Create New Value by

JANOG45
サービスプロバイダーでのEVPN導入事例

株式会社NTTPCコミュニケーションズ

大塚 悠平

NTT PC COMMUNICATIONS

01 プロフィール

大塚 悠平　株式会社NTTPCコミュニケーションズ

自社VPNサービスインフラの設計/開発/構築/運用

JANOG40くらい～一般参加
JANOG45　初登壇＆初スタッフ(ハッカソン)

01 はじめに

モチベーション
- 自身が携わってきた技術についてフィードバックしたい
- VPN技術に関する情報、特に失敗事例はあまり見かけない

共有したいこと
- 具体的にどんなNWを作ったか
- EVPN導入あたって悩んだあるいは苦労したこと

議論したいこと
- EVPNは利用者によって要求事項や利用方法が大きく異なる（と思われる）ため、様々な導入事例や知見を伺いたい
- EVPNが要件に合わなかった、あるいはSRv6など特定機能の実装待ち、など導入を見送っている方のご意見も伺いたい

01 EVPNとは？

EVPNはMAC等をマルチプロトコルBGPを介して伝搬するVPN技術

NTTコミュニケーションズさん資料 (MPLS Japan2015)
https://mpls.jp/2015/presentations/NTTCom_EVPN.pdf

古河電工さん資料 (Internet Week 2016)
https://www.nic.ad.jp/ja/materials/iw/2016/proceedings/t05/t5-kamitani-2.pdf

ネットワンシステムズさんブログ
https://www.netone.co.jp/report/column/column1/20160308.html

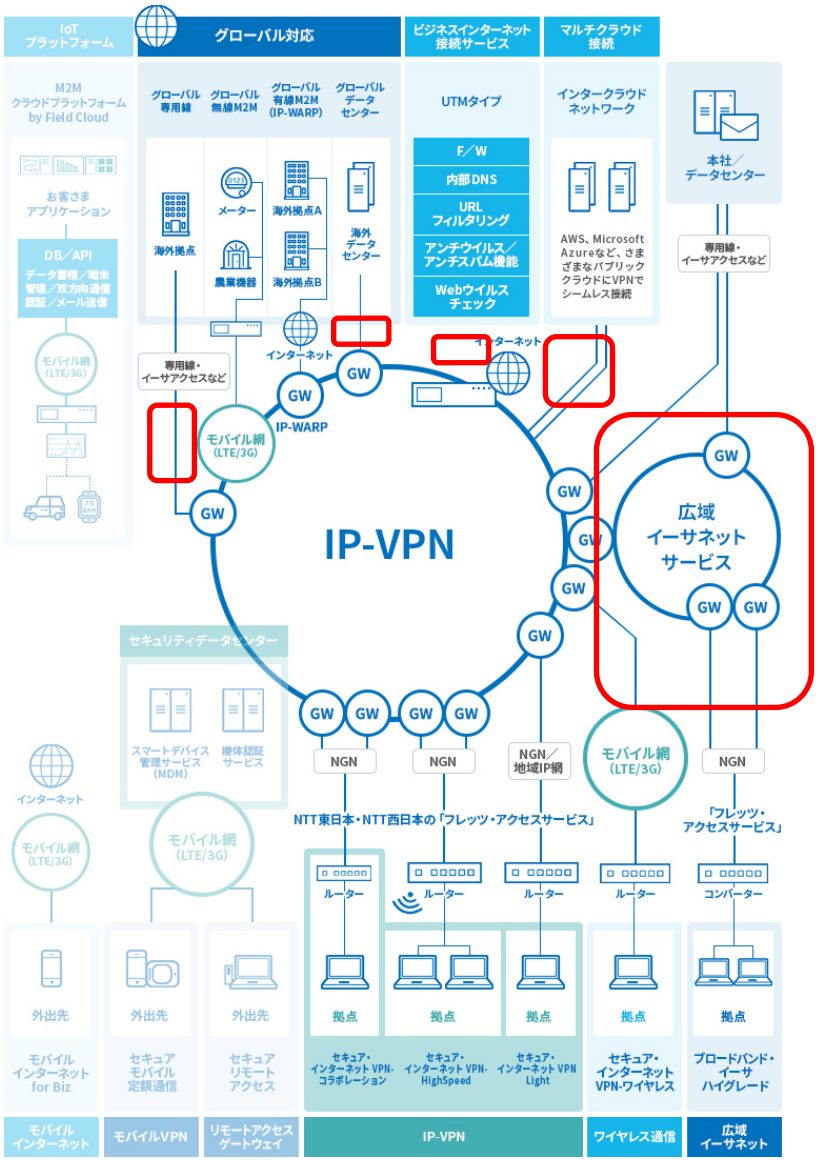
Yahoo!さんブログ
https://techblog.yahoo.co.jp/infrastructure/evpn/

