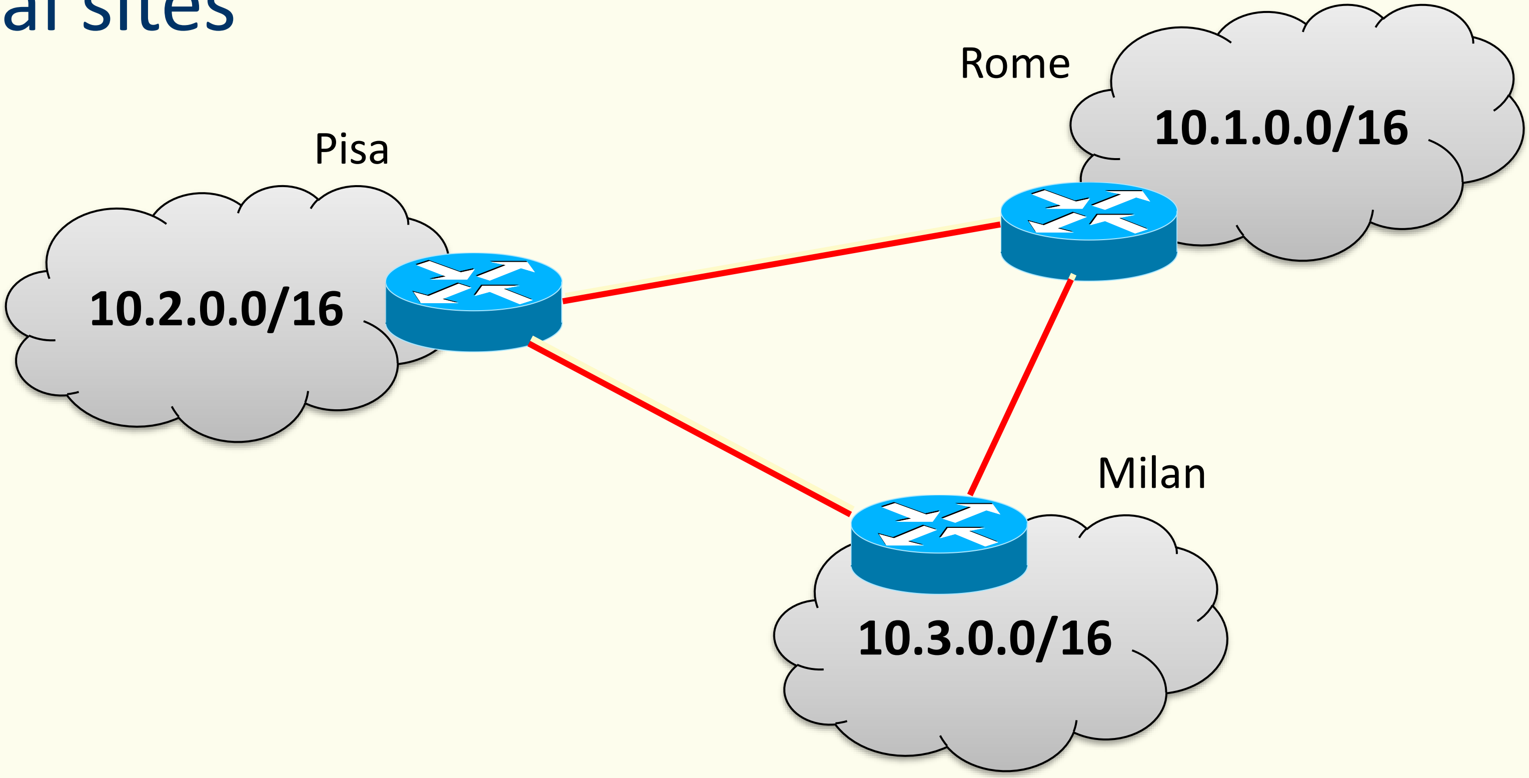
Virtual Private Networks

BGP/MPLS IP VPNs

Enzo Mingozzi
Professor @ University of Pisa
enzo.mingozzi@unipi.it

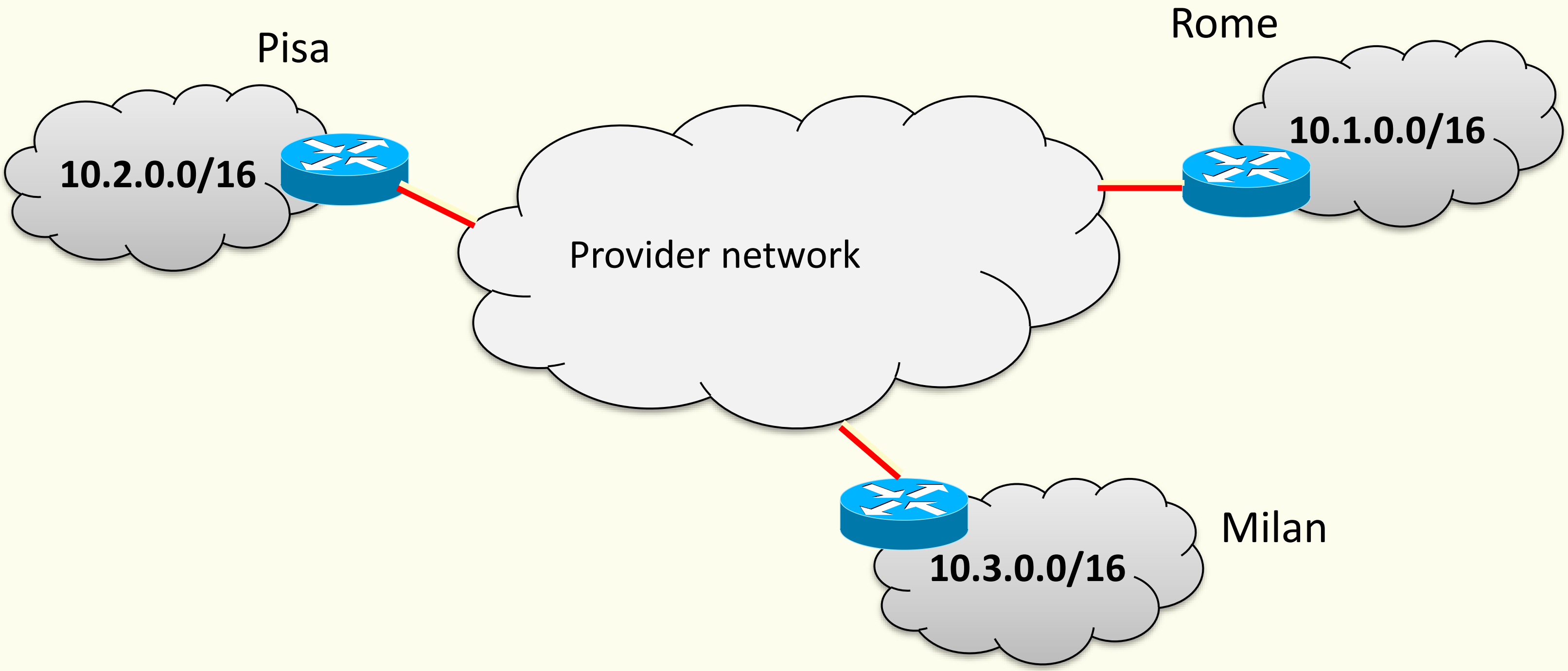
Corporate WANs

- The Wide Area Network (WAN) infrastructure is the set of links interconnecting border routers at local sites



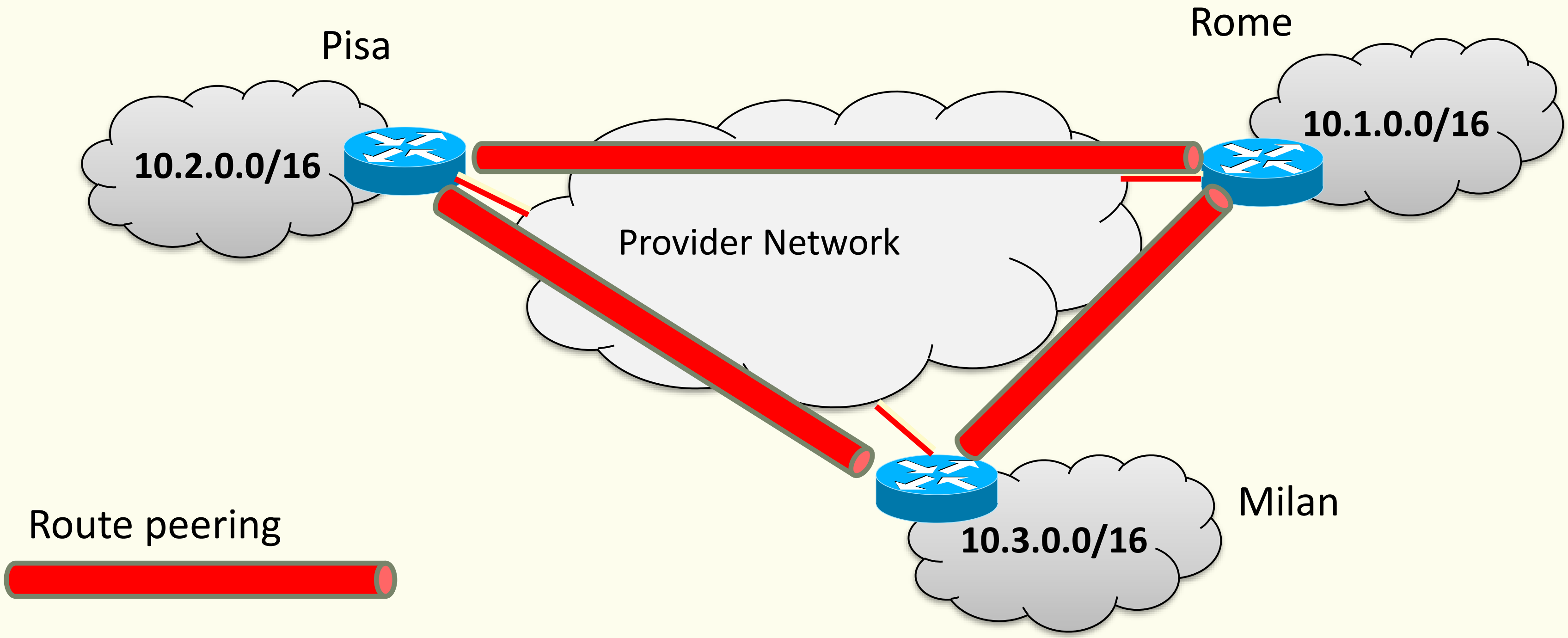
Virtual Private Networks

- **Private**: exclusive use, independent addressing and routing
- **Virtual**: the actual infrastructure is shared



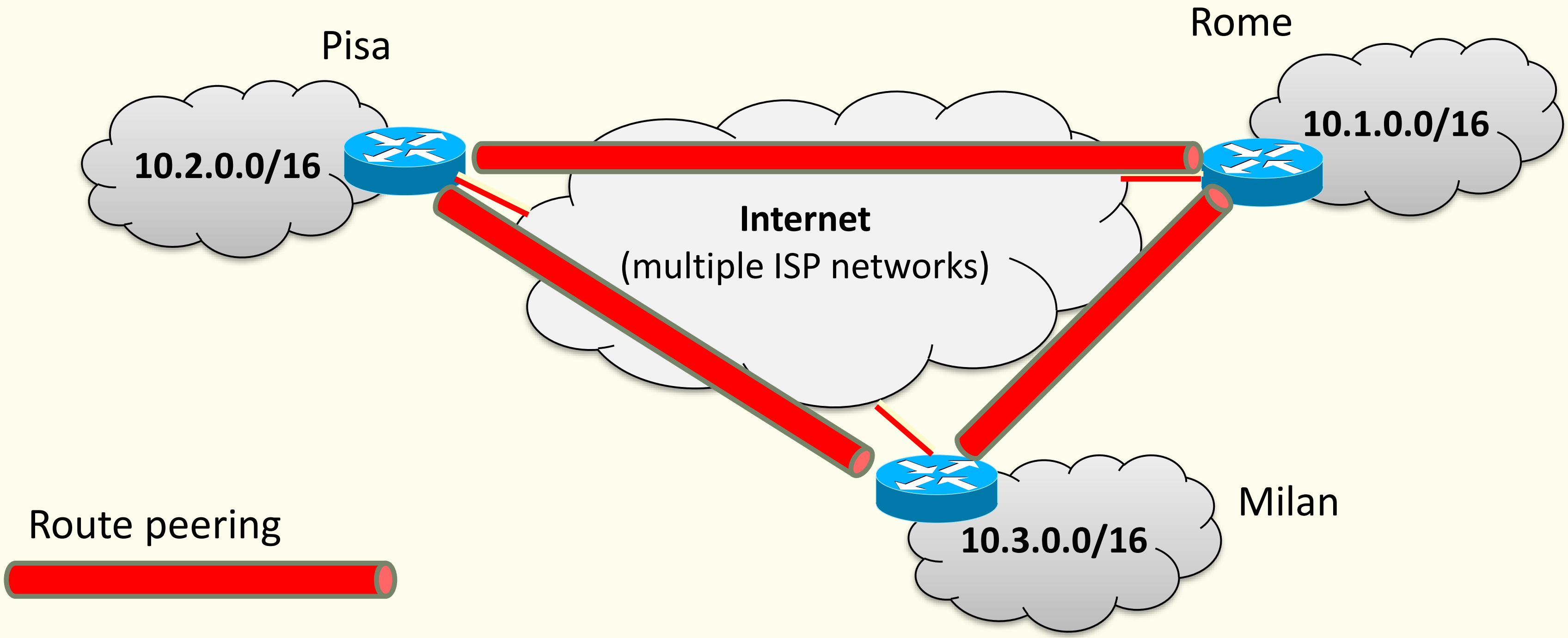
Virtual CE-CE links

- **Virtual backbone overlay** on top of PN
  - Leased lines (L1, dedicated circuit)
  - Frame Relay (L2, pachet switched)



