

IP-VPN (BGP MPLS/VPN)

InternetWeek 2004

アジアネットコムジャパン株式会社
石井 秀雄
<hishii@asianetcom.net>

IP-VPN

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

IP-VPN Agenda

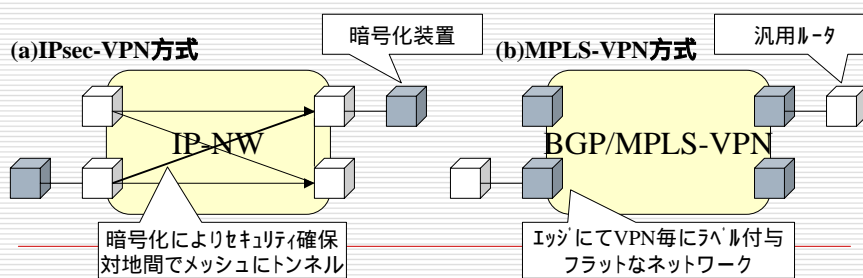
- BGP/MPLS-VPNとは
 - BGP/MPLS-VPNの動作概要
 - BGP/MPLS-VPNのラベルパス決定方法
 - BGPにおけるVPN経路情報
 - BGP/MPLS-VPN設定例
 - BGP/MPLS-VPNユーザ構築事例
 - BGP/MPLS-VPNまとめと新技術
-

BGP/MPLS-VPNとは

- RFC2547bisに記されたISPサービスとしての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データ部を暗号化で実現
 - MPLS-VPN = MPLS(ラベルによるカプセリング)により、論理的なクローズドなネットワークを実現
 - 昨今のMPLSを使った他のIP-VPN技術と区別してBGP/MPLS-VPNと呼ばれる。
-

BGP/MPLS-VPNとは

- ルータによる、多様なIFによる提供が可能(ATM~HSDなどの非対称構成も可能)
- 暗号に頼らないセキュリティの確保が可能(FRなどと同等の機能をIPネットワークで実現)
- お客様側への特別な装置が不要

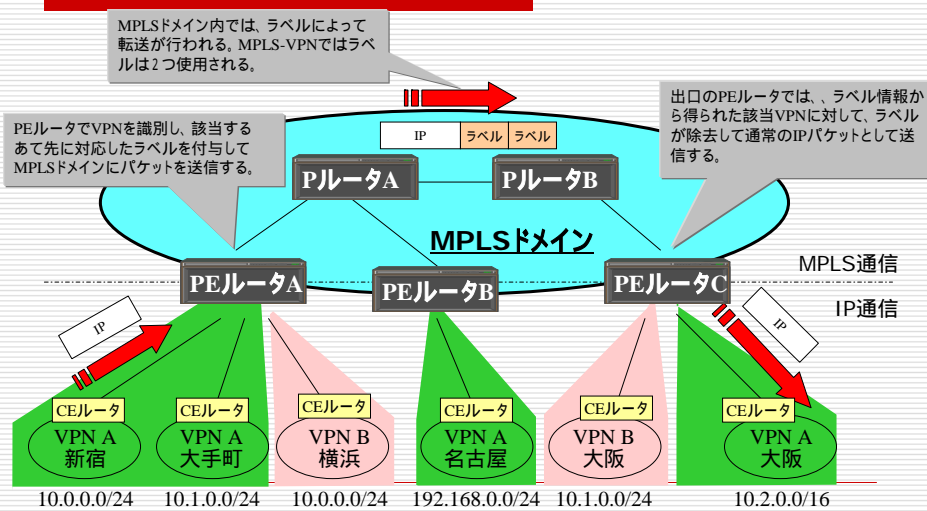


BGP/MPLS-VPNとは

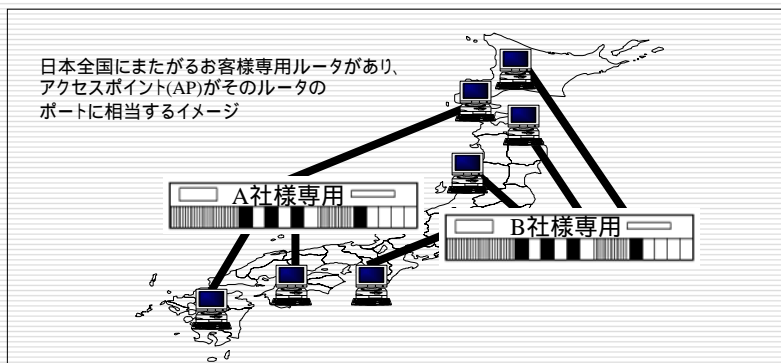
- 網内パケット転送にMPLS(LDP/RSVP)、VPN経路情報交換にBGP(mpBGP:RFC2858, RFC3107)を使用
- ルーティングプロトコルがエッジで終端されるPeerモデルのLayer3 IP-VPN
- VPNごとに異なるルーティングテーブルを持ちユーザルータとルーティング情報を交換する。
- Layer3ルーティングのISPへのアウトソーシング

BGP/MPLS-VPNの動作概要

BGP/MPLS-VPN動作概念



BGP/MPLS-VPN動作概念



- 日本全国にまたがるお客様専用ルータを提供するイメージとなる。複数のVPNでバックボーンを共用するが、お互いのVPNは論理的に独立している。

特徴(ユーザ側)

