BGP/MPLS IP VPNs

Ina Minei

ina@juniper.net

Disclaimer

- The views presented are of the author and do not necessarily represent Juniper Networks.

Topics

1. **VPN basic concepts**
2. **Hierarchical and recursive applications**

Part 1 – Basic concepts

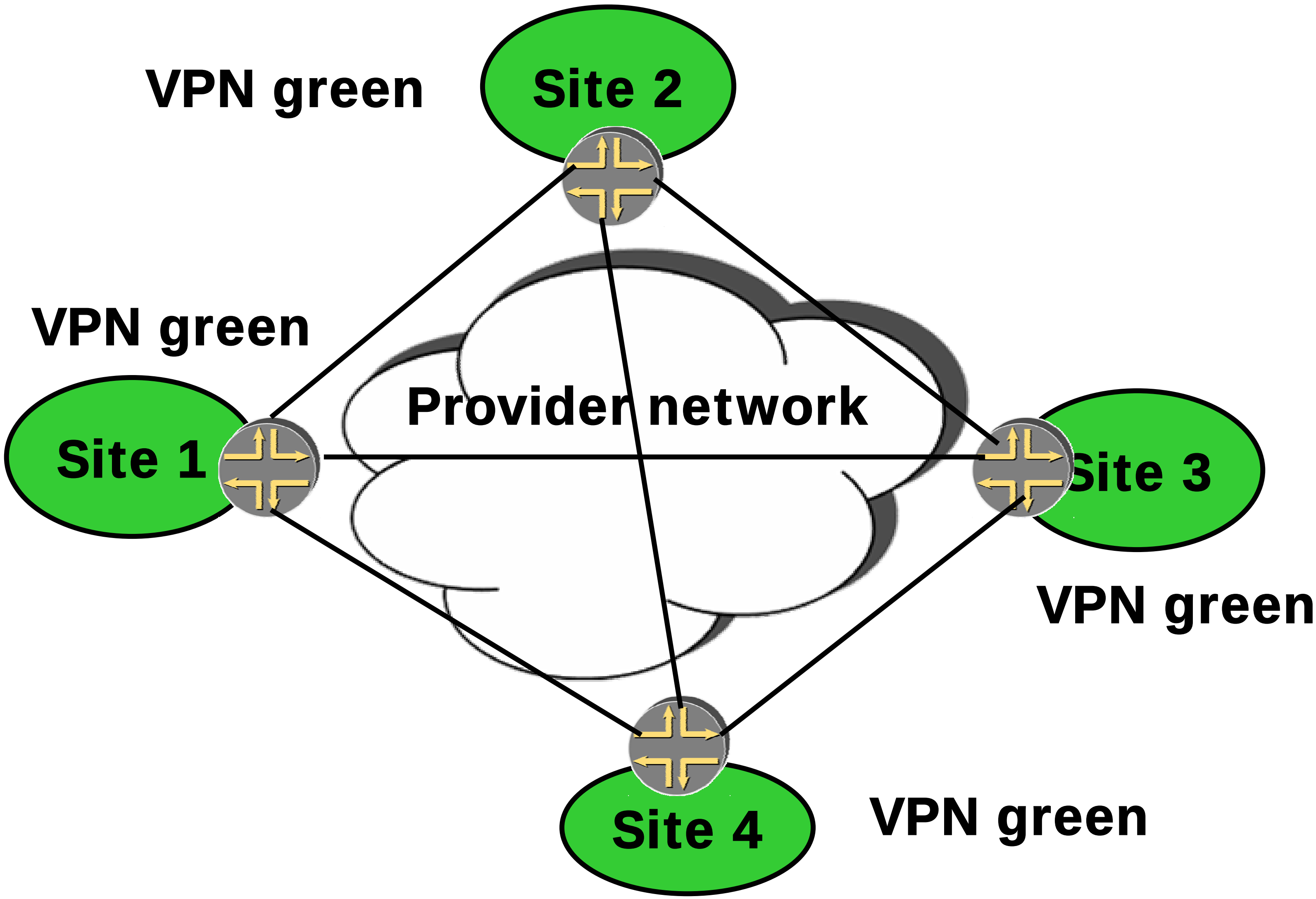
- Introduction
- How it works
- Scalability
- Connectivity models

VPNs

- Virtual Private Networks – provide a private network over a shared infrastructure.

- Interconnect geographically separate sites, with the same privacy and guarantees as a private network.

VPNs


VPN green
Site 2
VPN green
Provider network
Site 1
Site 3
VPN green
Site 4
VPN green

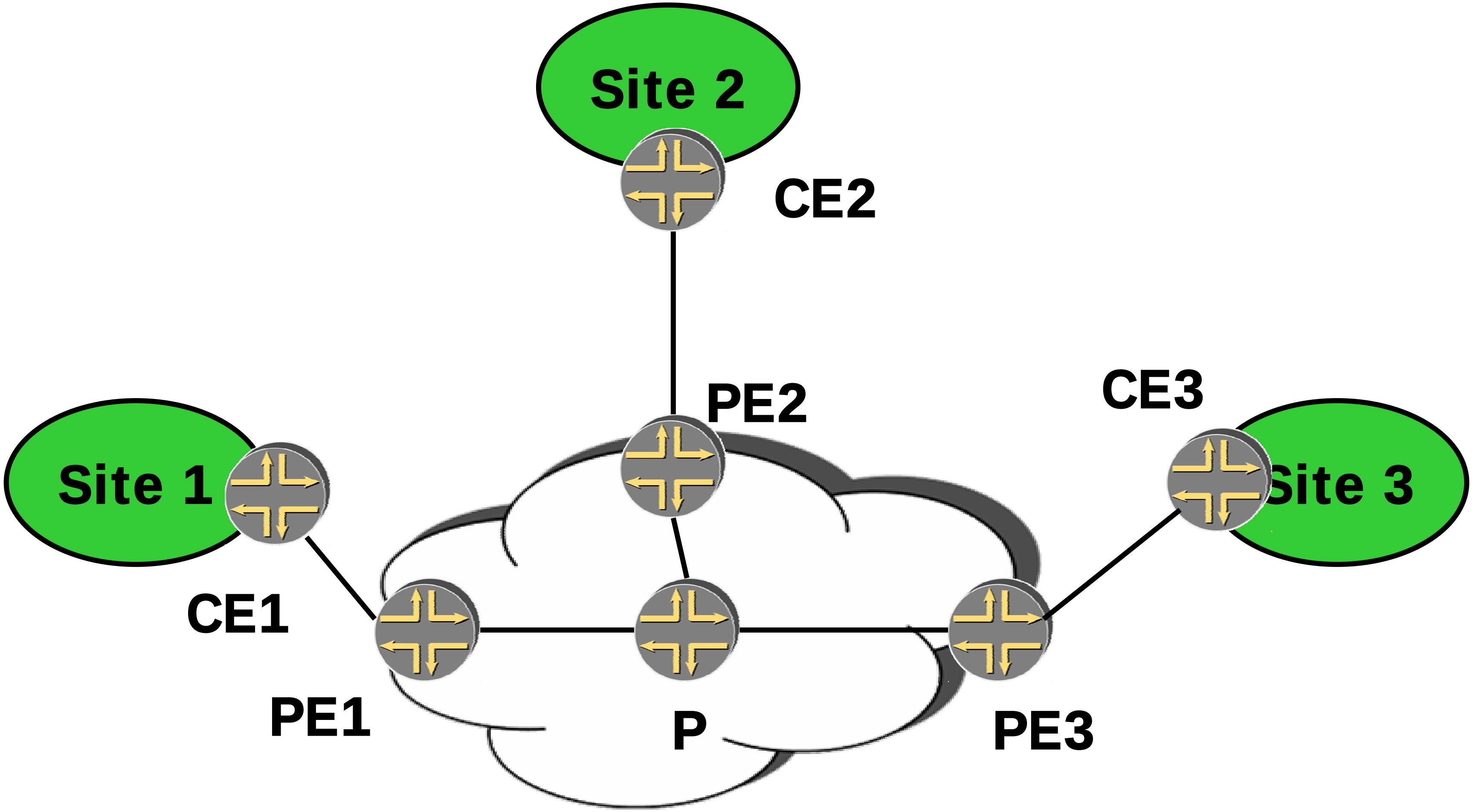
The Overlay Model for VPNs

- Sites are connected with p2p links – leased lines, FR circuits, ATM circuits, GRE, IPsec.
- Customer routers peer with customer routers.
- The provider needs to design and operate “virtual backbones” for all the customers – scaling issue.
- Problem with VPNs that have a large number of sites.
- Adding a new site requires configuring all the existing sites.

BGP-MPLS VPNs

- **Goal: solve the scaling issues. Support thousands of VPNs, support VPNs with hundreds of sites per VPN, support overlapping address space.**
- **Peer model – customer routers peer with provider routers.**

Terminology



Properties of the model

- CE router peers with a PE router, but not with other CE routers.
- Adding/deleting a new site requires configuring the PE router connected to the site.
- A PE router only needs to maintain routes for the VPNs whose sites are directly connected.

Goals

- Achieve intersite connectivity
- Privacy – don't allow traffic from one VPN to be seen in another VPN
- Independent addressing – private addresses in each VPN.

Part 1 – Basic concepts

- Introduction
- How it works
- Scalability
- Connectivity models

BGP-MPLS VPNs - areas

- Separation of forwarding
- Distribution of routing information
- New address type
- Forwarding with MPLS

Operation – separation of forwarding

- **Goal: control connectivity and ensure privacy by segregating the forwarding information.**
- **PE router connected to CEs from several VPNs.**
- **With a single forwarding table, it is possible to forward packets from one VPN to another.**

Multiple forwarding tables

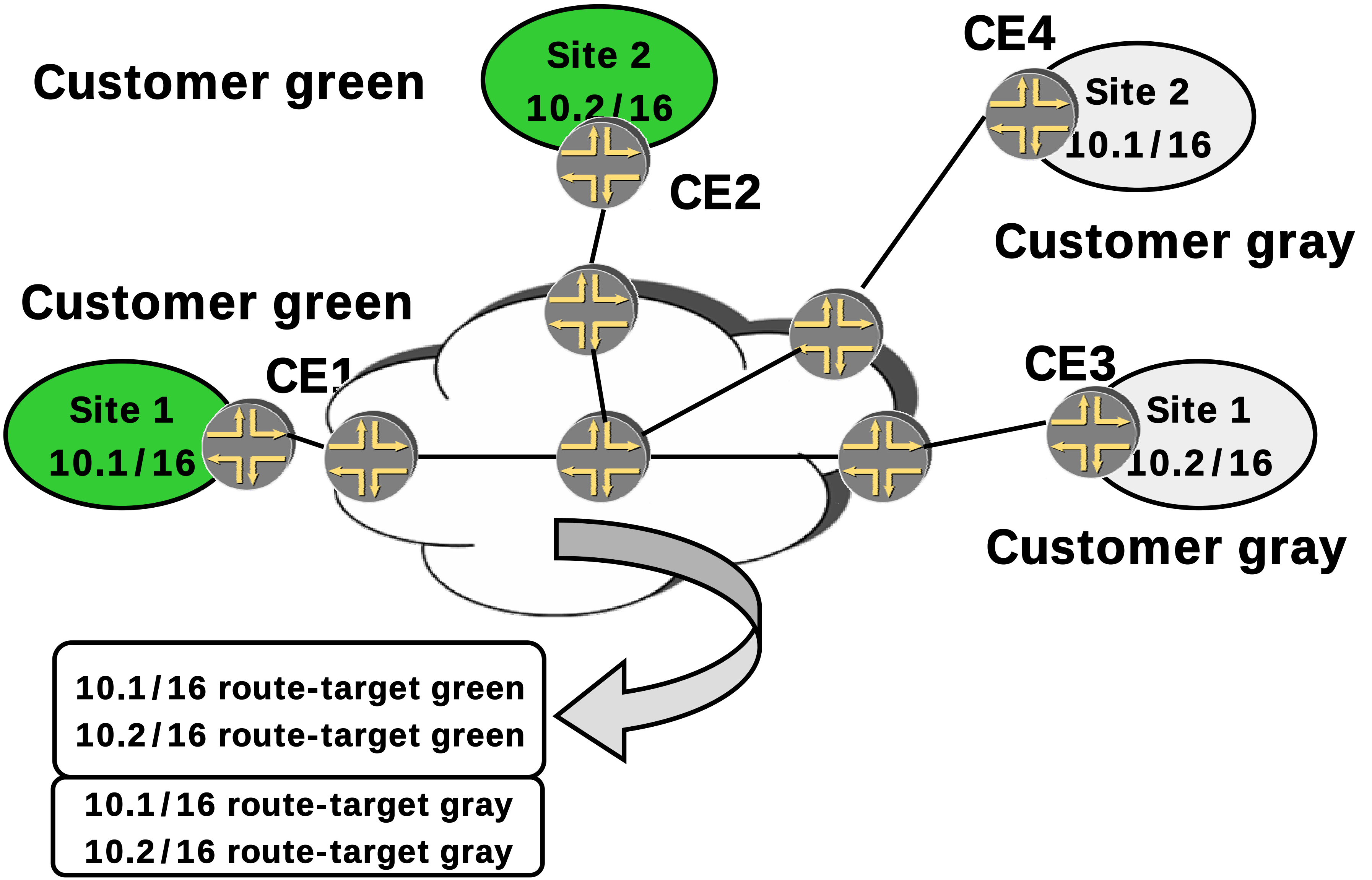
- Multiple forwarding tables – each table associated with a site.
- Packets from the customer are identified based on the incoming port, which identifies the forwarding table.
- Contents: routes received from the CE, and routes received from remote PEs with constrained routing.
- Called VPN routing and forwarding table – VRF.

Operation – Constrained distribution of routing information

- **The idea:**

1. **CE advertises routes to the local PE via some routing protocol.**
2. **The local PE marks these routes with a particular extended community (route target) and advertises them in BGP.**
3. **The routes are distributed to all remote PE by BGP.**
4. **Remote PE receives BGP routes, filters them based on the community and advertises them to the CE.**

Constrained route distribution – the need for unique addresses



The model so far (1)

- **The P routers carry all VPN routes, so the addresses used in the VPNs need to be unique in the provider's network.**

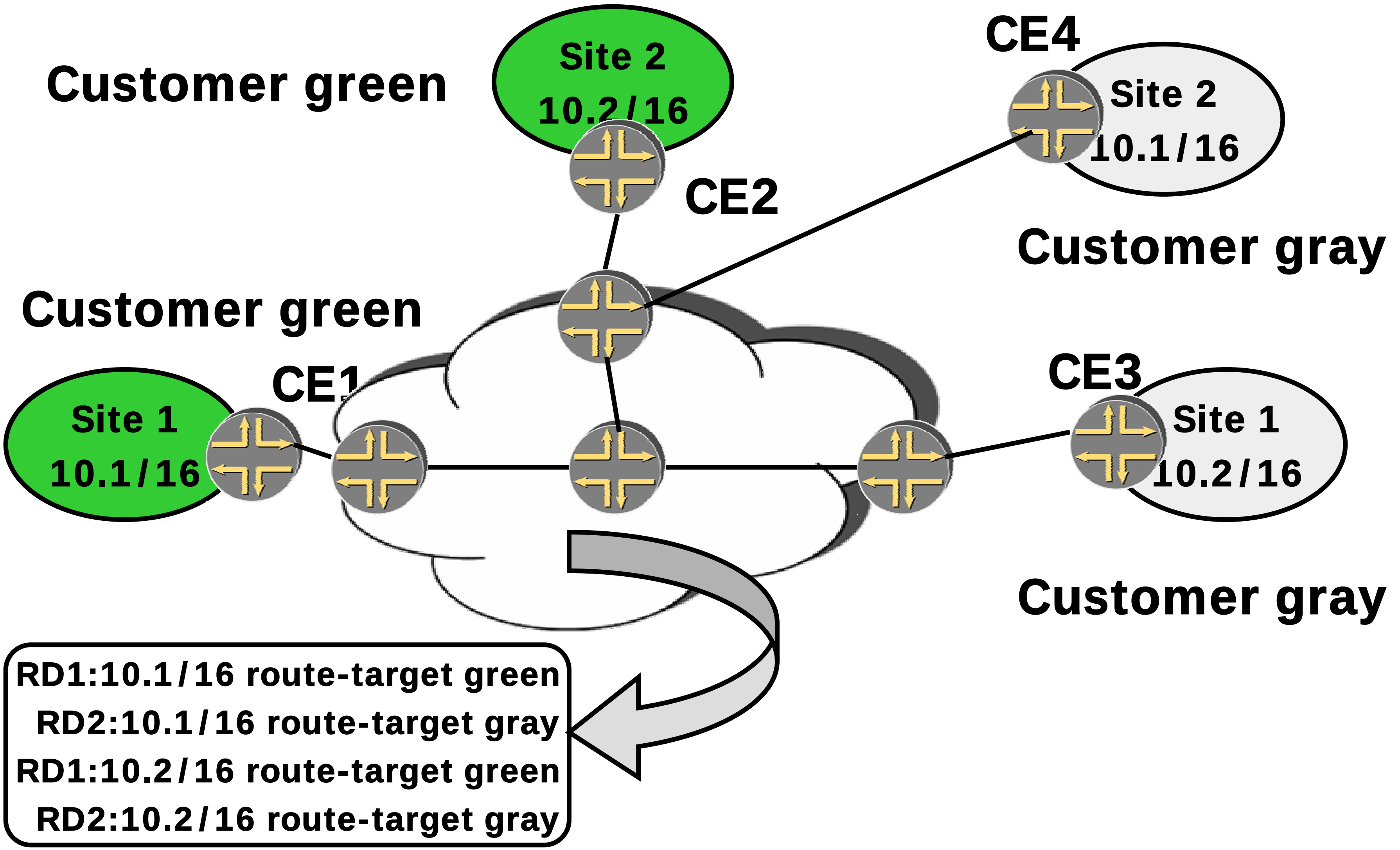
Operation: overlapping address space and VPN-IP addresses

- Goal: turn non-unique addresses into unique addresses.
- Constructed by concatenating an IP address and an 8 byte unique identifier called the route distinguisher.
- Route Distinguisher – 8 bytes – doesn’t have to be the same for all routes in the VPN. Typical values: either AS:number or IPaddress:number.

VPN-IP addresses (cont)

- Advertised in a special address family by BGP (MP-BGP)
- Used only in the provider's network.
- Used only in the control plane.
- The translation from IP addresses to VPN-IP addresses happens on the PE.
- Not used for route filtering (we use communities for that).

Example using VPN-IP addresses



The model so far (2)

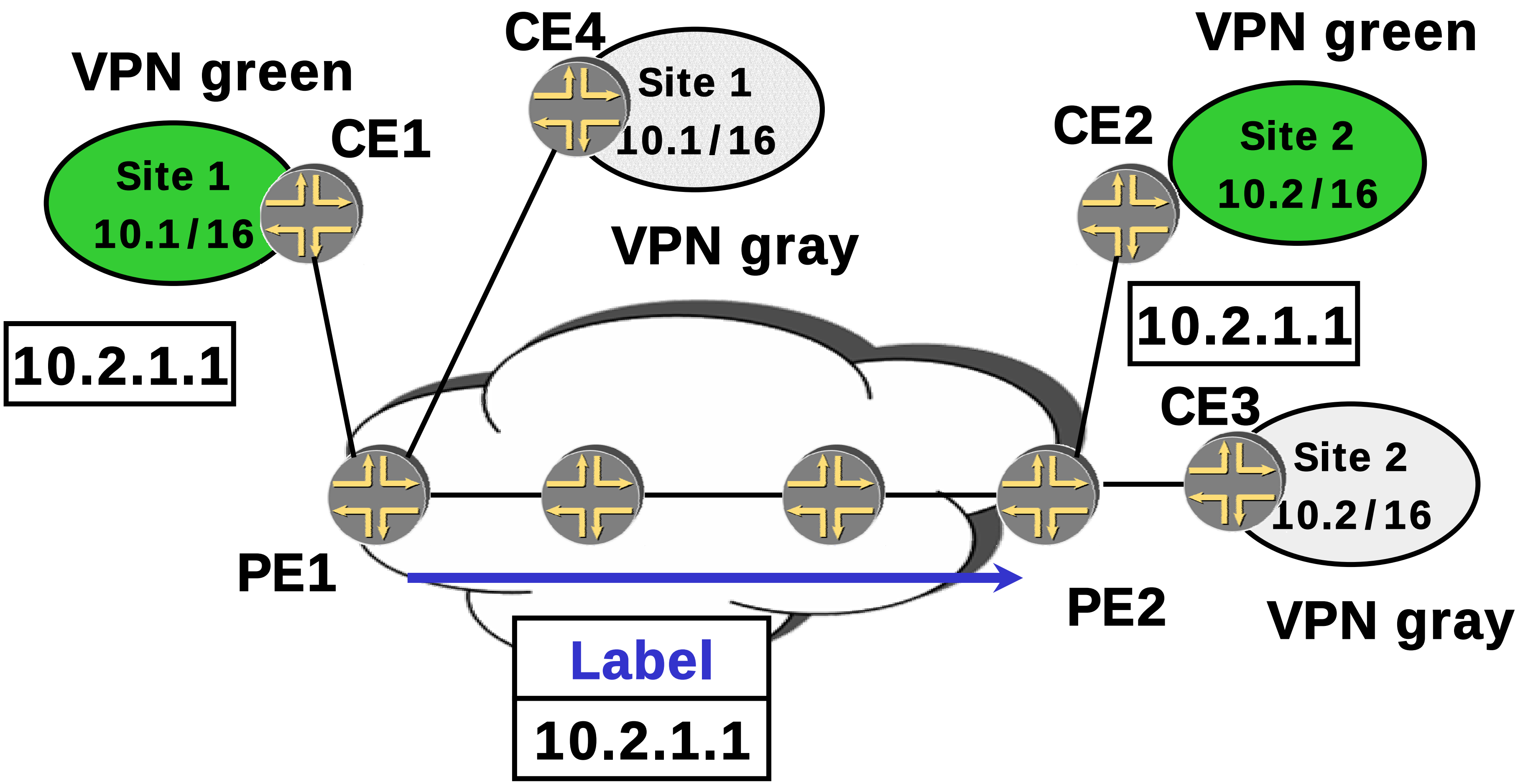
- Can use overlapping address space.
- How to forward based on VPN-IP addresses?
- The P routers still carry all the VPN routes.

Why MPLS?

