



“Segment Routing”

**Chasmを越えてついに実用段階へ、そして
これからのNetwork Programmability**

MPLS編

KDDI 株式会社

宮坂 拓也

本発表では…

- MPLS-TEを利用したバックボーンネットワークへSegment Routing(SR)を導入する上での

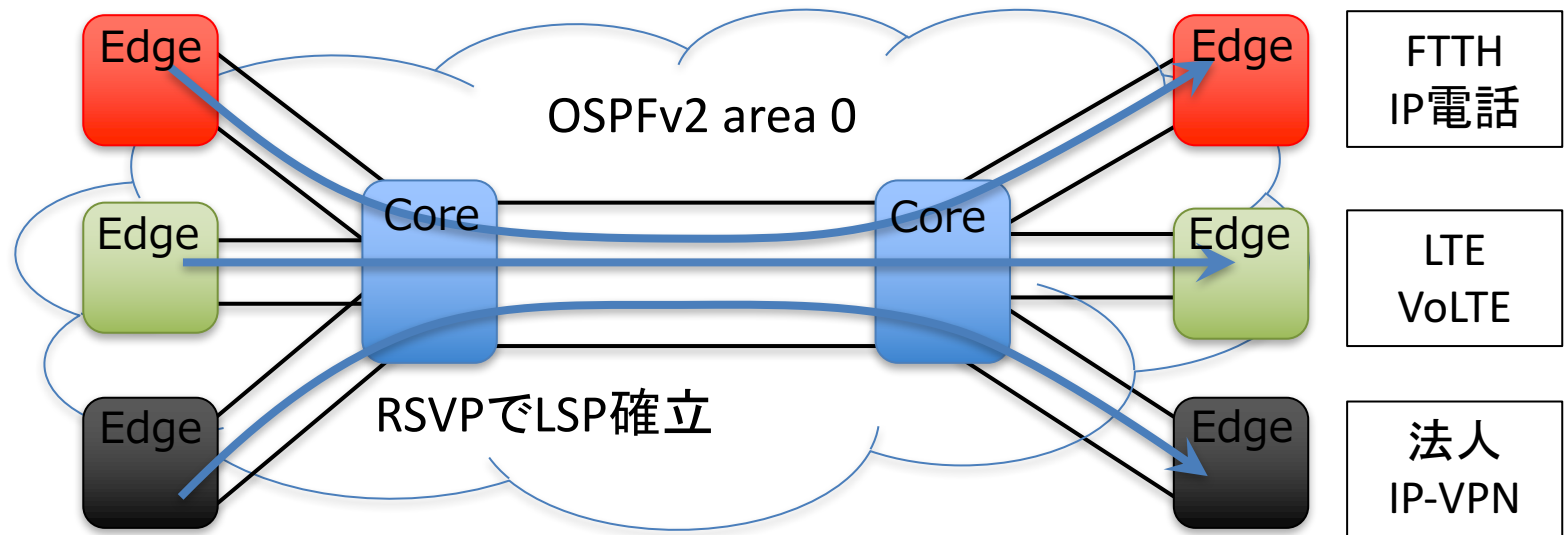
- **モチベーション**：どうして入れたいのか？
- **マイグレーション**：どうやって入れるのか？
- **オペレーション**：運用で気をつけることは？

を紹介します

- 最後に、**ネットワークアーキテクチャのあるべき姿**について思いを馳せたいと思います



現在のKDDIバックボーンネットワーク



1. 各サービスのエッジルーターが他拠点のエッジルーターへ MPLS-TE LSPを確立し、パケットを転送
 - LSPを確立するため、全ルータは同一のOSPFv2 area 0に所属
2. コアルータでは全てのサービスのパケットをMPLSラベルに従って転送する

