

# マルチ AS かつマルチベンダに おける SR-MPLS TE の実現

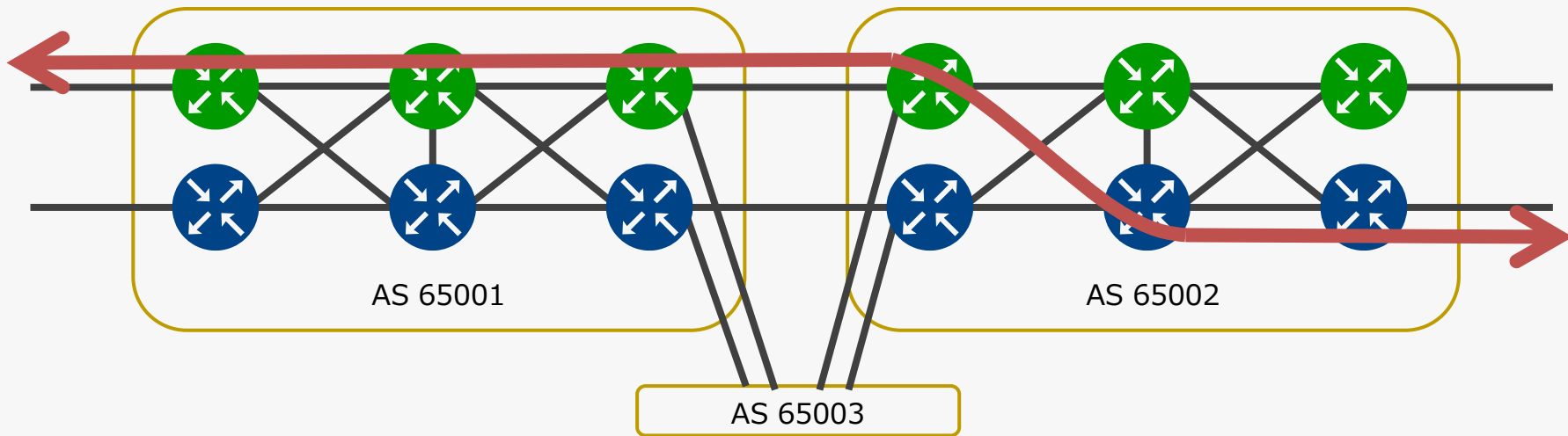
田島 照久

NTTコミュニケーションズ / イノベーションセンター  
teruhisa.tajima@ntt.com

2021/11/01 MPLS JAPAN 2021 (オンライン開催)

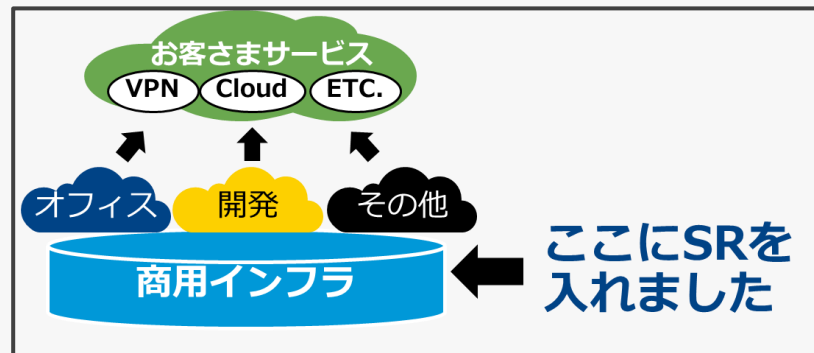
# 発表要旨

- 複数のAS、複数のSRドメインをまたがるVPNの実装について
- E2EでSR-TEを行う際の実装について
- SR-MPLS/SRv6相互接続SFCの実装について



# SR(-MPLS)の活用状況・発展

- SR-MPLSを用いたL2/L3VPN網はすでに実用段階
  - シングルSRドメイン、シングルベンダでは問題なく動作
  - 参考：MPLS JAPAN 2019 “SR-MPLSの導入事例と今後の展望について”



- 適用範囲や網の機能を増やす際の実装を検討したい
  - 適用範囲 = underlayの広さ、多様さ
  - 網の機能 = Traffic Engineering、Fast Re-Route