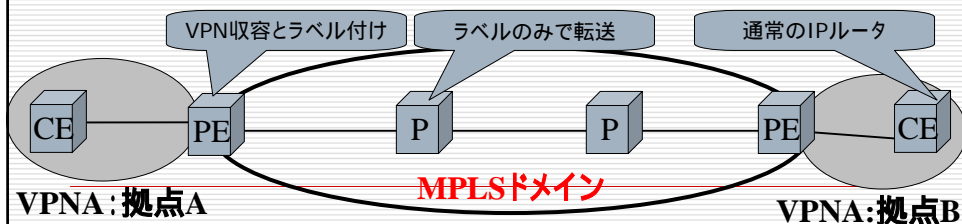
- お客様宅に設置されるルータは通常のIPルータで良い(MPLSやIP-Sec等の機能はいらぬ)
- FRやATM等のようなパスの管理が必要ない(利用して、理論的に2つのチャネルを確立することも可能)
- IPアドレスはお客様にて任意に設定可能でありIPv4プライベートアドレスを自由に持ちこめる。
- VPN同士の通信は、ルータ内及び網内にて完全に分離されておりFR、ATMと同等のセキュリティが保たれている。

特徴(ISP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)

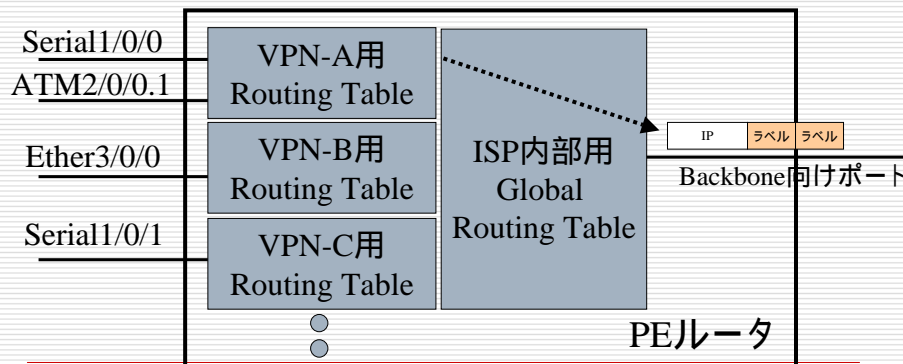


PEルータのしくみ

- 複数のVPNを1台のPEルータに収容するために
 - 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
 - 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
 - VRF同士はルータ内部で分離されており、またバックボーンには、ラベルでカプセルリングしてパケットを送出するので、ATM/FRと同等レベルのセキュリティが確保できる。

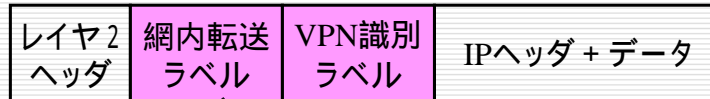
PEルータのしくみ

- VPNごとにルーティングテーブルを保持する。
- 一部の実装では、VR(Virtual Router)の場合も



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)

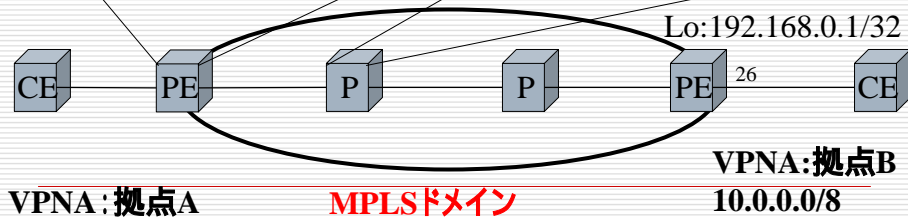
PEルータで挿入され、出口のPEルータに到着するまでは、コアネットワーク内では参照されず値も変わらない。(mpBGPでPEルータ同士で情報交換)

- MPLSラベルスタックを2つ使う
- 32bit固定長ラベル × 2

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

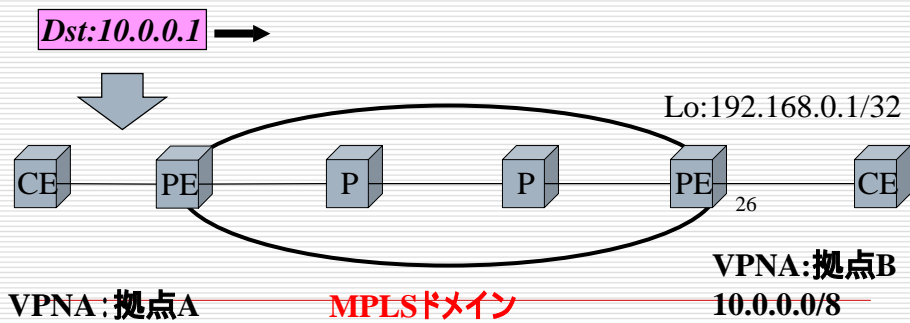
in転送用ラベル	出口のPEルータのアドレス	out転送用ラベル
42	192.168.0.1/32	32



BGP/MPLS-VPN動作概要(cont.)