こんなMPLS-TEネットワークにSRを入れるモチベーションは？

<https://www.janog.gr.jp/meeting/janog36/program/kddi>

モチベーション



マイグレーション



オペレーション

SRを導入するモチベーション

- サービスの観点

- SR ready

- エッジ側でSRを利用したサービスをいつでも展開できるように、コア側は準備しておく
 - SRはSource Routing = エッジ側が主導！

- ネットワークの観点

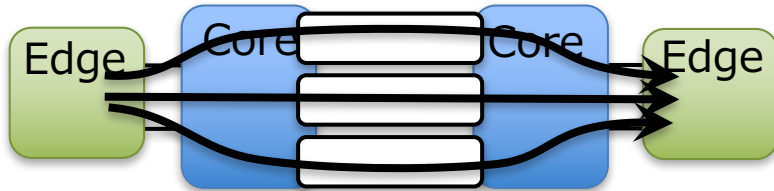
- ネットワークのシンプル化

- RSVP/LDPを使わず、IGP(OSPF, IS-IS)だけ
 - ECMPが簡単（詳細は次ページから）

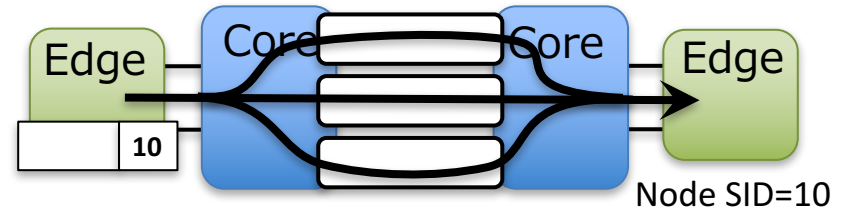


ECMP : MPLS-TE vs SR

MPLS-TEにおけるECMP



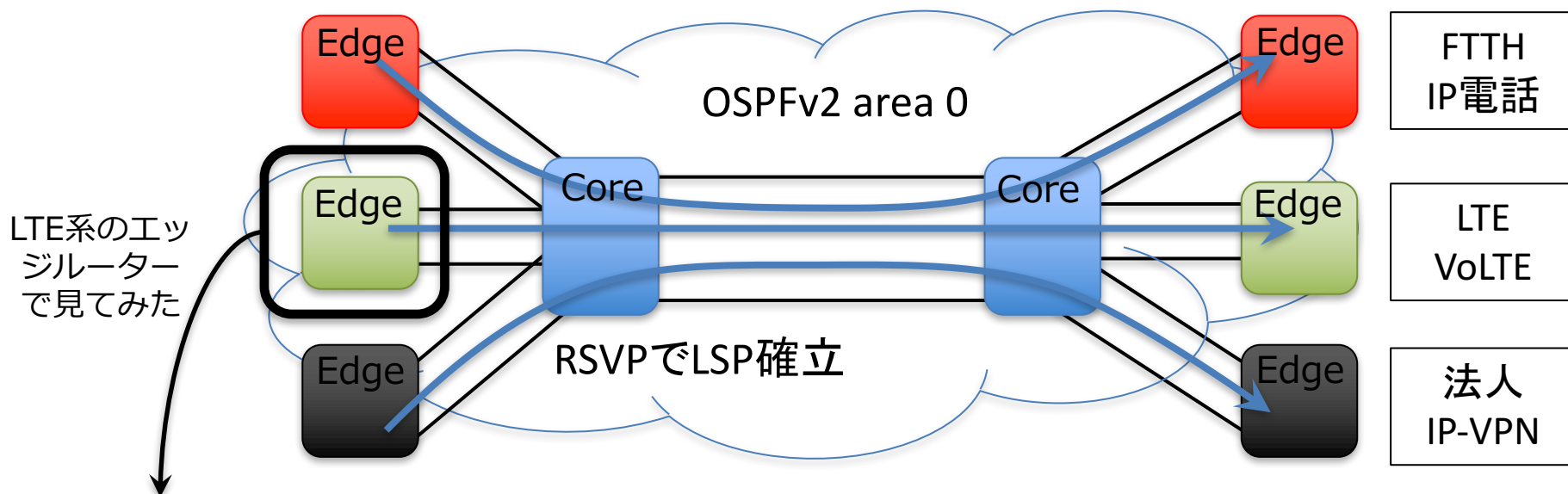
SRにおけるECMP



- MPLS-TEでは、複数リンクにトラフィック分散したい場合、全リンクを通るようにLSPを確立しないといけない
 - LSPの数が膨大となり、複雑なネットワークとなる…
 - Core間でLAGをしてしまうと、特定のリンクだけ通したいという要求に応えられない…
- SRでは、Node-SIDを使うことで勝手にECMPに従い分散してくれる！
 - 特定リンクに乗せたければAdj-SIDを使えばいい！



