第三章 MPLS技术 PART-2

张喆
zhezhang@njupt.edu.cn
通信与信息工程学院

第三章 MPLS技术

- 3.1 MPLS的标记交换原理及LDP协议(回顾)
- 3.2 MPLS中的流量工程
- 3.3 MPLS VPN
- 3.4 VPLS
- 3.6 GMPLS
- 3.6 高速路由器的设计

多地健康码系统“频崩”专家呼吁：临近春运 应尽快推广刷身份证自动核验健康码

2022-01-12 18:17:00 来源：央广网

央广网北京1月12日消息（记者王晶）眼下，随着疫情反复，近期多个省市的本土“健康码”先后出现系统崩溃等问题，不少网友发出质疑，在经历了两年多的使用后，“健康码”系统为何仍问题频出？记者就此展开调查。

原因？

不过，每个地方的“健康码”系统并不统一，一位通信行业相关人士告诉记者，出现崩溃的原因也是多样的。“很多时候，是当日最高查询峰值激增导致系统阻塞。”他称，比如，和广东一样，西安“一码通”用户访问量激增时，出现每秒访问量达到以往峰值的10倍以上。

西安电子科技大学网络与信息安全学院教授杨超则引用了一个更为恰当的比喻，来形容这种“高峰”，“比如节假日前一天的晚高峰，这是近似全城的拥堵；而二维码的网络流量拥塞，更是一种定向拥堵。潜在的千万级别的并发请求流量同时涌向目标服务器，而网络带宽容量是有限的，必然会造成拥塞，进而导致一部分人不能正常访问服务，他解释说，感觉是“系统崩了”，其实是“网络塞了”。

MPLS的流量工程（MPLS-TE）

- 传统IP：路由协议虽然能够实现基本的路由选择，但缺乏对网络流量进行灵活调度和优化的能力，难以解决网络拥堵的问题。
- 人们的愿望：希望能够寻求一种更加高效、灵活的网络流量工程技术。
- MPLS-TE：应运而生

MPLS的流量工程

- 流量工程：
  - From ISP：流量工程（TE）可以保证网络资源得到充分、合理利用，从而避免了整个网络在某个地方网络资源过度利用，而在另外一些地方网络资源被闲置不用的不良情况；
  - From users：流量工程（TE）可以保证用户所申请的服务质量得到满足。

MPLS的流量工程

- MPLS是一种交换和路由的综合体，它集成了链路层的交换技术与网络层的路由技术。
- MPLS流量工程提供了完整的流量管理方法。
- MPLS流量工程根据业务流所需的资源和网络中的可用资源来引导业务流有效通过网络，并采用“基于约束路由的选路”方法，这条约束路由对要调节的业务来说是满足约束条件的最好路由。
- MPLS流量工程可以平滑地将失效链路或节点上的业务流利用新的约束转移到网络的其他通路上进行传输，从而有效地对发生故障的节点和链路进行恢复。

MPLS-TE的主要内容

- 1. 路径选择：
  - **通过采用MPLS的显式路由的选择方式，可以根据网络资源的合理利用来引导业务的流向，以便使一条拥挤路径上的一部分流量转移到一条负荷较轻的或不太拥挤的路径上，从而避免网络拥塞。**

MPLS-TE的主要内容

- 2.路径优先级的选择：
  - **通过设置LSP建立优先级和保持优先级来实现高优先级的业务流，即使已为某一业务建立了LSP，也应空出网络资源给高优先级的业务使用，以便在网络资源匮乏的时候，也能对优先级高的业务提供服务保证。**

MPLS-TE的主要内容

- 3．负载均衡
  - MPLS可以使用两条和多条LSP来承载同一个用户的IP业务流，**合理地将用户业务流分摊在这些LSP之间。**

- 4．路由备份
  - MPLS可以配置两条LSP，一条处于激活主用状态，另一条处于备份状态，一旦主LSP出现故障，业务立刻倒换到备份的LSP，直到主LSP从故障中恢复，业务再从备份的LSP切回到主LSP。

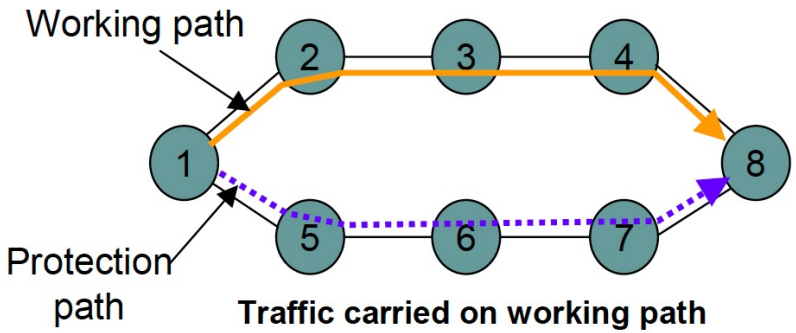
MPLS-TE的主要内容

- 5．故障恢复
  - **当一条已经建立的LSP在某一点出现故障时，故障点的MPLS会向上游发送消息，以通知上游LER重新建立一条LSP来替代这条出现故障的LSP，由此上游LER就会重新发出消息，建立另外一条LSP来保证用户业务的连续性。**

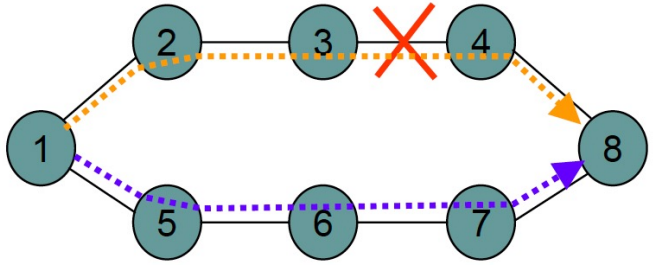
MPLS-TE的主要内容-MPLS PROTECTION

- IP routing recovers from faults in seconds to minutes
- SONET recovers in 50 ms
- MPLS targets in-between path recovery times
- Basic approaches:
  - Restoration: slower, but less bandwidth overhead
  - Protection: faster, but more protection bandwidth
- Repair methods:
  - Global repair: node that performs recovery (usually ingress node) may be far from fault, depends on failure notification message
  - Local repair: local node performs recovery (usually upstream from fault); does not require failure notification

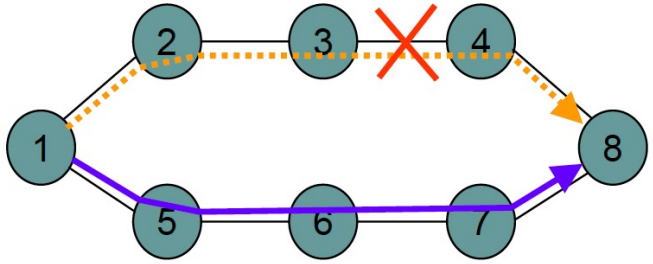
MPLS-TE的主要内容-MPLS PROTECTION



Traffic carried on working path



Failure on working path is detected



Traffic is switched to the protection path

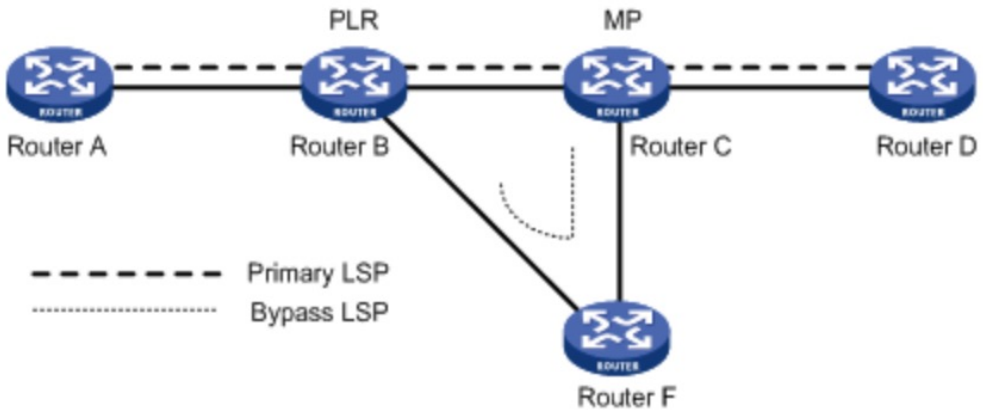
- Protection paths are setup as backups for working paths
  - 1+1: traffic is transmitted simultaneously on two paths
  - 1:1: traffic is transmitted on working path only
- Protection paths selected so that they are disjoint from working path
- Faster recovery than restoration

MPLS-TE的主要内容-MPLS PROTECTION

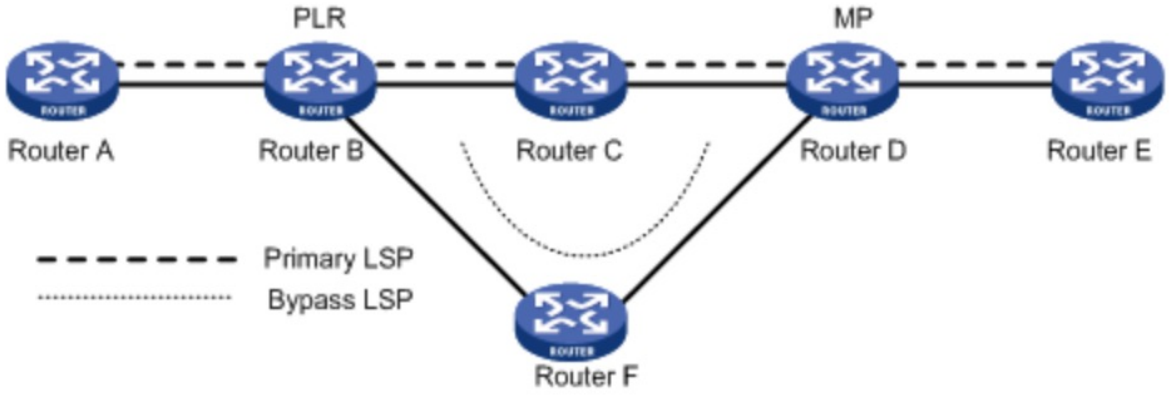
- 带宽预留的考虑
  - 一般说来，没有必要规定要为备份路径分配多少带宽资源。主通路的保持优先权可以作为备份路径的流量触发通路抢占优先权。
  - 流量触发：只有数据流切换到备份路径上传输，备份路径才能使用网络资源。