□ パケット転送(CEルータからのパケット到着)

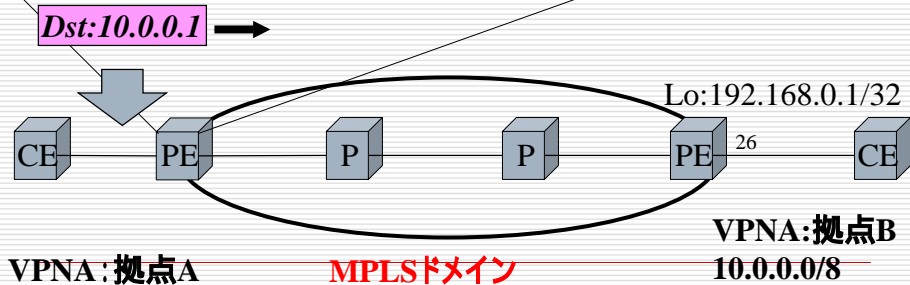
VPNA;拠点B:10.0.0.1行きパケット到着



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

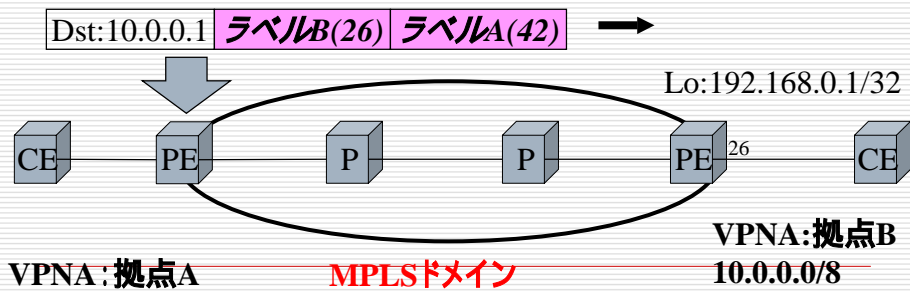


BGP/MPLS-VPN動作概要(cont.)

□ PEルータでのパケットへのラベル付与

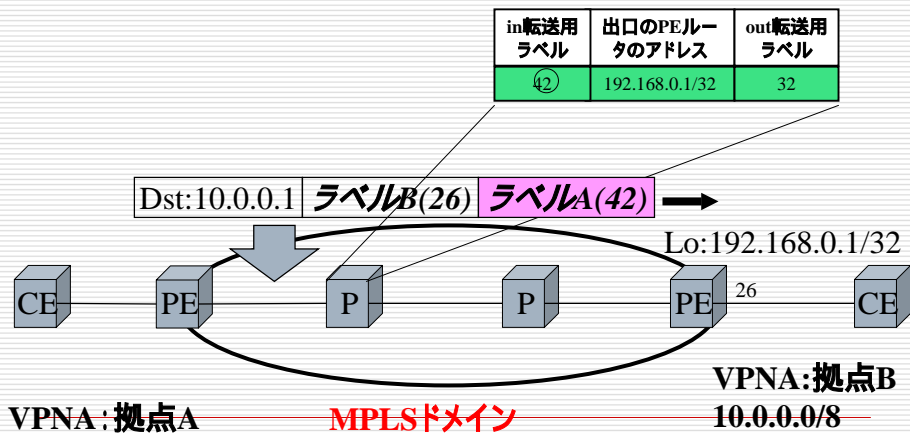
出口のPEルータより得たVPNA:10.0.0.0/8に相当するVPN識別用ラベルBを付与する。

- (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。



BGP/MPLS-VPN動作概要(cont.)

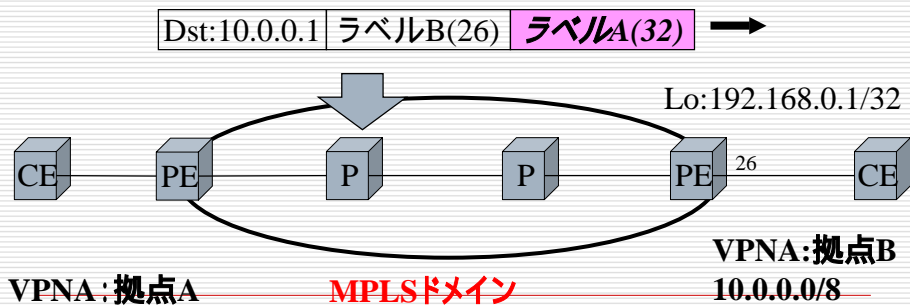
□ Pルータでのラベルテーブルルックアップ



BGP/MPLS-VPN動作概要(cont.)

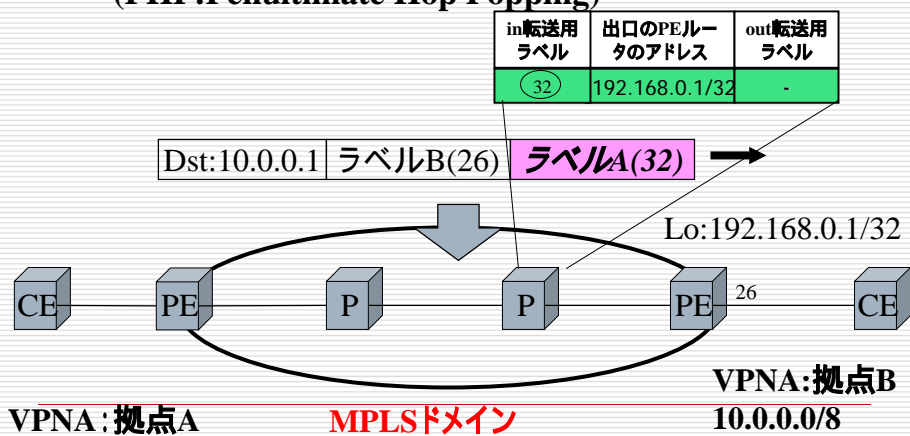
□ Pルーターでのラベルスワップ

バックボーン内のPルーターでは、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.)

