IP VPN構築の理論と実践

～ネットワークベースVPN最新動向～

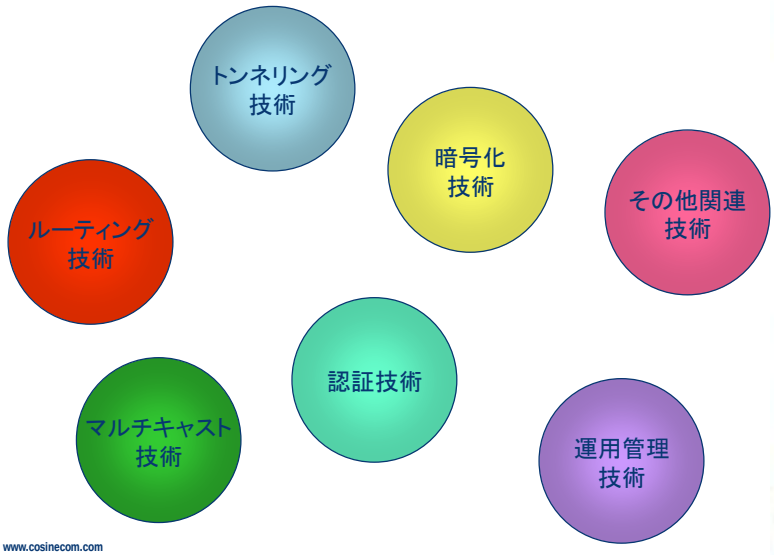
コサインコミュニケーションズ(株)
シニアシステムズエンジニア　進藤　資訓
mshindo@cosinecom.com

VPNはいまだに・・・

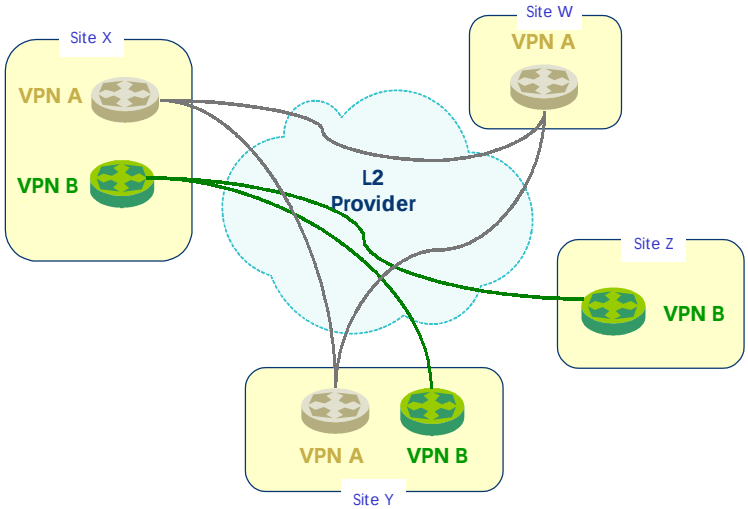
```
% mkdir vpn-doc
% cd vpn-doc
% wget -q ftp://ftp.ietf.org/internet-drafts/¥*vpn¥*
% ls | wc
   72    72    2570
% cd ..
% du -s vpn-doc
1863   vpn-doc
%
```

