g to buy and probably dive

Session Agenda

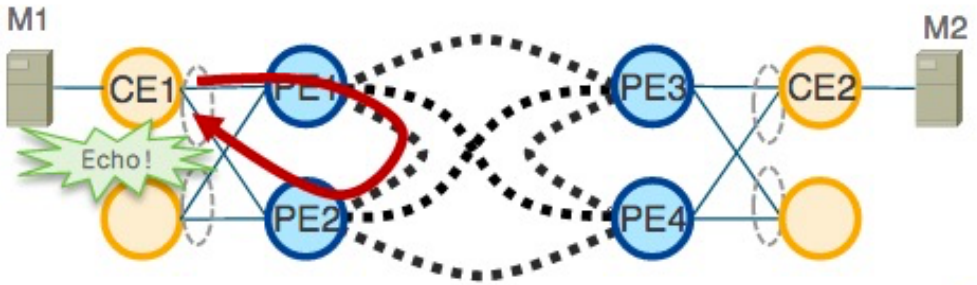
- Why EVPN ?
- EVPN Terminology
- Message Exchanges
- Useful links

Why EVPN ?

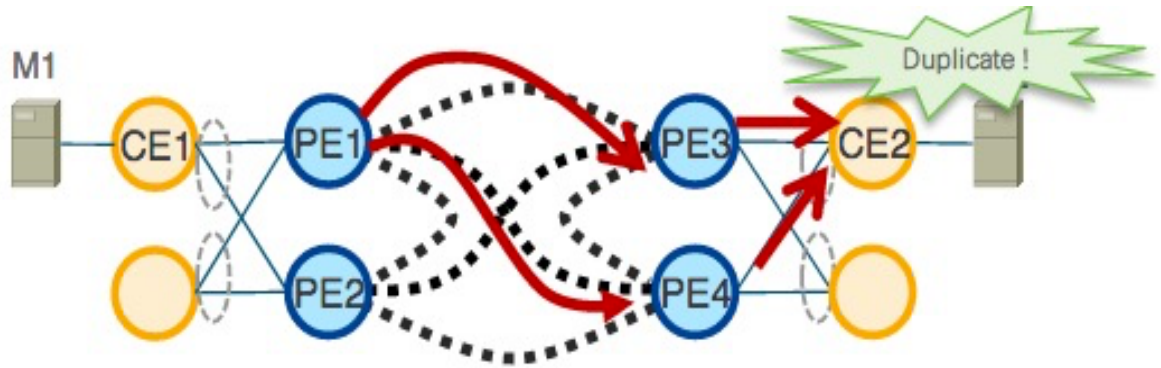
- Existing VPLS solutions do not offer an all-active per flow redundancy.
- Looping of traffic flooded from PE when a site is multihomed.
- Duplicate frames from floods from the core.
- MAC flip flopping over pseudowires.
- Separate services for layer 2 and layer 3.
- Scalability

Why EVPN ? (cont.)

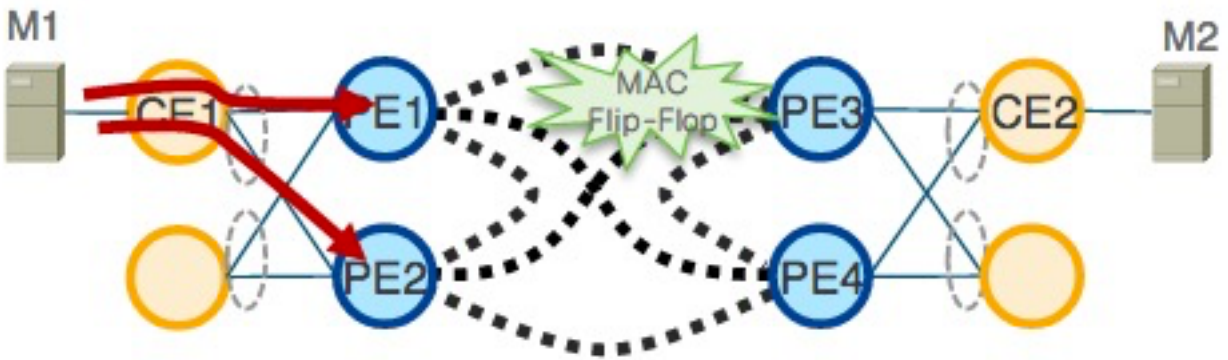
Traffic Looping



Duplicate frames



MAC Flip-Flopping

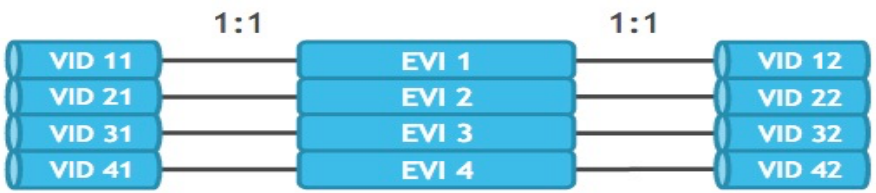


EVPN - RFC 7432

- All active forwarding for multihomed sites without the complexities of MC-LAG.
- Can be used to implement E-Line, E-LAN and even L3VPN.
- Optimized load-balancing between PEs.
- Optimized failure signaling during a network event reducing the amount of messages exchanged between the PEs.
- Layer3 VPN like operation for scalability and control.
- Reduction of control-plane protocols from LDP, BGP, MC-LAG, HSRP/VRRP & **STP** to just BGP MP.
- Straightforward configuration.
- Fits nicely with segment routing based network.

EVI

- EVPN can be broken down into different concepts for easier understanding.
- EVPN Instance (EVI)
  - Each EVPN instance is identified by a EVI. This is similar to pseudowire ID in a traditional L2VPN. EVIs can be of different types.
    - Port Based: There are no VLANs and the entire port is used. This is the simplest type.
    - VLAN Based: There is 1:1 mapping between the VLAN ID (bridge-domain in XR) and EVI.

