IP-VPN
(BGP MPLS/VPN)

InternetWeek 2005

アジアネットコムジャパン株式会社
石井秀雄 (hishii@ancip.net)

IP-VPN

- MPLSで実現できる代表的なサービスとして、IPVPNを取り上げます。
- また、そのVPNの種類としては、L3VPN (Layer3VPN)およびL2VPN(Layer2 VPN)がありますが、ここでは、現在普及しているL3VPNを中心に説明します。

IP-VPN Agenda

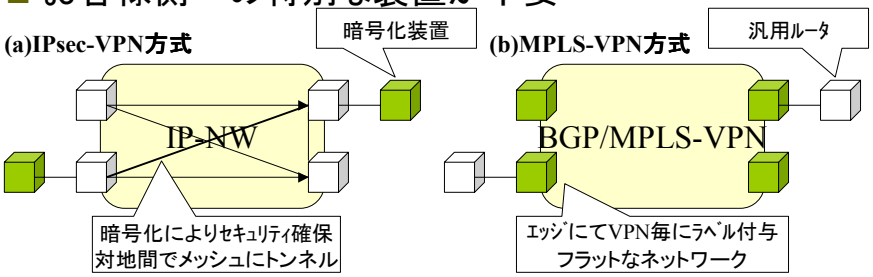
- BGP/MPLS-VPNとは
- BGP/MPLS-VPNの動作概要
- BGP/MPLS-VPNのラベルパス決定方法
- BGPにおけるVPN経路情報
- VPNにおけるQoSの提供
- BGP/MPLS-VPNネットワーク相互接続
- BGP/MPLS-VPNユーザ構築事例
- BGP/MPLS-VPNまとめ（実際と新技術）

BGP/MPLS-VPNとは

- RFC2547bisに記された通信事業者としてのIP-VPN実現技術
- インターネットVPN ＝ オープンなネットワーク上でIPデータ部を暗号化し、閉域ネットワークを実現
- MPLS-VPN ＝ MPLS label（ラベルヘッダをつかってカプセリング）により、理論的な閉域ネットワークを実現
- 昨今、MPLSを使ったほかのIP-VPN技術と区別してBGP/MPLS-VPNと呼ばれる

BGP/MPLS-VPNとは

- 多様なインターフェンスタイプでの提供が可
  - ATM—SDHといいた非対称構成も可能
- 暗号に頼らないセキュリティの確保が可能
  - Frame-Relayなどと同等の機能をIP網で実現
- お客様側への特別な装置が不要

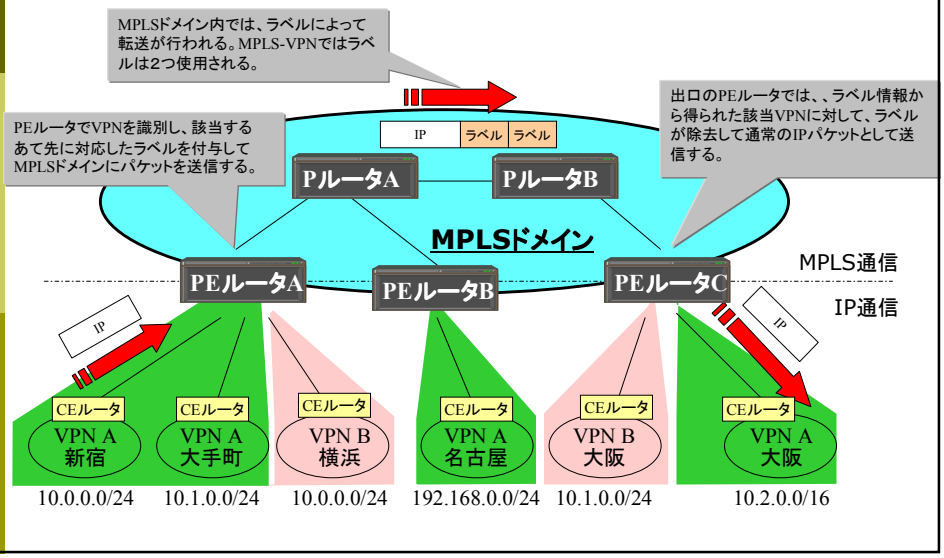


BGP/MPLS-VPNとは

- 網内のパケット転送にMPLS(LDP/RSVP)、VPN経路情報交換にBGP(multiprotocol-BGP:RFC2858, RFC3107)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
- プロバイダ側ルータで、VPNごとに異なるルーティングテーブル持ち、ユーザ側ルータとルーティング情報の交換をする。
- Layer3ルーティングをプロバイダにアウトソーシングしているといえる

BGP/MPLS-VPNの動作概要

BGP/MPLS-VPN動作概要



特徴（ユーザ側）

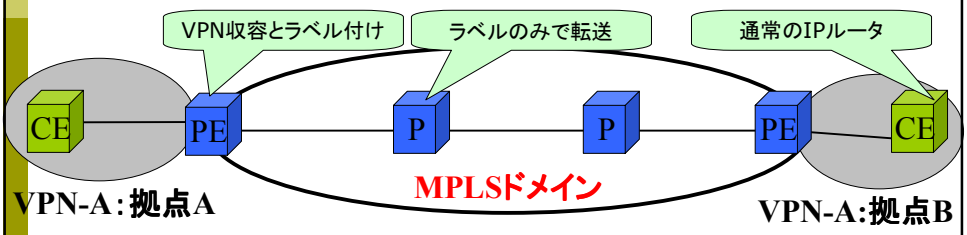
- お客様宅に設置されるルータは通常のIPルータで良い(MPLSやIP-Sec等の機能はいらない)
- FRやATM等のようなパスの管理が必要ない（利用して、理論的に2つのパスを確立することも可能）
- IPアドレスはお客様にて任意に設定可能でありIPv4プライベートアドレスを自由に持ちこめる。
- VPN同士の通信は、ルータ内及び網内にて完全に分離されておりFR、ATMと同等のセキュリティが保たれている。

