

Interop Tokyo 2019 ShowNetにおける サービスチェイニング構築を通して見る、 SRv6の実践的知見

Interop Tokyo 2019 ShowNet NOCチームメンバー
上野 幸杜

本日の発表内容

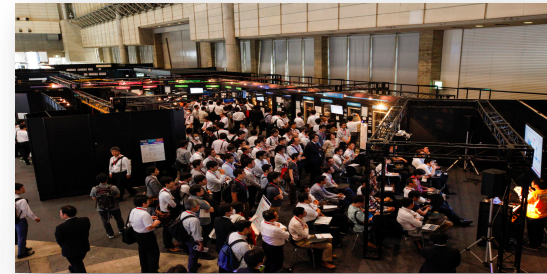
1. ShowNet 2019における
SRv6 Service Chainingのアーキテクチャ
2. SRv6最新動向と標準化の場へのフィードバック
3. ShowNet 2020 STMプログラム

本日の発表内容

1. ShowNet 2019における
SRv6 Service Chainingのアーキテクチャ
2. SRv6最新動向と標準化の場へのフィードバック
3. ShowNet 2020 STMプログラム

Interop Tokyo

- 世界最大のネットワーク機器と技術の展示会
 - 1986年米国モントレで開催されたカンファレンスイベント「TCP/IP Vendors Workshop」が始まり
 - 日本では毎年6月に幕張メッセで開催
 - 来場者約14万人



SHOWNET

EVOLVE INTO
THE NEXT GENERATION



- “I know it works because I saw it at Interop”
 - 産業界、学術、研究機関からトップエンジニアが集まり、Interopで構築される世界最大のデモンストレーションネットワーク
 - 2年後、3年後に業界に浸透する技術に先駆けて挑戦
 - 様々な技術の相互接続性検証の場
 - 最新技術を実装しながら安定したサービスを出展ブース・来場者に提供
 - 出展社や来場者へのネットワーク提供
 - “Live” Network

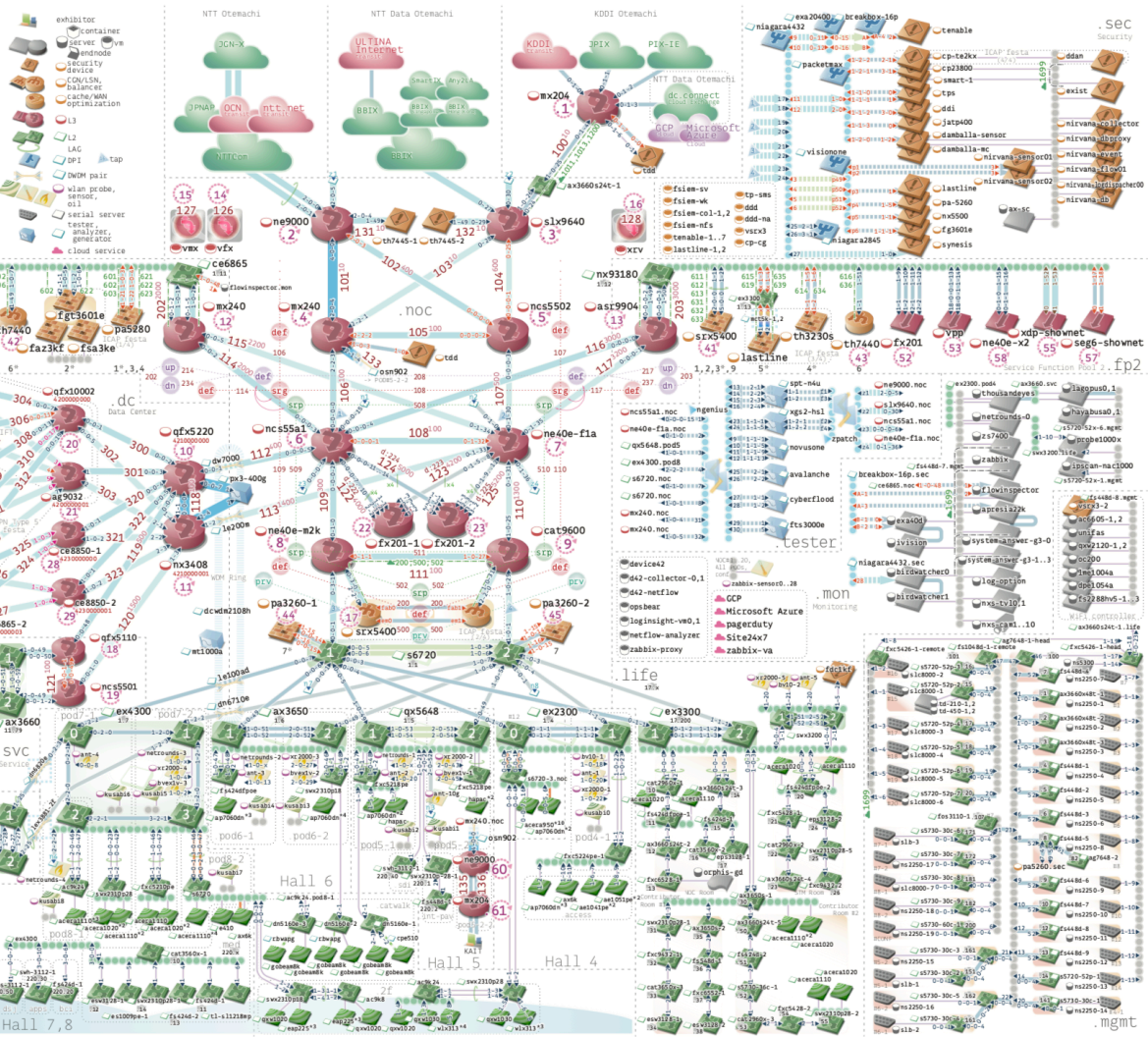


SHOWNET
 EVOLVE INTO
 THE NEXT GENERATION
 as of 2019/6/10 17:45

Device ID: 45.0.0.X/32
 Loopback Address: 45.0.0.X
 P2P Address: 45.0.X.XYZ
 Block: 45.0.X.XYZ/30
 P2P VLAN ID: XYZ (Link ID)

