MPLS-VPNのしくみ

~RFC 2547モデル~

NTTコミュニケーションズ株式会社
池尻 雄一
ikejiri@ntt.ocn.ne.jp

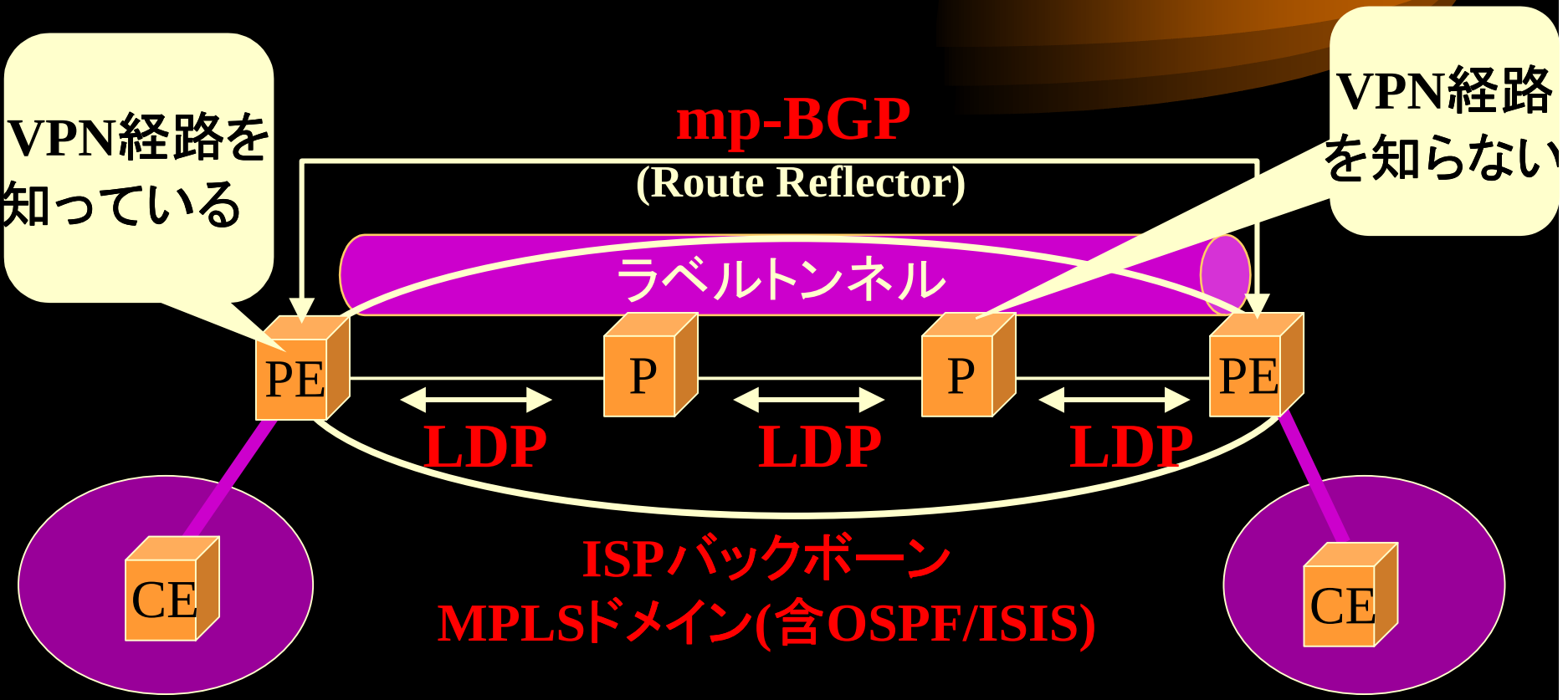
MPLS-VPNとは

- C社を中心としてRFC2547(Informational)に記されたISPサービスとしてのIP-VPN実現技術
- 網内パケット転送にMPLS(LDP/TDP)、VPN経路情報交換にBGP(mpBGP:RFC2283)を使用
- ルーティングプロトコルがエッジで終端されるPeerモデルのIP-VPN

RFC2547に登場するルータたち

- PEルータ：Provider Edge Router(お客様を収容するルータ、MPLSエッジルータ)
- Pルータ：Provider Router(MPLSコアルータ)
- CEルータ：Customer Edge Router(PEルータにつながるお客様ルータ)