特徴(プロバイダ側）

- 既存のルータによるIPネットワークをそのまま使ってIP-VPNサービスを提供できる。
- 複数のルーティングプロトコルを使ってお客様を収容できるので柔軟なサービスが提供できる。
- 複数のVPNを1台のルータに収容できるため効率の良いIP-VPNサービスを提供できる。
- 異なるVPN間で同じアドレスが使えるためサービス性が良い
- 論理的に分離されたネットワークなのでQoSなどのサービスも実現しやすい。

BGP/MPLS-VPN構成ルータ

- **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)**

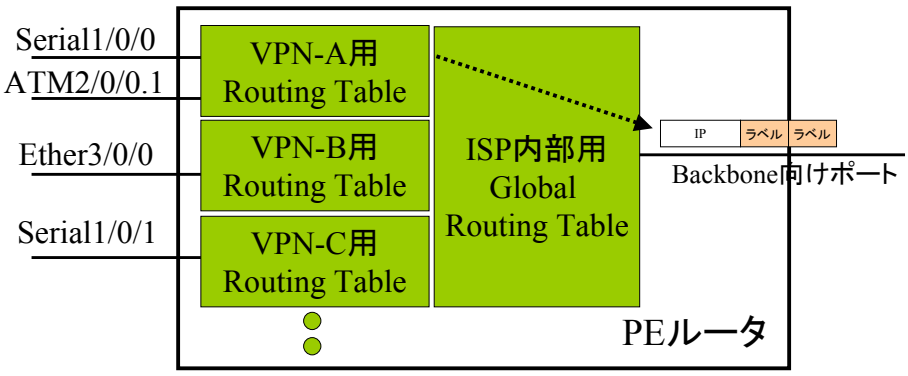


PEルータの仕組み

- 複数のVPNを1台のPEルータに収容するために
  - VPNごとに異なるVRF (VPN Routing and Forwarding table)をもつ
  - それぞれのCEルータを接続するインターフェイスを該当するVRFに括り付ける
  - 個々のVRFはルータ内部で分離されており、またバックボーンでは、パケットがラベルでカプセリングされて転送されるので、ATM/FRと同等レベルのセキュリティが確保できる

PEルータの仕組み

- VPNごとに別々のルーティングテーブルを保持



BGP/MPLS-VPNラベル構成

- Shimヘッダ形式

PEルータで挿入され、出口のPEルータを目指してPルータをホップするたびにラベルの値は変わっていく(LDPでhop by hopに決定)

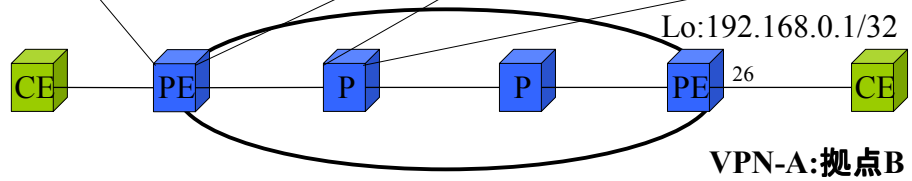
PEルータで挿入され、出口のPEルータに到着するまでは、コアネットワーク内では参照されず値も変わらない。(mpBGPでPEルータ同士で情報交換)

  - 32bit固定長MPLSラベルが2つ (MPLS label dual stack)

BGP/MPLS-VPN動作概要

VPN名	Route Dist.	あて先アドレス	VPN識別ラベル	出口のPEルータのアドレス	転送用ラベル
A社	12	10.0.0.0/8	26	192.168.0.1/32	㊷
A社	12	11.0.0.0/8	989	192.168.0.1/32	42

in転送用ラベル	出口のPEルータのアドレス	out転送用ラベル
㊷	192.168.0.1/32	32

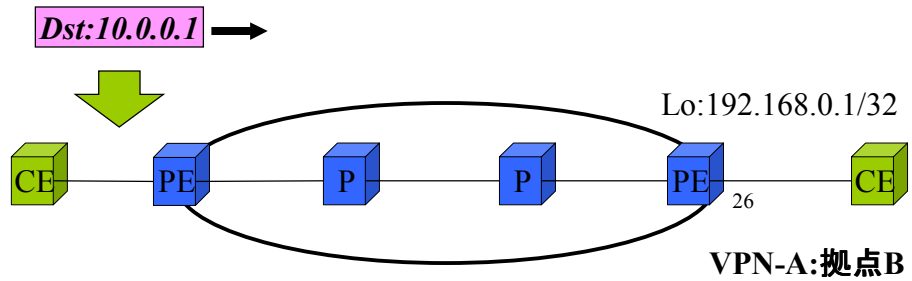


VPN-A:拠点A　　MPLSドメイン　　VPN-A:拠点B 10.0.0.0/8

BGP/MPLS-VPN動作概要(cont.)