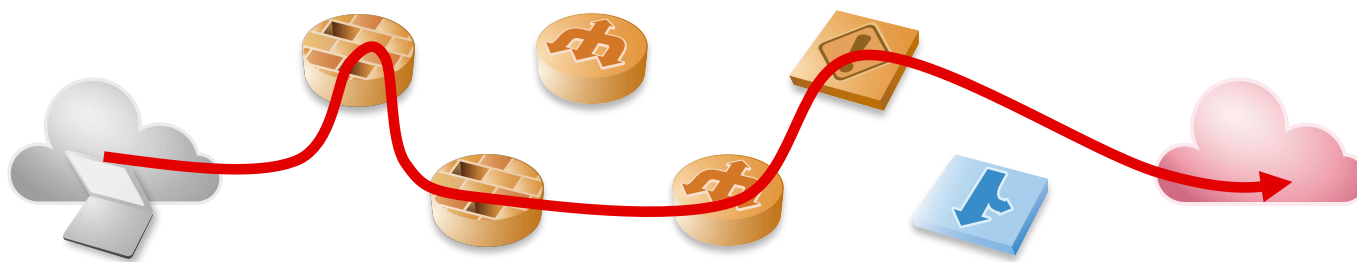
SRV6 VRF Link IDs
 up: upside
 dn: dnside
 bak: backhole
 def: default
 scp: srv6 private
 prv: private
 stg: srv6 global

NHMT Address: 172.16.XY/16 or 172.16.0.X



サービスチェイニング

- ネットワークの機能(Service Function)を鎖のように連携させるアーキテクチャ
 - サービス構成とネットワーク構成の分離
 - ユーザごとのきめ細かなサービスの適用と制御
 - 自動迂回・フェイルオーバーなどの耐障害性



ShowNetにおける挑戦

- ShowNet 2016までは柔軟性に分のあるSDN/NFV技術を採用
- 2017-2018ではシンプルさ・スケーラビリティに分のあるIP経路制御技術に立ち返り、サービスチェイニングに最適なバランスを模索

SDN/NFV@ShowNet 2017の概要

- VPEEによる出展者依存
 - NAT, サービス識別子の付与
 - Juniper Networks, vSRX
- 3DのVNFの描
 - DPI/トラフィック分析
 - Palo Alto Networks, PA-VM
 - Cisco Systems, CSR1000V
 - DDoS Mitigation
 - A10 Networks, Thunder 6435Sys
- OpenFlow Switch
 - NEC PF5248, PF5459

OpenFlow Security

SDN Security

OpenFlowで"連結"

BGP Flowspecで"乗り換え"

ASを超えたネットワーク接続の自動化

ShowNet 2017の選択: サービスチェイニングを前提としたバックボーン

- バックボーンはシンプルで安定したIPネットワークで構築
- IP経路制御によるService Chaining@ShowNet 2018

ファンクションプールとは

- ネットワーク機能を集約した"貯水槽"
- ShowNetでは冗長のため2つのファンクションプールを構築

個別のネットワークとチェーンの構築

- VRFを用いてFunctionごとに個別のIPネットワークを構築
- 上流のVRFには0.0.0.0/0を、下流のVRFにはユーザのPrefixをRoute Leak
- 各VRFのペアはOSPFで経路交換、Leakされた経路をさらに再配布
- 各VRFのペア間には迂回用のバックアップリンク(高コスト)を用意

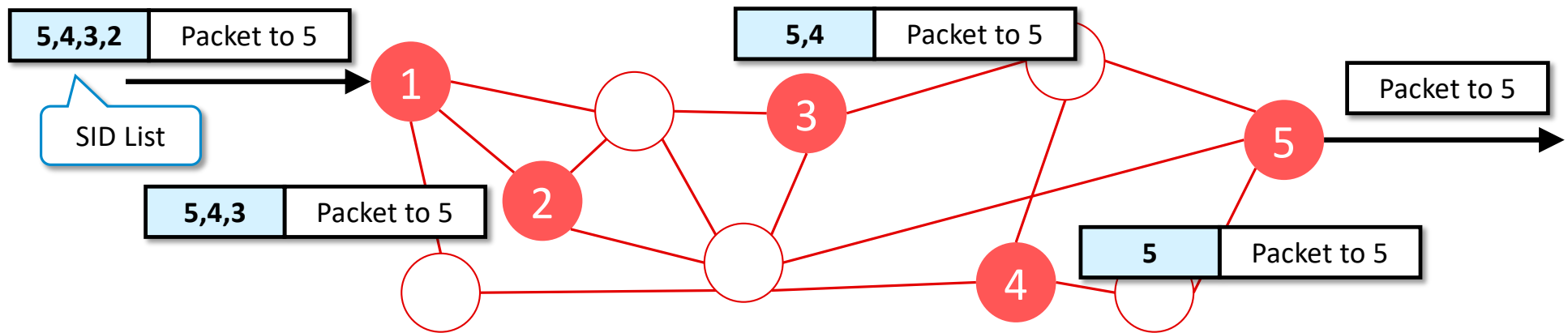
BGP Flowspecによる経路の"上りの場合"