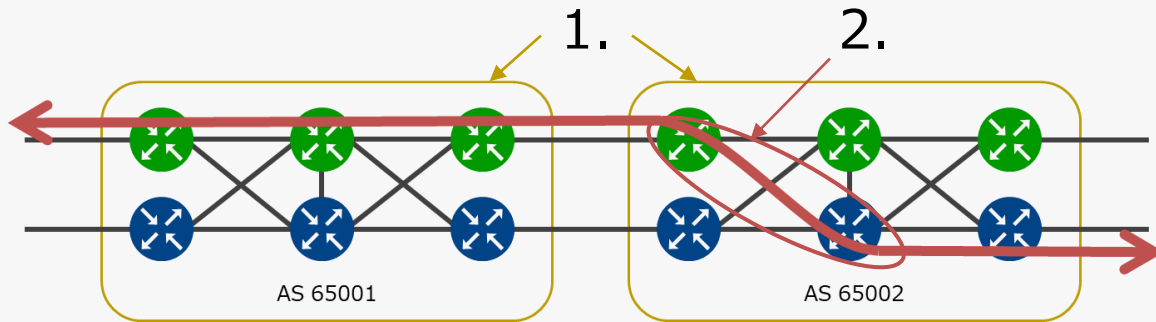
# より広範囲に適用可能なSR-MPLS網へ

## 1. Underlayの拡張性担保

- 網全体で共有するのではなく、underlayトポロジ情報の隠蔽が必要

## 2. ユースケースを満たせるTEの実装

- 単にE2Eの到達可能性だけでなく、SLAに応じたoverlayの提供が必要



# 1. Underlayの拡張性担保

## ■ シングルSRドメインの拡大とともに制約が増加

- レベル/エリアはあるがIP重複などはできない
- 設定の影響範囲を絞れるほうが嬉しい

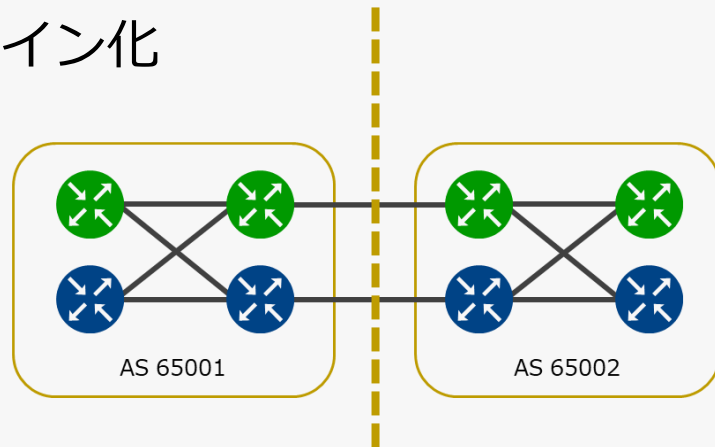
→ UnderlayのマルチAS化とマルチSRドメイン化

## ■ ASを区切り、AS内の実装を隠蔽

- ルータIDやP2P IPの重複が可能
  - Route Distinguisherへの影響は要考慮
- IGP, iBGPの設定影響がAS内に閉じる

## ■ AS境界でSRドメインを区切る

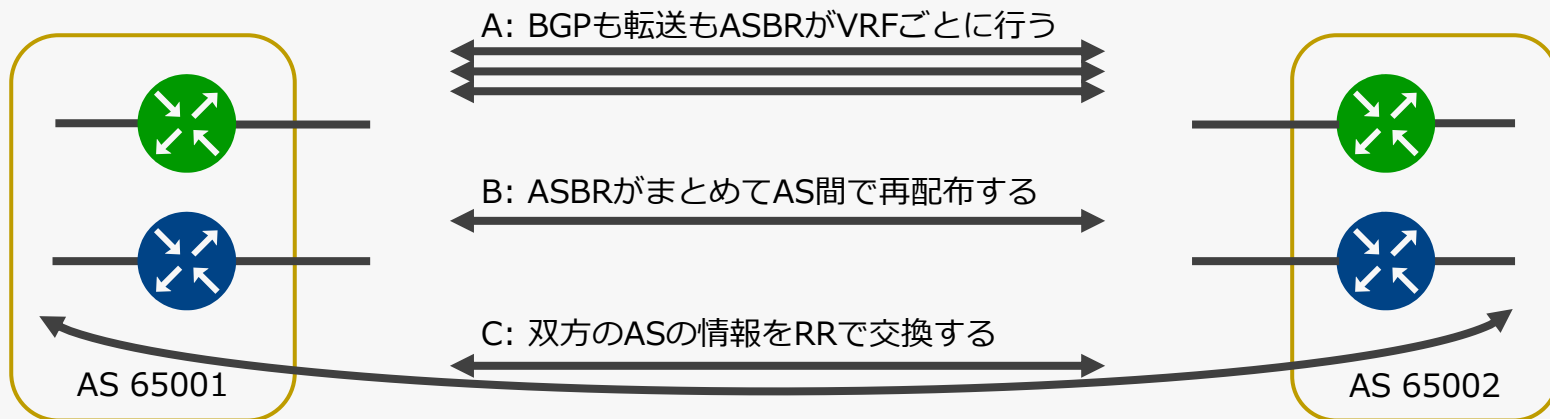
- SR関連の設定もAS内に閉じる
- ただし、分断されることによりE2EでのSRによる設定が不可に