As of 11/11/2001

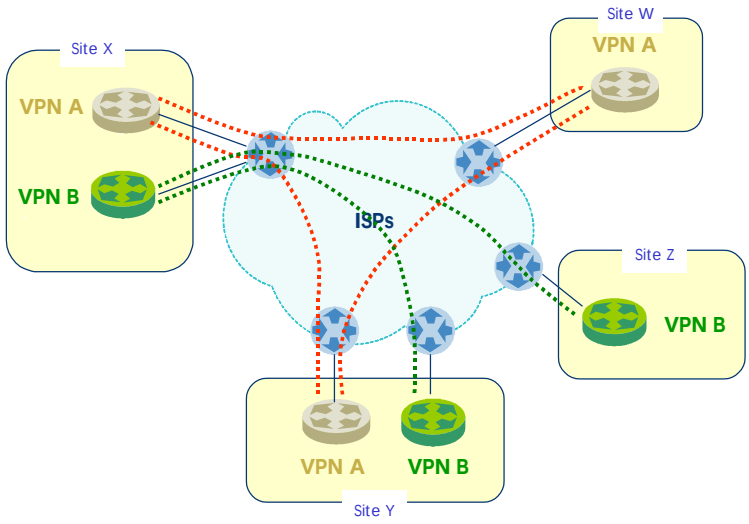
VPN関連技術



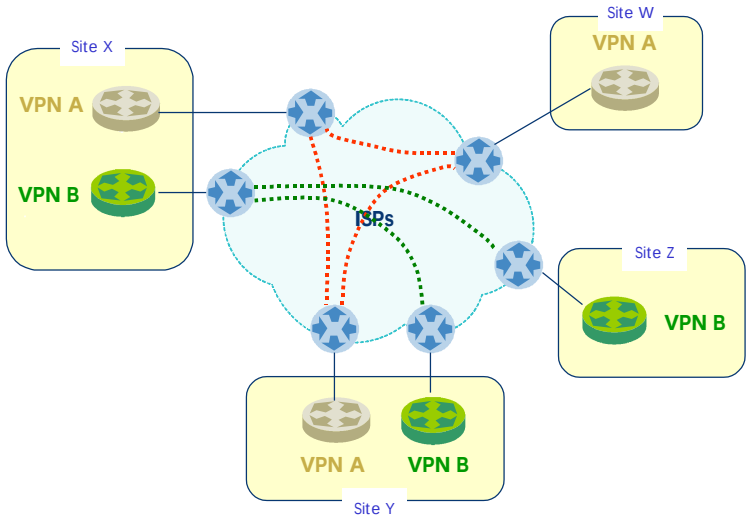
VPNの変遷（1） ～(!V)PN時代～



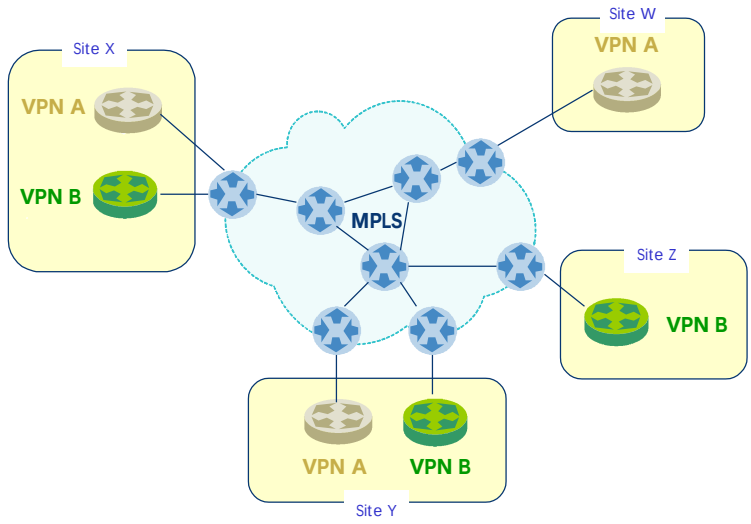
VPNの変遷（2）　～ CPE-based VPN ～



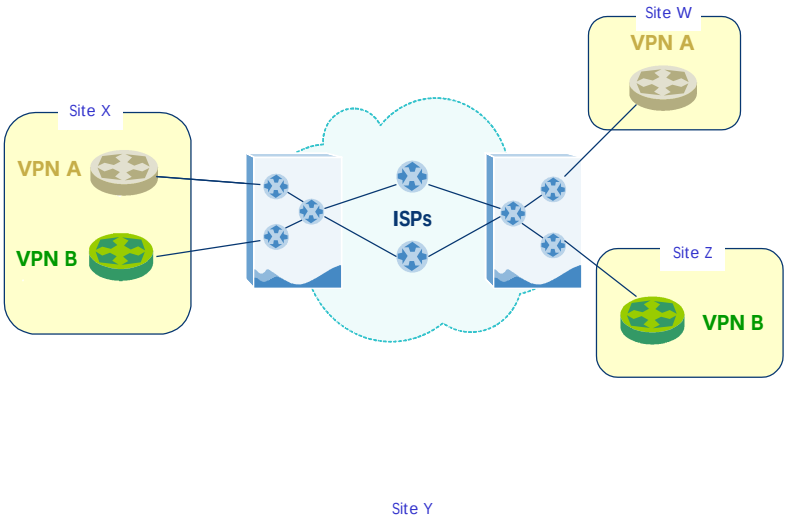
VPNの変遷（3）　～ Managed Router Solution ～



VPNの変遷（4a） ～ Network-based VPN ～



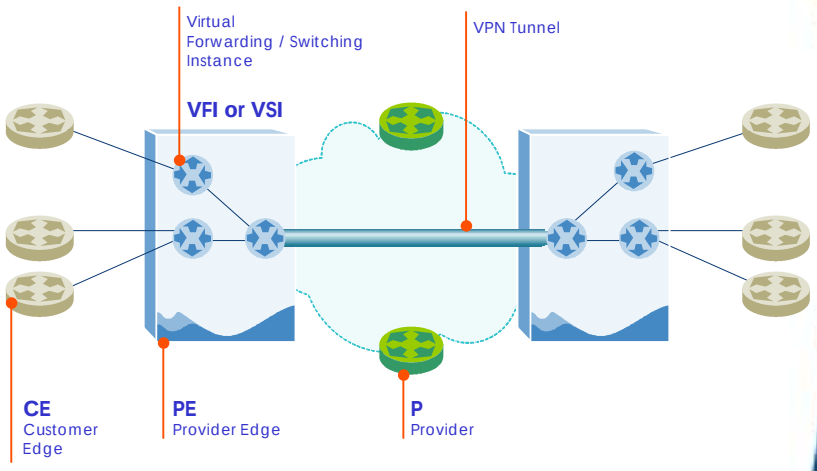
VPNの変遷（4b） ～ Network-based VPN ～



IETF / ITUの活動

- **Network-based VPN (NBVPN)**
  - August 3, 2000 – 48th IETF @ Pittsburgh - NBVPN BOF (Routing Area)
- **Provider Provisioned VPN (PPVPN)**
  - December 14, 2000 – 49th IETF @ San Diego - PPVPN BOF (Routing Area)
  - March 23, 2001 – 50th IETF @ Minneapolis – PPVPN BOF/WG (Sub-IP Area)
  - August 8, 2001 – 51st IETF @ London – PPVPN WG

トポロジーと用語（L2 and L3）



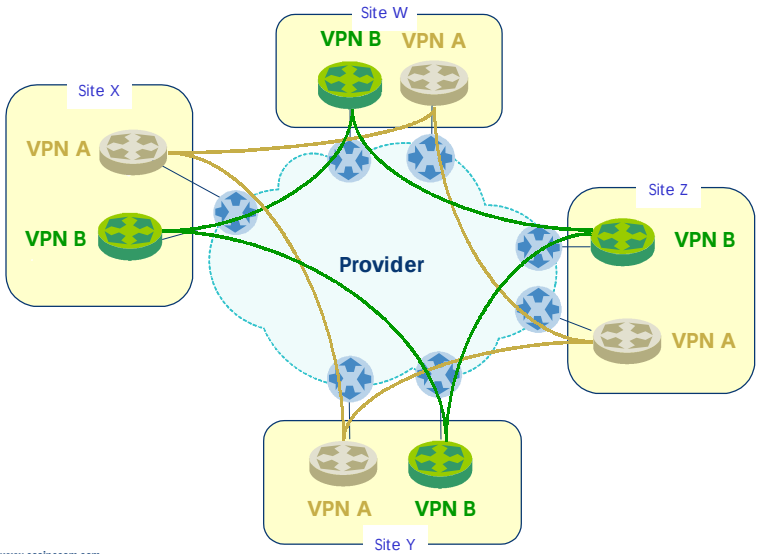
VPNの分類

- **Base**
  - CPE-based
  - Network-based
- **Model**
  - Overlay model
  - Peer model
- **Types / Applications**
  - Virtual Leased Lines (VLL)
  - Virtual Private Routed Networks (VPRN)
  - Virtual Private Dial Networks (VPDN)
  - Virtual Private LAN Segments (VPLS)
- **…**

