



TIAGORA.COM

# Fundamentos de MPLS/VPN

**Rogelio Alvez**

ralvez@tiagora.com

**CERTuy**

 **AGESIC**

  
PRESIDENCIA  
REPUBLICA ARGENTINA



**CERTuy**

 **AGESIC**



# MPLS/VPN: Conceptos generales



# Qué es una VPN ?

Es una red privada en términos lógicos, montada sobre un medio potencialmente compartido

Un conjunto de sitios a los que les es permitido comunicarse mutuamente

Es el “ámbito alcanzable” por una tabla de rutas

# VPN basada en la idea de “peers”

- Router de borde del backbone y router del “cliente” usan el mismo protocolo de red para dialogar
- Routers de cliente (CE) y routers de backbone (PE) arman una adyacencia en los términos del protocolo común
- Los routers del backbone conocen la información de direccionamiento de los routers de “cliente”

# MPLS-VPN

- Su base es el modelo de peers, pero
- PEs reciben y mantienen información de rutas de las VPNs directamente conectadas
- Reduce la cantidad de información que tiene que almacenar un PE
- Se usa MPLS para rutear en el backbone (no se necesita conocer información del cliente)

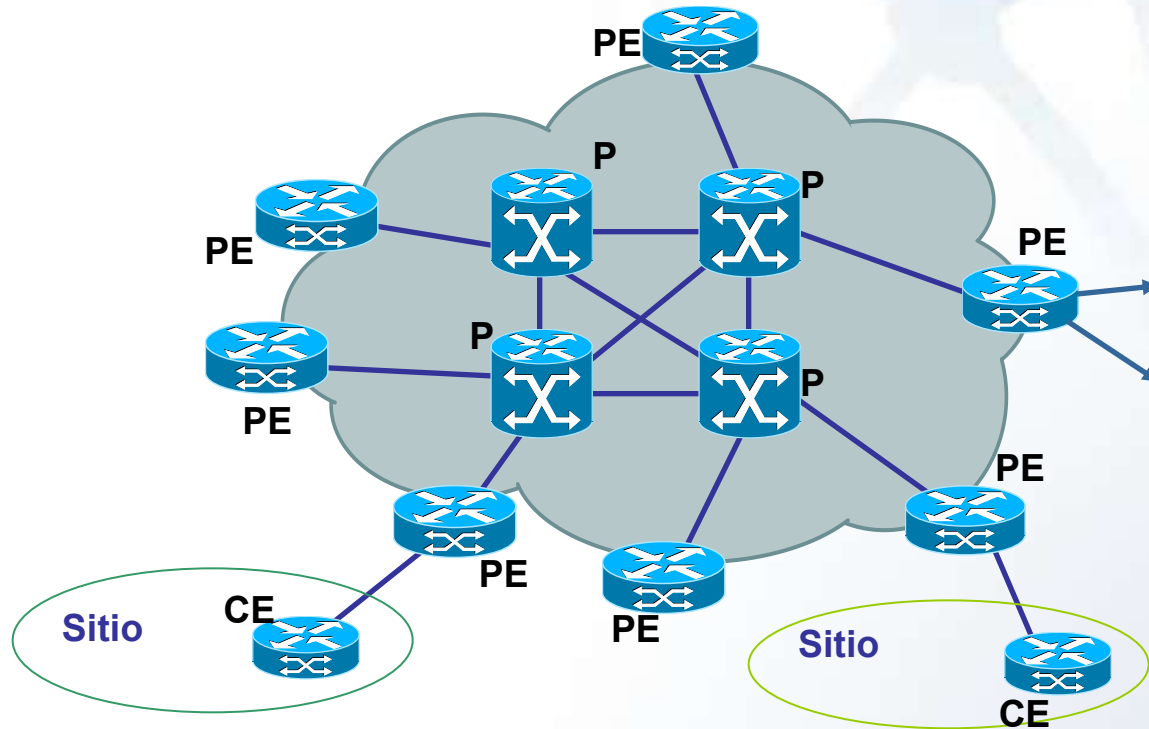
# MPLS-VPN: Terminología

- Red de Proveedor (Red P)
  - Backbone controlado por un “Proveedor MPLS”
- Red de Cliente (Red C)
  - Red bajo el control del cliente
- Router CE
  - Customer Edge. Parte de la Red C, que hace interfaz con la Red P

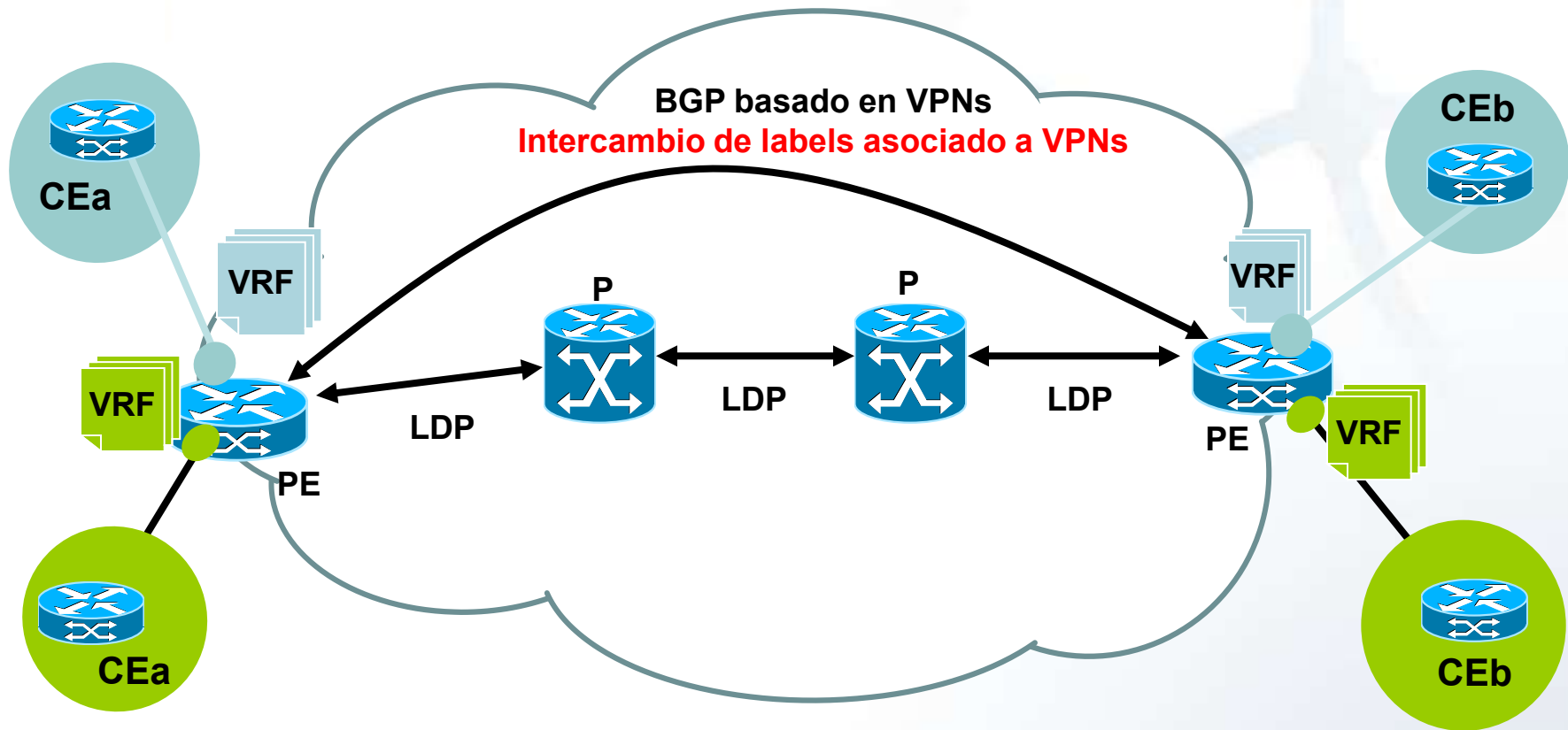
# MPLS-VPN: Terminología

- Sitio
  - Conjunto de redes de la Red C, ubicadas en el ámbito de un PE
  - Un sitio se conecta al backbone MPLS a través de uno o más enlaces PE/CE
- Router PE
  - Provider Edge router. Parte de la Red P; hace interfaz con los routers CE
- Router P
  - Provider (core) router; no tiene conocimientos de las VPNs

# Componentes de una red MPLS



# Plano de control MPLS/VPN

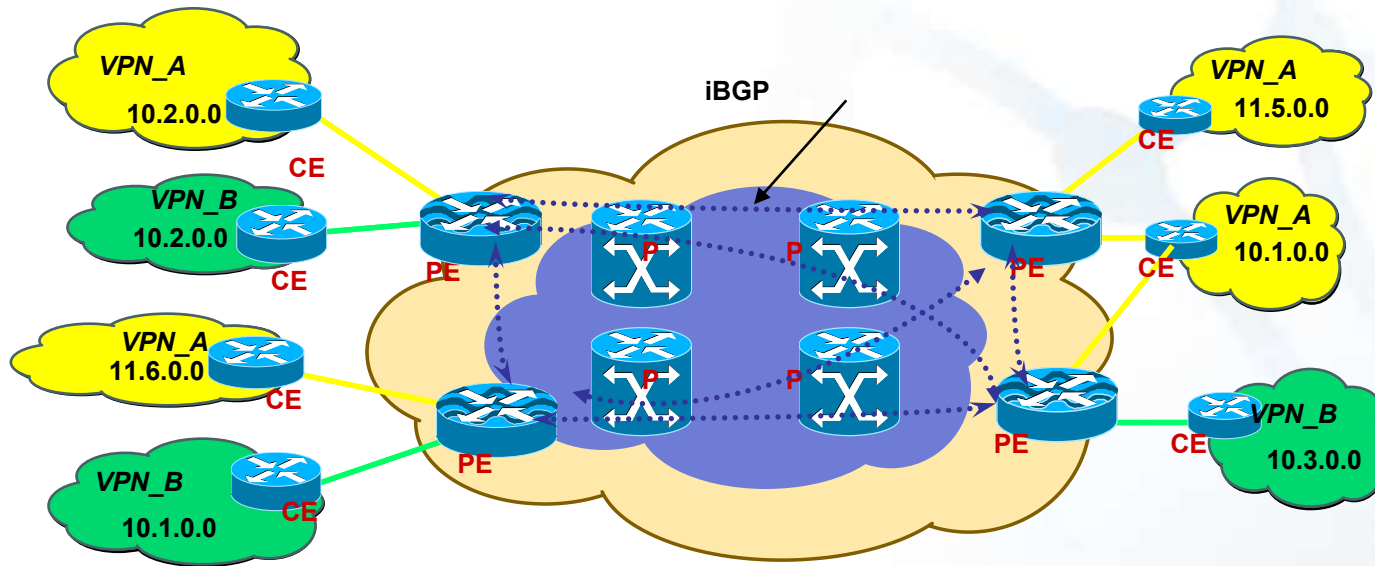


1. Activación de VPNs (tantas VRFs por site como VPNs distintas)
2. CEs “enseñan” sus rutas a cada PE, que las ubica en la VRF correspondiente
3. PE incorpora información de cada VPN al proceso BGP
4. PE asigna una etiqueta diferente a cada VPN, e informa lo que conoce a otros PEs
5. PE receptor distribuye adecuadamente lo recibido de otro PE a cada CE, según corresponda a la VPN de destino