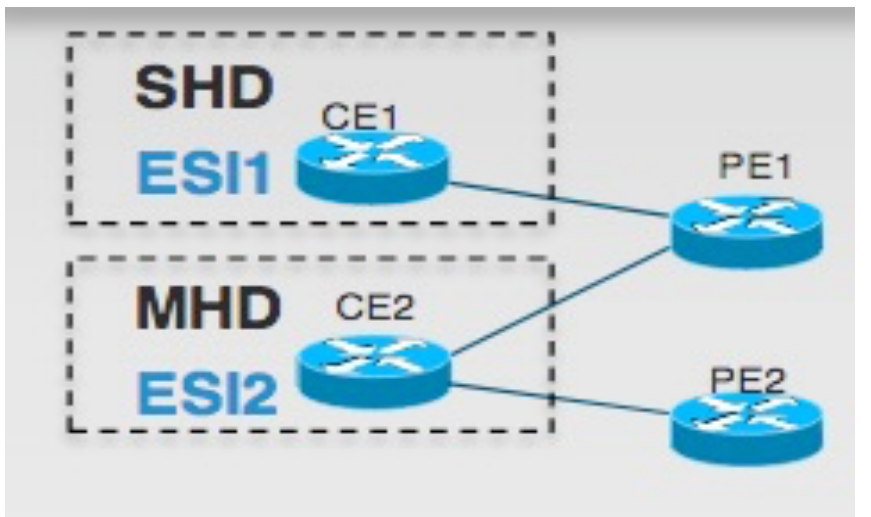


    - VLAN Bundling: There is N:1 mapping between the VLAN ID and EVI

Ethernet Segment

- Ethernet Segment
  - When a customer site is connected to one or more PEs via a set of Ethernet links, then this set of Ethernet links constitutes an "Ethernet segment"
  - For a multihomed site, each Ethernet segment (ES) is identified by a unique non-zero identifier called an Ethernet Segment Identifier (ESI).
  - An ESI is encoded as a 10-octet integer in the EVPN context. This value can be autogenerated or configured by the user.
  - XR only supports Type0 and Type1 out of the six types defined in the RFC.

ESI Type	Explanation	ESI Format
Type 0	Arbitrary ESI value based on configuration	1 octet ESI Type 0x00 9 octet ESI value
Type 1	Auto-generated ESI value based on LACP	1 octet ESI Type 0x01 6 octet CE LACP MAC address 2 octet CE LACP Port Key 1 octet value of 0x00



DF Election in a Ethernet-Segment

- Each multi-homed site, needs to elect a designated forwarder whose responsibility is to forward BUM for a particular VLAN segment.
- Each PE builds an ordered list of the IP addresses of all the PE nodes connected to the Ethernet segment (including itself), in increasing numeric value.
- Each PE gets an position based on the above ordered list.
- Assuming a redundancy group of n PE nodes, the PE with position **i** is the DF for an EVI when (EVI mod n) = **i**.

For example:

Let's say that PE1 and PE2 originator IP addresses are 1.1.1.1 and 2.2.2.2 respectively.

Ordered list on each PE will be:

0: PE1 with 1.1.1.1

1: PE2 with 2.2.2.2

Modulus Operation: (EVI ID % 2) as we have two PEs in this ESI

PE1 becomes DF for EVI ID 300 and PE2 becomes DF for EVI ID 301.

Route Distinguisher

- A unique RD must be set for each EVI instance on the PE.
- It is recommended to use the Type1 RD format which is 4 byte: 2 byte value.
- The 4-byte corresponds to the loopback address of the PE while the 2 byte value is the EVI ID in case of IOS-XR devices.

```
RP/0/RP0/CPU0:lab-1#show evpn evi vpn-id 1003 detail | i RD
Tue Mar  5 11:25:06.456 UTC
  RD Config: none
  RD Auto  : (auto) xxx.xxx.xxx.xxx:1003
RP/0/RP0/CPU0:lab-1#show running-config interface lo0
Tue Mar  5 11:25:12.374 UTC
interface Loopback0
 description Transport/Management Loopback Interface
 ipv4 address xxx.xxx.xxx.xxx 255.255.255.255
 ipv6 address 2a02:xxx:fc00::b/128
!
```

Route Target

- The EVPN route MAY carry one or more Route Target (RT) attributes.
- RTs may be configured (as in L3VPNs) or may be derived automatically.
- The RTs are auto-derived as follows:
  - The Global Administrator field of the RT MUST be set to the Autonomous System (AS) number with which the PE is associated.
  - The 12-bit VLAN ID MUST be encoded in the lowest 12 bits of the Local Administrator field, with the remaining bits set to zero. Cisco XR uses the EVI ID instead of the VLAN ID.

```
RP/0/RP0/CPU0:lab-1#show evpn evi vpn-id 1003 detail | i RT
Tue Mar  5 11:34:41.460 UTC
  RT Auto  : 210278:1003
!
RP/0/RP0/CPU0:lab-1#show bgp | i AS
Tue Mar  5 11:35:00.949 UTC
BGP router identifier xxx.xxx.xxx.xxx, local AS number 210278
```

MP-BGP Routes

- BGP Routes (AF: 25, SAFI: 70)
  - This is used to build the control-plane in EVPN