実際にどれだけ減るのか実機で見てみた



- とあるエッジルーターのMPLS-TE LSP数 : **248本**
 - ECMPするために同一ノード間でも複数LSPを確立
- SR-TEを使った場合のLSP数 : **60本**
 - 同一ノード間は1つのLSPのみでOK
 - 約**1/4**の本数ですむ！

(再)SRを導入するモチベーション

- サービスの観点

- SR ready

- エッジ側でSRを利用したサービスをいつでも展開できるように、コア側は準備しておく
 - SRはSource Routing = エッジ側が主導！

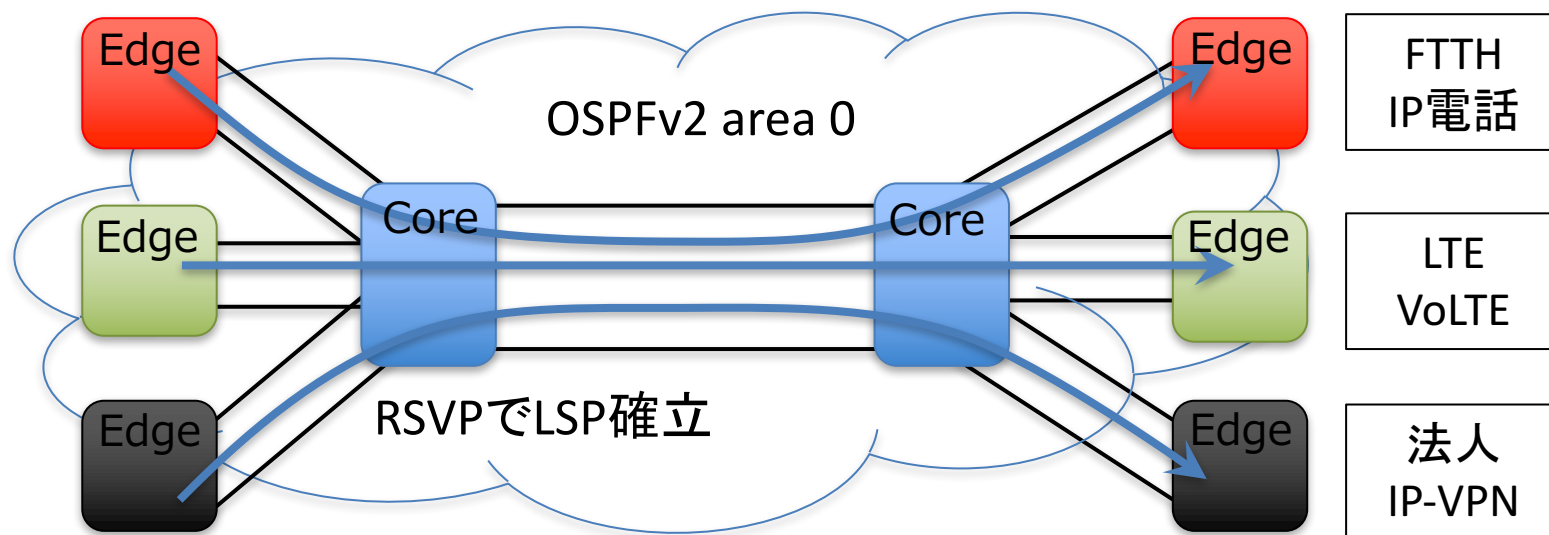
- ネットワークの観点

- ネットワークのシンプル化

- RSVP/LDPを使わず、IGP(OSPF, IS-IS)だけ
 - ECMPが簡単：LSPの本数が激減するみたい！



MPLS-TEからのマイグレーション



- 各サービスのエッジルーターが他拠点のエッジルーターへMPLS-TE LSPを確立し、パケットを転送
 - LSPを確立するため、全ルータは同一のOSPFv2 area 0に所属