- VPN-IP addresses are used by the routing protocols, but do not appear in headers of IP packets.

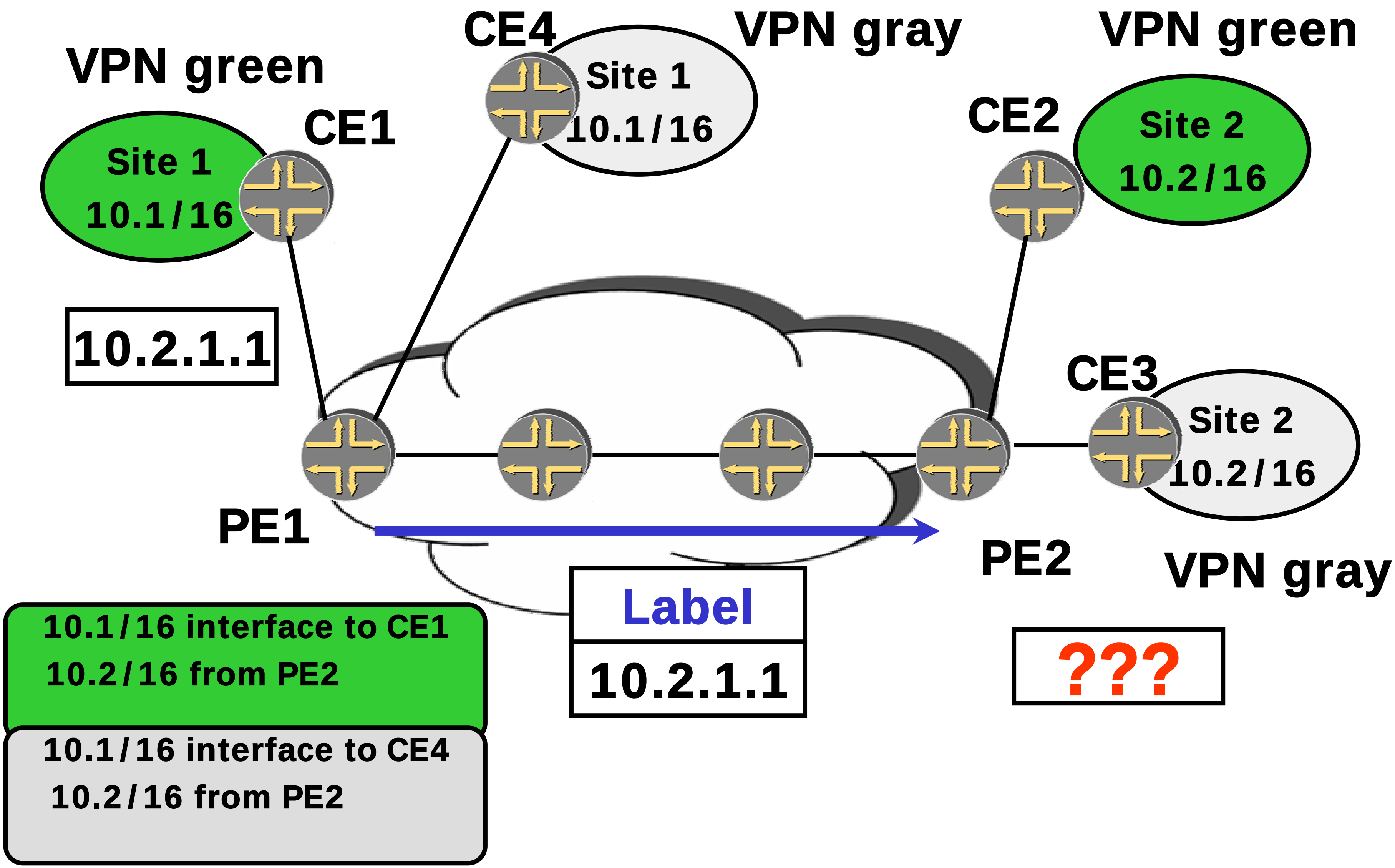
- Need a way to forward traffic along routes to VPN-IP addresses. MPLS decouples forwarding from the destination information.

Forwarding traffic - so far (1)



10.1/16 interface to CE1
10.2/16 from PE2

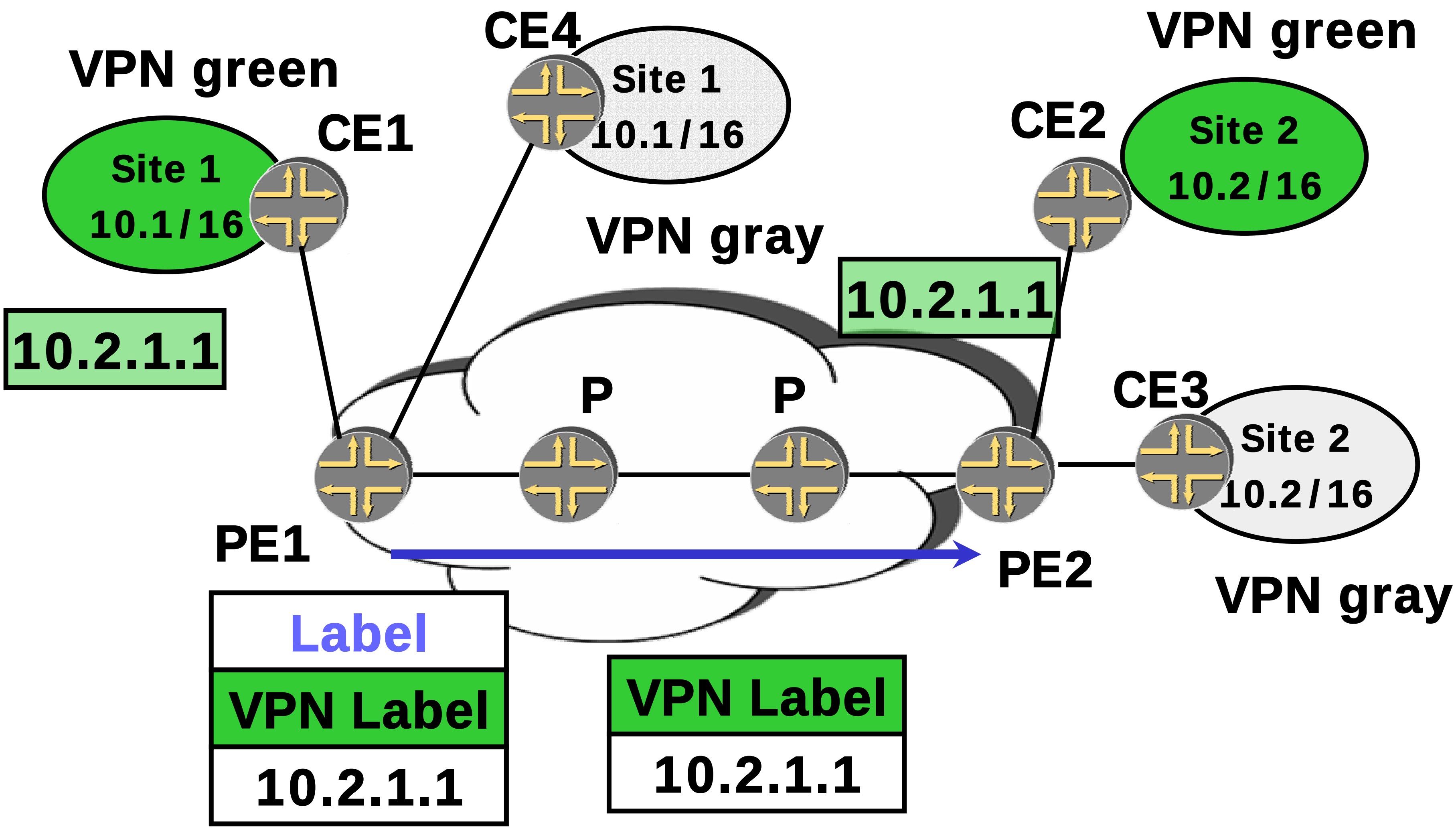
Forwarding traffic - so far (2)



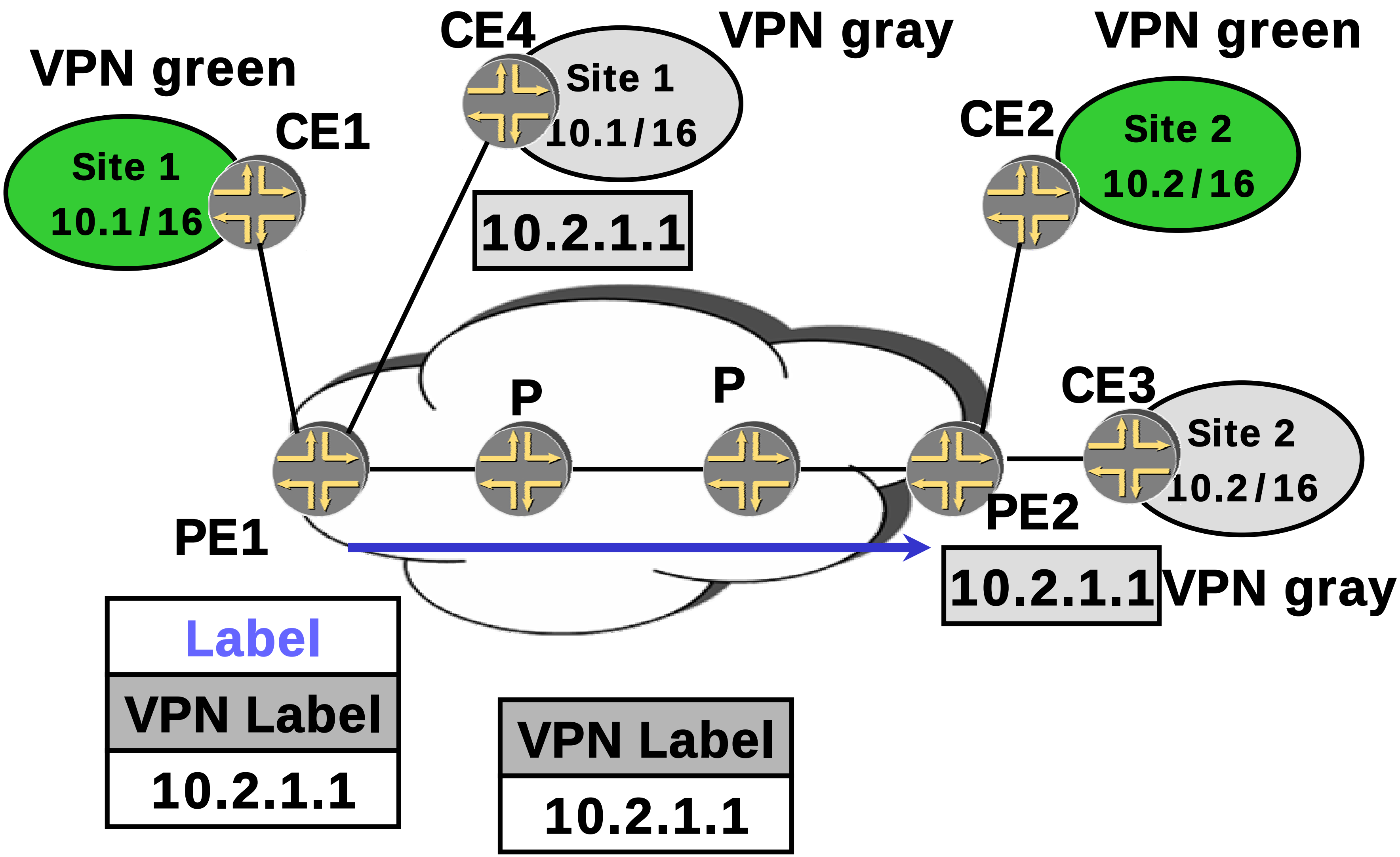
VPN labels

- The idea: Use a label to identify the next-hop at the remote PE. Also called VPN label.
- The label is distributed by BGP, along with the VPN-IP address.
- Traffic will carry two labels, the VPN label and the LSP label.
- The remote PE makes the forwarding decision based on the VPN label.

Forwarding traffic - revisited



Forwarding traffic - revisited



The VPN model - summary

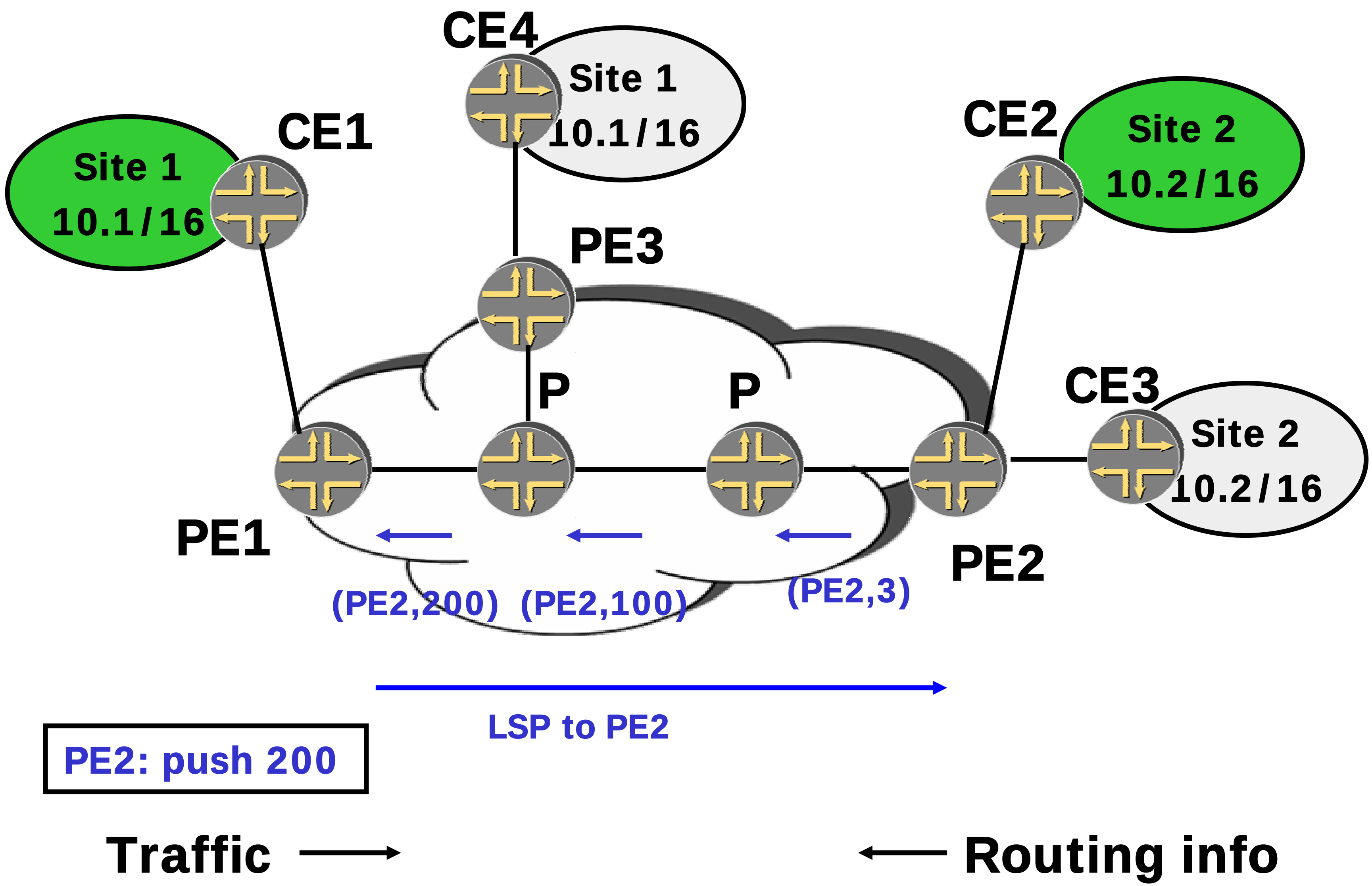
- P routers don’t need to maintain VPN routes at all. Only need to maintain routes to other P and PE routers.
- PE routers maintain VPN routes, but only for VPNs that have sites attached to them.
- VPNs can have overlapping address spaces.

Routing exchanges / traffic forwarding

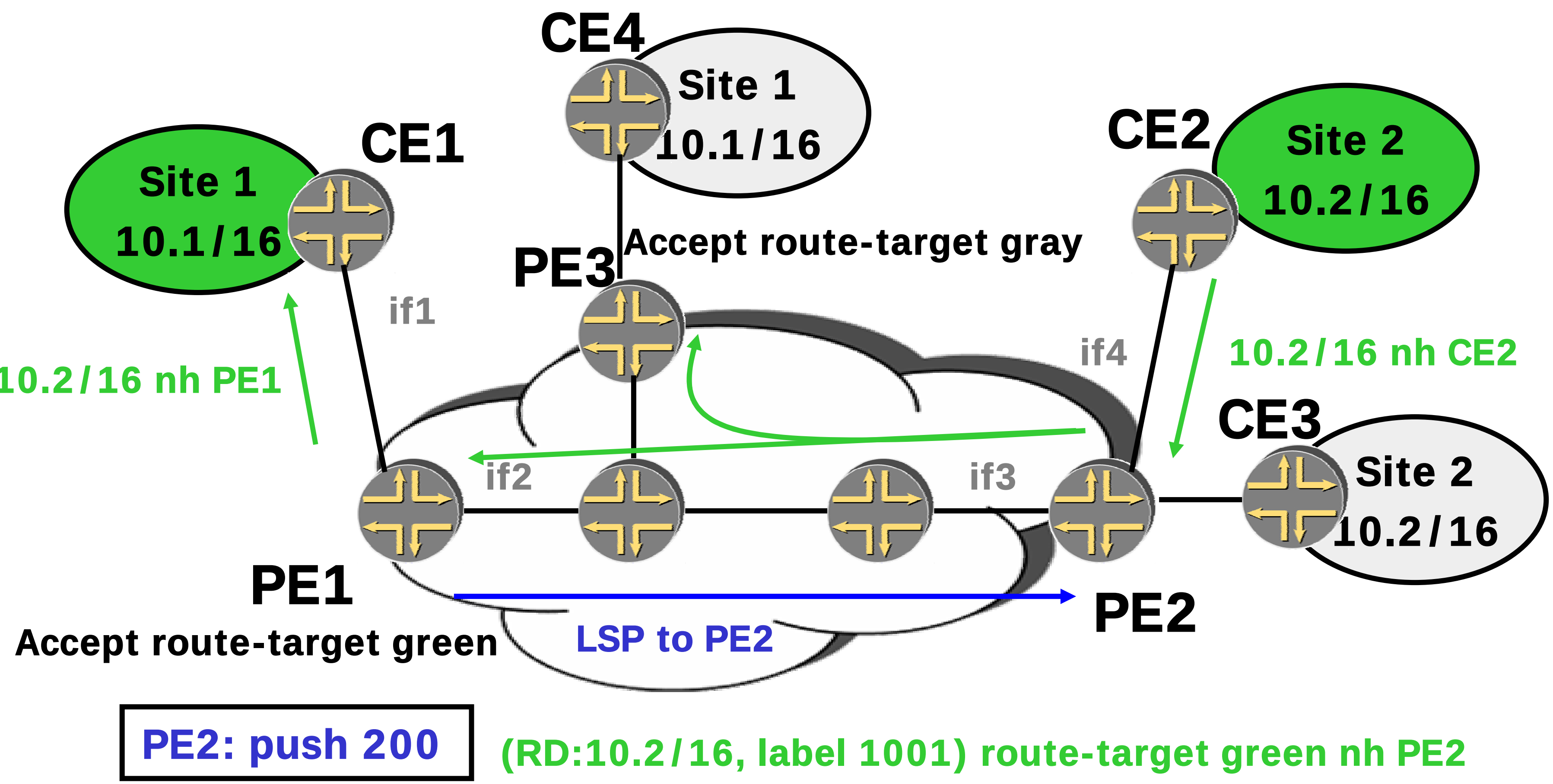
← Routing info

Traffic →

The whole picture 1 – LSP setup



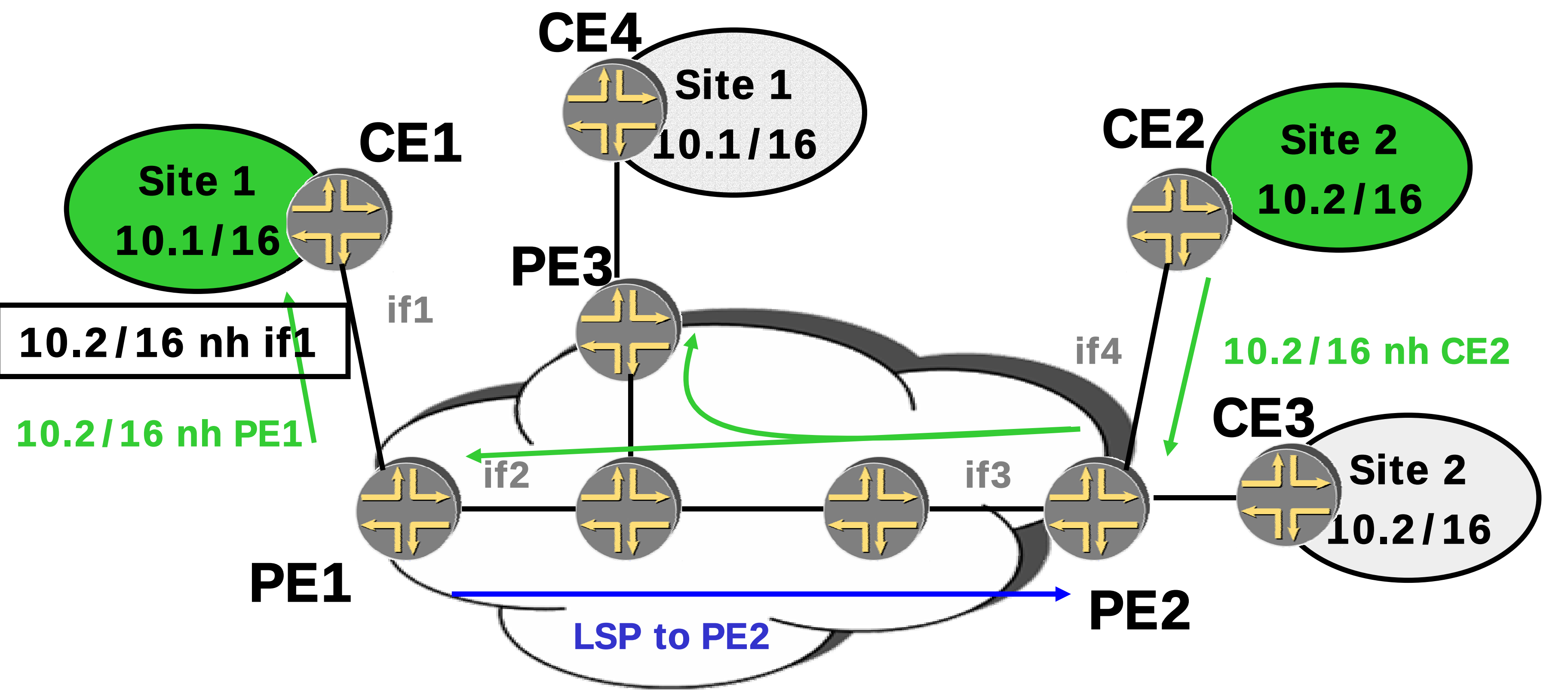
The whole picture 2 - route distribution