# MPLS-VPN: Terminología

- VRF (Virtual Routing and Forwarding)
  - Tabla de rutas más su tabla de forwarding
  - Alimentada por “procesos de routing contextualizados”
- VPN-Aware network
  - El backbone de un proveedor, en el cual se ha implementado MPLS-VPN

# MPLS VPN: Modelo



- P (LSR) pertenece al núcleo del backbone MPLS
- PEs (LSRs) utilizan MPLS dentro del backbone
- PEs intercambian información de VPNs entre sí, luego de interactuar con los routers CEs de los diferentes sitios



**CERTuy**

 **AGESIC**



# MPLS/VPN: Modelo de conexión

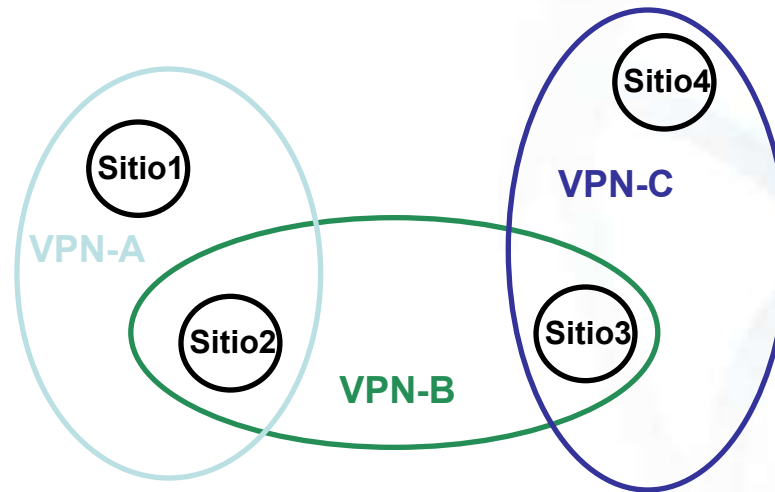


CCIE, CCSP y CSCI son marcas registradas de Cisco Systems, Inc

# MPLS VPN: Modelo de conexión

- Una VPN es un conjunto de sitios que comparten información de ruteo
- Un sitio puede ser parte de diferentes VPNs
- Una VPN hay que pensarla como una “comunidad de interés”
- Varias VRFs (Routing/Forwarding instances) presentes en los routers PE

# MPLS VPN: Modelo de conexión

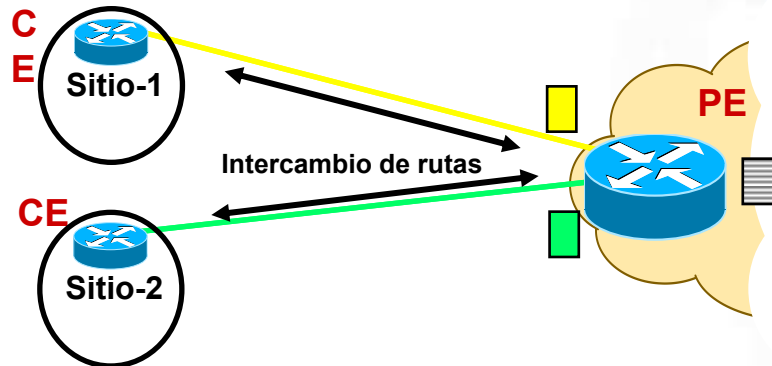


- Un sitio que pertenece a varias VPNs podría (o no) ser usado como tránsito entre las VPNs
- Si dos o más VPNs tienen un sitio en común, el espacio de direcciones debe ser único entre estas VPNs

# MPLS VPN: Modelo de conexión

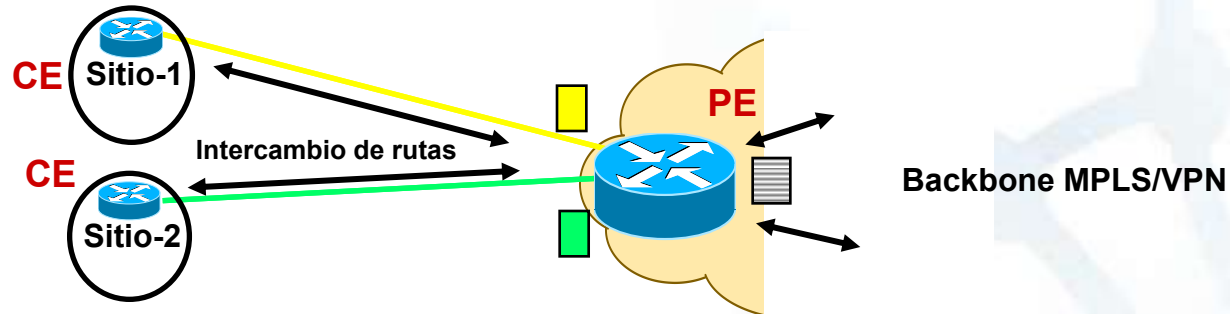
- Backbone VPN = conjunto de LSRs MPLS
  - PE routers (edge LSRs)
  - P routers (core LSRs)
- PEs hablan con CEs y propagan información de VPN con MP-BGP a otros PEs
  - Direcciones VPN-IPv4, Extended Community, Label
- **Routers P no corren BGP ni conocen VPNs**

# MPLS VPN: Modelo



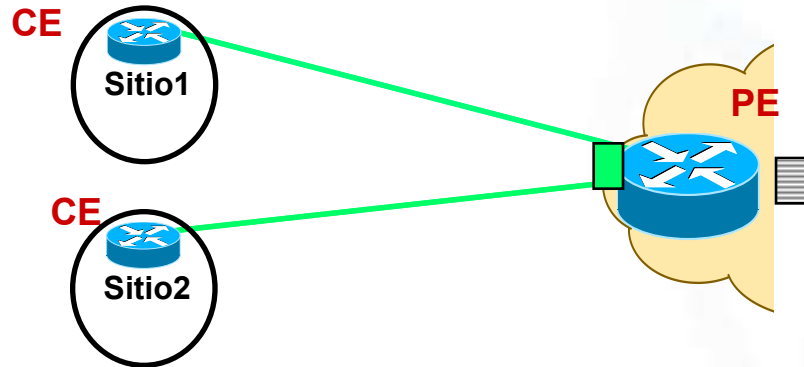
- PEs y CEs intercambian rutas
- CEs corren un software IP tradicional

# MPLS VPN: Modelo



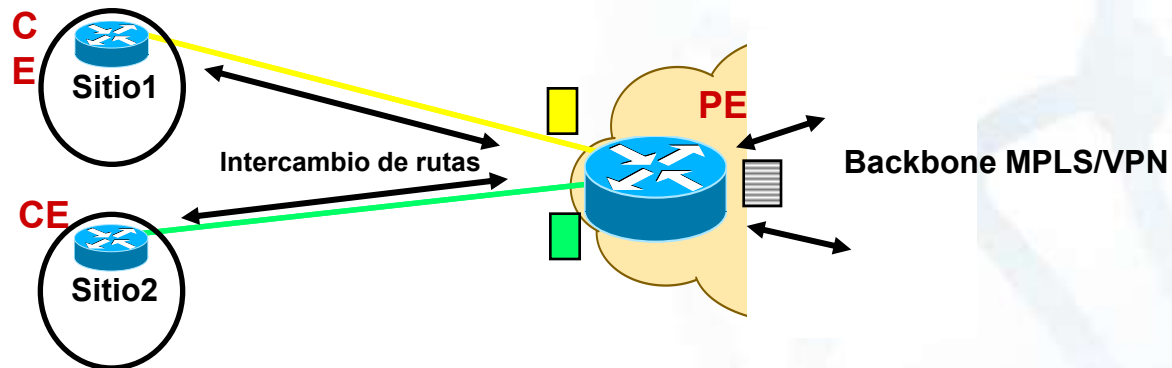
- PEs mantienen tablas de rutas separadas
  - Tabla de rutas global
    - Contiene a todos los P y PEs
    - Contiene el “mapa” del backbone
  - VRF (VPN Routing and Forwarding)
    - Tabla de Routing y Forwarding asociada a uno o más sitios directamente conectados (CEs)
    - VRF son asociadas con interfaces conectadas a los CEs
    - Interfaces pueden compartir la misma VRF si los sitios conectados pueden compartir la misma información de rutas

# MPLS VPN: Modelo



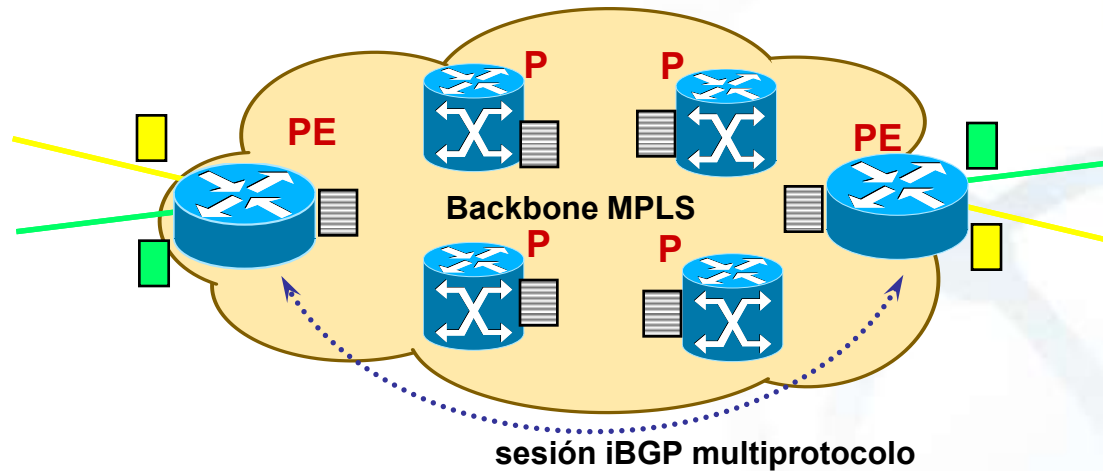
- Sitios diferentes que comparten la misma información, pueden compartir la misma VRF
- Las interfaces que conectan a estos sitios usarán la misma VRF
- Sitios que pertenecen a la misma VPN “podrían” usar la misma VRF

# MPLS VPN: Modelo



- Las rutas que el PE recibe de los CEs las ubica en una VRF apropiada
- Las rutas que el PE recibe desde el IGP del backbone las pone en la tabla IP “tradicional”
- Las direcciones de las VPNs no tienen que ser mutuamente exclusivas, ya que son ubicadas en diferentes VPNs

# MPLS VPN: Modelo



- PEs y Ps comparten un IGP (OSPF, ISIS, EIGRP)
- PEs arman sesiones MP-iBGP entre ellos
- PEs usan MP-BGP para informar sobre sitios y VPNs conectadas
  - Direcciones VPN-IPv4, Extended Community, Label

