

Interdomain VPLS and deployment experiences



Red IRIS



17th TF-NGN meeting – April 14th, 2005 (Zürich)

Laura Serrano <laura.serrano@rediris.es>

Introduction

Concepts

Intradomain environment

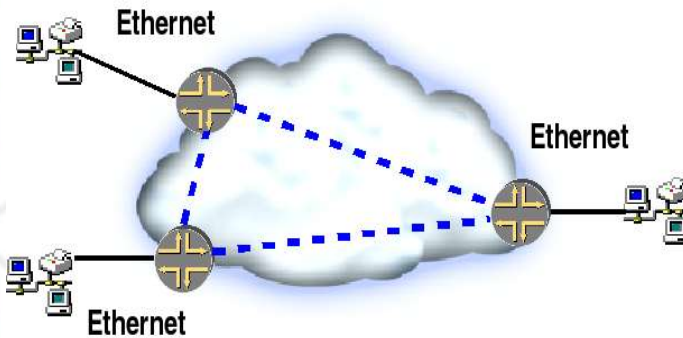
Interdomain environment

Scalability

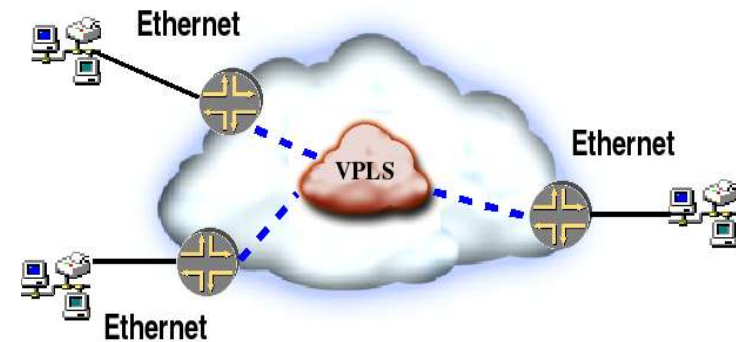
Summary and troubleshooting

What is VPLS?

- MPLS (*Multiprotocol Label Switching*) solutions:
 - Layer 3 VPNs -> RFC 2545-bis
 - Layer 2 VPNs ->



point to point service (Martini/kompella)



multipoint service (VPLS)

VPLS is a way of providing Ethernet/MPLS virtual private networks between sites located in different geographical points

Why VPLS?

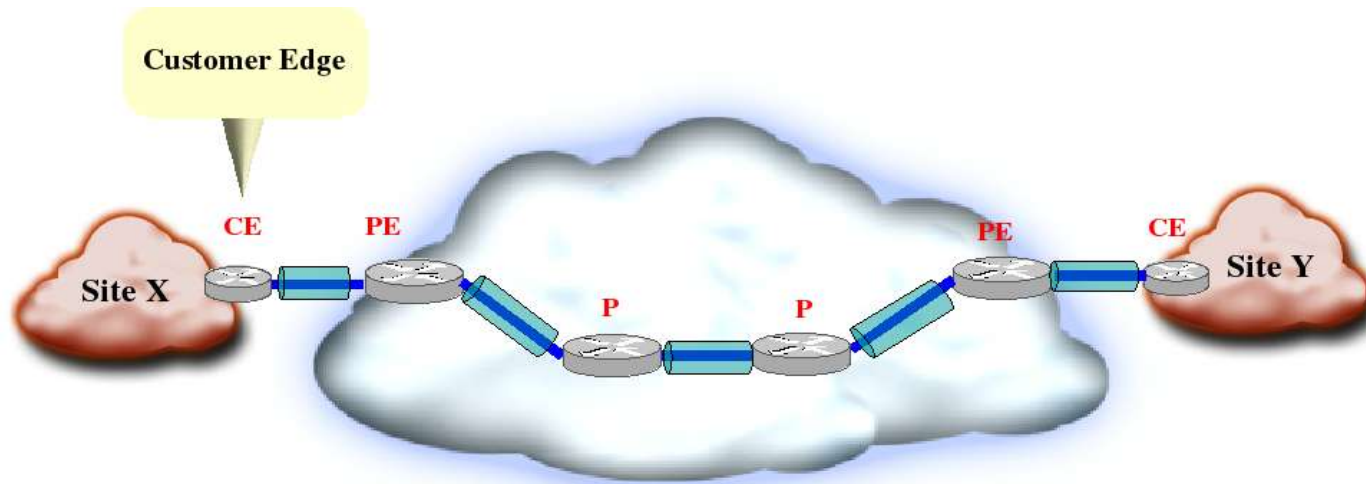
- Useful for advanced applications and distributed computing over L3 networks even in different domains
 - Grids and clustering
 - Corporate Networks
 - Advanced client-server applications
 - Remote data storage
 - Remote data calculation
 - Layer 2 tools: bootp? , license servers?
 - New technologies networks
 - ...

red.es

How does VPLS work?

- The provider network works as a single LAN
 - *Provider Edge* device learns MAC addresses from *Customer*
 - New forwarding table -> *new-vpn.vpls*
 - MAC addresses are “dynamically” mapped and sent to the other *provider edge* devices
- Standards for VPLS
 - Draft-ietf-l2vpn-vpls-bgp-X.txt
 - Uses MP-BGP for signaling -> SUPPORTED BY JUNIPER
 - Draft-ietf-ppvvpn-vpls-ldp-X.txt
 - Used LDP for signaling -> SUPPORTED BY CISCO
- Why MP-BGP signaling?
 - Autodiscovery & scalability properties (RR, confederations, ...)
 - **BWHJB :-)**

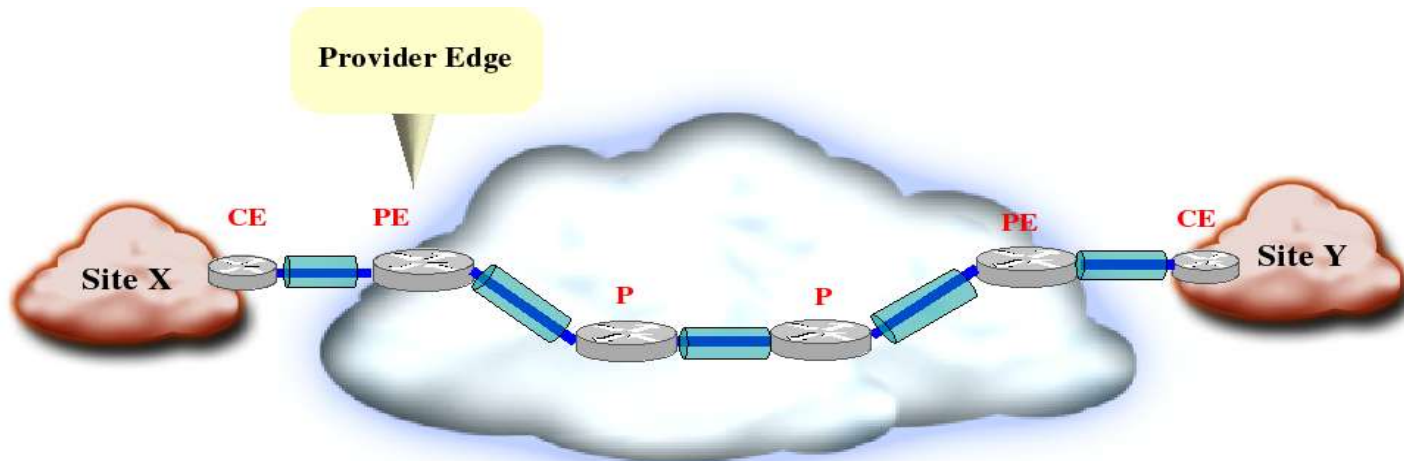
Devices involved in VPLS (I)