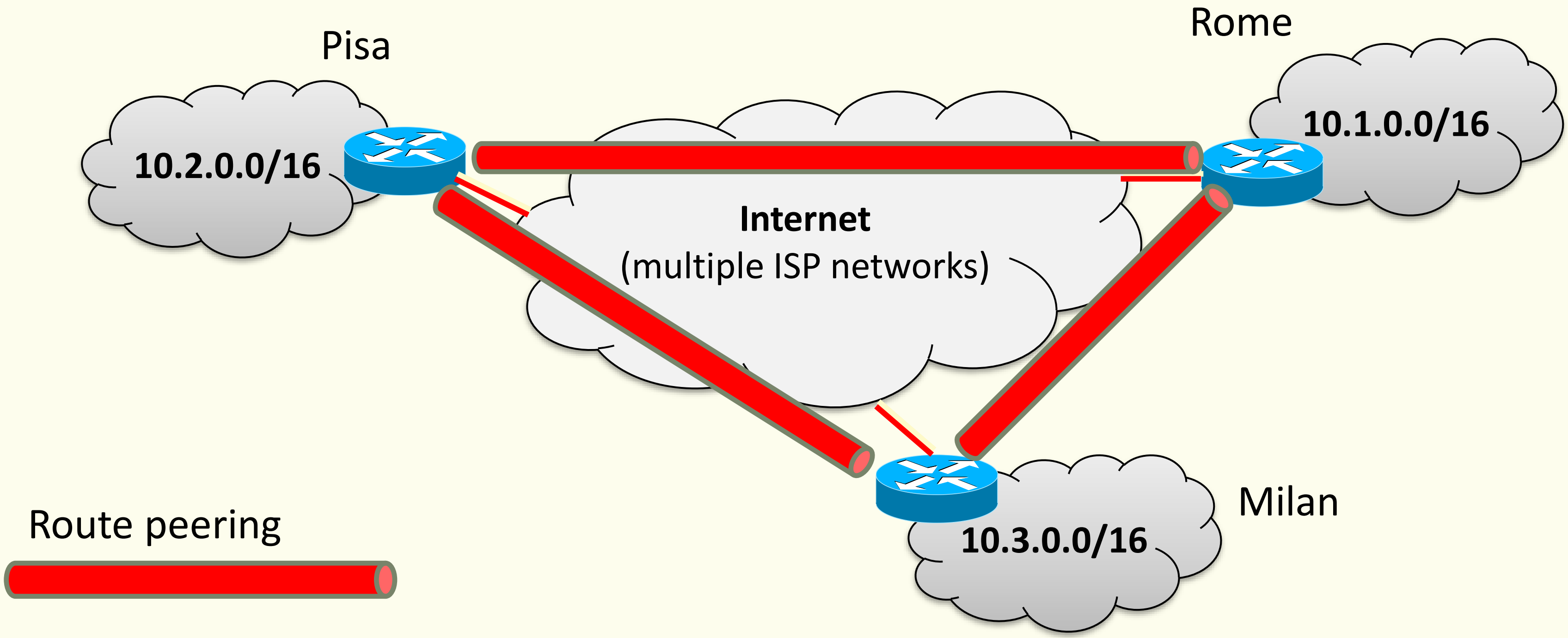
Virtual CE-CE links

- **Virtual backbone overlay** on top of PN
  - GRE or IPSec tunneling over Internet



Virtual CE-CE links

- CE routers at different sites **peer with each other**
- The overlay is ***visible*** to the VPN's **routing algorithm**

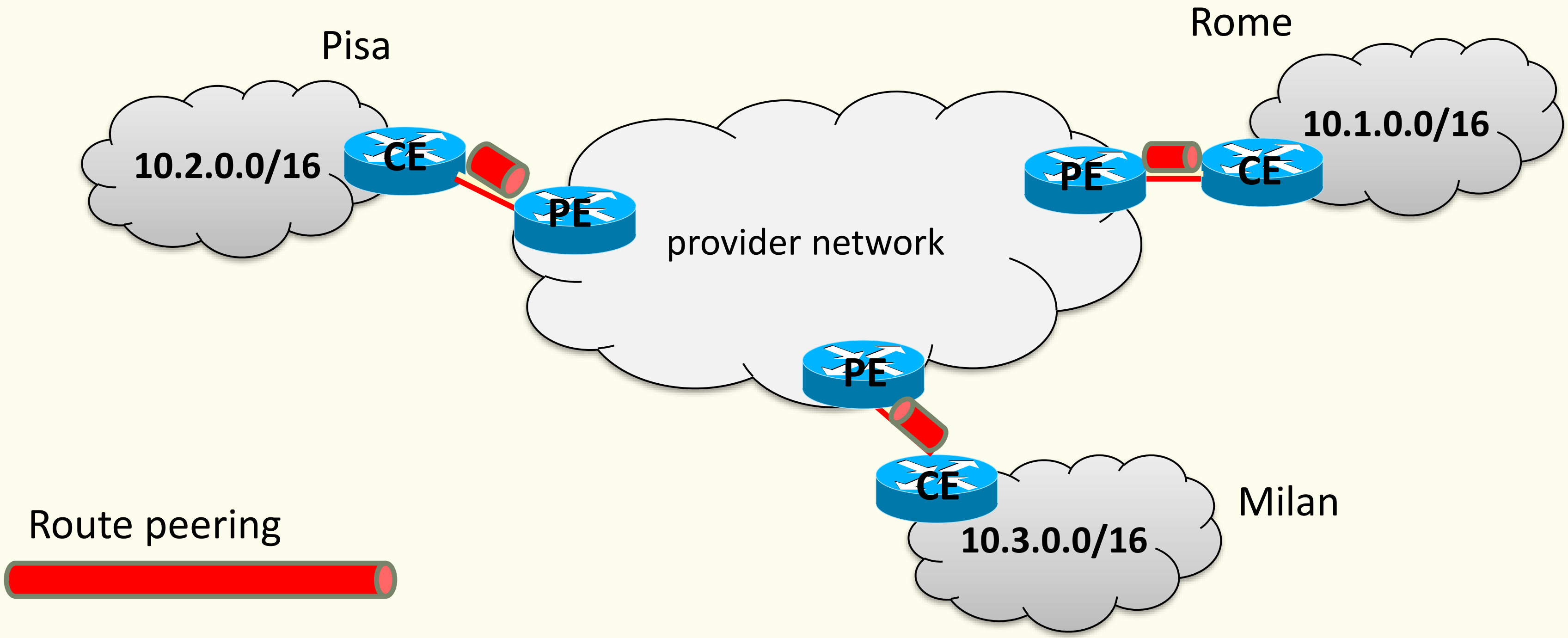


Virtual CE-CE links

- **Pros**
  - Achieves the fundamental goals of a VPN
    - Connectivity, private addressing, privacy of traffic
- **Cons**
  - VPN is implemented by Customer Edge (CE) routers
    - Requires network management expertise
  - IGP scaling routing limitations (mesh of CE peers)
  - Amount of configuration required for adding a new site
  - If provider managed
    - Scaling of management limitations with multiple customers

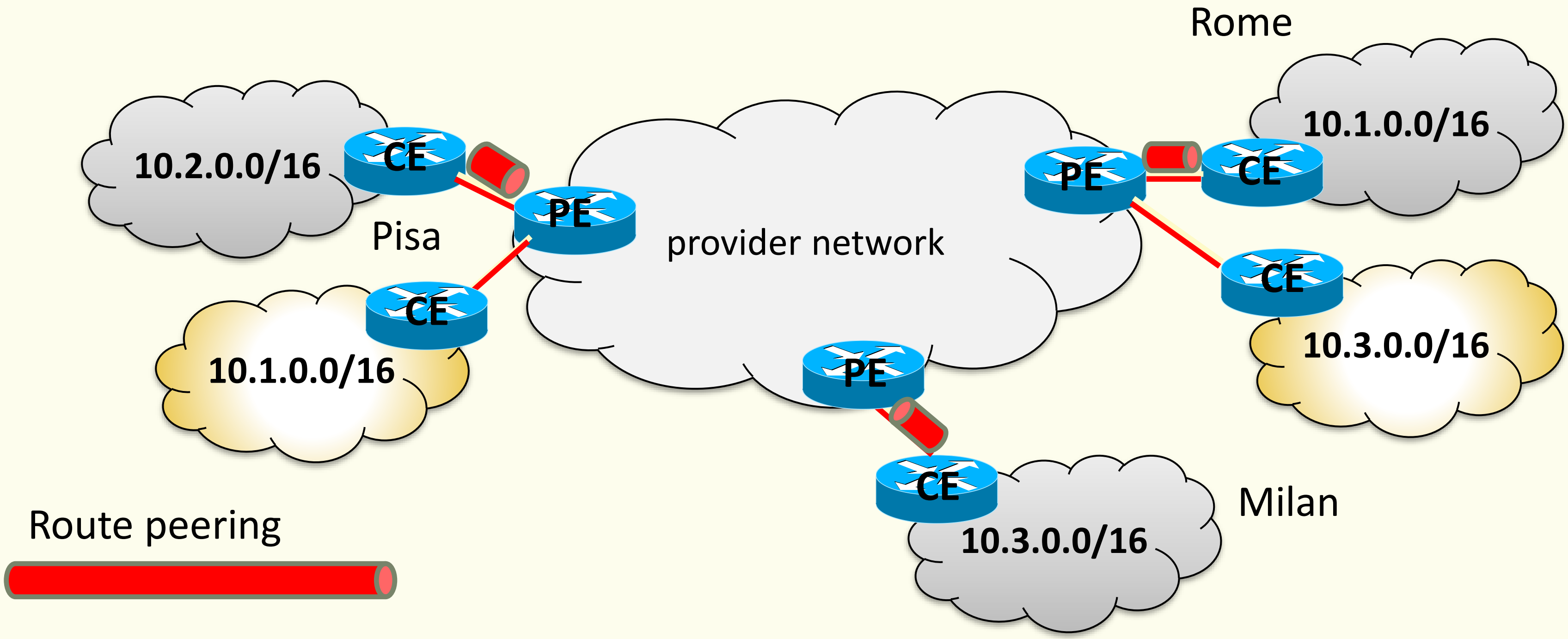
CE-PE peering approach

- **PE-based VPN**
  - – scales with the number of customers
- Simple routing config at CEs



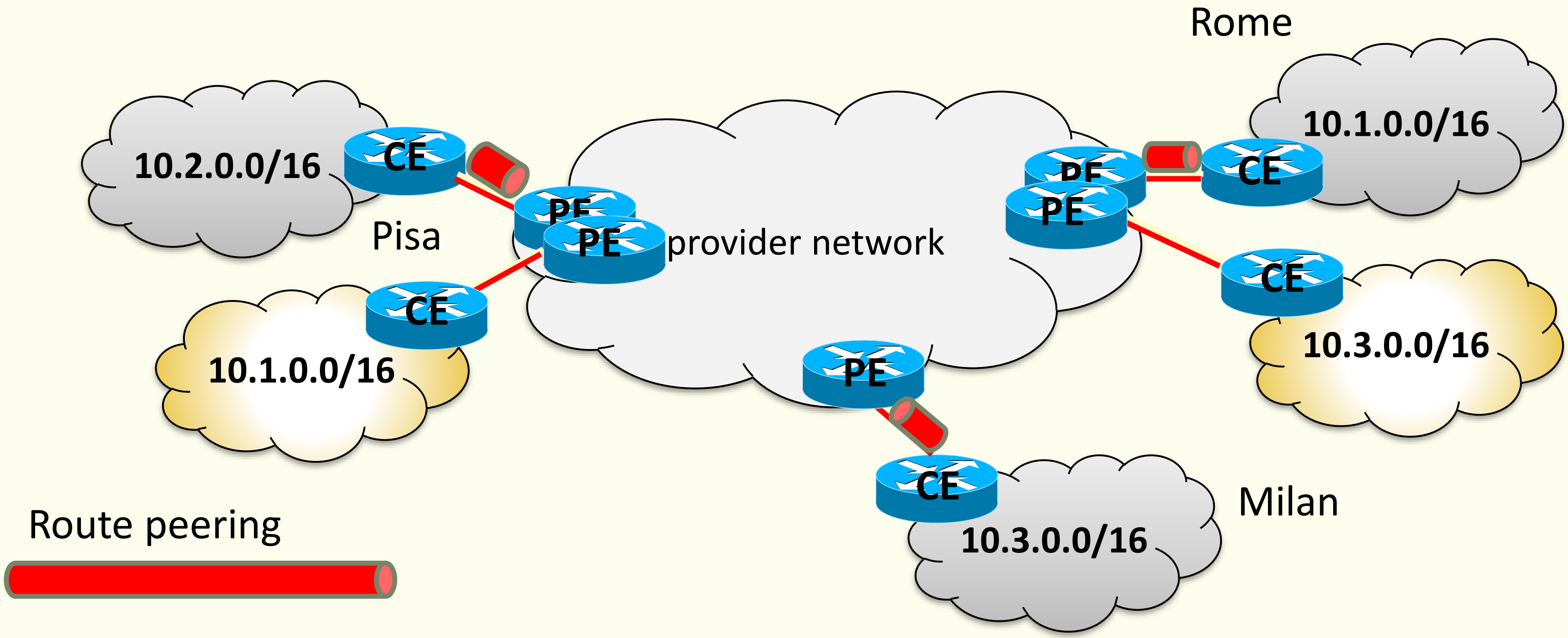
CE-PE peering approach

- How to achieve VPN goals?
  – private addressing, isolation of traffic
- Constrain traffic at forwarding time with ACLs ☹