Traffic →

← Routing info

The whole picture 3- forwarding tables



(RD:10.2/16, label 1001) route-target green nh PE2

10.2/16: nh PE2, 1001

VRF green

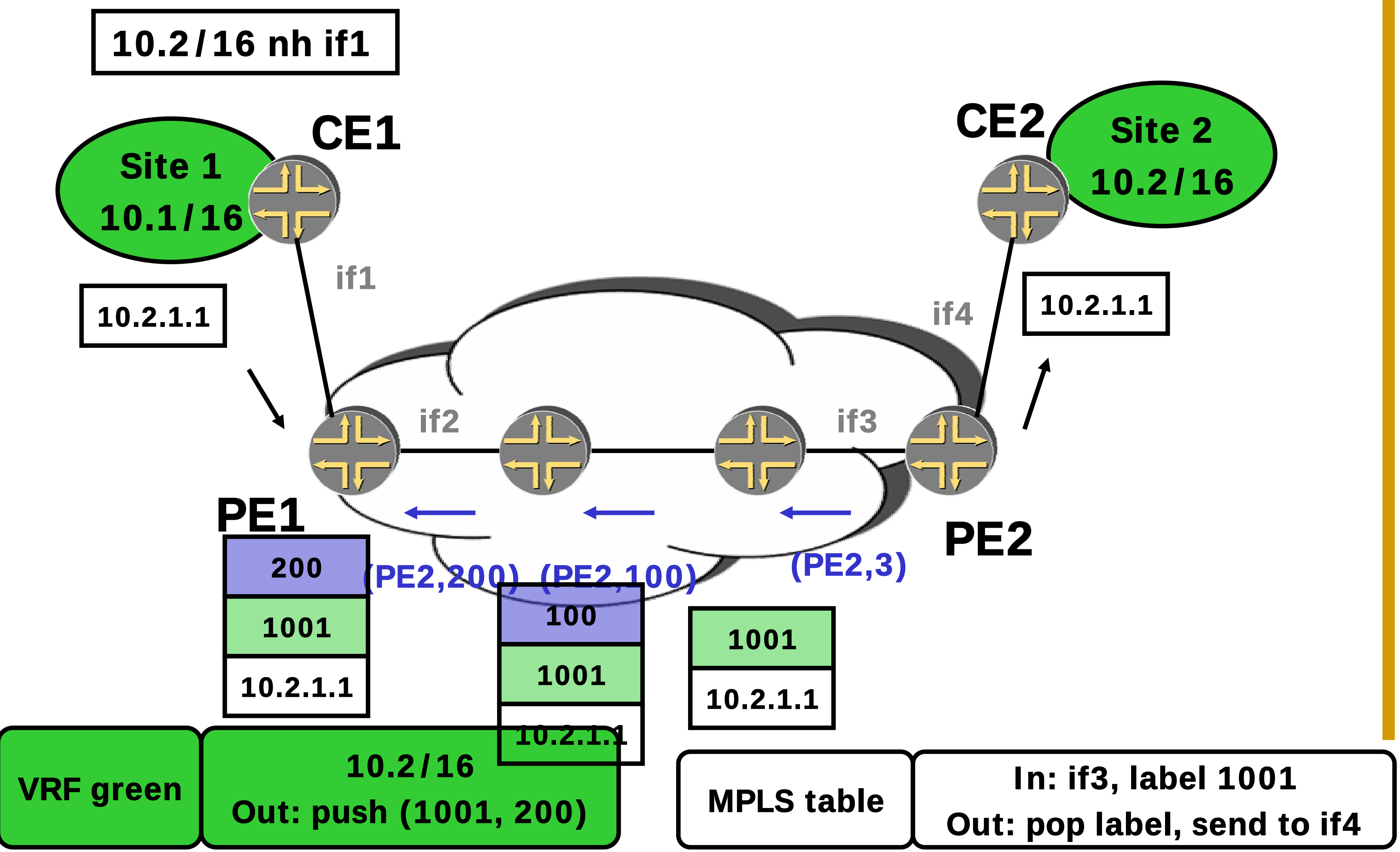
10.2/16

Out: push (1001, 200)

PE2: push 200

MPLS table	In: if3, label 1001 Out: pop label, send to if4

The whole picture 4- forwarding traffic



The whole picture - summary

- Full mesh of BGP between all PEs.
- MPLS connectivity between all PEs.
- BGP advertises a label along with the VPN-IP address. This determines the next-hop to use when receiving traffic.

Concepts

1. Use MPLS to forward traffic across nodes that don’t have routing information for the packet’s final destination.
2. Use a label to mark the traffic. Use this marking to determine the next-hop.
3. The address of the next-hop in the BGP advertisement provides coupling between the VPN routes and the internal routing to the remote PE.

Part 1 – Basic concepts

- Introduction
- How it works
- Scalability
- Connectivity models

Scaling properties

- Only one routing peering (CE-PE), regardless of the number of sites in the VPN.
- The customer doesn’t need routing skills. A customer doesn’t need to operate its own backbone.
- Adding a new site requires configuration of one PE regardless of the number of sites (constant # of changes required to add a new site)

Scaling properties

- PE has to maintain routes only for the VPNs to which it is connected.
- P routers don’t have to maintain VPN routes at all.

Scaling properties

- Can use overlapping address spaces – efficient use of private IP addresses.
- Route distinguishers are structured so that each service provider can manage its own number space.

Part 1 – Basic concepts

- Introduction
- How it works
- Scalability
- Connectivity models

Intersite connectivity

- Achieved through constrained distribution of routing information.
- Done by the PE:
  - No expertise required from the customer.
  - No configuration necessary on the customer box.
- Extended communities allow definition of very flexible policies.

Intersite connectivity models

- Connectivity models
  - Any-to-any
  - Hub and spoke
- Any other combination also possible.

Hub and spoke

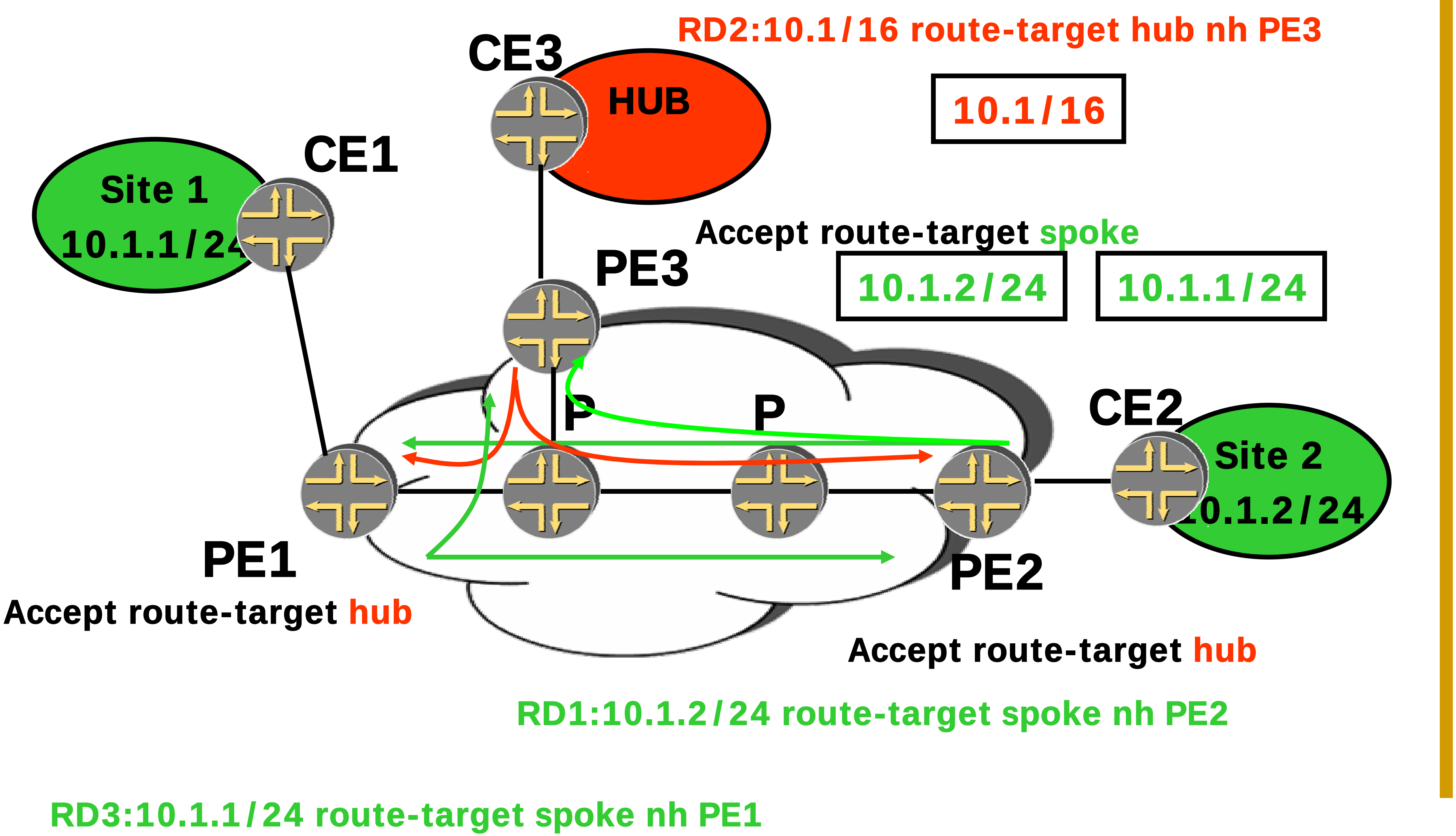
- The goal: make all the traffic originated at spoke sites go through one hub site (e.g. for implementing a firewall)
- The hub site has knowledge of all destinations in the spoke sites.

Hub and spoke

- Spoke sites export routes to the hub site using the “spoke” route target.
- The hub site re-exports these routes with a “hub” route target.

- Spoke sites only import routes with community “hub”.
- Traffic will flow from the spoke sites through the hub.

Hub and spoke



Part 2 – Hierarchical and recursive applications

- Introduction
- ISP as a VPN customer
- VPN service provider as a VPN customer
- VPN services across AS boundaries

Introduction

- VPN customer is himself a service provider: ISP or VPN service provider.
- Carriers carrier – all customer sites are in the same AS.
- Multi-AS operations – the customer sites have different AS numbers (VPN service spans two providers)

Introduction - terminology

- External routes – learned from peering points or from customers. Carried in BGP.

- Internal routes – include the provider's internal links (including BGP next-hops) and loopbacks. Carried in the IGP.

Concepts we saw previously

- Use MPLS to forward traffic across nodes that don’t have routing information for the packet’s final destination.
- Use a label to mark traffic. Use this marking to pick the correct next-hop.
- The BGP next-hop is the glue between external routes and internal routes.

Part 2 – Hierarchical and recursive applications

- Introduction
- ISP as a VPN customer
- VPN service provider as a VPN customer
- VPN services across AS boundaries

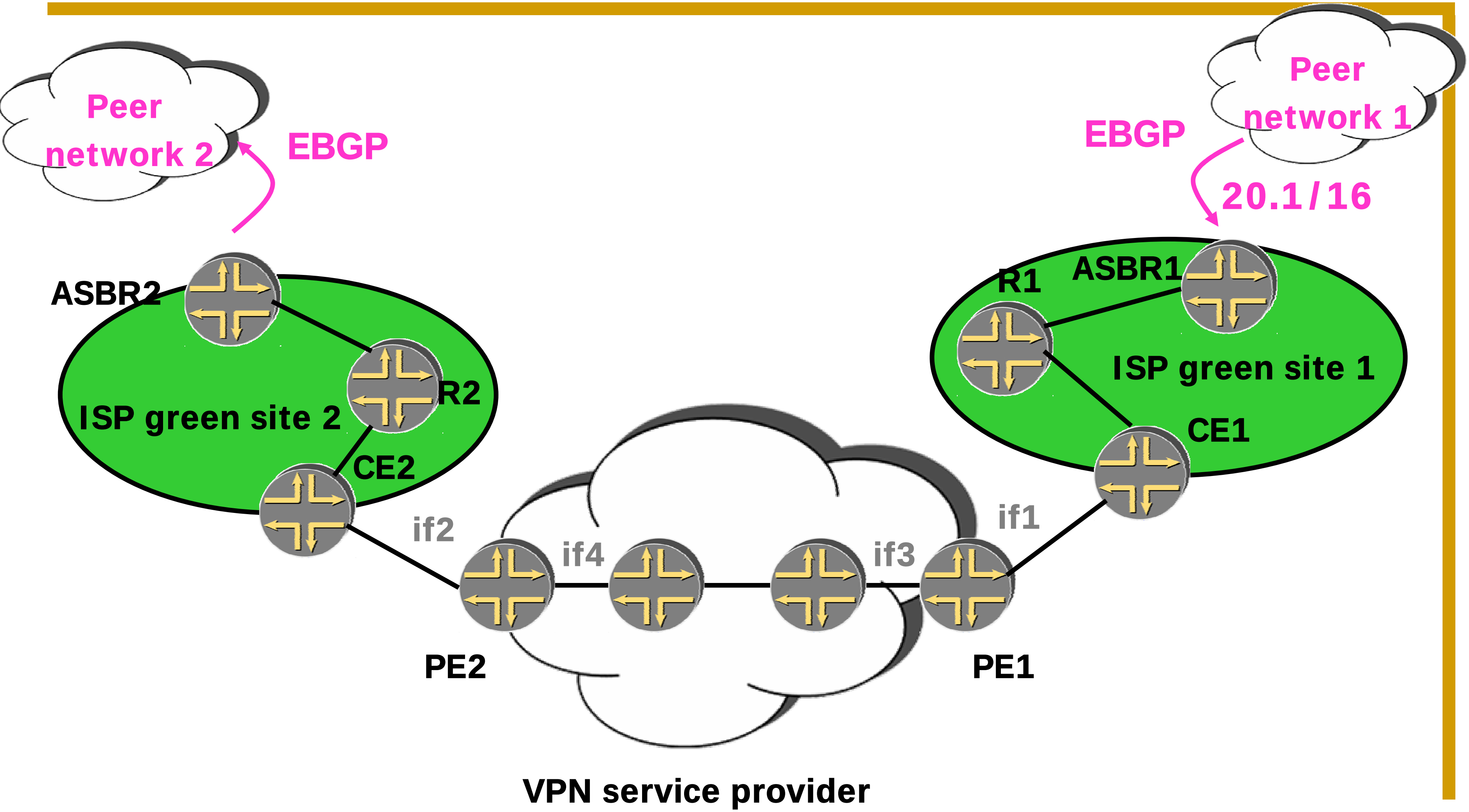
ISP as a VPN customer

- Goal – interconnect geographically separate sites of the ISP (e.g. POPs).
- Also known as “Carriers carrier”, section 9 of 2547bis.
- Two scenarios:
  - No MPLS within the sites – within a site, forward based on IP.
  - With MPLS within the sites – can use MPLS to forward within a site.

ISP as a VPN customer – step (1)

- **The problem:**
  - Requires the PE routers to carry a full set of internet routes as VPN-customer routes… for each such customer…
  - Requires the VPN provider to distribute the routes for each of the customers throughout the network (large amount of routing information).
- **The solution:** let the customer be responsible for the external routes.

ISP as a VPN customer – no MPLS within sites



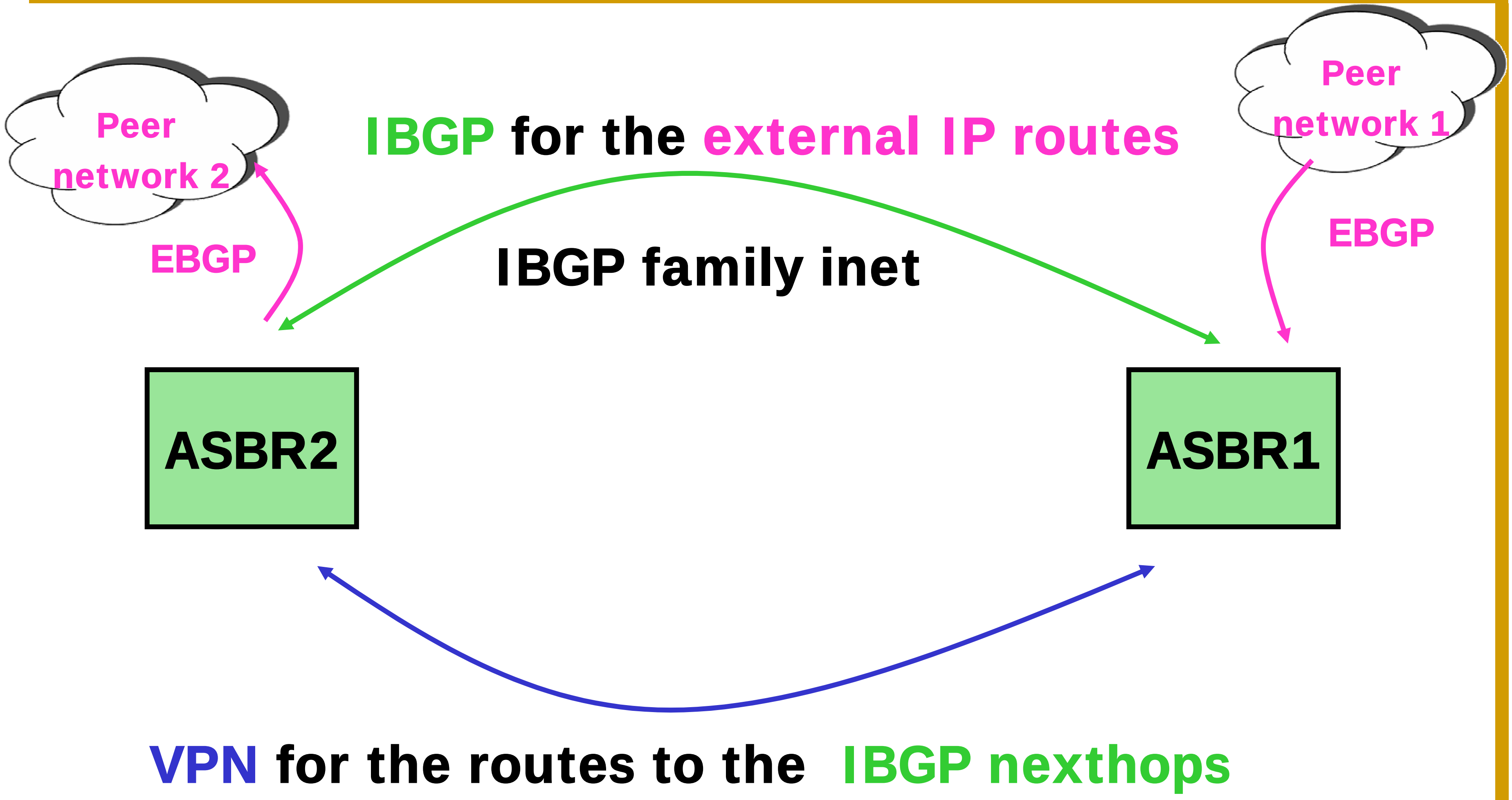
ISP as a VPN customer – step (1)

- **External routes are exchanged via BGP between the two geographically dispersed sites.**
- **Need to be able to establish BGP sessions across the VPN provider => must have routes to the routers in the other POP.**
- **Advertise the internal routes as the VPN customer routes.**

ISP as a VPN customer – scenario 1 – no MPLS within the customer sites

- No MPLS in the customer sites.
- Goal – the provider doesn't want to carry the customer's external routes.