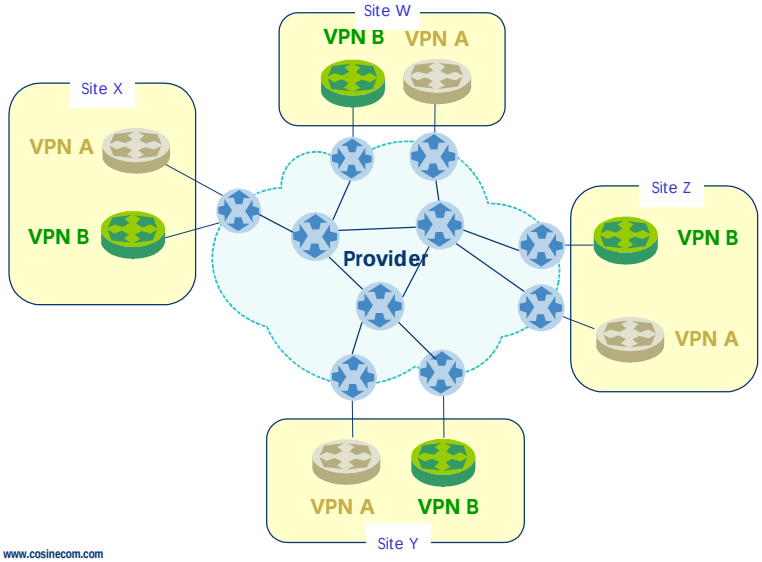
CPE-based vs Network-based

- **CPE-based**
  - プロバイダはVPNの存在を意識しない
  - プロバイダとの切り口
    - Layer 2
      - ATM / FR
    - Layer 3
      - Tunnel by GRE, IP-in-IP, IPsec, L2TP or MPLS, etc.
  - カスタマーからプロバイダのネットワークがどう見えるか？
    - Point-to-Point
    - Broadcast LAN
- **Network-based**
  - プロバイダは“サービス”としてVPNを提供
  - PE間のトンネルで実現
    - GRE, IP-in-IP, IPsec, MPLS, etc.
  - Layer 2 or Layer 3

Overlayモデル



Peerモデル



Network-based Layer 2 VPN

- ユーザからみると(巨大な)スイッチに見える
- 顧客の Layer 3 (特にルーティング)に関与しない
- Overlay model
- 利点
  - 管理の分解点が明確
  - 既存のLayer 2ネットワークからの移行がスムーズ
  - 自然なルーティングの分離
  - Layer 3独立(マルチプロトコルのサポート)
  - マルチキャスト
  - PEのスケーリング
  - 設定の容易さ
- 欠点
  - N^2問題
  - 単一のLayer 2に縛られる

Network-based Layer 3 VPN

- ユーザからみるとルーターに見える
- ユーザ側アドレスは通常PEのみexposeされる
- Overlay model or Peer model
- 利点
  - N^2 問題の回避
  - ルーティングを“サービス”として提供できる
- 欠点
  - カスタマーネットワーク(特にルーティング)に関わる必要がある
  - プロトコルの限定(典型的にはIP only)

NB-VPNに要求される各種機能（ステージ）

- **Discovery stage**
  - Membership discovery
    - 共通のVPNを持っているかを各PEがチェックする
  - Topology
    - Full-mesh / Hub & Spoke / Others
  - Capability discovery
    - PE間でトンネルやルーティングの方法に関して合意する
- **Tunnel establishment stage**
  - E.g. IPsecならIKEがこれにあたる
- **VPN routing stage**
  - Static, IGPs (OSPF/ISIS/RIP) and/or EGPs (BGP4)
- **Per-VPN or 全VPN共通**
- **これらのstageが個別に実現されている必要はない**
  - E.g. Layer-2 VPNの場合はVPN routing stageは必要ない

VPN discovery実現方法

- **NMS（Network Management System）**
  - 独自手法
  - 標準的手法（e.g. SNMP）
- **Directory (Database) server**
  - 例）LDAP
    - 組み込みの認証
    - Persistent search capability (proposed)
    - Replication (proposed)
- **ルーティングプロトコルにピギーバック**
  - 例）BGP4
    - VPN discovery とルーティングの統合
    - 幅広い実績とスケーラビリティー
    - きめの細かい制御が可能

Tunnel Establishment

- 通常、トンネリングメカニズムはVPN discoveryやVPN routingを実現するメカニズムとは独立している
- 目的
  - アドレスの重複
  - VPNトラフィックの差別化（e.g. QoS）
- IPsec、MPLS、GRE、IP-in-IP、L2TP、etc.
- 要素
  - Encapsulation
  - トンネルの多重化
  - トンネル作成
  - スケーラビリティーと階層化トンネル
  - トンネル維持管理

Tunneling encapsulation

- Per-tunnel stateの管理が必要か？
  - IPsec, IP-in-IP, GRE : 不要
  - MPLS : 必要
    - 削減するには、
      - 階層的多重化
      - Multi-Point to Point
- Encapsulation（ヘッダ）オーバーヘッド
  - IPsec >> MPLS
  - Fragmentation

トンネルの多重化

- ひとつのトンネルで複数のVPNをカバーするには、何かしらの多重メカニズムが必要
  - IPsec : SPI field
  - MPLS : Label
  - GRE : Key
  - IP-in-IP : outer IP address ?