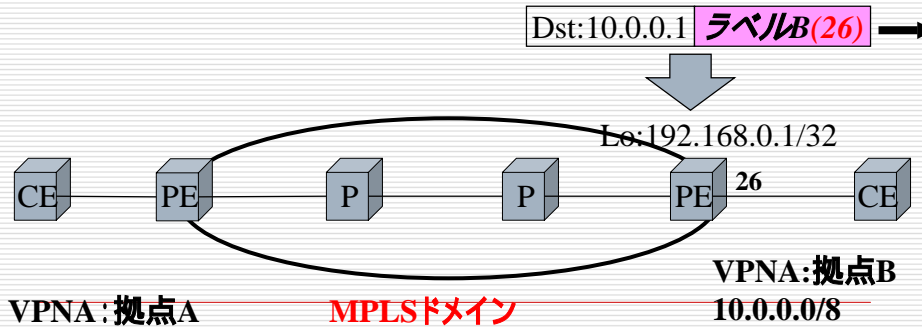
□ 最後のPルーターでは転送用のラベルを取ります (PHP:Penultimate Hop Popping)



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

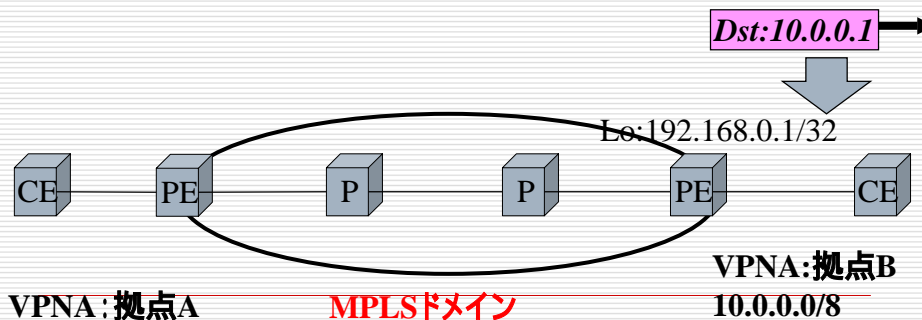
出口のPEルータでは、ラベルBの値を頼りにVPNを識別
& 出力インタフェースを決定しCEルータへパケットを転送



BGP/MPLS-VPN動作概要(cont.)