The abstraction – routing (no MPLS within the customer sites)



Routing exchanges / traffic forwarding

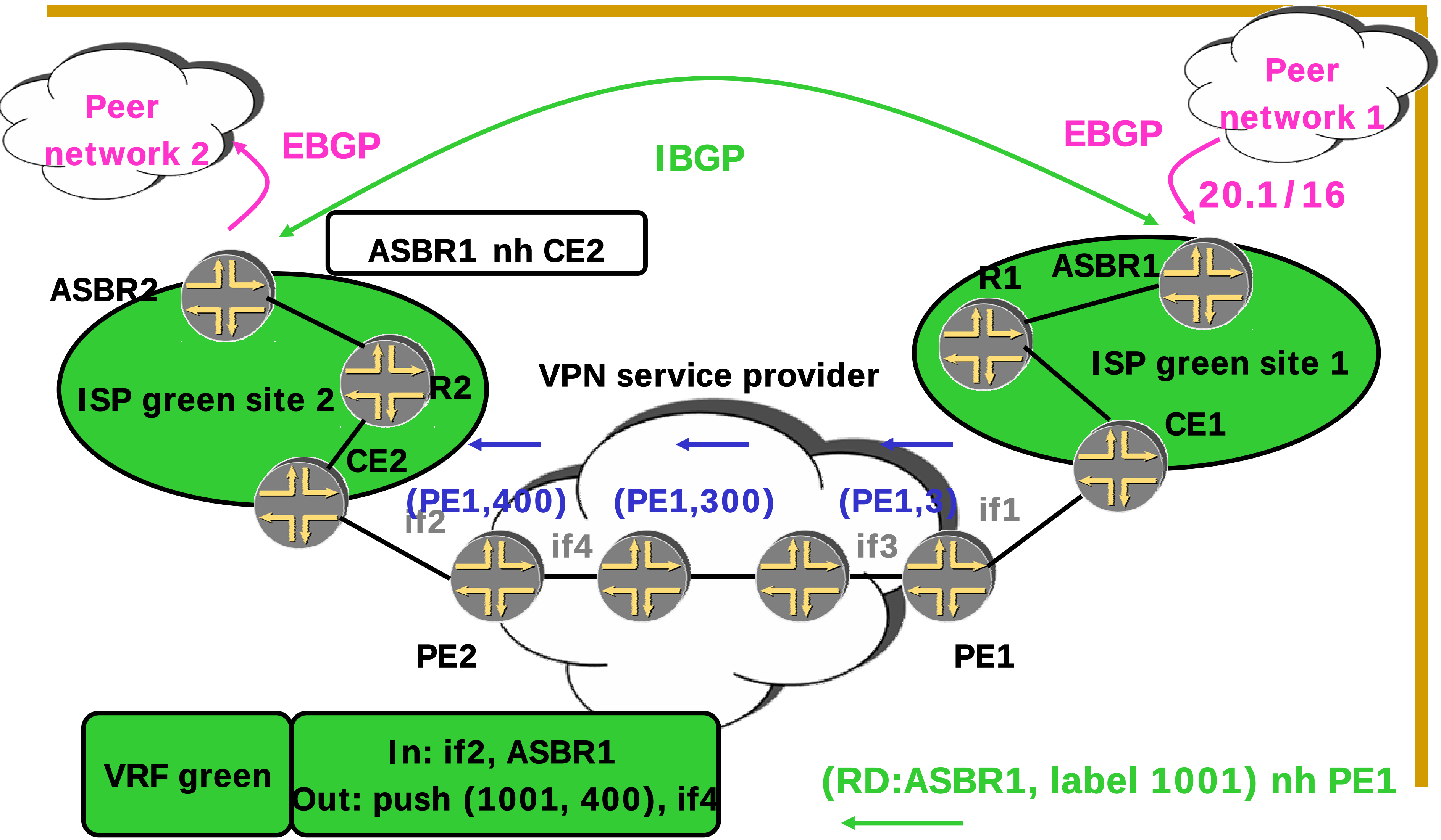
Site 2

Site 1

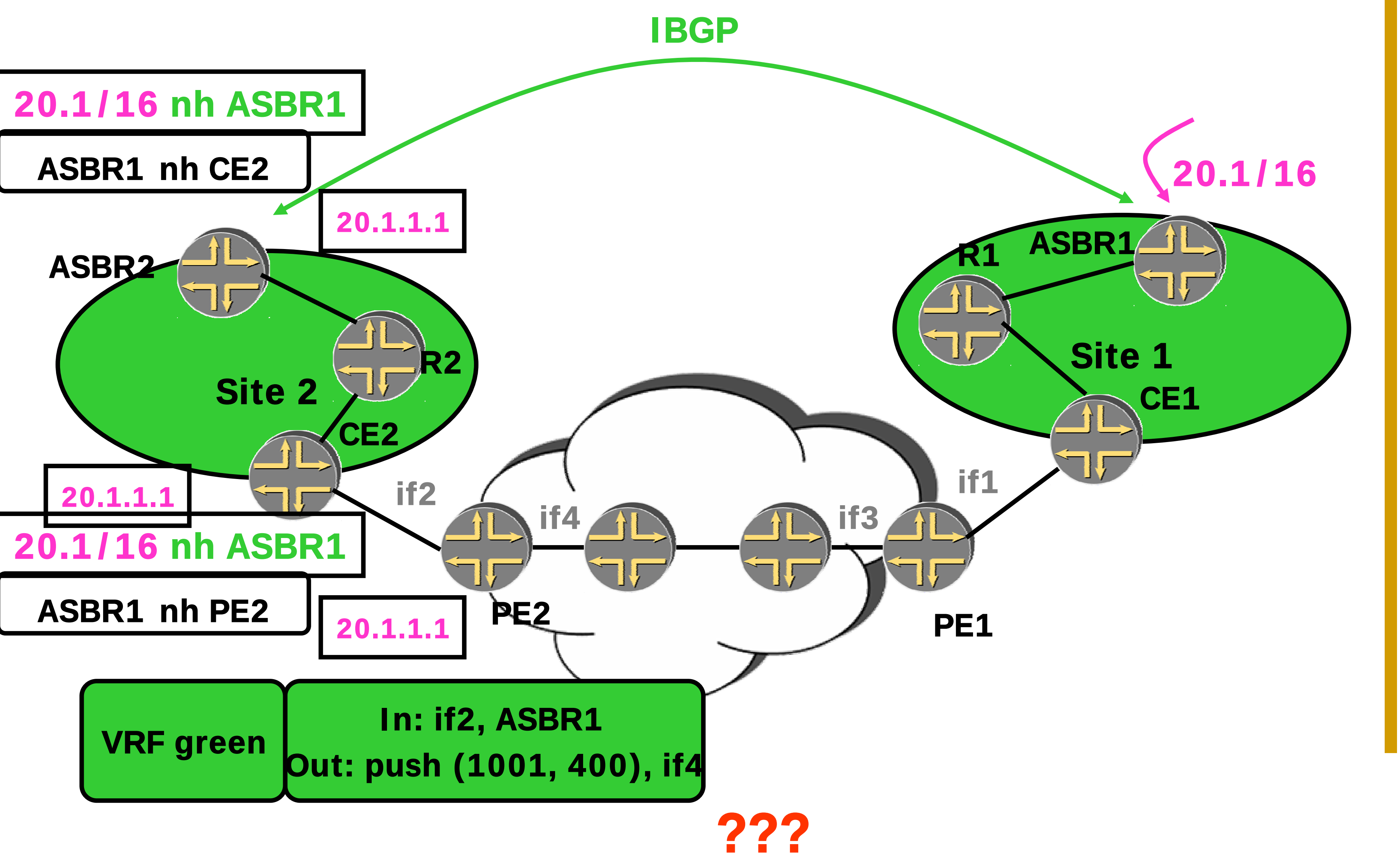
← Routing info

Traffic →

ISP as a VPN customer – exchange of routing information – (no MPLS in sites)



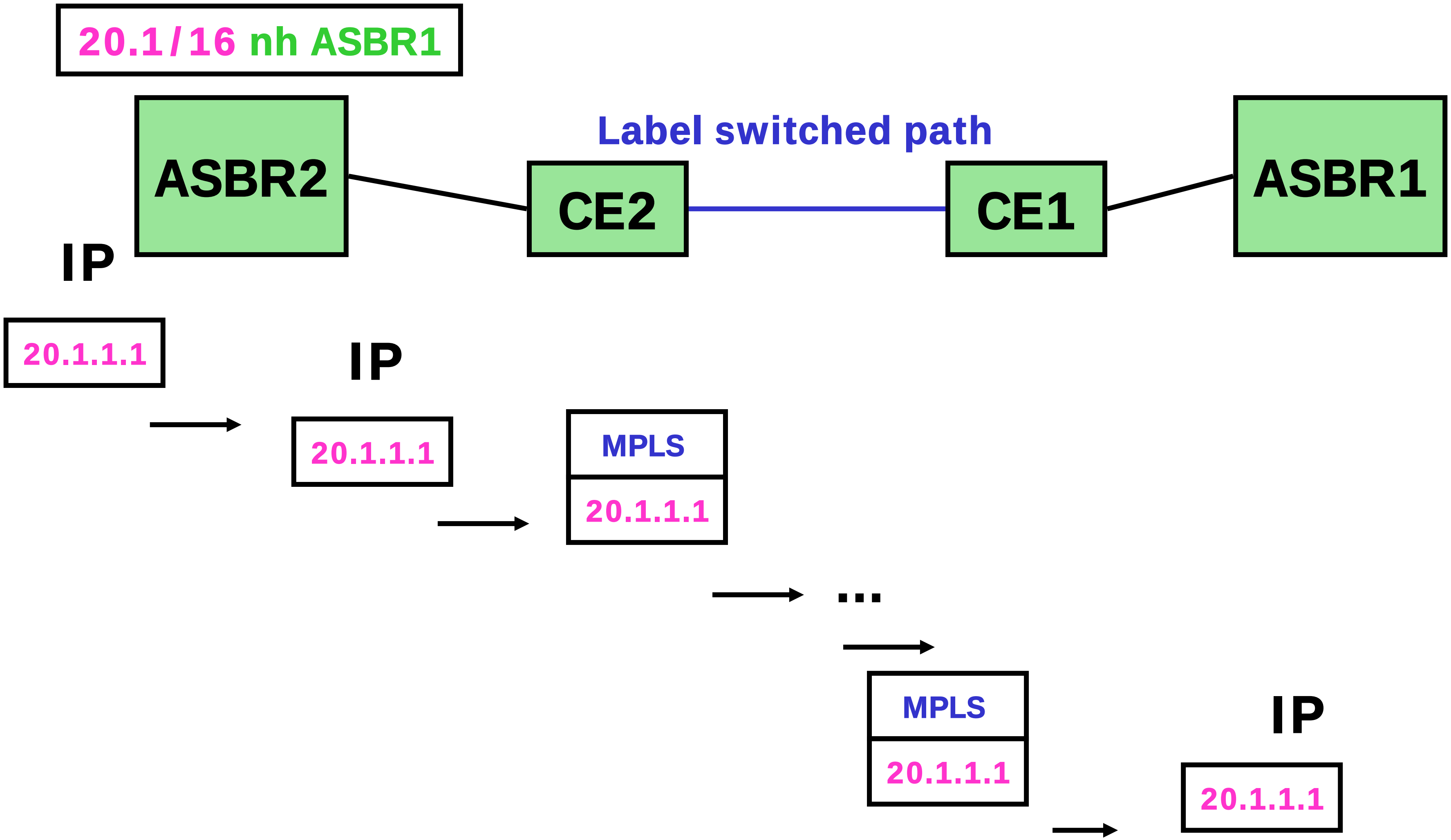
ISP as a VPN customer – forwarding traffic – the solution so far



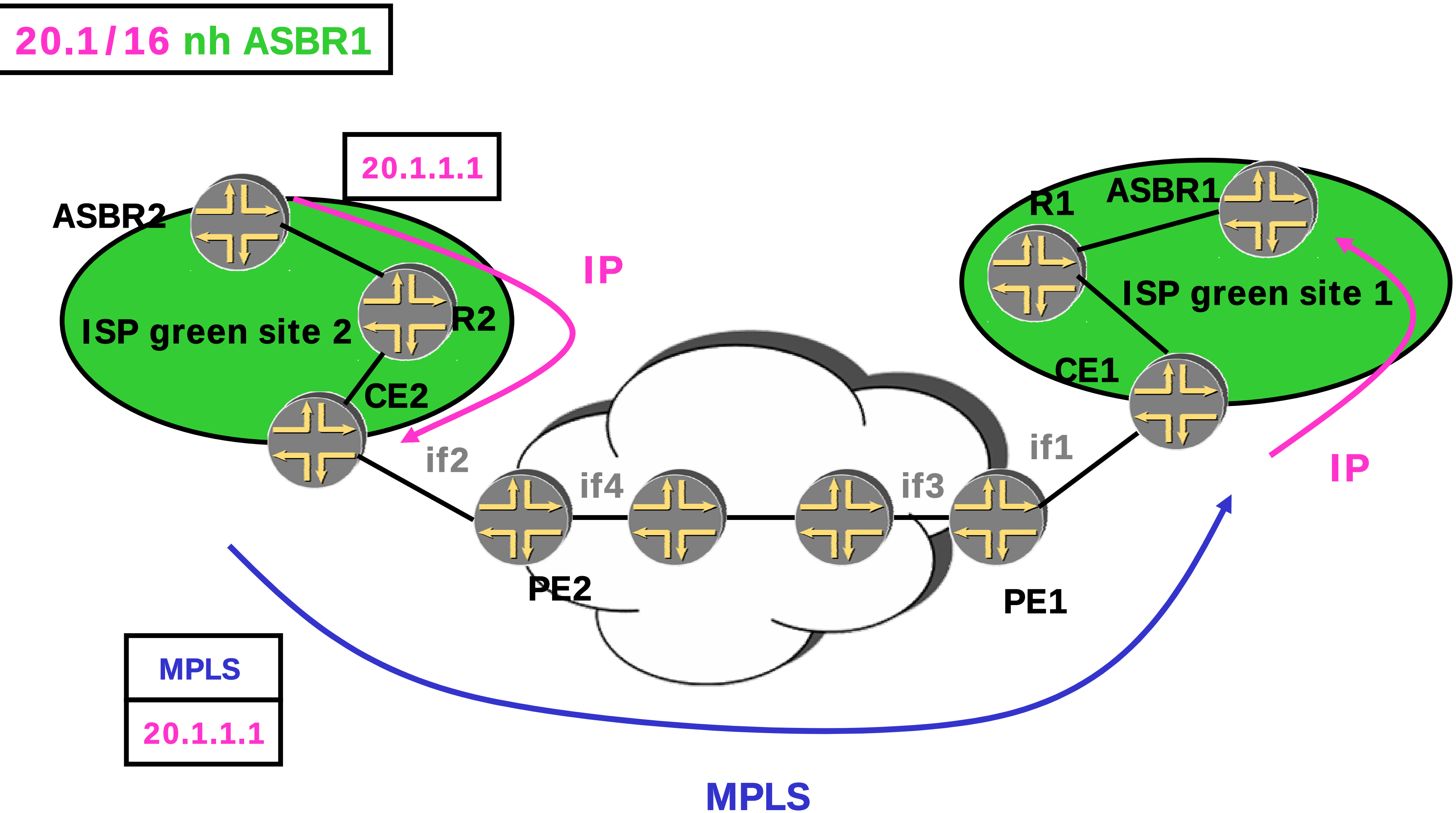
ISP as a VPN customer – step 2

- New problem – When forwarding customer traffic to an internet destination, the PE doesn’t have a route.
- The PE only has routes for the customer’s internal routes.
- The solution – Use MPLS to forward traffic across nodes that don’t have a route to the destination. Need to extend MPLS to the CE.

The abstraction – forwarding



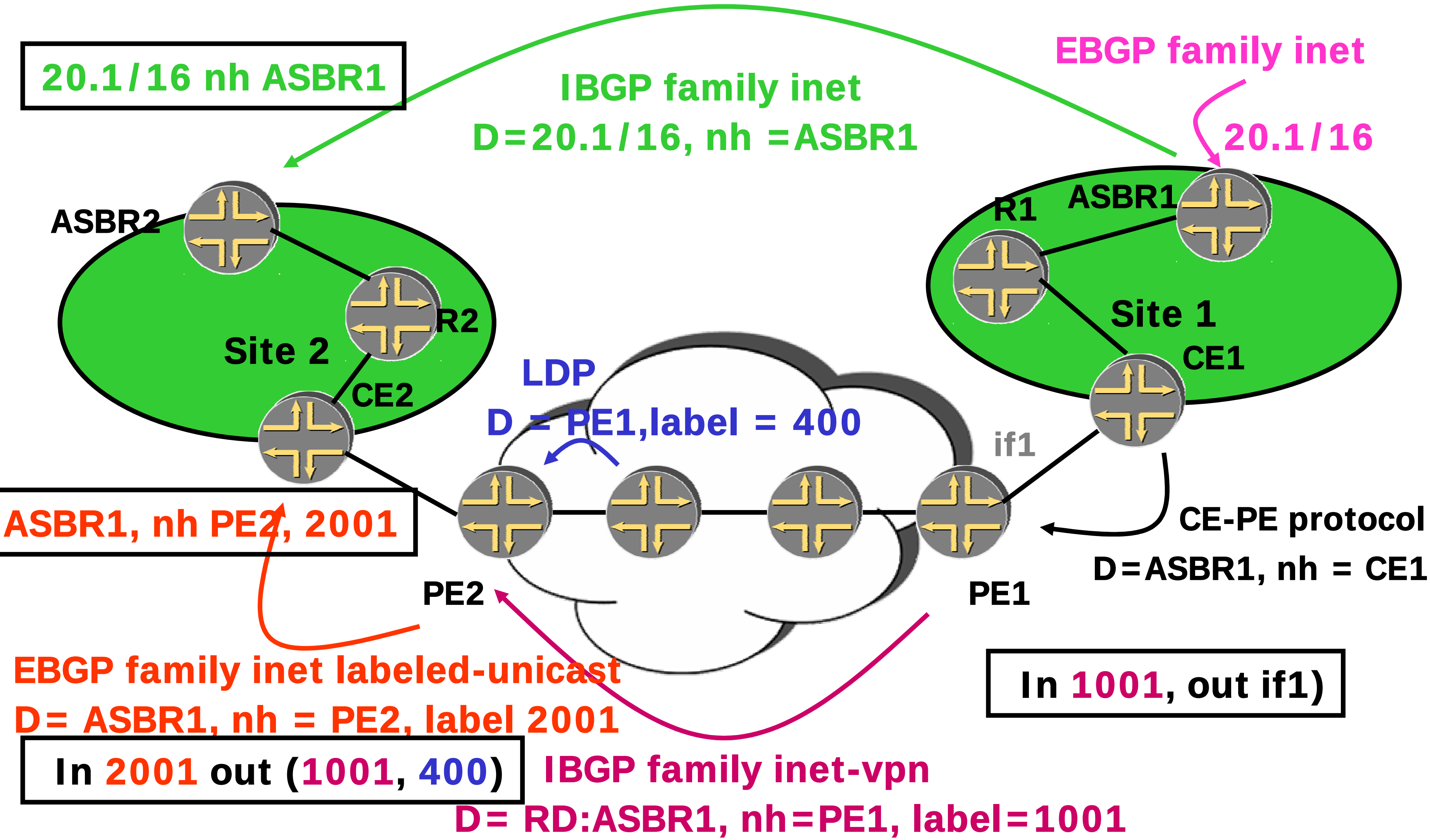
ISP as a VPN customer – no MPLS within sites - conceptual model



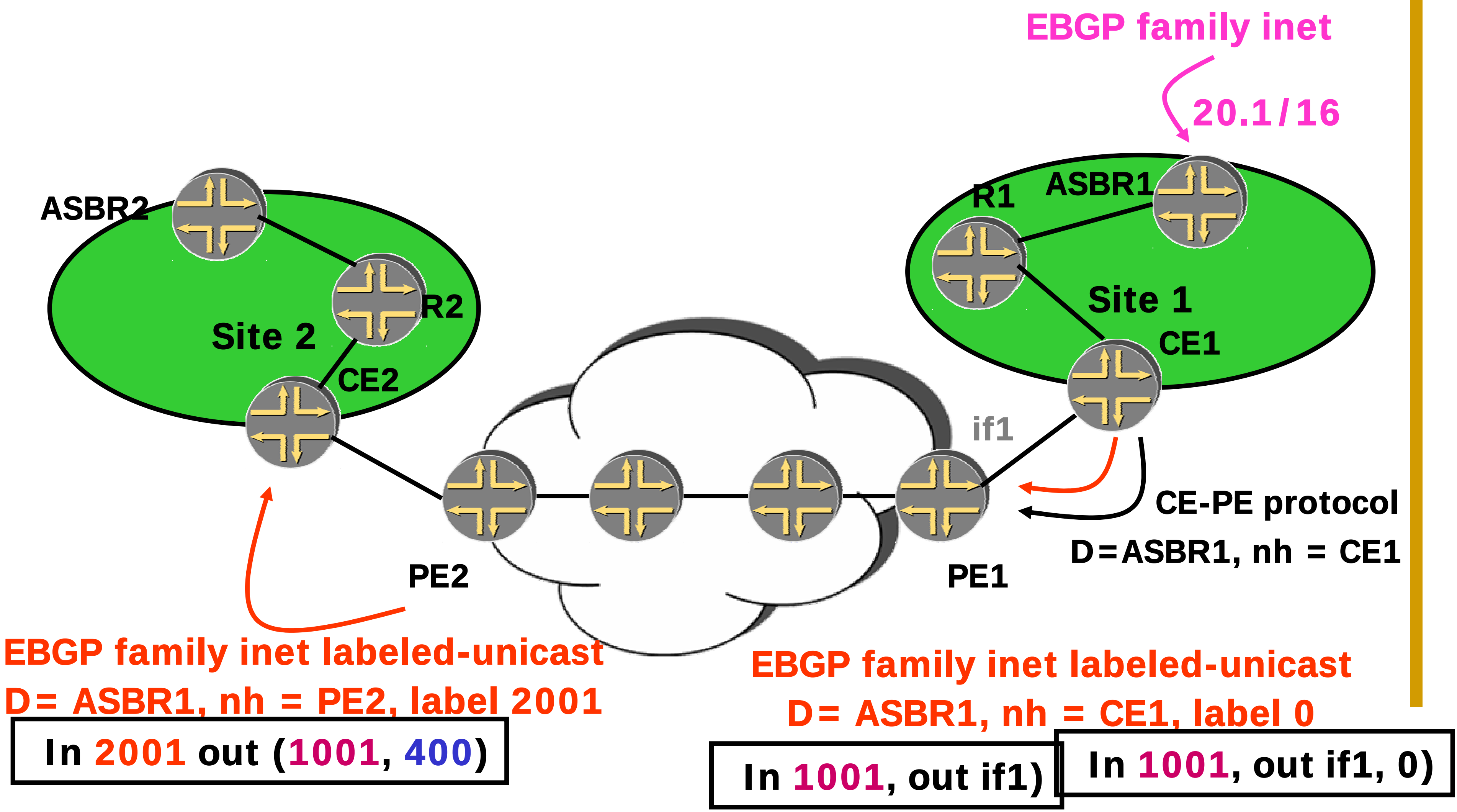
ISP as a VPN customer – step 2

- The CE forwards the traffic over MPLS to the remote CE that will have an IP route for the external route.
- The local CE needs a label-switched-path to the remote CE.
- When the PE advertises the VPN routes to the CE, it also advertises a label for them. This extends MPLS to the CE. We are using this label to pick the next-hop on the PE.