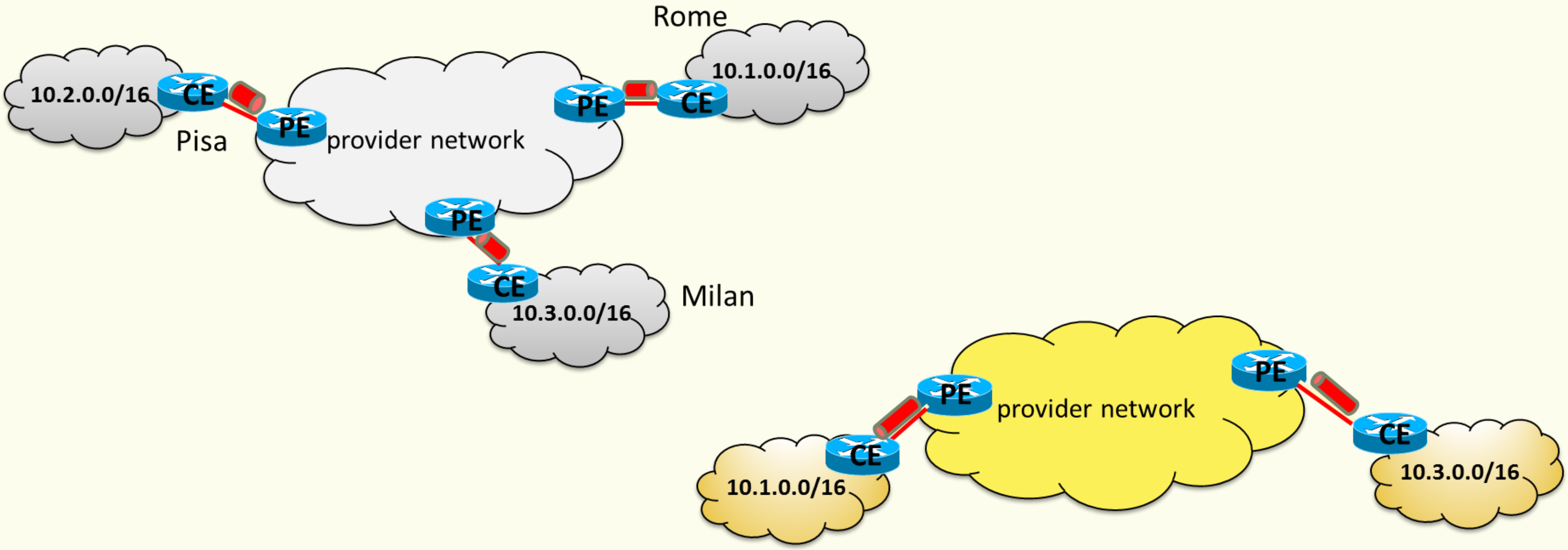
CE-PE peering approach

- How to achieve VPN goals?
  - private addressing, isolation of traffic
- Constrain routing information distribution ☹



CE-PE peering approach

- How to achieve VPN goals?
  - – private addressing, isolation of traffic
- BGP/MPLS IP VPNs

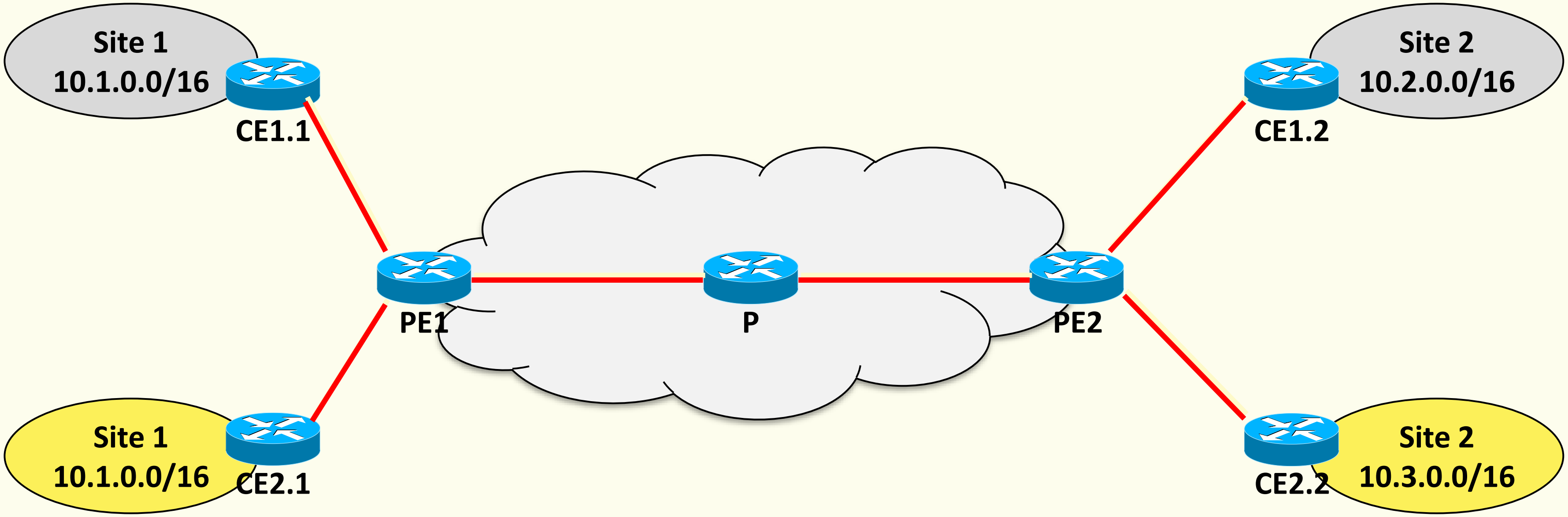


BGP/MPLS IP VPNs

- Originally developed as a Cisco solution for provider-provisioned VPNs
- Following its success, standardized afterwards as RFC 4364
- Also known as **L3VPNs**

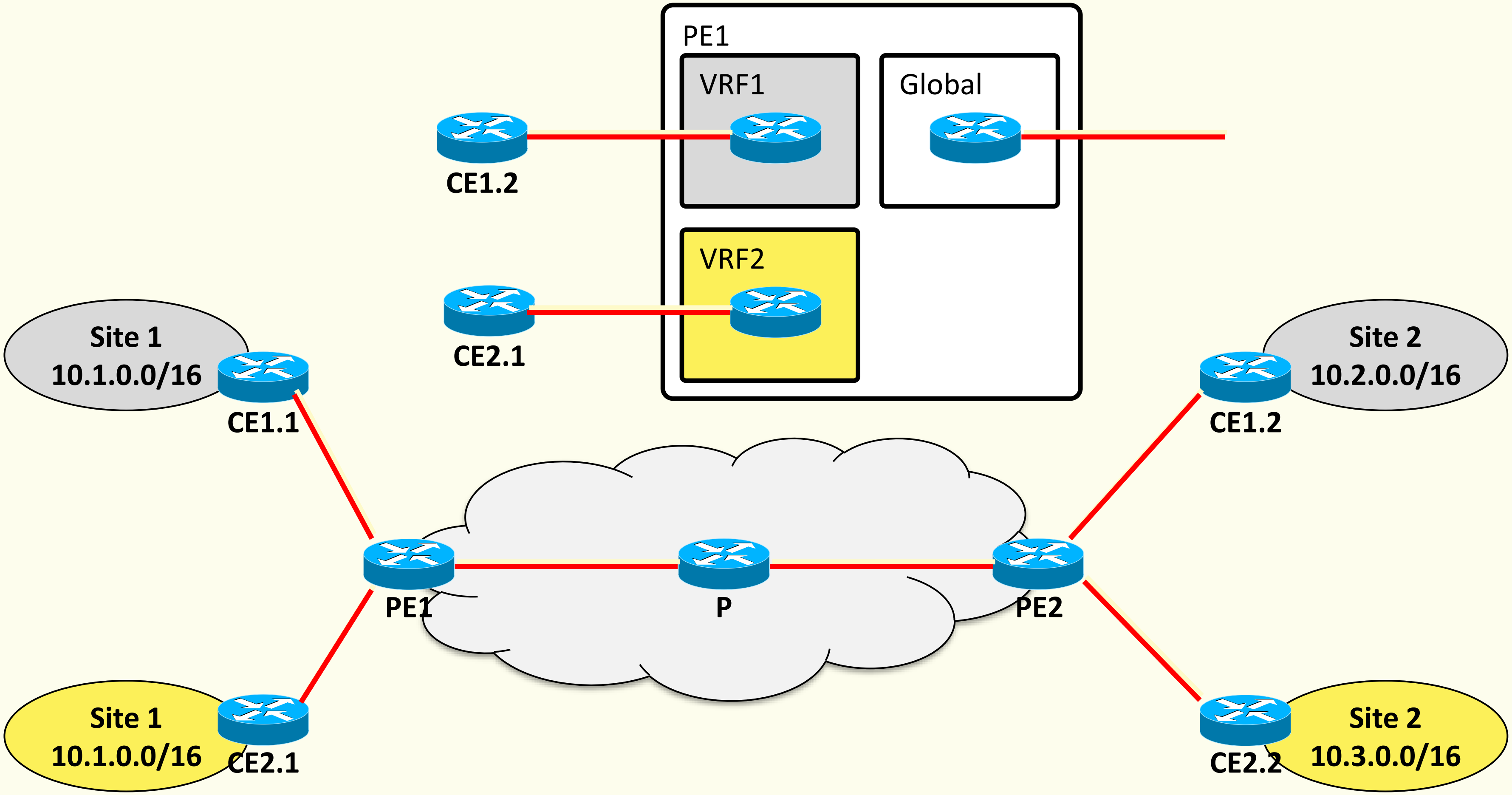
Example network

- Customer sites
  - Multiple sites may be attached to the same PE
  - One CE may be attached to multiple PEs
  - Multiple networks within each site



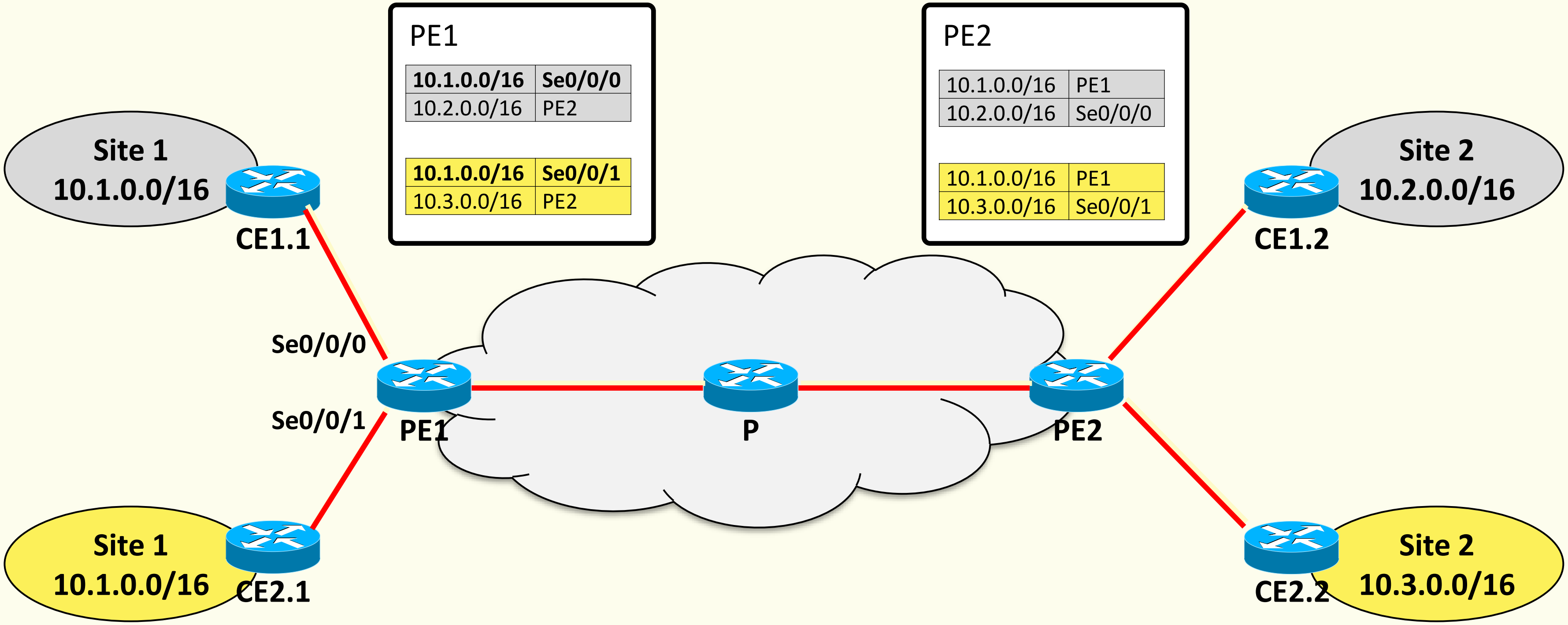
Isolation of traffic

- Per-VPN routing and forwarding tables (VRF)