01 Agenda

- 当社サービスとL2VPNについて
- 既存NWが抱える課題
- 方式検討・比較
- 新NW
- 既存NWとの接続、マイグレーション方式
- 検証や検討で苦労した点
- 今後の展望
- 議論

当社サービスとL2VPNについて

広域イーサネットサービス、各基盤間接続のためL2VPNを利用



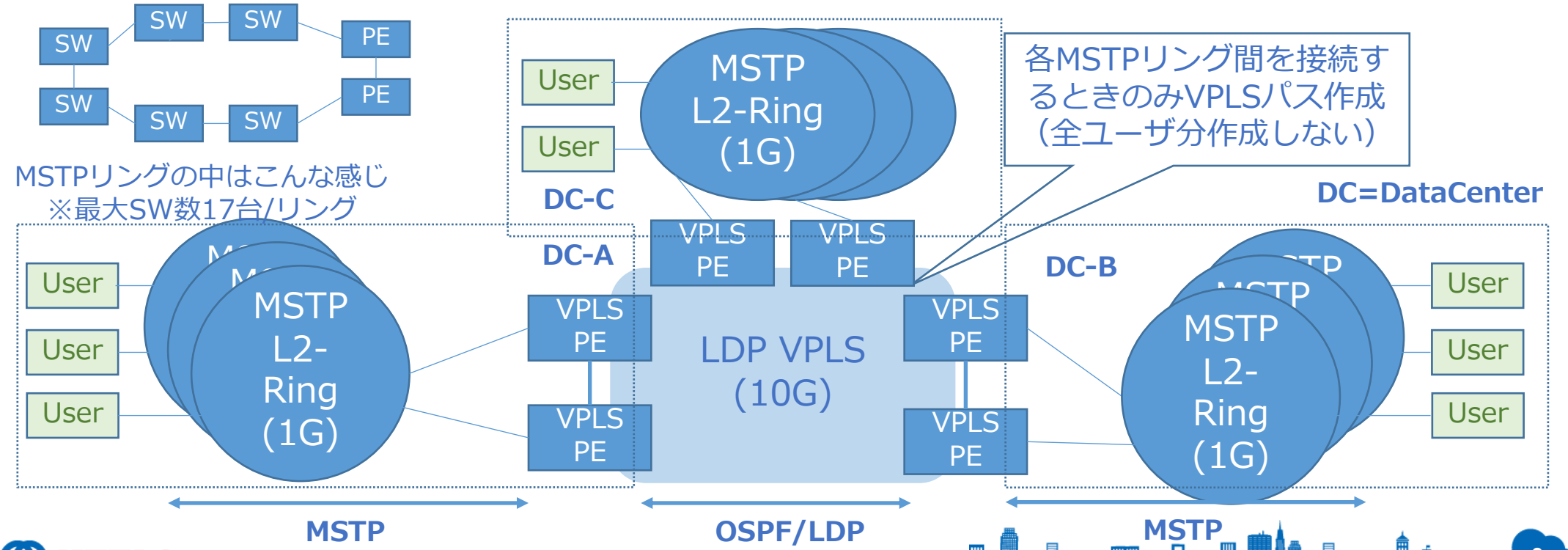
L2VPN入ってます

01 既存NWが抱える課題

1.スケーラビリティ
　各MSTPリングにおけるトラフィック容量不足やMAC数過多によるFIB逼迫

2.運用
　止められない需要対応(リング増設)とVPN毎のPseudowire設計/設定(N^2問題)