トンネル作成

- 多重化識別子をどう配布するか？
  - 明示的なシグナリング
    - 利点 : セキュリティーやQoSなどの特性をper-tunnelで指定できる
    - 欠点 : スケーラビリティー問題
  - 暗黙的方法
    - VPN membership や VPN routes の配布時にpiggybackする
    - シグナリングを持たないトンネリングプロトコルの場合

スケーリング

- CoreでStateを管理しなければいけないトンネリングを使う場合、単にVPN毎にすべてトンネルを張るのはスケールしない！！
- スケールさせるには階層的トンネルが必要
  - 例) RFC2547
    - PE間のLSPはLDPで作る
    - Per-VPNのラベルはBGPにPiggyback

Tunnel Maintenance

- トンネルの生死の確認
  - VR (per-VPN Routing) 方式
    - VPNルーティングで検出可能
  - Aggregated routing方式
    - Per-VPNのRouting Instanceがないので、別の方式を用意する必要がある

トンネリングプロトコル比較

	MPLS	GRE	IPsec	IP-in-IP	L2TP	PPPoE
多重化	可	なし（Key Field使用可）	可	なし	可	可
マルチプロトコル	可	可	IP only（*1）	IP only（*1）	可	可
QoS / SLA	本質的にはなし	Delivering Protocolに依存	本質的にはなし	本質的にはなし	Delivering Protocolに依存	なし
トンネル確立＆維持	LDP / RSVP	なし	IKE	なし	L2TP	PPPoE
MTU	制限なし	制限なし	制限なし	制限なし	64K	1500
トンネルオーバーヘッド	小	比較的小（Bit Vector）	大	小	中（Bit Vector & HC）	比較的小
In-Order Delivery	保証	可能（Sequence Field）	なし(*2)	なし	保証はないが可能	なし

VPN Routing

- **Layer2 Network-based VPN**
  - シンプル！
  - 顧客のルーティングにはサービスプロバイダは関与しない
- **Layer3 Network-based VPN**
  - VFI
  - Scaling Issue
  - Per-VPN routing vs Aggregated routing

Per-VPN Routing

- VPNごとにルーティングプロトコルのインスタンスを生成＆実行
  - VPNごとに自由にルーティングプロトコルを選択できる
- VPN membership と reachability は独立している
- PE間でトンネルする必要がある
  - 通常、データをトンネルするのに使われるトンネルを共用
    - Data Plane = Control Plane
- VRモデルとの親和性が高い

Aggregated Routing

- サービスプロバイダ内のネットワークで単一のルーティングプロトコルを使用
- 理論的にはどのようなルーティングプロトコル（IGP/EGP）でも使用可能だが、
  - Link State Protocol（like OSPF / ISIS）は不向き
- BGPがベストチョイス
  - 柔軟なポリシー設定
  - 高いスケーラビリティー
  - RR
  - マルチプロトコルサポート
  - 高い浸透率
- Data Plane != Control Plane
- Peerモデルとの親和性が高い

Case Study

- BGP / MPLS VPN
- IPsec / BGP VPN
- IPsec + 2547 VPN
- VR VPN
- L2 MPLS VPN

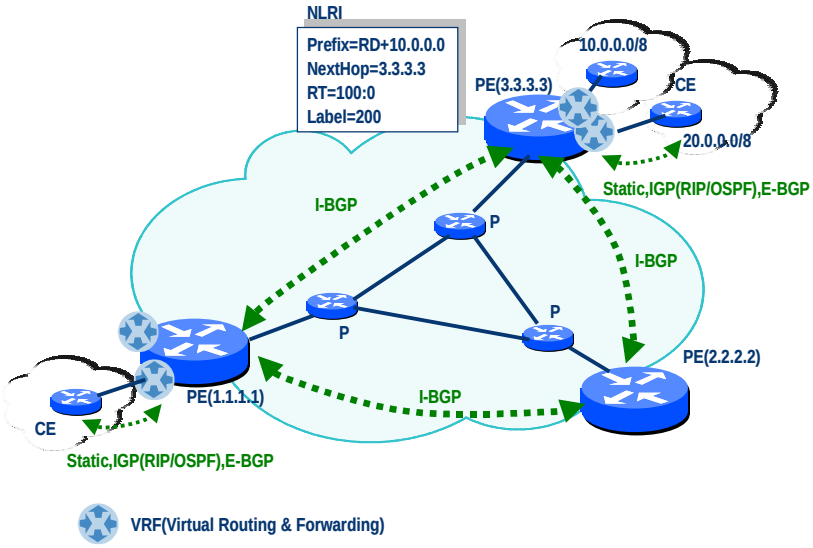
MPLS/BGP VPN

- RFC 2547 (Informational)
- draft-ietf-ppvpn-rfc2547bis-00.txt
- Layer 3 Network-based VPN
- Peerモデル
- MPLS (LSP)トンネル
- Aggregated Routing

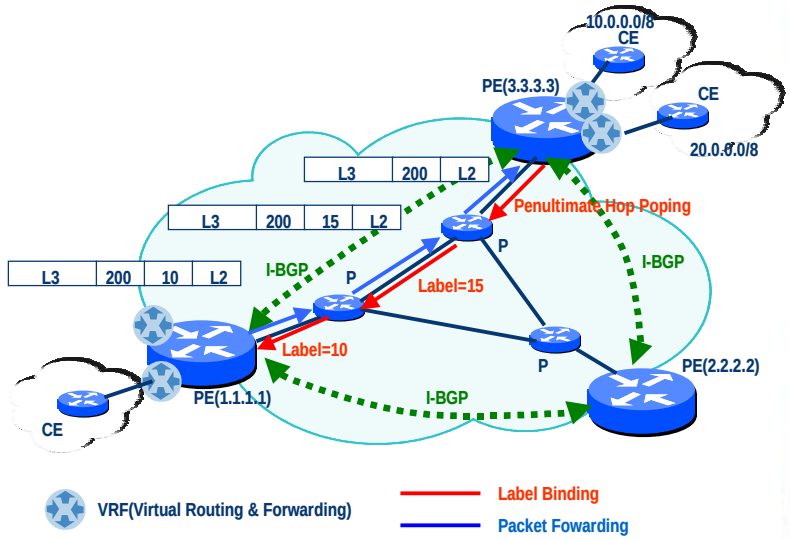
基本的な発想

- ルーティング情報の配布制御にBGP を使おう
  - スケールするし
  - Community で Filter するのがいいかも
    - でも、空間が足らないので Extended Community でencodeしよう！
- ルーティング情報を運ぶのにもBGPを使おう！
  - でも、アドレスは一意じゃないな～
    - VPN IP address = Route Distinguisher + IP address
  - そのままじゃ運べないよな～
    - マルチプトロコルなBGPを使おう！
- アドレスの重複はVRFで解決
- そのままじゃパケットを運べないので
  - MPLSを使おう
  - ラベルをスタックさせてスケールさせよう