Isolation of traffic

- Per-VPN routing and forwarding tables (VRF)
  - VRF look-up based on associating interfaces to CE (physical or logical) to VRFs by configuration



PE1

10.1.0.0/16	**Se0/0/0**
10.2.0.0/16	PE2

10.1.0.0/16	**Se0/0/1**
10.3.0.0/16	PE2

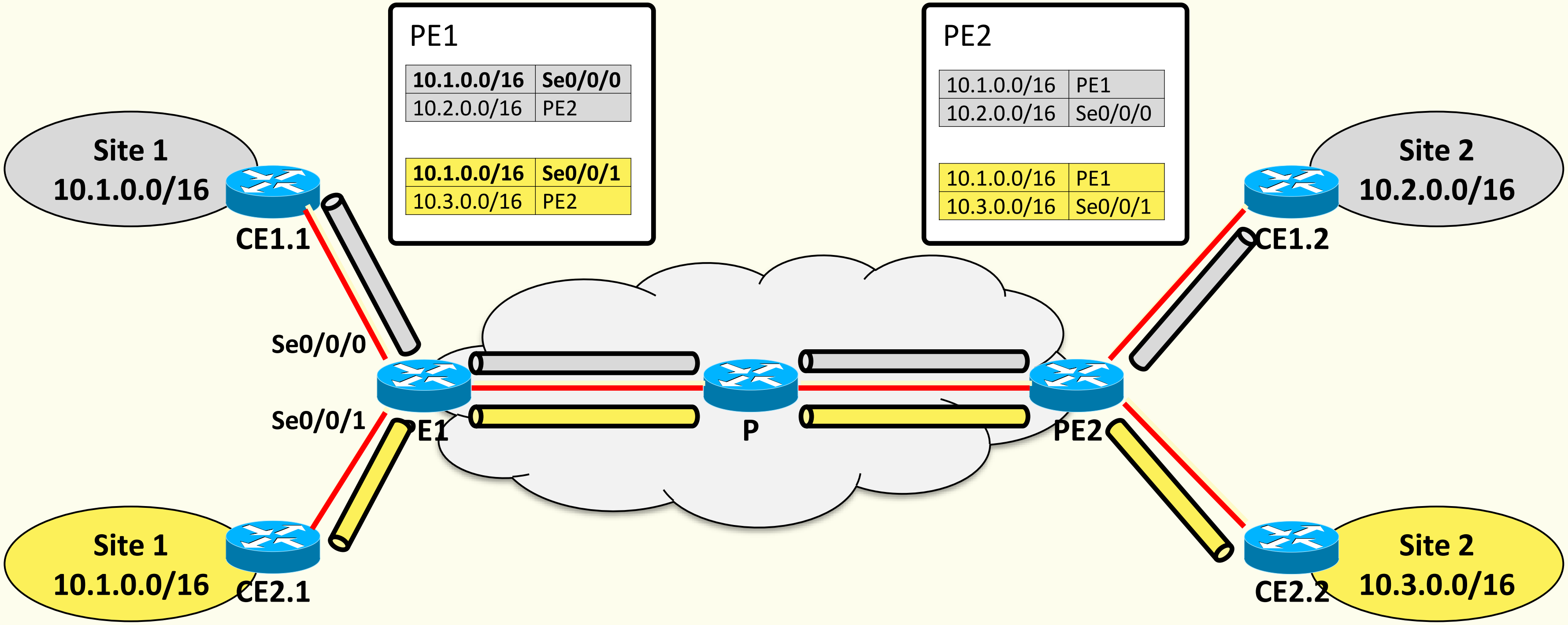
PE2

10.1.0.0/16	PE1
10.2.0.0/16	Se0/0/0

10.1.0.0/16	PE1
10.3.0.0/16	Se0/0/1

Constrained route distribution

- Run one routing protocol instance per VPN ☹



PE1

10.1.0.0/16	Se0/0/0
10.2.0.0/16	PE2

10.1.0.0/16	Se0/0/1
10.3.0.0/16	PE2

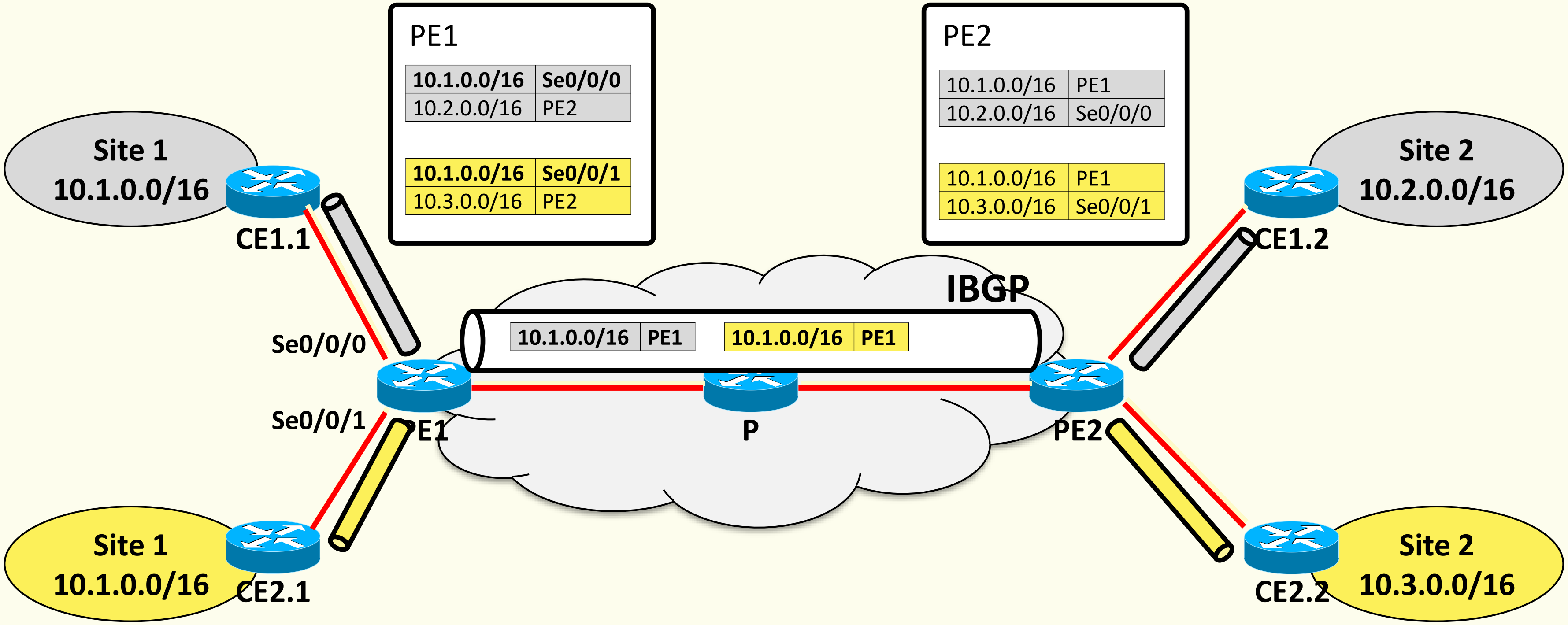
PE2

10.1.0.0/16	PE1
10.2.0.0/16	Se0/0/0

10.1.0.0/16	PE1
10.3.0.0/16	Se0/0/1

Route distribution

- Use iBGP to carry all routes
  - Supports route filtering
  - Supports route distribution between remote routers



PE1

10.1.0.0/16	Se0/0/0
10.2.0.0/16	PE2

10.1.0.0/16	Se0/0/1
10.3.0.0/16	PE2

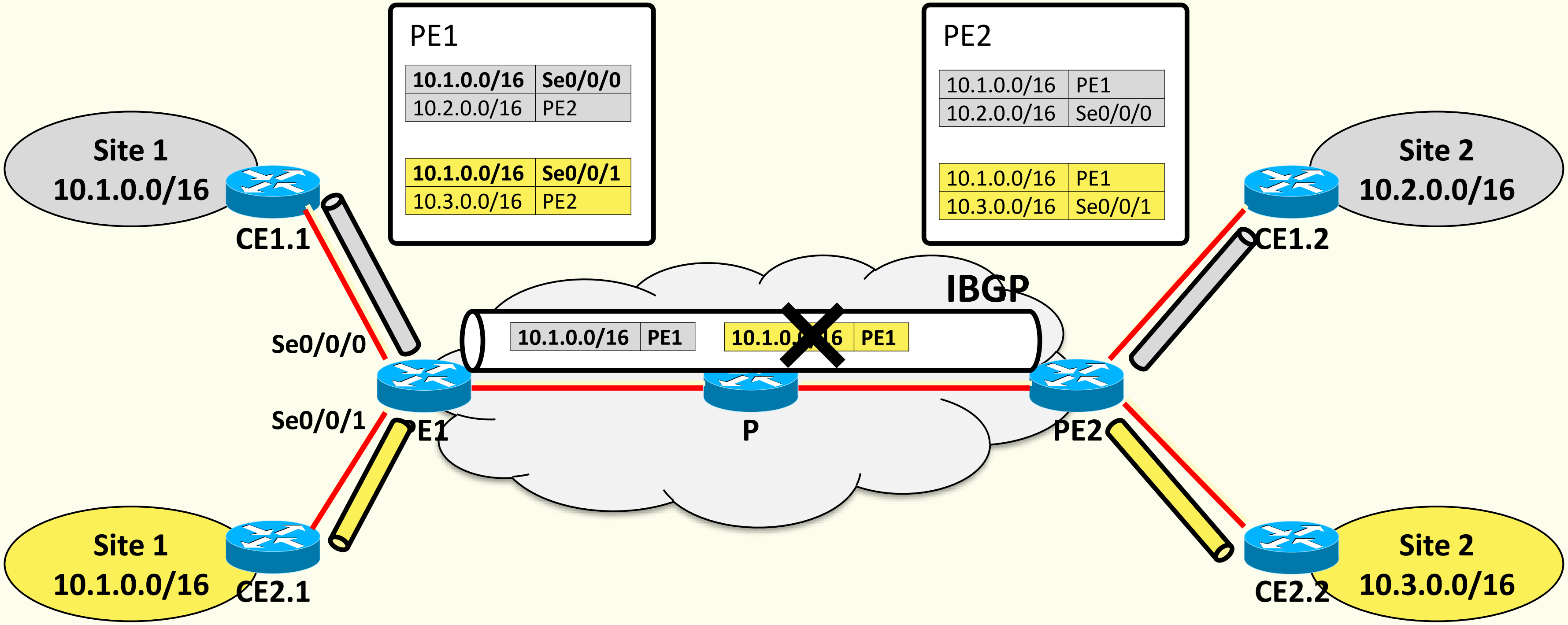
PE2

10.1.0.0/16	PE1
10.2.0.0/16	Se0/0/0

10.1.0.0/16	PE1
10.3.0.0/16	Se0/0/1

Route distribution

- Use iBGP to carry all routes
  - Can only distribute one route to a given address prefix



PE1

10.1.0.0/16	Se0/0/0
10.2.0.0/16	PE2

10.1.0.0/16	Se0/0/1
10.3.0.0/16	PE2

PE2

10.1.0.0/16	PE1
10.2.0.0/16	Se0/0/0

10.1.0.0/16	PE1
10.3.0.0/16	Se0/0/1

Multiprotocol extensions for BGP

- BGP-4, originally supporting IPv4 only, has been **extended** to carry routing information for **multiple Network Layer protocols** (e.g. IPv6) [RFC 4760], referred as **MP-BGP**
- New attributes: **Multiprotocol Reachable NLRI**, and **Multiprotocol Unreachable NLRI**
- Network Layer protocol identified by the pair
  - **Address Family Identifier (AFI)**: e.g. 1 (IPv4), 2 (IPv6)
  - **Subsequent AFI**: e.g. 1 (unicast), 2 (multicast)

VPN-IP addresses

- Define a new address family **VPN-IPv4**
  - AFI=1 (*IPv4*), SAFI=128 (*MPLS-labeled VPN address*)
- VPN-IP addresses are obtained from customer site addresses by pre-pending an 8-byte identifier named **Route Distinguisher (RD)**
- RDs must be unique **globally**
  - E.g.,