RFC2547プロトコル構成



VPNA:拠点A

VPNA:拠点B

RFC2547でのラベルの使い方

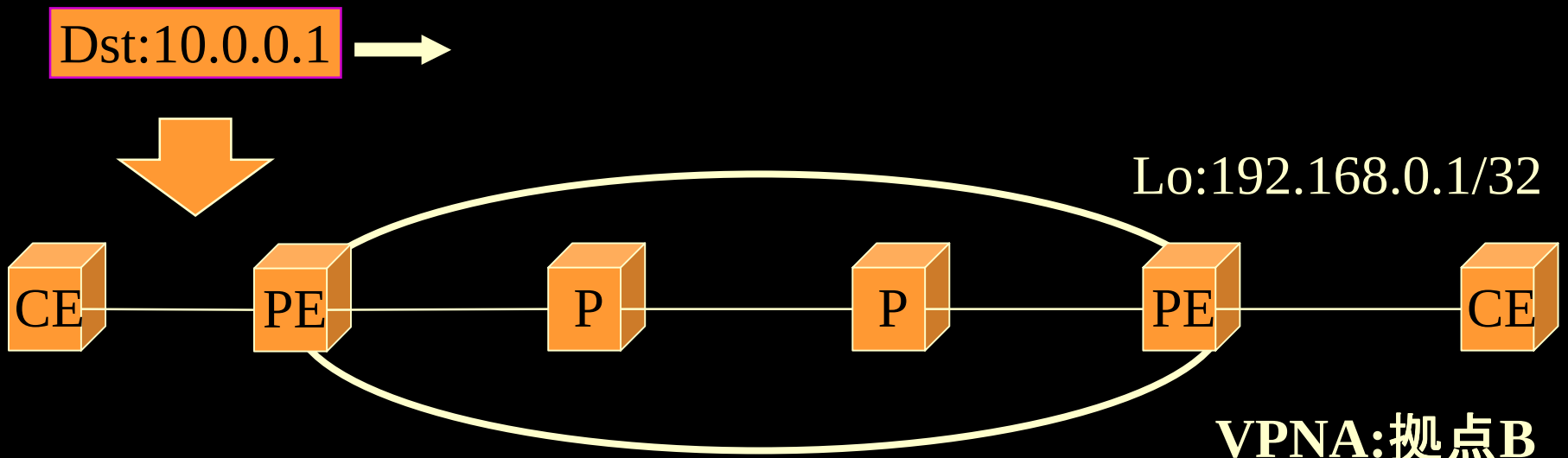
- MPLSラベルスタックを2つ使う

LDPで決定
(松嶋さんプレゼン参照)

mpBGPで決定
(PEルータ間のフルメッシュで)