- パケット転送　：　**CEからパケット到着**

VPNA;拠点B:10.0.0.1行きパケット到着

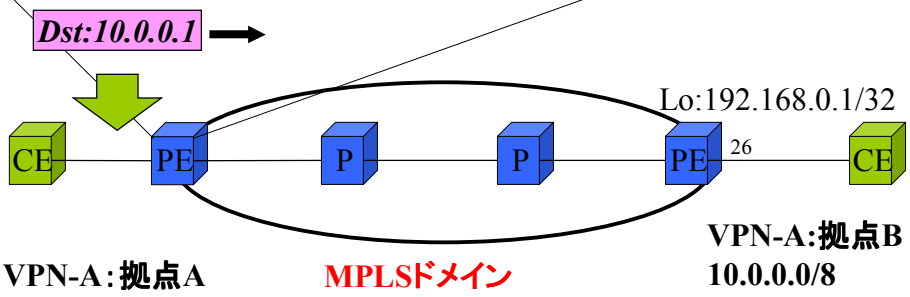


VPN-A:拠点A　　MPLSドメイン　　VPN-A:拠点B 10.0.0.0/8

BGP/MPLS-VPN動作概要(cont.)

- **PEルータでのラベルテーブルルックアップ**

VPN名	Route Dist.	あて先アドレス	VPN識別ラベル	出口のPEルータのアドレス	転送用ラベル
A社	12	10.0.0.0/8	26	192.168.0.1/32	㊷
A社	12	11.0.0.0/8	989	192.168.0.1/32	42

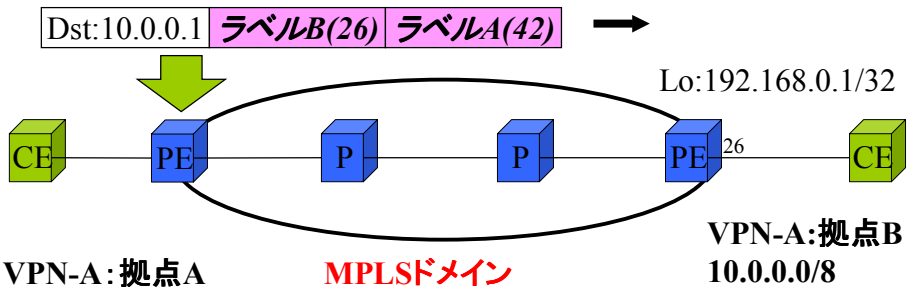


VPN-A:拠点A　　MPLSドメイン　　VPN-A:拠点B 10.0.0.0/8

BGP/MPLS-VPN動作概要(cont.)

- **PEルータでのパケットへのラベル付与**

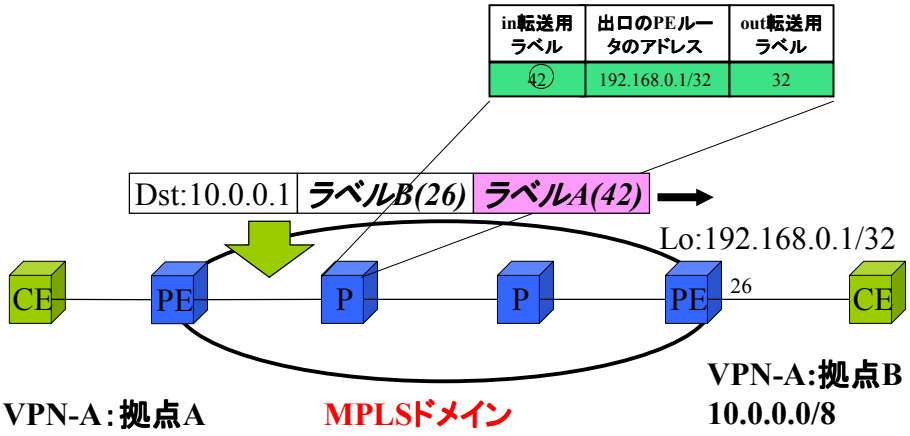
①出口のPEルータより得たVPNA:10.0.0.0/8に相当するVPN識別用ラベルBを付与する。
②(1)VPNA:10.0.0.0/8の出口のPEルータをBGP next-hopで知る。
(2)該当するBGP next-hopに対応した転送用ラベルAを付与する。



VPN-A:拠点A　　MPLSドメイン　　VPN-A:拠点B 10.0.0.0/8

BGP/MPLS-VPN動作概要(cont.)

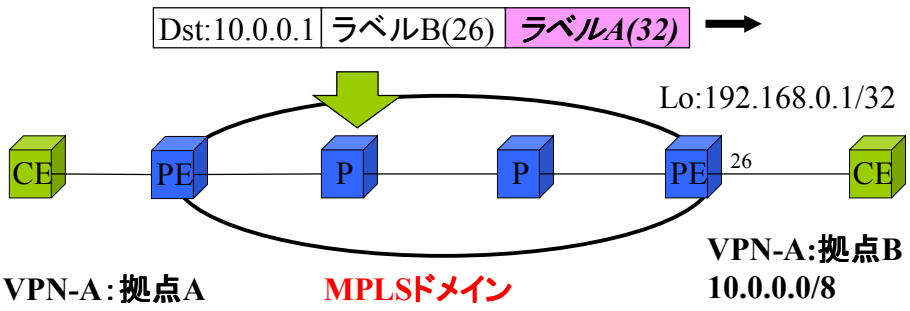
- **Pルータでのラベルテーブルルックアップ**



BGP/MPLS-VPN動作概要(cont.)