# Multi-ASの技術選定

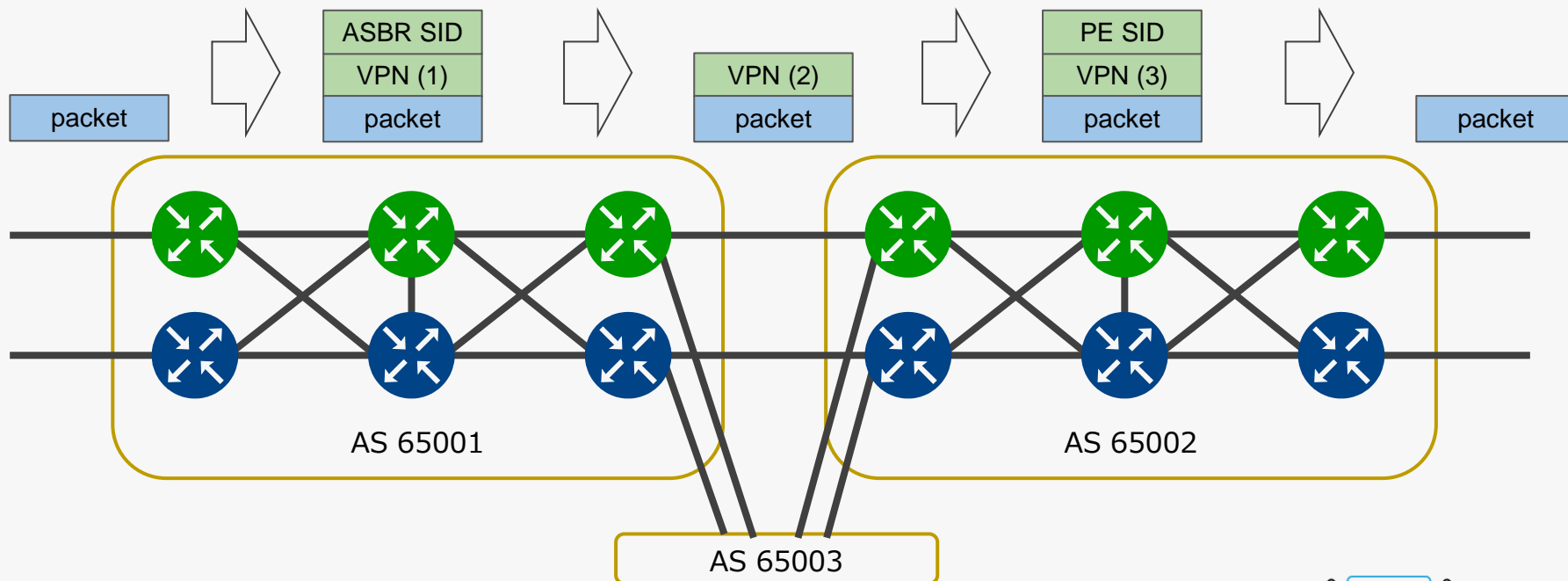
## ■ Inter-AS MPLS option-Bが要望を満たす

- A > VRFの制限や設定の煩雑さの問題あり
- B > 技術的にデメリットがないが、対応状況が微妙 ← これを選択
- C > underlayが見えるので今回の目的には不適
- D > Aの問題点が半分しか解消されない