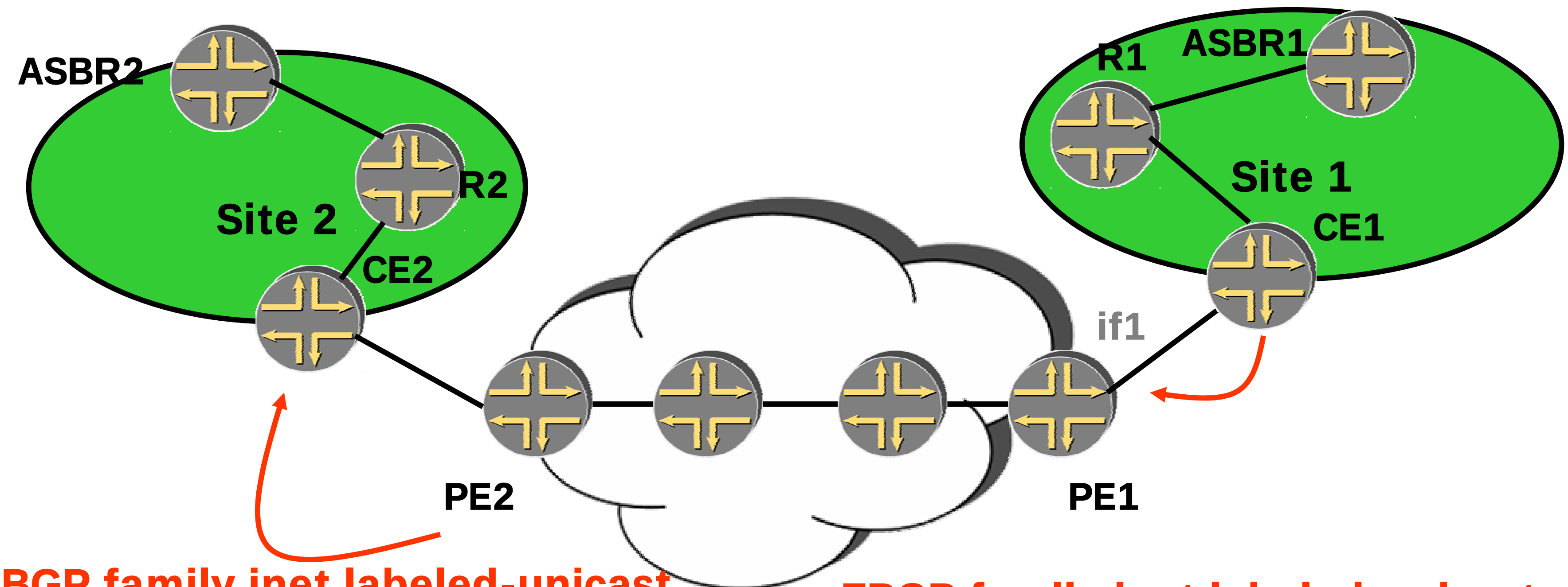
ISP as a VPN customer – exchange of routing information - revisited



ISP as a VPN customer – exchange of routing information - revisited



ISP as a VPN customer – exchange of routing information - revisited



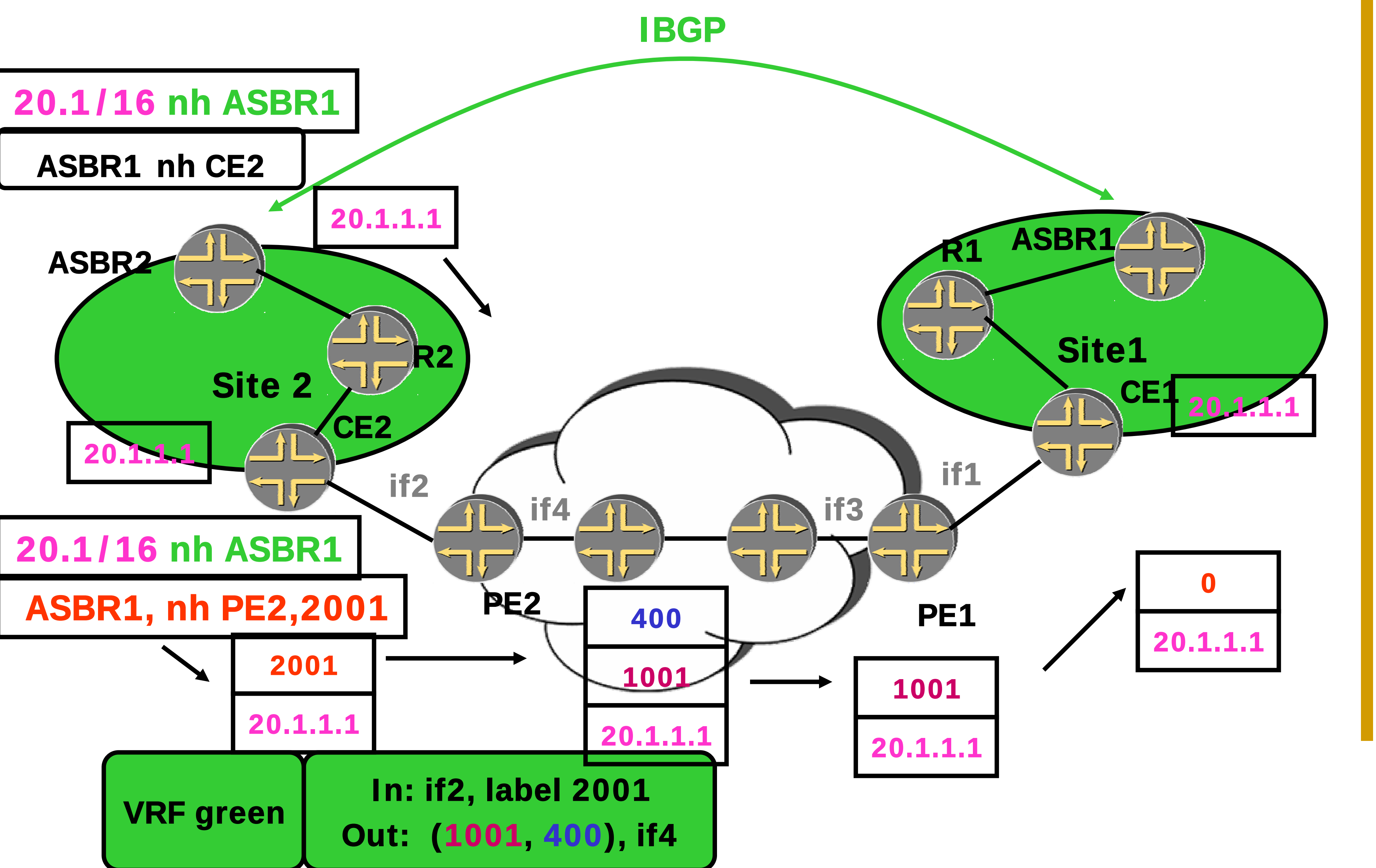
EBGP family inet labeled-unicast
D= ASBR1, nh = PE2, label 2001

In 2001 out (1001, 400)

EBGP family inet labeled-unicast
D= ASBR1, nh = CE1, label 0

In 1001, out if1, 0)

ISP as a VPN customer – forwarding traffic



New concepts

- The label is meaningful for the box that assigned it (it identifies the next-hop to be used for forwarding).

- When assigning a new label, must install MPLS forwarding state. This stitches the two LSPs together.

ISP as a VPN customer – no MPLS within sites – summary

- The VPN provider doesn’t carry the customer’s external routes in its backbone, it only carries the customer internal routes (BGP next-hops).
- A labeled-switched path is established between the remote CEs.
- The IP traffic to external destinations travels over this label-switched-path to the remote CE.

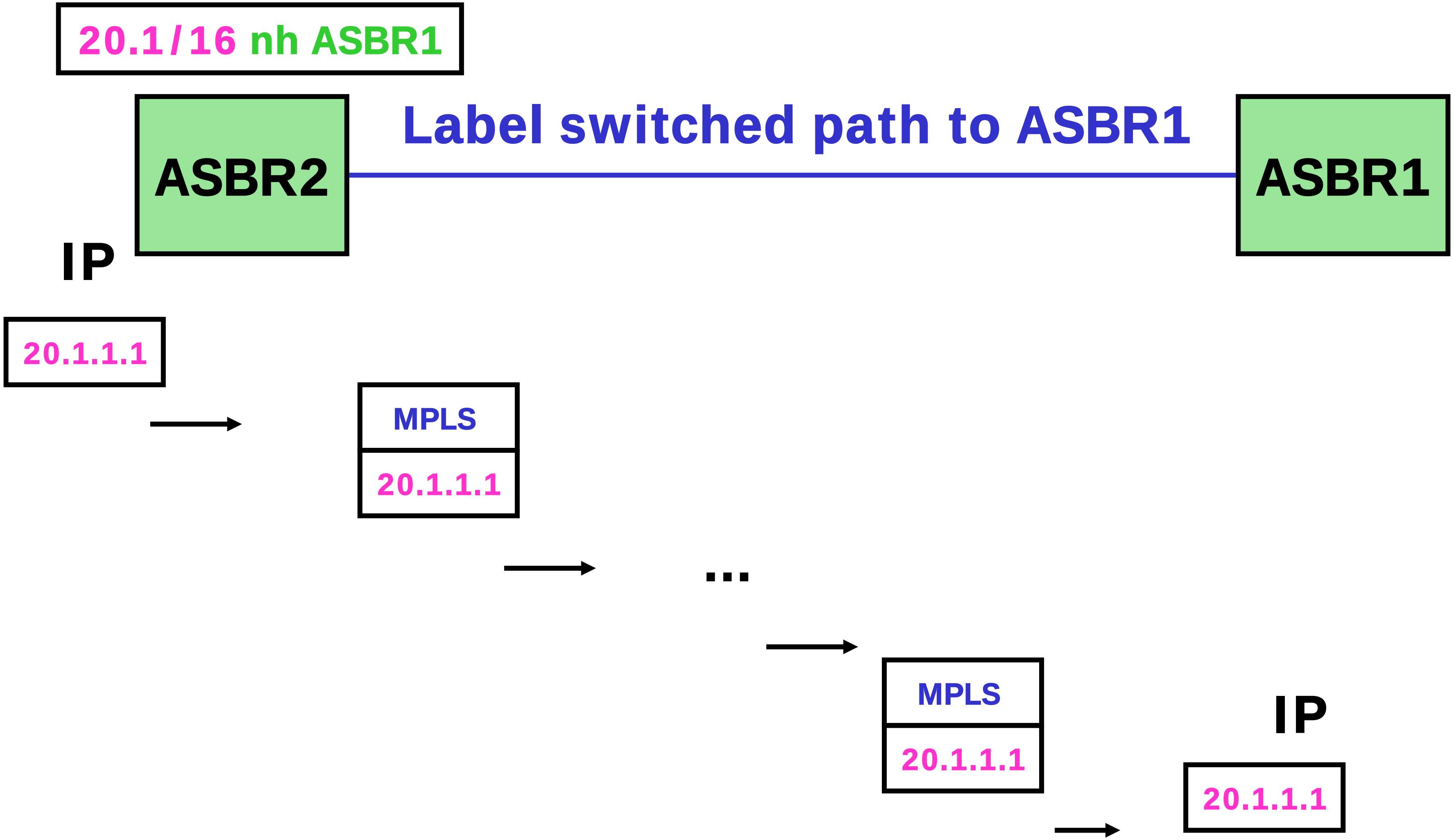
ISP as a VPN customer – scenario 2

- With MPLS in the customer sites.
- Goal – the provider doesn't want to carry the customer's external routes.

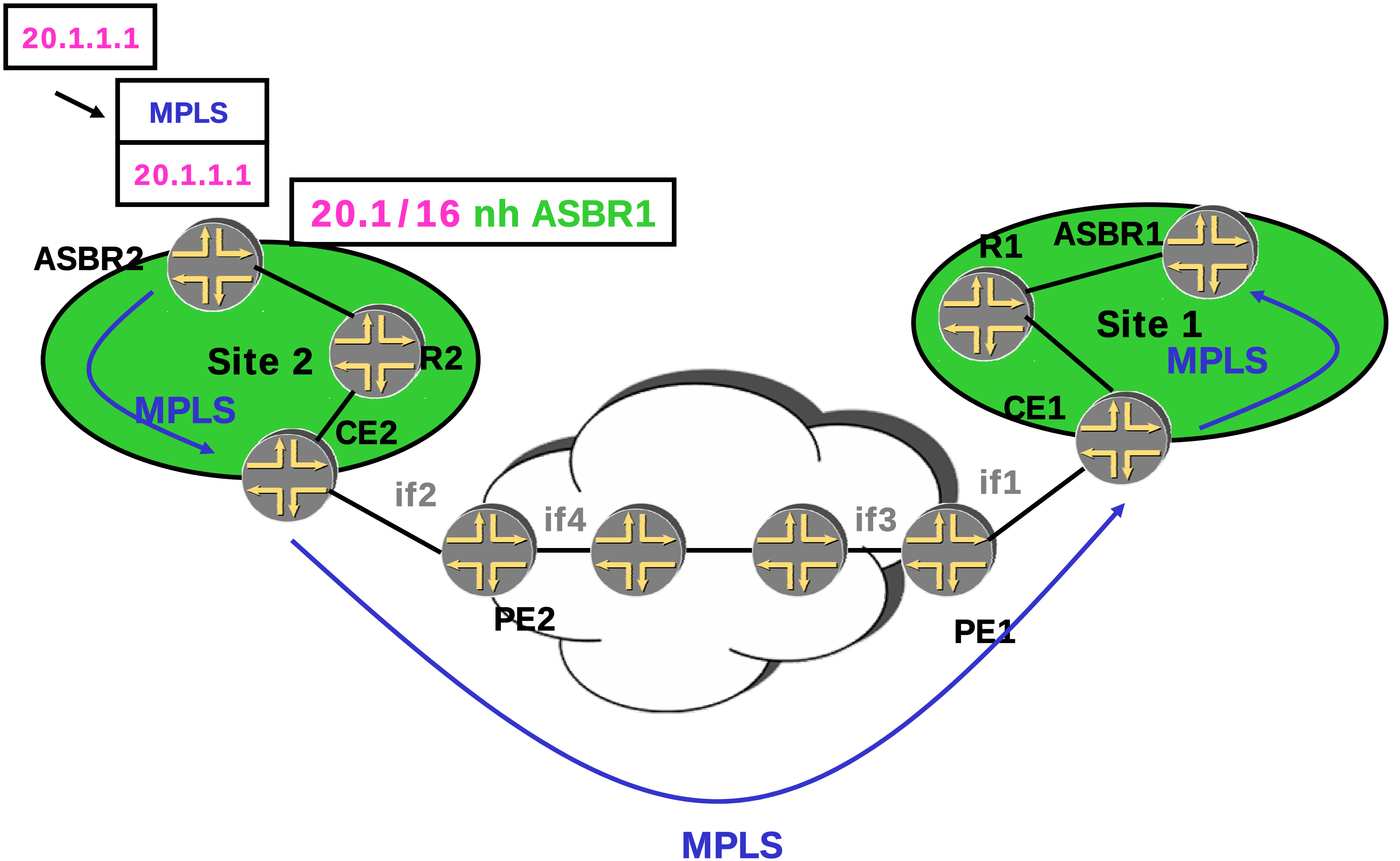
ISP as a VPN customer – MPLS in the customer sites – The idea

- Can use MPLS to forward traffic in the customer's sites.
- No need for all the routers to carry the external routes. Rely on MPLS to forward traffic to destinations for which the transit routers don't have routing entries.
- Need a label-switched path between the routers that carry the external prefixes.

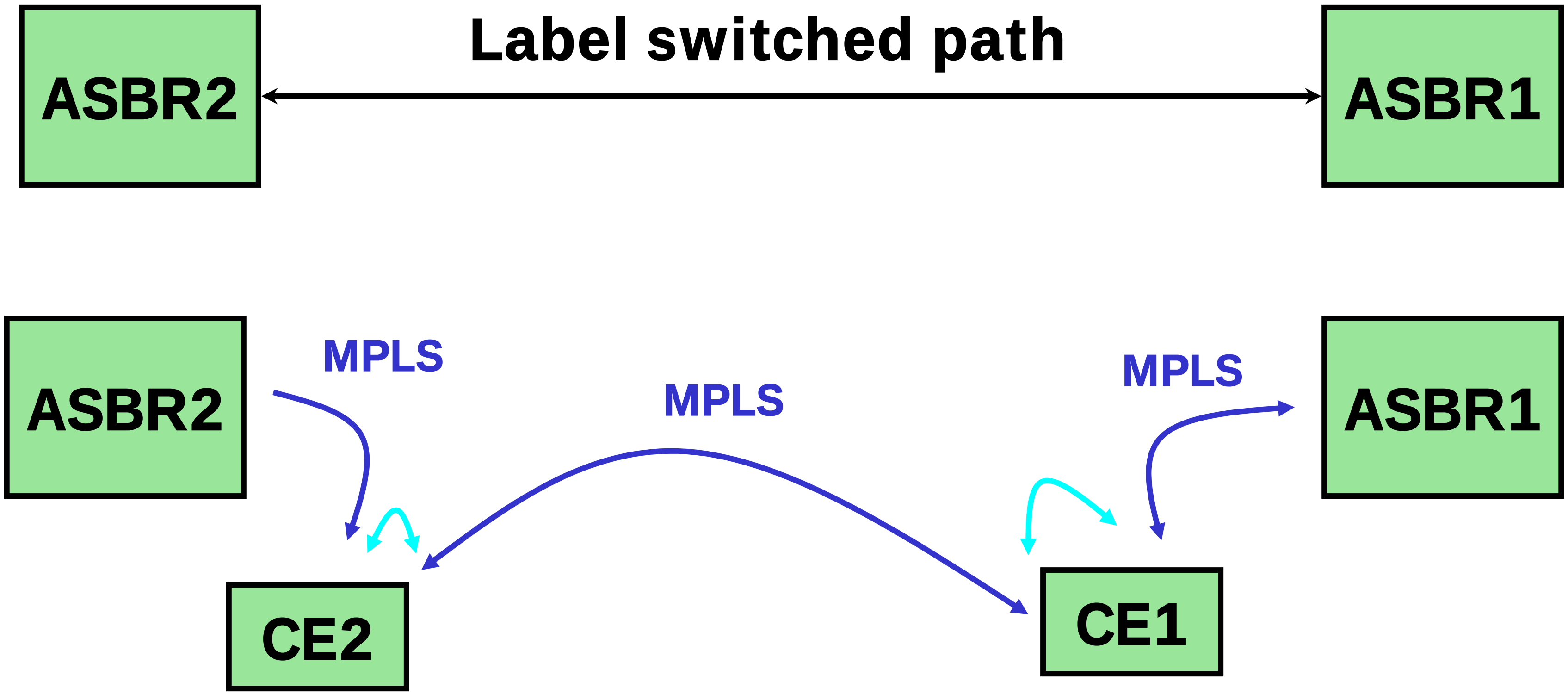
The abstraction – forwarding



ISP as a VPN customer – MPLS in the customer sites – conceptual model



The abstraction – the label-switched-path



ISP as a VPN customer – with MPLS in the customer sites

- The label-switched path between the ASBRs is made up of several segments.
- In the previous scenario we saw how to establish a CE-CE label-switched path.
- Need to stitch the CE-CE path with the CE-ASBR paths.