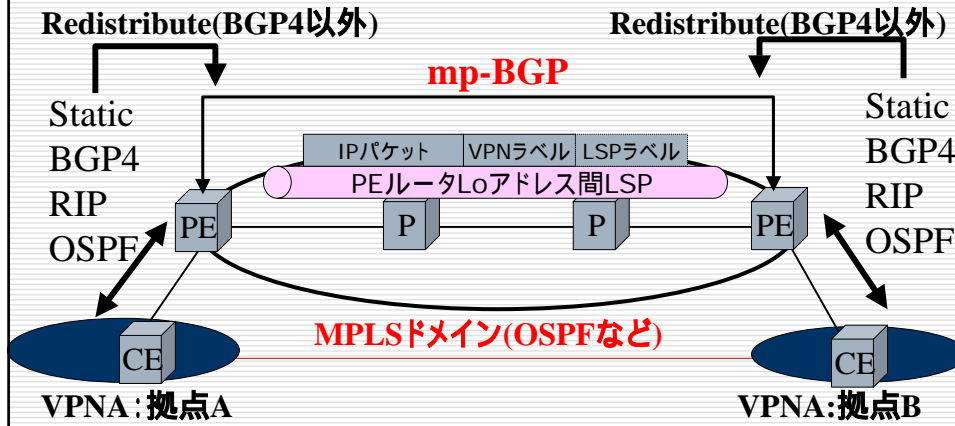
□ 目的のCEルータへ到着

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



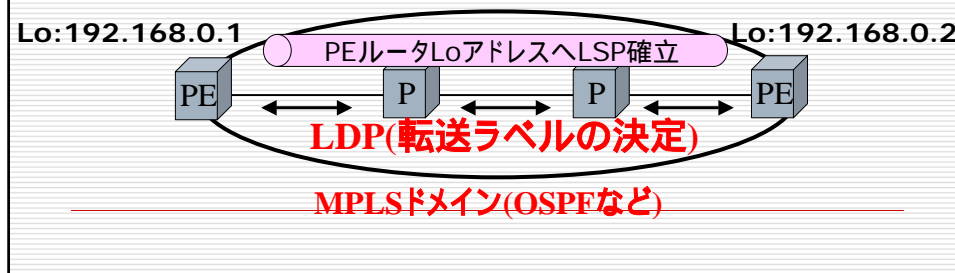
BGP/MPLS-VPNラベル決定方法

- 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



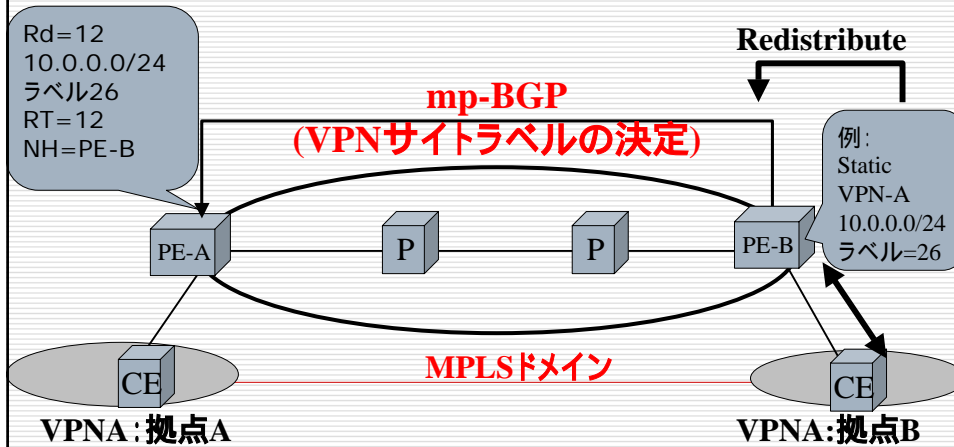
BGP/MPLS-VPNラベル決定方法

- 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

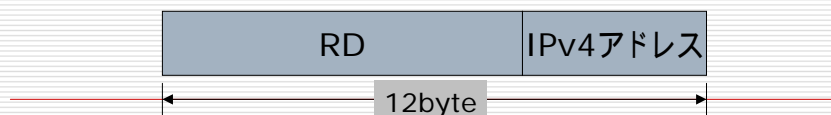
BGPにおけるVPN経路

- 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

Type	Value
2byte	6byte

■ 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: 18084:1

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

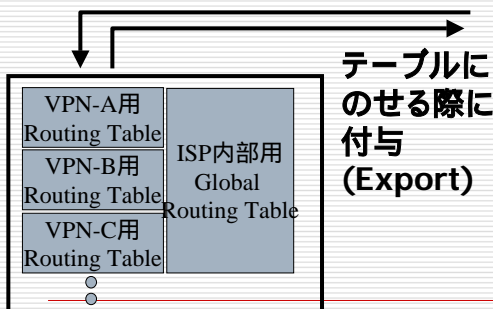
例: 192.168.0.1:1

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

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

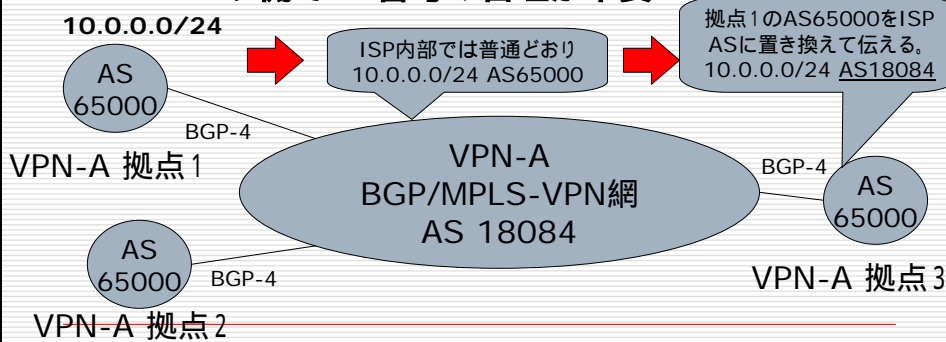


BGPテーブル

```
RD:18084:1(VPN-A)
 10.0.0.0/24 RT:18084:1(Export)
 10.0.1.0/24 RT:18084:1(Export)
RD:18084:2(VPN-B)
 10.0.0.0/24 RT:18084:2(Export)
 10.0.1.0/24 RT:18084:2(Export)
RD:18084:3(VPN-C)
 10.0.0.0/8 RT:18084:3(Export)
.
.
.
```

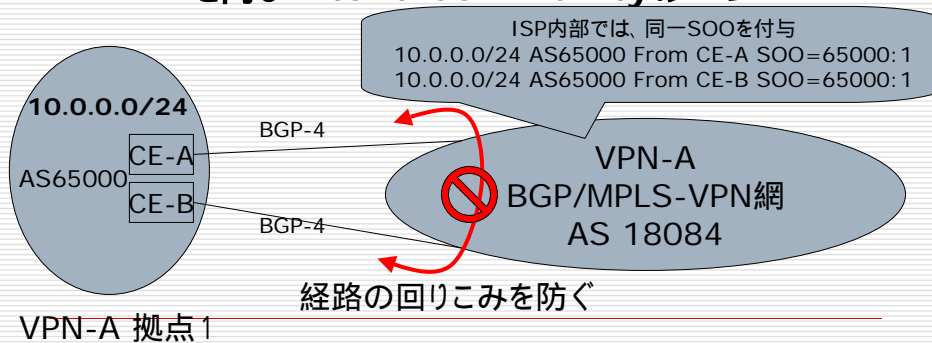
AS Override

- 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



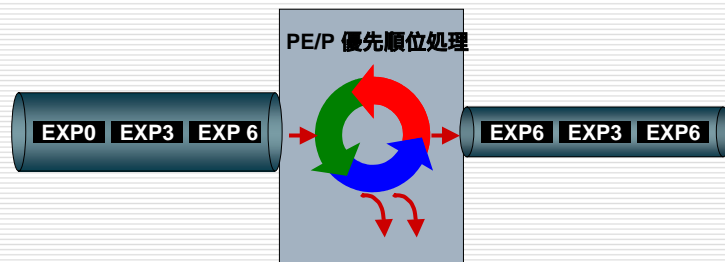
VPNにおけるQoSの提供

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に敏感な、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。
-

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

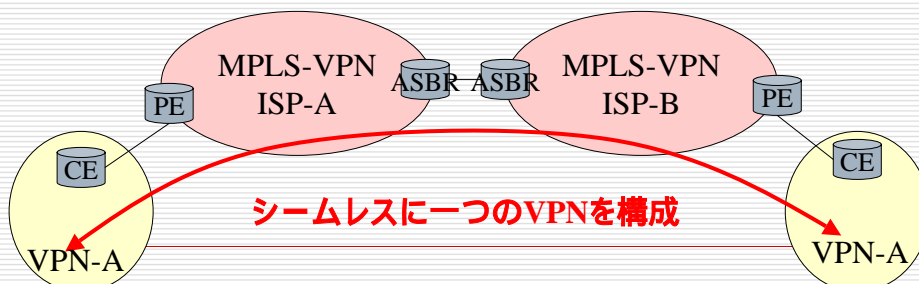
- サービス提供者の管理体制
 - SAA (Service Assurance Agent)をつかって、POP-POP,またはEND-to-ENDの品質を管理
 - SNMPの情報でQueueの使用率や状況を確認
 - SNMPの情報を使って、Ingress/Egressのポートの状況を管理
 - SNMPの情報を使って、バックボーンのリソース利用状況を管理