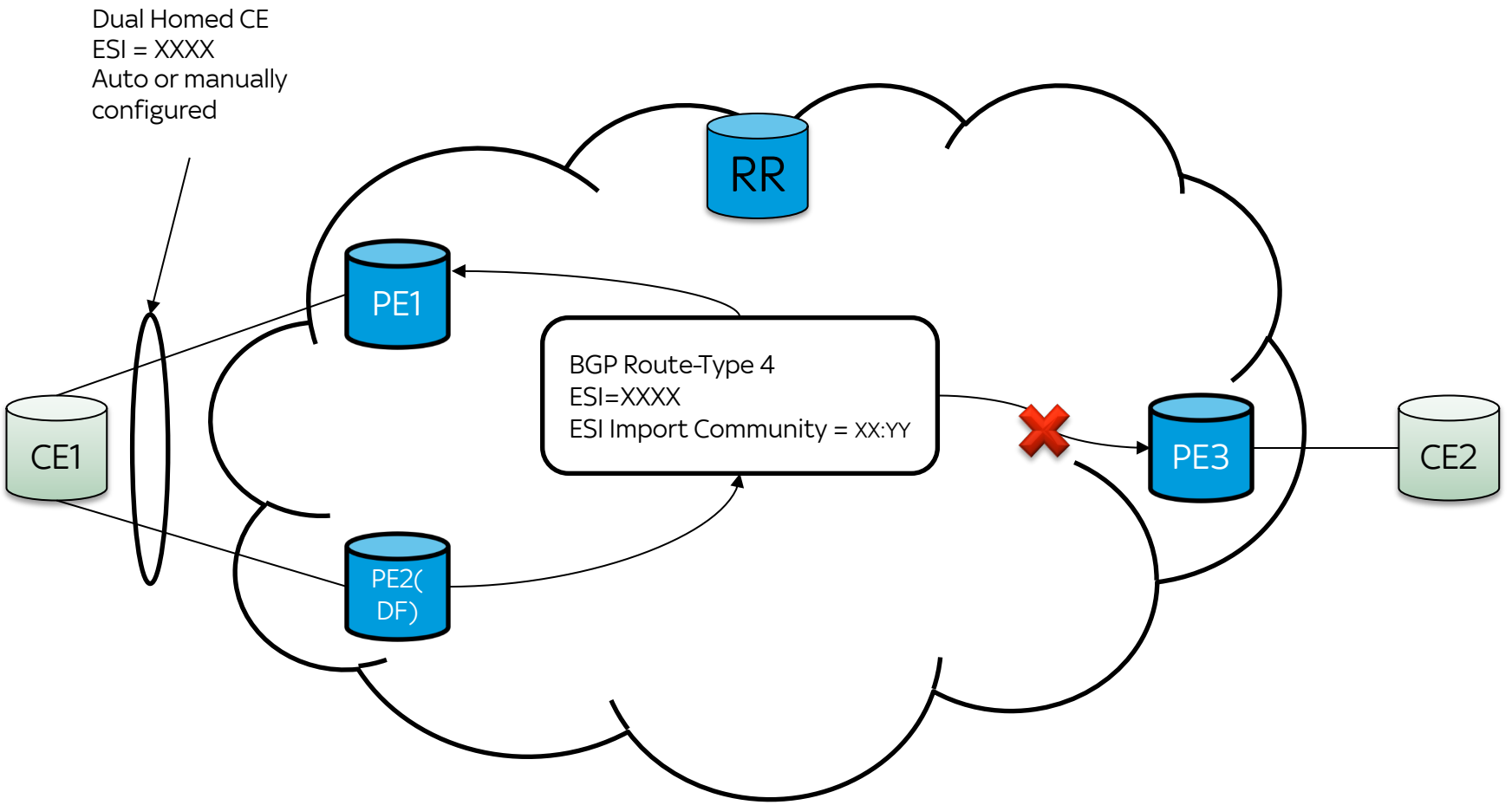
Route Type	Name	Usage
1	Ethernet Auto-Discovery (AD) Route	Type 1 routes are used for achieving split horizon, fast convergence and aliasing.
2	MAC/IP Advertisement Route	Advertise MAC, address reachability, advertise IP/MAC binding
3	Inclusive Multicast Ethernet Tag Route	This route establishes the connection for BUM traffic from a source PE to a remote PE. This route is advertised on per VLAN and per ESI basis
4	Ethernet Segment Route	This route is used for auto-discovery of dual homed sites and for DF election.
5	IP Prefix Advertisement Route	This route is used to advertise the IP prefixes and the L3 gateway address.