求められることは、トラフィックロスなく
SRへマイグレーションすること

モチベーション

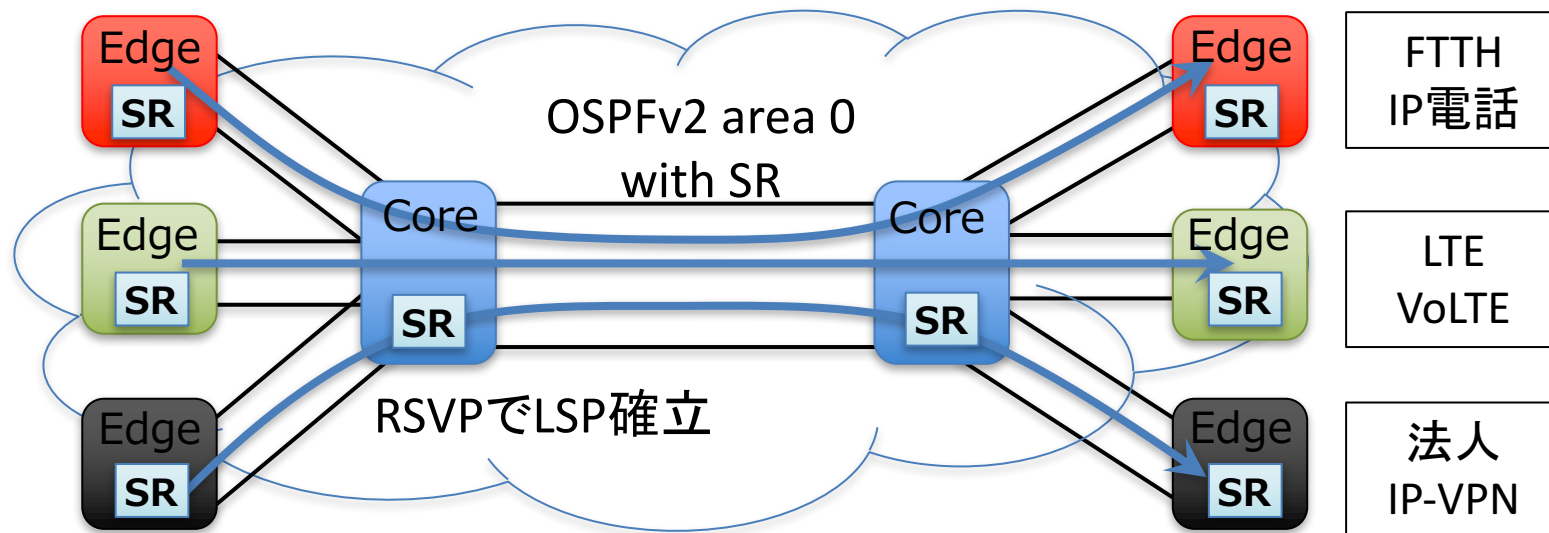


マイグレーション



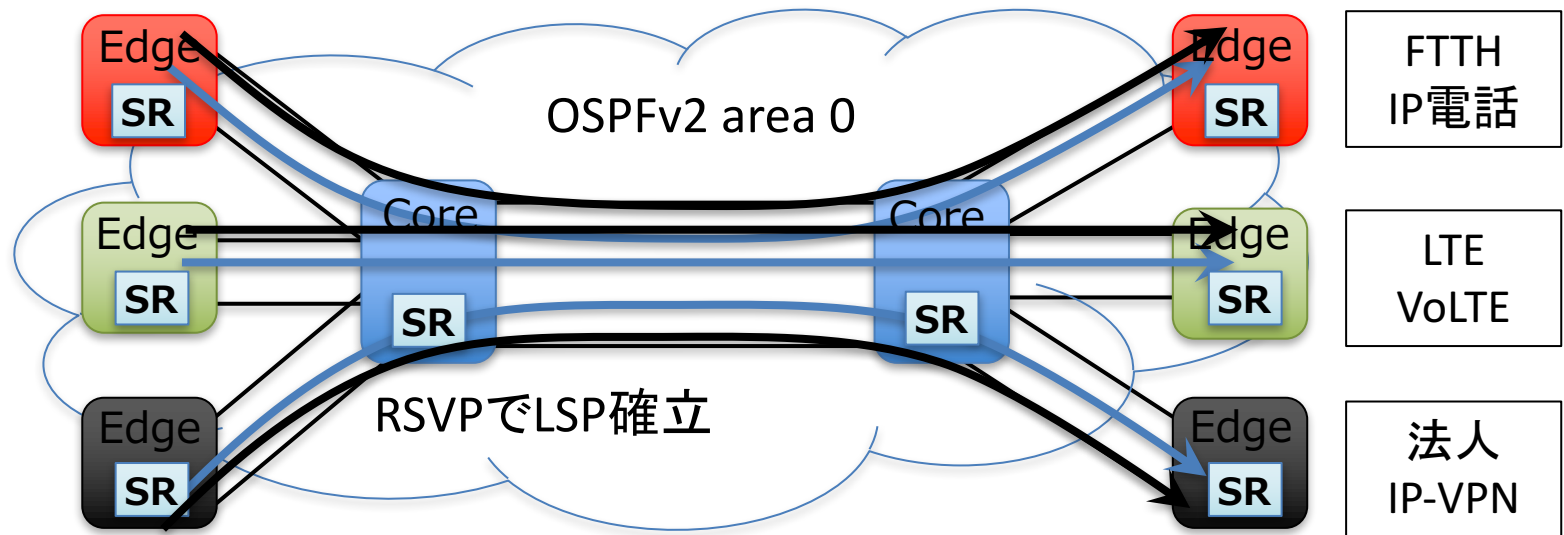
オペレーション

1. IGPでSRを動作



- 各ルーターのOSPFv2において、Segment Routingを有効化し、Node-SID/Adj-SIDなどを広報する