- Customer edge devices (CE)
 - Router or switch connected to provider network
 - No particular requirement

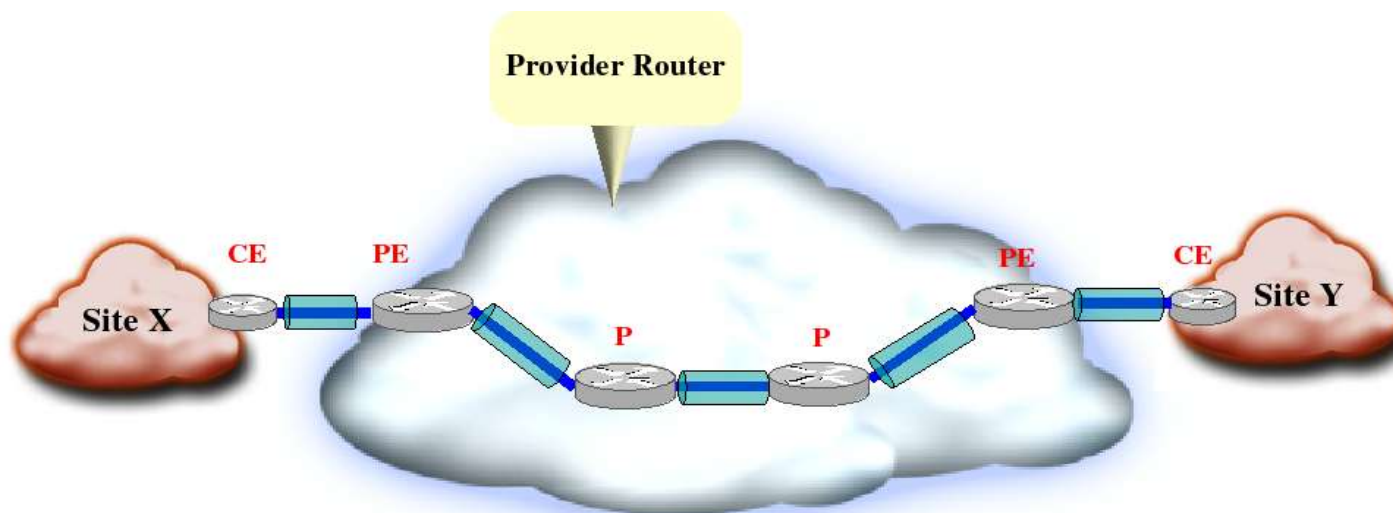
Note: By default, both ends of the VPN must use the same Layer 2 technology (ethernet or ethernet with vlan tagging)

Devices involved in VPLS (II)



- Provider Edge Routers
 - Maintain VPN information
 - Responsible of MAC address learning process
 - Exchange information with the other PEs (MP-BGP)
 - Require MPLS LSPs to forward VPN traffic to other PEs

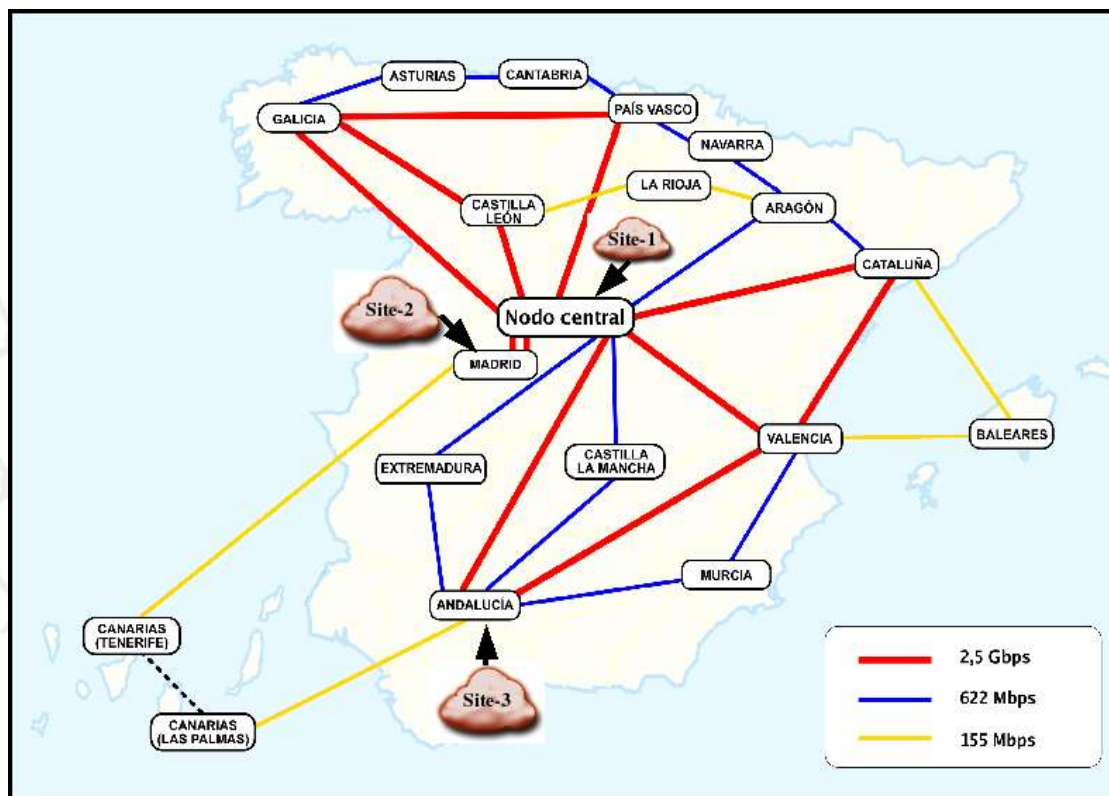
Devices involved in VPLS (III)



- Provider Routers
 - Forward traffic transparently over LSPs
 - No particular requirement

What do we want to do?