Routing exchanges / traffic forwarding

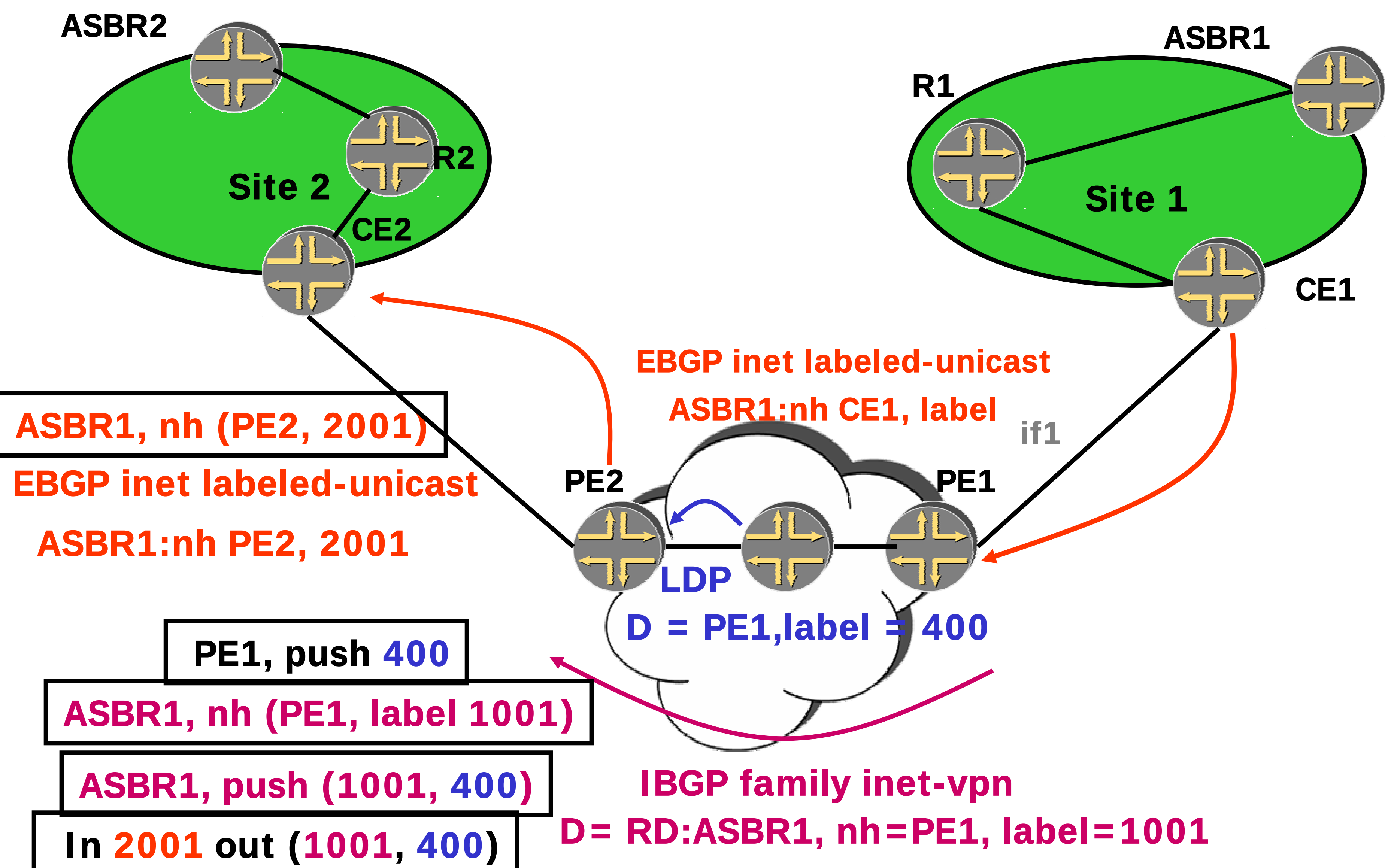
Site 2

Site 1

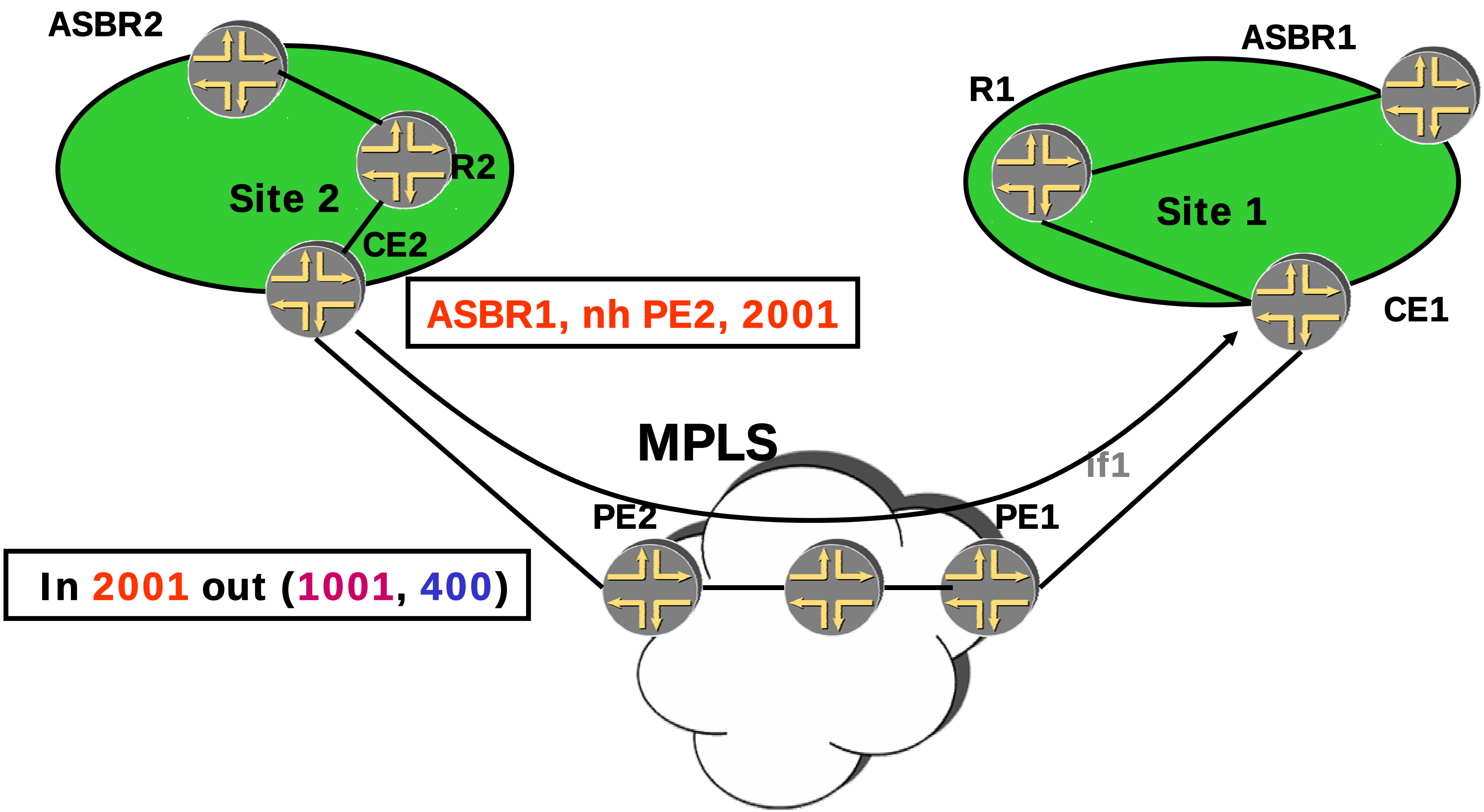
← Routing info

Traffic →

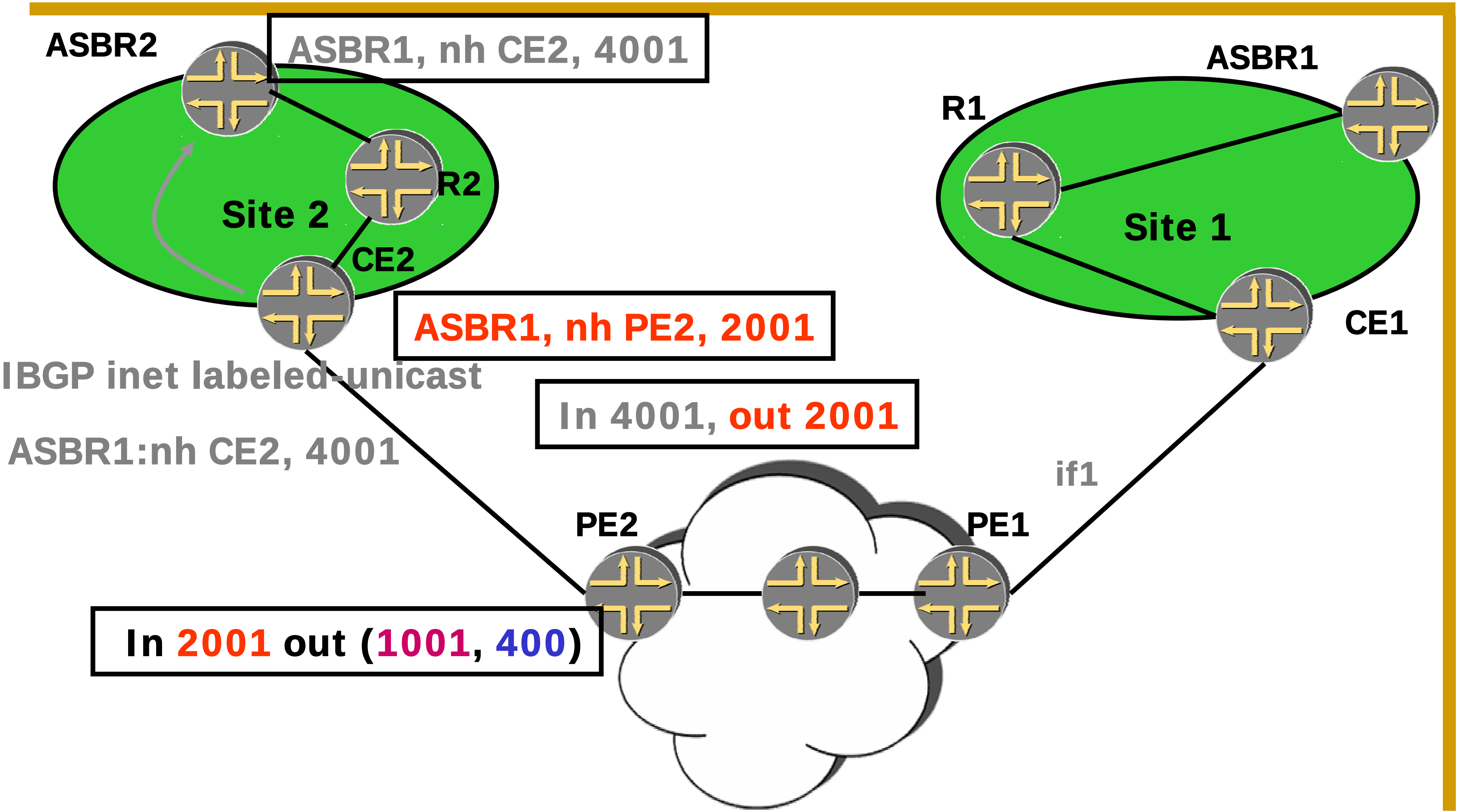
Establishing a path between the ASBRs (1) CE2-to-CE1



Establishing a path between the ASBRs (2) CE2-to-CE1

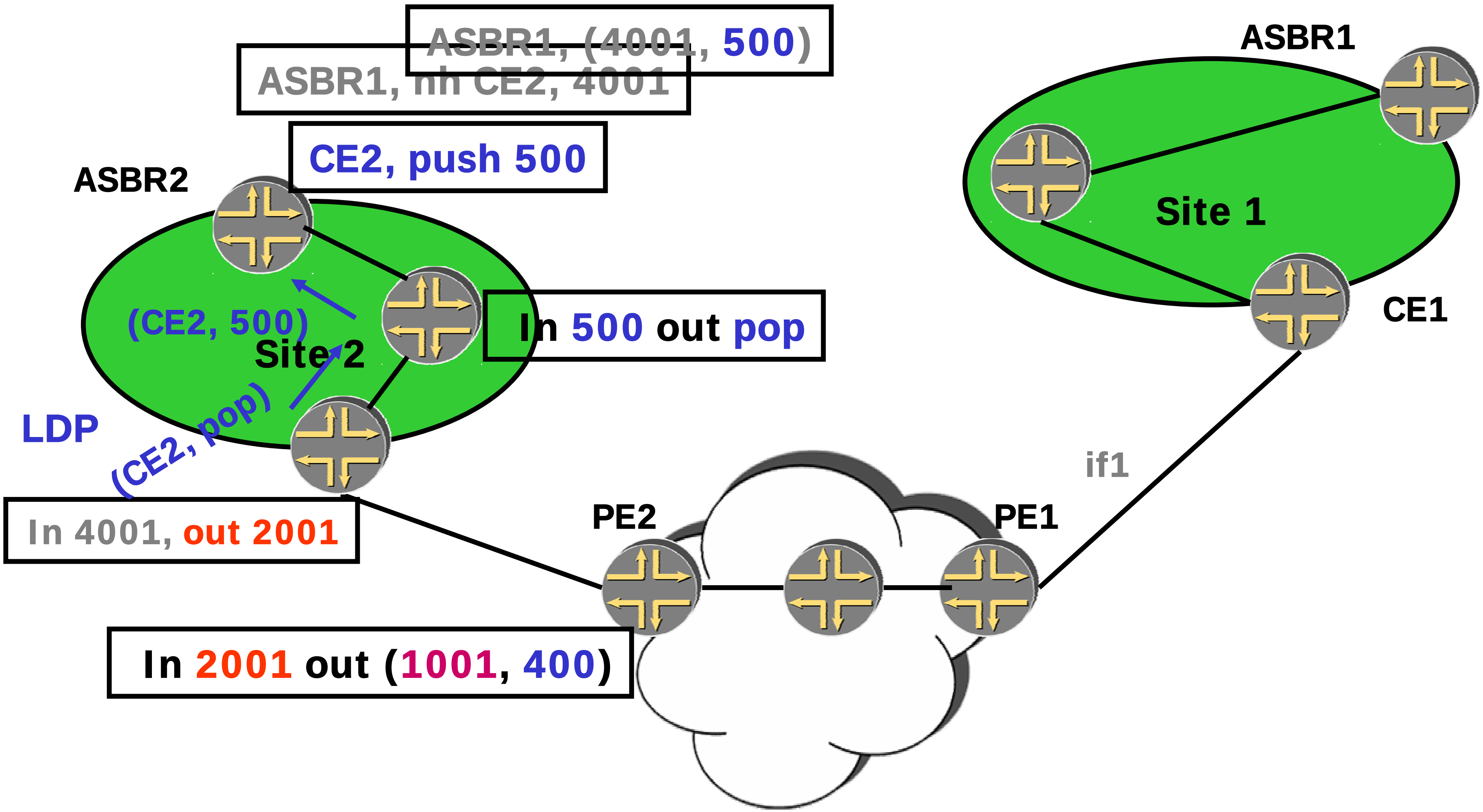


Establishing a path between the ASBRs (3) ASBR2-to-CE2



The problem – need an MPLS path to CE2.

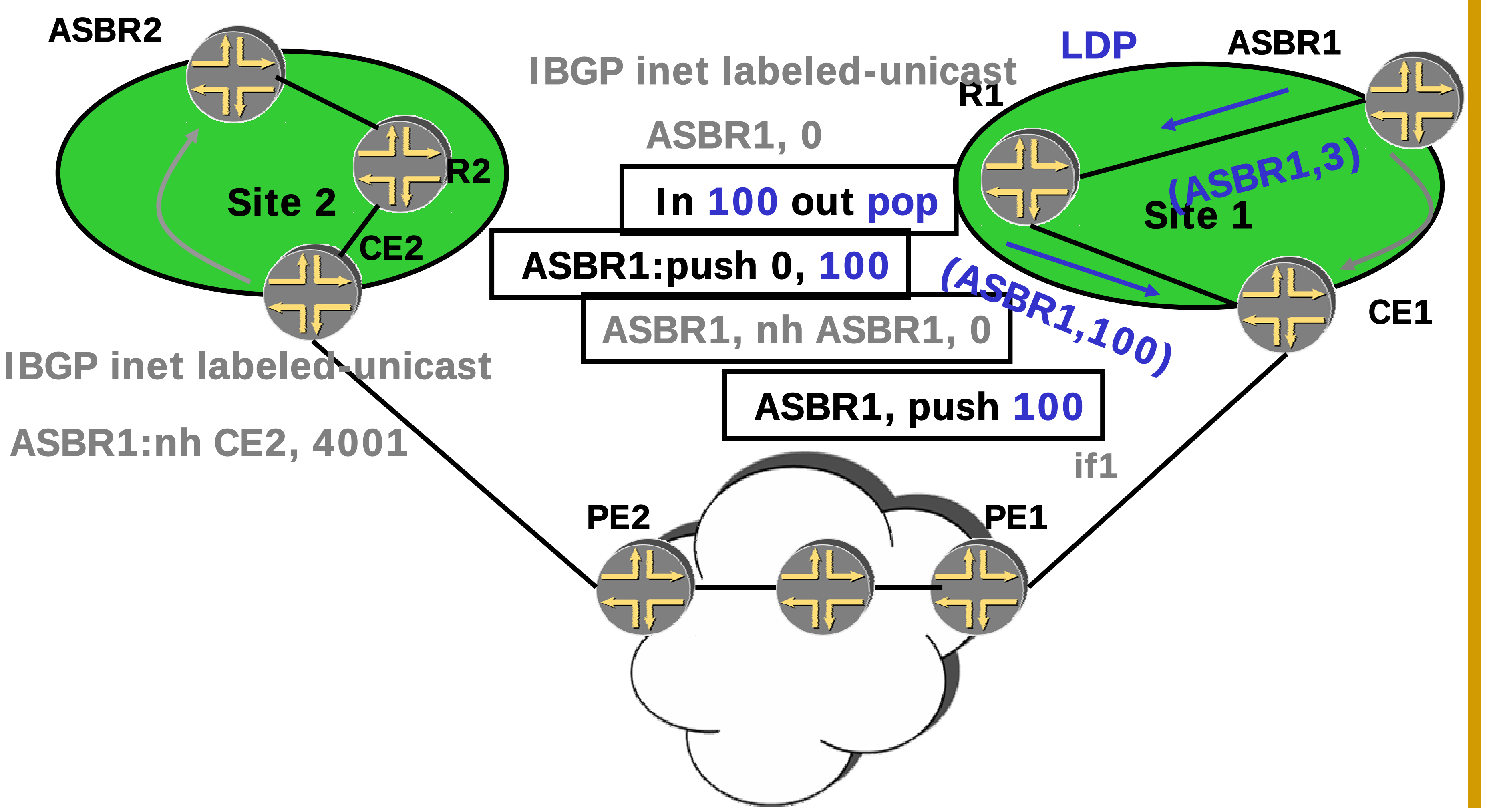
Establishing a path between the ASBRs (4) ASBR2-to-CE2



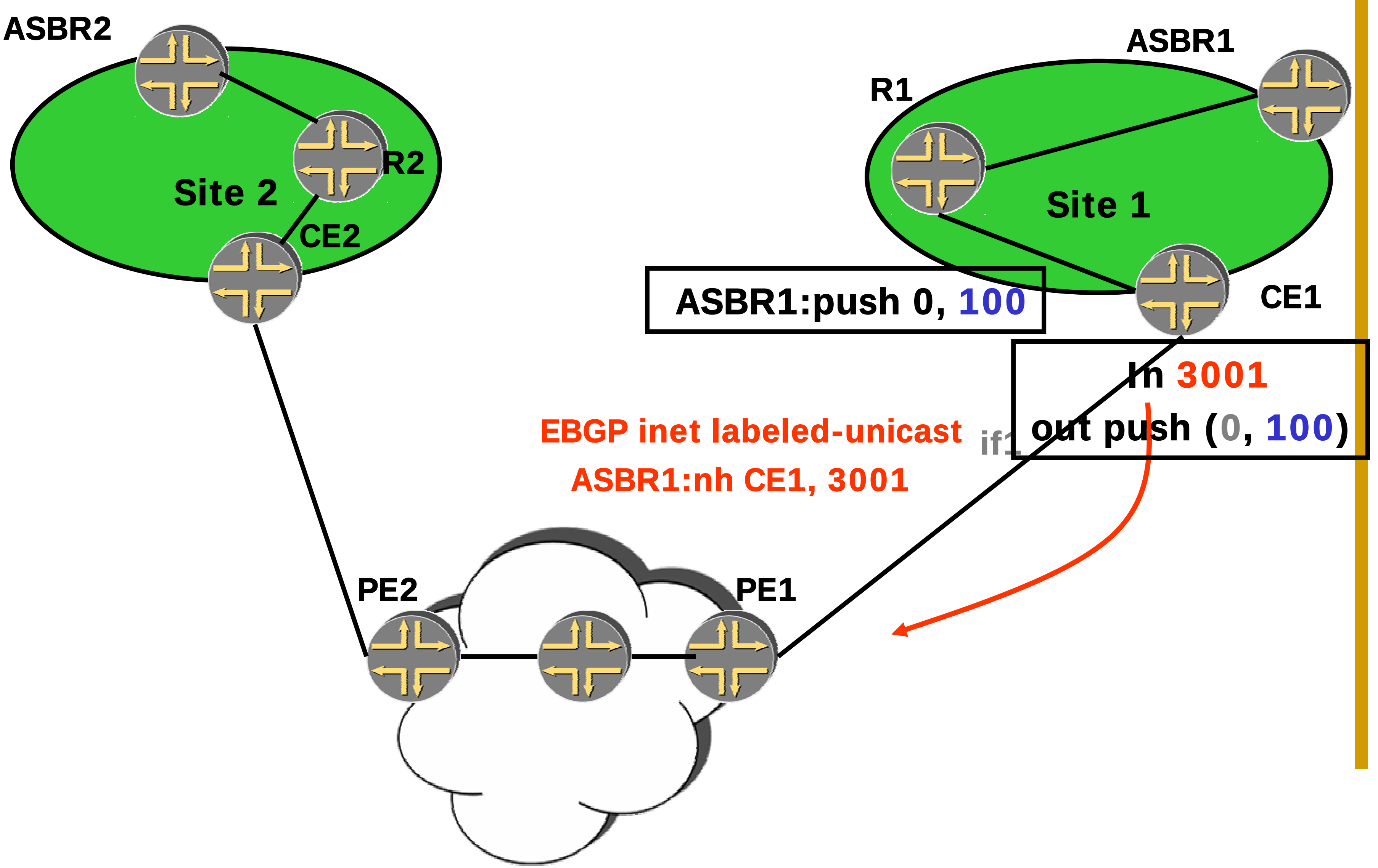
ASBR2-to-CE2 – discussion

- **Requires a two label push at ASBR2.**
- **One label identifies ASBR1, and the other label identifies the path to CE2 which is the router that injected ASBR1.**
- **The route for ASBR1 doesn’t need to be known inside Site2.**

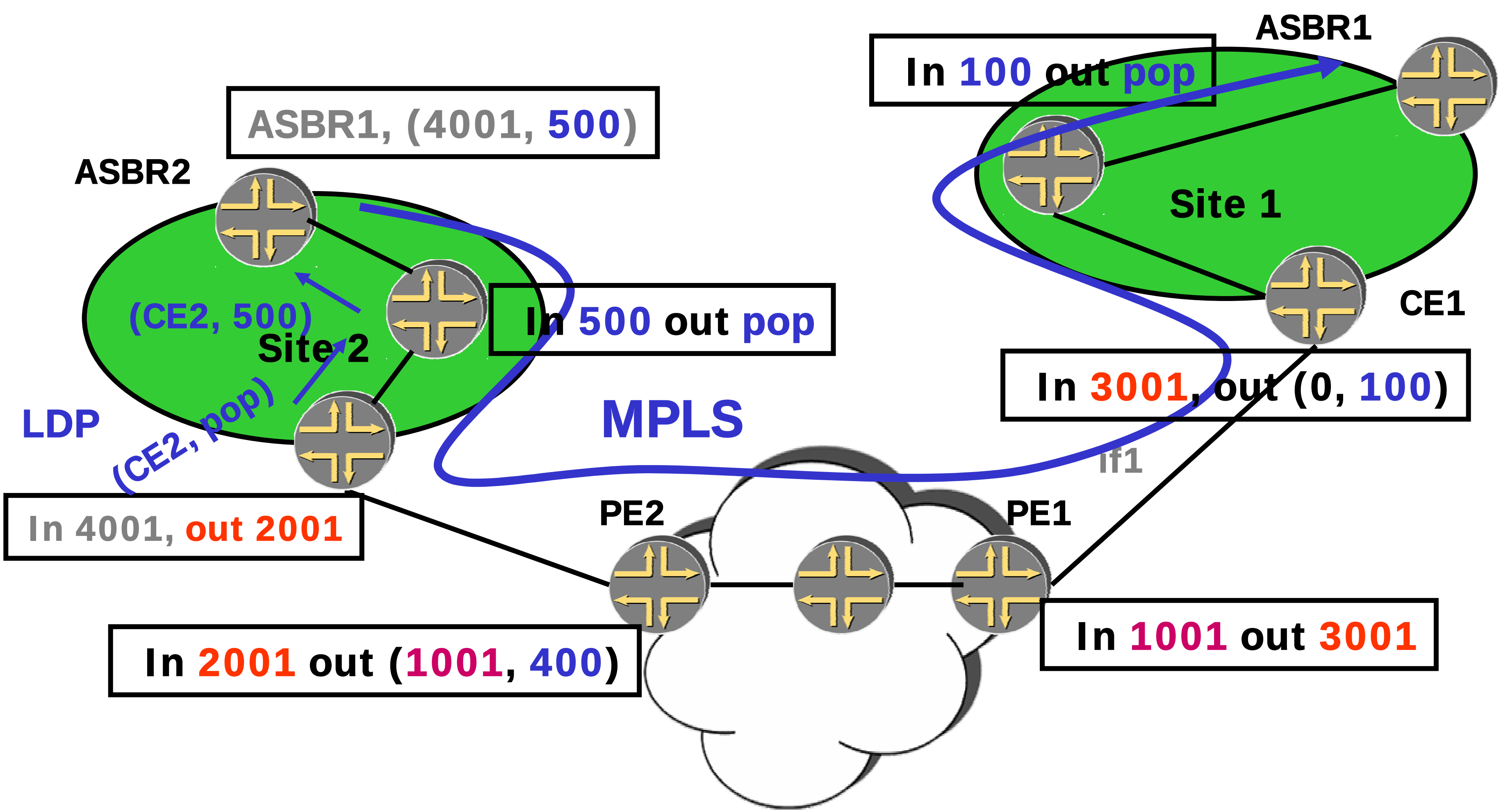
Establishing a path between the ASBRs (5) CE1-to-ASBR1