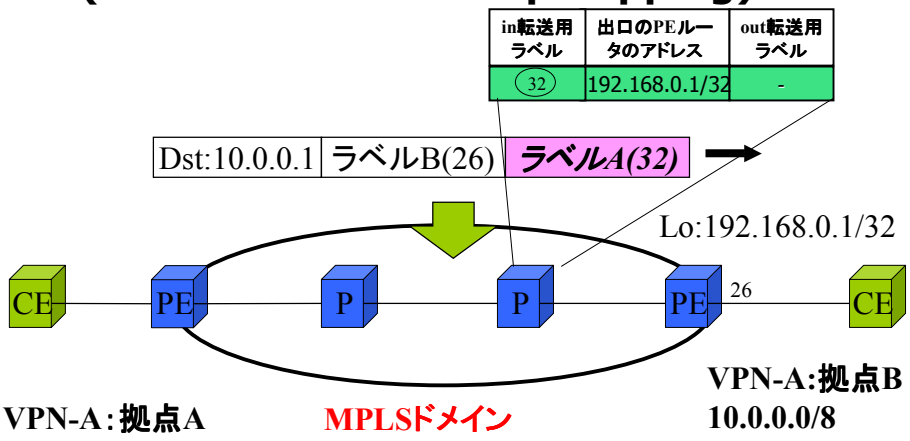
- **Pルータでのラベルスワップ**

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



BGP/MPLS-VPN動作概要(cont.)

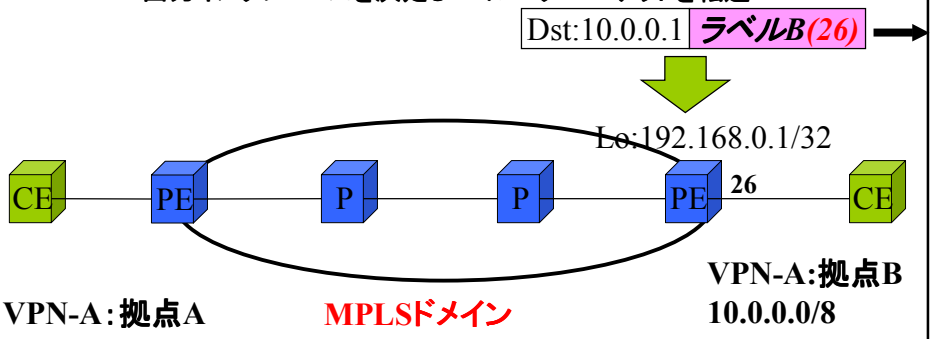
- 最後のPルータでは転送用のラベルを取ります
  (PHP:Penultimate Hop Popping)



BGP/MPLS-VPN動作概要(cont.)

- 最終PEルータでのラベルテーブルのルックアップ

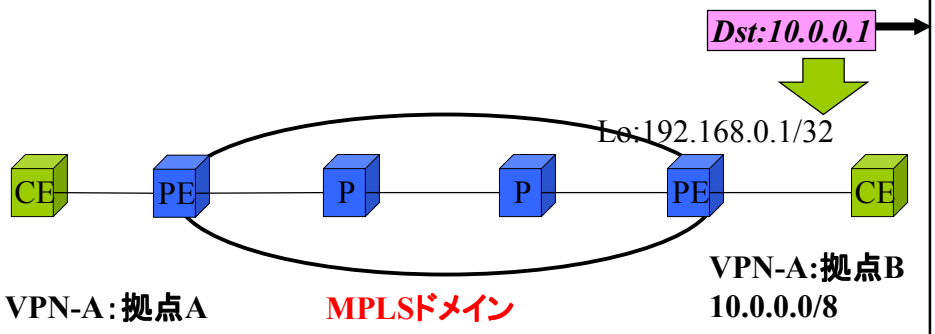
出口のPEルータでは、ラベルBの値を頼りにVPNを識別
＆出力インタフェースを決定しCEルータへパケットを転送



BGP/MPLS-VPN動作概要(cont.)