- Three customers connected through one VPLS
 - Two of them in Madrid and the last one in Andalucía

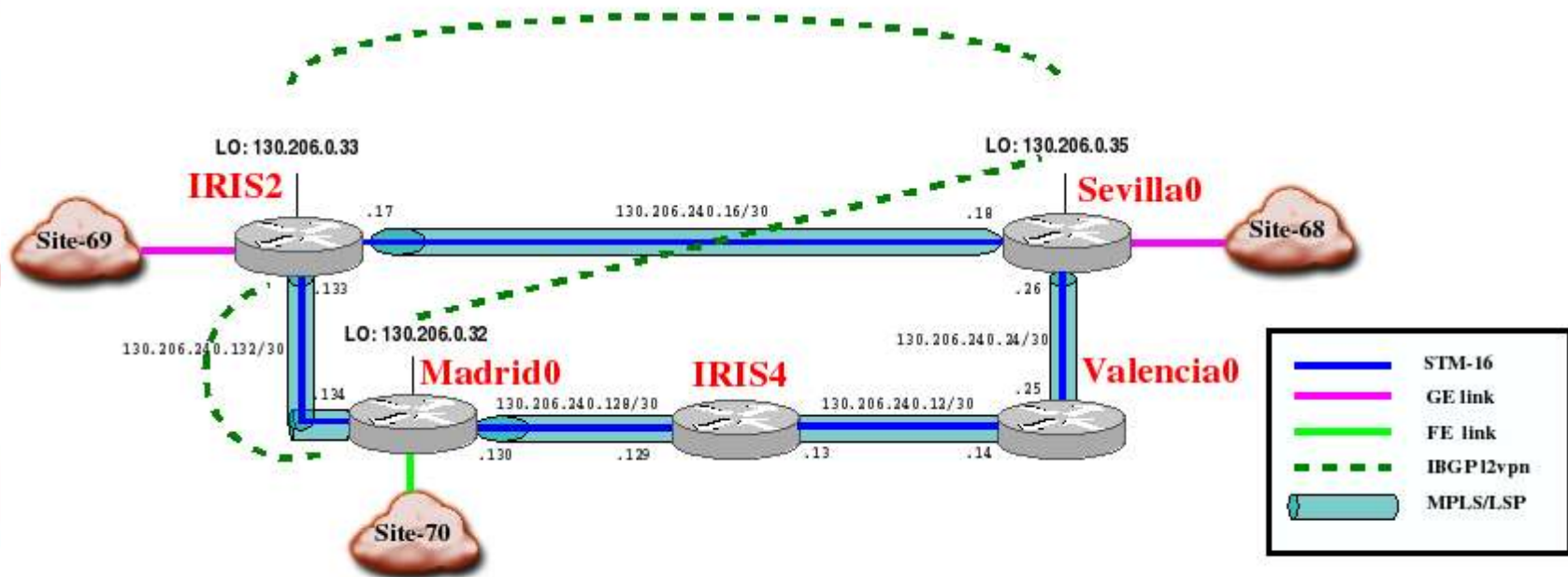


Note: The VPLS connection has been configured over the RedIRIS, full-production network

Configuration steps

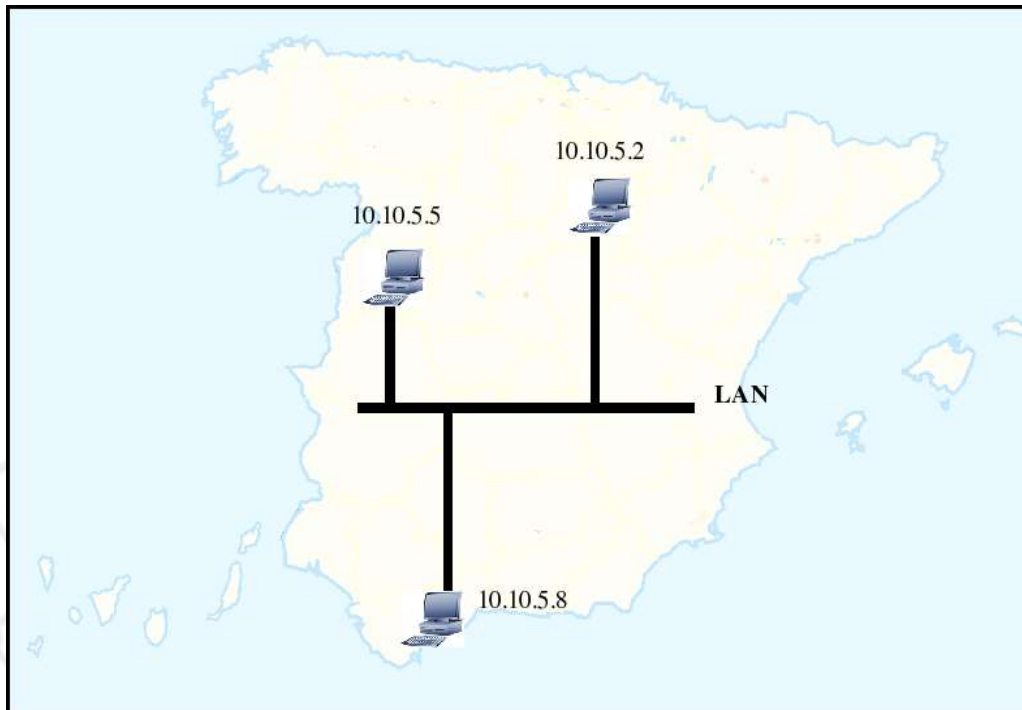
- MP-IBGP between PE routers (family I2vpn)
- MPLS/LSPs between PE routers (RSVP or LDP)
- One instance for each site (VPLS support)

IBGP vpls



Note: The PE routers must have a Tunnel Services Physical Interface Card or similar