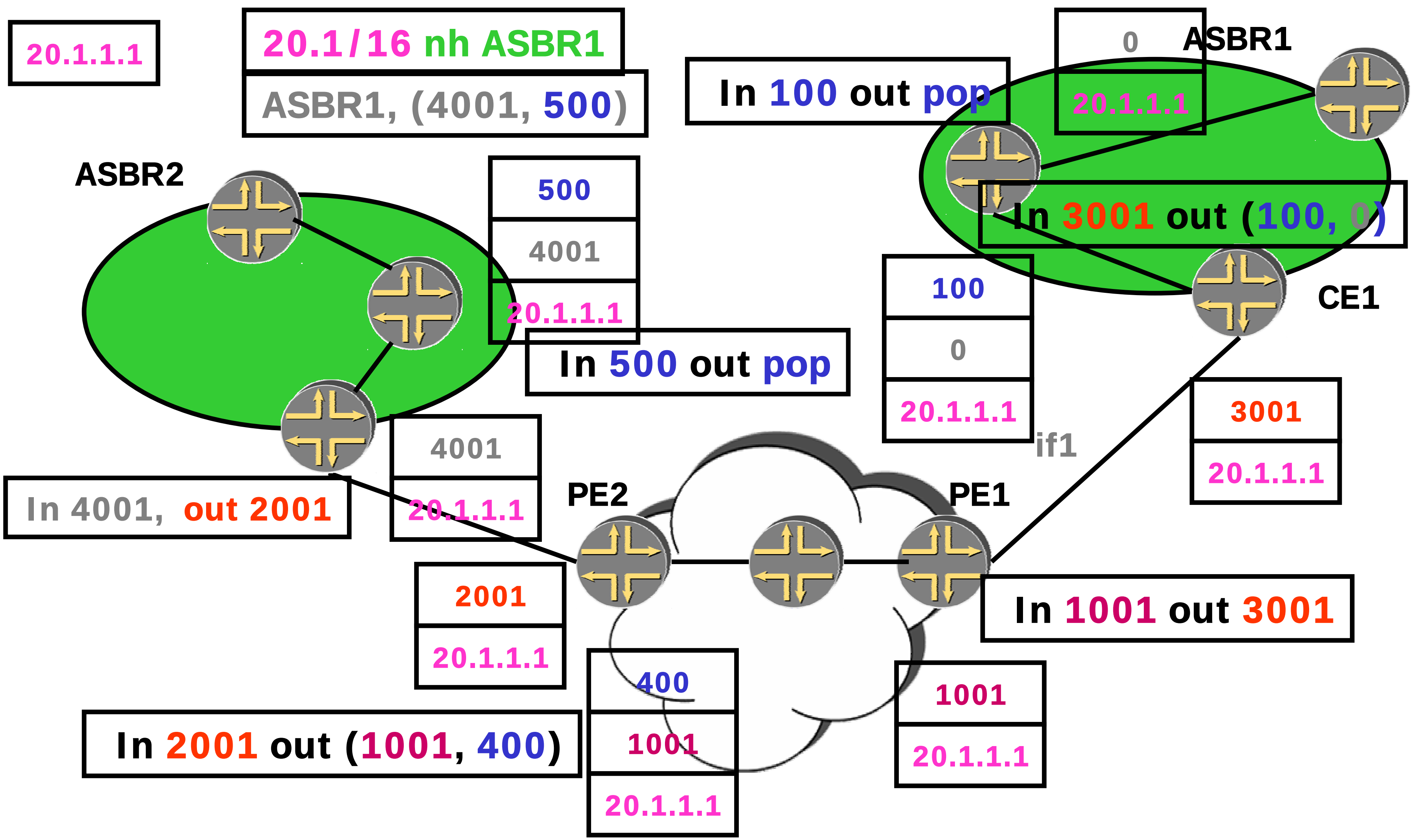
Establishing a path between the ASBRs (6) PE1-to-ASBR1



Establishing a path between the ASBRs (7) ASBR2-to-CE2-to-CE1-to-ASBR1



Forwarding traffic along the ASBR2-ASBR1 path



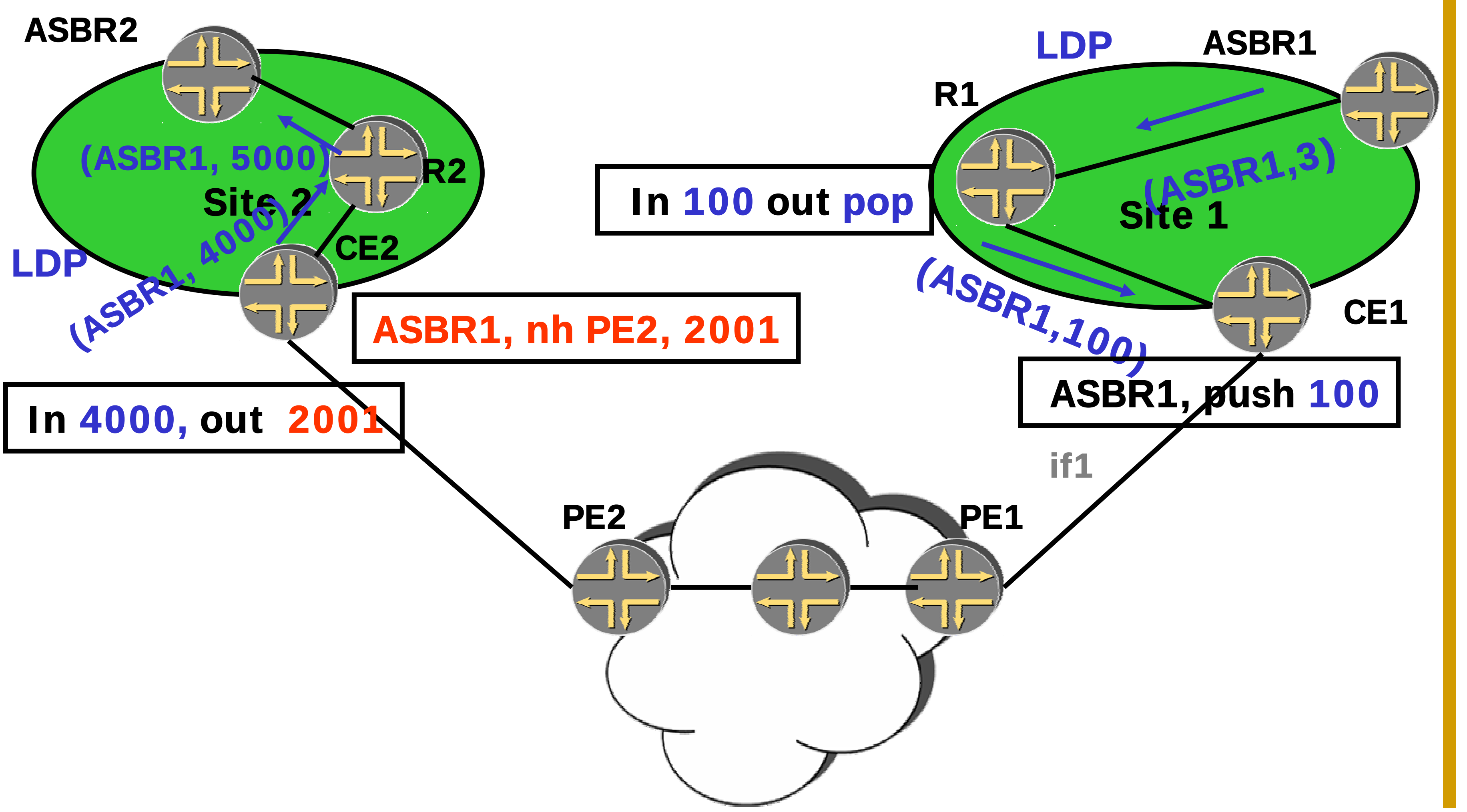
ISP as a VPN customer – with MPLS in the customer sites

- Can use MPLS to forward traffic in the customers sites.
- Can isolate the knowledge of the external routes to the ASBRs.

ISP as a VPN customer – scenario 2 – revisited – using LDP

- **Can use either LDP or labeled-BGP inside the sites (so far the example only showed labeled-BGP).**

Establishing a path between the ASBRs – LDP instead of labeled BGP



Make LDP advertise a FEC for ASBR1.

LDP instead of labeled-BGP

- **Can use LDP to advertise the route to ASBR1. (requires support of this behavior in LDP).**
- **Requires redistribution of the route to ASBR1 into the IGPs. (redistribution from BGP to IGP).**
- **All routers in site2 will carry an IGP route for ASBR1.**

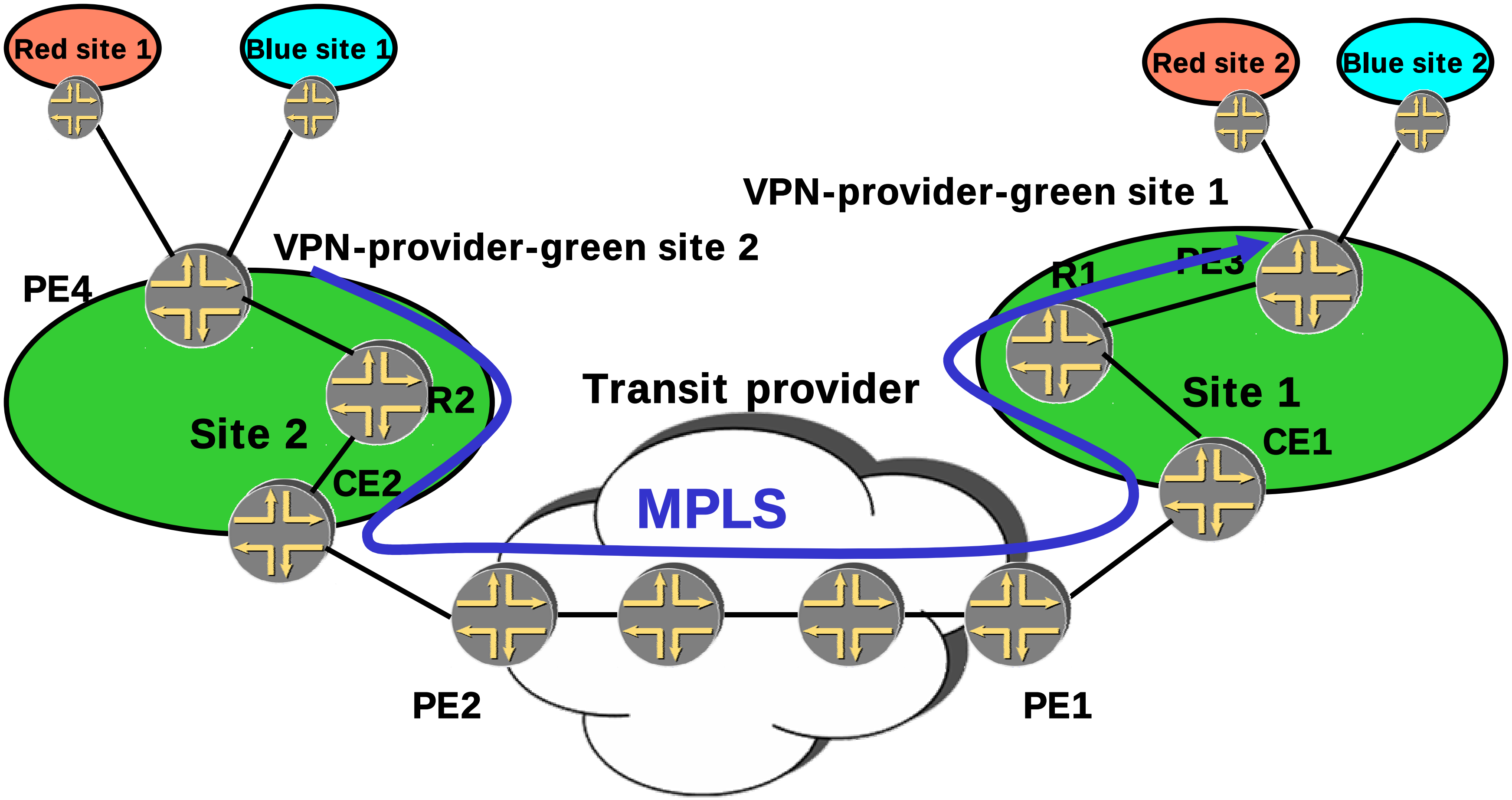
Part 2 – Hierarchical and recursive applications

- Introduction
- ISP as a VPN customer
- VPN service provider as a VPN customer
- VPN services across AS boundaries

VPN service provider as a VPN customer

- Goal: provide connectivity for geographically dispersed sites of a VPN service provider.

VPN service provider as a VPN customer

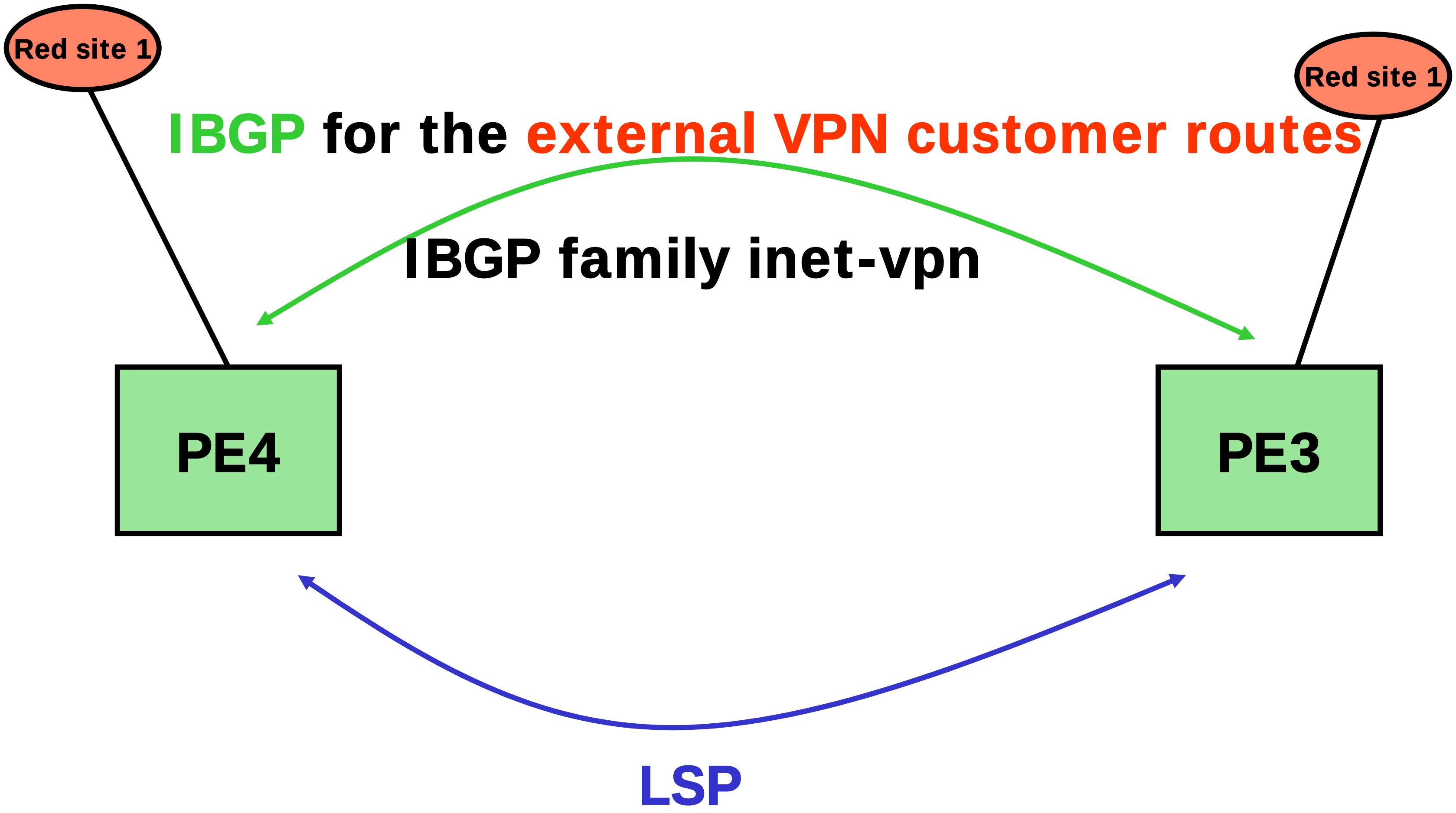


VPN service provider – provides transit for VPN-provider-green

VPN service provider as a VPN customer

- **The problem:**
  - **Want to avoid having to carry the VPN routes (red, blue routes) of the VPN-customer (VPN-provider-green) in the VPN-provider network (transit provider).**
- **The solution:**
  - **Let the VPN customer (VPN-provider-green) be responsible for its VPN routes (which are in effect external routes).**

The abstraction - routing



VPN service provider as a VPN customer

- The same as the ISP as a customer scenario with MPLS in the customer sites.
- All customer sites are in the same AS.

- Differences:
  - The routes exchanged between the customer routers are VPN-IP routes instead of IP routes.
  - Three label push (when labeled-BGP is used)