BGP/MPLS VPN ネットワーク相互接続

BGP/MPLS-VPN網間接続

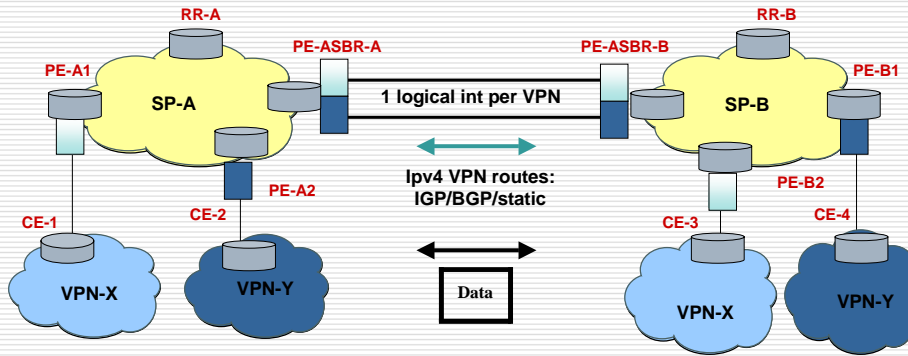
- ネットワークの規模を大きくして、カバレッジを増やすために、VPN網の相互接続が行われている。

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。



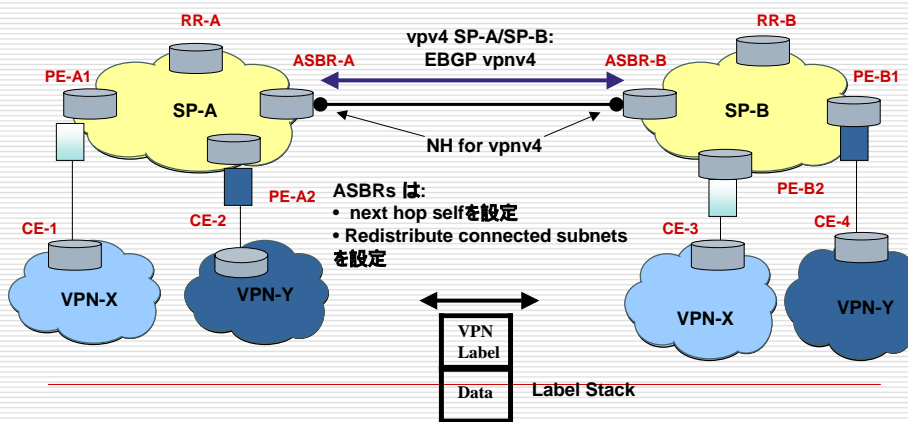
Option A

VRF to VRF : FRやSerialを使ってVPNを分ける

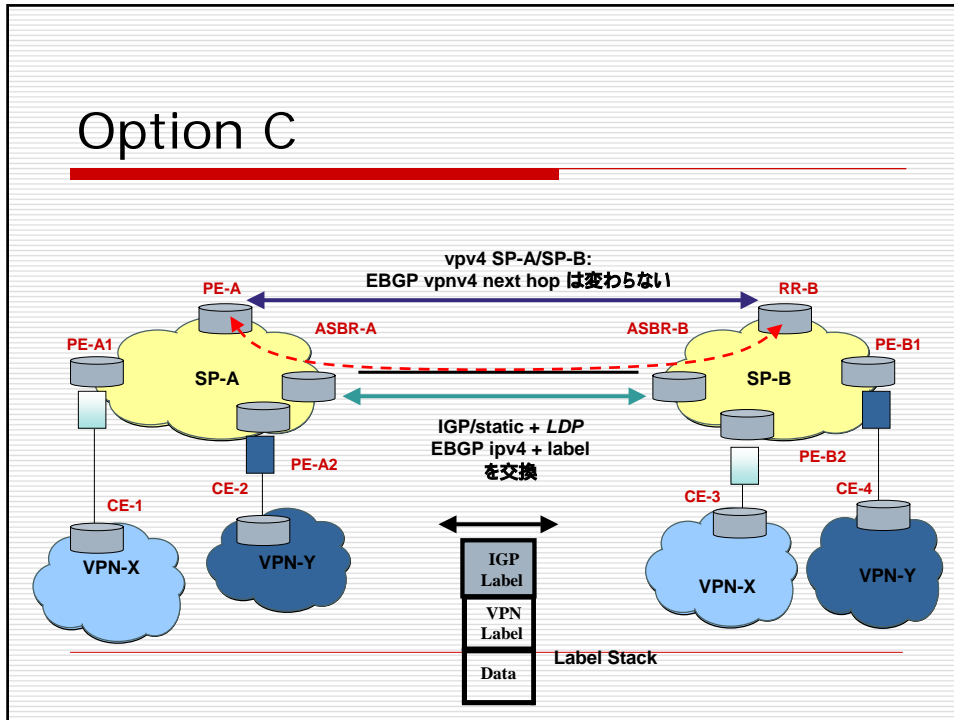


Option B

ASBR-A to ASBR-B for ipv4 & vpnv4:



Option C



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

BGP/MPLS-VPN設定例

Cisco PEルータConfig例

□ VPNの定義

```
ip vrf VPN-TEST  
rd 203.100.1.1:1  
route-target import 18084:1  
route-target export 18084:1
```

□ インタフェースのVPNへ括り付け

```
Interface Serial1/0/0  
ip vrf forwarding VPN-TEST  
ip address 10.0.0.1 255.255.255.252
```


Cisco PEルータConfig例(Cont.)