新しいアプローチ

- 次世代のネットワークには
シンプルさ・スケーラビリティと柔軟性の"両立"が
求められる
 - 2019年時点で期待される解の一つがSRv6
 - シンプルさ・スケーラビリティ:
世界規模で運用されるIPv6そのまま
 - 柔軟性:
途中経路を制御することでネットワークに任意の機能を付加可能

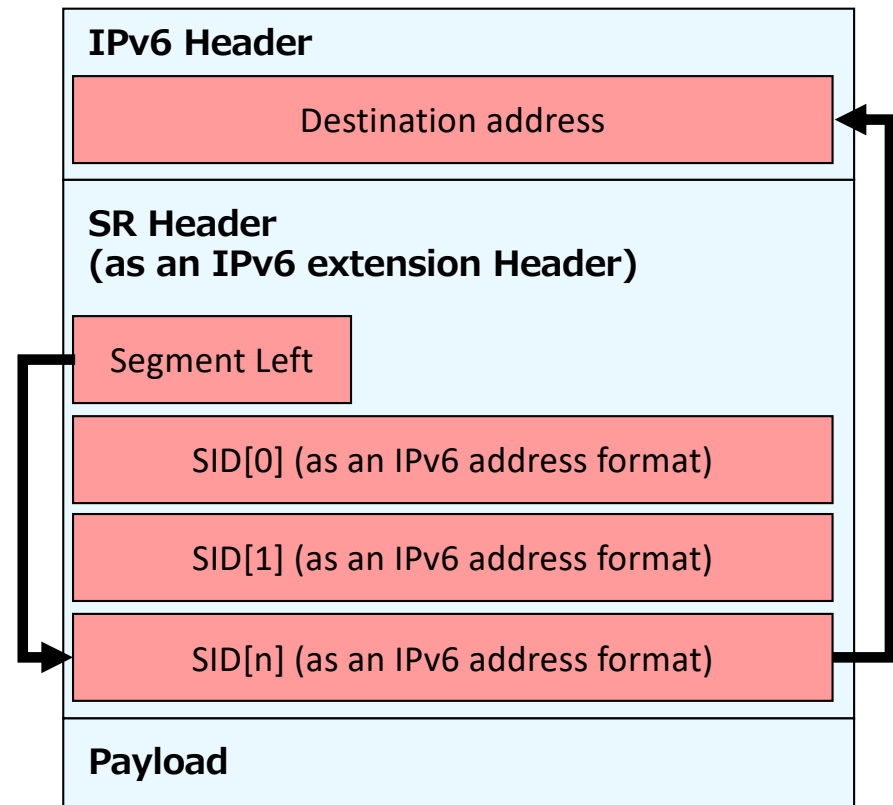
Segment Routing

- ネットワークをSegmentで表現し、ヘッダに埋め込まれたSegment Listによってパケットを転送する新しいルーティング技術



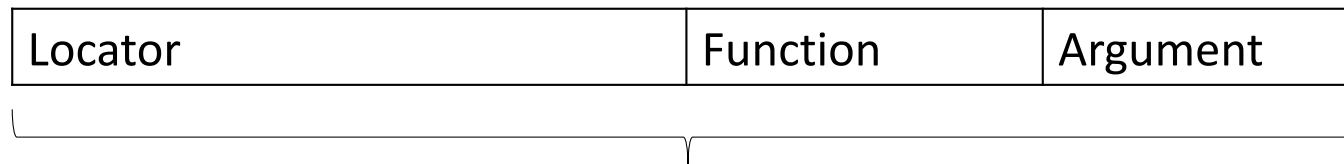
Segment Routing IPv6とは

- Segment RoutingのデータプレーンにIPv6を用いる方式
 - SRv6ノードはSegment Leftが指し示すSIDをDestination Addressに上書きする
 - SIDは次のSRv6ノードまたは最終的な宛先を示す



SRv6 SIDの構造

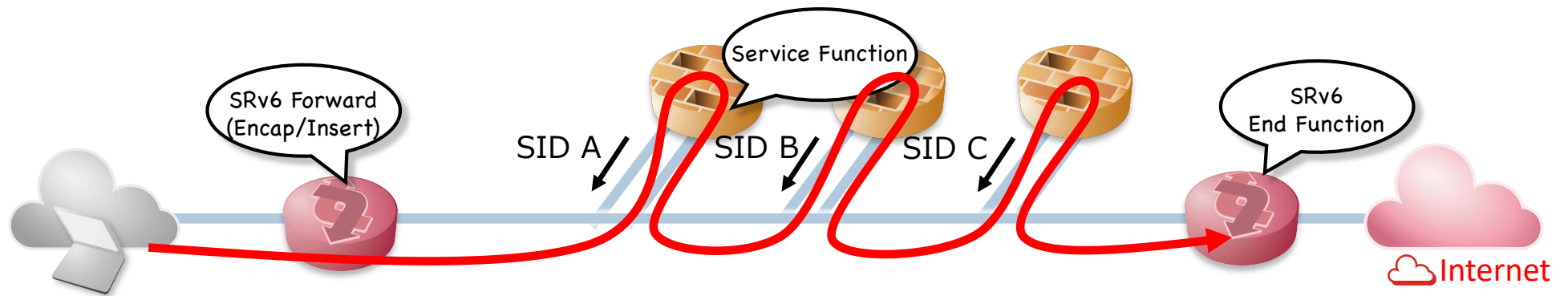
- SRv6 SID は128ビットのIPv6アドレス表記
 - Locator: SRv6のノードへルーティングするためのビット
 - Function: SRv6のノードにおいて取られるアクションを示すビット
 - Argument [optional]: アクションで使用される引数