3.品質
　特大L2ループ障害発生に伴いL2ループ対策が急務に



01 課題解決に向けた方式検討

1.スケーラビリティ
　大容量化（100Gベース）かつ拡張性に優れたNWアーキテクチャの採用

2.運用
　大容量化による増設対応の削減、アーキテクチャ変更による設計・設定稼働の削減

3.品質
　徹底したL2ループ対策による品質向上、スパツリ排除

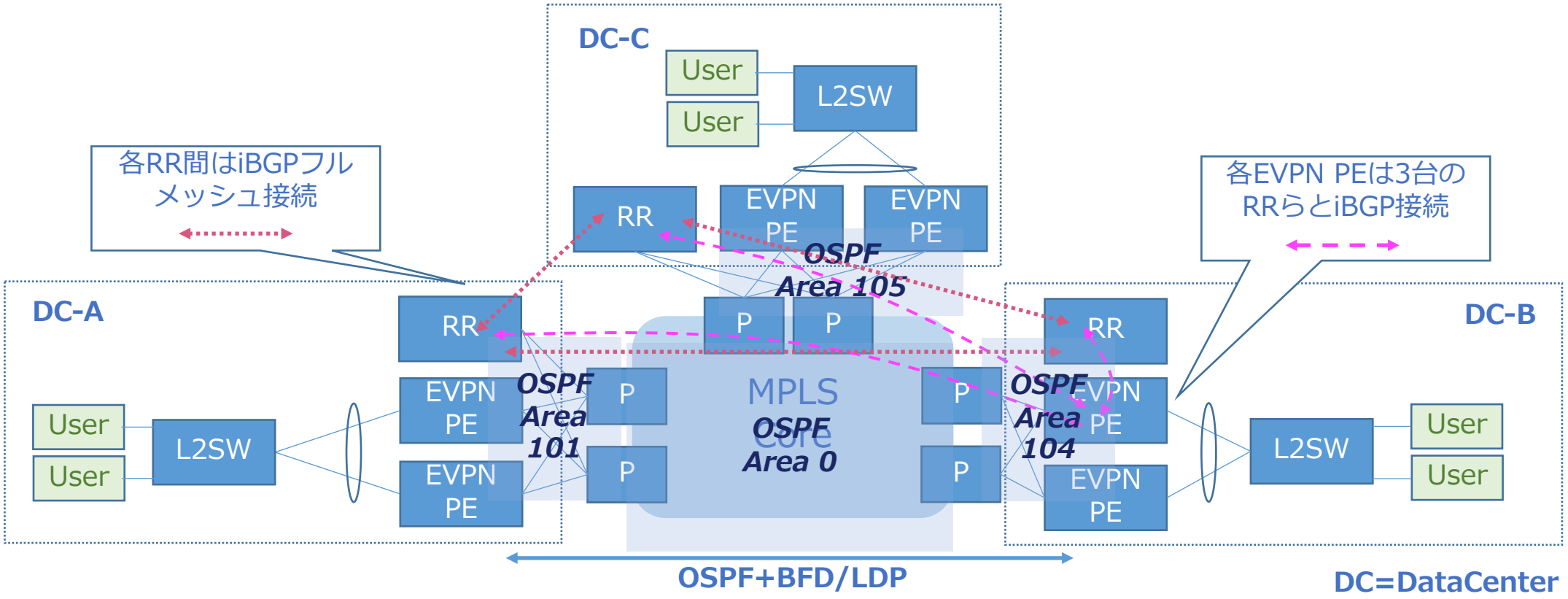
方式	Pros	Cons
VPLS	・成熟したプロトコル	・PW数N^2問題　※LDP方式の場合
EVPN-MPLS	・EVPN-Multihoming利用可 ・既存MPLSインフラを利用可	・実装機器がやや高価
EVPN-VxLAN	・EVPN-Multihoming利用可 ・多くのOSでサポート（Whitebox NOSも視野に）	・VxLAN用IPNWを新たに用意する必要がある ・VTEPスケール問題 ・VxLANのみ実装製品はタグ処理機能が不足
EVPN-PBB	・EVPN-Multihoming利用可 ・MAC抑制による高スケール	・実装機器が少なくやや高価