RFC2547の動き（1）



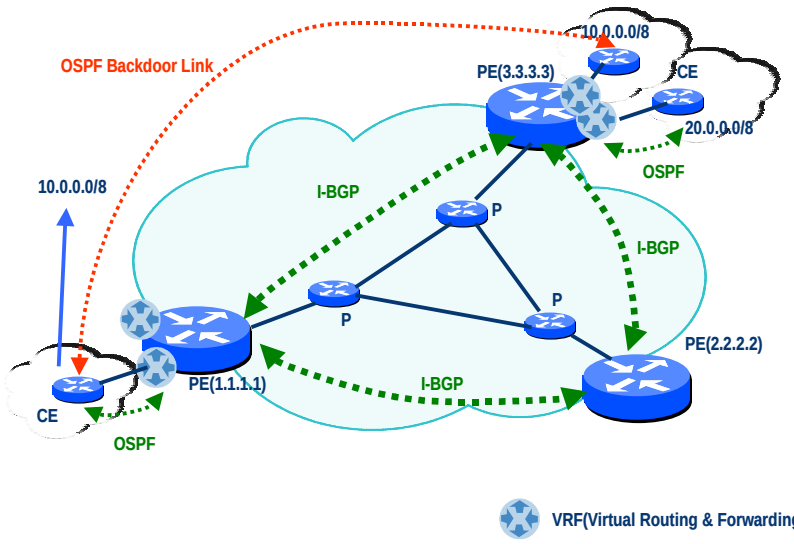
RFC2547の動き（2）



PE – CE routing in RFC2547

- PE (VRF) ～ CE間のルーティングプロトコルは自由に使うことができる
  - BGP-4 / RIP / OSPF / ISIS / Static, etc.
- ただし、
  - ループができやすいRIPなどは事故のもと
  - OSPF / ISIS のような Link State Protocol で、他サイトのルートは AS External なルートになってしまう
- 結局、現実的なのは Static か BGP-4！！

OSPF backdoor 問題 in RFC2547



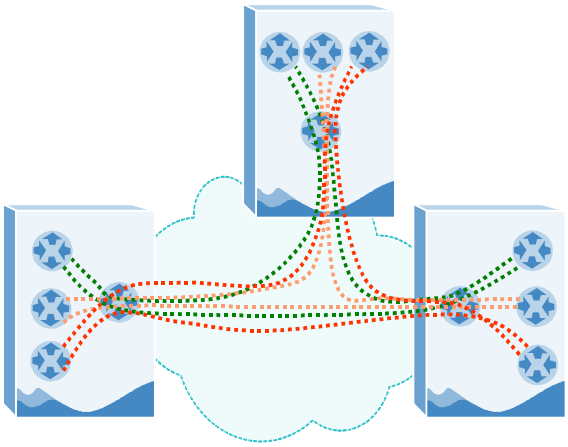
IPsec BGP VPN

- draft-declercq-bgp-ipsec-vpn-01.txt
- Layer 3 Network-based VPN
- Peerモデル
- Aggregated Routing
- 基本的な考え方はRFC2547(bis)を踏襲
  - ただし、MPLS LSPトンネルの代わりにIPsecトンネルを使用
  - セキュリティーの向上
  - MPLS core を仮定しなくてもよい（ただし、否定もしない）
  - スケーラビリティーを考慮

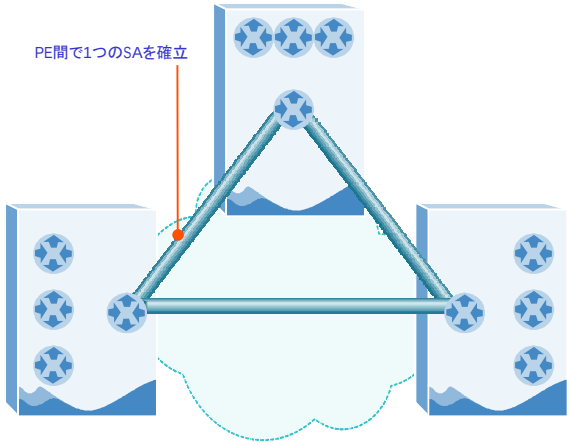
発想

- トンネル手法とVPN実現の仕組みは独立しているべきである！！
  - MPLS LSPトンネルの代わりにIPsecトンネルを使おう
- ただ、IPsecにはMPLSのLabel Stackingのような仕組みはないので、単純に考えると各VFI間でfull meshにIPsecトンネルを張ることになる
  - それではスケールしない！！
- では、トンネルはPE間だけにして、トンネル中に複数のVPNトラフィックを多重できるようにしよう！
- SPI (Security Parameter Index)を使うしかないね

SAトポロジー（VFIメッシュ）



SAトポロジー（PEメッシュ）



IPsec SA（おさらい）

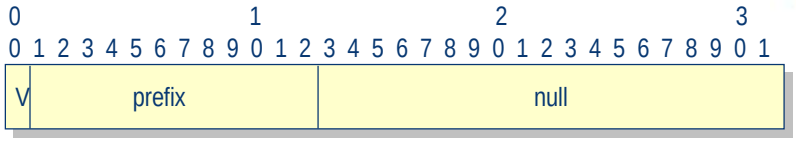
- 3つのパラメータで規定される
  - SPI (Security Parameter Index)
  - 送信先IPアドレス
  - セキュリティープロトコル（AH or ESP）
- SPIのフォーマットは規定されていない
  - Pseudo Random な 32bit Valueでありさえすればよい

手法

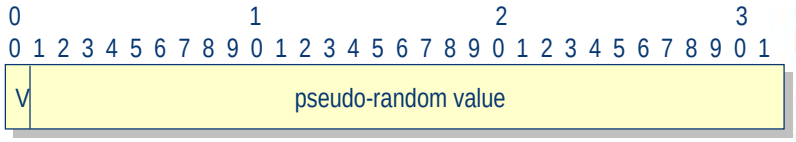
- 複数のSPI (SPI pool)を一つのSAにマップする手法
- BGP/IPsec VPNと通常時のコンテキストを区別する