IPv6 Address(128bit)

SRv6の理想的なアーキテクチャ

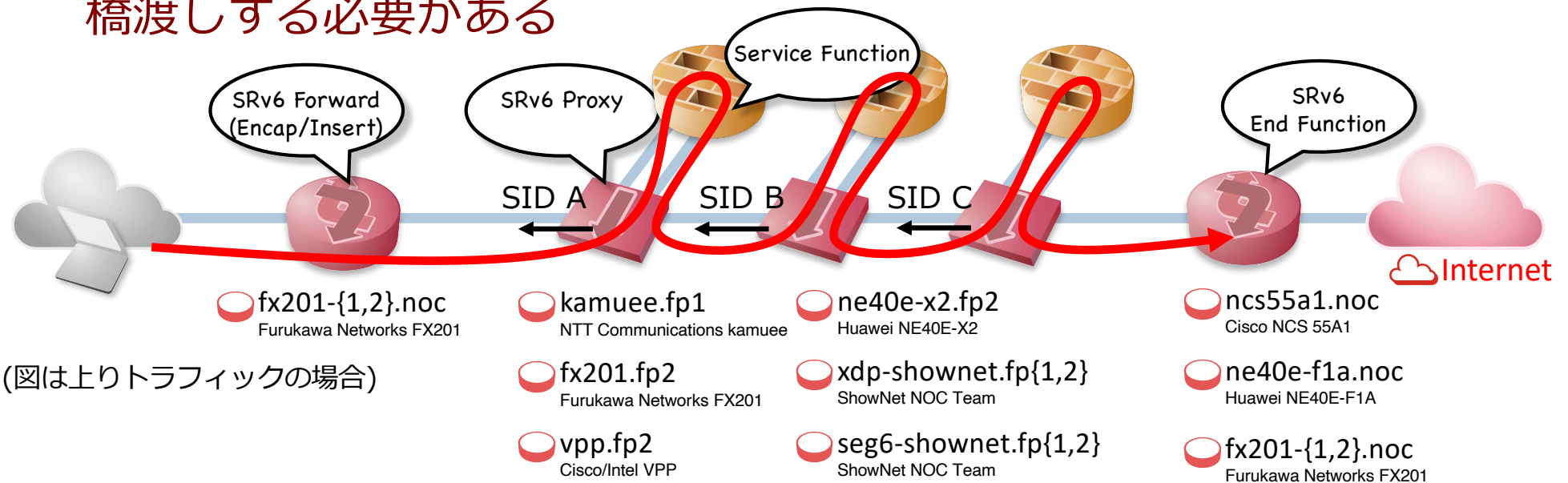
- 各Service FunctionがSIDを広告、SRv6ヘッダを解釈し、それに基づいてトラフィックを転送する



(図は上りトラフィックの場合)

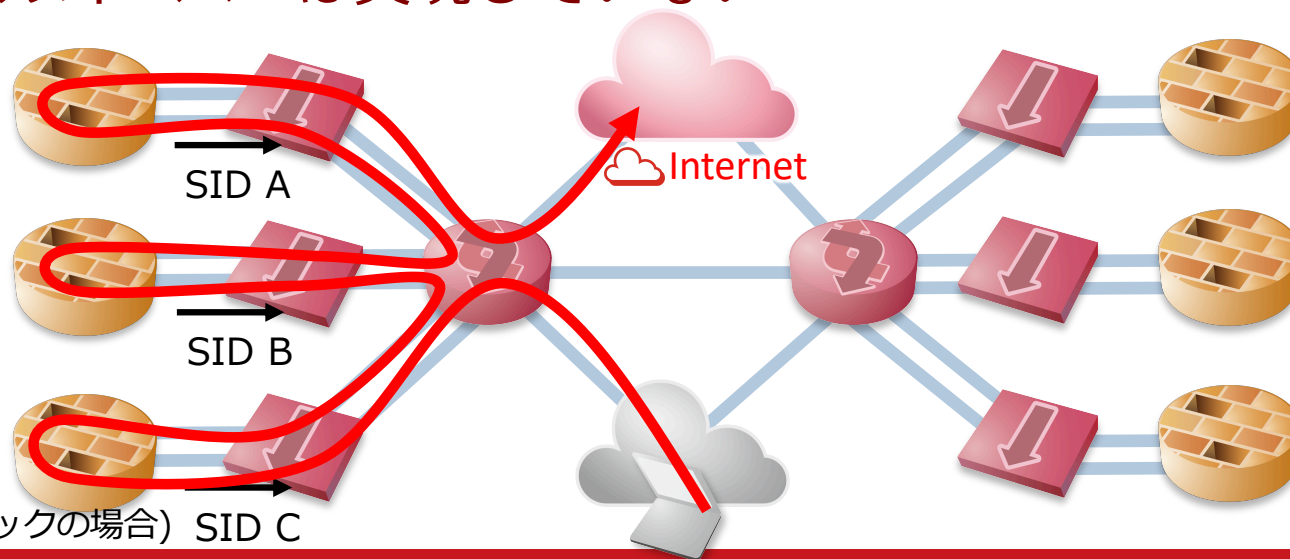
現時点で実現可能なアーキテクチャ

- SRv6ヘッダを解釈できるService Functionは限られているため、SRv6 ProxyによってSRv6の世界と普通のIPv4/IPv6の世界を橋渡しする必要がある