コントロールプレーンはEVPN一択、検討の末「EVPN-MPLS」方式を選択

01 新NW

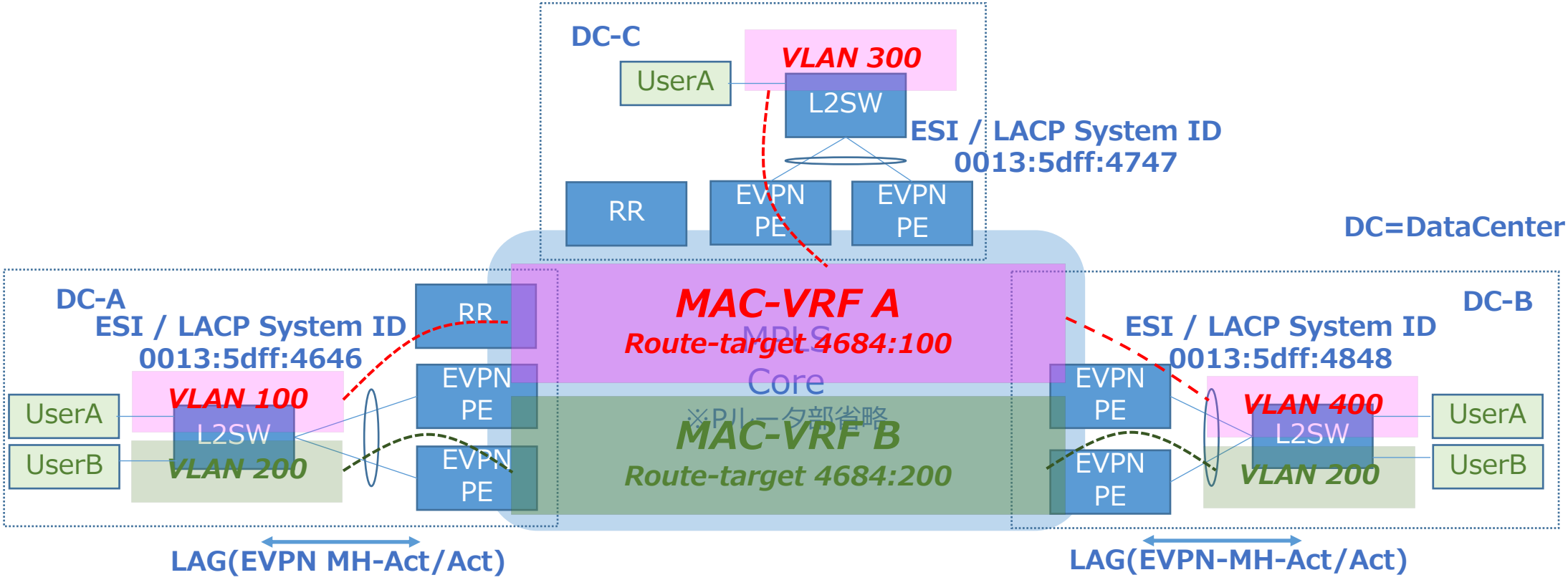
- ルーティング設計
- EVPN設計

01 新NW（ルーティング設計）



- P～PE間はLDP接続、MPLS Core仕様によりOSPFマルチエリア設計とした
- 網内バースト耐性のためOSPFはECMP(網内完全帯域２倍設計)とした
- 各EVPN PEでは網内の誤ハッシュによるフレームリオーダリング回避のため MPLS PW Control Word付与必須とした
- MP-BGP(EVPN)は全てiBGP接続、RRはPEと機能分離・階層化せず

01 新NW（EVPN設計）

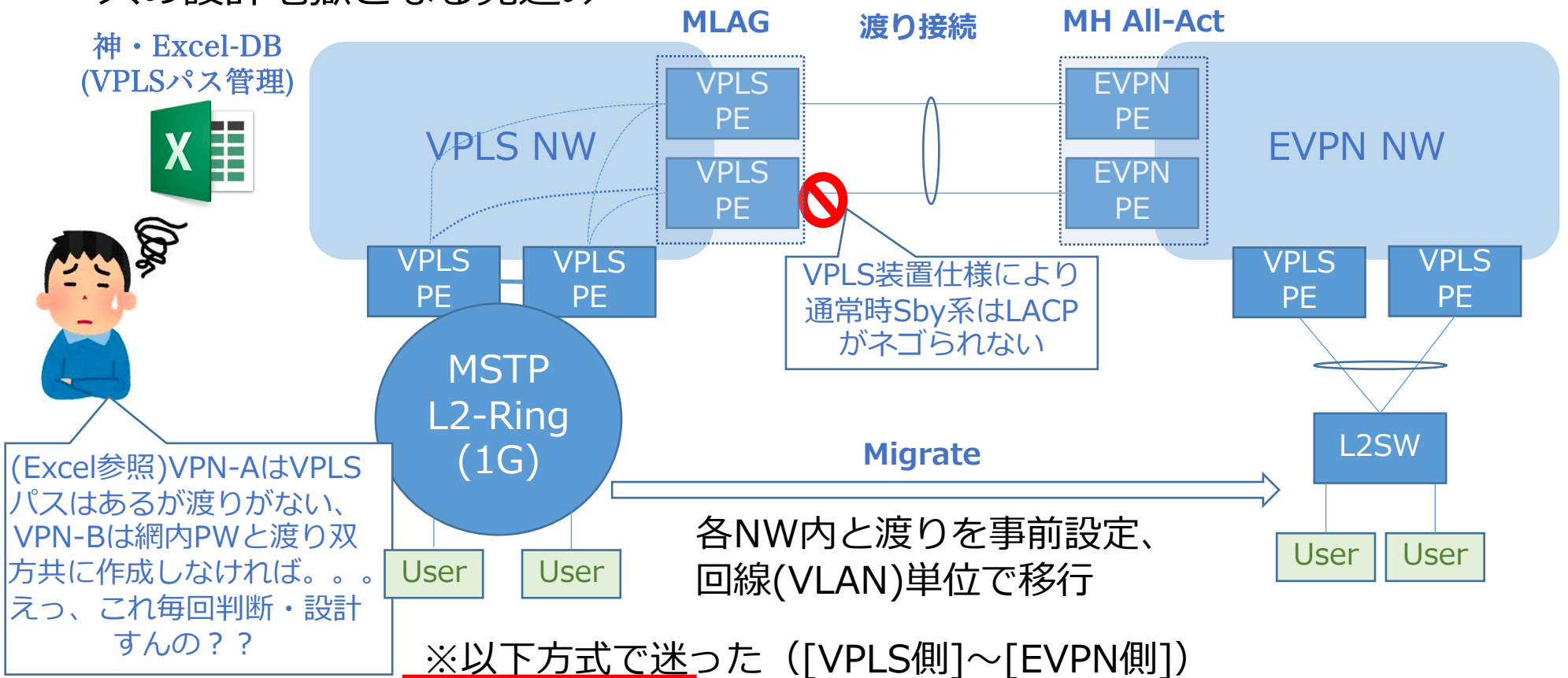


- ユーザ要求に応じて自由な接続形態を提供する必要があるためEVPN Service Interfaceは「VLAN-Based」を採用、MAC-VRFあたり1つのRoute-Target値を付与