TYPE (2)	**AS number (2)**	**Locally assigned number (4)**

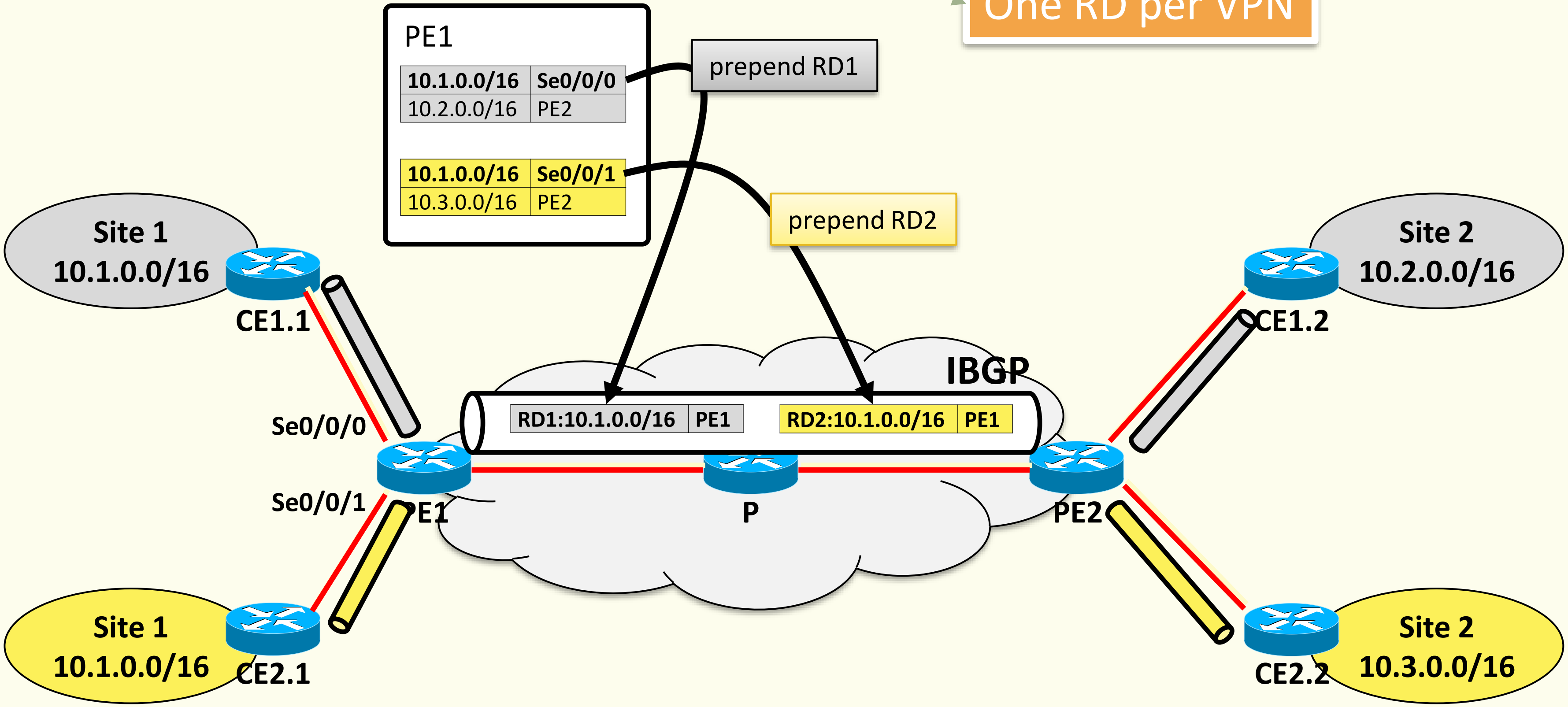
VPN-IP addresses

- RD – route association
  - By configuration only!!!

One RD per VPN per route

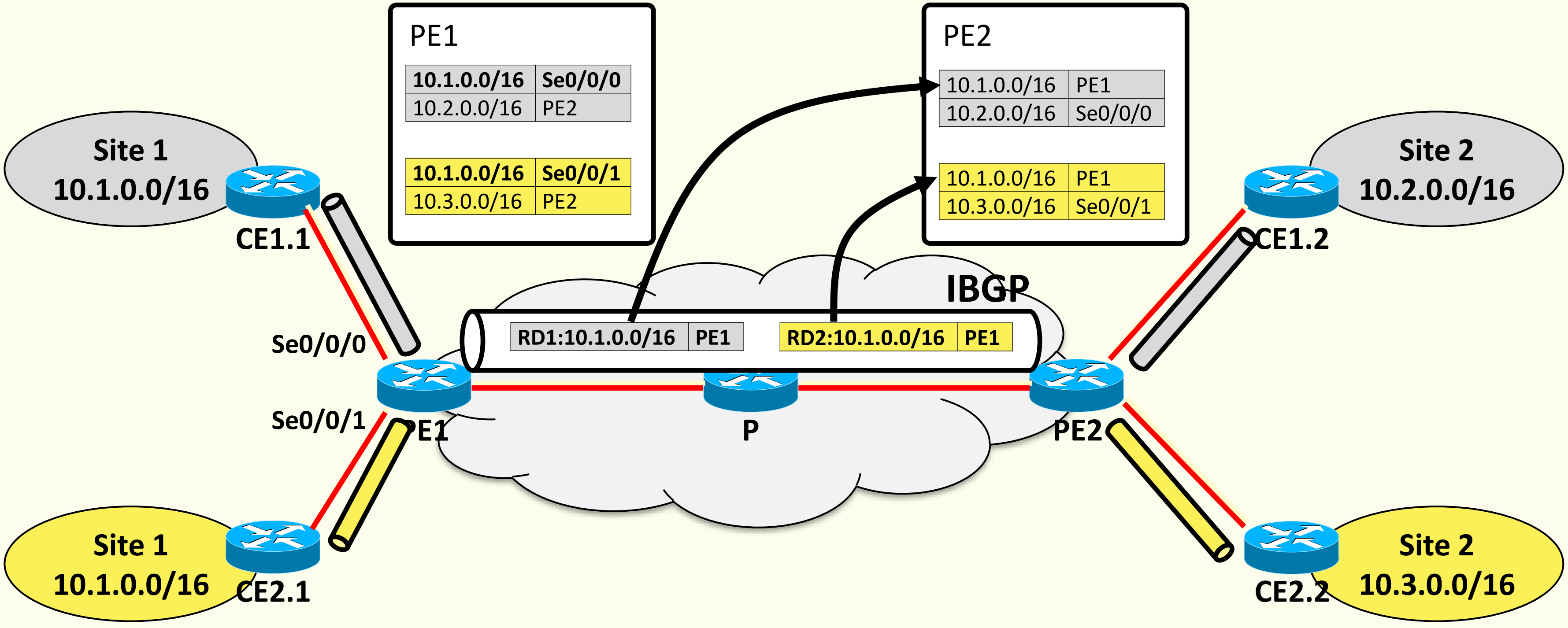
One RD per VPN per PE

One RD per VPN



Route distribution

- Use iBGP to carry VPN-IP routes
  - RDs are stripped off when **redistributing** BGP updates into VRFs



PE1

10.1.0.0/16	**Se0/0/0**
10.2.0.0/16	PE2

10.1.0.0/16	**Se0/0/1**
10.3.0.0/16	PE2

PE2

10.1.0.0/16	PE1
10.2.0.0/16	Se0/0/0

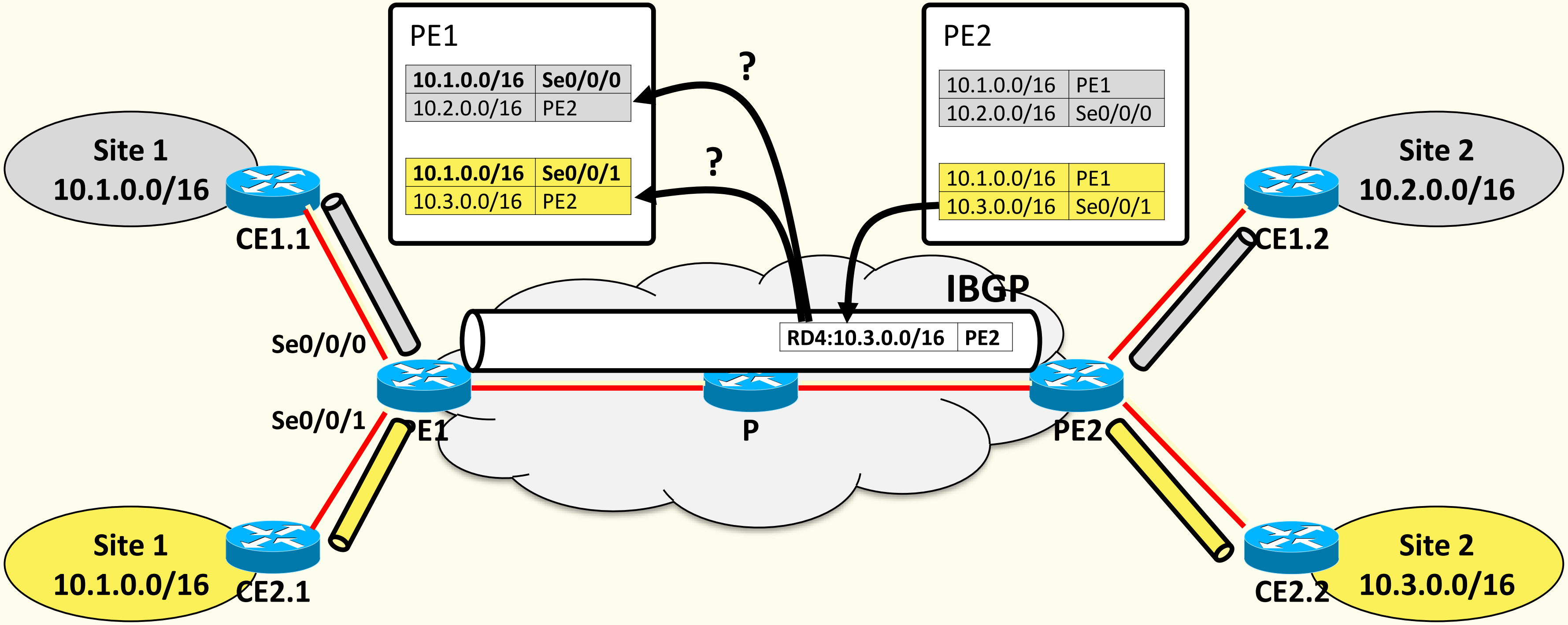
10.1.0.0/16	PE1
10.3.0.0/16	Se0/0/1

RD1:10.1.0.0/16	**PE1**	**RD2:10.1.0.0/16**	**PE1**

Route distribution

- The **only purpose** of RDs is to make the VPN routes unique

The target VRF **is not inferrable** in any manner from the RD



PE1

10.1.0.0/16	**Se0/0/0**
10.2.0.0/16	PE2

10.1.0.0/16	**Se0/0/1**
10.3.0.0/16	PE2

PE2

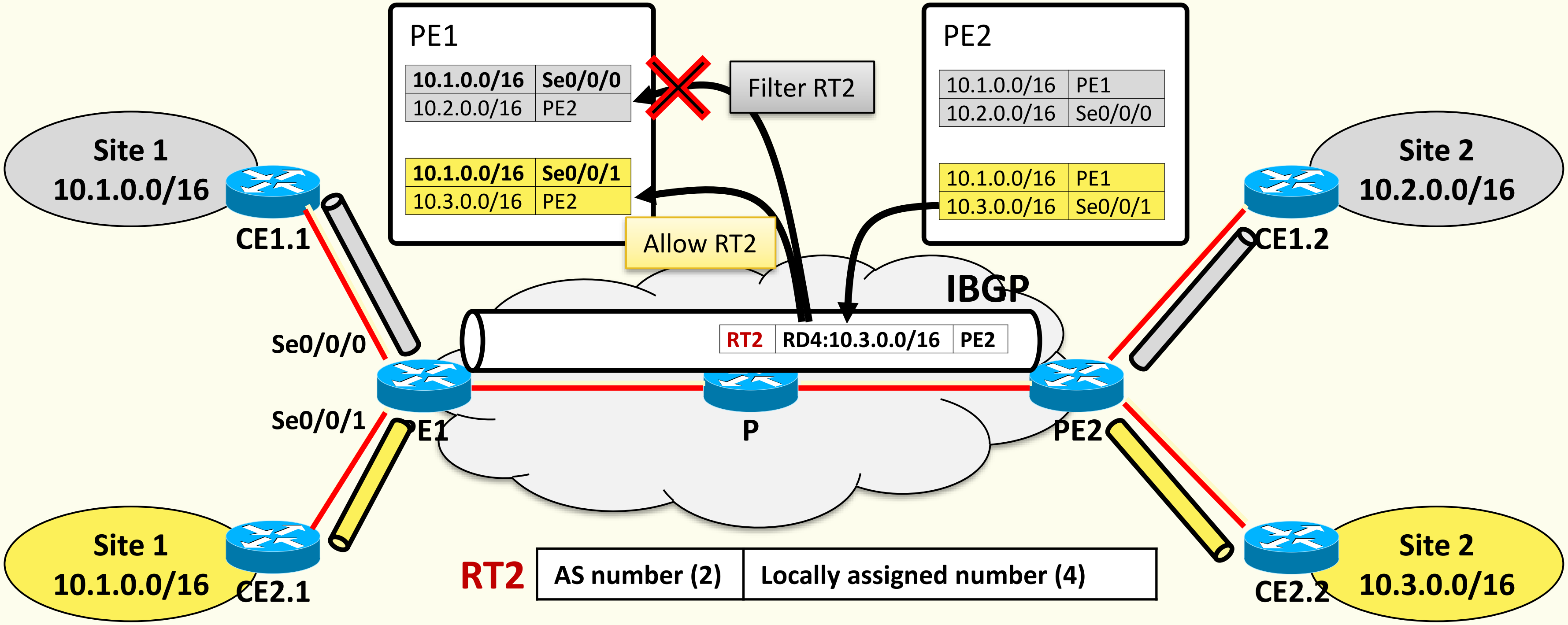
10.1.0.0/16	PE1
10.2.0.0/16	Se0/0/0

10.1.0.0/16	PE1
10.3.0.0/16	Se0/0/1

RD4:10.3.0.0/16	PE2

Constrained route distribution

- Use of the BGP (Extended) Community attribute to carry a **Route Target** to **filter** routes out