サービスチェーンの可用性

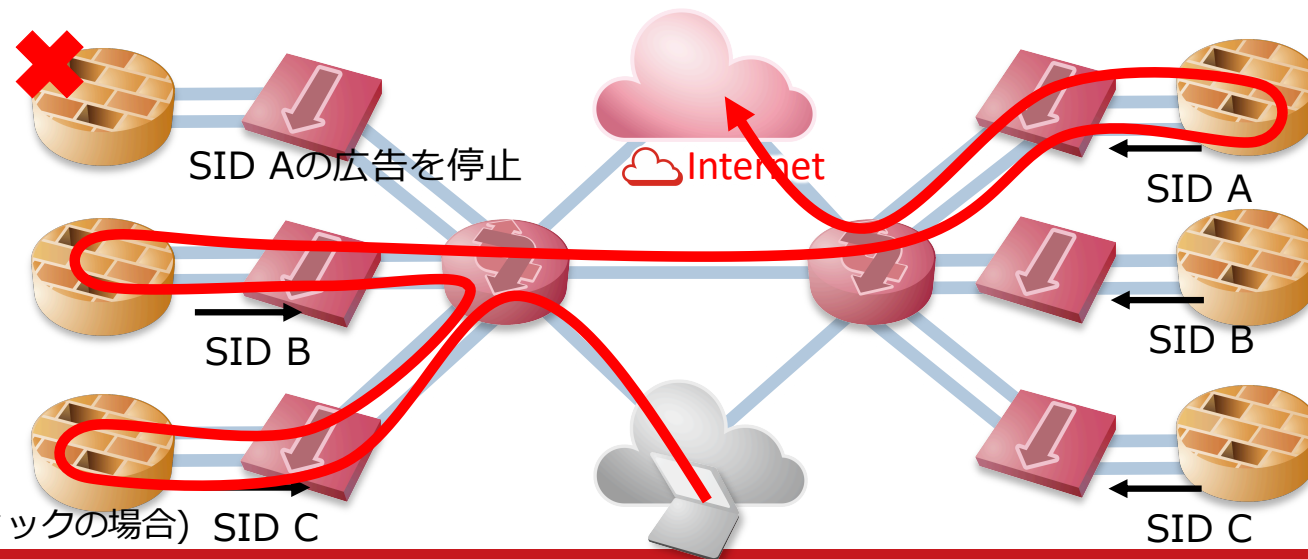
- チェインのどこかが故障した場合に備え何らかの対策が必要
- これまでのShowNetでは迂回を実現しているがフェイルオーバーは実現していない



(図は上りトラフィックの場合) SID C

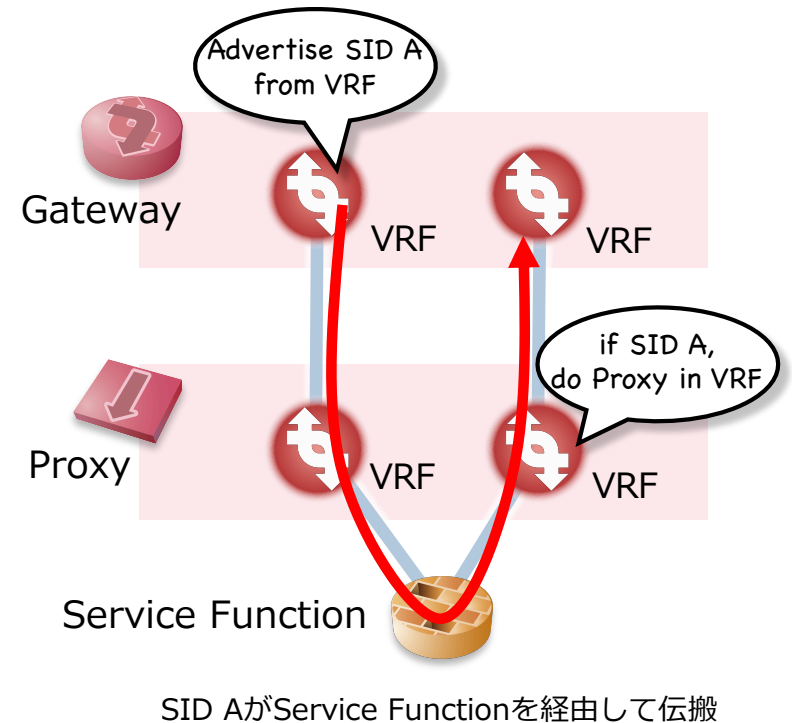
Service Functionの自動フェイルオーバー

- Functionの故障時に対応するSIDの広告を停止することで、anycastされているバックアップのService Functionに自動フェイルオーバーを実現