VPN-SPI

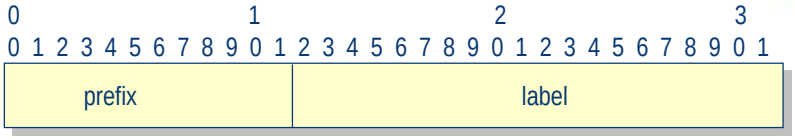
- V-flag = 1



- V-flag = 0

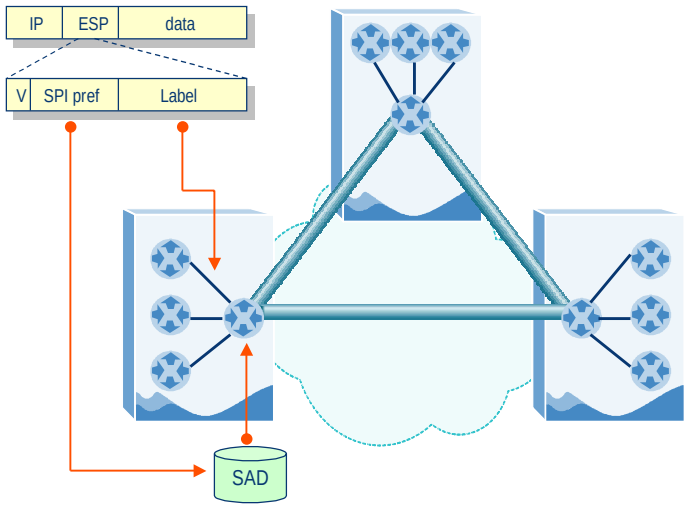


SPI in IPsec Processing



- “label”はSAを特定するためには使われない
- 20bit はMPLSラベルを使うことを想定

BGP / IPsec VPN



その他の部分

- RFC2547bisと同様（BGPを使う）
  - VPN-IPv4 AFとRoute Distinguisher (RD)
  - Route Target (RT) と Extended Community
  - Route Reflector (RR)はほぼ必須（大きなネットワークの場合）
  - …
- スケーリングに関する（理論上の）特性も同じ

P routerへの要求

- VPNに関する“知識”はなくてよい
- MPLSを要求しない（否定もしない）
  - 単なるIPルーター
  - 各PEに対するhost routeさえあればよい

セキュリティー

- MPLS-based
  - Link Layerでのセキュリティー
- IPsec-based
  - PE – PEのエンドツーエンドセキュリティーを実現
  - SPまたがりのサービス時に有用

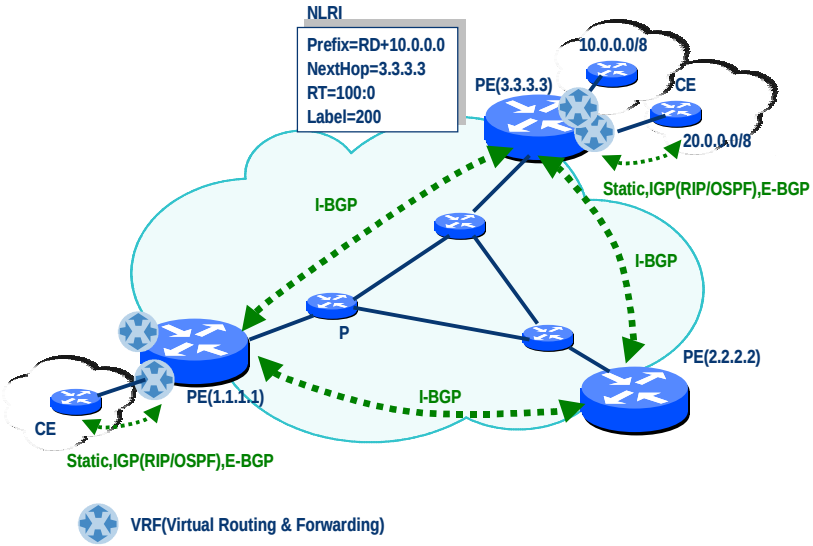
IPsec + 2547 VPN

- draft-ietf-ppvpn-ipsec-2547-00.txt
- Layer 3 Network-based VPN
- Peerモデル
- Aggregated Routing
- IPsec / BGP VPNの改良？

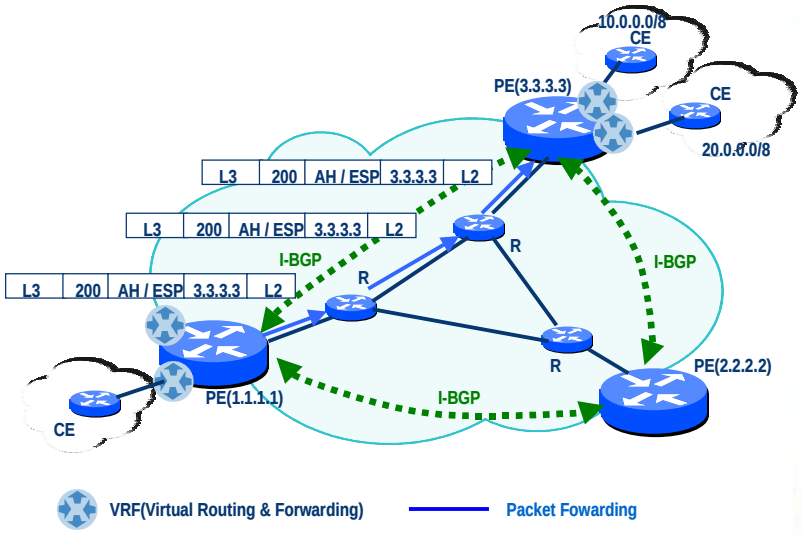
発想

- IPsec / BGP VPN の有用性には同意
- ただし、SPIに新たなセマンティックスを導入するのは、IPsec および IKE への変更につながる
- ならば outer なラベルに IPsec を使い、inner なラベルにMPLSを使えばいいんじゃない？
  - MPLS-in-IP でencapすればいいじゃん！！