- 網内バースト耐性とL2ループ耐性を考慮しEVPN-MHは2台のPEによる「All-Active方式」のみ採用、※Single-Active方式の場合、L2SW(CE)にてPE毎にLAGを分ける必要があるため、L2SW側が単なるLAGで良いAll-Activeの方が安全と考えた

01 既存NW(VPLS NW)との接続、マイグレーション

- マイグレーションのためNW間を接続、方式としてMLAG～MH接続を選択(※)
- ユーザ(VLAN)を1つずつ、またはある程度纏った単位で移設していく
- VPLS NWでは全VPNがVPLSを保持しない（MSTPリング間を渡るユーザのみVPLSパスを作成）ため、移設の都度VPN毎に人手でExcelDBを元にしたVPLSパスの設計地獄となる見込み



※以下方式で迷った（[VPLS側]～[EVPN側]）
①MLAG～MH、②STP-MH、③直繋ぎ～MH、④VPLS延伸

01 苦労話

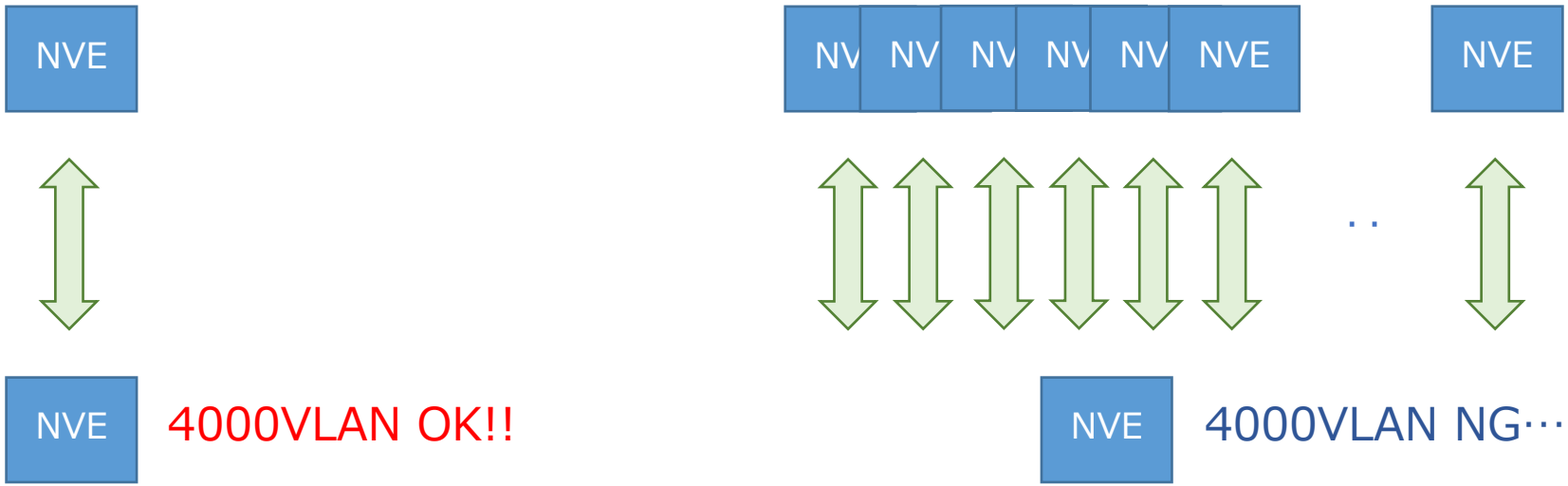
①機種選定の苦労：
　マーチャントシリコン搭載製品のVxLAN性能

②検証の苦労：
　切り替え時の性能

③運用開始後の苦労：
　ヘアピンできない

①機種選定の苦労
マーチャントシリコン搭載製品のVxLAN性能

問題：マーチャントシリコン搭載製品のEVPN-VxLAN実装ではスケールが我々の要求を満たさなかった（最大リモートVTEP数と最大VLAN収容数が反比例）



※2017-2018年頃の状況であるため、現状改善されている可能性があります
※要求スケールによっては当動作が全く問題とならないケースがあります

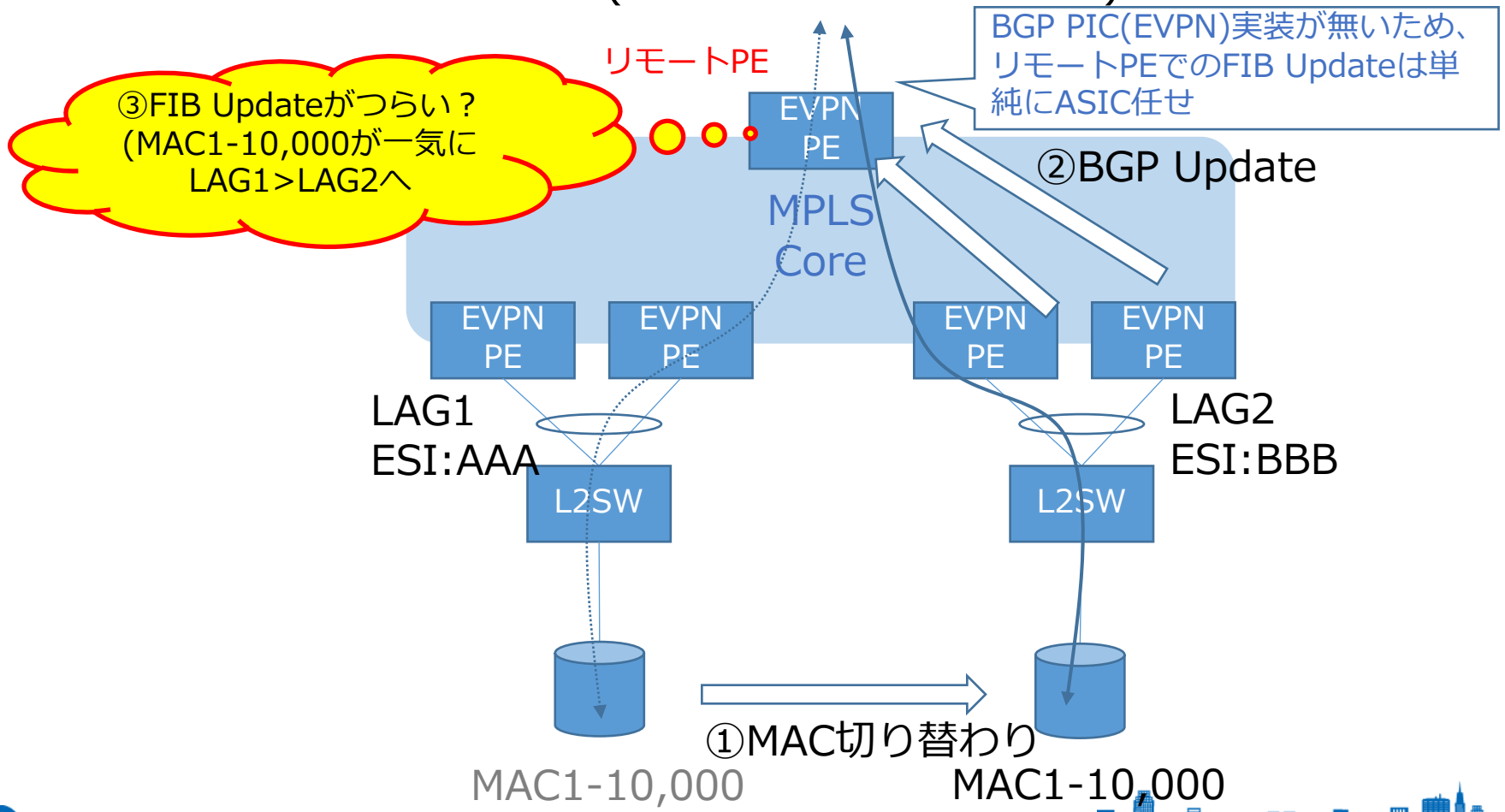
①機種選定の苦労
マーチャントシリコン搭載製品のVxLAN性能

対処：マーチャントシリコン搭載製品のEVPN-VxLAN実装を採用しなかった（EVPN-MPLSを選択）