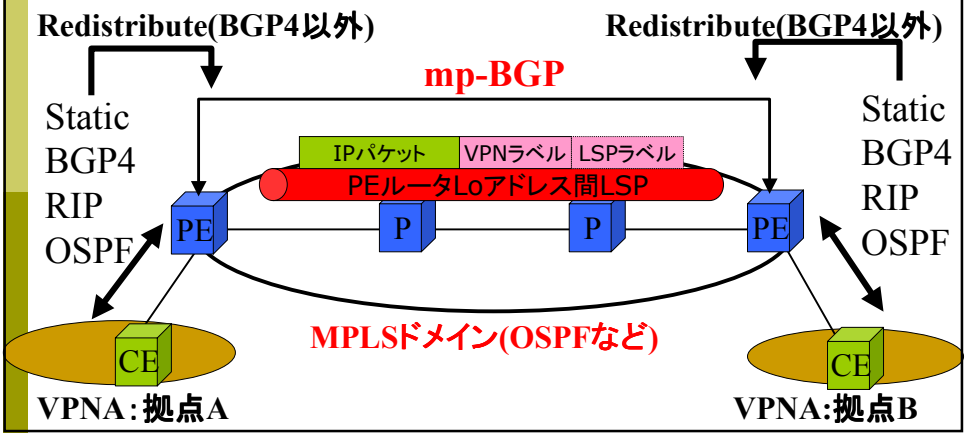
- 目的のCEルータへ到着

  ラベルがはずされ通常のIPパケットとして
  CEルータに到着する



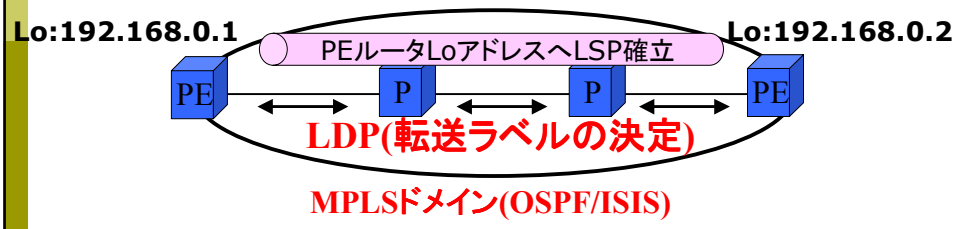
BGP/MPLS-VPN動作概要

- PEルータ間のLSPをVPN識別用ラベルでカプセル化されたパケットが通るイメージ



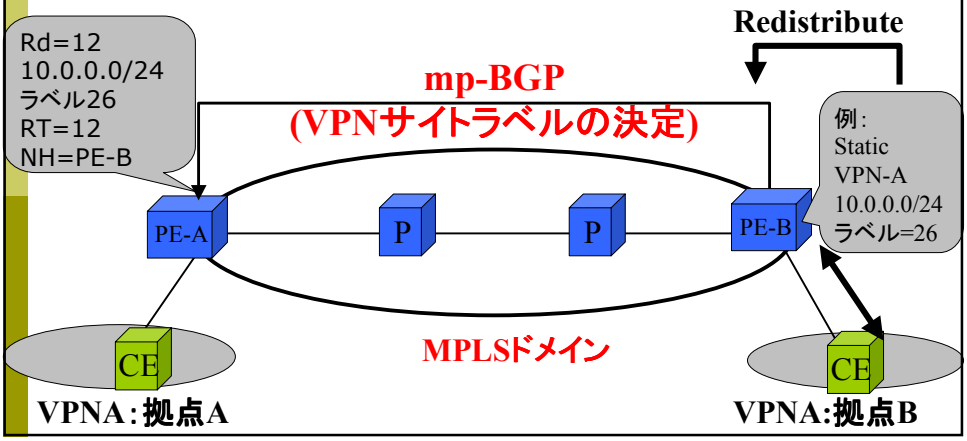
BGP/MPLS-VPNラベル決定方法

- PEルータ・Pルータ間でOSPF/ISISにて経路のやり取りをし、その経路情報にラベル情報を対応(LDP/RSVP-TE)
- 特にPEルータのLoopbackアドレスが最終的にVPNの出口を示すので重要



BGP/MPLS-VPNラベル決定方法

- PE-CE間のルーティングプロトコルで得たVPN経路情報をラベルの情報とともにPEルータ間で交換



BGP/MPLS-VPNラベル決定方法

- 個々のVPNを識別するためのラベルはmp-BGPを使ってPE間で交換される。（VPNラベル）
- MPLSドメイン内にある、PE-P,P-P間で使用されるラベルは、LDP、もしくはRSVP-TEでアサインされる。（転送ラベル）

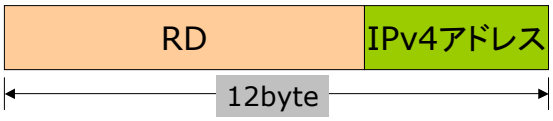
BGPにおけるVPN経路情報

BGPにおけるVPN経路

- RFC2858 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
- ラベル情報は、RFC3107に従ってEncoding

BGPにおけるVPN経路

- mp-BGPにおける経路扱い
  - 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
  - VPN-IPv4 Address(12byte)
    ＝RD(8byte)+IPv4(4byte)



BGPにおけるVPN経路

- mp-BGPにおける経路扱い
  - RD(8byte)のFormat

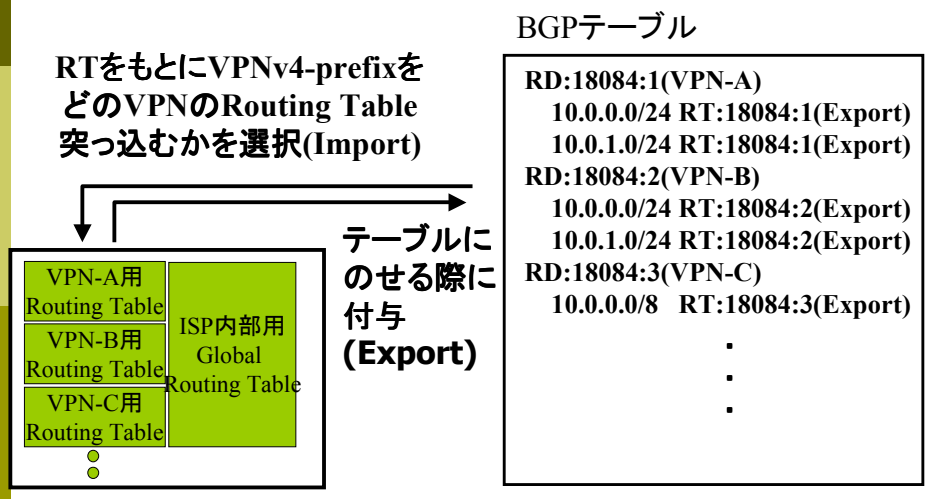
  - ISP間の識別も可能なValue Field Format

Type 0 = ASN(2-byte):任意の番号(4-byte)
例：65530:1
Type 1 = IP address(4-byte):任意の番号(2-byte)
例：192.168.0.1:1

Extended Community

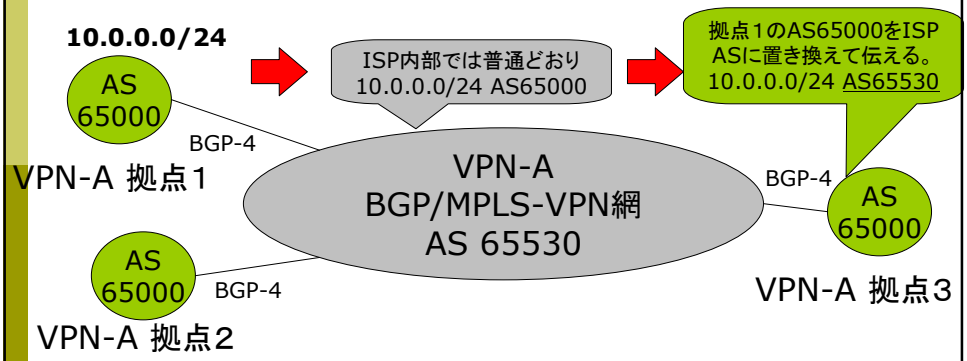
- Extended Community Attribute(Type Code 16)が新たに定義
- その中の一つがRoute Target(RT)
- VRFよりBGPにアナウンスされる経路には、必ず一つ以上のRTを付与する(Export Targets)
- リモートPEからの経路をローカルVRFに落とし込む際の選択に使用(Import Targets)
- VPN間通信、AS間通信の実現