パケット転送の流れ

VPNA；拠点B：10.0.0.1行きパケット到着

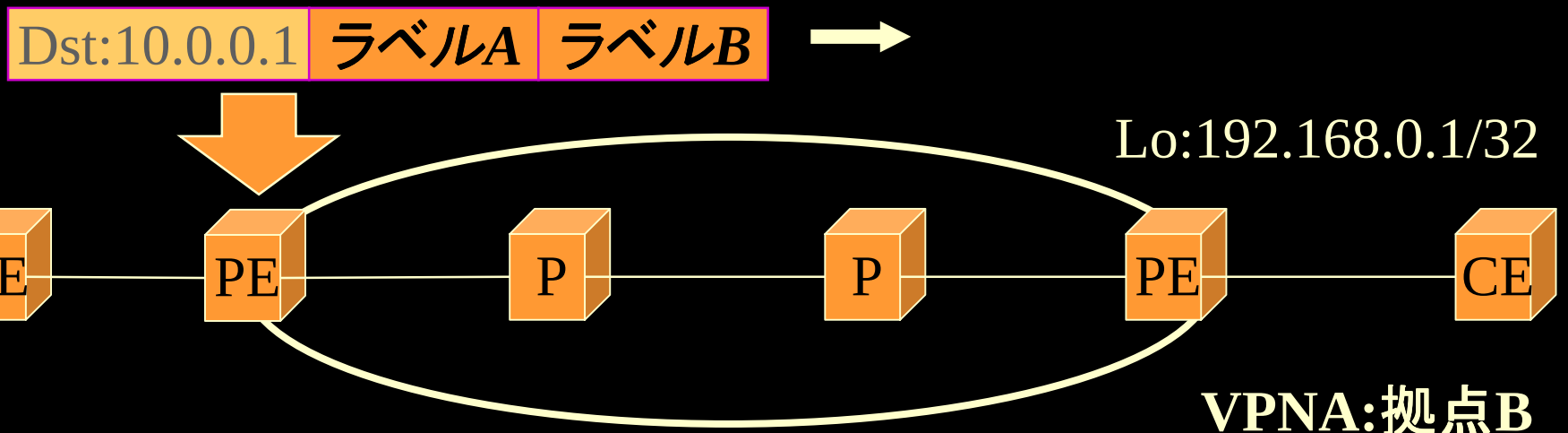


パケット転送の流れ(Cont.)

①入力インタフェースからVPNAであることを認識し、VPNA:10.0.0.0/8に相当するVPN識別用ラベルAを付与する。

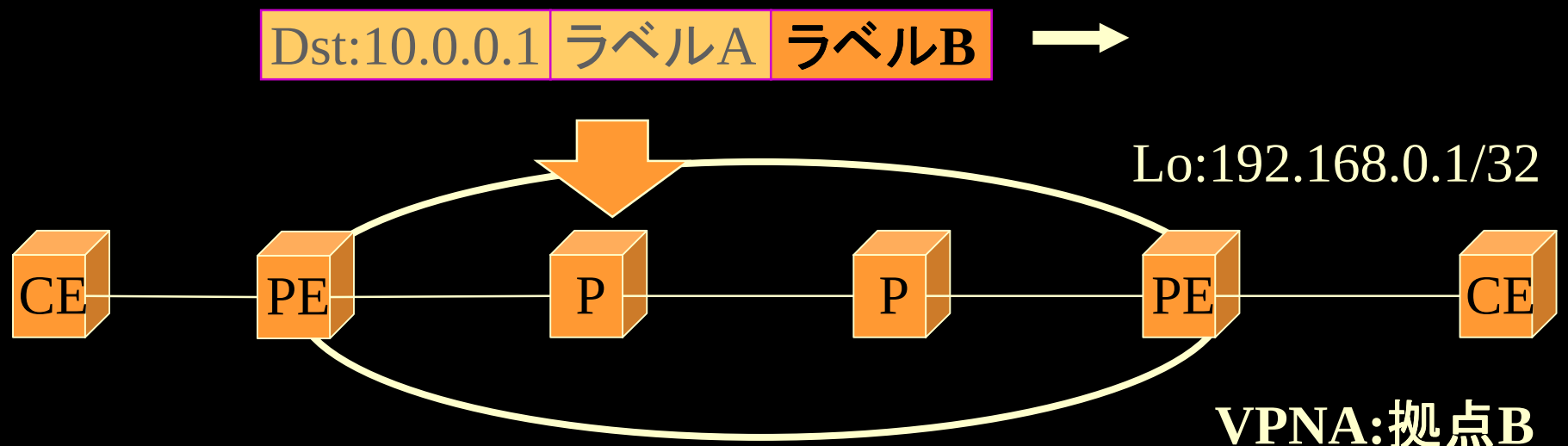
②(1)VPNA:10.0.0.0/8の出口のPEルータをBGP next-hop(192.168.0.1)で知る。