What do we have?



```
laura@10.10.5.5:~# ping 10.10.5.2
```

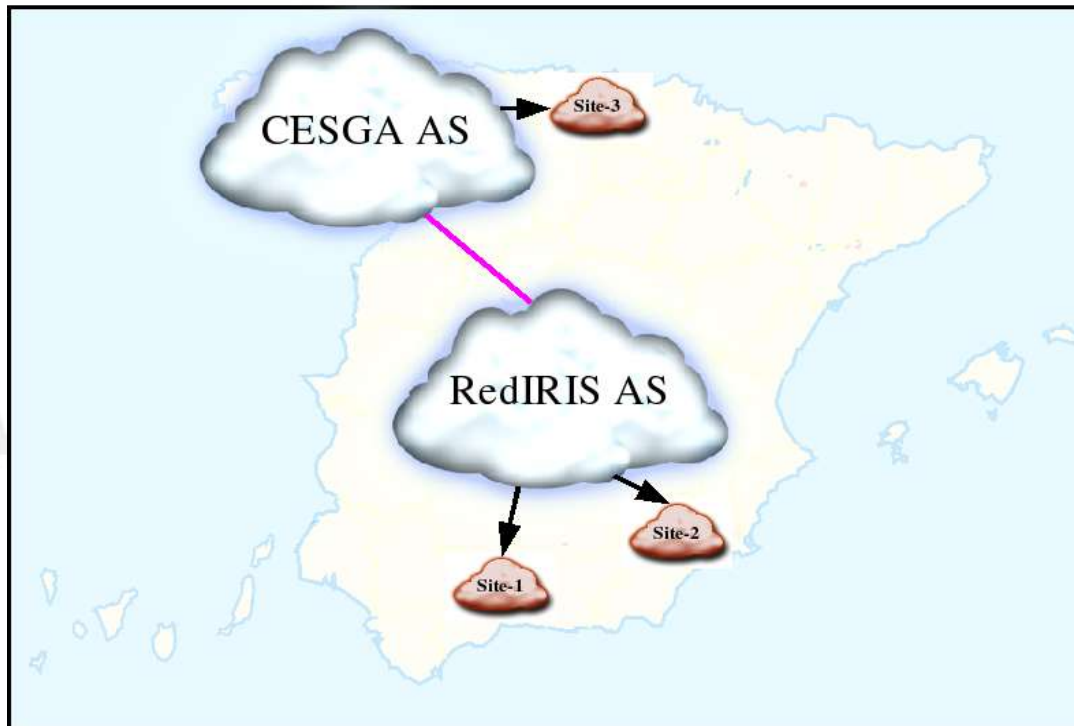
```
PING 10.10.5.2 (10.10.5.2) 56(84) bytes of data.  
64 bytes from 10.10.5.2: icmp_seq=0 ttl=64 time=0.095 ms  
64 bytes from 10.10.5.2: icmp_seq=1 ttl=64 time=0.037 ms
```

```
laura@10.10.5.5:~# traceroute 10.10.5.2
```

```
traceroute to 10.10.5.2 (10.10.5.2), 30 hops max, 38 byte packets  
1 10.10.5.2 (10.10.5.2) 0.160 ms 0.102 ms 0.068 ms
```

What do we want to do?

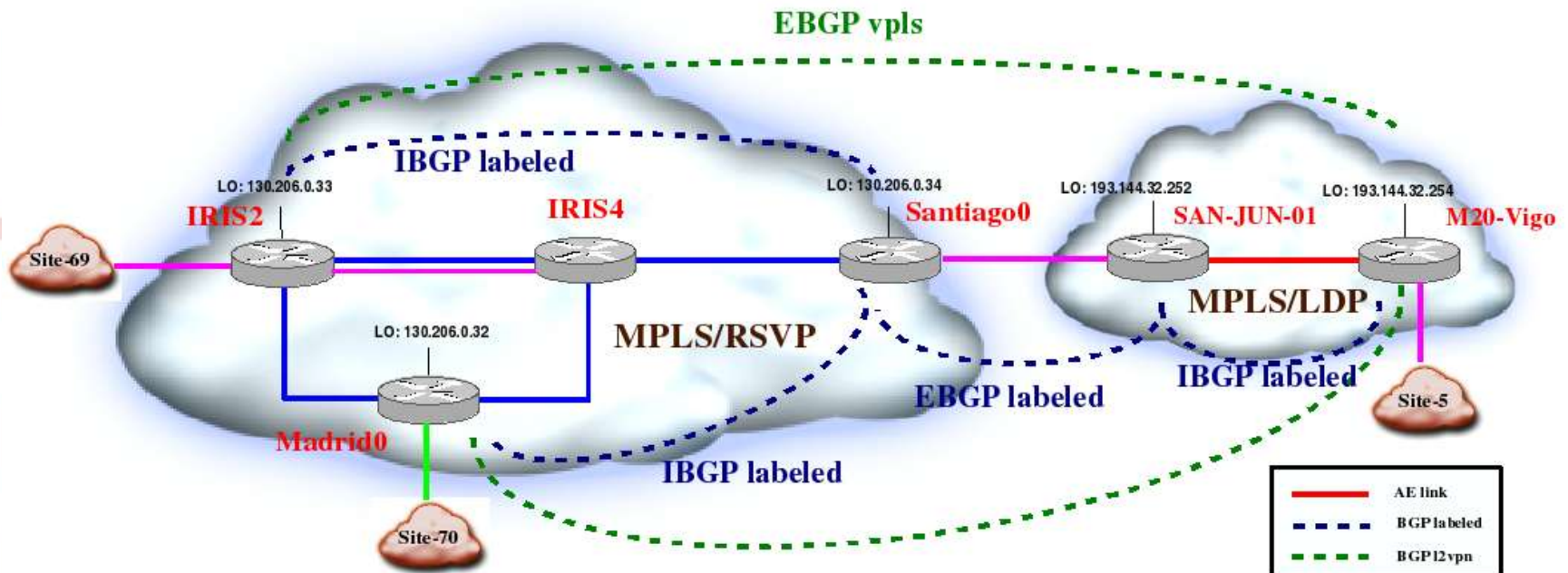
- Three customers connected through one VPLS
 - Two of them in RedIRIS domain and the last one in GESGA domain



Note: We are speaking about interprovider VPLS configuration

Configuration steps

- MP-BGP between PE routers (VPLS BGP)
- Extend the MPLS/LSPs between PE routers
 - Exchange labeled route -> BGP (labeled-unicast)