# Option-BによるL3VPN (VPNv4/6)

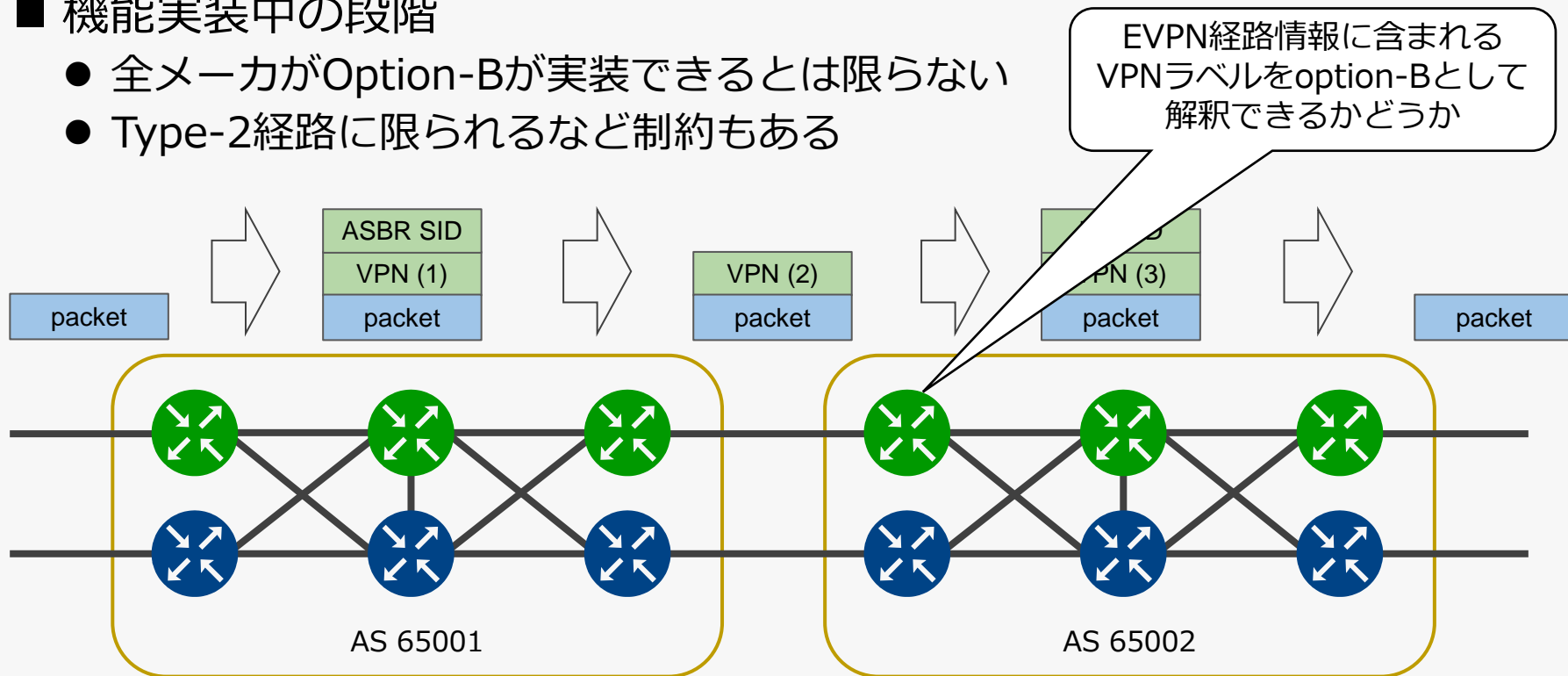
- ASBRはAS内とAS間のVPNラベルを相互にswapしてVPNを接続
- 3AS以上でも正しく動作



# Option-BによるL2VPN (EVPN)

## ■ 機能実装中の段階

- 全メーカーがOption-Bが実装できるとは限らない
- Type-2経路に限られるなど制約もある





# Multi-ASの工夫が必要な点

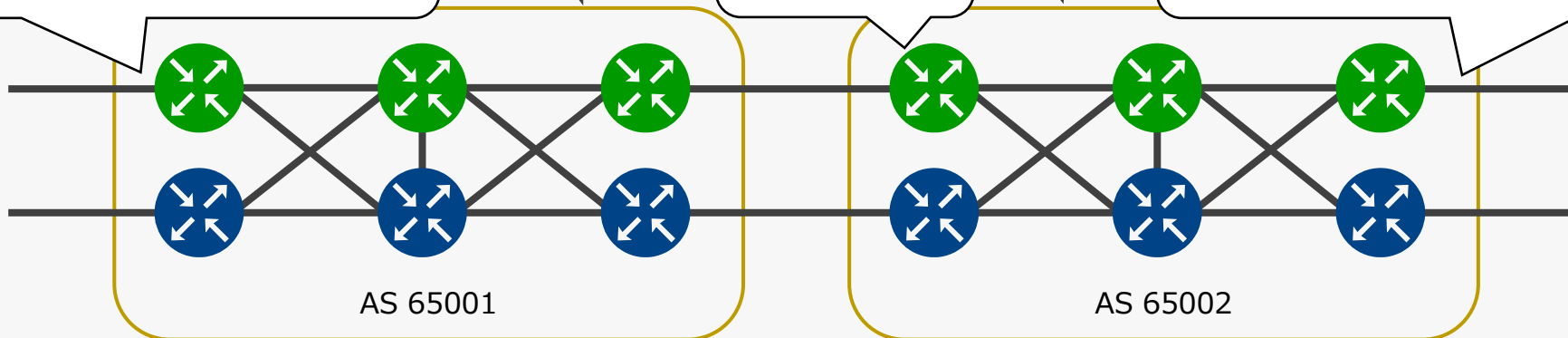
## ■ Route Target

- ingress-PEではASを超えた先のegress-PEの経路情報をインストールしないといけない
- ASの実装を隠蔽するためには、相互流通用ルールを作って対応

RT = (共通ASN):(識別番号)  
をインストール

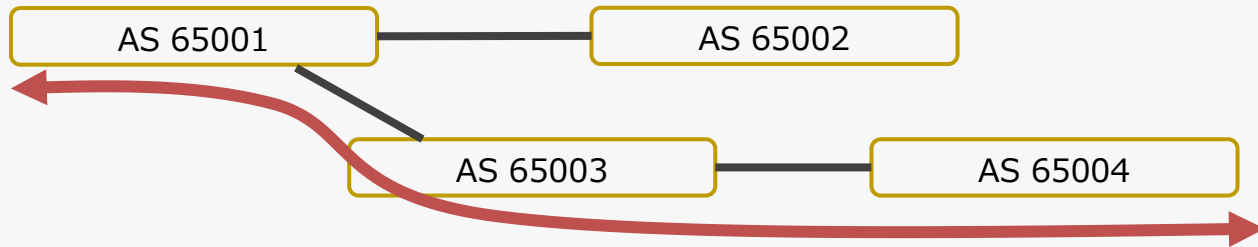
ASBRで付け  
替えてもよい

RT = (共通ASN):(識別番号)  
例) 64999:1000



# 1. Underlayの拡張性担保 のまとめ

- AS=SRドメインを分割し、Option-BでのVPN実装
- Underlayの設計自由度が確保
  - P2P IPの採番が自由
  - IGPコストの設計が自由
- AS間の接続の自由度も確保
  - AS間のつながりはSRに依存しないため、トランジットASの存在も可能