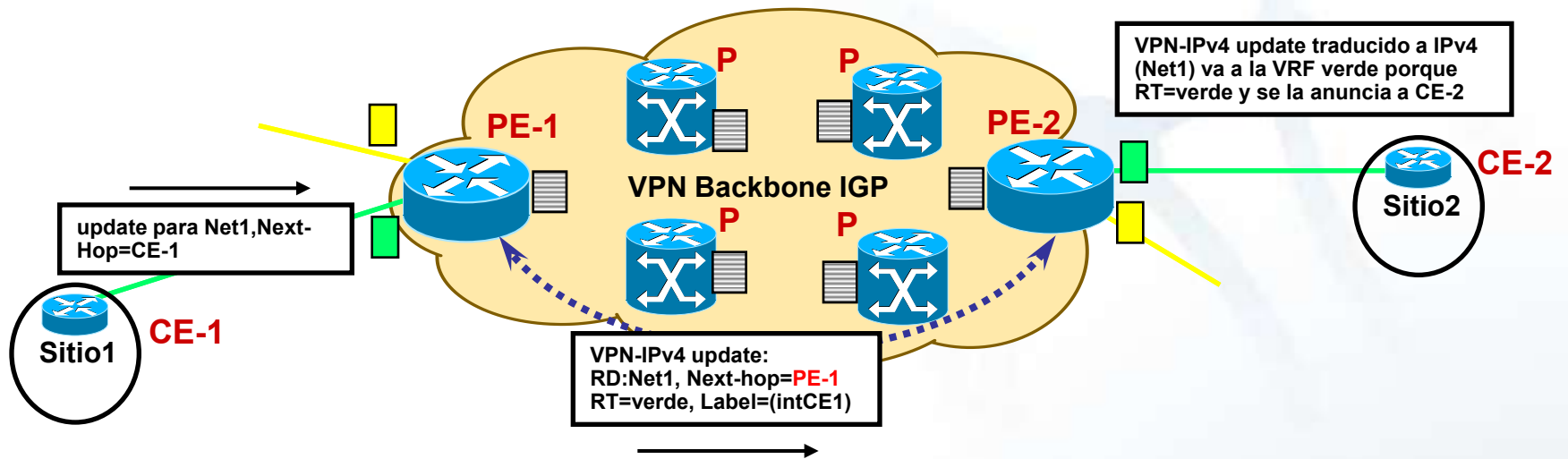
# MPLS VPN: anuncio MP-BGP

- Dirección VPN-IPV4: es la unión de las siguientes estructuras de datos
  - 1- Route Distinguisher
    - 64 bits
    - **Le da unicidad global al prefijo IPv4**
    - RD es configurado en el PE para cada VRF
    - RD puede estar (o no) relacionado a un sitio o una VPN
  - 2 - Direcciones IPv4 (32bits)
- Route Target (dato de 64 bits)
  - Identifica al conjunto de sitios a los que la ruta les debe llegar anunciada

# MPLS VPN: anuncio MP-BGP

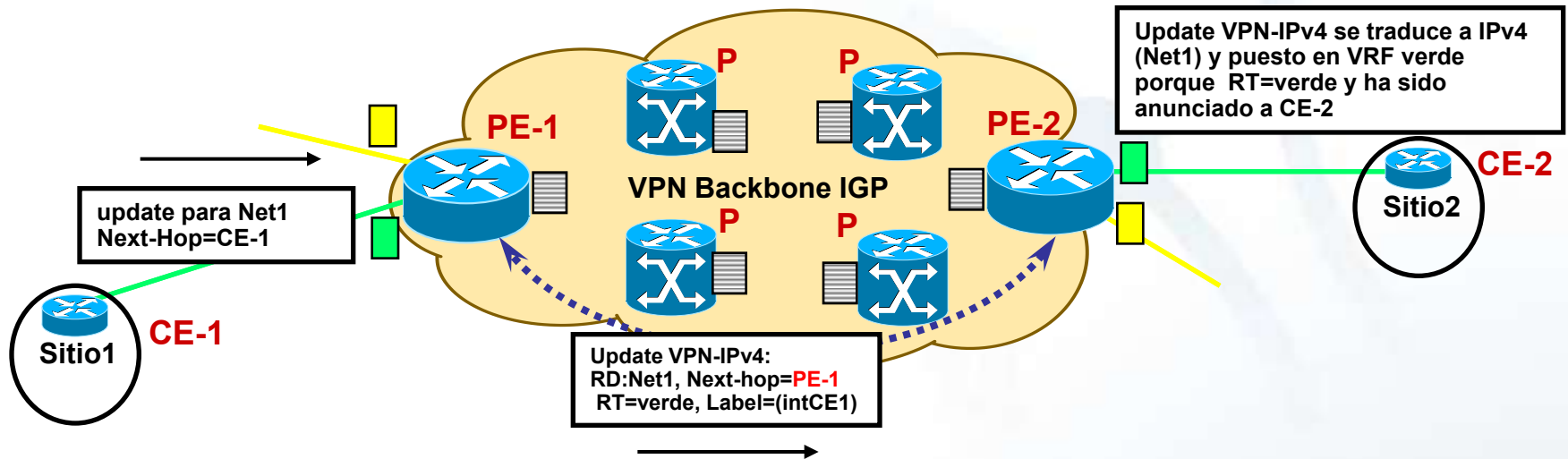
- Otros atributos BGP tradicionales
  - Local Preference
  - MED
  - Next-hop
  - AS\_PATH
- Un Label, que identifica
  - La VRF a la cual pertenece el anuncio (label de la VPN)

# MPLS VPN: Modelo



- PEs reciben updates IPv4s
- PEs traducen el update a formato VPN-IPv4
  - Asignan un SOO y RT según lo configurado
  - Re-escriben el atributo Next-Hop
  - Asignan un label basándose en la VRF y/o la interfaz
  - Envían anuncio MP-iBGP a sus vecinos PE

# MPLS VPN: Modelo



- PEs receptores traducen nuevamente a IPv4
  - Insertan la ruta en la VRF identificada por el atributo RT (de acuerdo a cómo haya sido configurado el PE)
- El label asociado a la dirección VPN-IPv4 será impuesto en los paquetes enviados hacia su respectivo destino

# MPLS VPN: Modelo

- Distribución de rutas a los sitios se basa en el atributo llamado “route-target”
  - Es una “Community” (atributo) del protocolo BGP
  - En el sitio receptor, se instalarán rutas en la VRF que “reclame” rutas que coincidan con atributo “route-target” de los anuncios que llegan desde otros extremos del backbone
  - Controlado a través de configuración en el PE
- Un PE que conecta sites de VPN distintas instala la ruta en la VRF del sitio si el atributo “route-target” contiene una o más VPNs al cual el sitio está asociado



**CERTuy**

 **AGESIC**



# MPLS/VPN: Forwarding



CCIE, CCSP y CSCI son marcas registradas de Cisco Systems, Inc

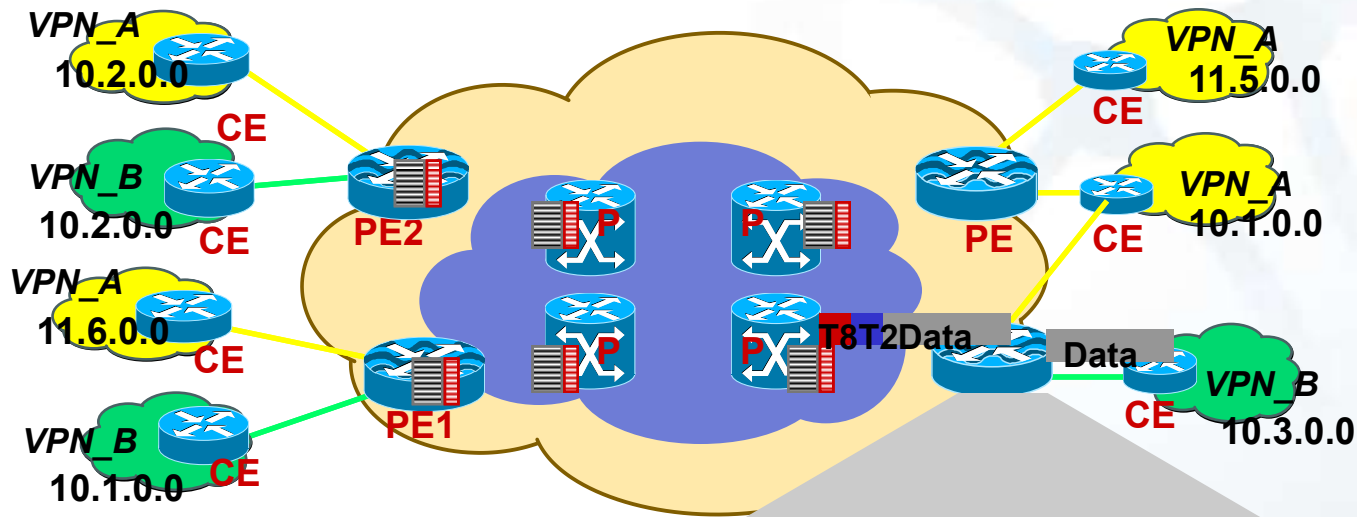
# MPLS: Forwarding de paquetes

- Los valores “next-hop” anunciados por el BGP de los PEs deben ser visibles por el IGP
- Labels se distribuyen con LDP (hop-by-hop), con el fin de lograr llegar a los Next-Hops de BGP
- Stack de labels para el forwarding de paquetes
  - Top label indica el next-hop BGP (label interior)
  - label de 2do nivel indica la interfaz saliente o VRF (label exterior)

## Forwarding de paquetes

- Los nodos MPLS conmutan con el label externo
- **P no saben nada de BGP ni de VPNs**

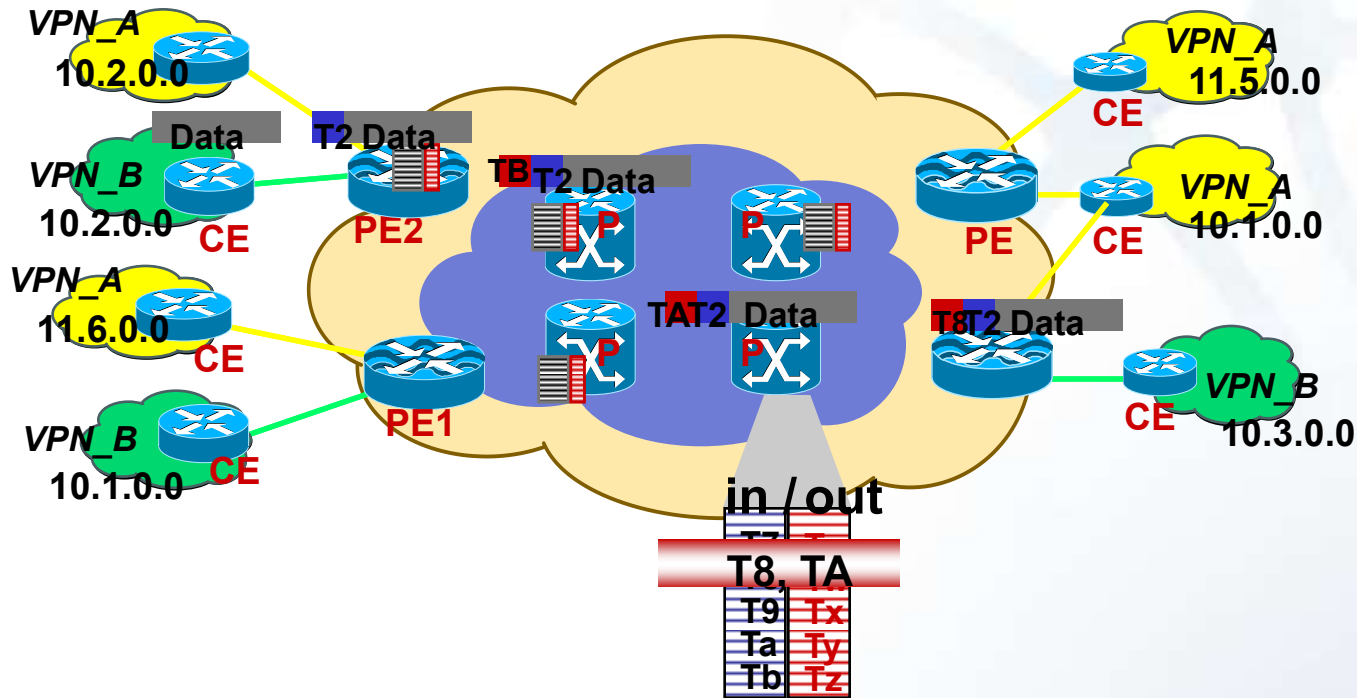
# Forwarding



<RD_B,10.2> , iBGP NH= PE2 , T2	T8
<RD_B,10.2> , iBGP next hop PE2	T2 T8
<RD_B,10.3> , iBGP next hop PE3	T3 T9
<RD_A,11.6> , iBGP next hop PE1	T4 T7
<RD_A,10.1> , iBGP next hop PE4	T5 TB
<RD_A,10.4> , iBGP next hop PE4	TB
<RD_A,10.2> , iBGP next hop PE2	T8