PE configuration review

```

laura@IRIS2_router# show
[edit protocols bgp]
group iBGP-VPLS-Test {
  type internal;
  local-address 130.206.0.33;
  family l2vpn {
    unicast;
  }
  neighbor 130.206.0.32;
}
group EBGP-VPLS-Test {
  type external;
  multihop {
    ttl 4;
    no-nexthop-change;
  }
  local-address 130.206.0.33;
  family l2vpn {
    unicast;
  }
  peer-as 64800;
  neighbor 193.144.32.251;
}
group iBGP-Labeled-VPLS {
  type internal;
  local-address 130.206.0.33;
  family inet {
    labeled-unicast {
      resolve-vpn;
    }
  }
  neighbor 130.206.0.34;
}

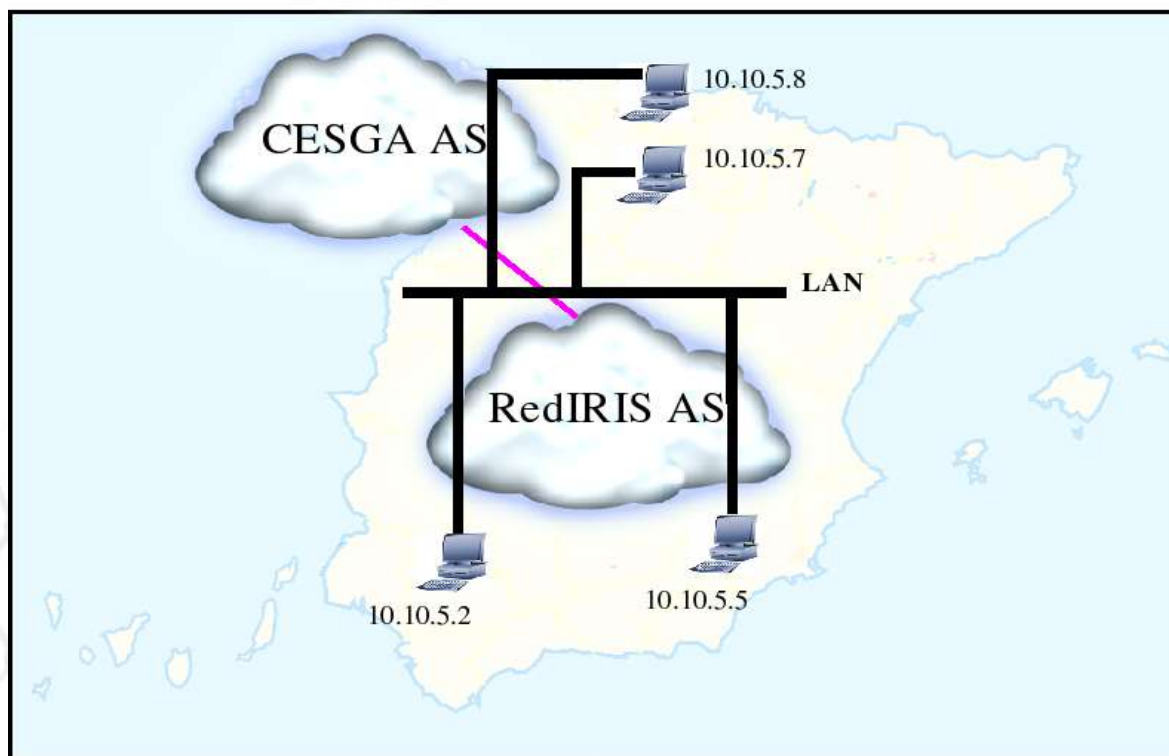
```

```

laura@IRIS2_router# show
[edit protocols mpls]
explicit-null;
label-switched-path LSP-IRIS2-Madrid {
  from 130.206.0.33;
  to 130.206.0.32;
  no-cspf;
}
label-switched-path LSP-IRIS2-IRIS4-Santiago0
{
  from 130.206.0.33;
  to 130.206.0.34;
  no-cspf;
}
interface all;
[edit routing-instances]
VPLS-RedIRIS-CESGA {
  instance-type vpls;
  interface ge-3/0/1.669;
  route-distinguisher 130.206.0.33:101;
  vrf-target target:100:2;
  protocols {
    vpls {
      site IRIS2 {
        site-identifier 69;
      }
    }
  }
}

```

What do we have? (I)

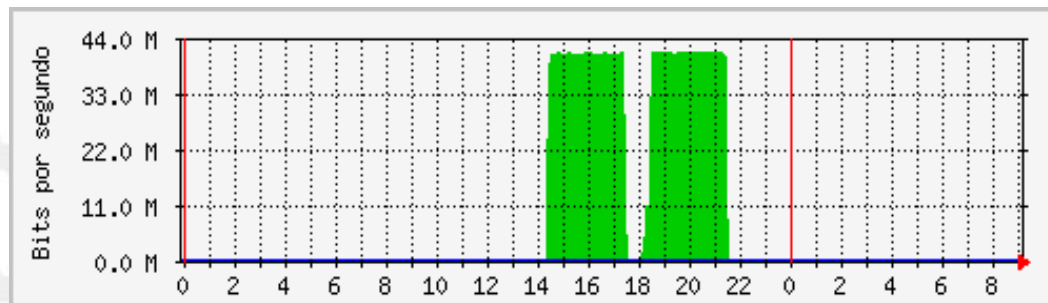


Customers located in different ASs as if they were in the same LAN

- Provider network has IPv4/IPv6/IP multicast and VPLS traffic with the same infrastructure (routers and links)

What do we have? (II)

- More security and flexibility
 - We get a trusted networks: only trusted hosts are included in the VPLS connection
 - Customer manages its network without provider control
- Hosts with full connectivity between them



```
laura@10.10.5.5:~# iperf -c 10.10.5.8 -u -b 40M -p 5002 -w 256k -l 1300 -i 10 -t 10800
[ 3] Server Report:
[ 3] 0.0-10793.3 sec 50.3 GBytes 40.0 Mbits/sec 0.003 ms 21991/41538463 (0.053%)
[ 3] Sent 41538463 datagrams
```

```
laura@10.10.5.5:~# traceroute 10.10.5.8
traceroute to 10.10.5.8 (10.10.5.8), 30 hops max, 38 byte packets
 1 10.10.5.8 (10.10.5.8) 11.883 ms 11.915 ms 11.766 ms
```