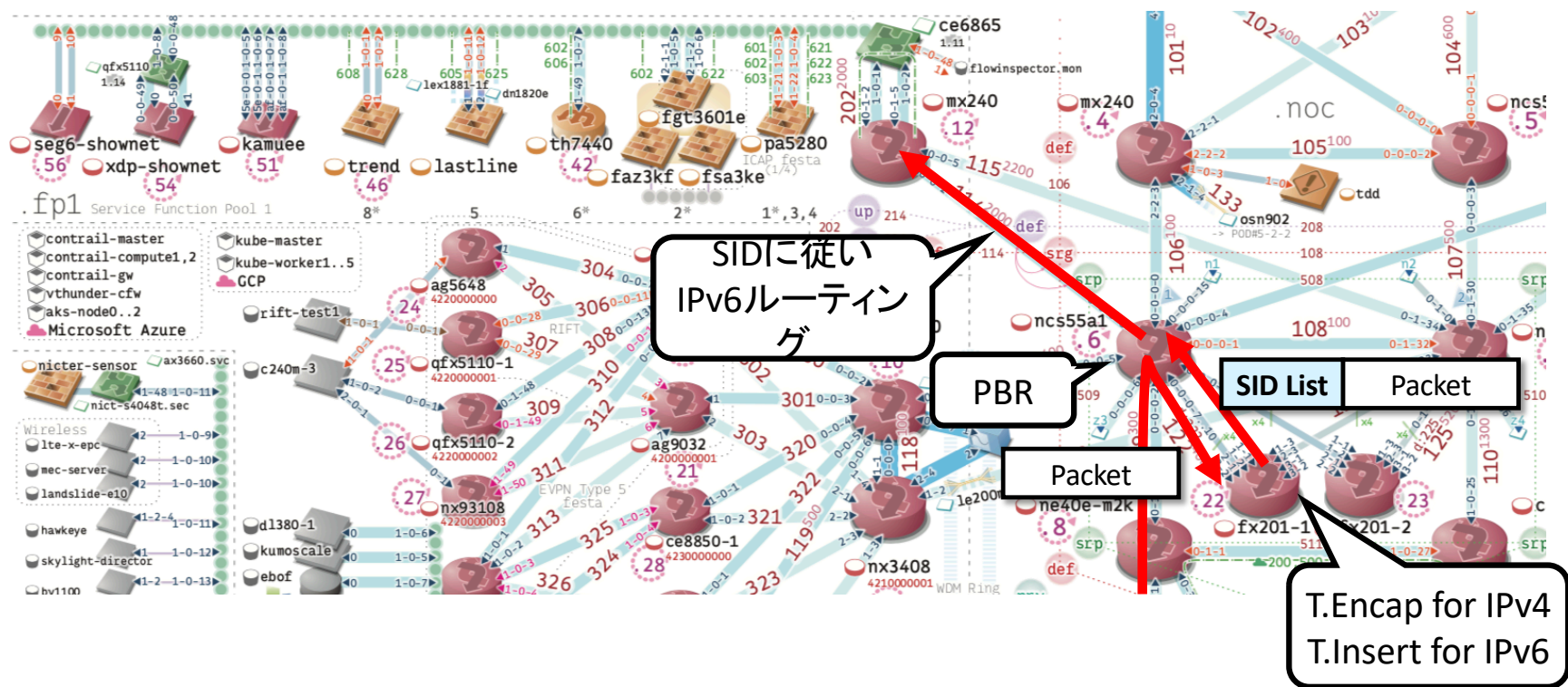


Service Function故障時の自動SID広告停止

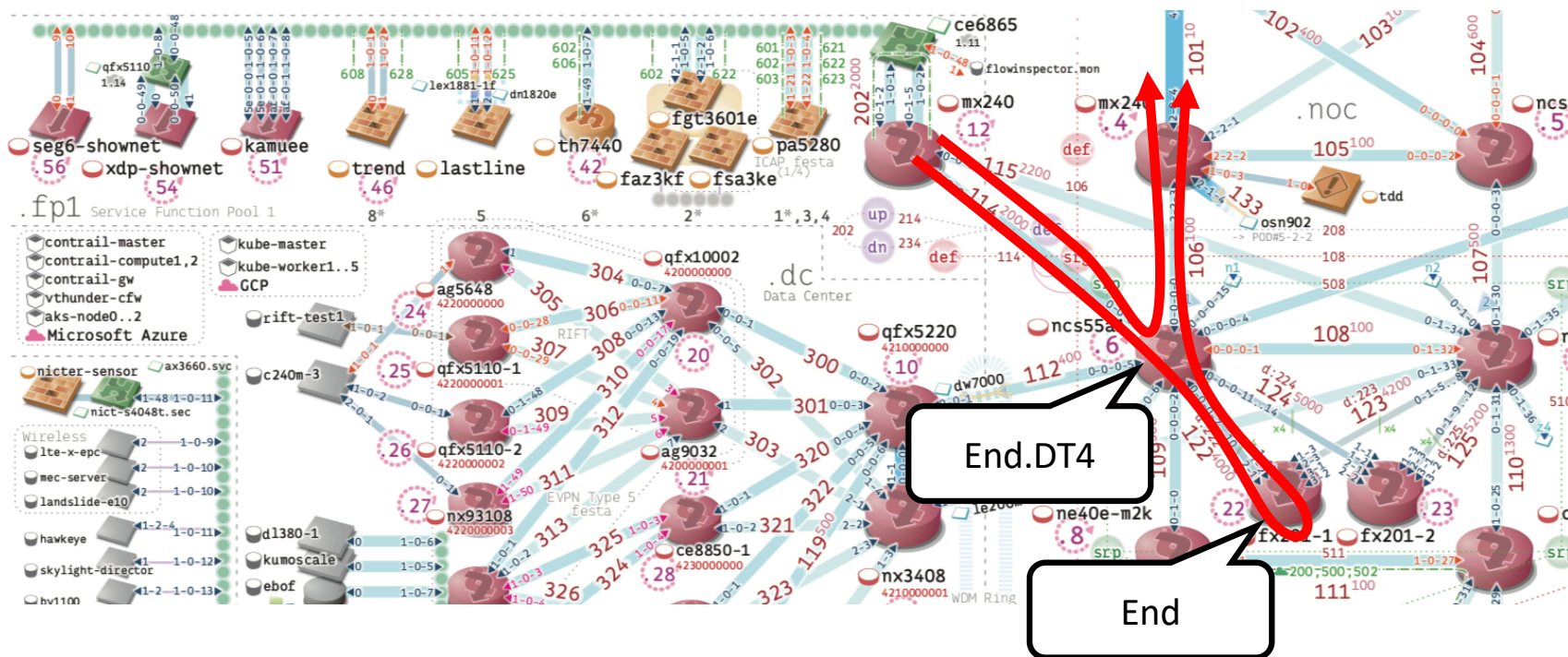
- 現時点では、OSPFv3 + VRF (Virtual Routing and Forwarding) によるワークアラウンド的な手法が限界
 1. VRFとService Functionからなるライントポロジを構成
 2. SID経路をGatewayのVRFの一方から広告
 3. Service Functionが止まるとSID経路の伝搬も止まる
 4. 副系のFunctionのSID経路が優先される (=フェイルオーバーする)