## 2. ユースケースを満たせるTEの実装

### ■ ユースケース

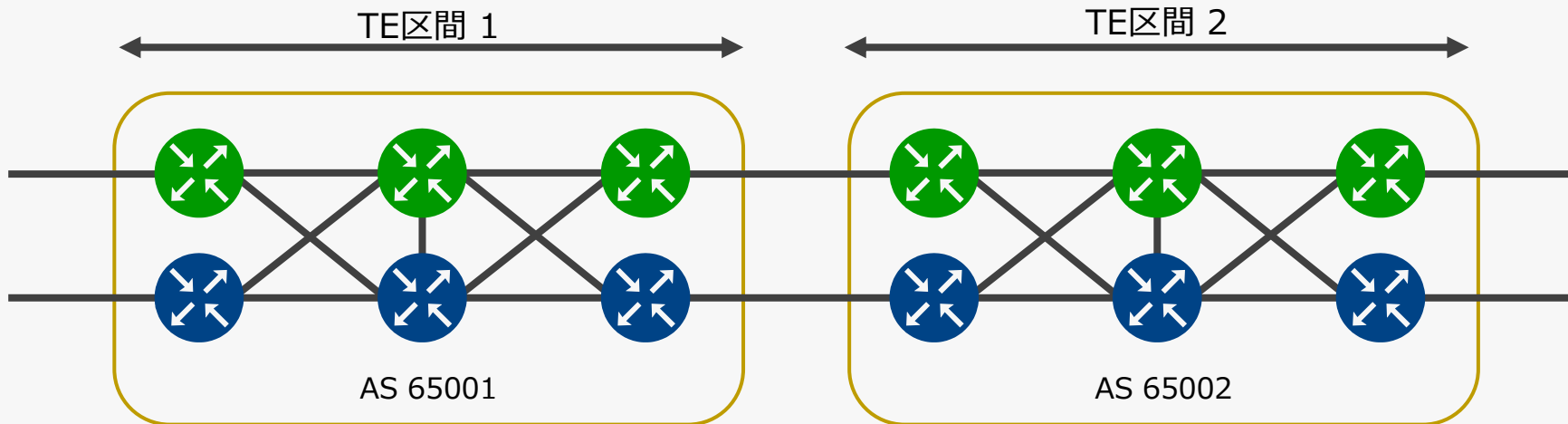
- このスライスは {特定の経路} で送りたい
- このスライスは {低遅延|広帯域} で送りたい
- このスライスは冗長系のうち {1系をメインに|2系をメインに} 使いたい
- このフローは {低遅延|広帯域} で送りたい