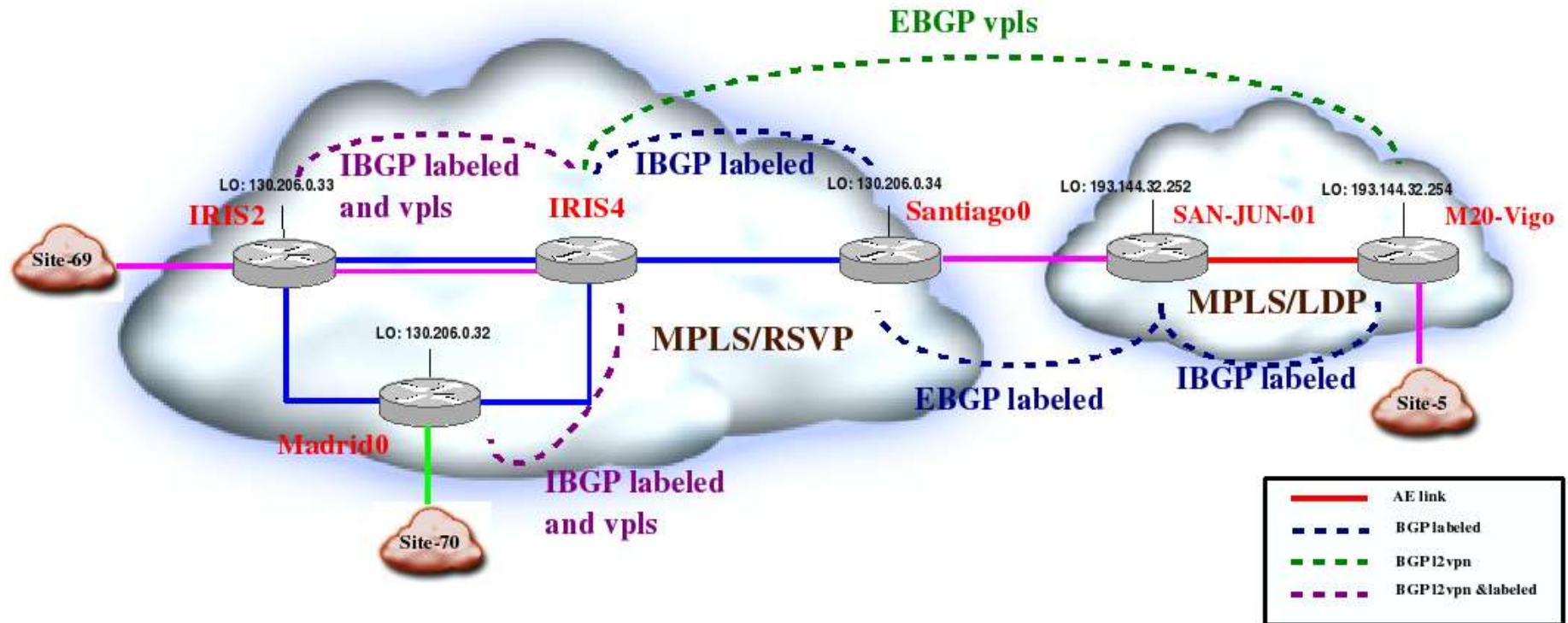
What happen if we want to add more extra sites or services like
l3vpn or point to point l2vpn?

- Benefit of using BGP as signalling protocol
 - Scalability -> **Router Reflector**
- **MP-BGP session between RRs, not between PEs devices**
 - Only one MP-EBGP session between domains
 - To add new sites or services we only need:
 - To establish one IBGP session between the RR and each new PE device for labeled and the correspond family
 - Configure the routing instance for that site
- **No additional configuration between domains is required**

*Once we have one VPLS connection between domains, we can add
as many sites as we want easily*

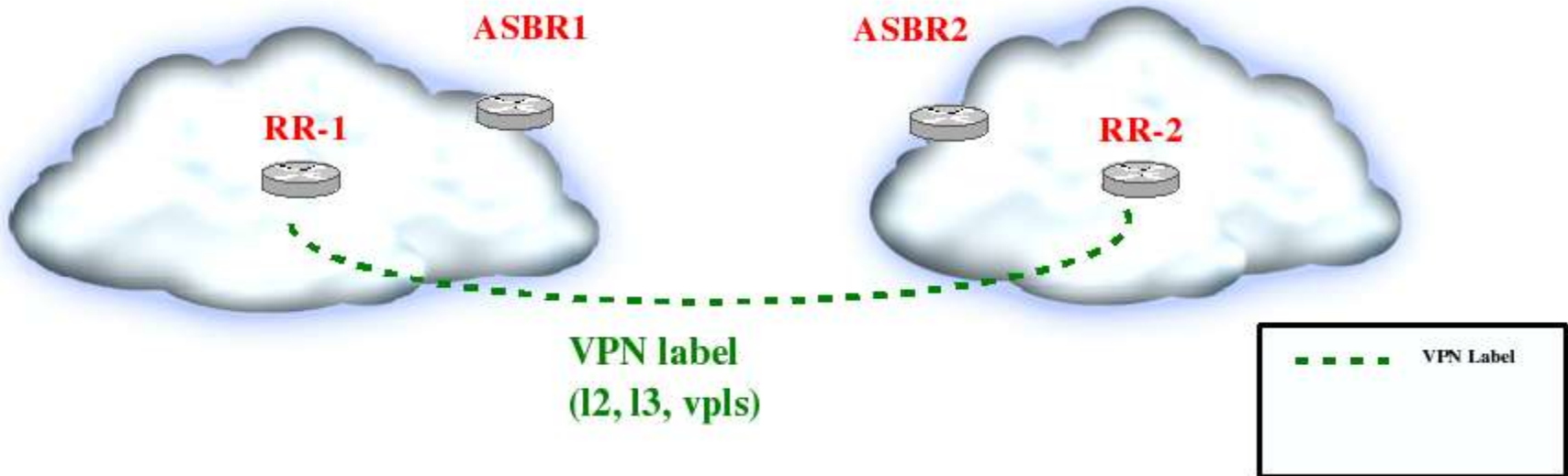
Configuration steps

- MP-BGP between PE routers (VPLS BGP)
 - **ONE** EBGP (vpls family) between RRs
- Extend the LSPs between domains ->BGP (labeled-unicast)



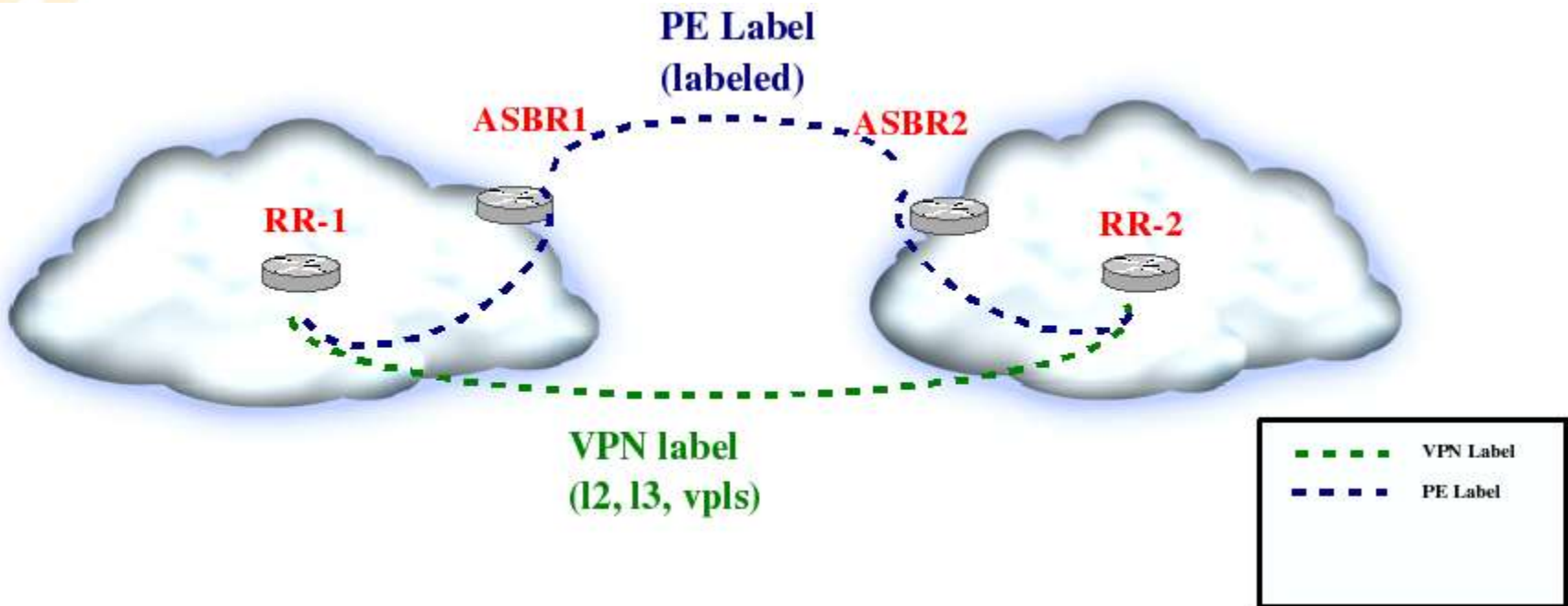
Summary of the configuration steps (I)