※今後一切EVPN-VxLANを採用しないというわけではない

②検証の苦労：切り替え時の性能

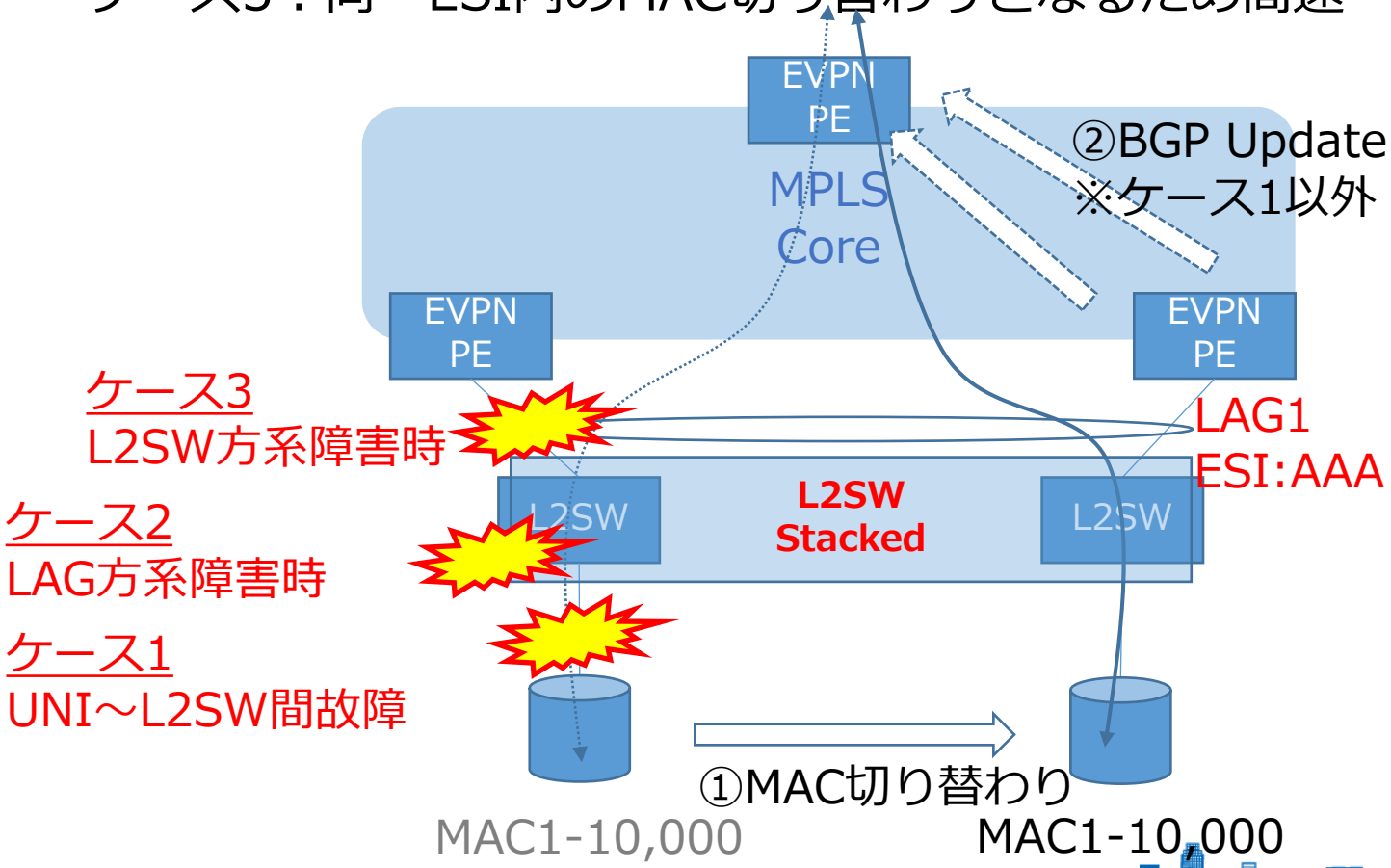
問題：一定数のMACが異なるESを跨いで切り替わる際に時間がかかった
→異なるESを跨ぐMAC切り替わりによるBGP Updateを受けたリモートPEでのFIB Update処理がつらそうだった。
→同ESI内で切り替わる(Mass-Withdrawされる)動作は問題無かった



②検証の苦労：切り替え時の性能

対処：一部L2SWをStack化＆LAGを単一ESI化、いずれのケースも問題無く切り替わった。

ケース1：そもそもローカルPEからBGP Update飛ばない
ケース2：同一ESI内のMAC切り替わりとなるため高速
ケース3：同一ESI内のMAC切り替わりとなるため高速