- この時点でネットワークとしてはSRを利用可能となる
 - 残っているのは**MPLS-TE LSPの移行**

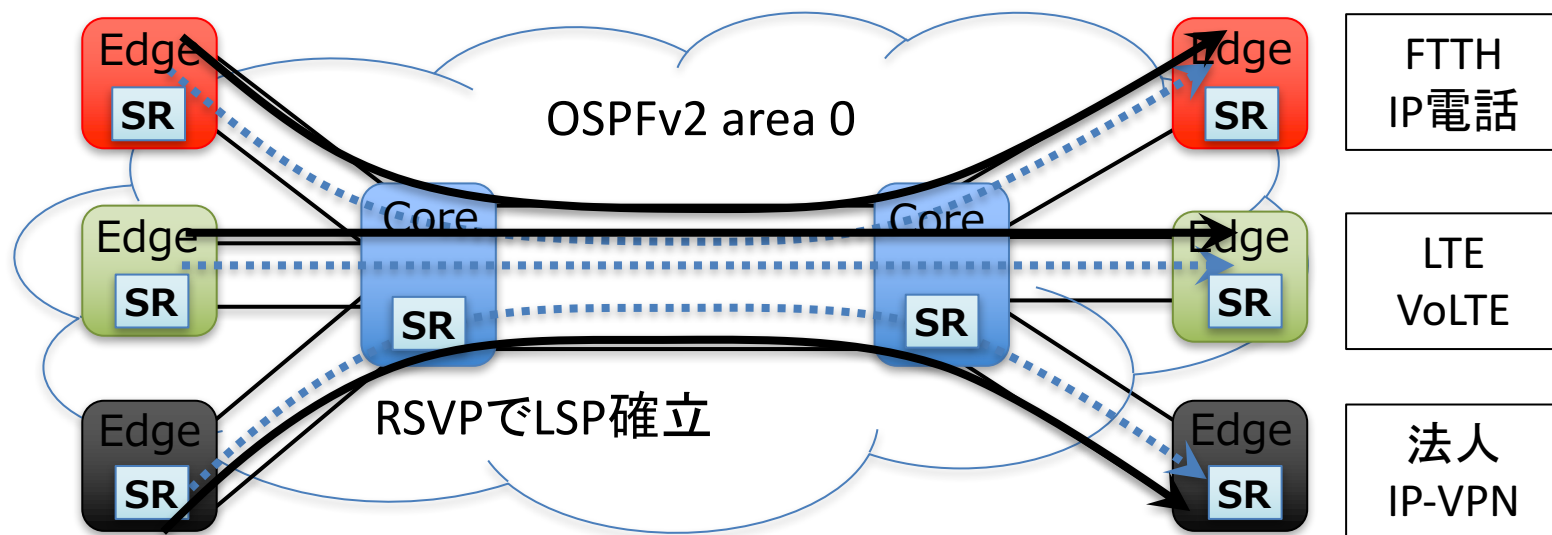
2. SR-TE LSPを設定する



- 各エッジルーターでSR-TEパスを設定する
- 設定方法としては以下 2 つ
 - ルーターにコンフィグとして設定する
 - PCEからSR-TEパスを設定する
- この時点ではSR-TEへトラフィックを流さないでおく