PE1

10.1.0.0/16	**Se0/0/0**
10.2.0.0/16	PE2

10.1.0.0/16	**Se0/0/1**
10.3.0.0/16	PE2

PE2

10.1.0.0/16	PE1
10.2.0.0/16	Se0/0/0

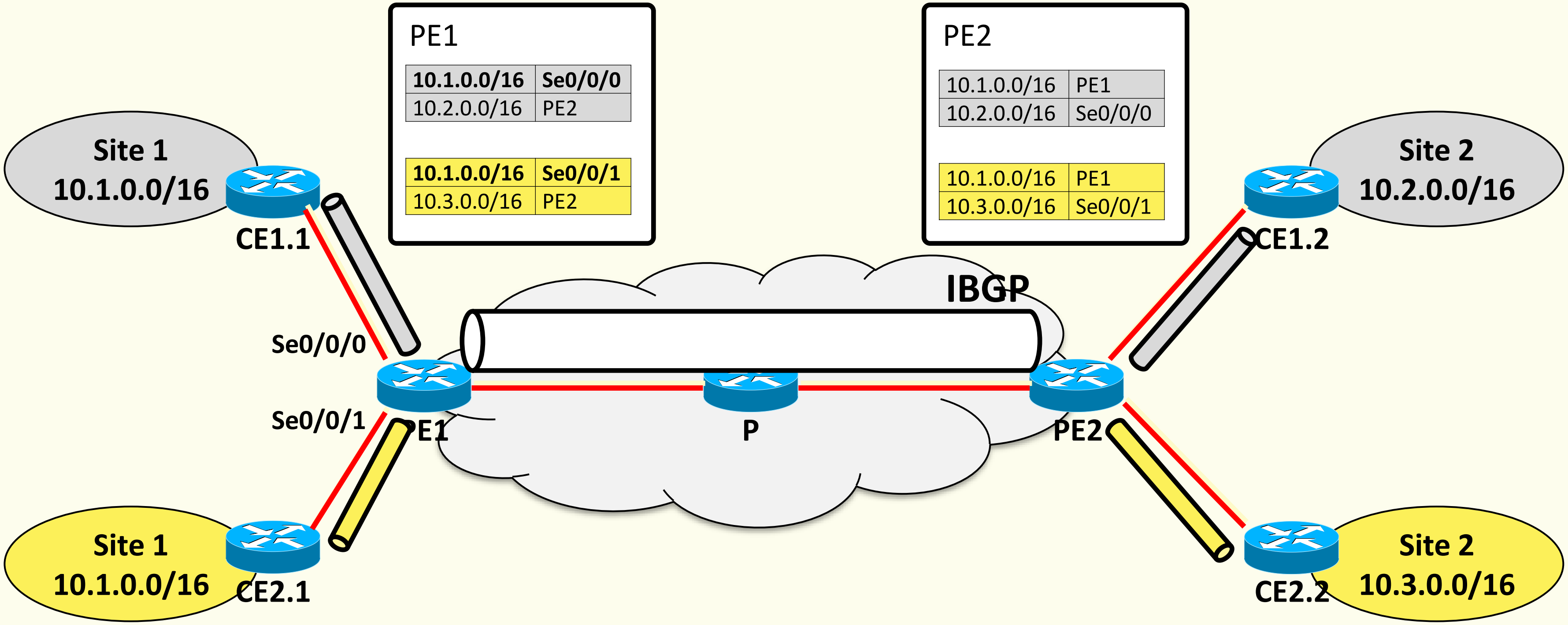
10.1.0.0/16	PE1
10.3.0.0/16	Se0/0/1

RT2	RD4:10.3.0.0/16	PE2

RT2	**AS number (2)**	**Locally assigned number (4)**

Constrained route distribution

- Broader requirements than simple VPN isolation
  - Support for **overlapping** VPNs (one site belongs to multiple VPNs)
  - Arbitrary and complex connectivity models



PE1

10.1.0.0/16	**Se0/0/0**
10.2.0.0/16	PE2

10.1.0.0/16	**Se0/0/1**
10.3.0.0/16	PE2

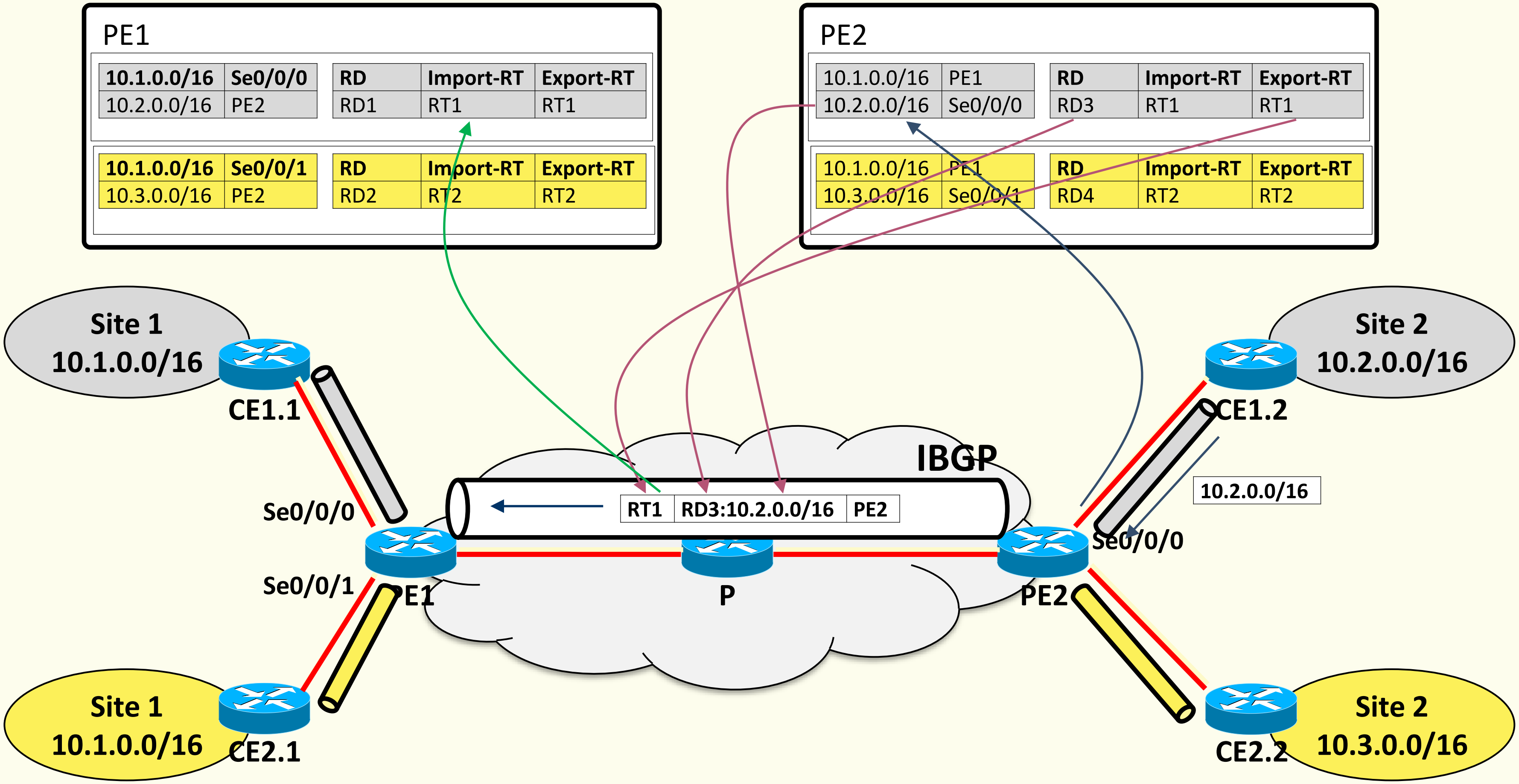
PE2

10.1.0.0/16	PE1
10.2.0.0/16	Se0/0/0

10.1.0.0/16	PE1
10.3.0.0/16	Se0/0/1

Constrained route distribution

- RT **import** and **export** policies on a per-VRF basis



PE1

10.1.0.0/16	**Se0/0/0**
10.2.0.0/16	PE2

RD	Import-RT	Export-RT
RD1	RT1	RT1

10.1.0.0/16	**Se0/0/1**
10.3.0.0/16	PE2

RD	Import-RT	Export-RT
RD2	RT2	RT2

PE2

10.1.0.0/16	PE1
10.2.0.0/16	Se0/0/0

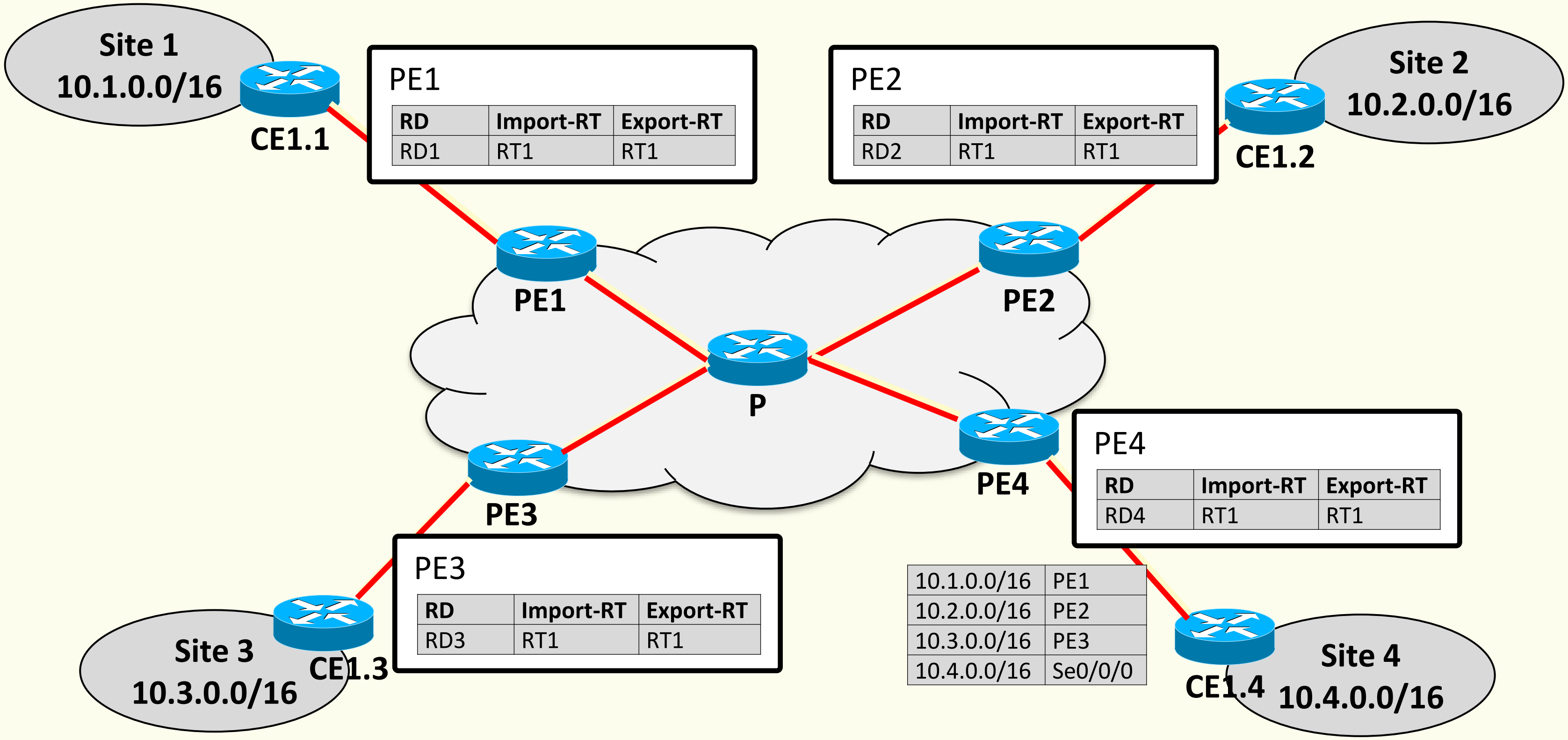
RD	Import-RT	Export-RT
RD3	RT1	RT1

10.1.0.0/16	PE1
10.3.0.0/16	Se0/0/1

RD	Import-RT	Export-RT
RD4	RT2	RT2

Constrained route distribution

- **Full-mesh**: single RT at all sites (in & out)