□ mpBGP部分の設定(CEルータStatic):抜粋

```
router bgp 18084
no bgp default ipv4-unicast
neighbor 192.168.0.1 remote-as 18084 他PEルータ向けPeer
!
address-family ipv4 vrf VPN-TEST      VPN用設定
redistribute static
no auto-summary
no synchronization
exit-address-family
!
address-family vpnv4                    route-target情報用
neighbor 192.168.0.1 send-community extended
!
```

Cisco PEルータConfig例(Cont.)

□ VPN用Static設定

```
ip route vrf VPN-TEST 10.0.0.0 255.0.0.0 Serial1/0/0 10.0.0.2
ip route vrf OTHER-VPN 10.0.0.0 255.0.0.0 Serial1/1/0 10.0.0.2
```

□ VPNが異なれば同じアドレスでも設定可

Juniper PE Router Config例

□ VPNの定義

```
routing-instance{
  VPN-TEST{
    instance-type vrf;
    interface t1-0/3/0.0;
    route-distinguisher 203.100.1.1:1;
    vrf-import VPN-TEST-import;
    vrf-export VPN-TEST-export;
    routing-options {
      static{
        route 10.10.10.0/24 next-hop
        203.100.254.2;
      }
    }
  }
}
```

Juniper PE Router Config例 (cont.)

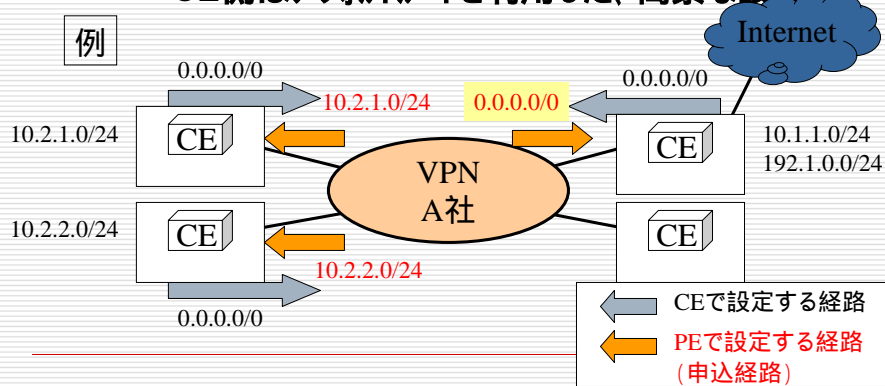
```
policy-options {
  policy-statement VPN-TEST-import {
    term 1 {
      from community VPN-TEST-import;
      then accept;
    }
    term 2 {
      then reject;
    }
  }
  policy-statement VPN-TEST-export {
    term 1 {
      from protocol static;
      then {
        community add VPN-TEST-export;
        accept;
      }
    }
    term 2 {
      from protocol direct;
      then {
        community add VPN-TEST-export;
        accept;
      }
    }
    term 3 {
      then reject;
    }
  }
}
community VPN-TEST-export members target:813:1;
community VPN-TEST-import members target:813:1;
```

BGP/MPLS-VPNユーザ構築例

MPLS-VPNユーザ構築例

□ Staticの考え方…主に拠点向き

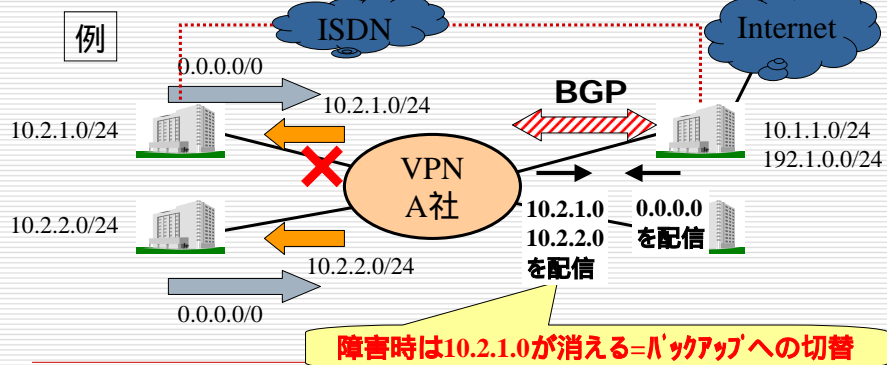
■ CE側はデフォルトルートを利用した、簡素な設定



MPLS-VPNユーザ構築例

□ BGPの考え方・・・主にセンタ向き

■ 動的ルーティングを生かしたバックアップ構成の実現



BGP/MPLS-VPNまとめ
(実際と新技術)

BGP/MPLS-VPN技術の実際

- Informational RFCからdraft-ietf-l3vpn-rfc2547bis-03.txtにて改定中
- 現在、L3VPN-WGにおいて、Provider Provisioned Layer3 VPNをサポート
- Cisco社中心から、現在では、複数のメーカーの実装が出ている。
- 同じMPLSを使ったVPN技術としてVR方式やルーティングのアウトソーシングを受けないLayer2 MPLS-VPN, EoMPLSなど新しい技術がどんどん現れ、IP-VPNを実現するための唯一の解ではなくなっている。