MPLS-TE的主要内容-MPLS PROTECTION

链路保护：



节点保护：



第三章 MPLS技术

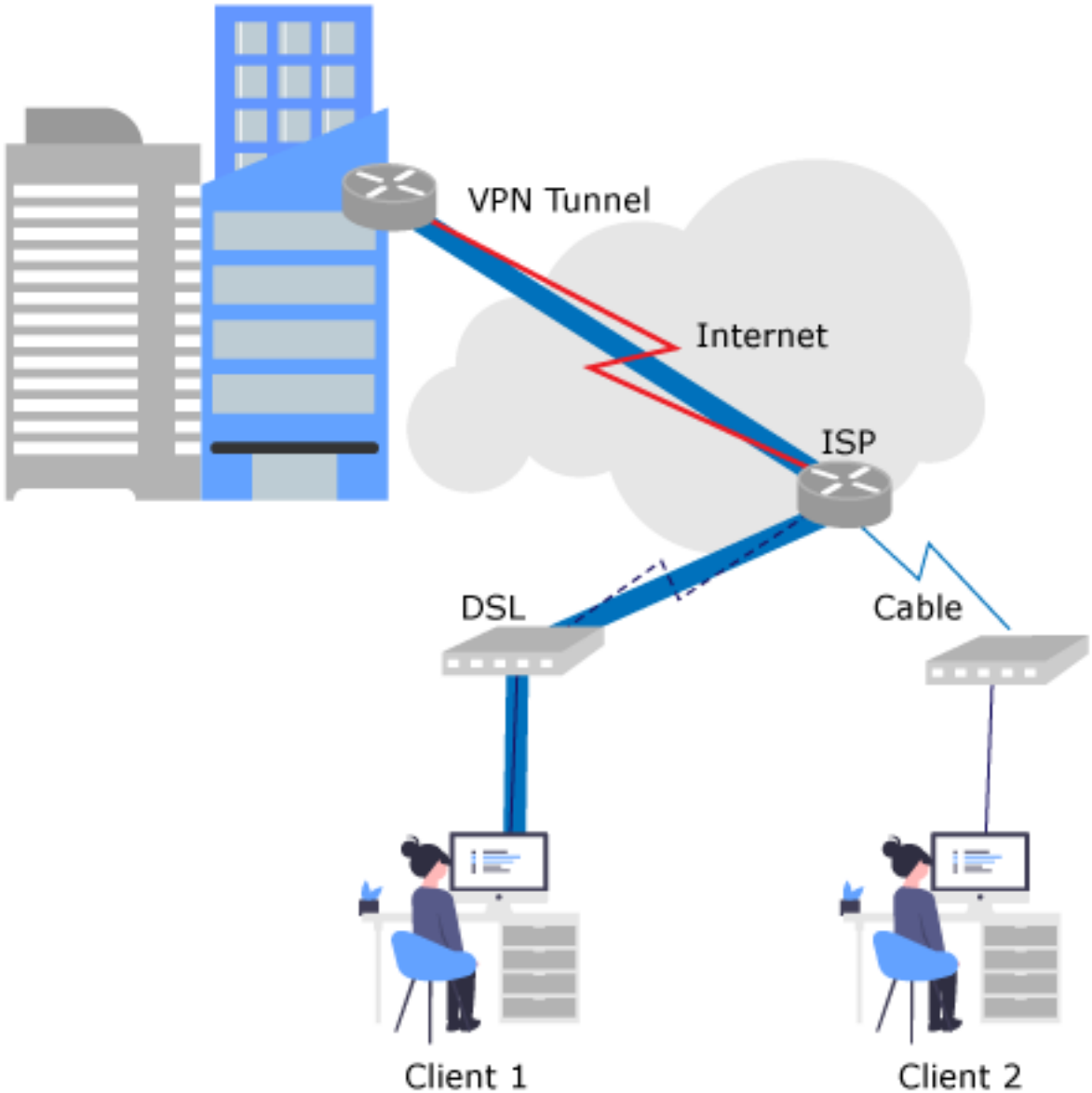
- 3.1 MPLS的标记交换原理及LDP协议(回顾)
- 3.2 MPLS中的流量工程
- 3.3 MPLS VPN
- 3.4 VPLS
- 3.6 GMPLS
- 3.6 高速路由器的设计

什么是VPN?

- VPN:利用公用网络来连接到企业私有网络。用安全机制来保障机密型，真实可靠行，完整性严格的访问控制。这样就建立了一个逻辑上虚拟的私有网络。

VPN定义

- **虚拟“virtual” 指没有物理的连接存在于2个网络间；事实上，连接是通过Internet的路由完成的；**
- **专用“private” 指传输的数据是保密的（通过加密和安全隧道）；**
- **网络“network” 指利用各种网络（私有、公用、有线无线等等）构成的通信手段；**


VPN Tunnel
Internet
ISP
DSL
Cable
Client 1
Client 2

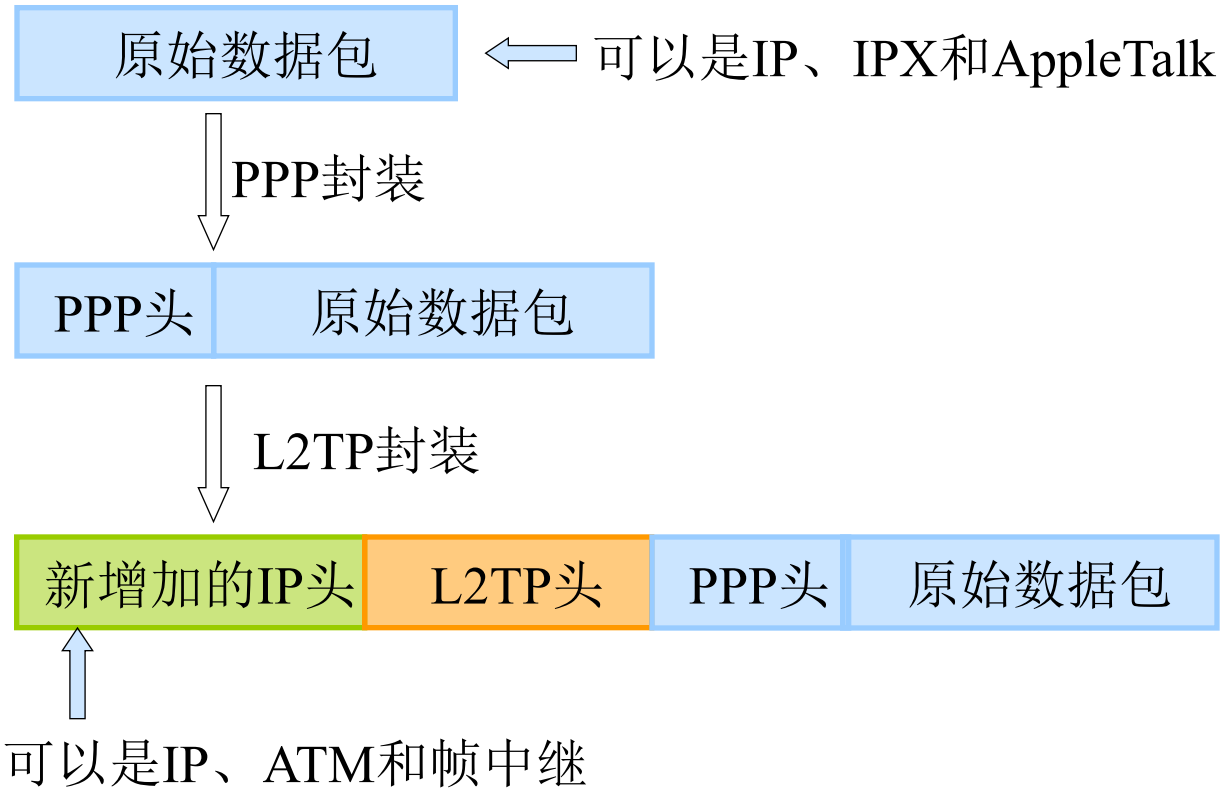
IP隧道实现技术

- **隧道技术：原始报文在A地进行封装，到达B地后把封装去掉还原成原始报文，这样就形成了一条由A到B的通信隧道。**

- **常用的IP隧道技术：**
  - **L2TP(Layer2 Tunneling Protocol) IETF, Cisco：2层**
  - **GRE(Generic Routing Encapsulation)：3层**
  - **IPSec: 3层**
  - **MPLS: 2层，3层**