MP-BGP Route Attributes

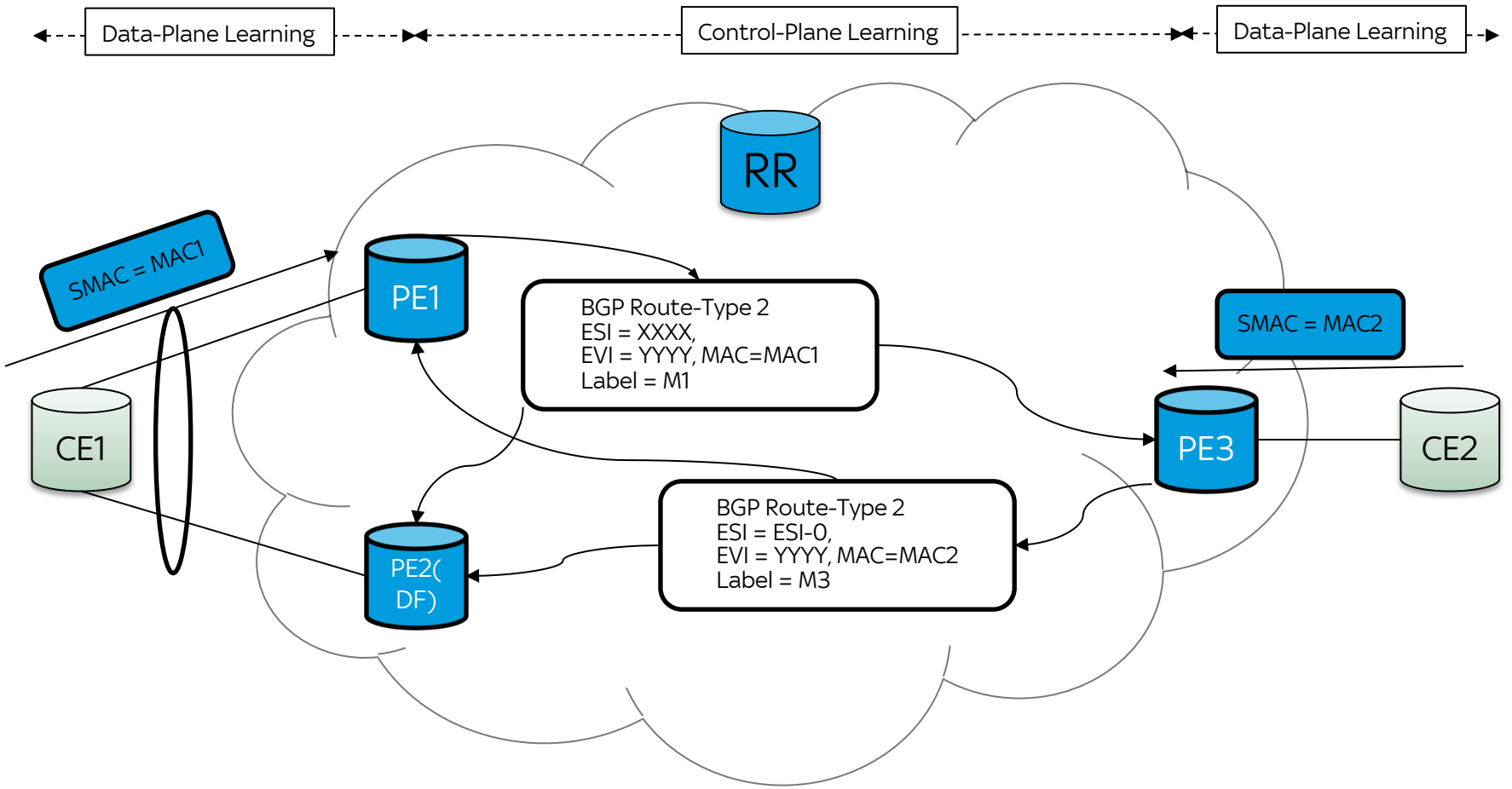
- BGP Route Attributes
  - New BGP attributes to expand information carried in the BGP routes

Extended Community	Used with which EVPN route(s)	Usage
ESI Label	RT1	This community encodes the split-horizon label used by the PEs in the same multi-homed sites for filtering.
ES-Import RT	RT4	It enables all the PEs connected to the same multihomed site to import the Ethernet Segment routes. The value is derived automatically for the ESI Types 1, by encoding the high-order 6-octet portion of the 9-octet ESI Value. It also indicates whether the site is single-active or active-active (flag)
Mac Mobility	RT2	This community contains a 32-bit sequence number used to ensure the PEs retain the correct MAC/IP route when multiple updates occur for the same MAC address.
Default Gateway	RT2	Each PE that acts as a default gateway for a given EVPN instance MAY advertise in the EVPN control plane its default gateway MAC address using RT2, and each such PE indicates that such a route is associated with the default gateway by carrying the default gateway extended community

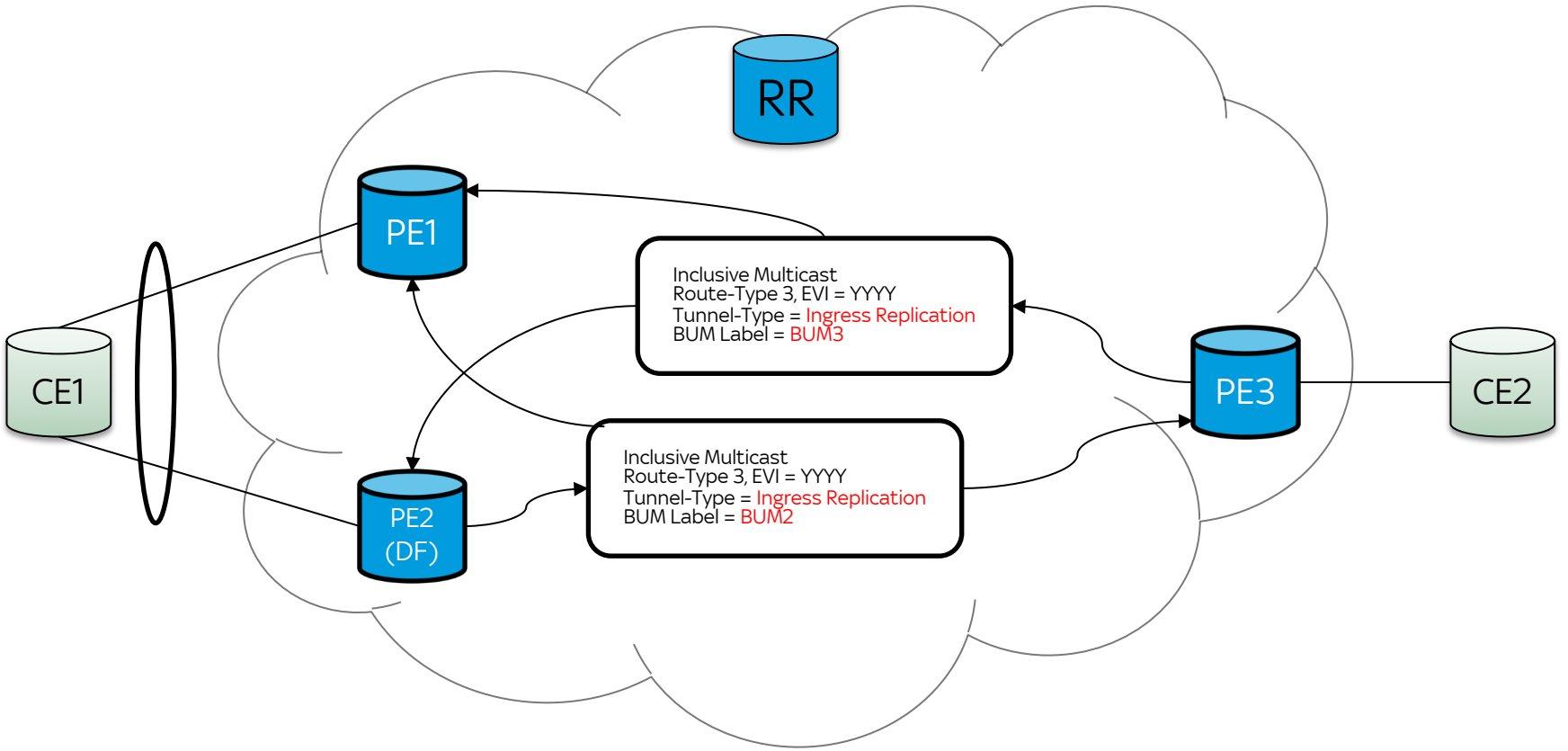
Discovering Ethernet Segment for Multi-homed Sites (RT4)