(2)この出口のPEルータに行くための転送用ラベルBを付与する。



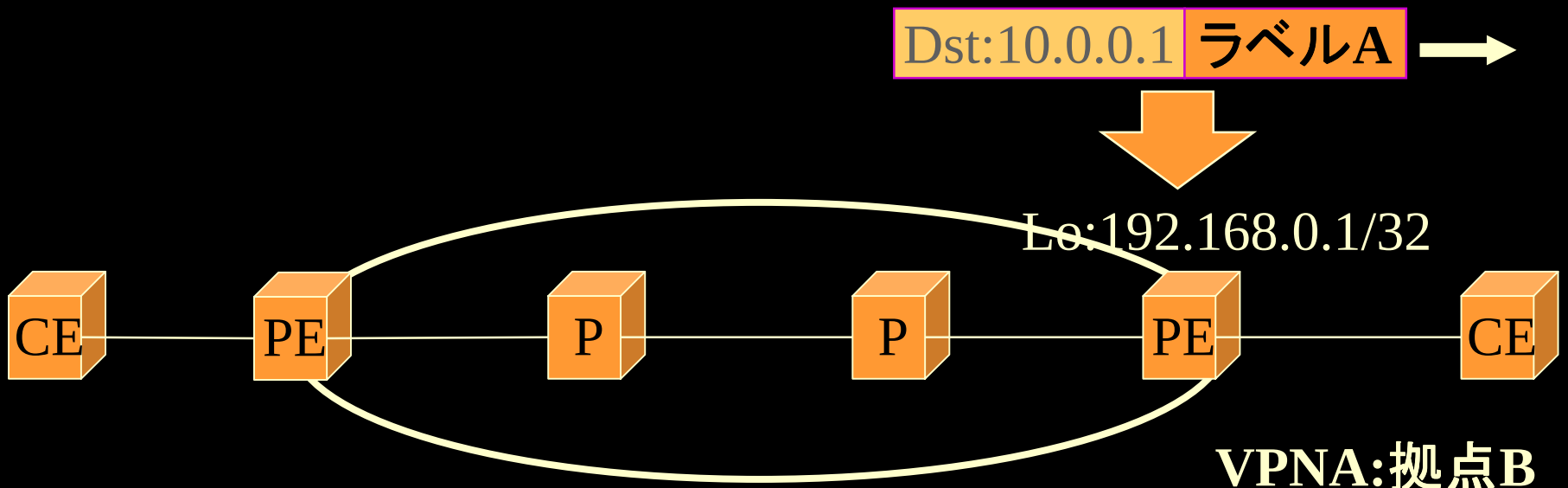
パケット転送の流れ(Cont.)

バックボーン内のPルータでは、転送用ラベルBだけを参照
※値はホップバイホップで変わります。



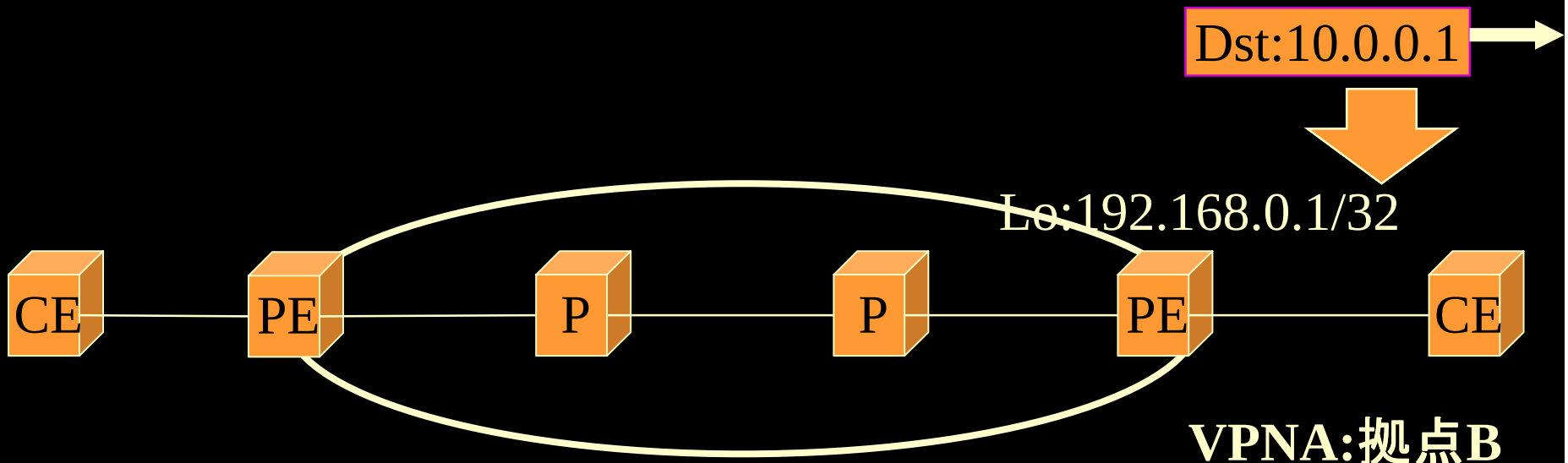
パケット転送の流れ(Cont.)