IPsec + 2547の動き（1）



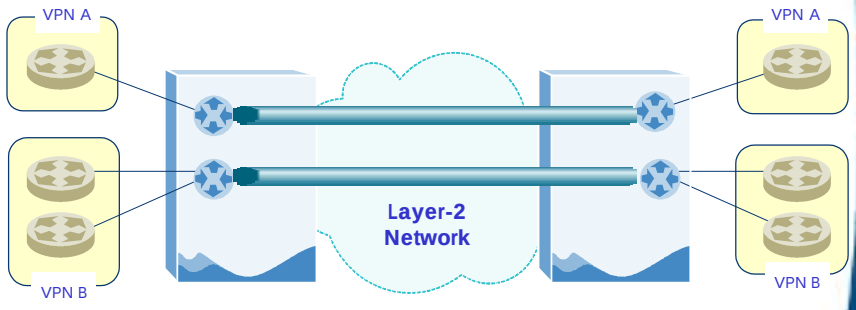
IPsec + 2547の動き（2）



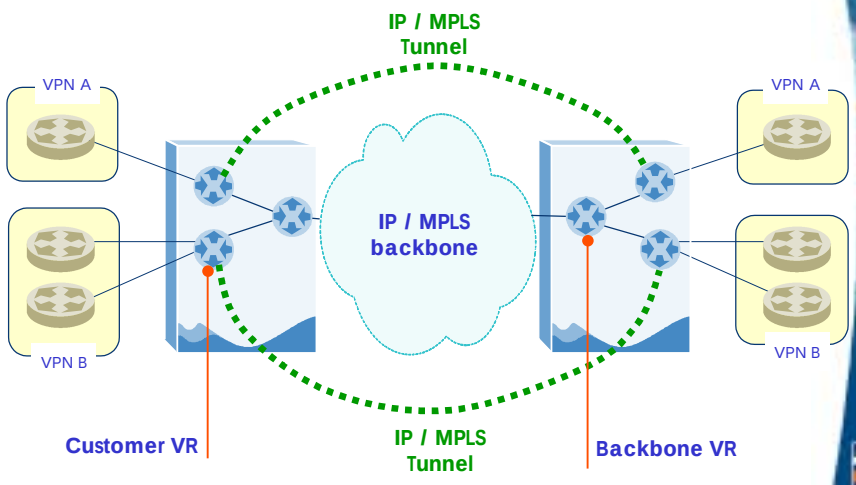
VR-based VPN

- draft-ietf-ppvpn-vpn-vr-00.txt
- Virtual Router（仮想ルーター）
- Layer 3 Network-based VPN
- Overlay モデル
- Per-VPN Routing

Layer-2 backbone-based VR-based VPN



Aggregated Backbone-based VR-based VPN



Layer-2 vs Aggregated Backbone

- **Layer-2 Backbone**
  - Per VPNでQoSをコントロールできる
  - 完全なルーティングの分離
- **Aggregated Backbone**
  - BackboneにLayer-2で多重できるメディアを仮定しなくてもよい
    - GbE, POS, etc.
  - スケーラビリティー
  - トンネル
    - MPLS and/or IPsec
    - Per VPN vs VPNで多重

MPLS L2 VPN

- **draft-kompella-ppvpn-l2vpn-00.txt**
- **Layer 2 Network-based VPN**
- **Overlay モデル**
- **ユーザのルーティングには関与しない**

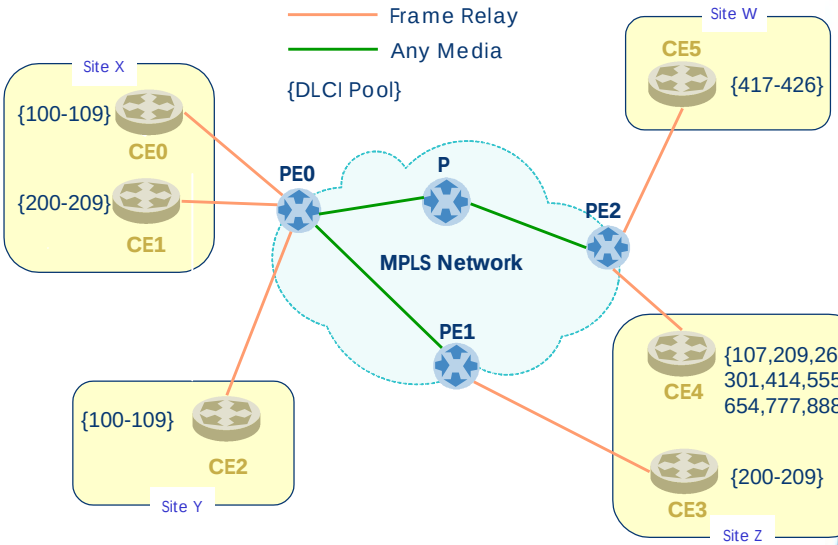
L2のモチベーション

- DSL
- Optical Ethernet
- Metropolitan Network

基本的な発想

- Layer 2 VPNを提供したい
  - L2 VPNの利点を参照
  - できるだけLayer 2 のセマンティックスを保ちたい
    - マルチプロトコル性
    - In-Order Delivery
    - Non address情報
- トランスポートにはMPLSを使おう！
  - ラベルスタック & CE-PE-P-PE-CEモデル
- でもN^2問題は避けたい
- Over Provisioning でProvisioningの負荷を軽減してやろう！（半自動 Provisioning）
  - Layer 2 ID (DLCI, VPI/VCI) および Label は cheap である！（という前提）