サービスチェーンの始点

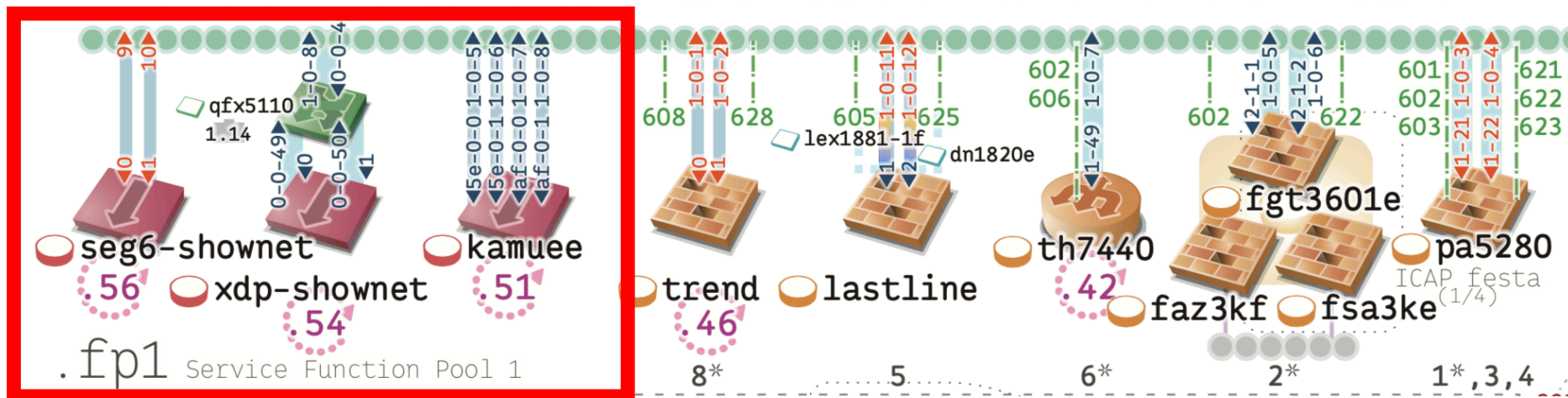


サービスチェーンの終点



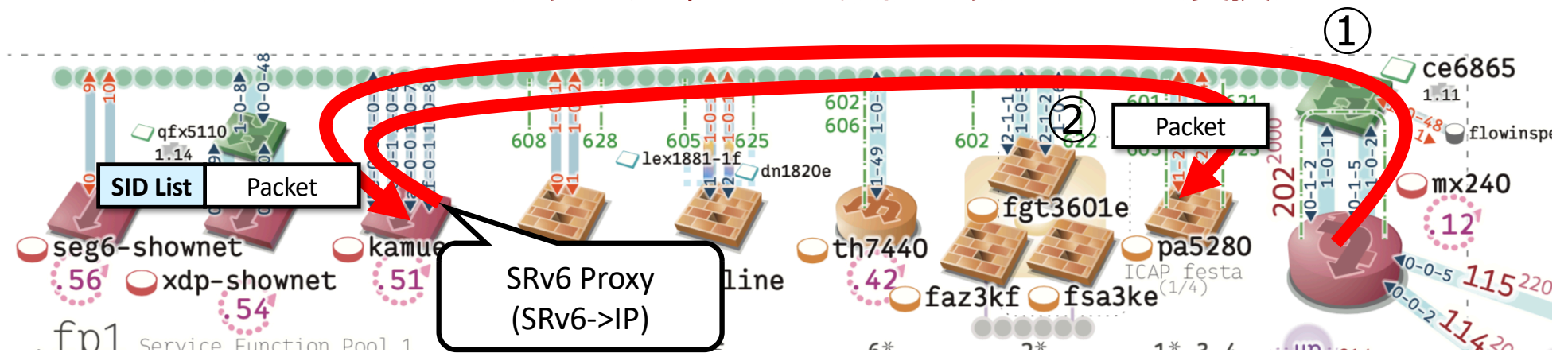
SRv6 Proxy

- SRv6ヘッダを持ったIPv6パケットを、SRv6を解釈しないService Functionで処理できる形に変換
- Service Functionから出る際にSRv6パケットに戻す