出口のPEルータでは、ラベルAの値を頼りにVPNを識別
＆出力インタフェースを決定しCEルータへパケットを転送



パケット転送の流れ(Cont.)

ラベルがはずされ通常のIPパケットとして
CEルータに到着！！



RFC 2547モデルの特長

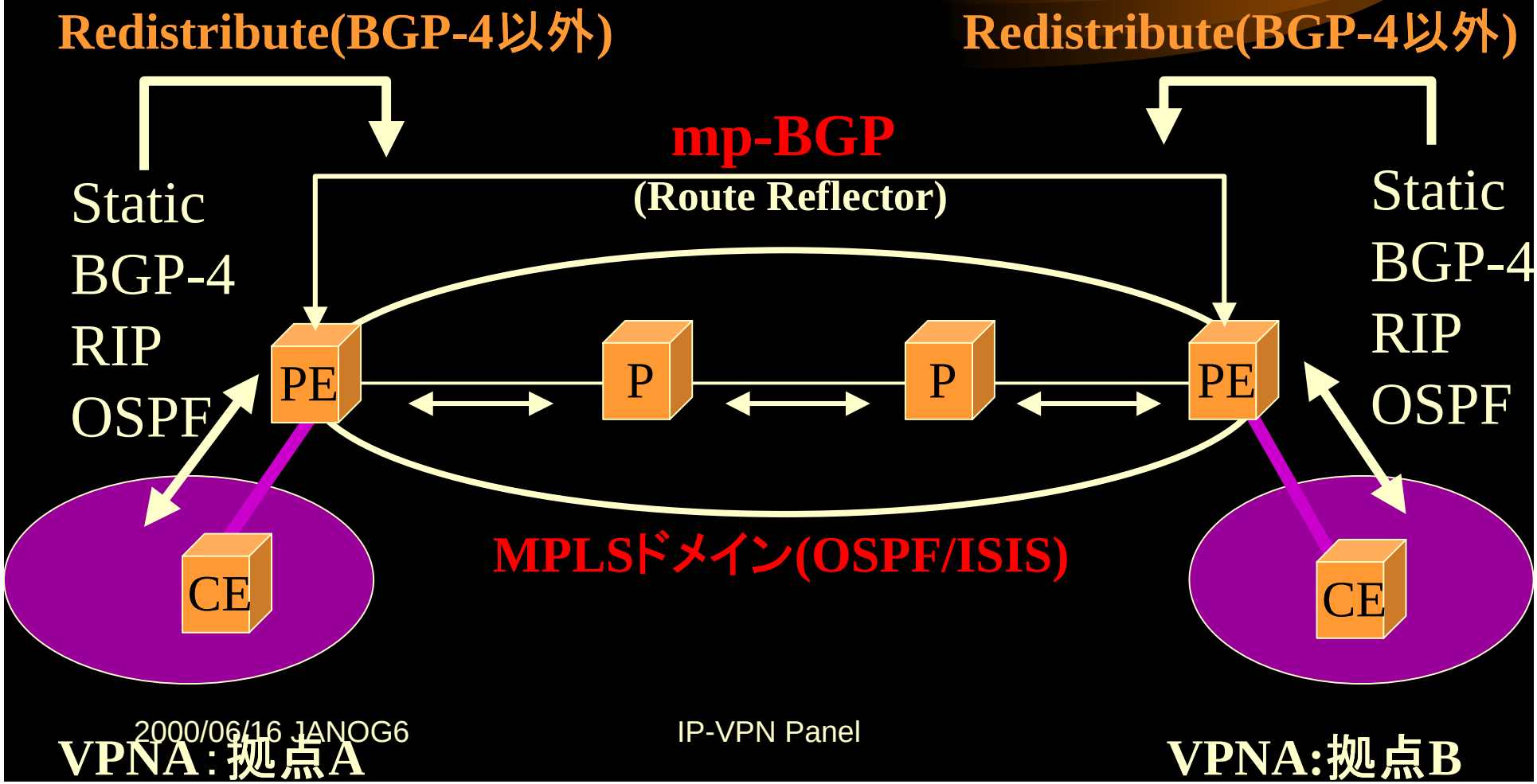
- CEルータは普通のルータで良い
  - パス管理もなく、セキュリティもFR/ATMなどと同等
- PEルータとCEルータの間は、複数のルーティングプロトコルが使える
  - Peerモデルではここがだいじ、サービスに直結
- 複数のVPNを1台のPEルータに収容可能
  - ISPサービスとしては収容効率はだいじ
- 異なるVPN間で同じアドレスが使える
  - みんなプライベートアドレスだし