3. SR-TEパスへの切り替え



- 各エッジルーターでMPLS-TEパスからSR-TEパスへ切り替えを行う
- 切り替え後に不要となったMPLS-TEパスは削除する



マイグレーションまとめ

- MPLS-TEからSR-TEへのマイグレーションは比較的シンプルにいけそう！
- しかしながら、MPLS-TEで帯域確保した運用をしている場合は注意が必要
 - **SR-TEそれ自身には帯域確保する機能は無い**が、一時的にMPLS-TEと共存する状態が出てくる
 - 以下の解決策がある
 - 静的にMPLS-TE/SR-TEで使う帯域を分ける
 - PCEで帯域管理を中央制御する
 - SRで利用される帯域をIGPで広報する
 - SR over RSVP-TE
 - SRで消費されるトラフィック量をTEDへ反映する

[draft-ietf-teas-sr-rsvp-coexistence-rec](#)

モチベーション



マイグレーション



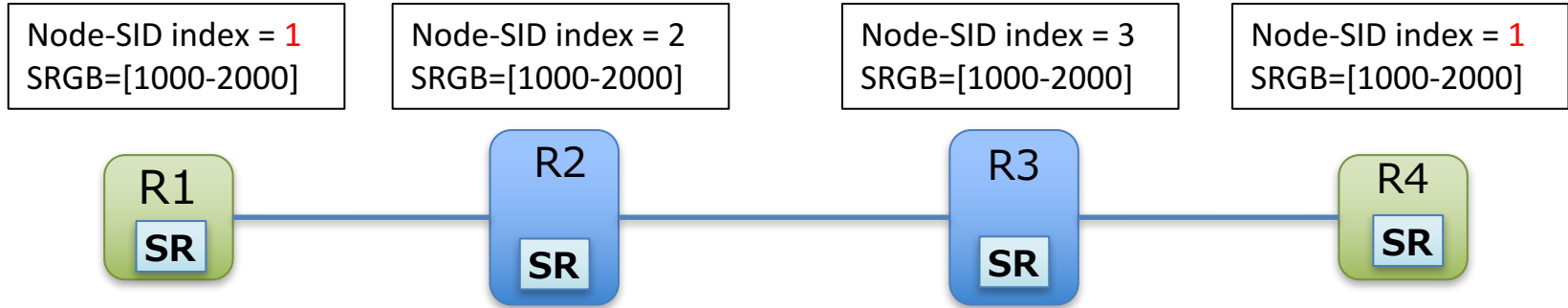
オペレーション

運用で気をつけないといけないことは？

- SR-MPLS運用上の注意点
 - SR-MPLSはMPLSデータプレーンをそのまま使うというが、微妙な違いはある
 - Global SID
 - SRGB
 - その違いに関して、運用上けっこう気をつけないといけない
- また、SRv6とSR-MPLSの相互接続はできる？