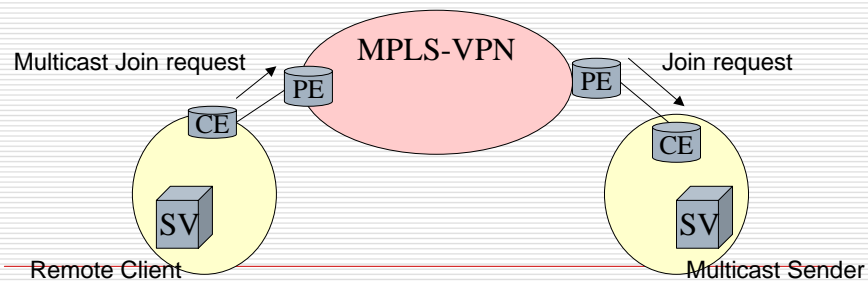
BGP/MPLS-VPN技術の実際

- 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
 - $1\text{VPN} * 1000\text{経路} \times 200\text{VPN} = 20\text{万経路}$ ！
 - 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

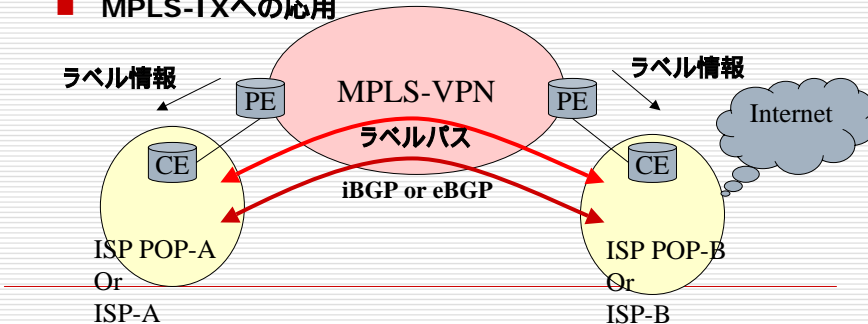
- Multicast PIM-SMをVPNユーザ単位に分けて機能を提供
- ユーザは個々に独立したMulticast網として利用可能
- 限定される設定や機能がまだ多い



BGP/MPLS-VPN関連の新しい技術

□ Carrier's Carrier機能

- IP-VPNを使ったISPバックボーンの構築
- 既存のIPネットワークをMPLS-VPNの統一プラットフォーム上に実現できる。
- CEルータでもMPLSを設定しラベルパスをCEから張る。
- MPLS-IXへの応用

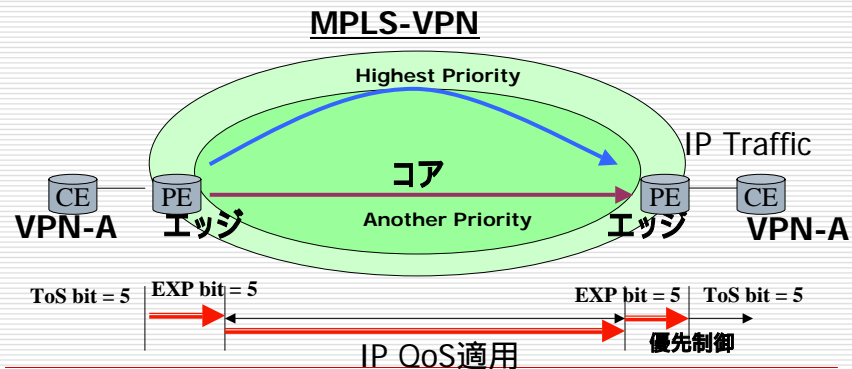


BGP/MPLS-VPN関連の新しい技術

- MPLS-VPNとTraffic Engineering,最新のQueue管理技術との組合せ
 - 特定のVPNやトラフィックに対してTEの機能を適用して、バックボーンを含めたQoS,CoS(MPLS-Diffserve, Diffserv-aware-TE)やFRR(FastReroute),特定クラス限定のLSP等の機能を提供

BGP/MPLS-VPN関連の新しい技術

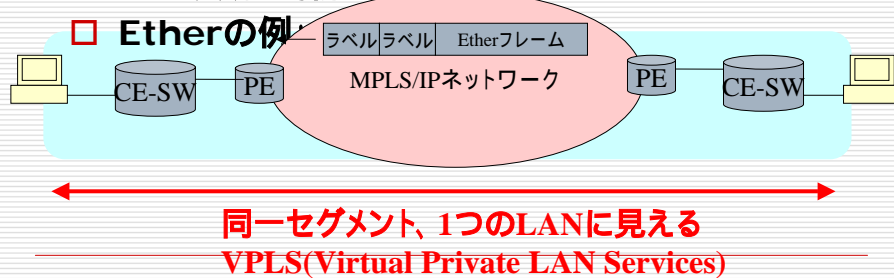
- MPLS-VPNとTraffic Engineering,最新のQueue管理技術との組合せ



参考: Layer2 VPN

□ IPネットワーク上で、レイヤ2フレームを転送する技術

- Ether/ATM/Framerelay/SDH : VPW(Wire)S, VPLS
- BGP/MPLS-VPNと組み合わせてさらに多様なVPNの実現が可能となる



L3VPN まとめ

- 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
- BGP/MPLS-VPNにおけるQoSの提供
- BGP/MPLS-VPN設定例
- BGP/MPLS-VPNユーザ構築事例
- BGP/MPLS-VPNまとめと新技術