### ■ 実装方針

- 各SRドメイン内でTEを実装
- SRドメイン間ではあらかじめTEメニューを設定しておく

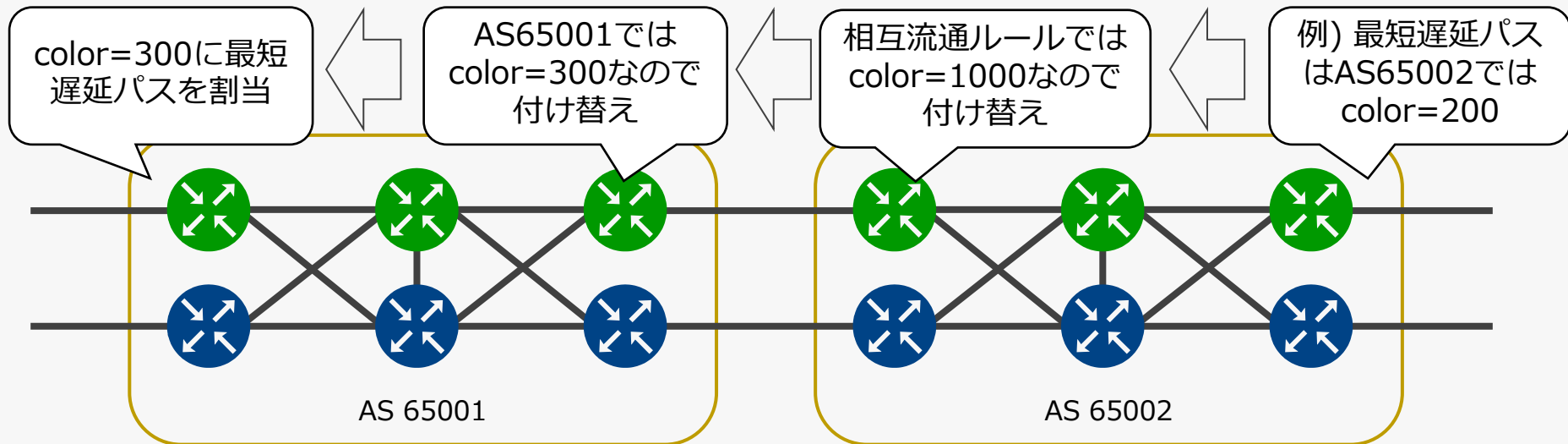
# あらかじめTEメニューを設定 とは

- マルチSRドメインで真のend-to-end TEは難しい
  - 達成するにはunderlayすべての情報が必要だが、意図的に隠蔽している
    - なお複数PCE連携等の検証は未着手
  - TEの実装は各SRドメインに閉じて、部分最適をつなぎ合わせることにした



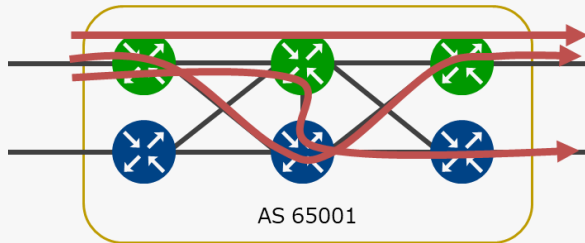
# ColorによるTEメニューの設定

- AS間はTEメニューに対応する相互流通用のColor表を作成
- Colorの値の意味はTE区間=SRドメインに隠蔽

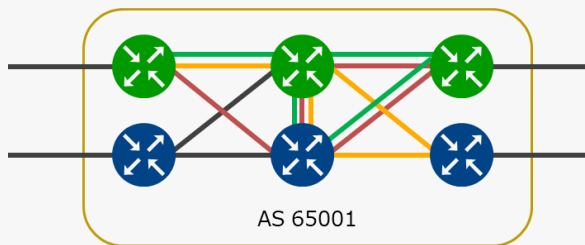


# SRドメイン内でのTEの実装パターン

2-a. Head-end(PE/ASBR)でTail-end(ASBR/PE)までのSIDを積む



2-b. 各ルータがFlex- Algoによる転送を行う

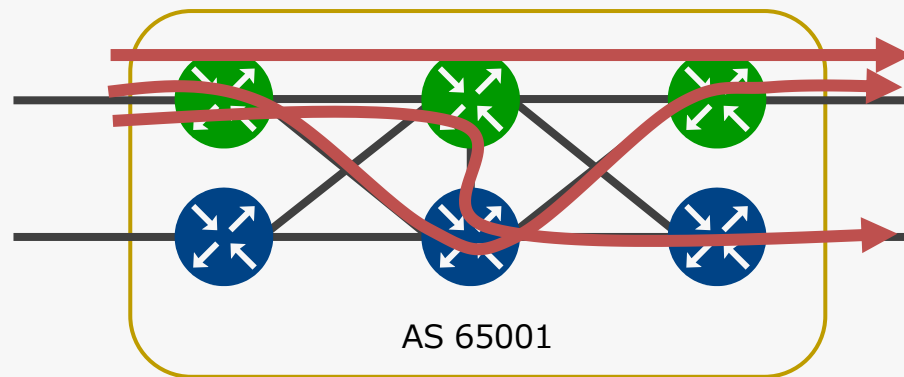


## 2-a. Head-endでSIDを積むパターン

- VPNのingress PE = Head-endで転送経路を計算しSIDで表現
  - PEの機能のみによって実現される
  - Pros: 中継がマルチベンダでも関係ない
  - Cons: TE経路の増減に合わせて各Head-endでの詳細な設定が必要

### ■ 指定できる経路の種類別

- Explicit-path > OK
- TE metric > OK
- Latency-based metric > OK

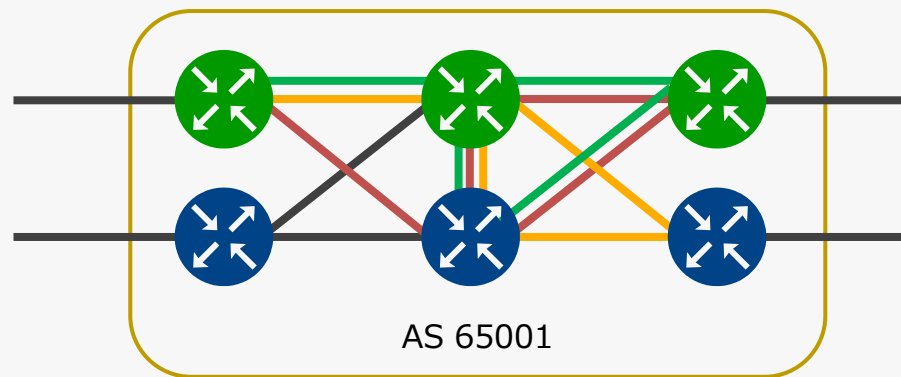


## 2-b. Flex-Algoで転送するパターン

- 詳細な経路をSIDの組で指定するのではなく、Algoに従い転送
  - Pros: Algoの指定だけで設定が統一される
  - Cons: Algo面が全ルータで正しく認識されている必要がある
    - 一部ルータのIS-IS TLV実装の差異で認識されない組み合わせがあった