- PE de ingreso recibe paquetes IP
- PE hace "IP Longest Match" en **VPN\_B FIB** , deduce el next hop **PE2** y agrega el stack de labels:  
Label exterior **T2** + Label interior **T8**

# Forwarding

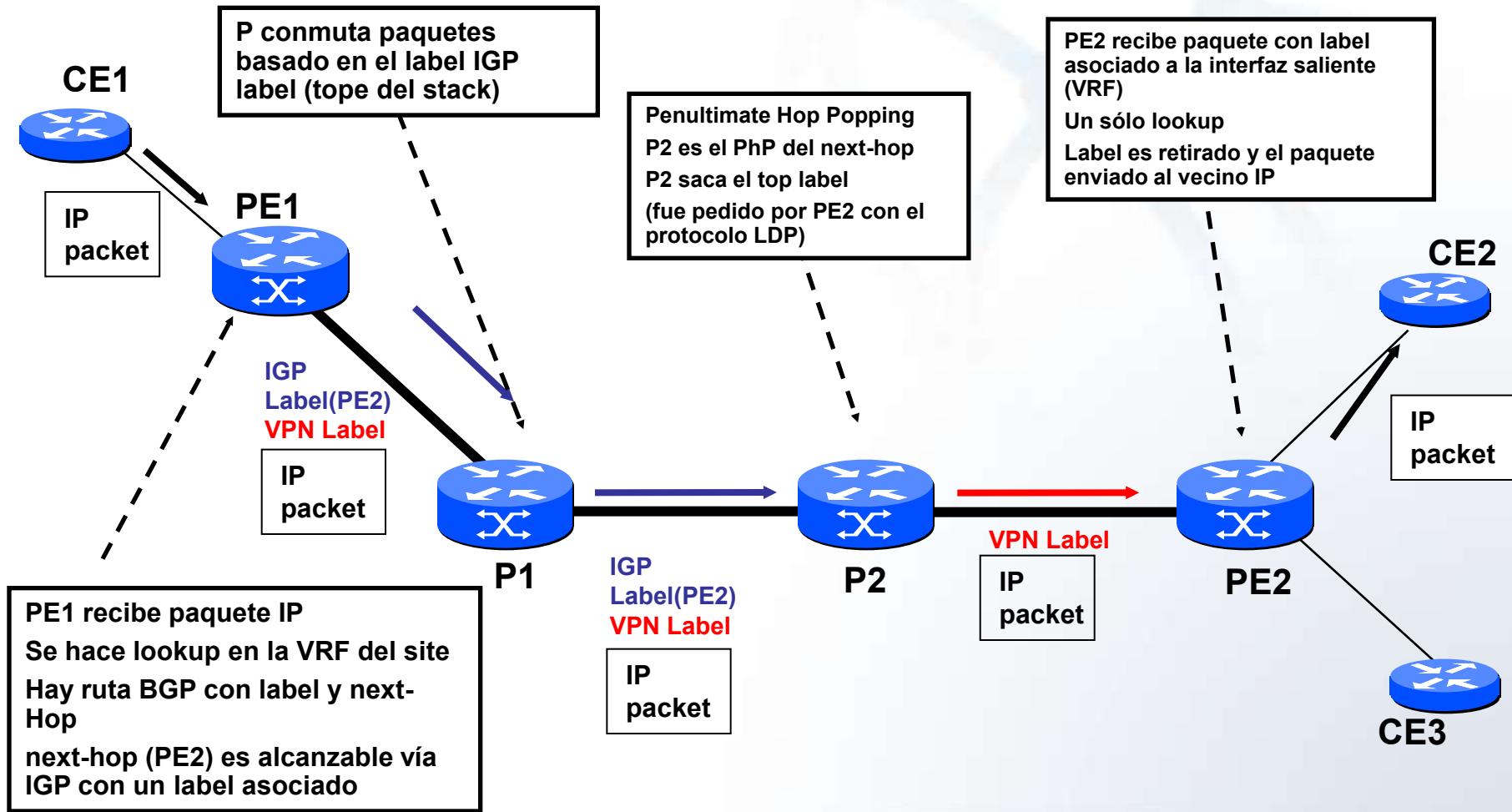


- Los P rutean usando el label interior
- El PE de egreso quita el label interior
- El PE de egreso usa el label exterior para decidir la VPN a la cual entregar el paquete.
- Se saca el label exterior y se envía el paquete al router CE

# Penultimate Hop Popping

- El router upstream LDP del next-hop de BGP (PE) descartará la etiqueta externa
  - Penultimate hop popping
- Es solicitado con el protocolo LDP
- El PE de egreso conmutará el paquete en base a la info de la etiqueta de 2do nivel (que explica la interfaz y/o la VPN de salida)

# Forwarding y Penultimate Hop Popping





**CERTuy**

 **AGESIC**



# MPLS/VPN: Condiciones de base para su funcionamiento



CCIE, CCSP y CSCI son marcas registradas de Cisco Systems, Inc

# Mecanismos: VRF y múltiples instancias

- Equipos con la capacidad de VRFs:
  - VRF Routing Protocol Context
  - VRF Routing Tables
  - VRF CEF Forwarding Tables

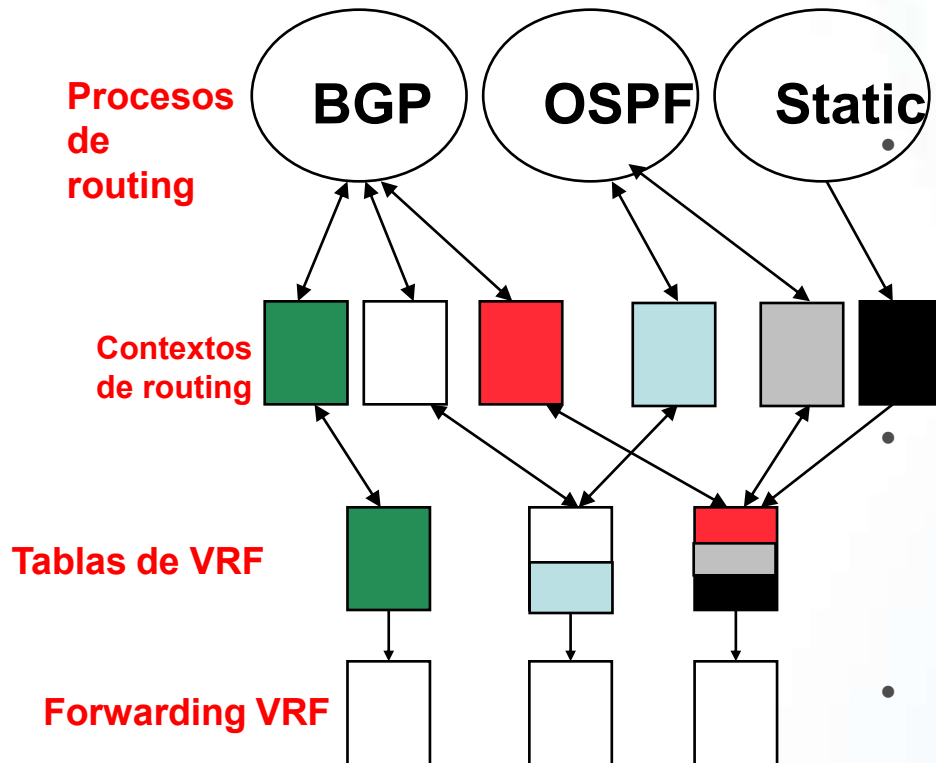
# Mecanismos: VRF y múltiples instancias

- Protocolos de ruteo “preparados para VRFs”
- Selección e instalación de rutas en la tabla apropiada de cada VRF o de cada VPN
- Los protocolos disponibles dependerán del fabricante

# Mecanismos: VRF y múltiples instancias

- La tabla VRF contiene rutas que deberían estar disponibles para un conjunto de sitios de una misma VPN
- Análoga a una tabla de rutas tradicional
- Interfaces que se conectan a routers de los sitios son asignados a VRFs
  - Una VRF por interfaz
  - Posiblemente, muchas interfaces para una VRF

# Mecanismos: VRF y múltiples instancias

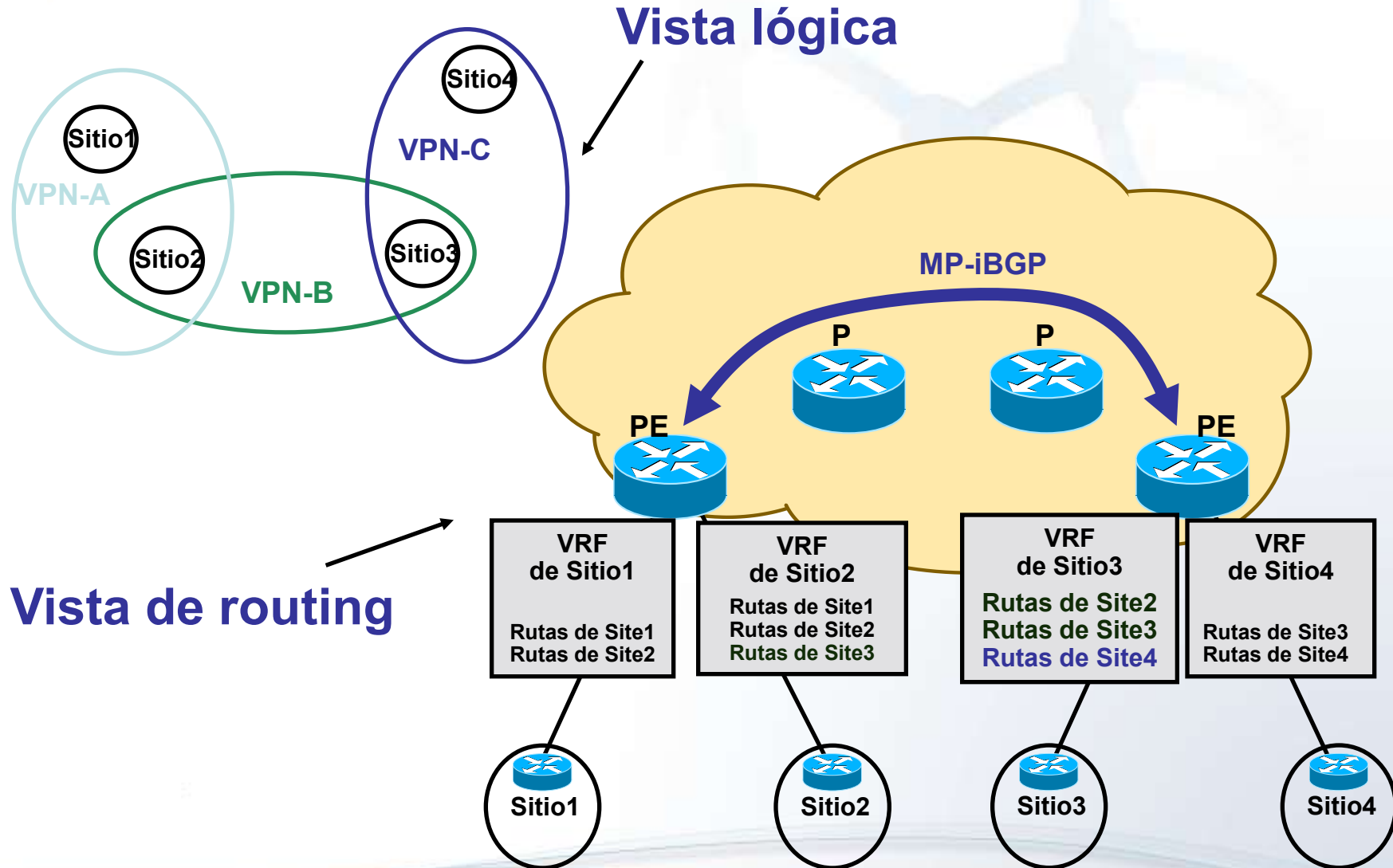


Procesos de routing corren en “contextos” específicos

- Pueblan tablas y FIBs específicas para c/VPN (VRF)