- Three level label stacks
 - **Level 1: VPN label -> (L2, L3, VPLS,...) -> MP-BGP**
 - Level 2: ASBR/PE label -> BGP labeled
 - Level 3: MPLS label -> MPLS-LSP



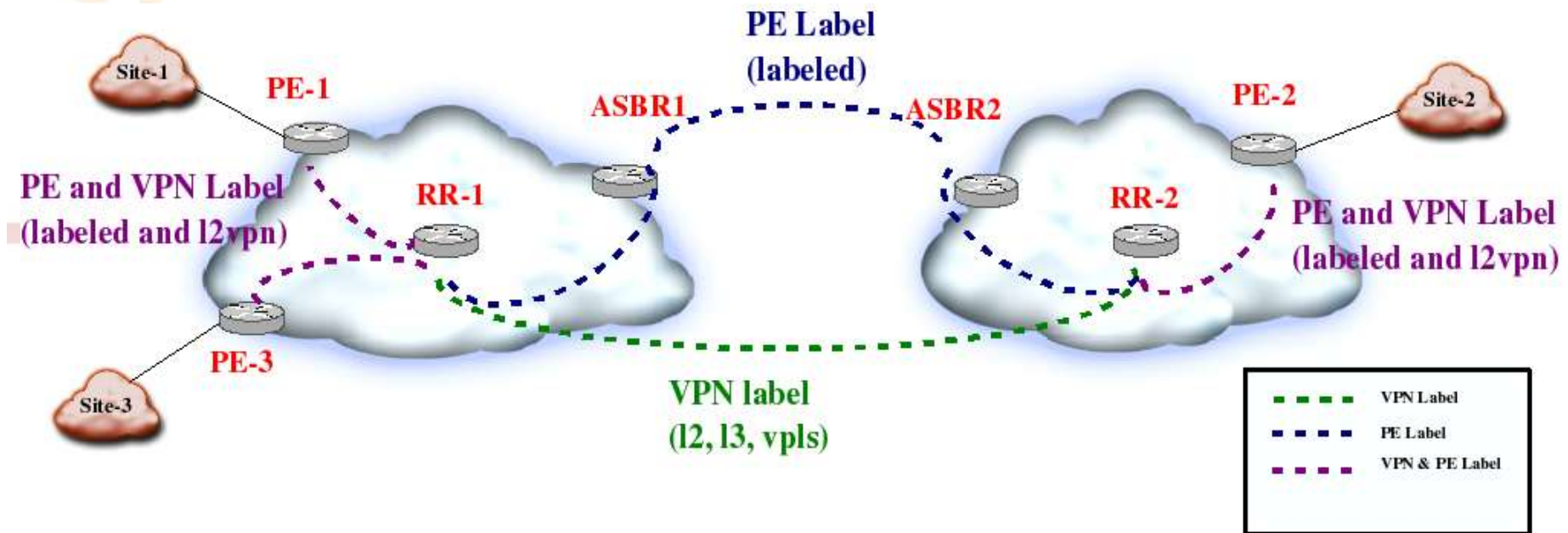
Summary of the configuration steps (II)

- Three control planes
 - Level 1: VPN label -> (L2, L3, VPLS,...) -> MP-BGP
 - Level 2: ASBR/PE Label -> BGP labeled
 - Level 3: MPLS label -> MPLS-LSP



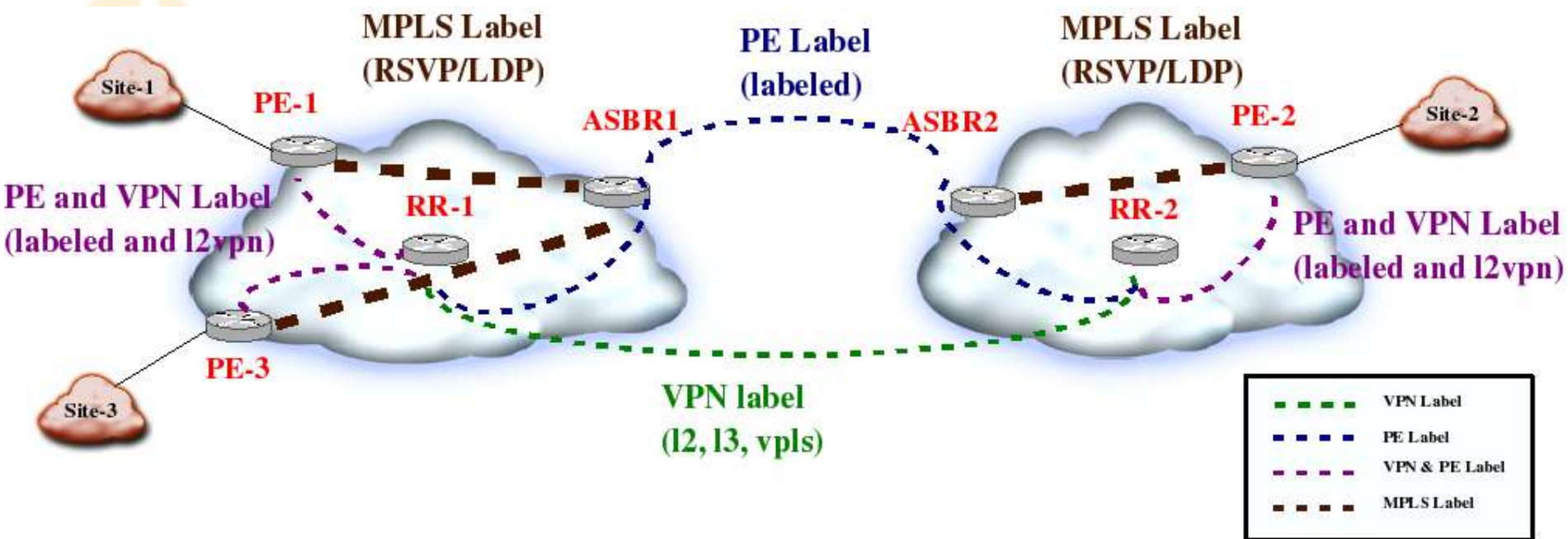
Summary of the configuration steps (III)