### ■ 指定できる経路の種類

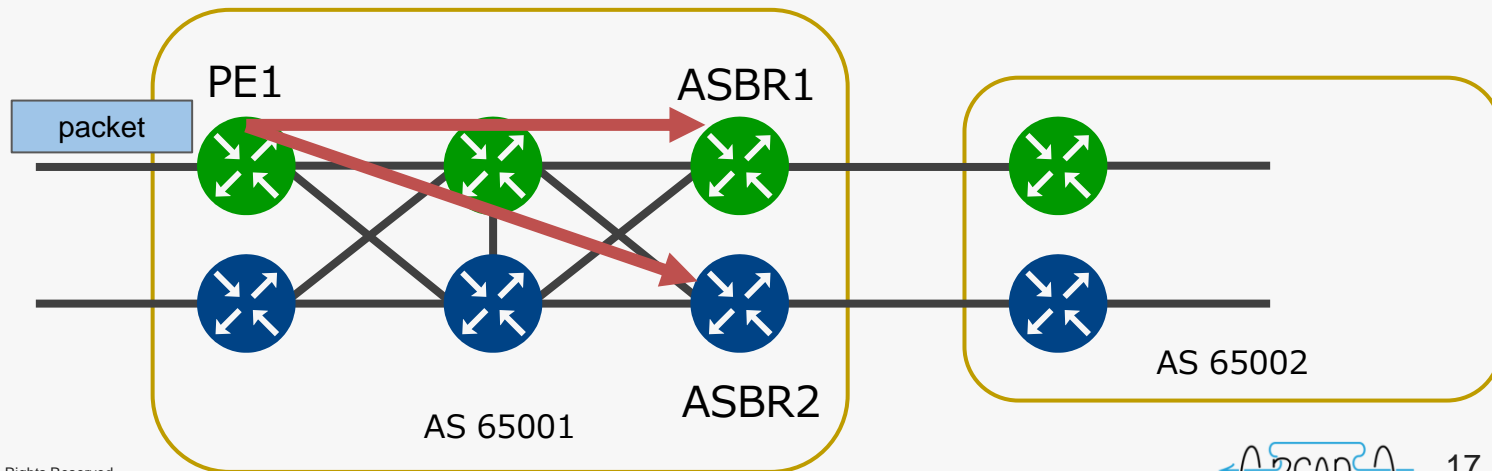
- 特定のAlgo面を通るpath > OK
- TE metric > OK
- Latency-based metric > OK