Routing exchanges / traffic forwarding

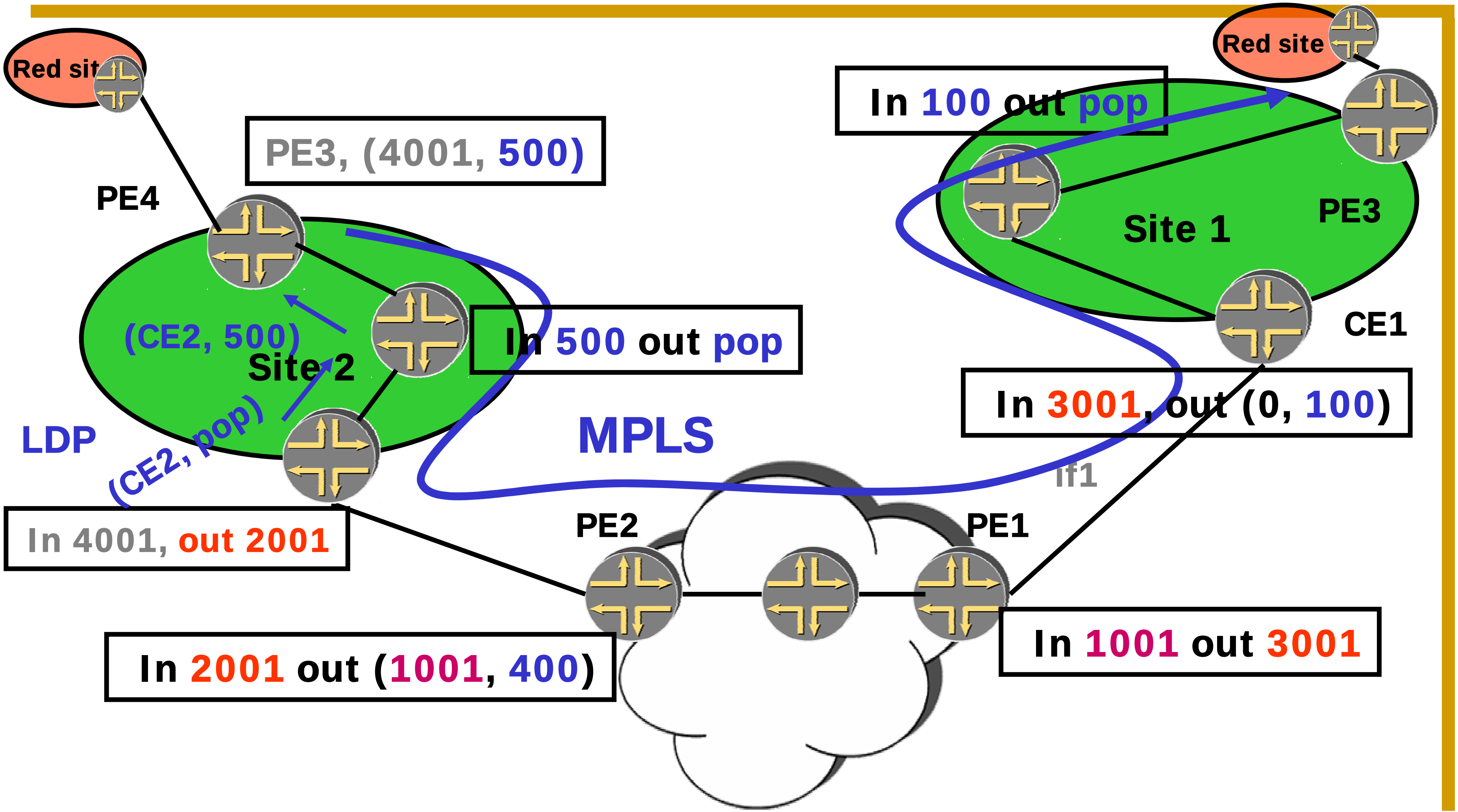
Site 2

Site 1

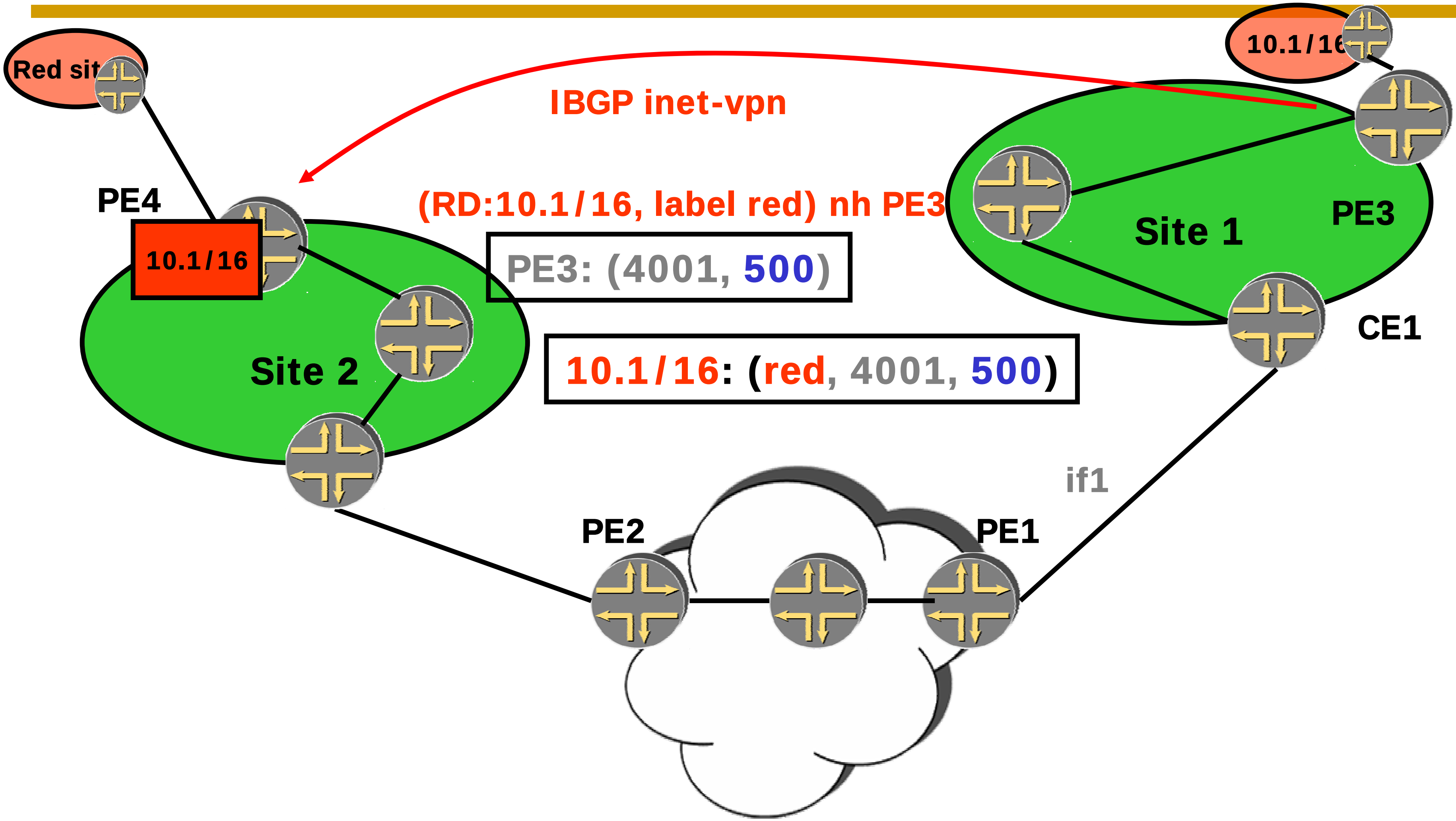
← Routing info

Traffic →

VPN service provider as a VPN customer



VPN service provider as a VPN customer



VPN service provider as a VPN customer

- Since the routes exchanged are VPN-IP routes, forwarding traffic to from one site to another will require a 3 label push:
  - One label identifying the VPN-IP route
  - Two labels to reach the remote PE (when labeled BGP is used within sites).

Part 2 – Hierarchical and recursive applications

- Introduction
- ISP as a VPN customer
- VPN service provider as a VPN customer
- VPN services across AS boundaries

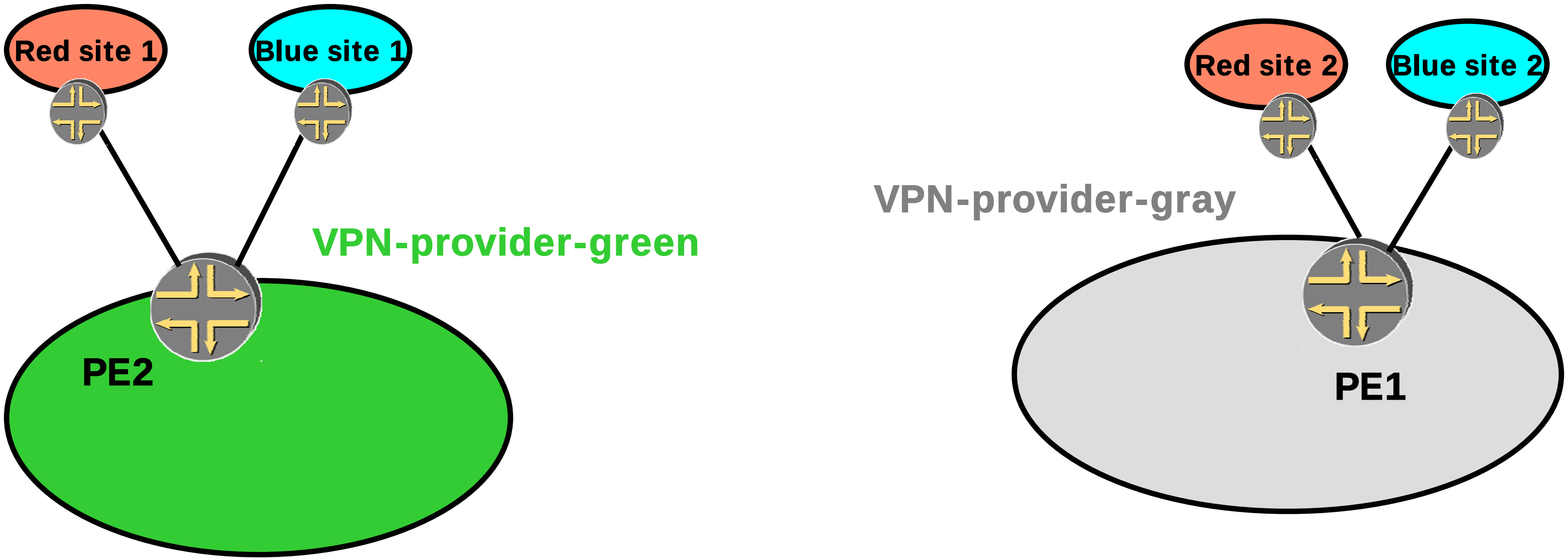
VPN services across AS boundaries

- So far we've seen examples where all sites in a VPN are connected to the same AS.

- What if not all sites are in the same AS?

- Useful if:
  - VPN sites are connected to different providers
  - The provider's backbone is partitioned among different AS.

Sites in different AS



VPN services across AS boundaries

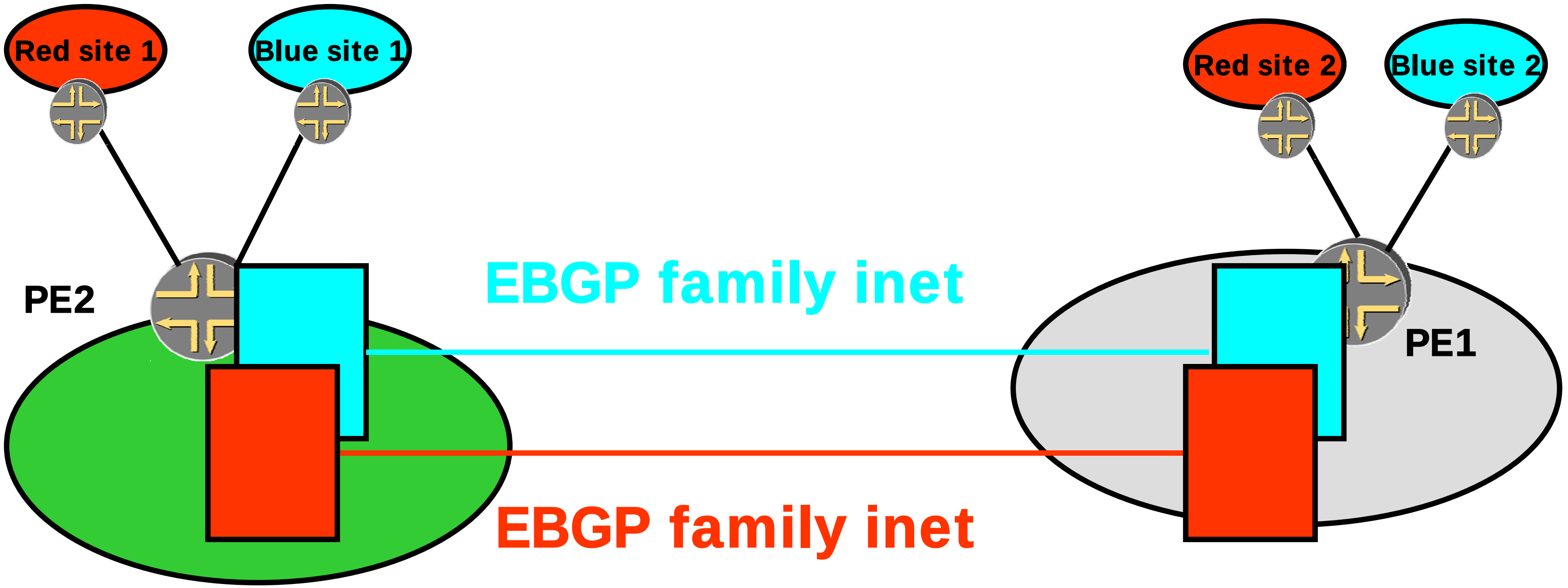
- **The problem – can't run IBGP between the remote sites anymore.**

- **The solutions:**
  - **Discussed in section 10 of 2547bis and referred to as "Option a", "Option b" and "Option c".**

Option A – VRF-to-VRF connections

- A PE router in one AS attaches directly to a PE router in another AS.
- There are several interfaces between the PEs, one for each VPN whose routes are passed between AS.
- Each PE treats the other as a CE and exchanges the VPN routes using EBGP on a per-VRF basis.

Option A – VRF-to-VRF connections



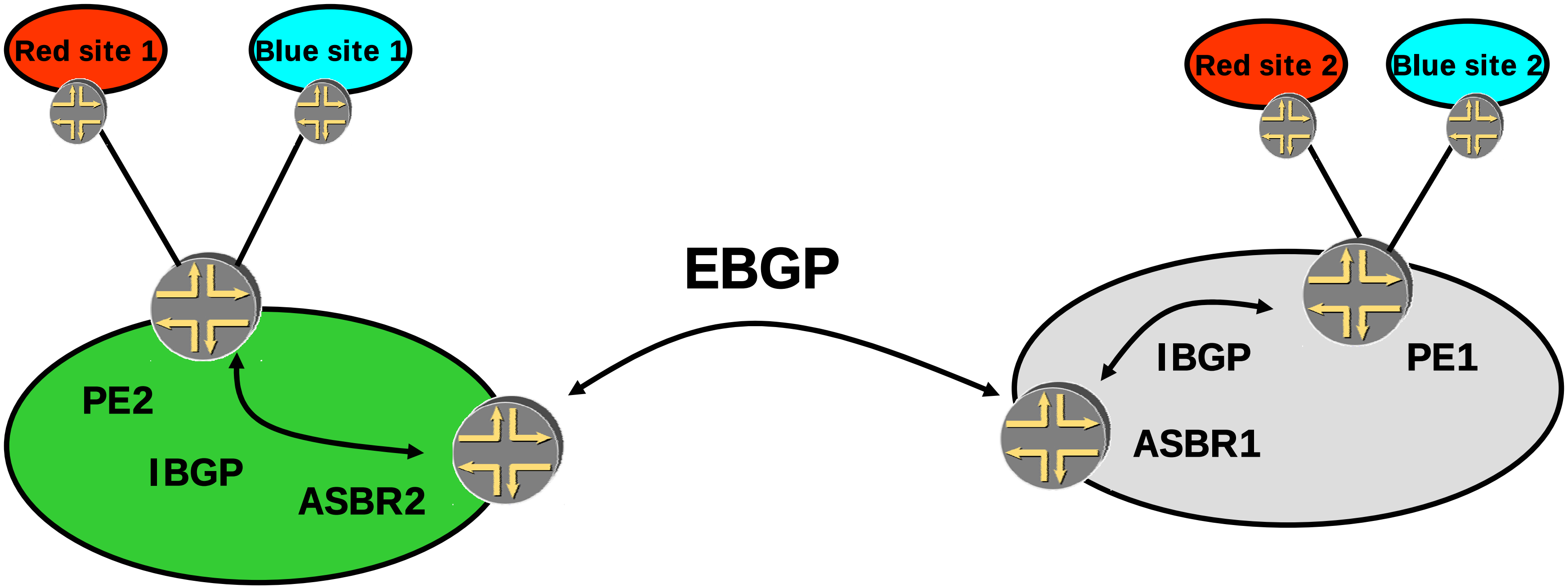
Option A – VRF-to-VRF connections

- **Major scaling issues:**
  - **All VPN routes are exchanged.**
  - **Multiple EBGP sessions need to be maintained.**
  - **The ASBRs must carry a large number of routes.**

Option B – EBGP redistribution of labeled VPN-IP routes between ASBRs

- The PE routers use IBGP to redistribute labeled VPN-IP routes to an ASBR.
- The ASBR uses EBGP to redistribute the labeled routes to an ASBR in a different AS.
- Requires a label-switched path across AS between the PEs.

Option B – EBGP redistribution of labeled VPN-IP routes between ASBRs



All routes exchanged are labeled VPN-IP.

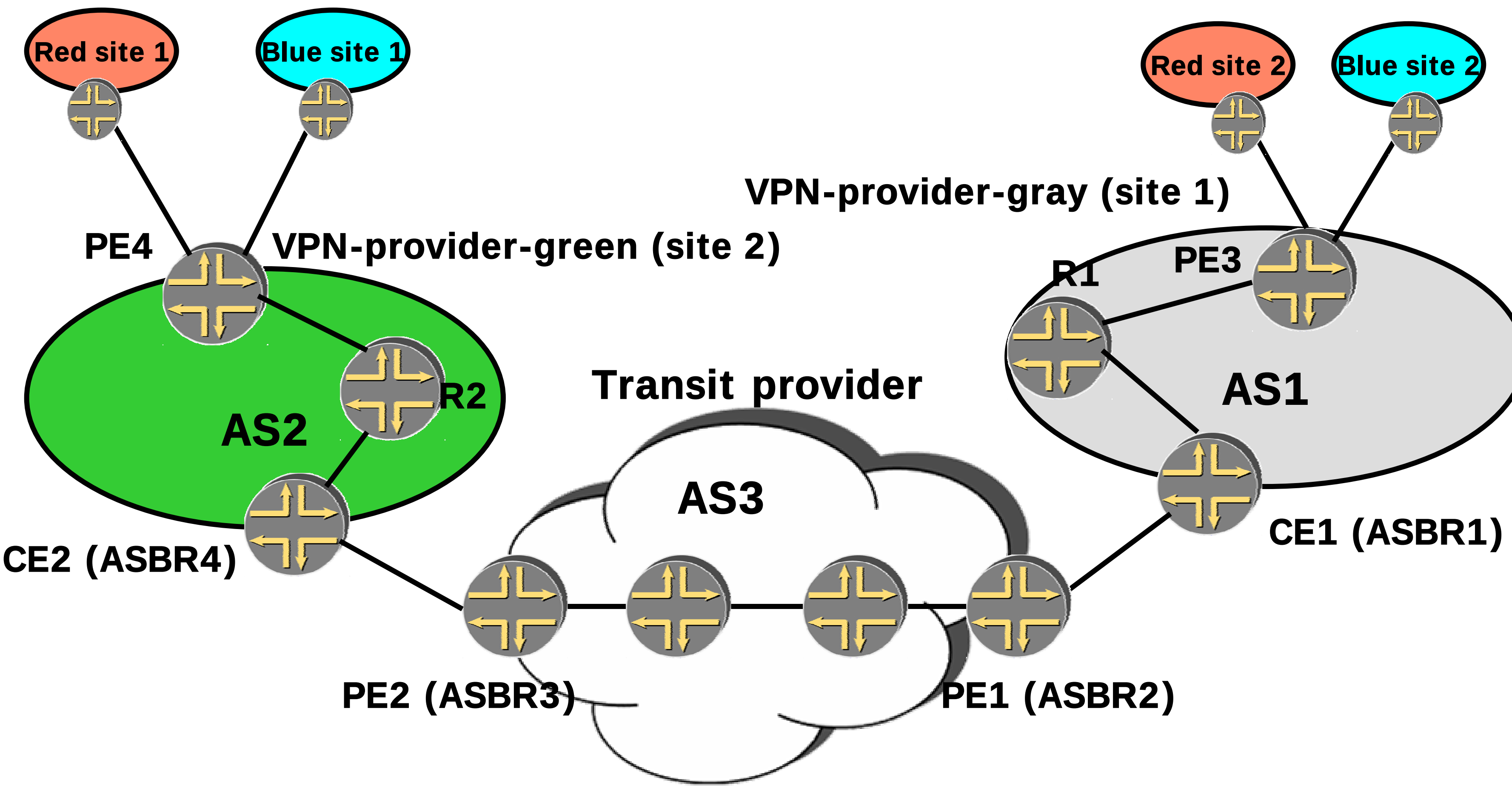
Option B – EBGP redistribution of labeled VPN-IP routes between ASBRs

- More scalable than option A:
  - No need for per/VPN configuration at the ASBRs.
- Still exchange all the VPN routes.
- Requires an inter-AS LSP between the two PEs.

Option C – EBGP redistribution of labeled VPN-IP routes between PEs

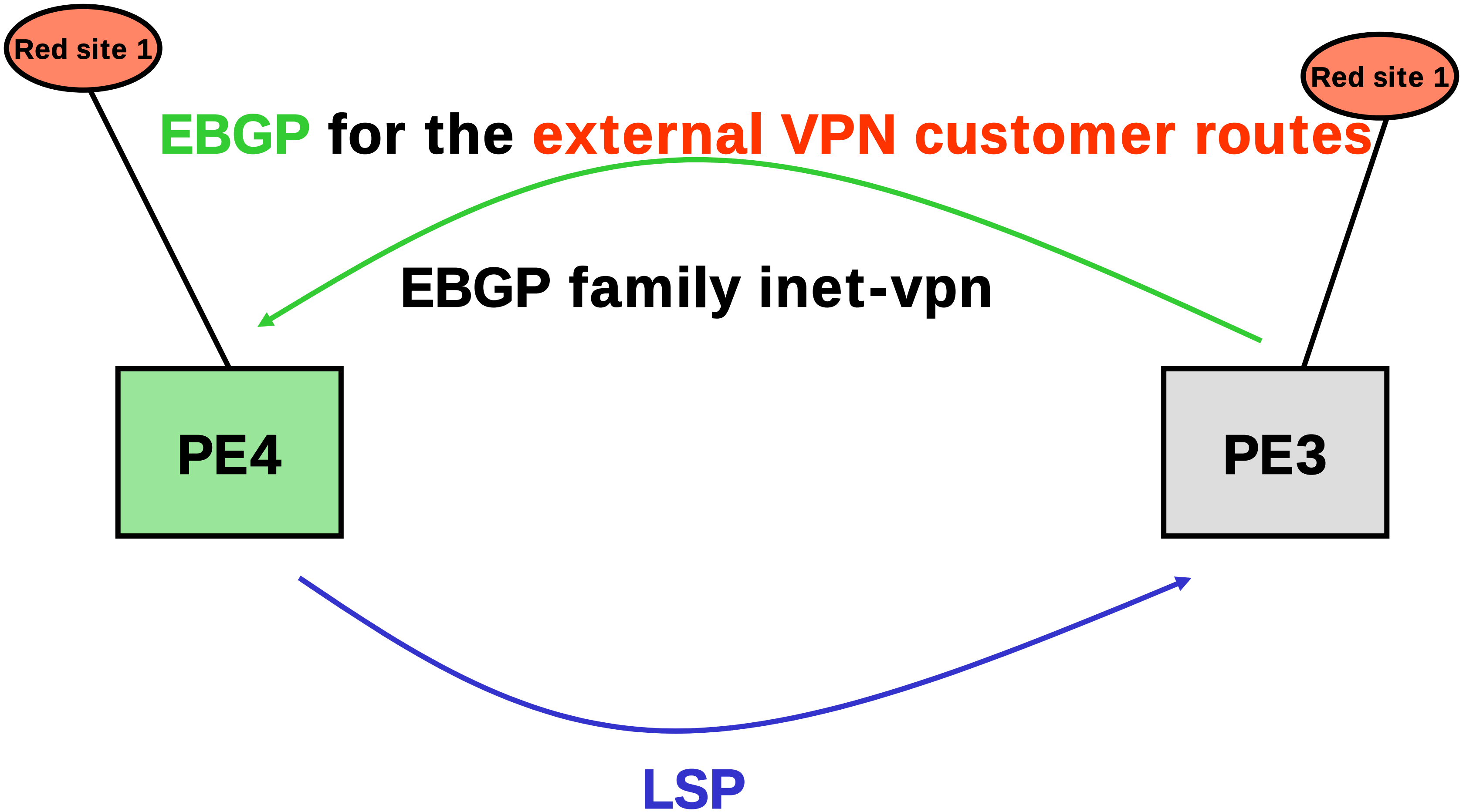
- In both option A and option B
  - All VPN routes are exchanged.
  - The scalability is determined by the amount of VPN routing information.
  - The load on the ASBRs is determined by the amount of VPN information carried.
- Option C – use multi-hop EBGP to distribute the VPN-IP routes between the PEs. The ASBRs exchange the internal routes, not the VPN routes.

Option C



Red site 1
Blue site 1
Red site 2
Blue site 2
VPN-provider-gray (site 1)
PE4
VPN-provider-green (site 2)
PE3
R1
R2
AS1
AS2
Transit provider
AS3
CE2 (ASBR4)
CE1 (ASBR1)
PE2 (ASBR3)
PE1 (ASBR2)

The abstraction



Option C

- **Looks like VPN-provider as VPN-customer**
  - **The same routing exchanges will happen.**
  - **Relies on the fact that the next-hop won’t be changed on the E-BGP session.**
- **Differences:**
  - **The two sites are attached to different AS.**
  - **Requires EBGP instead of IBGP between the customer PEs.**

Thank you!

Please send comments to
ina@juniper.net