PE-CE間ルーティングプロトコル

- Static
- BGP-4
  - Private ASNを使ういくつかの機能追加あり
- RIP(Version1, Version2)
  - 冗長構成時の経路ループに注意
- OSPF
  - multiple instance & デザインの制限に注意

どこまでサービスとして出すかはISPしだい

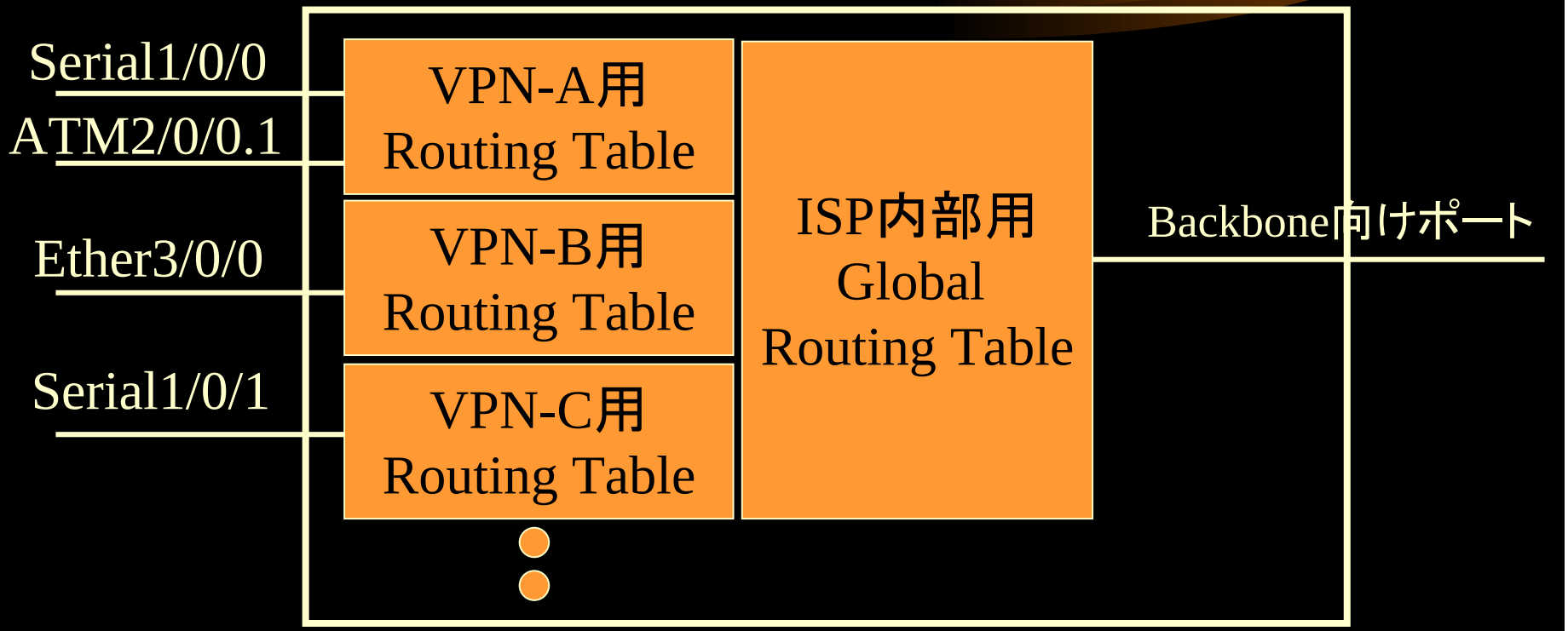
PE-CE間ルーティングプロトコル (Cont.)