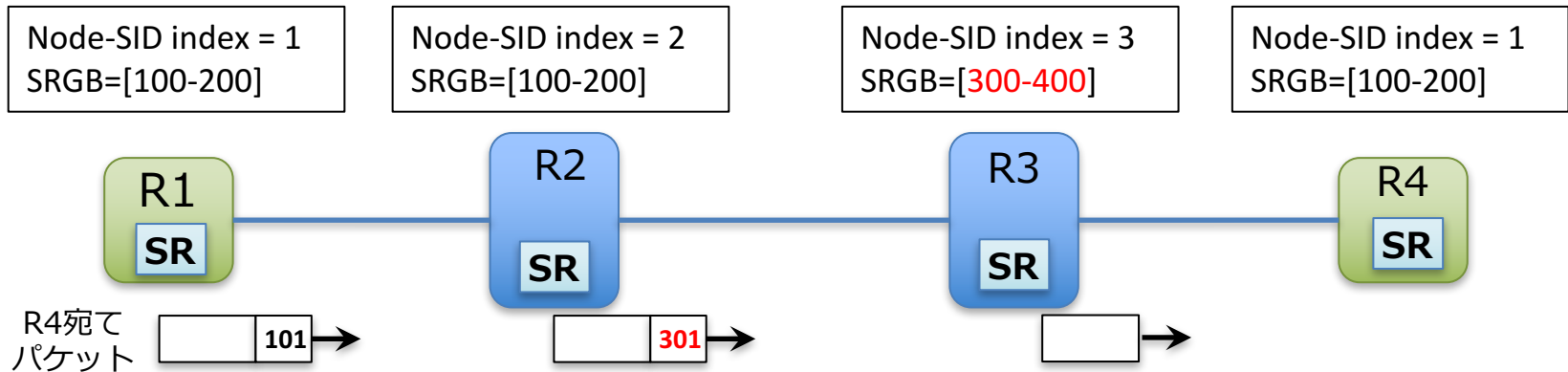
SIDを被らせちゃダメ！



- Node-SIDなどはドメイン内でGlobalな値である必要がある、つまり**“被っちゃダメ”**
 - 上の例ではR1とR4が被ってしまっている
 - 被った場合に、どちらを有効化するかという規定はある
[draft-ietf-spring-conflict-resolution](#)
- しかしながら、コンフィグなどで設定するindex値は、ただの整数値(上のように1とか)なので、管理が大変
 - excel管理表？(いやだー)
 - 何らかの自動プロビジョニングシステムが欲しいよね



SRGBは揃えるべき？



- SRGBが異なると、実際にパケットに付与されるMPLSラベルがホップによって異なってしまう
 - もちろん、これでも問題なく動作する
 - しかし、SRGBを揃えた方が運用は楽だし、SRのdraftにも推奨されている
 - Using the same SRGB on all nodes within the SR domain ease operations and troubleshooting and is expected to be a deployment guideline.
[draft-ietf-spring-segment-routing](#)
- ベンダーによってSRGBのデフォルト値は異なるので、マルチベンダー環境で揃えたい場合は、全ルーターでSRGBの設定が必要