- Interfaces se asignan a VRFs

# Mecanismos: VRF y múltiples instancias





**CERTuy**

 **AGESIC**

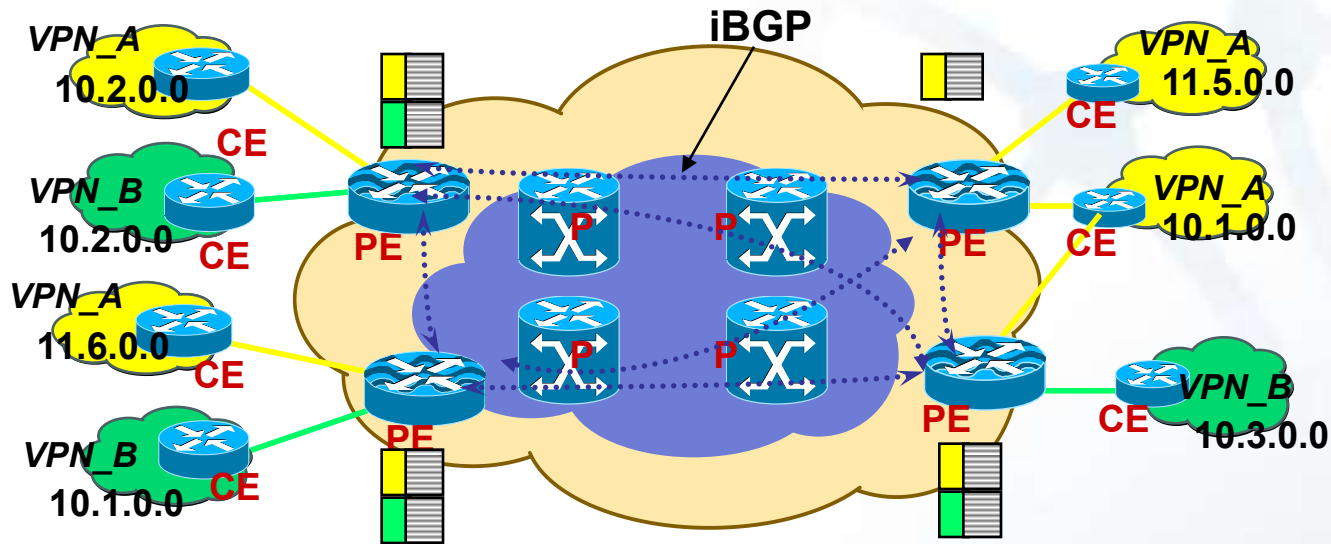


# MPLS/VPN: Topologías



CCIE, CCSP y CSCI son marcas registradas de Cisco Systems, Inc

# Topologías

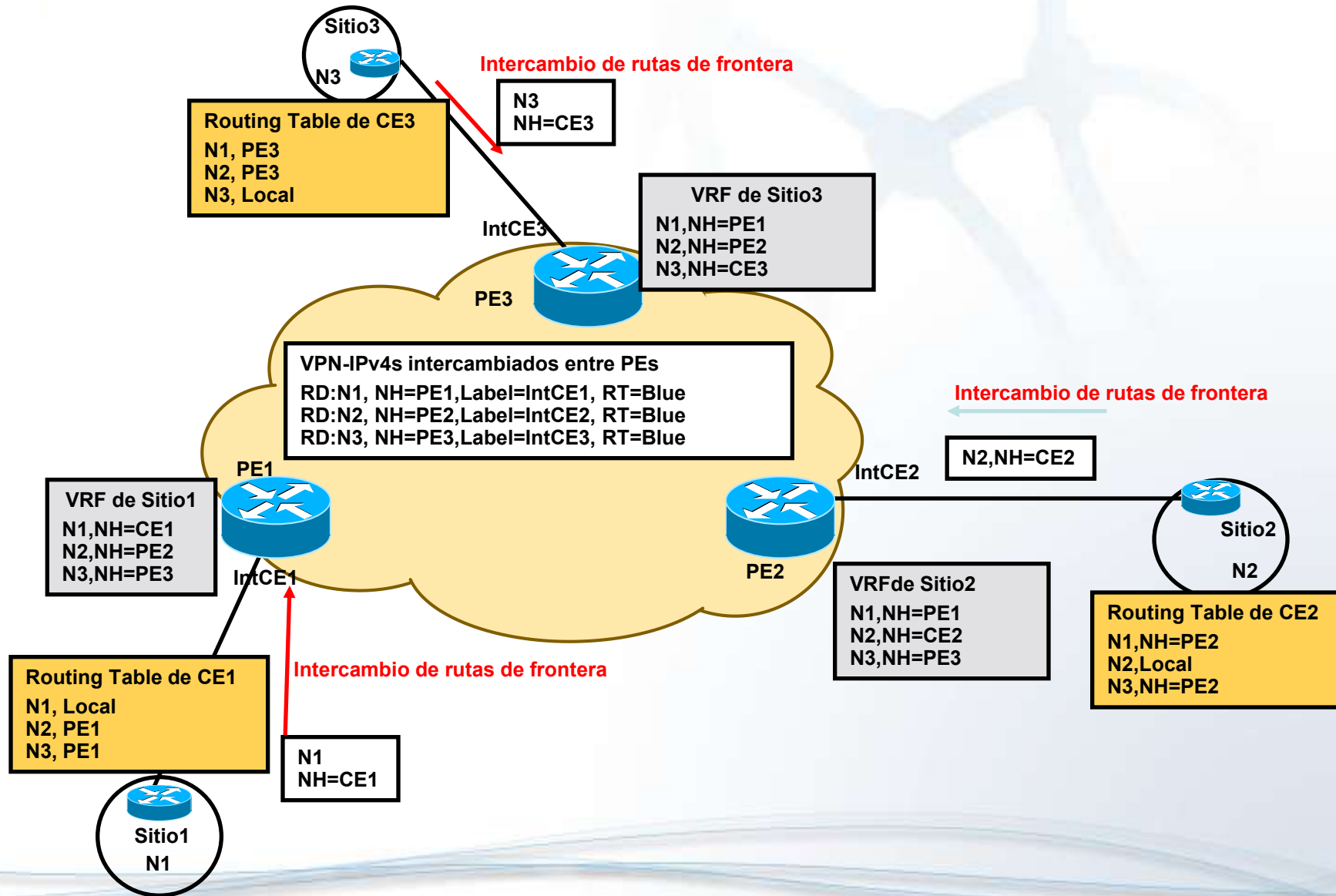


- Direcciones VPN-IPv4 se propagan junto al label asociado como una extensión “multiprocol BGP”
- Extended Community (route-target) se asocia a cada dirección VPN-IPv4, para poblar la VRF del sitio

# VPN Full Mesh

- Cada sitio conoce a todo otro sitio (de la misma VPN)
- Cada CE anuncia su propio espacio de IP
- Anuncios MP-BGP VPN-IPv4 entre PEs
- Ruteo es óptimo dentro del backbone
  - Cada ruta tiene el next-hop de BGP más cercano hacia el destino
- No se usa a ningún sitio como punto central de conectividad