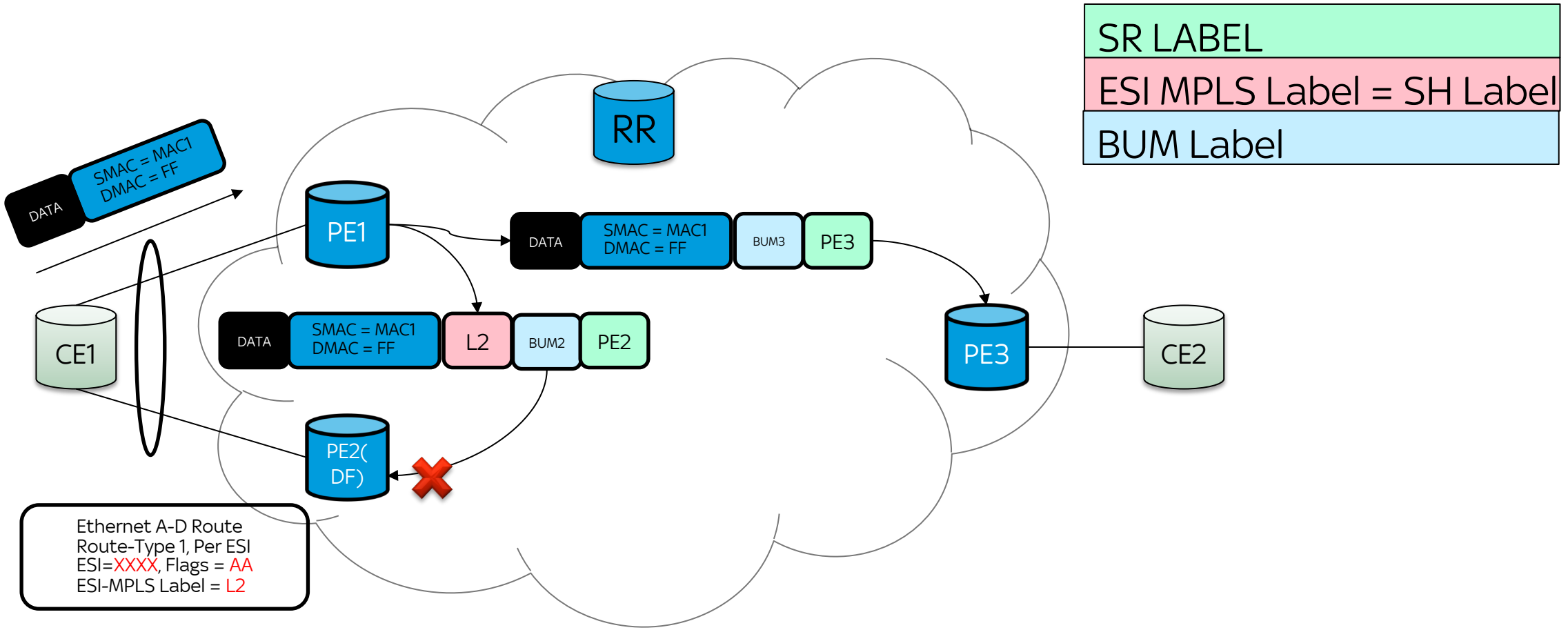
Advertising MAC address (RT2)