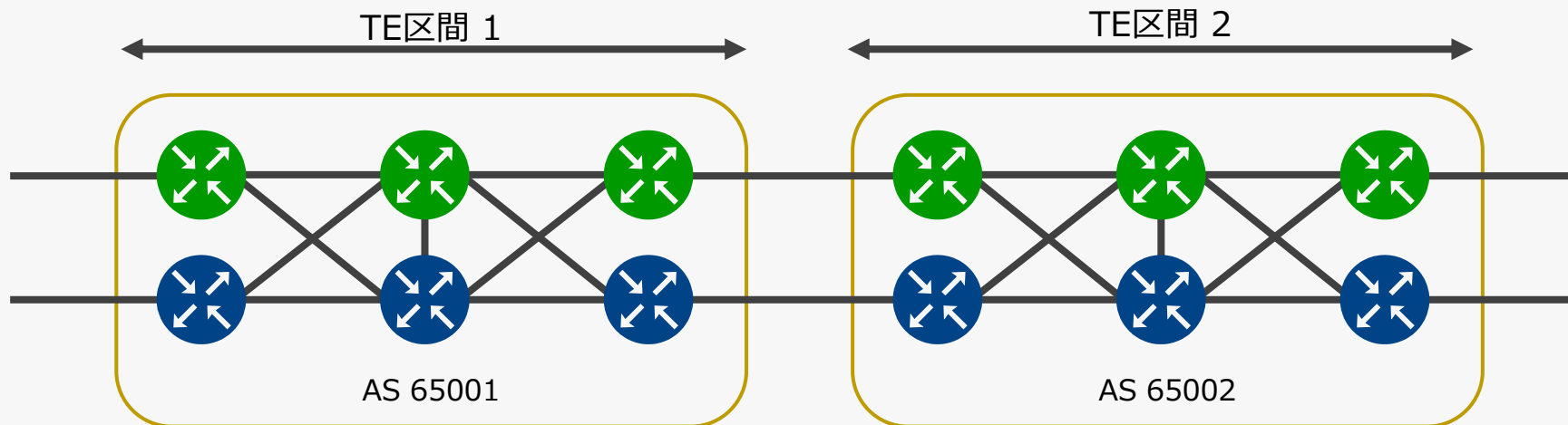
# Inter-AS部分のTE設計の難しさ

- Head-endから見た時のASBRの選択方法が難しい
  - PE1はASBR1/2どちらに送るか、先が見えないのでマクロな判断は不可
  - ASBR1/2もinter-ASの状況をPE1にSR的には伝えられない
    - EPEはVPNラベルと同時選択できない
    - BGP attribute (LP/MED) での制御はできる
    - SR-TEのメトリクスによるASBR選択も一部はできる



## 2. ユースケースを満たせるTEの実装 まとめ

- AS=SRドメインを分割したことにより必然的にTE同士をつなげる
- TEの実装がASで隠蔽されたが、E2Eの真の最適化ができない
- 特にInter-AS部分の選択はSRだけでは難しいこともある

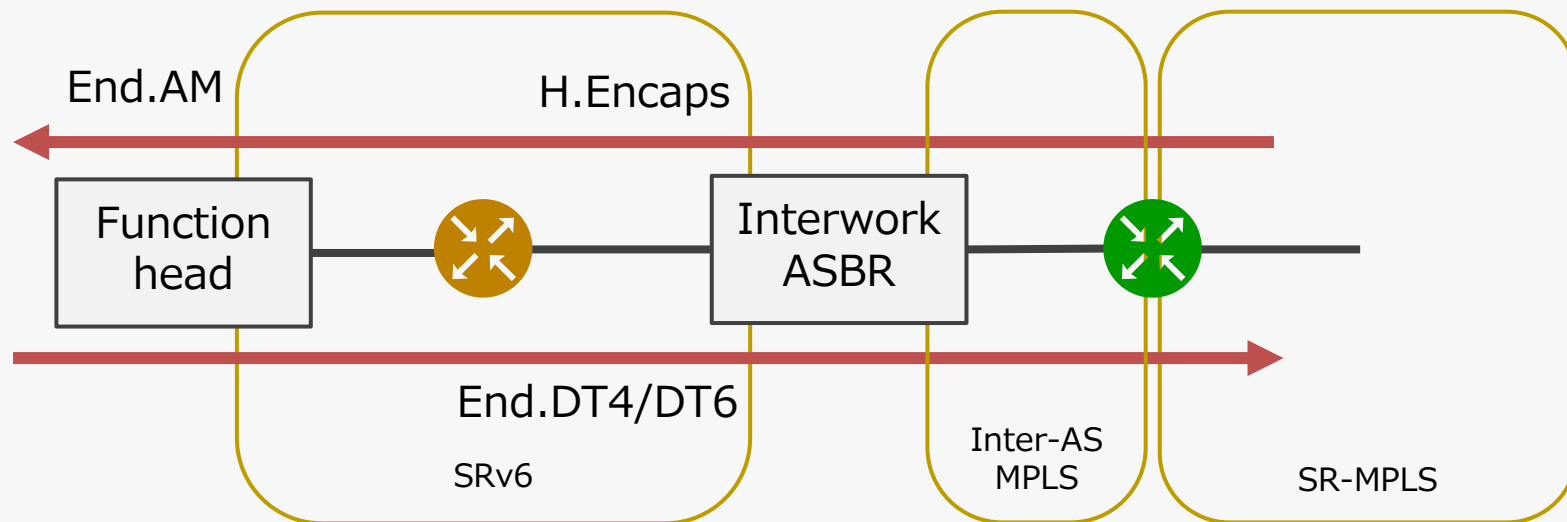


# その他細かい点のつまづき

- SRLBからのラベル割り当て値の差
  - Localなので動作上問題はなく単なる見た目の問題
  
- 遅延の測定結果がメーカーによりオーダーが違うほどの差
  - E2Eの遅延計算がリンク遅延の合計となると、正しく実情を反映できない
  
- FRRが意外に扱いつらいいうえ、inter-AS部分の存在によりE2Eの保護が難しい
  - FRRが動かない組み合わせがあった
  - inter-AS部分の保護ができない課題

# SR-MPLS/SRv6相互接続SFCの実装

- SRv6はEnd.DT4/DT6での終端、SR-MPLSはVPNv4/6の終端を行う
- LinuxにおいてVRFを2つ使い実装
  - 単にSRv6のVPN PEとの通信であれば1つで可能
  - SFCでMPLSに戻すことを考えると2つ必要



# まとめと今後の展望

- 複数AS、複数SRドメインでOption-Bを用いてVPNを実装可能
- プリセットメニューのTEは実装可能
- SRv6 interwork SFCの実装は可能

## 今後のさらなる課題

- PCE等外部ツールでのパスコントロールの可能性