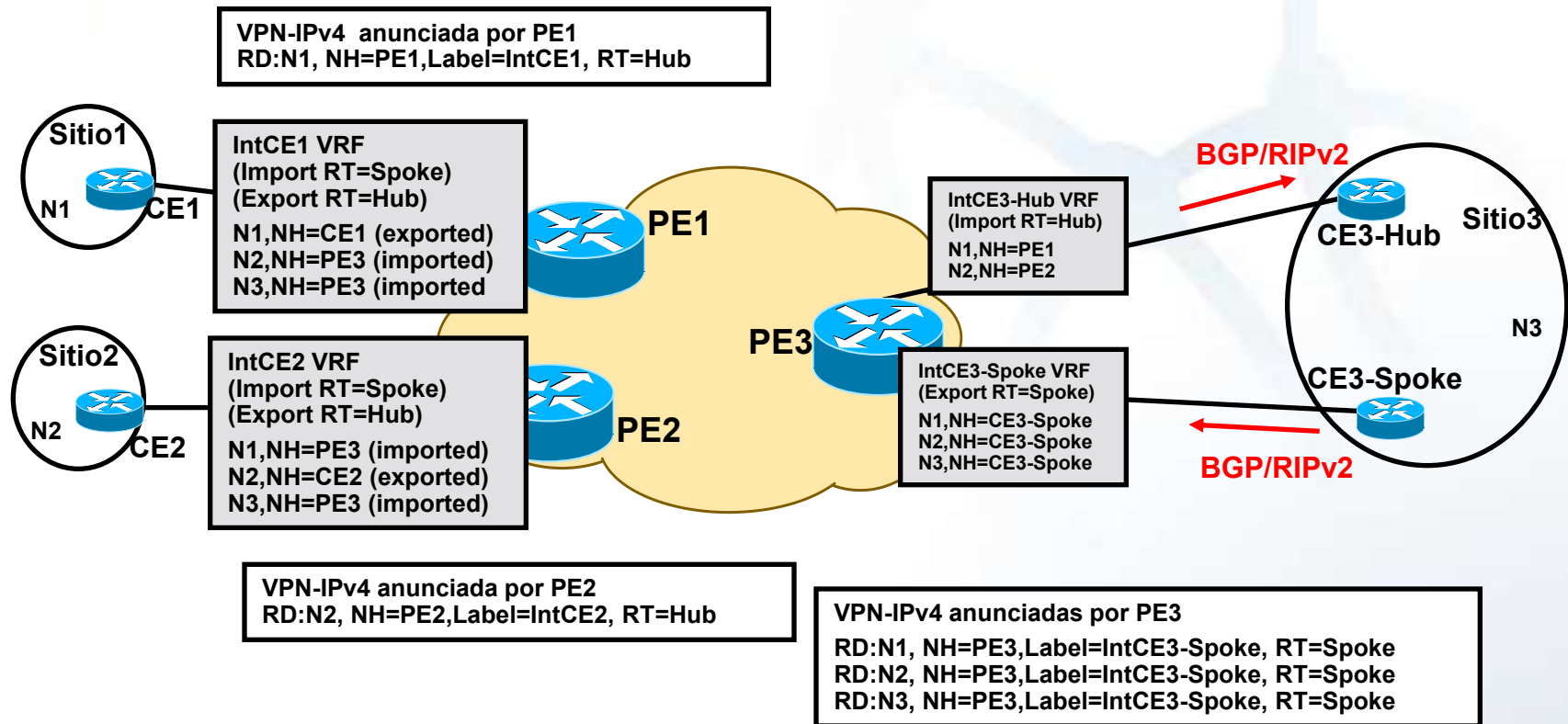
# VPN Full Mesh



# VPN Hub & Spoke

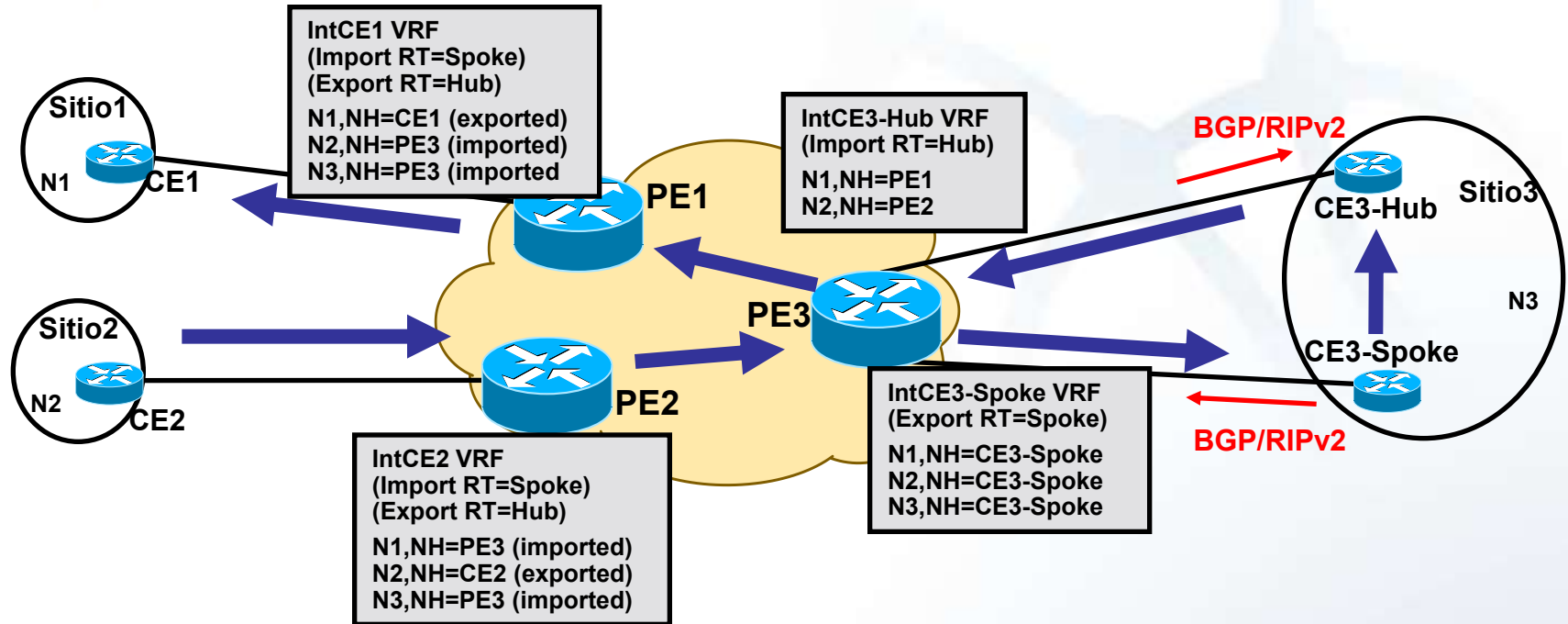
- Un sitio central conoce al resto (de la VPN)
  - Hub-Site
- El resto de los sitios habla solamente con el hub
  - Spoke-Sites
- **Hub-Site es el punto de tránsito entre los Spoke-Sites**

# VPN Hub & Spoke



- Rutas son importadas/exportadas en las VRFs según valor RT de los anuncios VPN-IPv4
- PE3 usa 2 (sub)interfaces con dos VRFs diferentes

# VPN Hub & Spoke



- Tráfico entre spokes circula por el hub



**CERTuy**

 **AGESIC**



# MPLS/VPN: Configuración en plataformas Cisco



CCIE, CCSP y CSCI son marcas registradas de Cisco Systems, Inc

# Configuración

- Conocimiento de VPNs en todos los PEs
- Al PE hay que configurarle
  - VRF y Route Distinguisher
  - VRF import/export policies (basadas en Route-target)
  - Protocolos para hablar con los CEs
  - MP-BGP entre PEs
  - BGP para rutas de Internet
    - Con otros PEs
    - Con algunos CEs

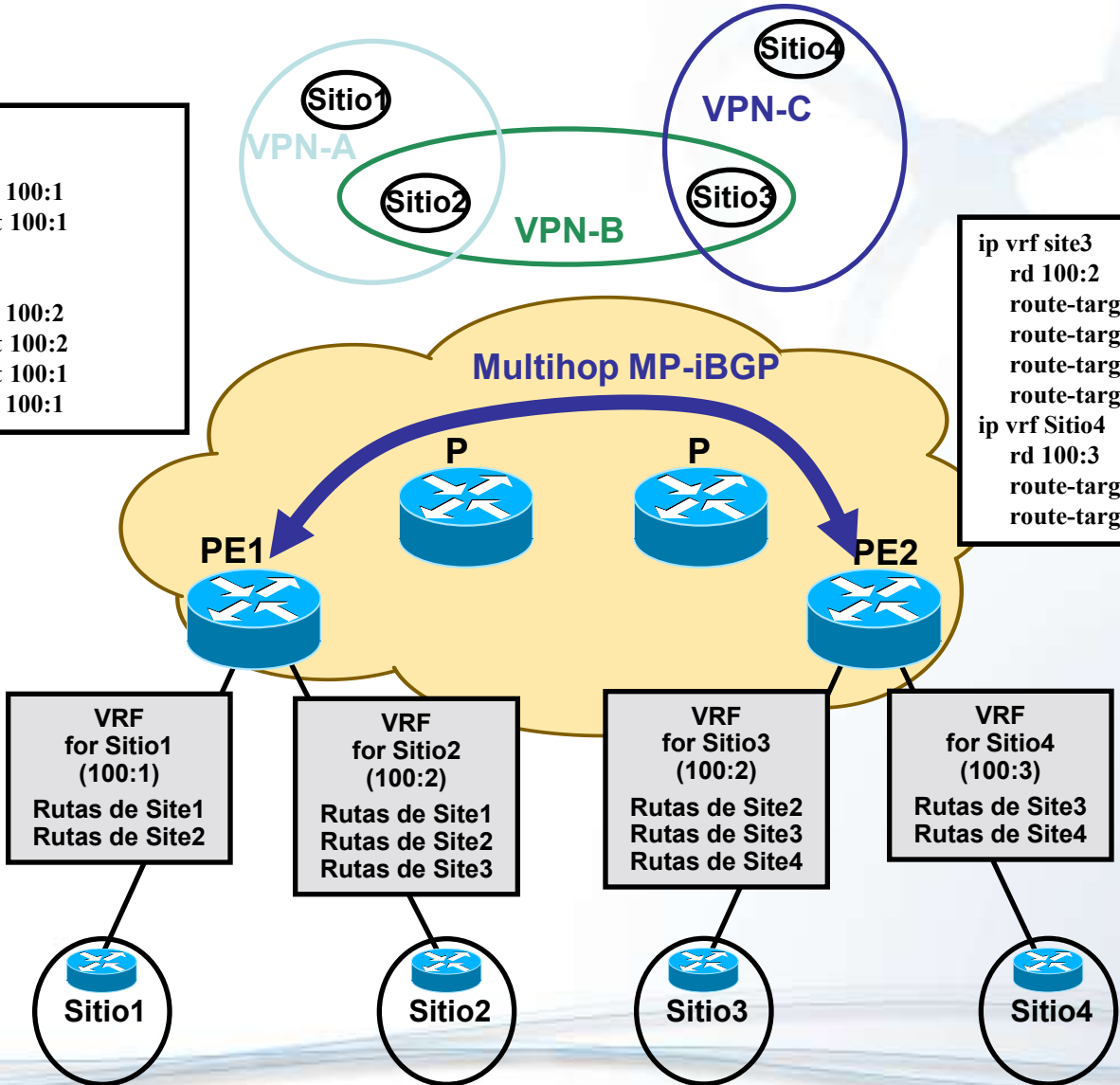
# Configuración: VRF y Route Distinguisher

- RD es configurado en los PE (para cada VRF)
- VRFs se asocian a los RDs en cada PE
- Consejo: mismo RD para misma VPN en todos los PEs
  - Pero no es obligatorio
- VRF configuration command
  - `ip vrf <vrf-symbolic-name>`
    - `rd <route-distinguisher-value>`
    - `route-target import <Import route-target community>`
    - `route-target export <Import route-target community>`

# CLI – configuración de VRFs

```
ip vrf site1
rd 100:1
route-target export 100:1
route-target import 100:1
ip vrf site2
rd 100:2
route-target export 100:2
route-target import 100:2
route-target import 100:1
route-target export 100:1
```

```
ip vrf site3
rd 100:2
route-target export 100:2
route-target import 100:2
route-target import 100:3
route-target export 100:3
ip vrf Sitio4
rd 100:3
route-target export 100:3
route-target import 100:3
```



# Configuración (PE/CE)

- PE/CE: casi cualquier protocolo de routing
- Se usa un “routing context” en cada VRF
- “Routing contexts” son definidos dentro de la instancia de un protocolo de routing
  - ```
router rip
version 2
address-family ipv4 vrf <vrf-symbolic-name>
...
    cualquier comando que aplique a esta instancia
...

```

# Configuración: PE/CE

- BGP usa el mismo comando “address-family”
  - router BGP <asn>  
...  
address-family ipv4 vrf <vrf-symbolic-name>  
...  
cualquier subcomando aplicable de BGP  
...
- Rutas estáticas: se configuran por cada VRF
  - ip route vrf <vrf-symbolic-name> ...