PE1

RD	Import-RT	Export-RT
RD1	RT1	RT1

PE2

RD	Import-RT	Export-RT
RD2	RT1	RT1

PE3

RD	Import-RT	Export-RT
RD3	RT1	RT1

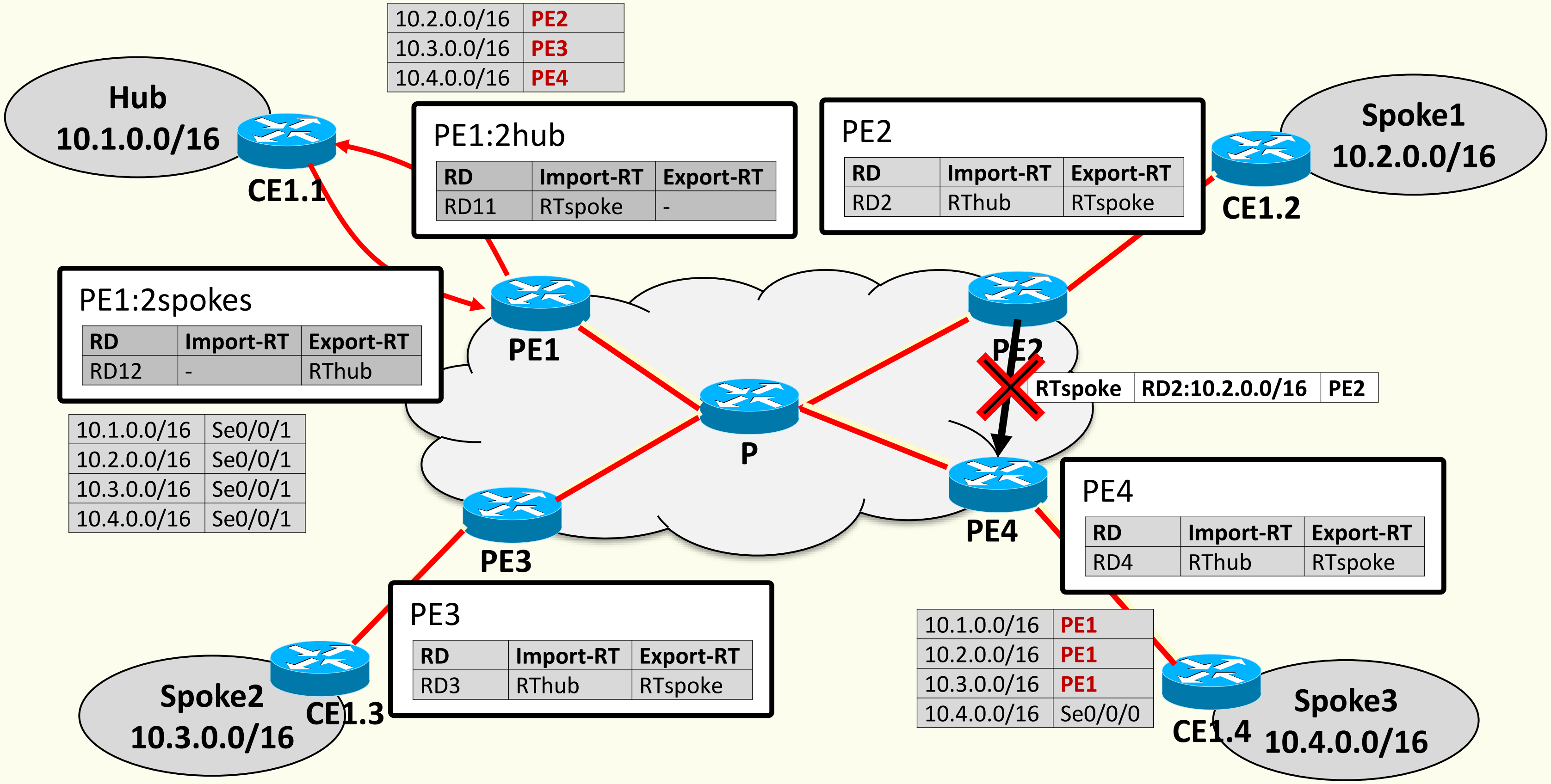
PE4

RD	Import-RT	Export-RT
RD4	RT1	RT1

10.1.0.0/16	PE1
10.2.0.0/16	PE2
10.3.0.0/16	PE3
10.4.0.0/16	Se0/0/0

Constrained route distribution

- **Hub-and-spoke**: RThub and RTspoke



10.2.0.0/16	**PE2**
10.3.0.0/16	**PE3**
10.4.0.0/16	**PE4**

PE1:2hub

RD	Import-RT	Export-RT
RD11	RTspoke	-

PE2

RD	Import-RT	Export-RT
RD2	RThub	RTspoke

PE1:2spokes

RD	Import-RT	Export-RT
RD12	-	RThub

10.1.0.0/16	Se0/0/1
10.2.0.0/16	Se0/0/1
10.3.0.0/16	Se0/0/1
10.4.0.0/16	Se0/0/1

RTspoke	**RD2:10.2.0.0/16**	**PE2**

PE4

RD	Import-RT	Export-RT
RD4	RThub	RTspoke

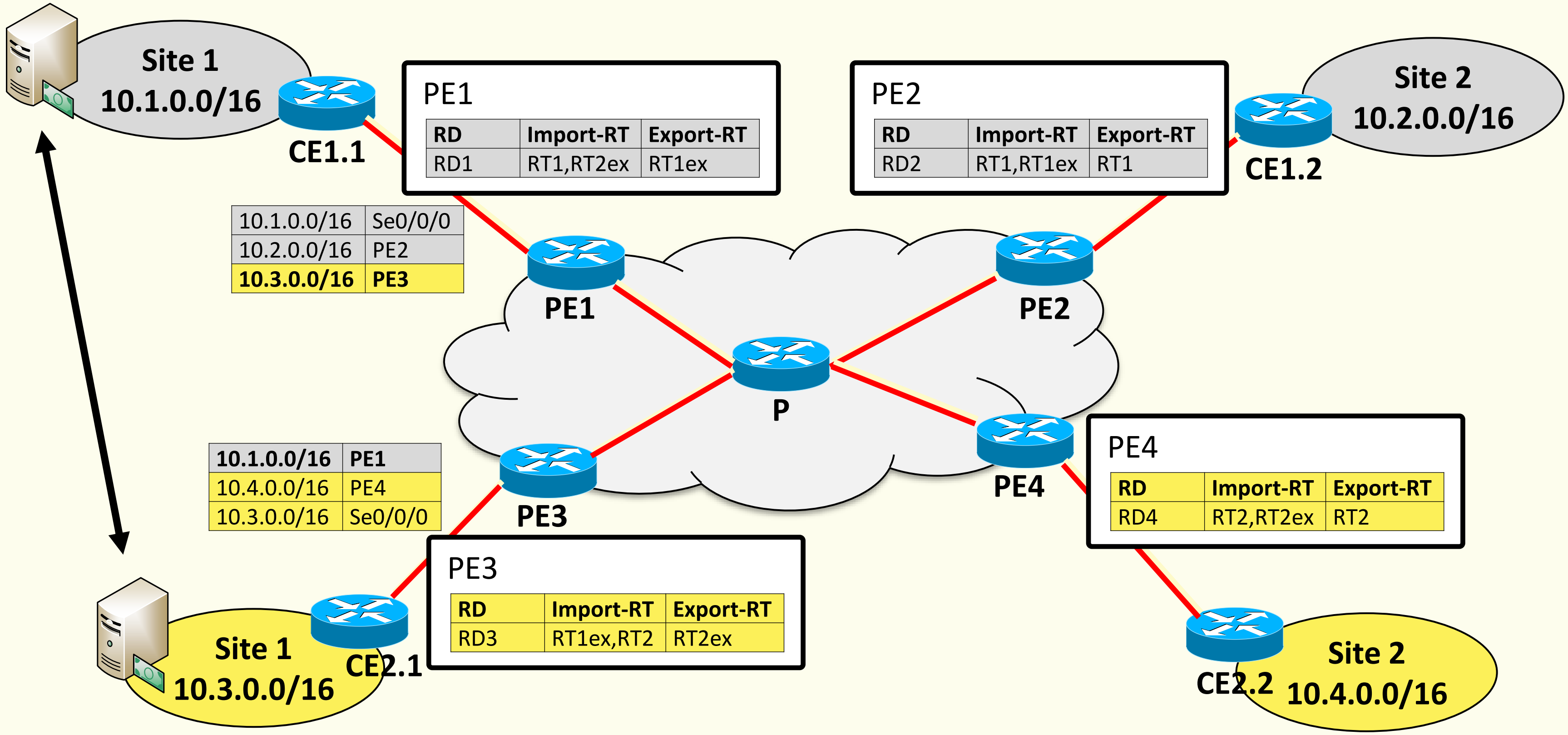
PE3

RD	Import-RT	Export-RT
RD3	RThub	RTspoke

10.1.0.0/16	**PE1**
10.2.0.0/16	**PE1**
10.3.0.0/16	**PE1**
10.4.0.0/16	Se0/0/0

Constrained route distribution

- **Overlapping VPNs** (extranets)



PE1

RD	Import-RT	Export-RT
RD1	RT1,RT2ex	RT1ex

PE2

RD	Import-RT	Export-RT
RD2	RT1,RT1ex	RT1

10.1.0.0/16	Se0/0/0
10.2.0.0/16	PE2
10.3.0.0/16	**PE3**

10.1.0.0/16	**PE1**
10.4.0.0/16	PE4
10.3.0.0/16	Se0/0/0

PE3

RD	Import-RT	Export-RT
RD3	RT1ex,RT2	RT2ex

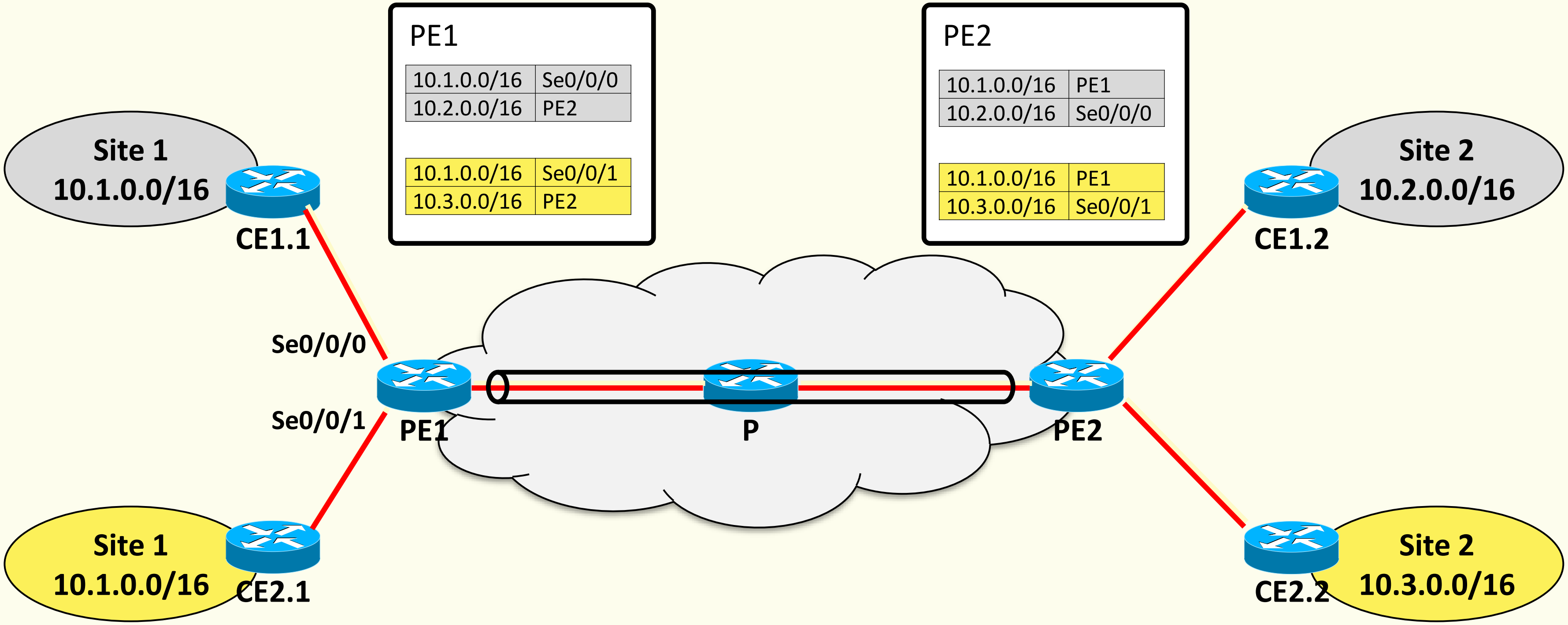
PE4

RD	Import-RT	Export-RT
RD4	RT2,RT2ex	RT2

Provider network forwarding

- The advertising PE is the next hop of a route
  - P has no information on the routes
  - VPN-IP addresses are not routable

Tunneling between PE is necessary



PE1

10.1.0.0/16	Se0/0/0
10.2.0.0/16	PE2

10.1.0.0/16	Se0/0/1
10.3.0.0/16	PE2

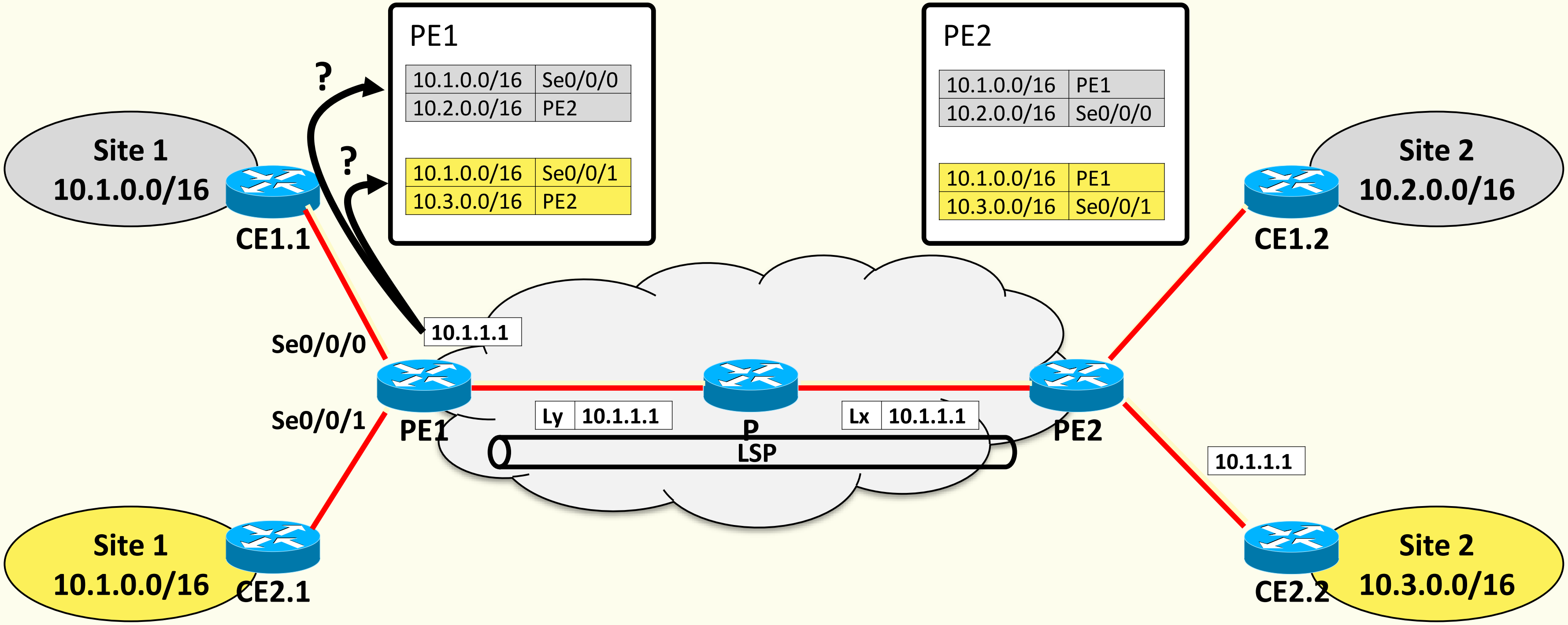
PE2

10.1.0.0/16	PE1
10.2.0.0/16	Se0/0/0

10.1.0.0/16	PE1
10.3.0.0/16	Se0/0/1

Provider network forwarding

- How traffic from a remote PE is demultiplexed?
  - One LSP per VPN is needed between PEs!