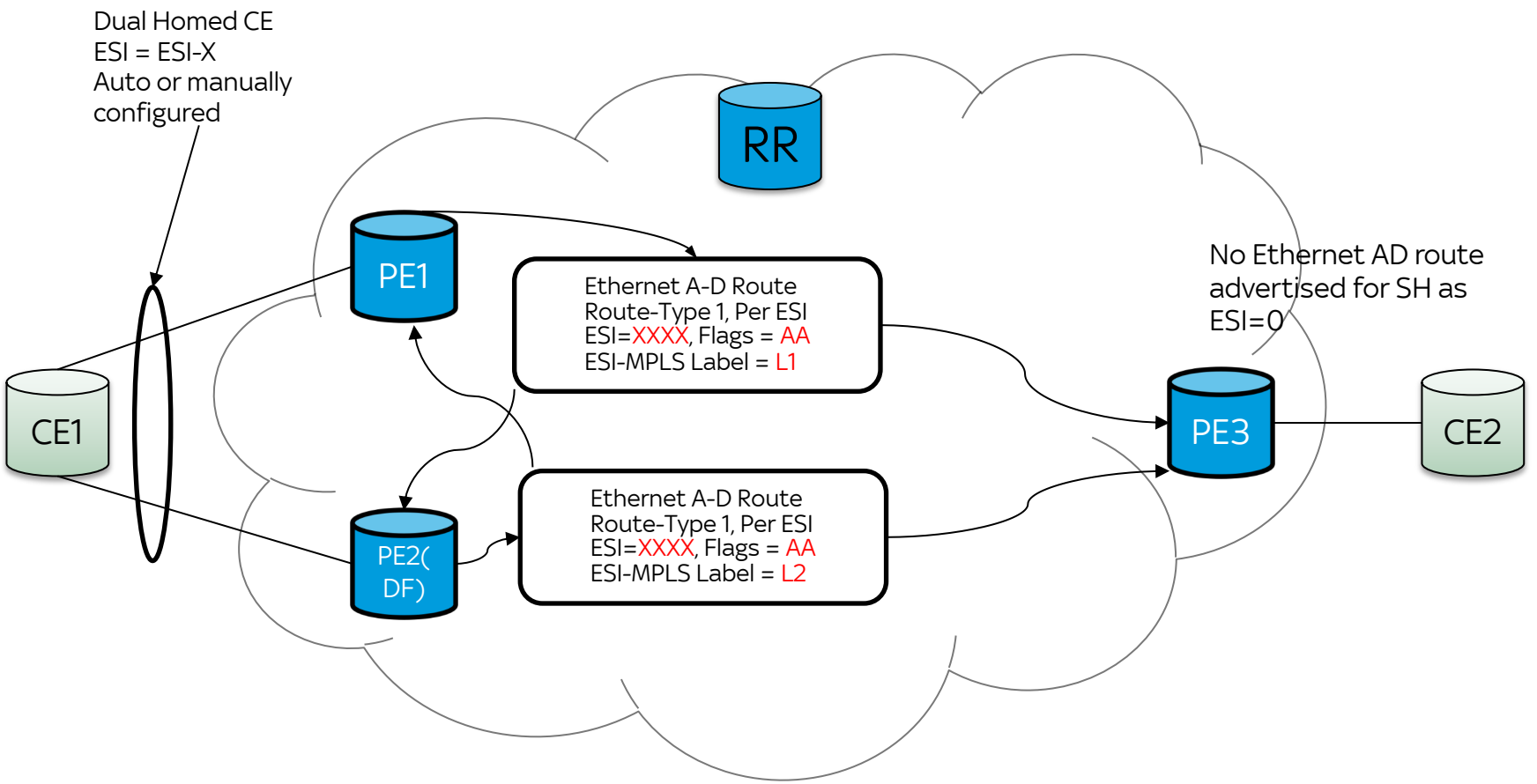
Building control-plane for BUM traffic (RT3)