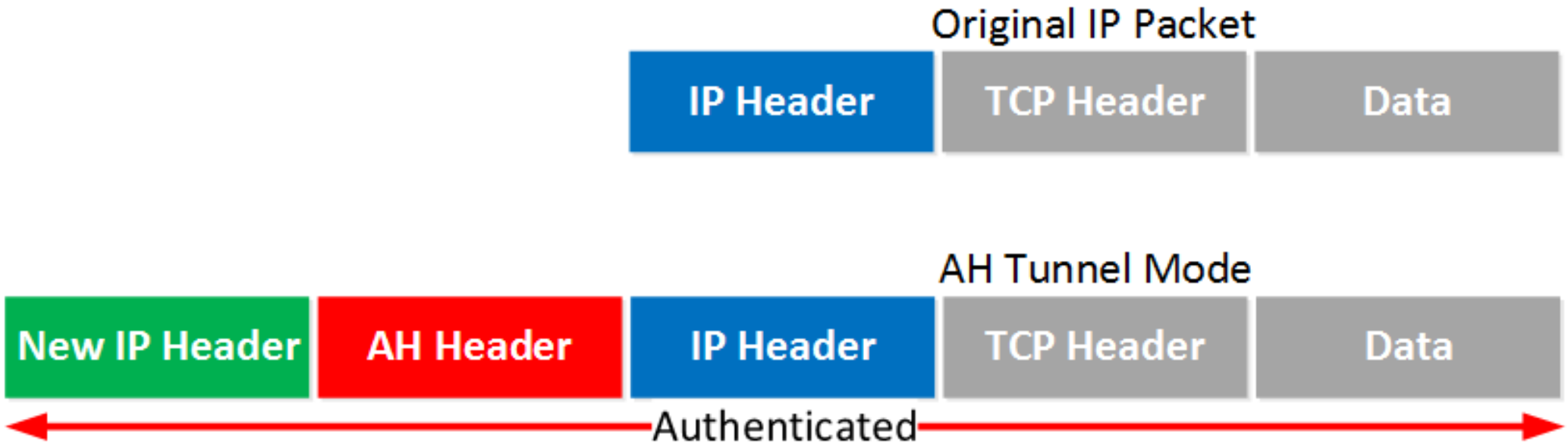
L2TP

- L2TP封装的乘客协议是位于第二层的PPP协议。



- L2TP没有对数据进行加密。

IPSEC


Original IP Packet
IP Header
TCP Header
Data
AH Tunnel Mode
New IP Header
AH Header
IP Header
TCP Header
Data
Authenticated

MPLS VPN

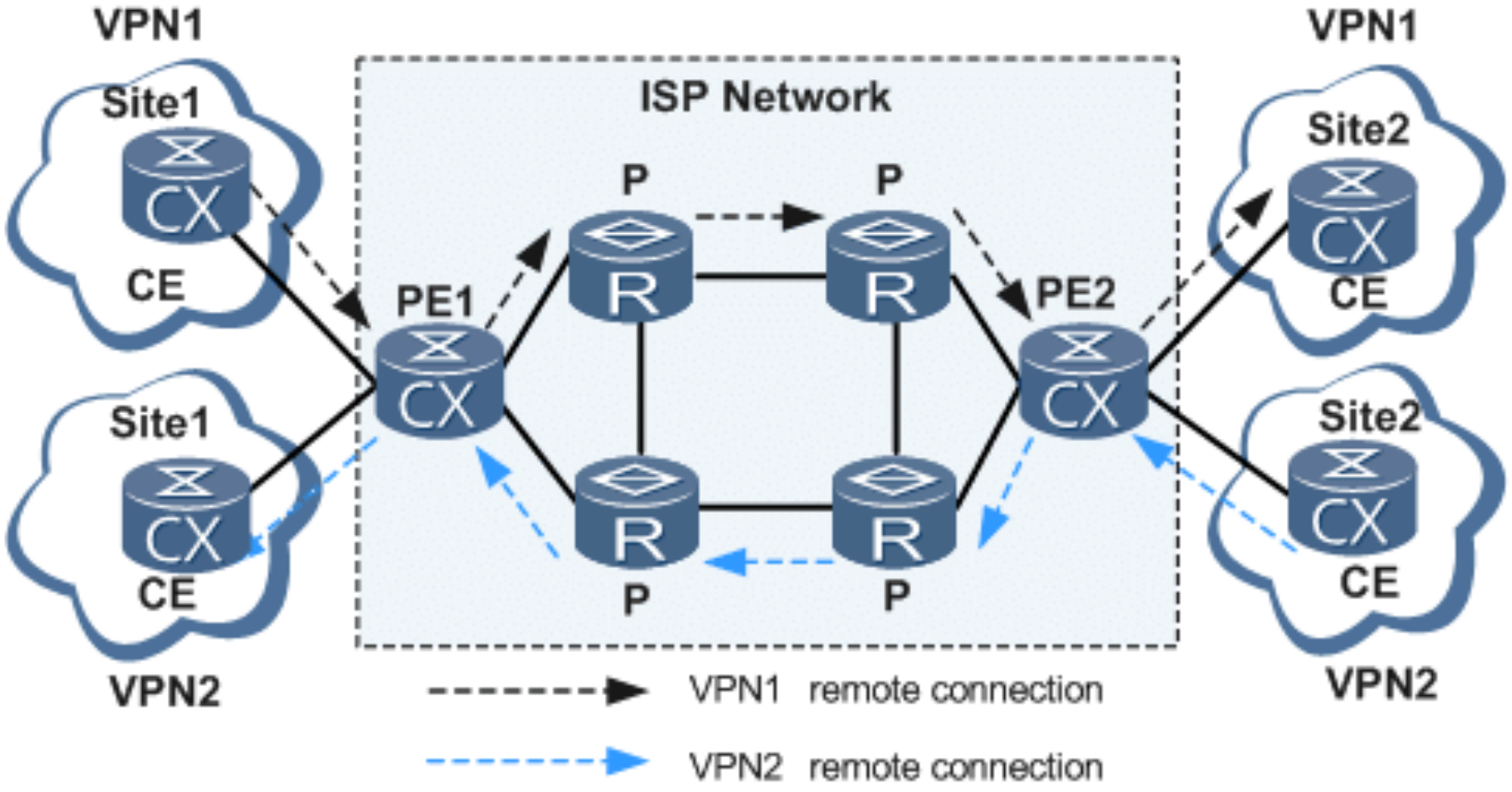
- MPLS的一个重要应用是VPN；
- MPLS VPN是一种基于MPLS技术的IP-VPN，根据PE（Provider Edge）设备是否参与VPN路由处理又细分为二层VPN和三层VPN。

MPLS/BGP VPN的模型

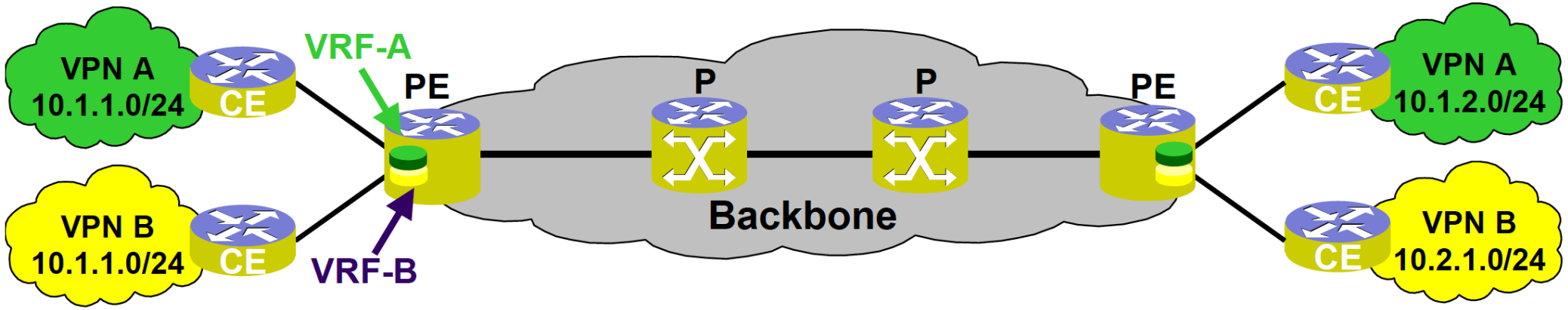
- MPLS/BGP VPN: 三层VPN。
- 在MPLS/BGP VPN的模型中，网络由运营商的骨干网与用户的各个Site组成，所谓VPN就是对site集合的划分，一个VPN就对应一个由若干site组成的集合。如MPLS/BGP VPN的实现如图所示；
- 利用标记堆叠来实现VPN，在一个IP分组上叠加两个MPLS标记头标进行转发，外侧标记用于转发，内侧标记用于VPN。

MPLS VPN

CE: Customer edge, PE: provider edge

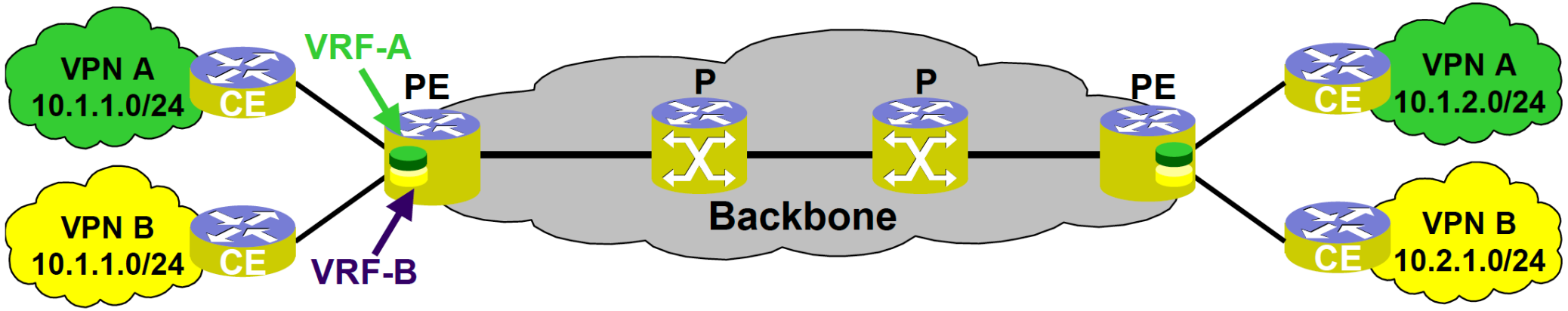


VPN 路由信息的通告



- 分三部分：本地CE 到入口PE、入口PE 到出口PE、出口PE 到远端CE;