③運用開始後の苦労：ヘアピンできない

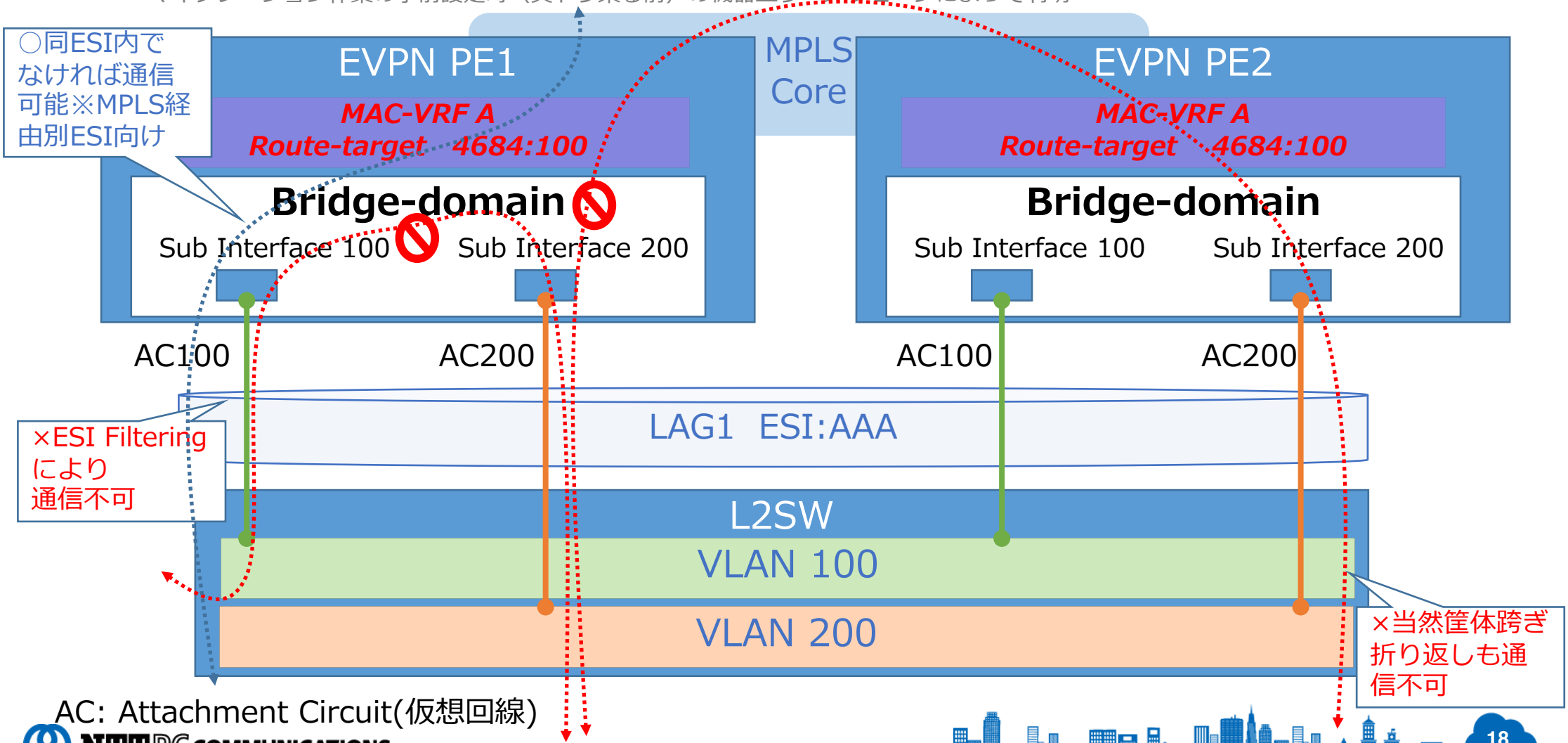
ヘアピンって何？イメージは↓まっすぐ進んでそのまま返ってくるよ

③運用開始後の苦労：ヘアピンできない

”今回で言うヘアピン”=「単一LAG（ESI）内における折り返し通信」

問題：EVPN-MHの機能(Split-Horizon / ESI Filtering)によって、EVPN Multihoming構成時は同LAG内AC(※)の折り返し通信（ヘアピン）が不可

マイグレーション作業の事前設定時（実トラ乗る前）の機器エラーメッセージによって判明



AC: Attachment Circuit(仮想回線)

③運用開始後の苦労：ヘアピンできない

対処：検討中。。。

- 恒久的にはヘアピンをなくす（設計で回避する）ことを検討中
- draft-sajassi-bess-evpn-ac-ware-bundlingがイイ線いってたがタグ処理に制約が我々の要件を満たさなかった、、、
- 独自のExt-Commを定義、同LAG内ACをユニークに判別＆指定したヘアピン通信を許容するような実装が実現出来れば解決？
- 当ケース（ヘアピン）を現L2VPNサービスで実現している事業者さんは少なくは無いはず。。？困りませんか？
- 良い案があれば是非ご意見・コメントください！！

01 今後の展望

Create New Value by CDI

L3VPNとの統合は視野に入れず

　L3機能はEVPNよりも既存IP-VPN技術(RFC4364)がシンプルかつ高機能、EVPNはあくまでL2用途に留めることとなりそう。
※ただしAnycast GWなど特定機能目的で部分的にEVPN(L3)導入する可能性有

自動化

　Stackstorm+Ansibleを用いた設定自動化など

モニタリング高度化

　Telemetryを用いたL2ループ検知＆自動対処など

01 おわりに

<u>共有させて頂いたこと</u>

- 当社での課題と解決方法、具体的にどんなNWを作ったか
- EVPN導入あたって悩んだあるいは苦労したこと

<u>議論したいこと</u>

- EVPNは利用者によって要求事項や利用方法が大きく異なる（と思われる）ため、様々な導入事例や知見について議論したい
- EVPNが要件に合わなかった、あるいはSRv6など特定機能の実装待ち、など導入を見送っている方のご意見も伺いたい

御清聴ありがとうございました！