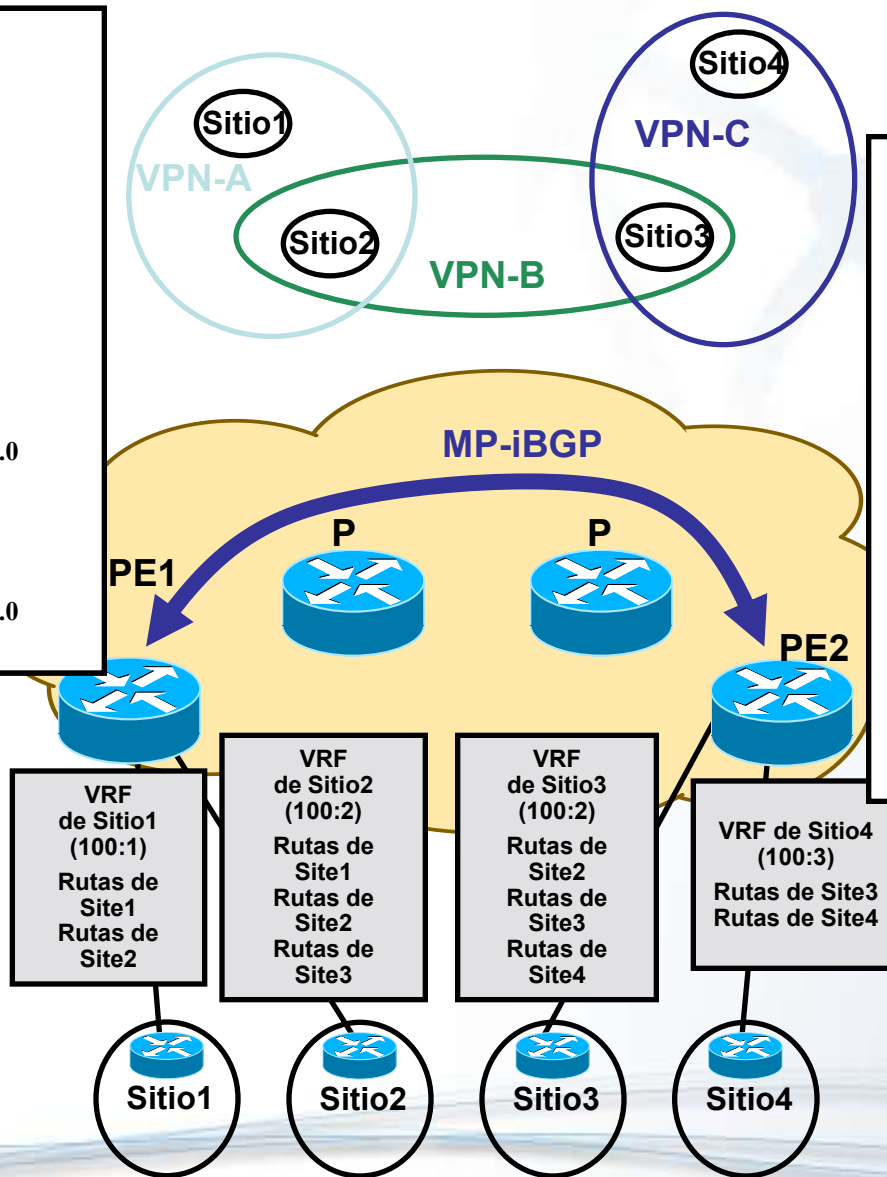
# Configuración: PE

- “show” pero basados en VRF
  - show ip route vrf <vrf-symbolic-name> ...
  - show ip protocol vrf <vrf-symbolic-name>
  - show ip cef <vrf-symbolic-name> ...
  - ...
- PING /Telnet basados en VRFs
  - telnet /vrf <vrf-symbolic-name>
  - ping vrf <vrf-symbolic-name>
  -

# Configuración: PE/CE

```
ip vrf site1
rd 100:1
route-target export 100:1
route-target import 100:1
ip vrf site2
rd 100:2
route-target export 100:2
route-target import 100:2
route-target import 100:1
route-target export 100:1
!
interface Serial3/6
ip vrf forwarding site1
ip address 192.168.61.6 255.255.255.0
encapsulation ppp
!
interface Serial3/7
ip vrf forwarding site2
ip address 192.168.62.6 255.255.255.0
encapsulation ppp
```

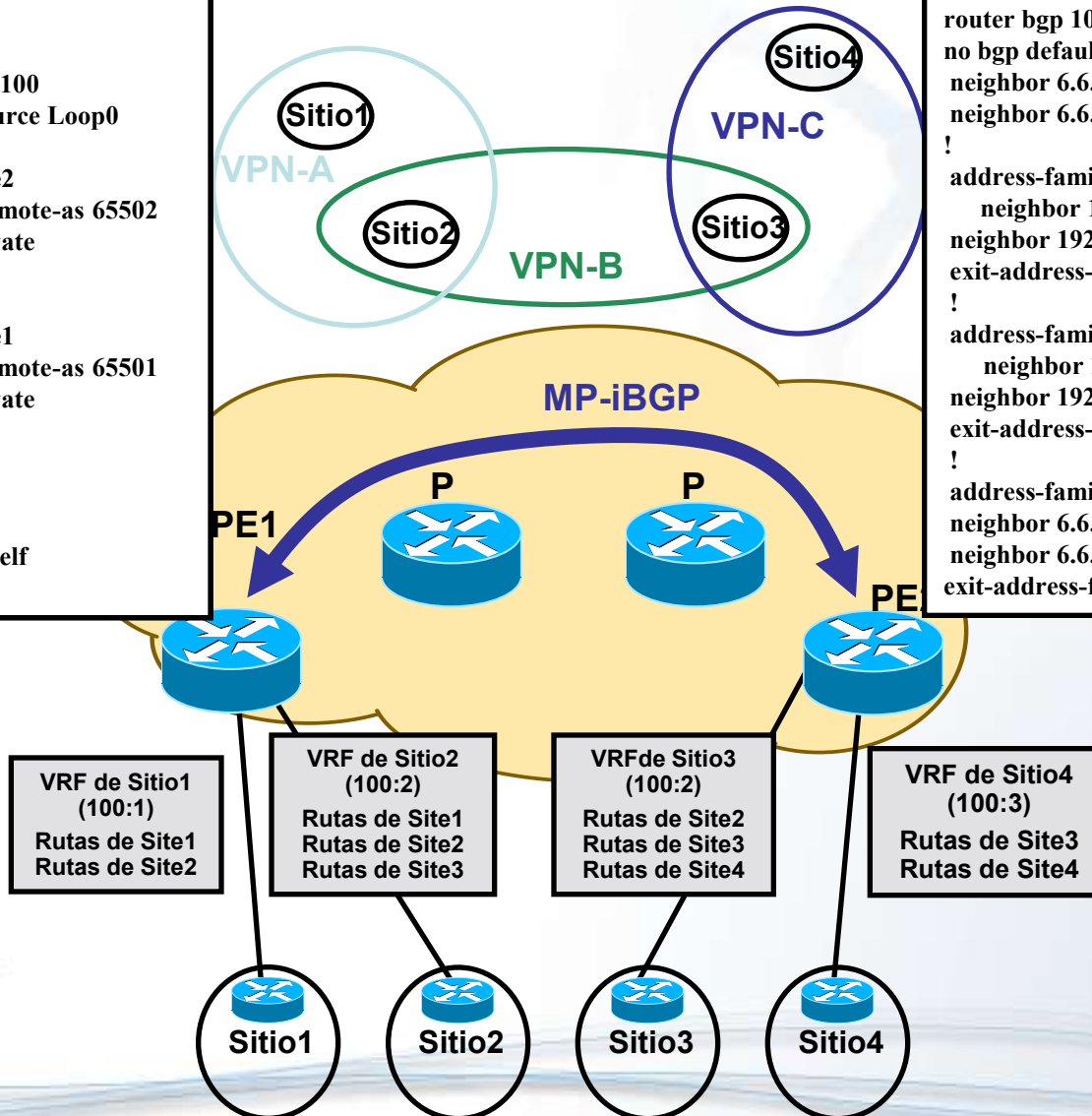
```
ip vrf site3
rd 100:2
route-target export 100:2
route-target import 100:2
route-target import 100:3
route-target export 100:3
ip vrf Sitio4
rd 100:3
route-target export 100:3
route-target import 100:3
!
interface Serial4/6
ip vrf forwarding site3
ip address 192.168.73.7 255.255.255.0
encapsulation ppp
!
interface Serial4/7
ip vrf forwarding site4
ip address 192.168.74.7 255.255.255.0
encapsulation ppp
```



# Configuración PE/CE

```
router bgp 100
no bgp default ipv4-unicast
neighbor 7.7.7.7 remote-as 100
neighbor 7.7.7.7 update-source Loop0
!
address-family ipv4 vrf site2
neighbor 192.168.62.2 remote-as 65502
neighbor 192.168.62.2 activate
exit-address-family
!
address-family ipv4 vrf site1
neighbor 192.168.61.1 remote-as 65501
neighbor 192.168.61.1 activate
exit-address-family
!
address-family vpnv4
neighbor 7.7.7.7 activate
neighbor 7.7.7.7 next-hop-self
exit-address-family
```

```
router bgp 100
no bgp default ipv4-unicast
neighbor 6.6.6.6 remote-as 100
neighbor 6.6.6.6 update-source Loop0
!
address-family ipv4 vrf site4
neighbor 192.168.74.4 remote-as 65504
neighbor 192.168.74.4 activate
exit-address-family
!
address-family ipv4 vrf site3
neighbor 192.168.73.3 remote-as 65503
neighbor 192.168.73.3 activate
exit-address-family
!
address-family vpnv4
neighbor 6.6.6.6 activate
neighbor 6.6.6.6 next-hop-self
exit-address-family
```





**CERTuy**

 **AGESIC**

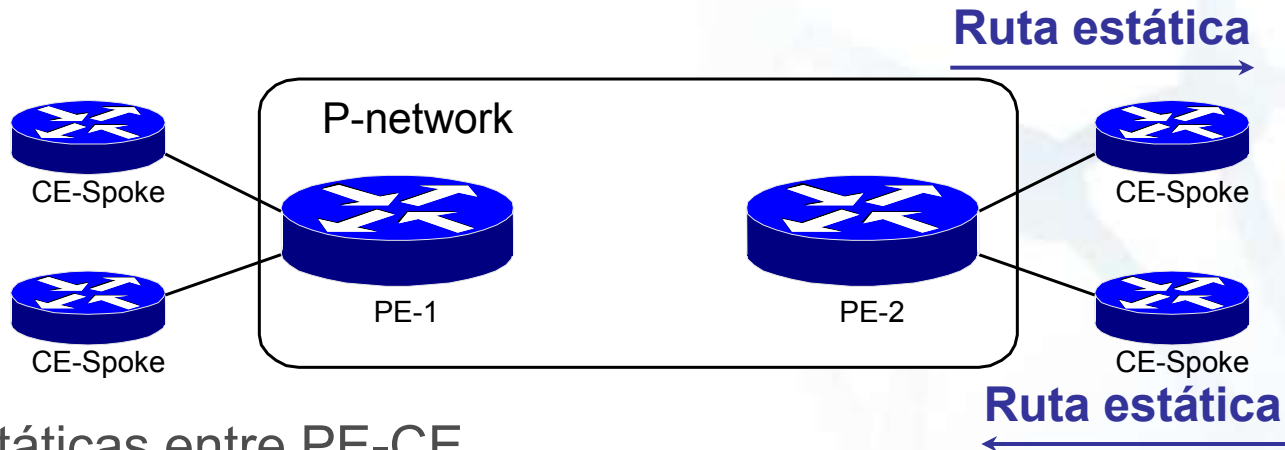


# Inyección de rutas estáticas en un PE



CCIE, CCSP y CSCI son marcas registradas de Cisco Systems, Inc

# Opciones de ruteo: estático

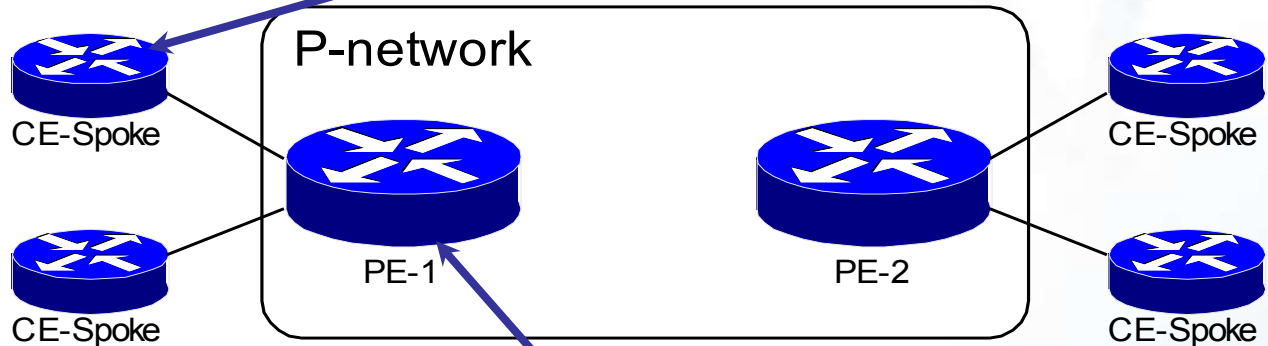


## Estáticas entre PE-CE

- Incrementa el trabajo del proveedor
- Pero es útil para cubrirse de desmadres en el dinámico del cliente
- 0.0.0.0 en el CE + estáticas en el PE
- Estáticas hay que redistribuirlas en MP-BGP