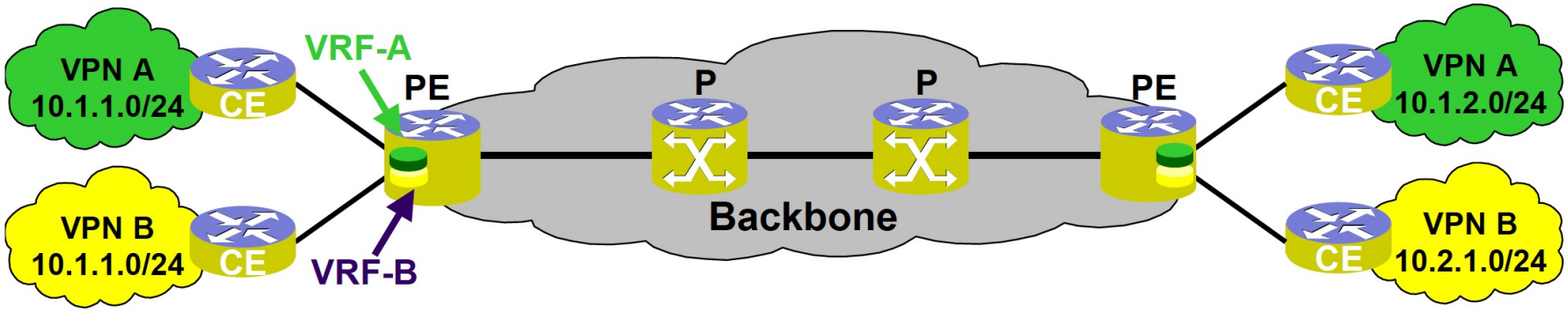
地址共享问题



OSPF，BGP，RIP等

- 本地CE 到入口PE

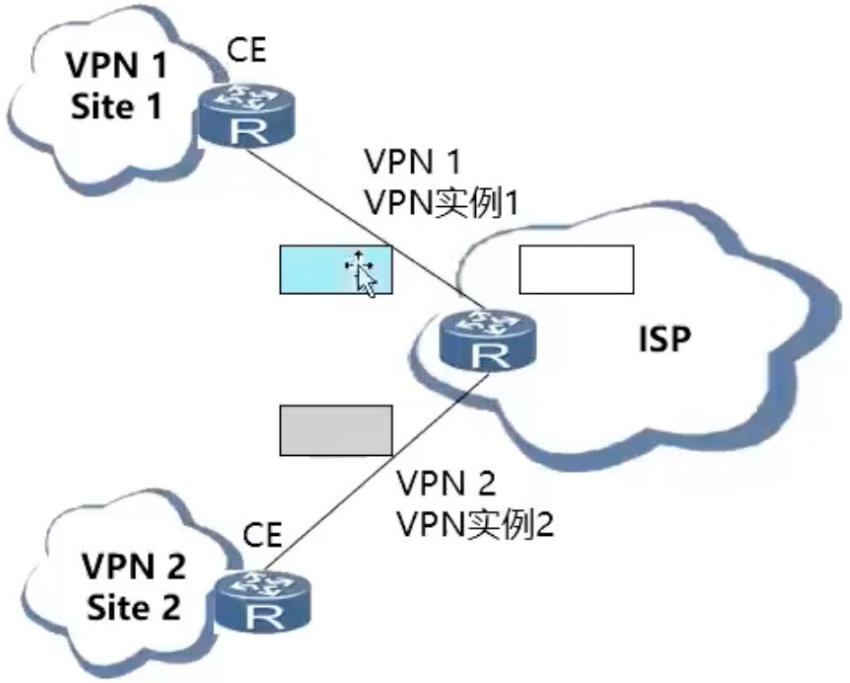
VPN ROUTING AND FORWARDING (VRF) TABLE



- Each VPN needs a separate VPN routing and forwarding instance (VRF) in each PE router to
  - Provides VPN isolation
  - Allows overlapping, private IP address space by different organizations

DETAIL

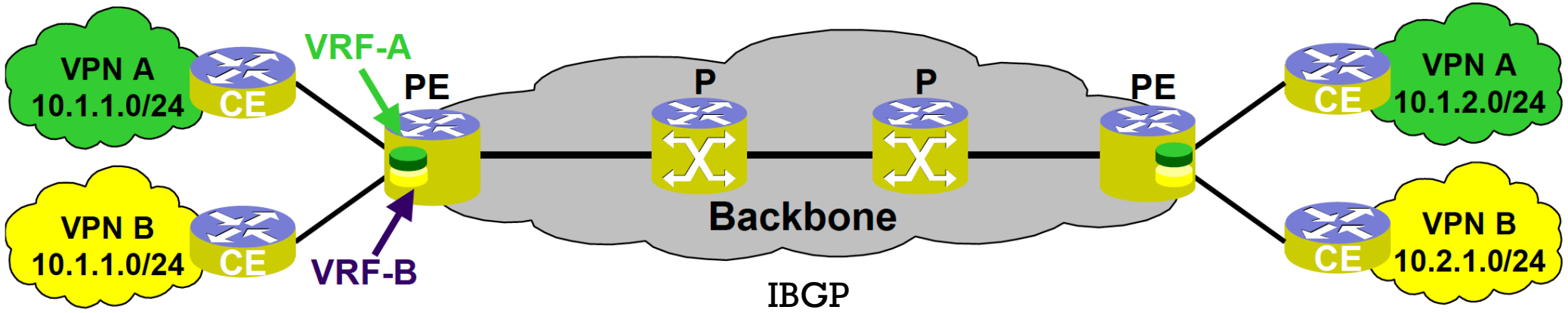
- 利用VRF与VPN绑定



VRF

- VRF（可以理解为虚拟路由器）包括以下元素：
  - 一张独立的路由表，从而包括了独立的地址空间；
  - 一组归属于这个VRF的路由器接口的集合；
  - 一组只用于本VRF的路由协议。
    - **VRF中定义的和VPN业务相关的参数：RT和RD。**

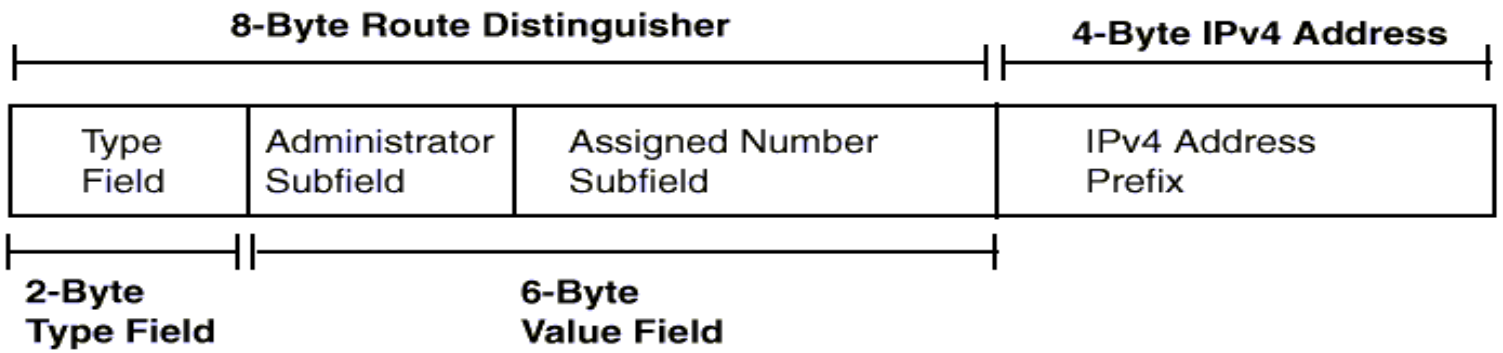
中间路由器P无VRF信息



- 入口PE到出口PE

RD (ROUTE DISTINGUISHER)

- 每个VRF中分别配置一个标识，称为RD。在PE发布VRF中的路由信息时，会在地址前面加上RD，以便接收方PE区分来自不同VRF的路由信息。



- RD的长度为8个字节。
- 在IPv4地址前加上RD之后，就称为VPN-IPv4地址族。

VPNV4 Address = Route Distinguisher + IPv4 地址

RD

- RD是本地PE路由器上VPN的唯一标识，主要用来区分具有相同地址的不同VPN路由(不同CE)
- 两个VRF中存在地址相同，由于RD不同，路由发布也不会混淆
- RD是给某VRF里面的路由打上标签；
- RD并不会影响不同VRF之间的路由选择以及VPN的形成，这些事情由RT完成。
- PE与CE之间传递的是IPv4路由， PE与PE之间传递的是VPN-IPv4路由

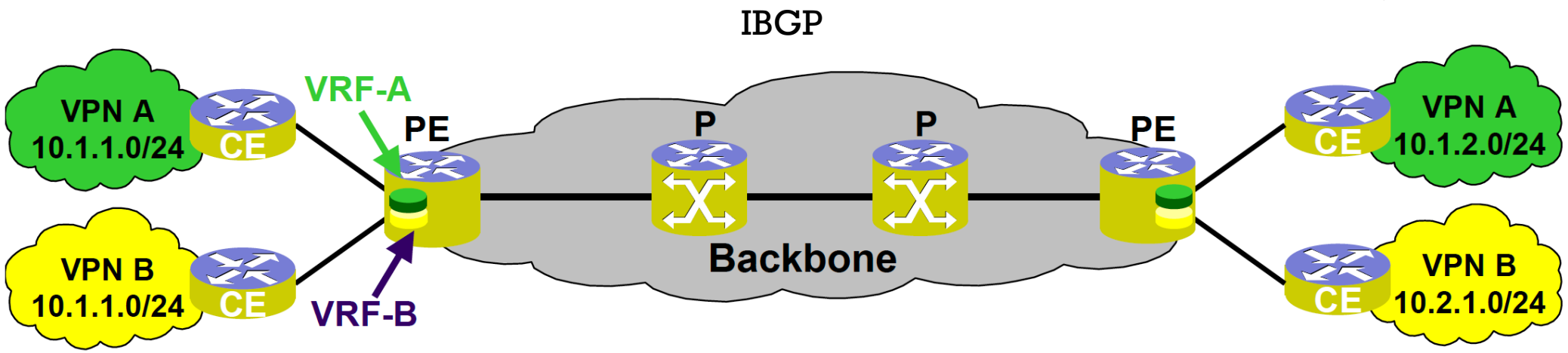
BGP的多协议扩展----MP-BGP

- BGP增加了两个扩展属性MP_REACH_NLRI和MP_UNREACH_NLRI。使用了这两种属性的BGP称为MP-BGP

MP_REACH_NLRI的结构

Address-family:指明使用了VPN-IPV4地址族
Next-hop:路由的下一跳地址
Label:24bit，与MPLS标签一样，但没有TTL字段
Prefix:64bit的RD＋IP前缀