Extended Community



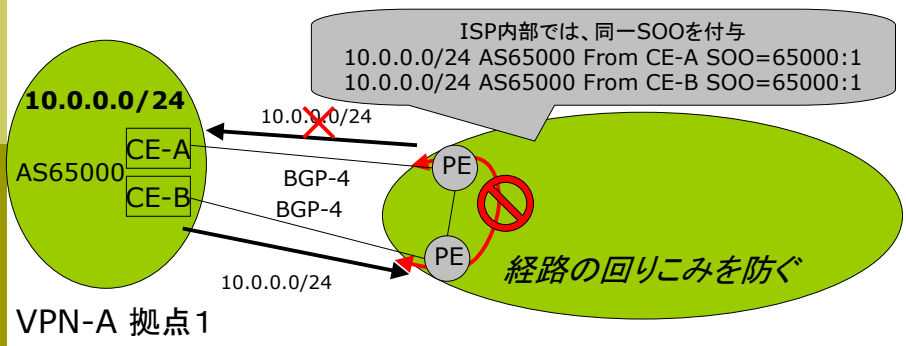
AS override

- 同一VPN内で複数の拠点で同一のAS番号を用いてPE-CE間を接続するための技術
- ユーザ側でAS番号の管理が不要



SOO (Site of origin)

- **AS Overrideと併用され冗長構成拠点の同一AS間のループを防ぐ**
- **RTと同じExtend Communityの一つ**



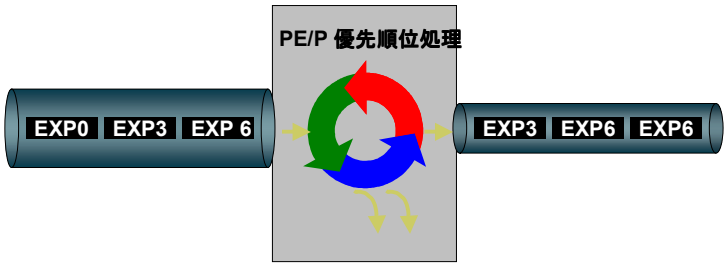
VPNにおけるQoSの提供

VPNにおけるQoSの提供

- 現在、VPNサービスの付加価値としてQoSの提供が進んでいる。
- JitterやdelayにVoIPやテレビ会議、画像のリアルタイム転送などのアプリケーションをVPNに統合したい、という要求

VPNにおけるQoSの提供

- MPLS ヘッダーのEXPフィールドを使ってClassわけを行い、すべてのP/PEで優先順位に基づいてパケットフォワーディングを行う



WRED/WFQ の処理によって、場合によっては低いプライオリティのパケットは廃棄される。

WRED:Weighted Random Early Detection
WFQ : Weighted Fair Queuing

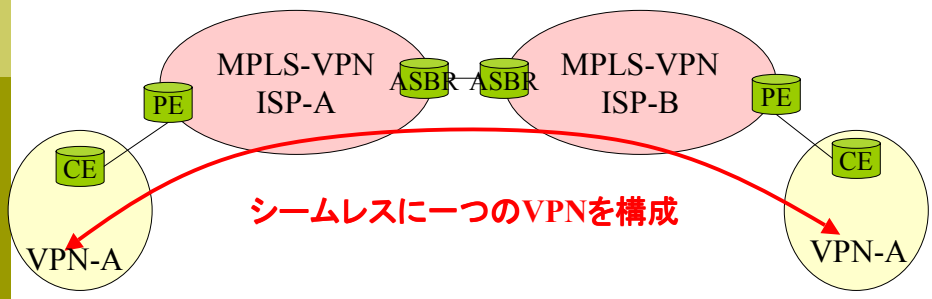
VPNにおけるQoSの提供

- サービス提供者の管理体制
  - SNMPの情報でQueueの使用率や状況を確認
  - SNMPの情報を使って、Ingress/Egressのポートの状況を管理
  - SNMPの情報を使って、バックボーンの回線利用率を管理
  - SAA (Service Assurance Agent)をつかって、POP-POP,またはEND-to-ENDの品質を管理

BGP/MPLS-VPNネットワーク相互接続

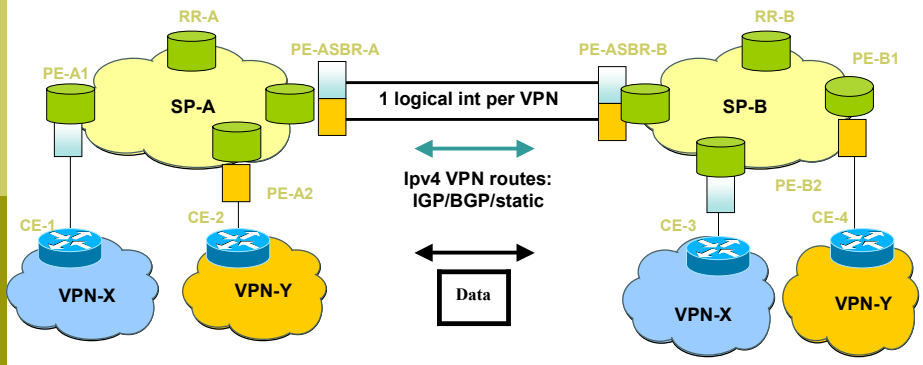
BGP/MPLS-VPN網間接続

- ネットワークの規模を大きくして、サービス提供範囲を広げるために、VPN網の相互接続が行われている。
  - RFC2547bis(draft-ietf-l3vpn-rfc2547bis-03.txt)において、3つのIPVPN AS間の相互接続の方法(option)が記載されている。