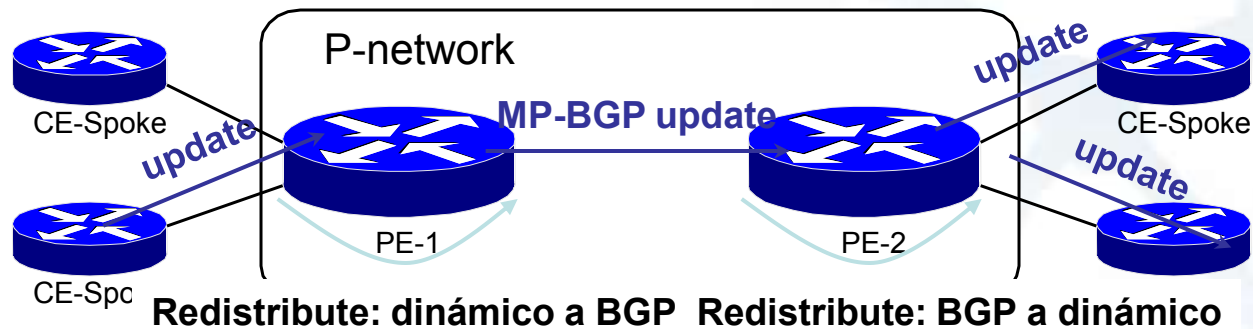
# Con rutas estáticas

```
ip route 0.0.0.0 0.0.0.0 serial 0
```



```
ip route vrf VPN_A 192.168.1.0 255.255.255.0 192.168.250.7 serial10/0
ip route vrf VPN_A 192.168.2.0 255.255.255.0 192.168.250.11 serial10/2
!
router bgp 213
 address-family ipv4 vrf VPN_A
  redistribute static
```

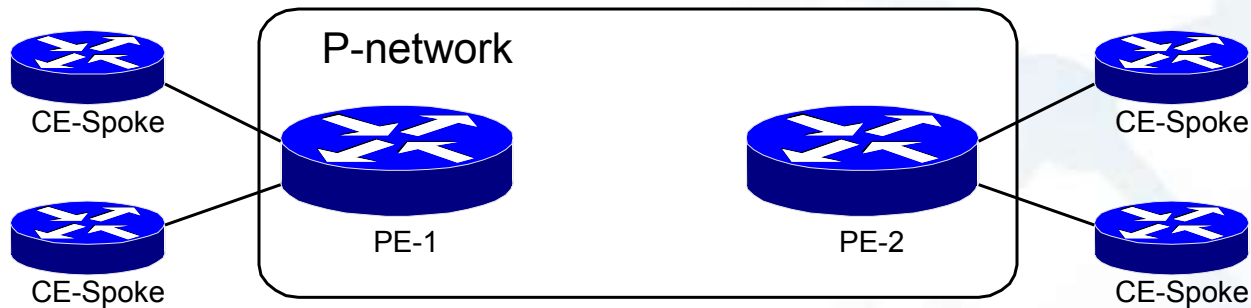
# Opción II: protocolo de rutas



## Dinámico entre PE y CE

- Rutas del CE redistribuidas en MP-BGP, pasadas por el backbone y redistribuidas de PE a CE
- Todo dinámico
- Util cuando los CEs necesitan todas las rutas

# Opciones de dinámicos PE-CE



- Soportados: RIPv2, OSPF, EIGRP, e-BGP
- RIP es sencillo si la convergencia no es un problema
- OSPF es más complejo
- BGP tiene algunas otra ventajas
  - Multihoming
  - Atributos que pasan end to end!



**CERTuy**

 **AGESIC**



# BGP como protocolo entre PE y CE



CCIE, CCSP y CSCI son marcas registradas de Cisco Systems, Inc

# Beneficios de usar BGP

- Continuidad de políticas entre sites
  - Se propagan todos los atributos  
AS\_PATH, Aggregator, Community
- No es necesario redistribución

# Beneficios de usar BGP

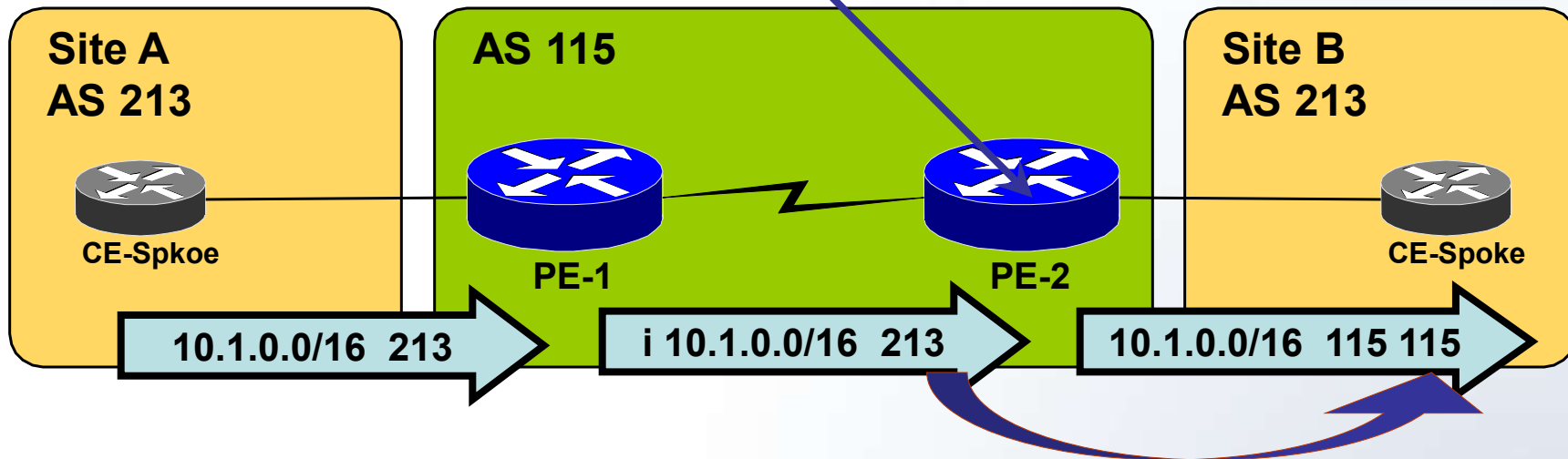
- Pueden usarse communities para definir políticas entre sites
- Filtros basados en atributos BGP
- Cliente puede definir sus políticas
- PE puede limitar nro de prefijos que aprende de CE

# Buenas prácticas de diseño

- AS privado y diferente en cada site
  - Sigue el modelo de conectividad Internet
- Mismo AS para los sites del cliente
  - Necesita el uso de AS-override para contrarrestar los efectos de control de loops

# Uso de “as-override”

```
router bgp 115
  address-family ipv4 vrf Customer_A
    neighbor 10.200.2.1 remote-as 213
    neighbor 10.200.2.1 activate
    neighbor 10.200.2.1 as-override
```



- PE-2 reemplaza AS del cliente con el propio y propaga el prefijo



**CERTuy**

 **AGESIC**

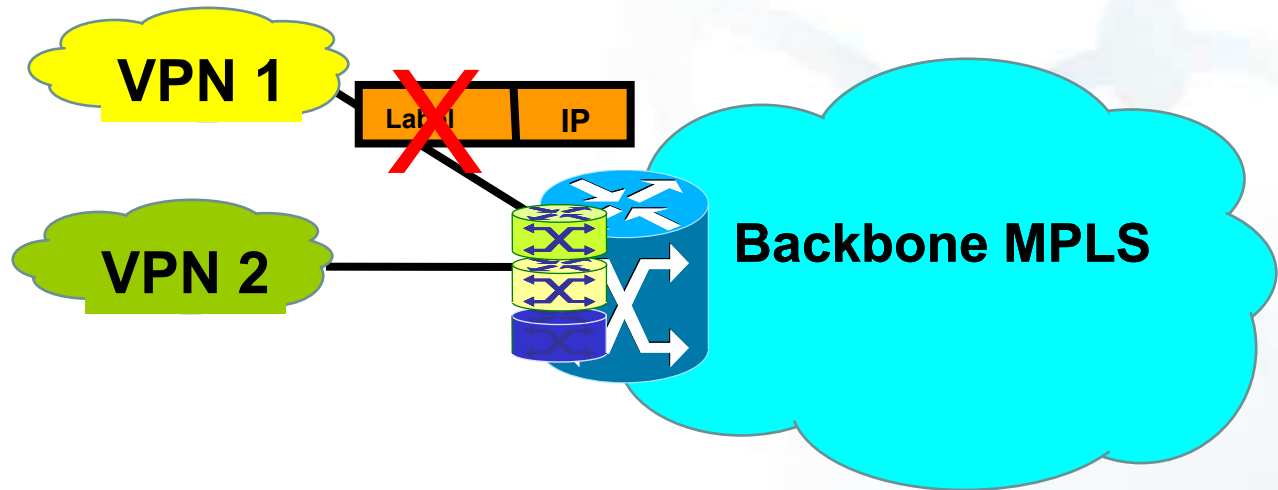


## Aspectos de seguridad del paradigma MPLS/VPN



CCIE, CCSP y CSCI son marcas  
registradas de Cisco Systems, Inc

# Podría ser susceptible de un ataque?



- No es posible que un atacante “inyecte” etiquetas (en el borde no existe soporte del protocolo LDP)
- En el borde, se deben tomar las mismas precauciones que para una arquitectura IP” (spoofing, DoS, etc)

# Cómo proteger un backbone MPLS

- Proteger el intercambio de rutas en la frontera PE/CE:
  - Optar por rutas estáticas si es posible
  - Filtrar lo que no haga falta
  - Aprovechar la presencia de autenticación MD5 en el intercambio
  - Usar las herramientas que proporciona el protocolo BGP (dampening, filtering, maximum-prefix)
- Protección de los PEs
  - Limitar el número de rutas aprendidas por VRF y por interfaz
  - Impedir tráfico de control innecesario en la frontera
  - Eventualmente, activar políticas de QoS para dejar pasar tráfico de control en momentos de congestión lícita o provocada por DoS