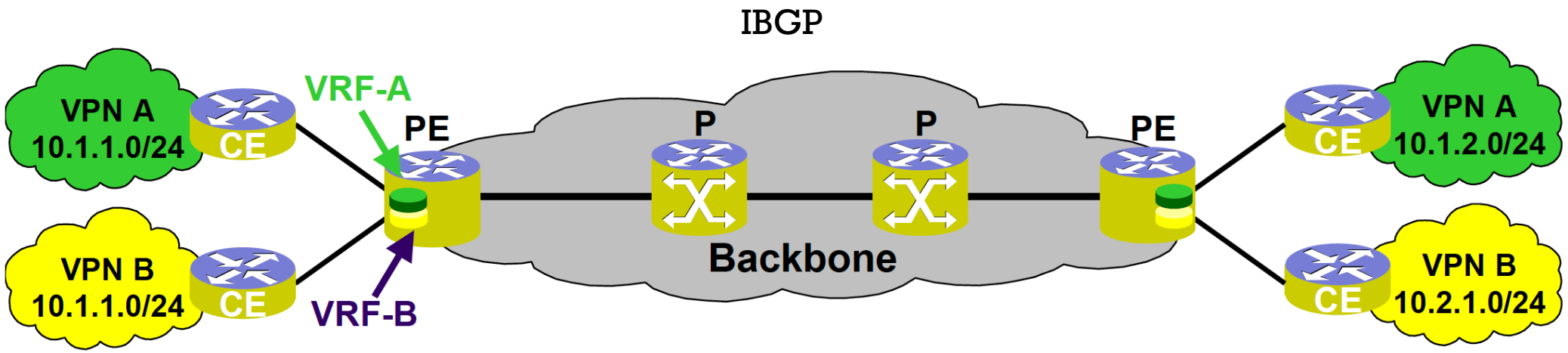
中间路由器P无VRF信息



RD+IPv4

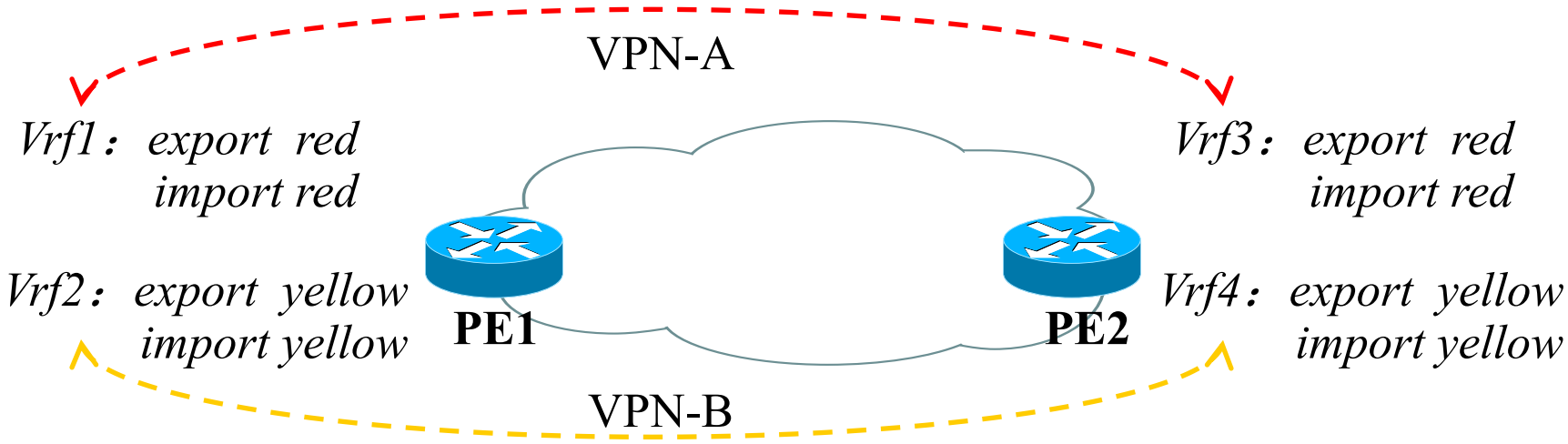
- **入口PE到出口PE**

出口PE该把路由放在哪个VRF里？

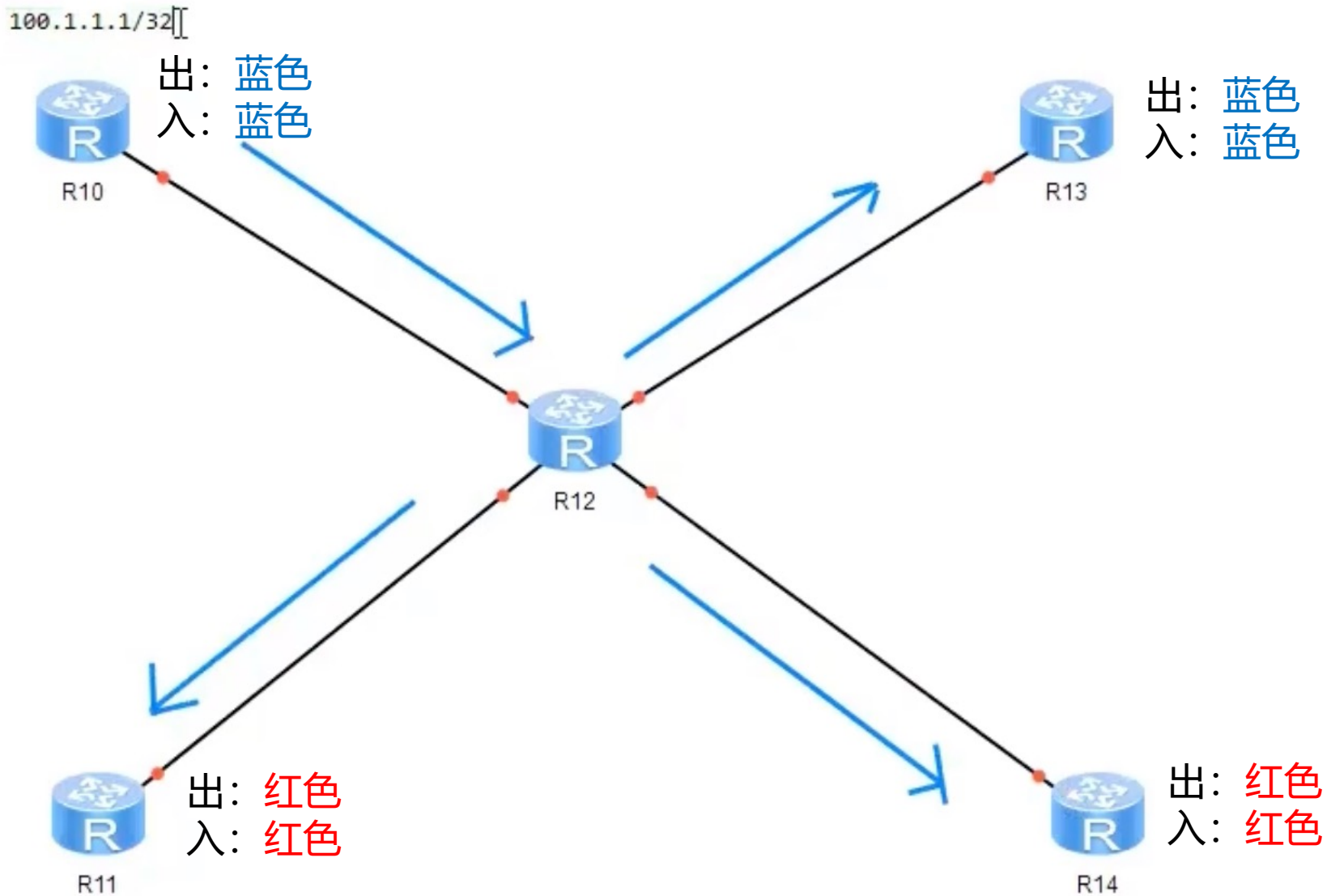


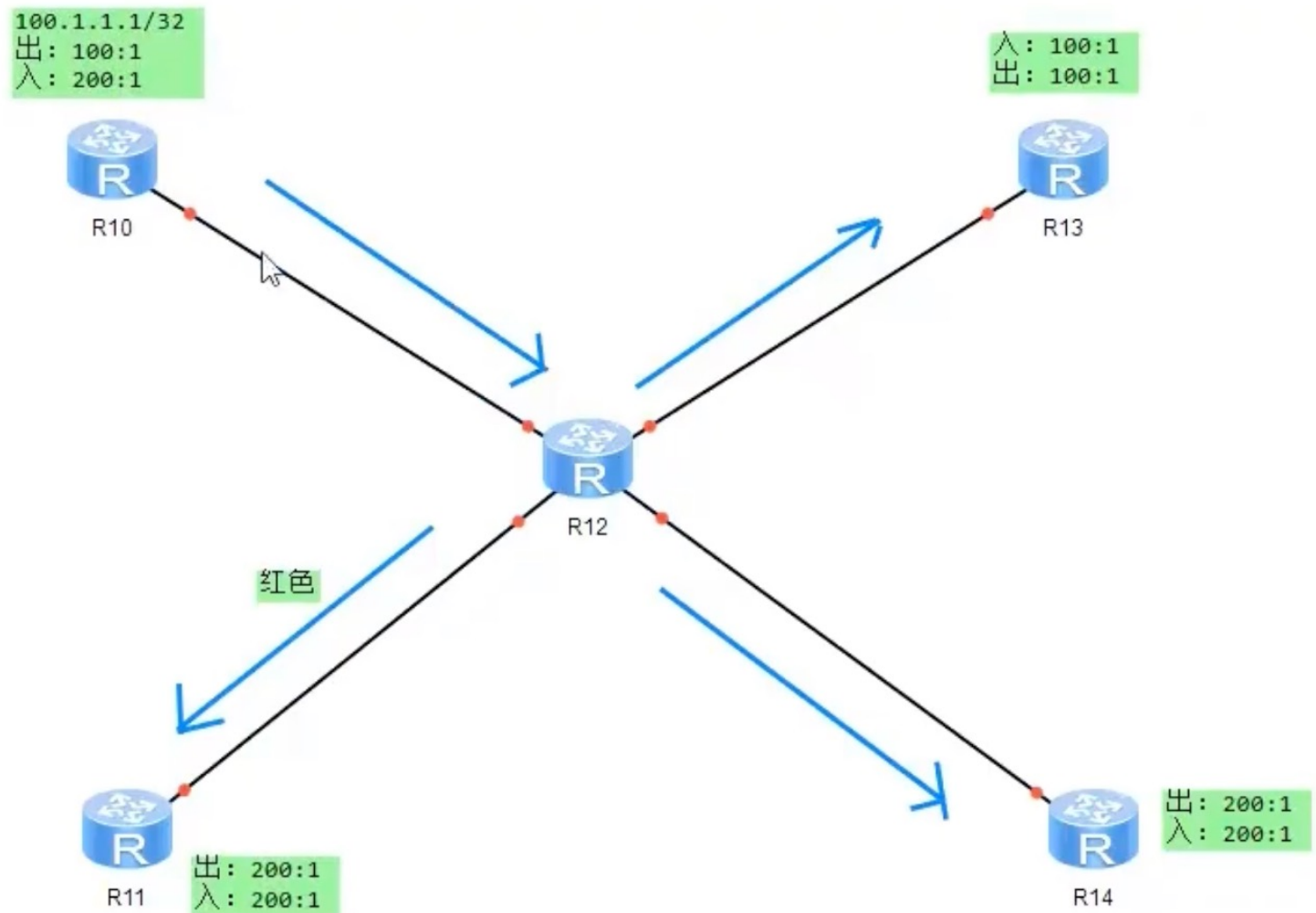
- 出口PE到远端CE

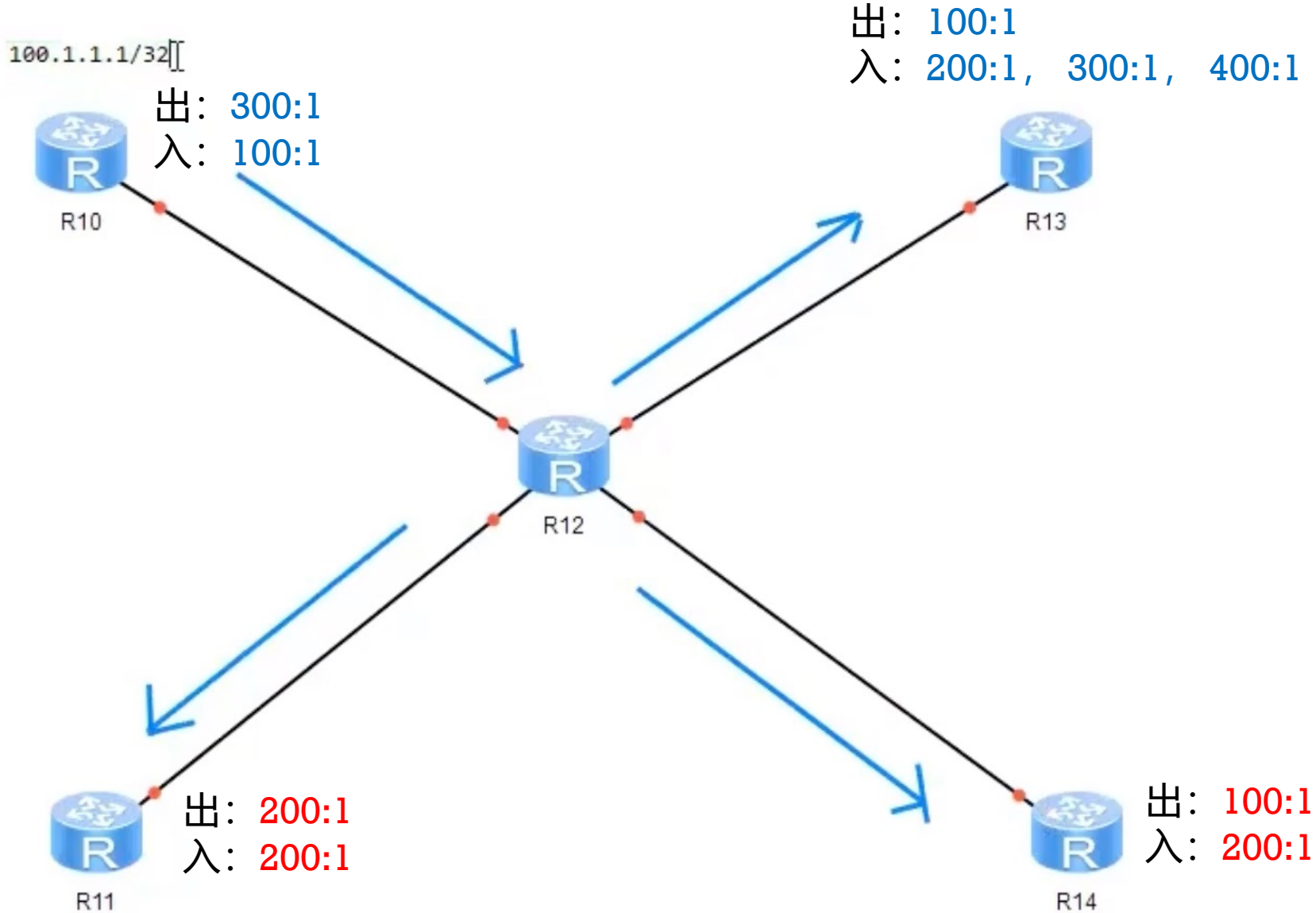
RT（ROUTE TARGET）-团体属性



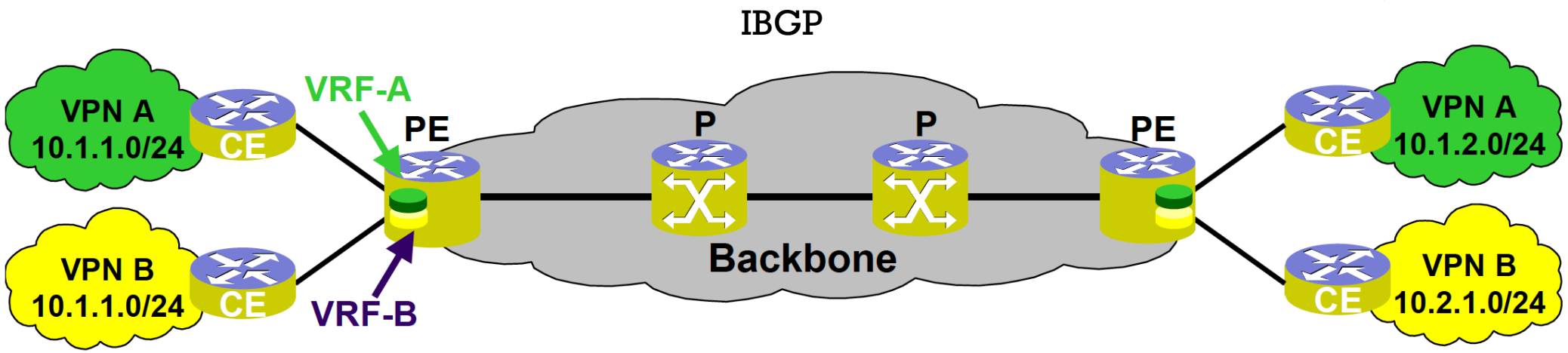
- RT的本质是每个VRF表达自己的路由取舍及喜好的方式（可以理解为“颜色”），分为两部分：
  - export target，表示发出路由的属性
  - import target，表示愿意接收什么路由
- RT是控制这个VRF里面可以发出和接受什么样的路由。
- RT格式与RD相似。


100.1.1.1/32
出：蓝色
入：蓝色
R10
出：蓝色
入：蓝色
R13
R12
出：红色
入：红色
R11
出：红色
入：红色
R14


100.1.1.1/32
出：100:1
入：200:1
R10
入：100:1
出：100:1
R13
R12
红色
R11
出：200:1
入：200:1
出：200:1
入：200:1
R14


100.1.1.1/32
出：300:1
入：100:1
出：100:1
入：200:1， 300:1， 400:1
R10
R13
R12
出：200:1
入：200:1
出：100:1
入：200:1
R11
R14