複数VPNを収容する技術

- VRFs:VPN Routing and Forwarding tables
- VPNごとに異なるルーティングテーブルを持つ
  - 経路数の増大に注意！！
- 各々CEルータを接続するインタフェースを該当するVRF(VPN)に括りつける

複数VPNを収容する技術(Cont.)



異なるVPN間で同じアドレスが使える技術

- VPN-IPv4 Address Family
- 通常のIPv4アドレスに8byteの識別子Route Distinguisher(RD)を付与し、12byteのアドレス空間に拡大
- これにより同じ10.0.0.0でもVPNが違えば別経路として扱われる(mpBGPは1プロセス)。
- VPN-IPv4 Address(12byte)

  = RD(8byte)+IPv4(4byte)

異なるVPN間で同じアドレスが使える技術 (Cont.)

- RD(8byte)のFormat

  Type Field : 2-byte

  Value Field : 6-byte

- ISP間の識別も可能なValue Field Format

Type 0 = ASN(2-byte):任意の番号(4-byte)

例→9598:1(将来を考えるとこっち?!)

Type 1 = IP address(4-byte):任意の番号(2-byte)

例：192.168.0.1:1

異なるVPN間で同じアドレスが使える技術 (Cont.)

- RDが違えば別経路(Best決定プロセスも別)

redistribute
※next-hopは自分自身に
(next-hop-selfがdefault)

BGPテーブル
#show ip bgp vpnv4 all

RD:9598:1(VPN-A)
10.0.0.0/24 next-hop 192.168.0.1/32
10.0.1.0/24 next-hop 192.168.0.1/32
RD:9598:2(VPN-B)
10.0.0.0/24 next-hop 192.168.0.1/32
10.0.1.0/24 next-hop 192.168.0.1/32
RD:9598:3(VPN-C)
10.0.0.0/8 next-hop 192.168.0.1/32
.
.
.

RFC2547技術詳細

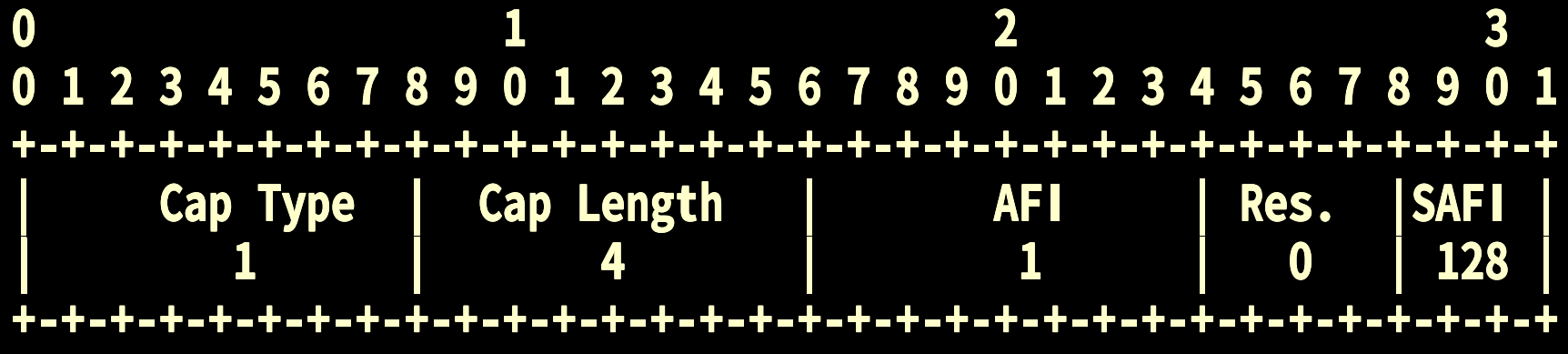
- VPN-IPv4アドレスってBGPで運べるの？
  - Multiprotocol Extensions for BGP-4
- 経路制御ってどこまでできるの？
  - AS-path, MED, LocalPref等の経路扱いはpure IPと同じ
  - Extended Communityの定義
- 直接所有していないVPN経路の扱いは？
  - Route Refresh Capability for BGP-4

Multiprotocol Extensions for BGP-4

- IPv6やマルチキャストと同じように RFC2283 Multiprotocol extensions for BGP-4を使用
- MP_REACH_NLRI(Type Code 14)
- MP_UNREACH_NLRI(Type Code 15)
- AFI=1 & SAFI =128
- MPLS-labeled VPN-IPv4 address

Multiprotocol Extensions for BGP-4(Cont.)