PE1

10.1.0.0/16	Se0/0/0
10.2.0.0/16	PE2

10.1.0.0/16	Se0/0/1
10.3.0.0/16	PE2

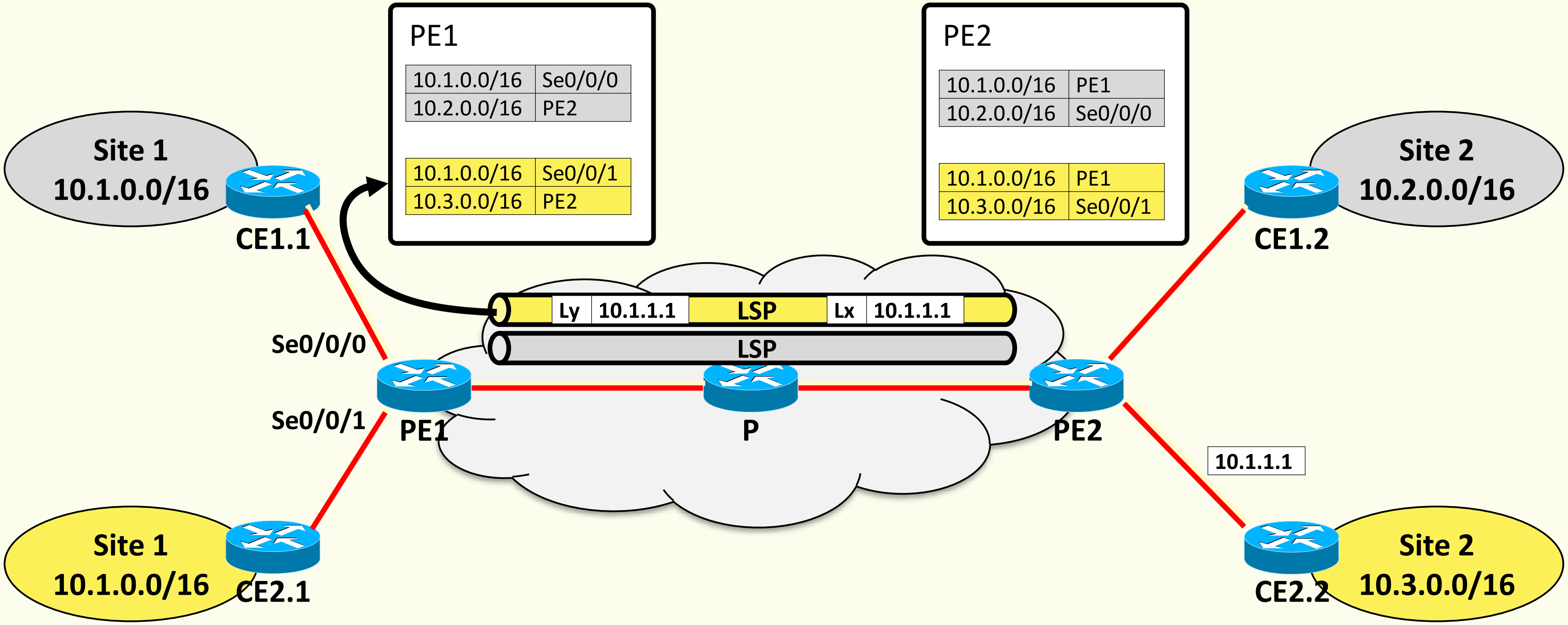
PE2

10.1.0.0/16	PE1
10.2.0.0/16	Se0/0/0

10.1.0.0/16	PE1
10.3.0.0/16	Se0/0/1

Provider network forwarding

- No separate state at P for each PE-PE VPN LSP
- VPN label distribution must be automatic



PE1

10.1.0.0/16	Se0/0/0
10.2.0.0/16	PE2

10.1.0.0/16	Se0/0/1
10.3.0.0/16	PE2

PE2

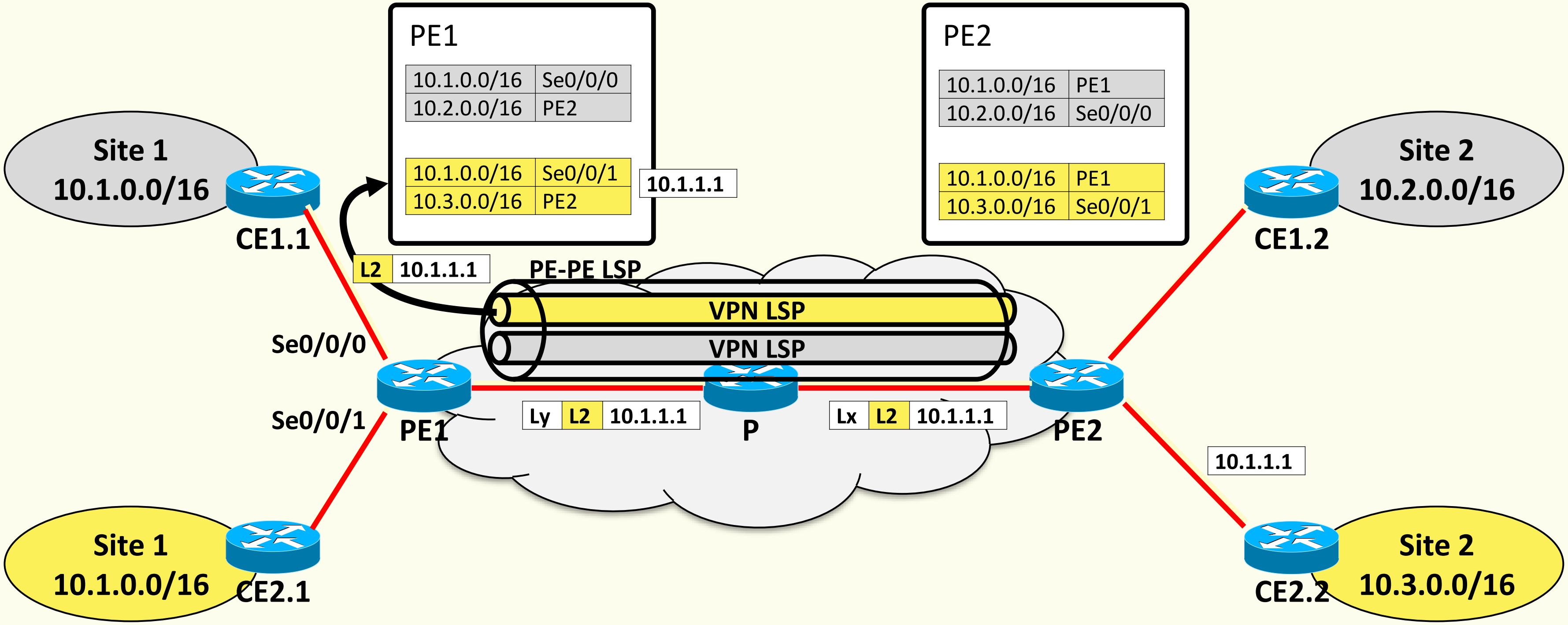
10.1.0.0/16	PE1
10.2.0.0/16	Se0/0/0

10.1.0.0/16	PE1
10.3.0.0/16	Se0/0/1

Provider network forwarding

- No separate state at P for each PE-PE VPN LSP

Use MPLS label stacking



PE1

10.1.0.0/16	Se0/0/0
10.2.0.0/16	PE2

10.1.0.0/16	Se0/0/1
10.3.0.0/16	PE2

PE2

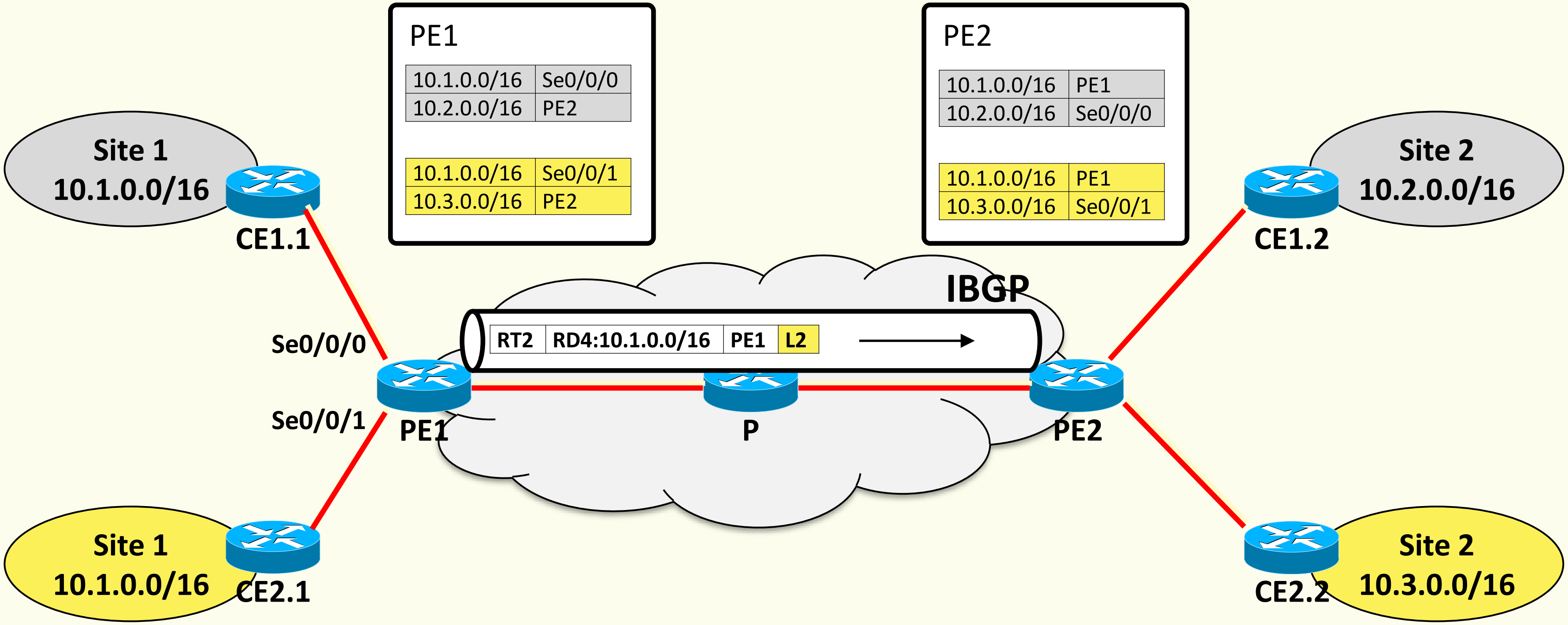
10.1.0.0/16	PE1
10.2.0.0/16	Se0/0/0

10.1.0.0/16	PE1
10.3.0.0/16	Se0/0/1

Provider network forwarding

- VPN label distribution must be automatic

Use MP-iBGP for label distribution



PE1

10.1.0.0/16	Se0/0/0
10.2.0.0/16	PE2

10.1.0.0/16	Se0/0/1
10.3.0.0/16	PE2

PE2

10.1.0.0/16	PE1
10.2.0.0/16	Se0/0/0

10.1.0.0/16	PE1
10.3.0.0/16	Se0/0/1

Benefits of BGP/MPLS IP VPNs

- **Customer**
  - Offload routing management to the provider
  - Access added-value services (firewall, auth)
- **Provider**
  - Service multiple VPN customers with a common infrastructure
  - VPN management is hidden to the core
  - Scale by adding PEs when needed
- MPLS tunneling plays a key **enabling** role

References

- I. Minei and J. Lucek, **MPLS-Enabled Applications: Emerging Developments and New Technologies**, 3rd Edition, Wiley, Dec. 2010
- RFCs
  - **RFC4364**, BGP MPLS IP Virtual Private Networks (VPNs), Feb. 2006
  - **RFC4760**, Multiprotocol Extensions for BGP-4, Jan. 2007