如何在出口PE区分VPN路由？

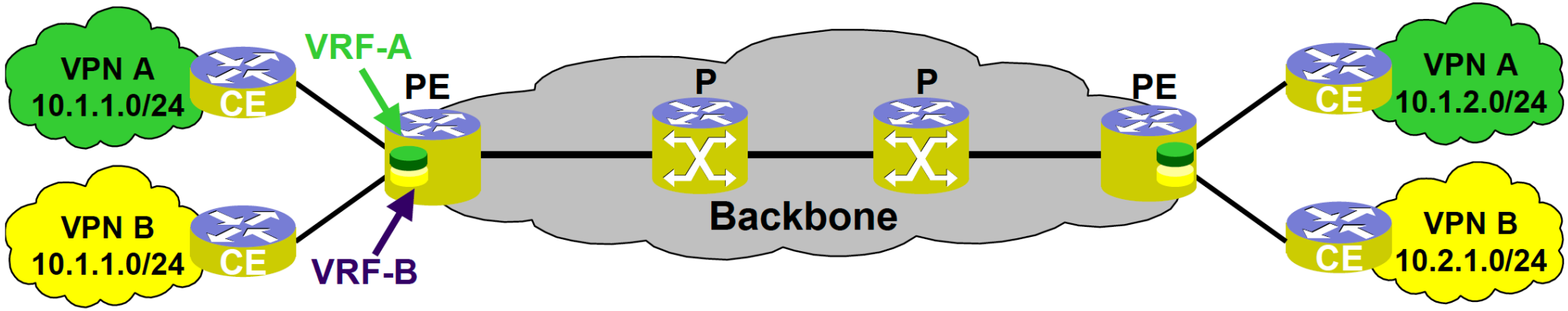


- 出口PE到远端CE

RT VS RD

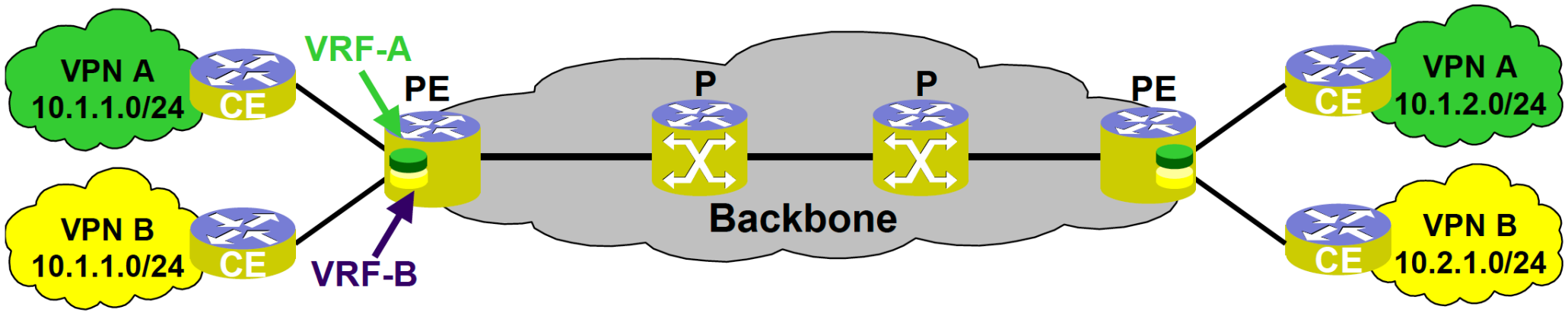
- **RD (Route Distinguisher)**
- 用于标识PE上不同VPN实例：实现VPN实例之间地址复用，与IP地址一起构成12 Bytes的VPNv4地址。
- RD与路由一起被携带在BGP Update报文中发送给对端。
- RD不具有选路能力，不影响路由的发送与接受。
- RD用来区分本地VRF，本地有效。
- **RT (Route Target)**：决定VPN路由的收发和过滤，PE依靠RT属性区分不同VPN之间路由。
- 当从VRF表中导出VPN路由时，要用Export RT
- RT是VPNv4路由携带的一个重要属性，对VPN路由进行标记。
- 当往VRF表中导入VPN路由时，只有所带RT标记与VRF表中任意一个Import RT相符的路由才会被导入到VRF表中。

VPN 路由信息的通告



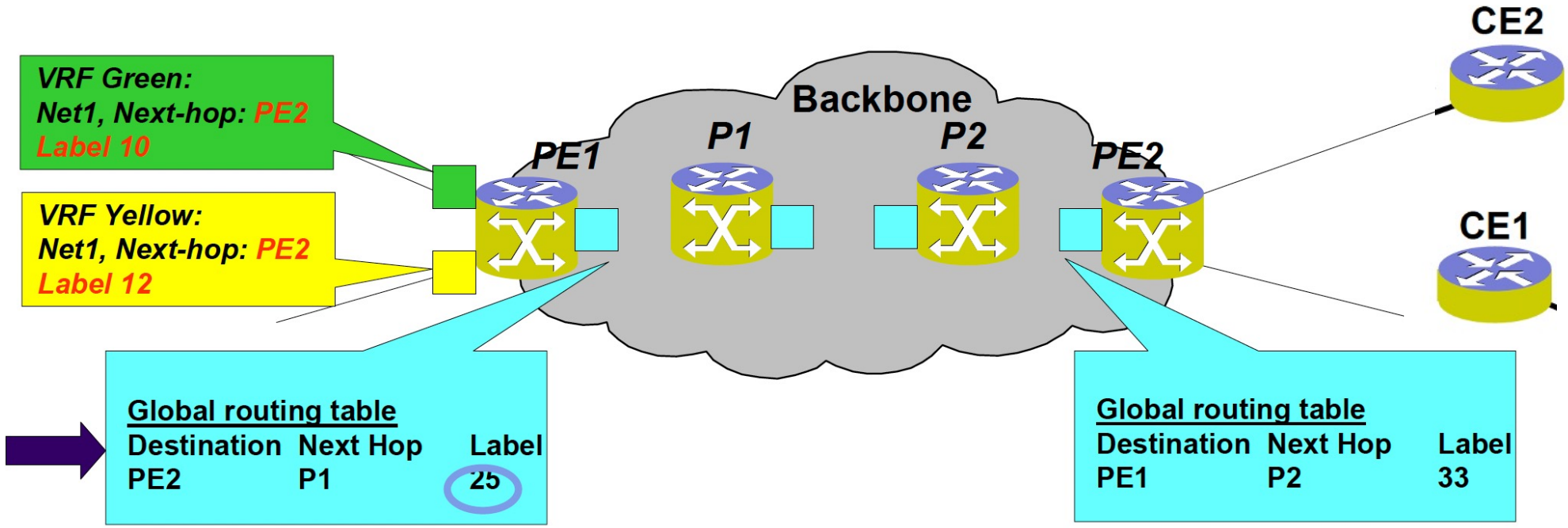
- 分三部分：本地CE 到入口PE、入口PE 到出口PE、出口PE 到远端CE;
- 入口PE 通过多协议扩展BGP（MP-BGP）把承载携带标记的VPN-IPv4 路由发布给出口PE;
- PE路由器使用MP-IBGP与其他PE路由器交换路由信息
- PE-CE之间交换路由信息可以通过静态路由、RIP、OSPF、IS-IS以及BGP等路由协议。通常采用静态路由，可以减少CE设备管理不善等原因造成对骨干网BGP路由产生震荡影响，保障了骨干网的稳定性。

路由黑洞问题



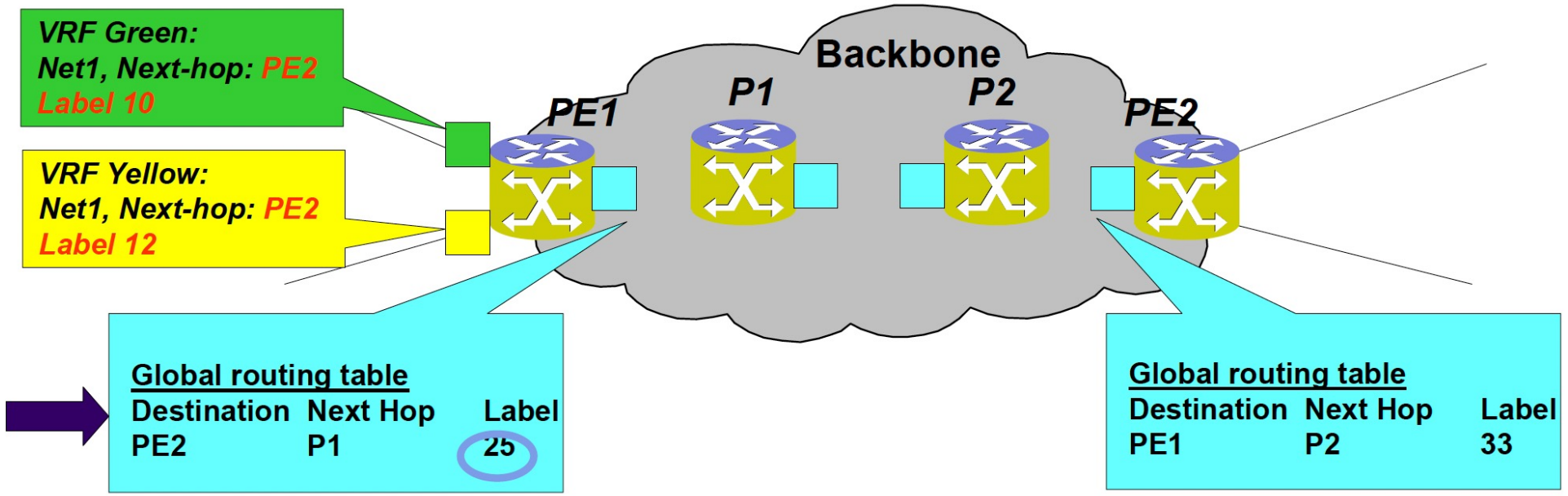
- P并没有出口PE的路由信息，收到来自入口PE的数据包会直接丢掉
- Solution?