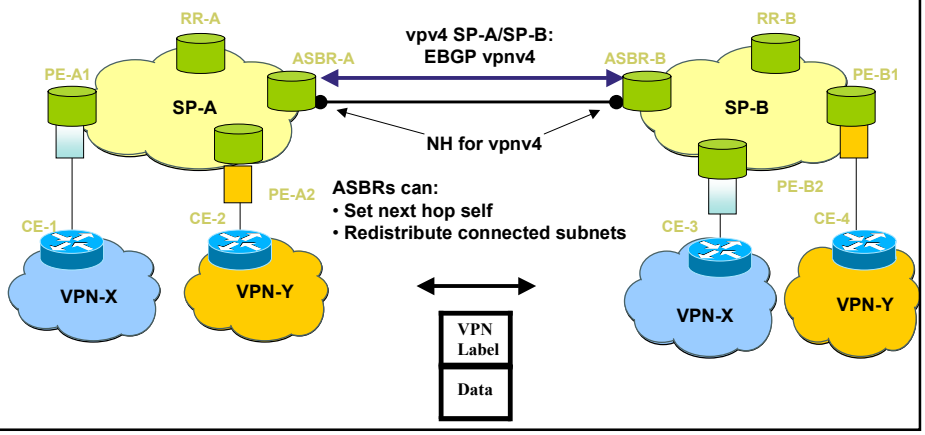
Option A

- VRF to VRF



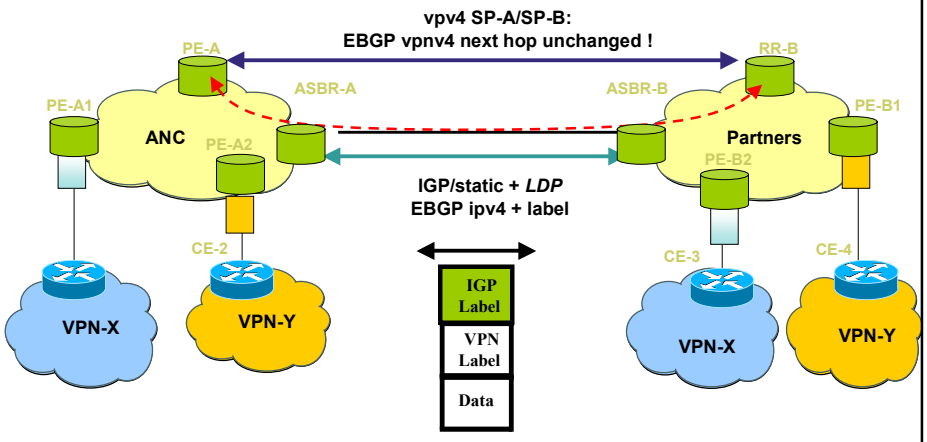
Option B

- ASBR to ASBR



Option C

- Route-Reflector to Route-Reflector

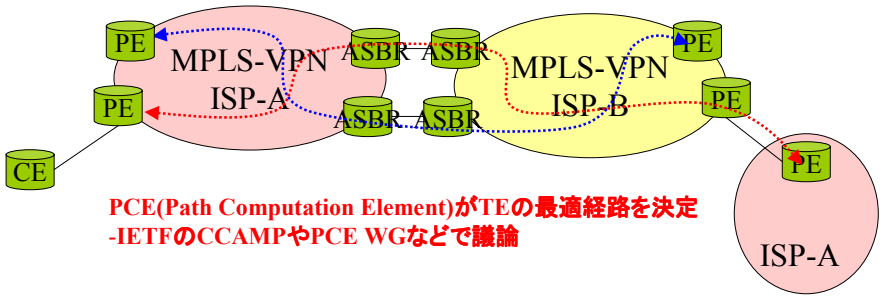


Optionの選択

- Option A (VRF to VRF)
  - NNIのインターフェイスでMPLSを使う必要がない。
  - TCP/IPのToSビットなどの書き換えが可能
  - 論理インターフェイス単位でトラフィック情報収集可能
- Option B (ASBR to ASBR)
  - PE-ASBR-ASBR-PEでMPLSで統一
  - 論理インターフェイスの管理不要
- Option C (RR to RR)
  - 経路情報の増大に対応
  - Plus Option Bの利点

Inter-AS Traffic Engineering

- VPN網の相互接続環境において、より一層の管理体制や品質確保のため、MPLS traffic-EngineeringのLSPをNNIパートナー網内のPE、もしくは網を通過し、出先の自社PEに構築する手法が議論されている。
- ASBR間においてのFastReRoute機能の要望