- Capabilities Advertisement with BGP-4
- MPLS-Labeled VPN-IPv4 addressを解釈できるかどうかをPeerを張る際に決定

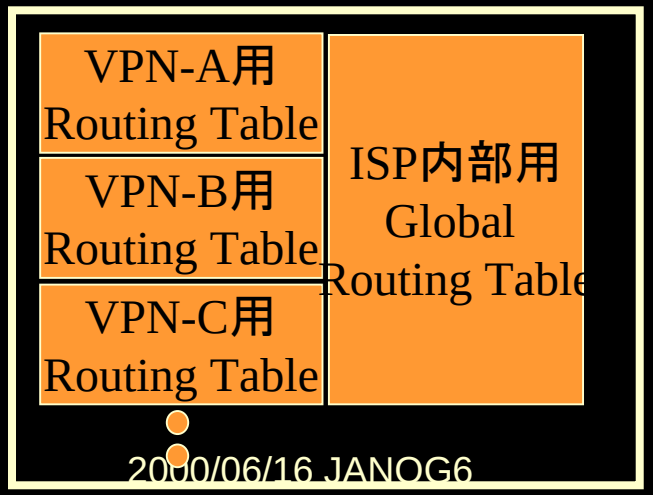


Extended Community

- Extended Community Attribute(Type Code 16)が新たに定義
- その中の一つがRoute Target(RT)
- VRFには必ず一つ以上のRTが付与される
- Export Targets:CEからの経路の付与
- Import Targets:他PEからの経路選択に使用
- RTを使ってVPN間通信、AS間通信の実現

Extended Community(Cont.)

RTをもとにVPNv4-prefixをどのVPNのRouting Table突っ込むかを選択(Import)



BGPテーブル

RD:9598:1(VPN-A)
10.0.0.0/24 RT:9598:1(Export)
10.0.1.0/24 RT:9598:1(Export)
RD:9598:2(VPN-B)
10.0.0.0/24 RT:9598:2(Export)
10.0.1.0/24 RT:9598:2(Export)
RD:9598:3(VPN-C)
10.0.0.0/8 RT:9598:3(Export)

Route Refresh Capability for BGP-4

- PEルータをスケールさせるための機能
- PEルータは直接自分に持ってくる必要が無いVPN経路は受け取らない。
- 新たにVPNが増えたときには、自動的に取りに行く。
- Route Reflectorは常にすべてのVPN経路情報を持っている必要がある。
  - VRFは持たないのでメモリは多少ラク

RFC2547の実際

- MPLS-VPNの枠組みはわかるが、細かい部分の規定がない(Informational)。
- よって実装はまだC社とあと一つだけ(^^;
  - もう少し広がってほしい。。
- Coreは軽くなったが、Edgeが重い。。
- 経路数が莫大に増える可能性
  - VPNにフルルートを持ちこまれたら。。。
  - 1VPN*1000経路×200VPN=20万経路！！

RFC2547の将来

- 現在draft-rosen-rfc2547bis-01.txtとしてupdate中
- まだまだC社中心だが技術を“Open”にする動きは見られる(authorにはJ社やCarrierの名前も。。)
- Carriers’ Carriers やInter-Provider BackbonesなどMPLS-VPNのISP相互接続(The Internetとどう共存するか。。）

RFC2547の将来(Cont.)

- MPLS-VPNとTE(Traffic Engineering)との組合せ
- IPSecなど他のIP-VPN方式との戦い。。

まとめ

- MPLS-VPN方式は、RFC2547に記述
- PeerモデルのISP提供VPN
- MPLSとmpBGPの組合せでVPNを実現
- VPNごとにルーティングテーブルを持つ
- RDを使ったVPN-IPv4アドレスへの拡張

References

- RFC2547/RFC2283
- draft-rosen-rfc2547bis-01.txt
- draft-ramachandra-bgp-ext-communities-03.txt
- draft-ietf-idr-bgp4-cap-neg-06.txt
- draft-ietf-idr-bgp-route-refresh-01.txt
- draft-ietf-idr-bgp4-multiporotocol-v2-05.txt
- draft-ietf-mpls-ldp-07.txt