モチベーション

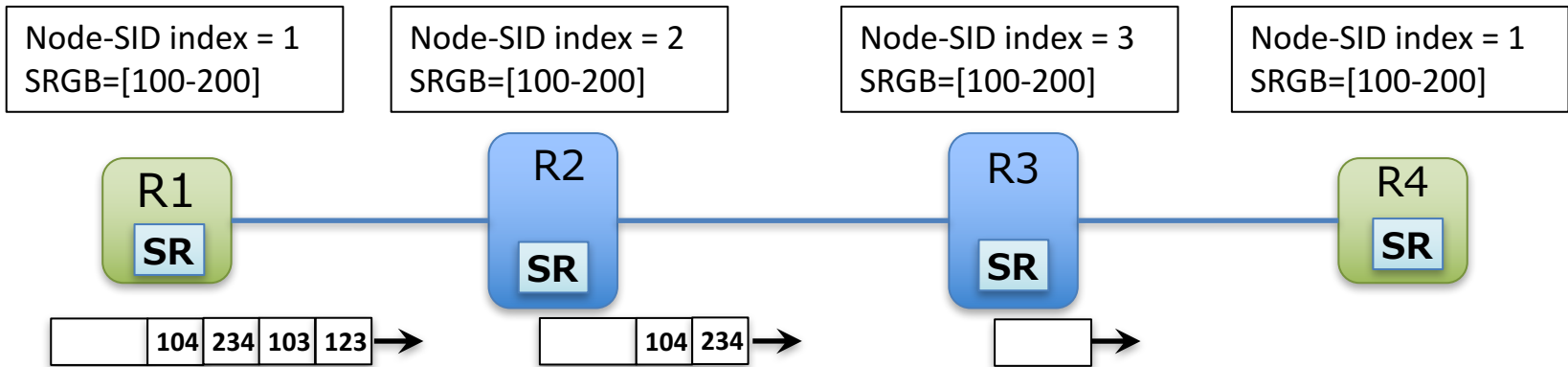


マイグレーション



オペレーション

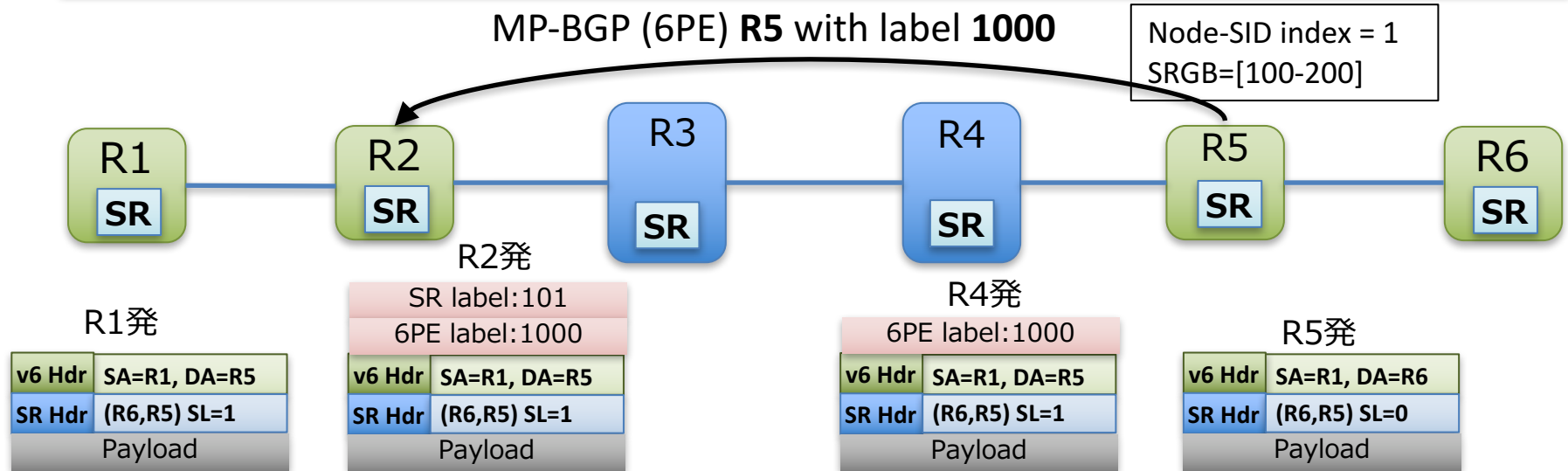
ラベルスタック数の増加



- SRではSource側で何段もラベルスタックできる
 - SR-TEやService Function Chaining(SFC)をSRでやるなど…
- MPLSラベルスタック数が増大することで以下の問題がある
 - ルーターが対応できる最大ラベルスタック数を超えてしまう
 - IGPなどで最大ラベルスタック数(MSD)を広報してそれを超えないようにラベルスタックするという提案は有る [draft-ietf-ospf-segment-routing-msd](#)
 - インターフェースのMTUを超えてしまい、dropする
 - MTUを十分大きくするしかなさそう？



SRv6→SR-MPLS→SRv6



- 両端のエッジでSRv6をおこない、コア側はSR-MPLSのみ動作する場合、6PEなどを用いて通常通りIPv6経路情報を広報すれば仕様上は問題なく動作するはず！
 - むしろ、SR-MPLSでなく通常のMPLSネットワークでも問題なくSRv6の“橋渡し”ができるはず！（これがSRv6のいいところでもある）

モチベーション



マイグレーション



オペレーション

SRオペレーションまとめ

- SR-MPLS運用上の注意点
 - SIDを被らないようにすること
 - SRGBはドメインで揃えた方がいい
 - ラベルスタッキングの増大への対処を気をつける
- SRv6とSR-MPLSの相互接続はできそう！



最後に：あるべきネットワークアーキテクチャ

- コアルーター側はなるべくシンプルに
 - そういう観点ではSegment Routingは最高だと思います
 - ProtocolはIGPだけでOK
 - LFIBは基本的にはNetwork Topologyの情報だけでOK
 - MPLS-TEではエッジから張られる全てのLSPのstateを持ってしまう
- 複雑なことはエッジルーター側で行う
 - サービスに特化したTEポリシーを実施する
 - ルーターだけでは厳しい場合は、PCEなどを用いる