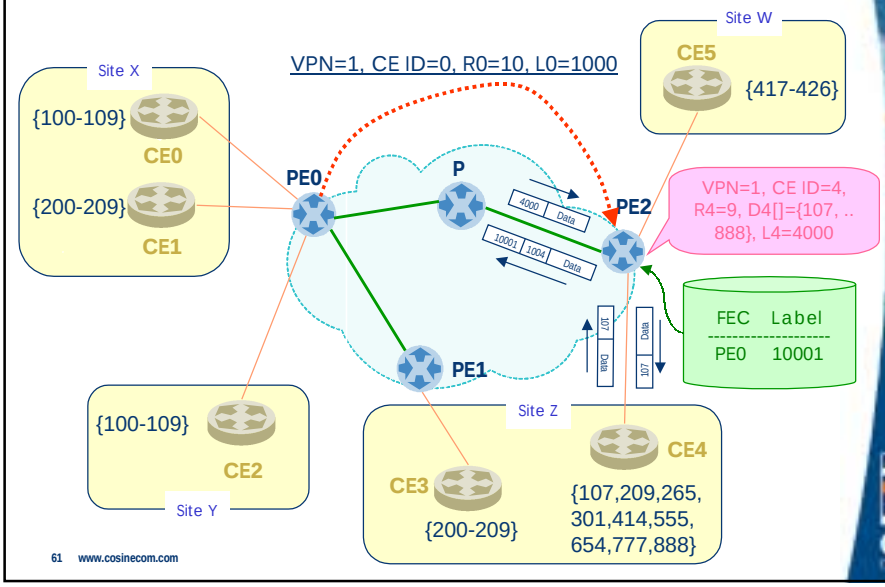
Topology



設定

- VPN ID
- CE ID
- Range
- Label Base

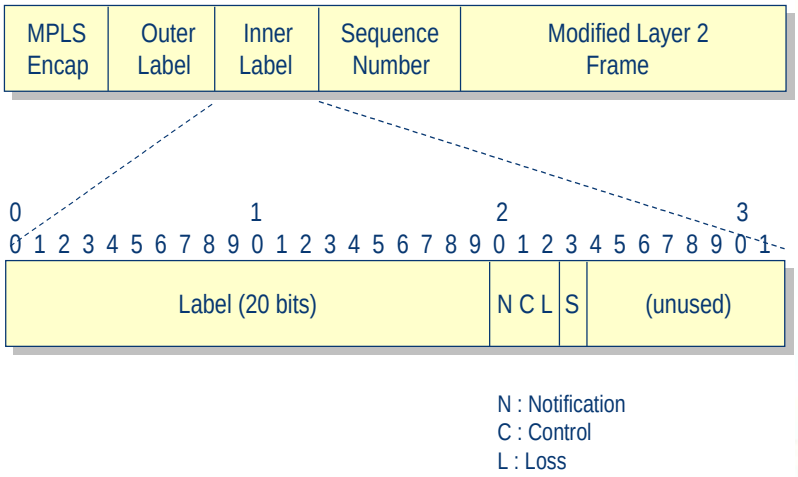
PE Advertisement



ポイント

- CE_a→CE_bなパケットを送る際にPE_kがinner labelとして割り当てたlabelは、PE_lがCE_a→CE_bなパケットを受け取る際のincomingなlabelとして割り当てたものと必ず等しくなる
- Layer2 ID（DLCI, VPI/VCI, etc.）とMPLS LabelのOver provisioningにより、CEの追加が発生しても、変更は局所的で済む（full mesh時でも）

フレーム & ラベルフォーマット



N : Notification
C : Control
L : Loss

シグナリング

- **BGP**
  - Multiprotocol BGP
  - L2-VPNのためのAFIおよびSAFIを新たに導入
  - L2-VPNのためのNLRIを規定
  - Extended Community
- **LDP**
  - 以前は提案されていたが、現状はとりあえずBGPにしぼる

BGP NLRI for L2-VPN

Length (2 octets)
Route Distinguisher (8 octets)
CE ID (2 octets)
Label-block Offset (2 octets)
Label-base (3 octets)
Variable TLVs (0 to N octets) ….

MPLS VPN : To be Layer 3 or not to be …

～Layer 3 MPLS VPN (rfc2547) vs Layer 2 MPLS VPN～

- **Peerモデルってそんなにいいの？**
  - Aggregated Routing になるよね
  - CE-PE間ルーティング問題（OSPF Backdoor Link問題）
    - draft-rosen-vpns-ospf-bgp-mpls-02.txt なんてのもあるけどさ
  - 顧客にルーティングを“サービス”する
    - 言葉を変えて言うと、「顧客からルーティングの自由を奪う」
  - スケールするの？
    - Provisioning的にはすると思う
    - でも、PEがユーザの経路持たなくちゃいけないよね！
    - ユーザーの経路がフラップするかもしれないし
- **とは言っても、2547は良くできたモデルだと思うし、Layer2の実装はこれから**
- **要は適材適所！**

各VPN手法の比較

	BGP / MPLS VPN	IPsec /BGP IPsec + 2547 VPN	VR VPN	L2 MPLS VPN
CPE / Network bas	Network-based	Network-based	Network-based	Network-based
Layer	Layer 3	Layer 3	Layer 3	Layer 2
Overlay / Peer mo	Peer	Peer	Overlay	Overlay
Tunnel	MPLS	IPsec	Layer 2 or Any IP/MPLS-based Tunnel	MPLS
Discovery	BGP	BGP	Any	BGP / LDP
Tunnel Establishme	LDP	IKE	Any	LDP / RSVP
VPN Routing	BGP (Aggregated)	BGP (Aggregated)	Any(Per-VPN)	N/A

まとめ

- **IP-VPN != RFC2547**
  - IPsec / BGP
  - VR
  - L2 MPLS
  - …
- **ユーザ側でルーティングに関して細かい要求がなければRFC2547は良いモデル**
  - モデル != 実装
- **L2 PPVPNは今後に期待**
  - ルーティングの自由度とマルチプロトコルが魅力
- **VPNの“P”を実現するにはIPsec**
- **ユーザーが選択できるのが望ましい**
  - 優れたProvisioning Toolが必要

参考資料

- RFC 2547 (BGP/MPLS VPN)
- RFC 2685 (VPN-ID)
- RFC 2764 (IP VPN Framework)
- RFC 2917 (CORE-MPLS-VPN)
- draft-ietf-ppvpn-framework-00.txt
- draft-ietf-ppvpn-rfc2547bis-00.txt
- draft-rosen-vpns-ospf-bgp-mpls-01.txt
- draft-declercq-bgp-ipsec-vpn-01.txt
- draft-ietf-ppvpn-vpn-vr-00.txt
- draft-ouldbrahim-bgpvpn-auto-01.txt
- draft-kompella-ppvpn-l2vpn-00.txt
- draft-shah-mpls-l2vpn-ext-00.txt
- draft-kb-ppvpn-l2vpn-motiv-00.txt
- draft-tsenevir-l2-req-00.txt
- draft-worster-mpls-in-ip-05.txt