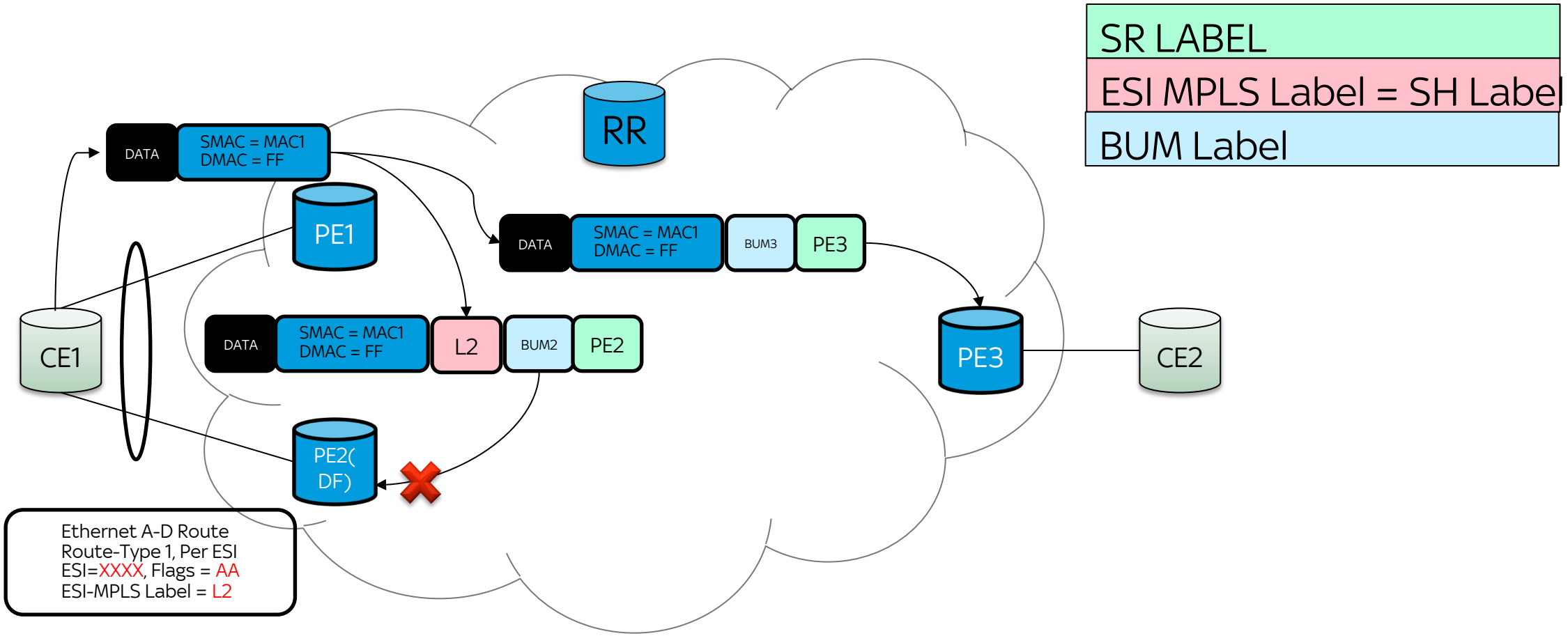
Forwarding BUM traffic



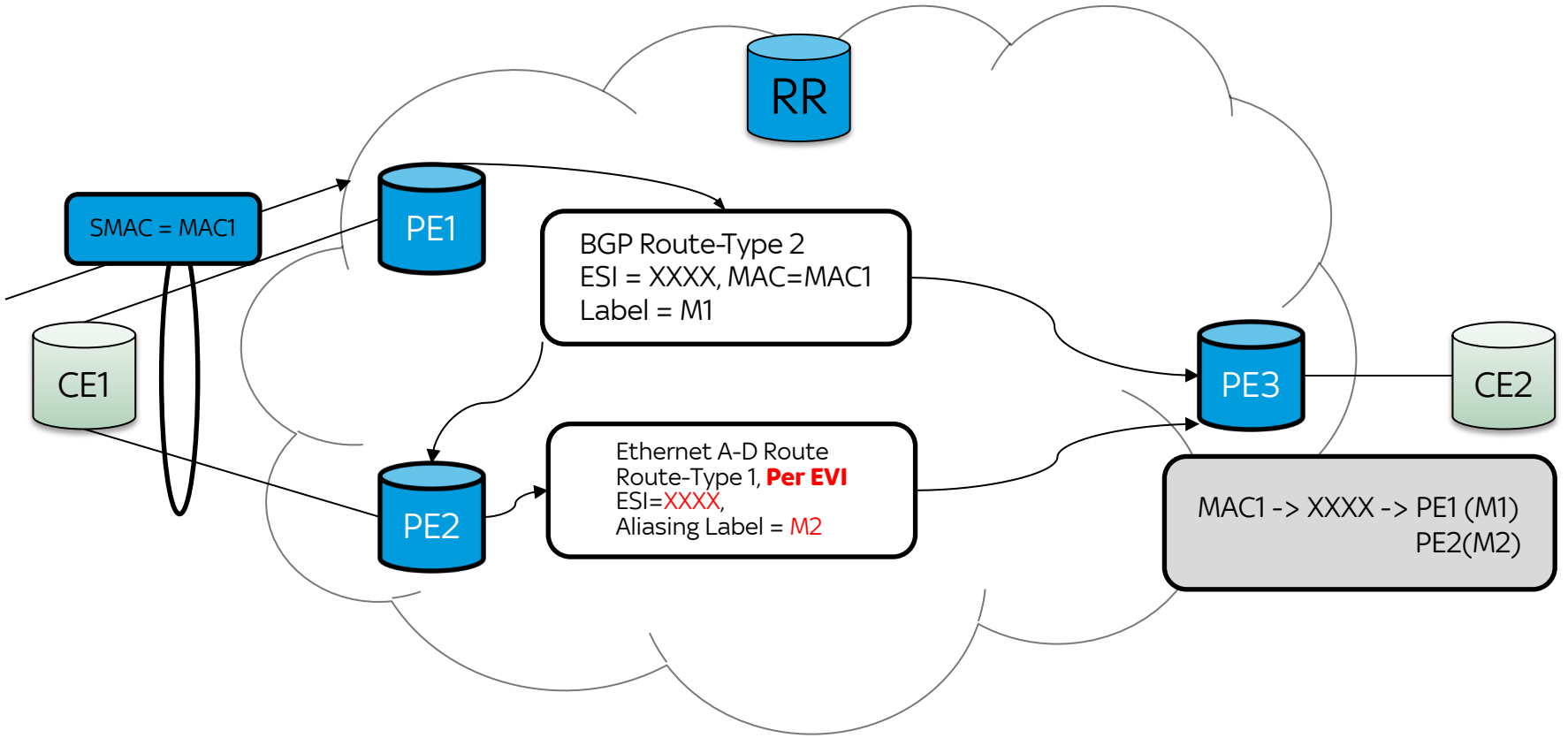
Achieving Split Horizon in Control-Plane (RT1)