- Three control planes
 - Level 1: VPN label -> (L2, L3, VPLS,...) -> MP-BGP
 - Level 2: ASBR/PE Label -> BGP labeled -> Add new sites
 - Level 3: MPLS label -> MPLS-LSP



Summary of the configuration steps (IV)

- Three control planes
 - Level 1: VPN label -> (L2, L3, VPLS,...) -> MP-BGP
 - Level 2: PE Label -> BGP labeled
 - Level 3: MPLS label -> MPLS-LSP



Verify the bottom label

- VPLS BGP (I and E) sessions are established and PE devices receive the remote PE information

```
laura@IRIS2_router# show bgp summary
```

Peer	AS	InPkt	OutPkt	OutQ	Flaps	Last Up/Dwn	State	#Active/Received/Damped...
130.206.0.32	766	2434	2435	0	2	20:14:40	Establ	
bgp.l2vpn.0: 2/7/0								
VPLS-1.l2vpn.0: 2/4/0								
130.206.0.34	766	10707	10703	0	1	1d 1:01:04	Establ	
inet.0: 1/1/0								
193.144.32.251	64800	2436	2438	0	0	20:14:50	Establ	
bgp.l2vpn.0: 5/5/0								
VPLS-1.l2vpn.0: 2/2/0								

```
laura@IRIS2_router# show route receive-protocol bgp 193.144.32.251
```

```
VPLS-RedIRIS-CESGA.l2vpn.0: 6 destinations, 8 routes (6 active, 0 holddown, 0 hidden)
```

Prefix	NextHop	MED	Lclpref	AS path
193.144.32.251:501:5:1/96				
*	193.144.32.251			64800 I
193.144.32.251:501:5:65/96				
*	193.144.32.251			64800 I

```
bgp.l2vpn.0: 7 destinations, 12 routes (7 active, 0 holddown, 0 hidden)
```

Prefix	NextHop	MED	Lclpref	AS path
193.144.32.249:1:2:1/96				
*	193.144.32.251			64800 I
193.144.32.250:1:3:1/96				
*	193.144.32.251			64800 I
193.144.32.251:1:1:1/96				
*	193.144.32.251			64800 I
193.144.32.251:501:5:1/96				
*	193.144.32.251			64800 I
193.144.32.251:501:5:65/96				
*	193.144.32.251			64800 I

Verify the middle label

- labeled-unicast BGP (I and E) sessions are established and PE have the remote PE addresses in its inet.0 and inet.3

```
laura@Santiago0_router# show route receive-protocol bgp 193.144.32.252
```

```
inet.0: 5232 destinations, 13880 routes (5231 active, 0 holddown, 1 hidden)
  Prefix                Nexthop                MED      Lclpref    AS path
* 193.144.32.251/32    193.144.32.252        1                64800 I
```

```
inet.3: 4 destinations, 6 routes (3 active, 0 holddown, 1 hidden)
  Prefix                Nexthop                MED      Lclpref    AS path
* 193.144.32.251/32    193.144.32.252        1                64800 I
```

```
laura@IRIS2_router# show route receive-protocol bgp 130.206.0.34
```

```
inet.0: 155178 destinations, 214993 routes (155173 active, 4 holddown, 1 hidden)
  Prefix                Nexthop                MED      Lclpref    AS path
* 193.144.32.251/32    130.206.0.34          1        100        64800 I
```

```
inet.3: 3 destinations, 3 routes (3 active, 0 holddown, 0 hidden)
  Prefix                Nexthop                MED      Lclpref    AS path
• 193.144.32.251/32    130.206.0.34          1        100        64800 I
```

Verify top label

- MPLS LSP between PE and ASBR device are correctly configured

```
laura@IRIS2_router# show mpls lsp
```

```
Ingress LSP: 2 sessions
```

To	From	State	Rt	ActivePath	P	LSPname
130.206.0.32	130.206.0.33	Up	1	IRIS2-M0	*	LSP-IRIS2-M0
130.206.0.34	130.206.0.33	Up	1	IRIS2-IRIS4-Santiago0	*	LSP-IRIS2-IRIS4-Santiago0

Total 2 displayed, Up 2, Down 0

```
Egress LSP: 2 sessions
```

To	From	State	Rt	Style	Labelin	Labelout	LSPname
130.206.0.33	130.206.0.34	Up	0	1 FF	0	-	LSP-Santiago-IRIS4-IRIS2
130.206.0.33	130.206.0.32	Up	0	1 FF	0	-	LSP-M0-IRIS2

Total 2 displayed, Up 2, Down 0

```
laura@IRIS2_router# show route 130.206.0.34
```

```
inet.0: 155180 destinations, 214993 routes (155167 active, 12 holddown, 1 hidden)
+ = Active Route, - = Last Active, * = Both
```

```
130.206.0.34/32    *[RSVP/7] 1d 01:46:49, metric 51
> via so-1/0/0.0, label-switched-path LSP-IRIS2-IRIS4-Santiago0
[IS-IS/18] 1d 01:46:51, metric 51
> to 130.206.240.2 via so-1/0/0.0
```

```
inet.3: 3 destinations, 3 routes (3 active, 0 holddown, 0 hidden)
+ = Active Route, - = Last Active, * = Both
```

```
130.206.0.34/32    *[RSVP/7] 1d 01:46:49, metric 51
> via so-1/0/0.0, label-switched-path LSP-IRIS2-IRIS4-Santiago0
```

Do not forget

- The PE devices need to have a **Tunnel Services Physical Interface Card**
- The technology in both ends of the VPLS connection **must be the same**
 - This restrictions can be solved with some QPP PIC to do translating (vlan swap, push and pop operations)
- To have in mind the “nex-hop” value -> no-nextthop-change

red.es



Keep working

- Install a complete GRID testbed using nodes connected using VPLS interdomain. Find some other AS interested.
- Measure the impact of the different network technologies in Grid terms.
- Install a remote boot server and user it to boot remote nodes using a VPLS connection.
- Establish mechanism and procedures to manage and monitoring this testbed..
- ...

red.es

Thanks for your attention!!
¡Gracias por su atención!
Merci pour votre attention!
¡Grazas pola sua atención!

red.es