MPLS VPN PACKET FORWARDING -LABEL STACKING



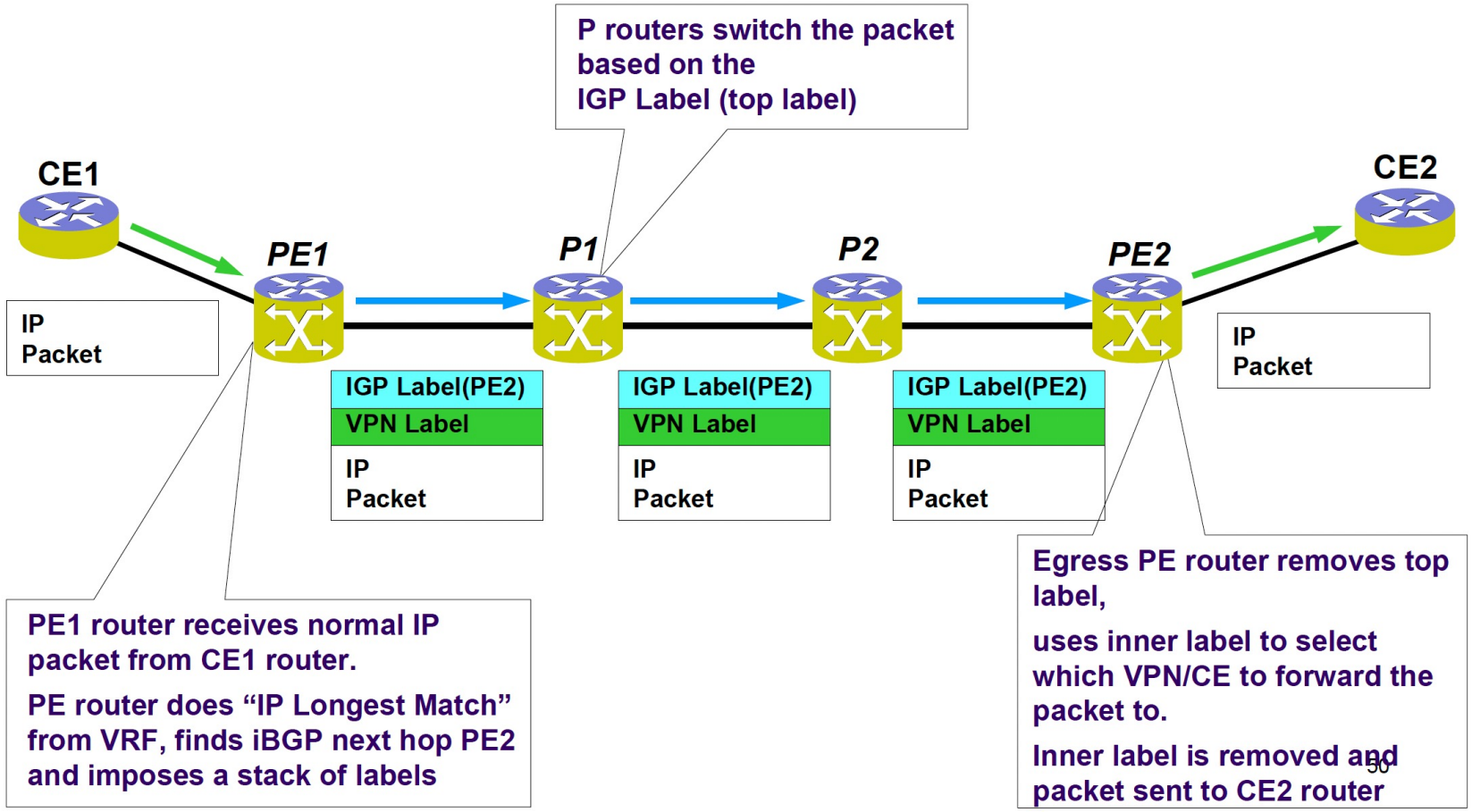
- 使用label来完成转发，此处的label称为top label
- 当PE2 pop掉label后，如何区分目的地？RD 和 RT?
- RD跟RT属于控制面，此处为转发面，看不到RD和RT

MPLS VPN PACKET FORWARDING -LABEL STACKING



- 在MPLS VPN中，属于同一的VPN的两个site之间转发报文使用两层标签；
  - 在入口PE上为报文打上两层标签，外层标签(PE router负责加注)在骨干网内部进行交换，代表了从PE到对端PE的一条隧道，VPN报文打上这层标签，就可以沿着LSP到达对端PE；
  - 使用内层标签(egress PE router负责分配)决定报文应该转发到哪个site上。

MPLS VPN PACKET FORWARDING -LABEL STACKING



第三章 MPLS技术

- 3.1 MPLS的标记交换原理及LDP协议(回顾)
- 3.2 MPLS中的流量工程
- 3.3 MPLS VPN
- 3.4 VPLS
- 3.6 GMPLS
- 3.6 高速路由器的设计

VPLS

- VPLS: Virtual Private Lan Service，也就是虚拟专用局域网业务。目前比较热门的一种MPLS二层VPN技术。VPLS技术用于在MPLS网络上提供LAN互联能力。（LAN OVER MAN/WAN）
- 目前业界有两个标准，
  - 一个标准以LDP作为信令协议，由Alcatel发起，得到业界大部分厂家的支持（包括Cisco）；
  - 另一个标准由Juniper发起，以BGP作为信令协议，目前只有Juniper和华为支持。
- It allows you to connect geographically dispersed LAN sites to each other across an MPLS backbone.

WHY VPLS?

- 既然已经有了MPLS VPN，为什么还需要VPLS?

- MPLS VPN工作在三层
- 广播与多播受限制
- 需要二层的VPN

VPLS

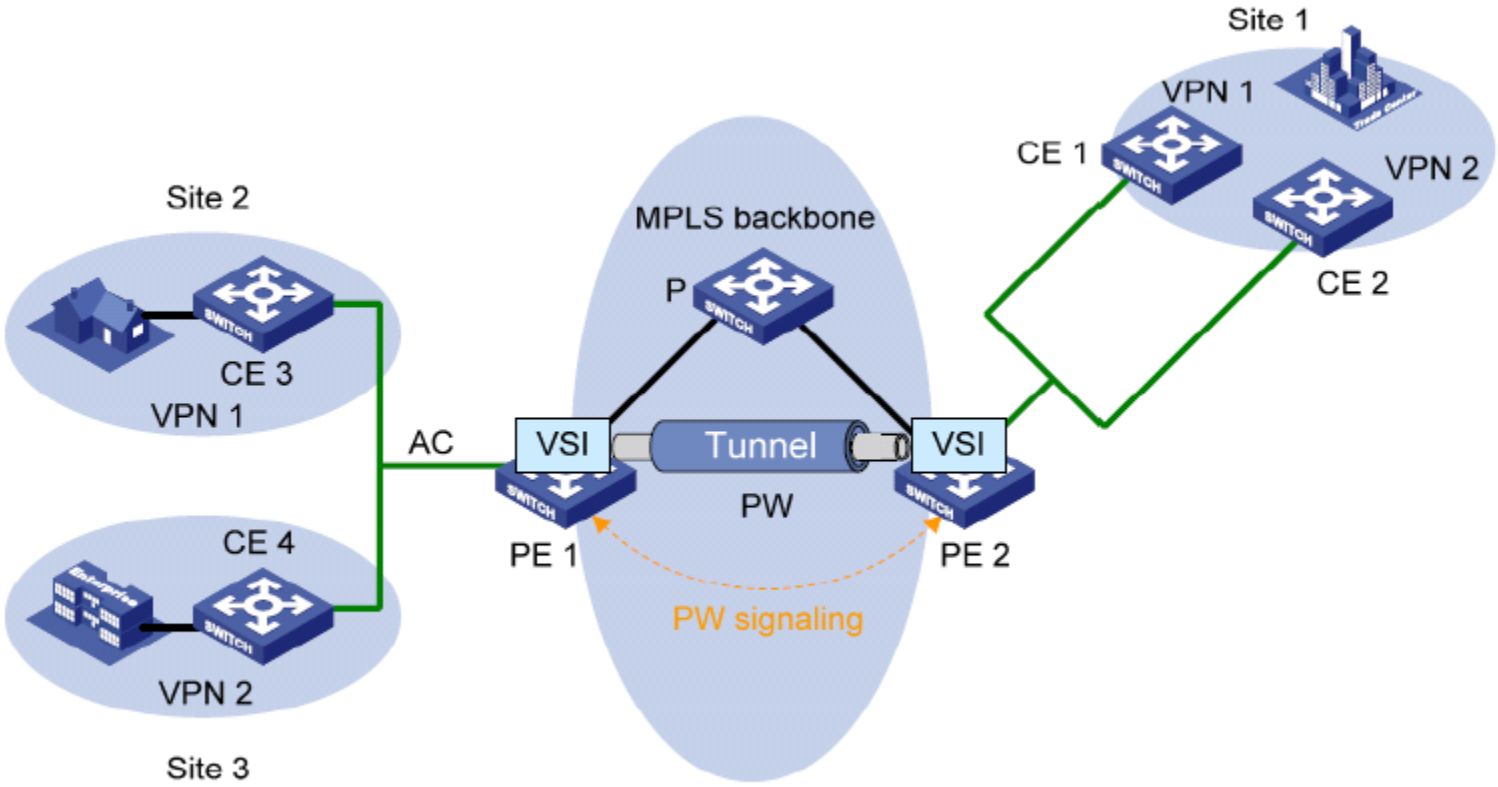
- 虚拟专用局域网业务(VPLS)是分组交换网(PSN)提供的一项业务，旨在通过预先建立的隧道和隧道中的伪线连接为用户提供专用的局域网(LAN)互联服务，属于二层VPN(L2VPN)的范畴。

- 原则上VPLS可以使用任何类型的隧道

VPLS

- 本质上是一种基于IP/MPLS和以太网技术的L2VPN（二层虚拟专用网）技术。
- 其核心思想是利用信令协议在VPLS实例中的PE（运营商边缘路由器）节点之间建立及维护PW（伪线），将二层协议帧封装后在PW上传输、交换，使广域范围内多个局域网在数据链路层面被整合为一张网络，向用户提供虚拟的以太网服务。
- VPLS技术有效地结合了IP/MPLS、L2VPN以太网交换等多种技术的特点，支持点到点、点到多点、多点到多点的业务类型，能够在较大网络规模下支持电信级以太网服务。

VPLS的网络结构(华为的文档)



GMPLS（1）

- 为了能适应未来智能光网络动态地提供网络资源和传送信令的要求，通用多协议标志交换协议GMPLS，（GMPLS：Generalized Multiprotocol Label Switching）：MPLS向光网络扩展的产物，它在支持传统的分组交换、时分交换、波长交换和光纤交换的同时，对原有的路由协议、信令协议作了修改和扩展。

GMPLS（2）

- GMPLS是对MPLS的扩展，从而涵盖了对时分交换（例如，SONET/SDH，PDH，G.709）和空分交换（如入端口或光纤到出端口或光纤）的支持。由于网络的不同层次可以实质上使用不同的数据或转发平面
- 为了支持电路交换（主要是SDH）和光交换（包括LSC和FSC），GMPLS设计了专用的标签格式，标签应该支持对光纤、波带、波长甚至时隙的标识。

Thank You