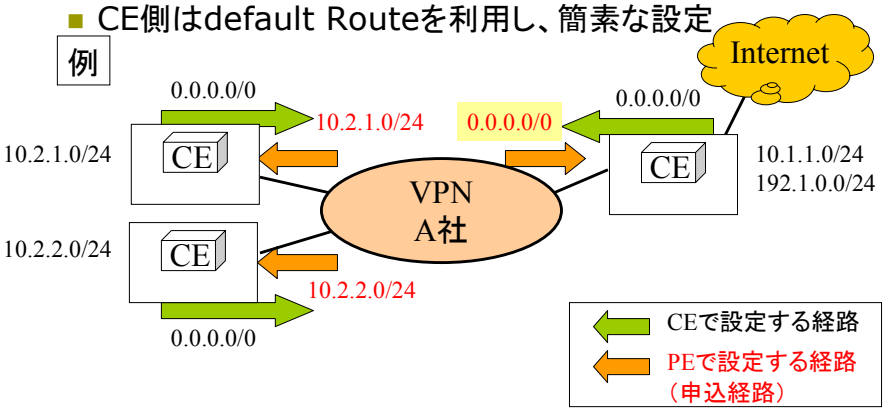
PCE(Path Computation Element)がTEの最適経路を決定
-IETFのCCAMPやPCE WGなどで議論

BGP/MPLS-VPNユーザ構築事例

BGP/MPLS-VPNユーザ構築事例

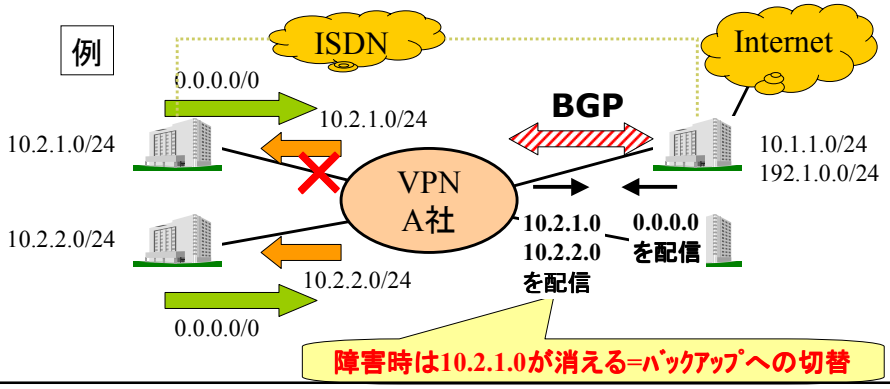
- Staticによるネットワーク構築
  - 主に拠点向き
  - CE側はdefault Routeを利用し、簡素な設定

例



BGP/MPLS-VPNユーザ構築事例

- BGPやOSPFなどを利用する
  - 動的ルーティングを生かしたバックアップ構成の実現
  - 大規模な社内網をサポート



BGP/MPLS-VPNのまとめ（実際と新技術）

BGP/MPLS-VPN技術の実際

- RFC2547bisがInformational RFCから、draft-ietf-l3vpn-rfc2547bis-03.txtにて改定中
- 現在、IETFのL3VPN-WGにおいて議論されている
- 複数のメーカが実装
- 同じMPLS技術をつかったVPNとして、VR(Virtual Router)方式や、ルーティングをアウトソースしないL2VPN(Layer2VPN)、VPLSなど新しいサービス形態がどんどん現れ、IP-VPNを実現するための唯一の解ではなくなっている。

BGP/MPLS-VPN技術の実際

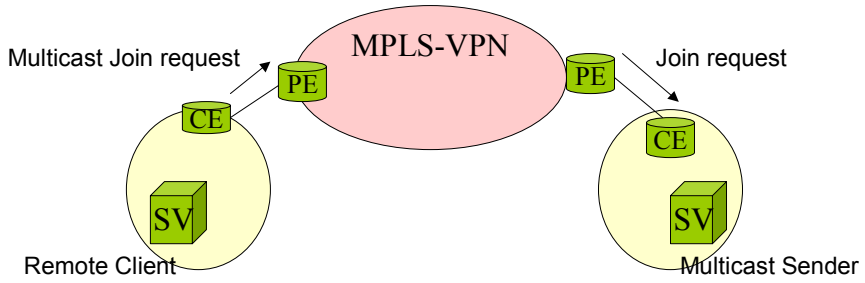
- ISP内部の設計に関しては、バックボーンは軽くなったが、エッジルータはVPNをハンドルするため負荷がかかる傾向
- ISPにてルーティングのアウトソーシングを受けるためISPとしては、経路数が莫大に増える可能性
  - 1VPN*1000経路×200VPN=20万経路!!
    - InternetのGlobalRouteでも16～17万経路
  - リフレクタを分ける、PEルータ収容を分ける、BGP Peer構成を分ける等のスケーラビリティ対応要

BGP/MPLS-VPN技術の実際

- BGP/MPLS-VPNに対して、高い品質を要求されている。
  - バックボーンやコアノードの障害時に、短時間でbackupへ迂回させるFast ReRoute機能をISP網内に導入
  - 障害箇所の判別や、PE-PE間のパス、遅延時間などの測定のためのLSP-ping/LSP-tracerouteやVRF単位でのPING、など、IP同様の機能がVPNの網内でもサポートされている。

BGP/MPLS-VPN関連の新しい技術

- **Multicast over MPLS-VPN**
  - Multicast PIM-SMをVPNユーザ単位に分けで機能を提供
  - ユーザは個々に独立したMulticast網として利用可能
  - 複数のベンダーから、いくつかのソリューションが提案されている
  - 限定される設定や機能がまだ多い



BGP/MPLS-VPN関連の新しい技術

- IPv6 for IPVPN
  - IPv6をVPN網で利用したい要求に対して考えられた新しい技術
  - BGP-MPLS IP VPN extension for IPv6 VPNとして、draft-ietf-l3vpn-bgp-ipv6-07.txtで規定
  - IPv4のVPN同様に、IPv6にもRD（route-distinguisher)を付与する
  - 6PEは1つのGlobal IPv6 Routing tableをもつが、VPNの場合は複数のIPv6 VRF routing tableをもつ必要あり