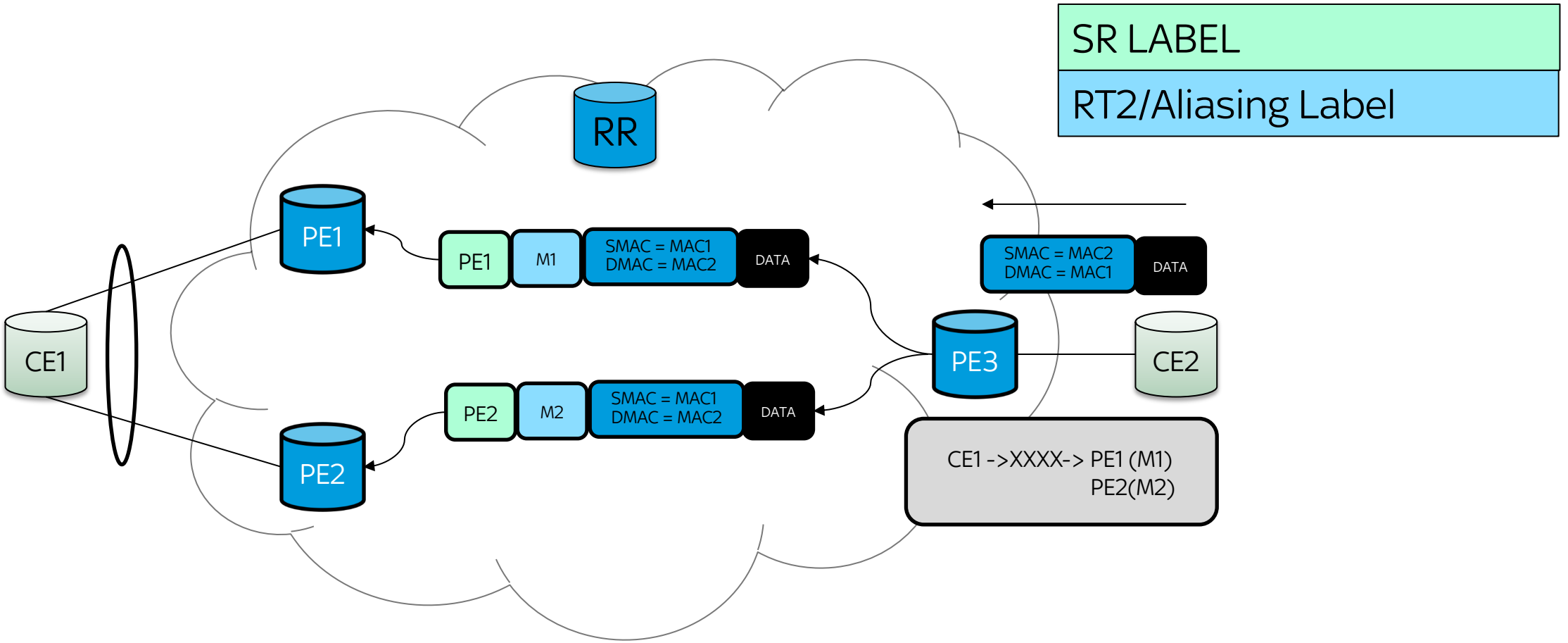
Split Horizon in Data Plane



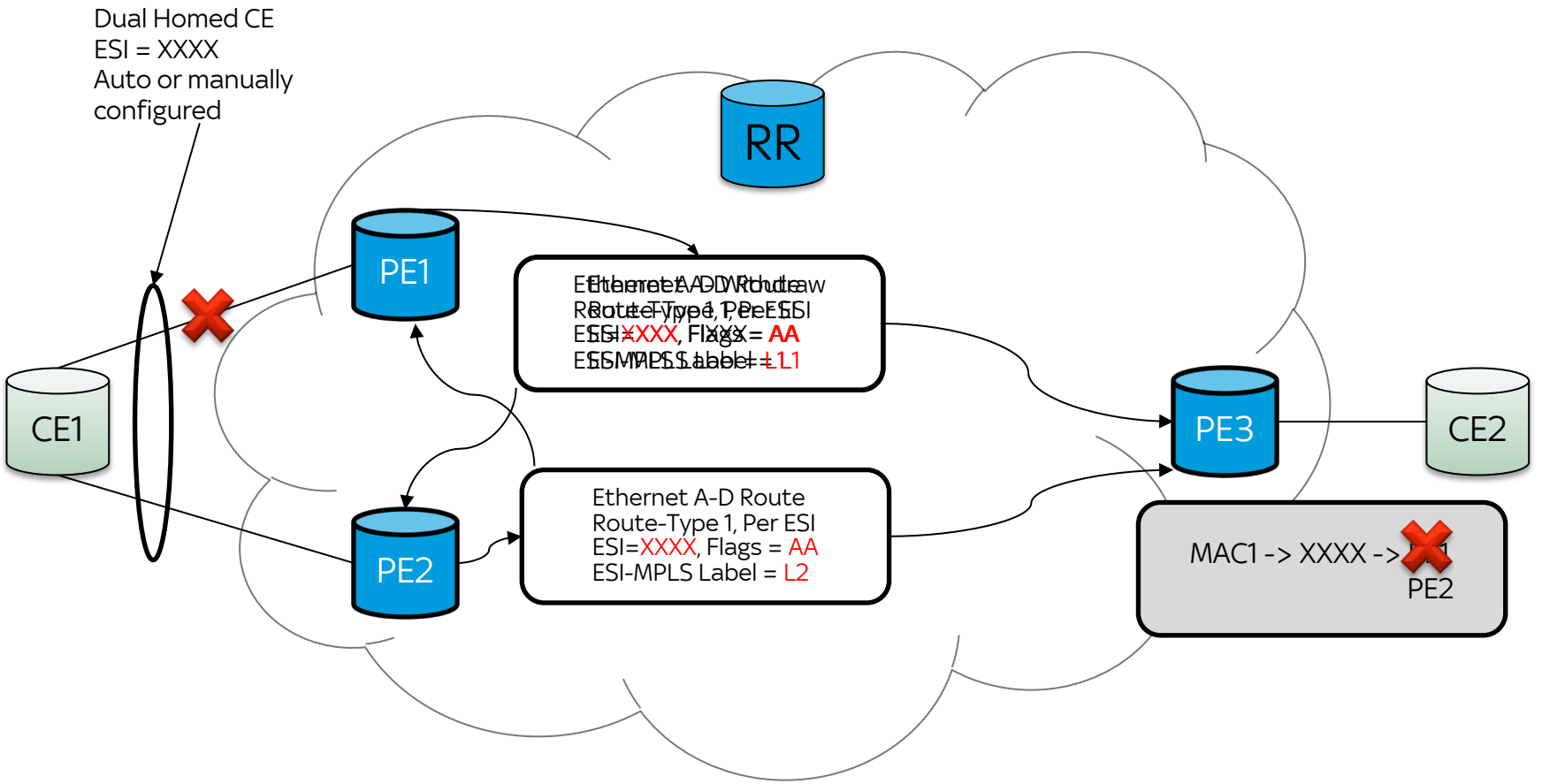
Achieving load-balancing via aliasing (RT1)



Load-balancing in data-plane



Achieving faster Convergence



Useful links

- https://tools.ietf.org/html/rfc7432
- https://clnv.s3.amazonaws.com/2019/eur/pdf/BRKSPG-2322.pdf
- https://clnv.s3.amazonaws.com/2019/eur/pdf/LTRSPG-2968.pdf
- https://clnv.s3.amazonaws.com/2019/eur/pdf/BRKSPG-3965.pdf
- https://conference.apnic.net/data/37/2014-02-24-apricot-evpn-presentation_1393283550.pdf
- https://e-vpn.io/tech/tech2/tech2.html
- http://karneliuk.com/tag/evpn/
- http://www.bgphelp.com/tag/evpn/

Questions/Beers ?