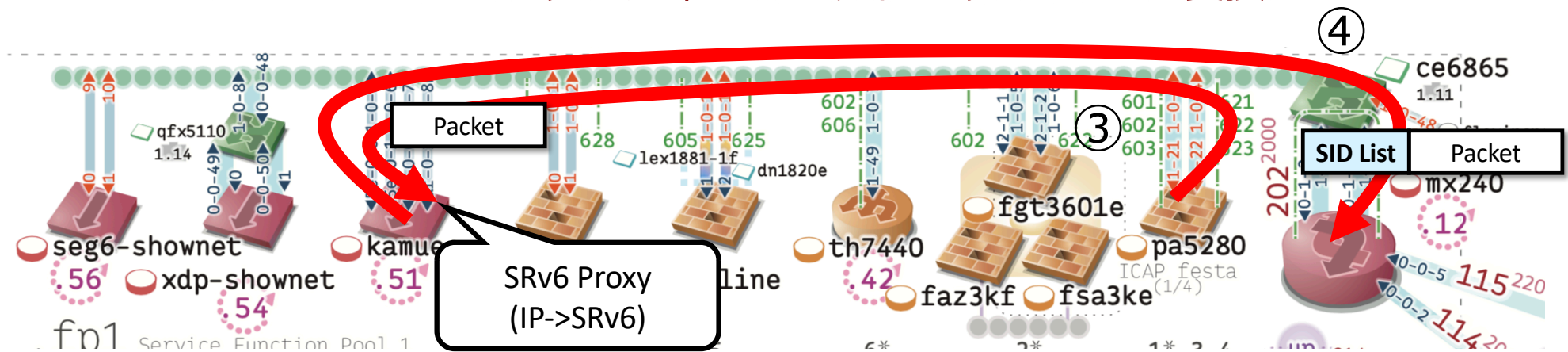
Function Poolの内部

- パケットがSRv6ヘッダに基づいてゲートウェイルータとService Functionを往復
 - ただし、Service Function前後でSRv6パケットをService Functionから見て通常のパケットに"見える"よう変換



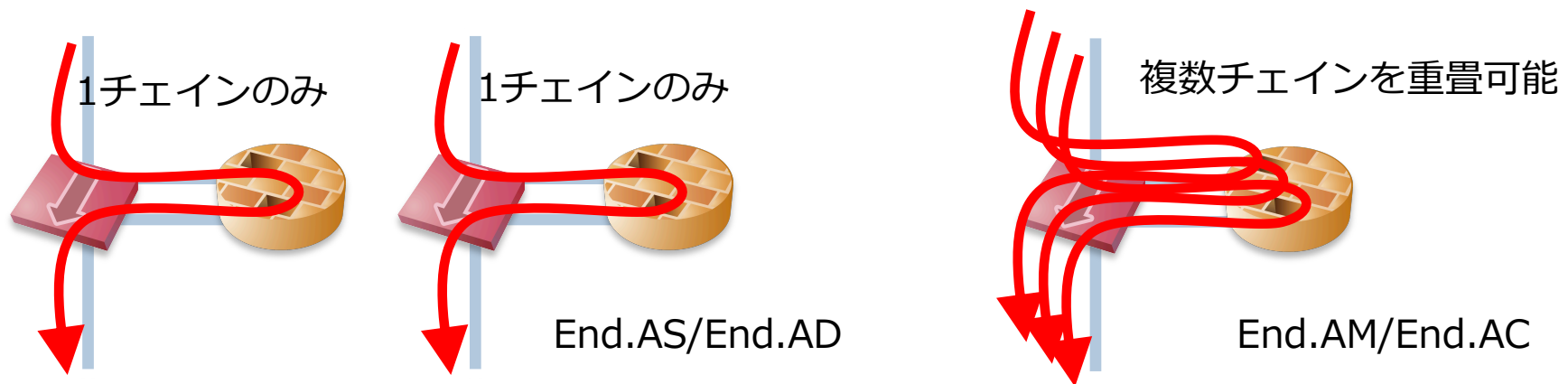
Function Poolの内部

- パケットがSRv6ヘッダに基づいてゲートウェイルータとService Functionを往復
 - ただし、Service Function前後でSRv6パケットをService Functionから見て通常のパケットに"見える"よう変換



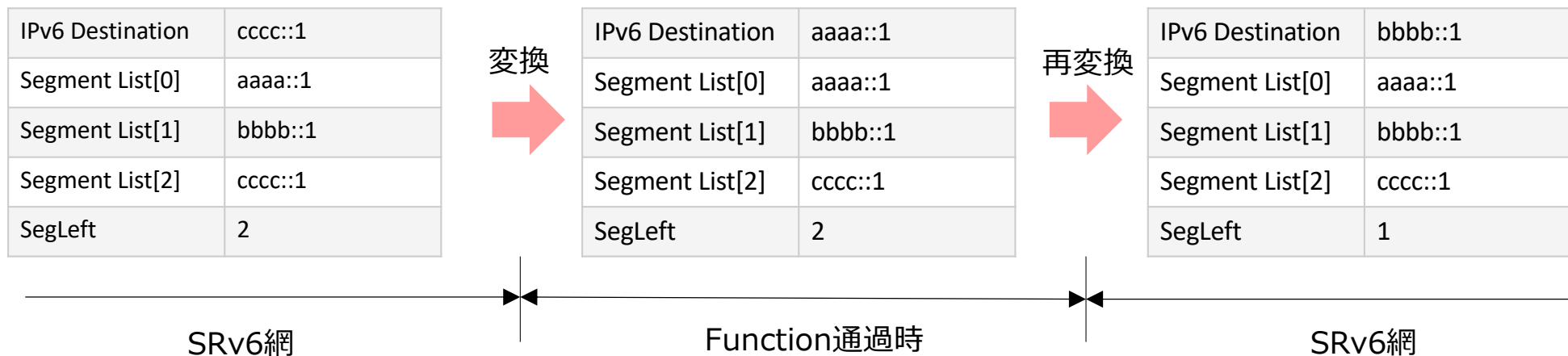
SRv6 Proxyの方式

- いくつか種類があるが、機能的な観点で見ると以下2種類に大別
- ShowNetのような大規模環境では、チェーンそれぞれにSRv6 ProxyおよびFunctionを用意することが現実的に不可能
 - ただしEnd.AMはIPv6トラフィックのみに適用可能



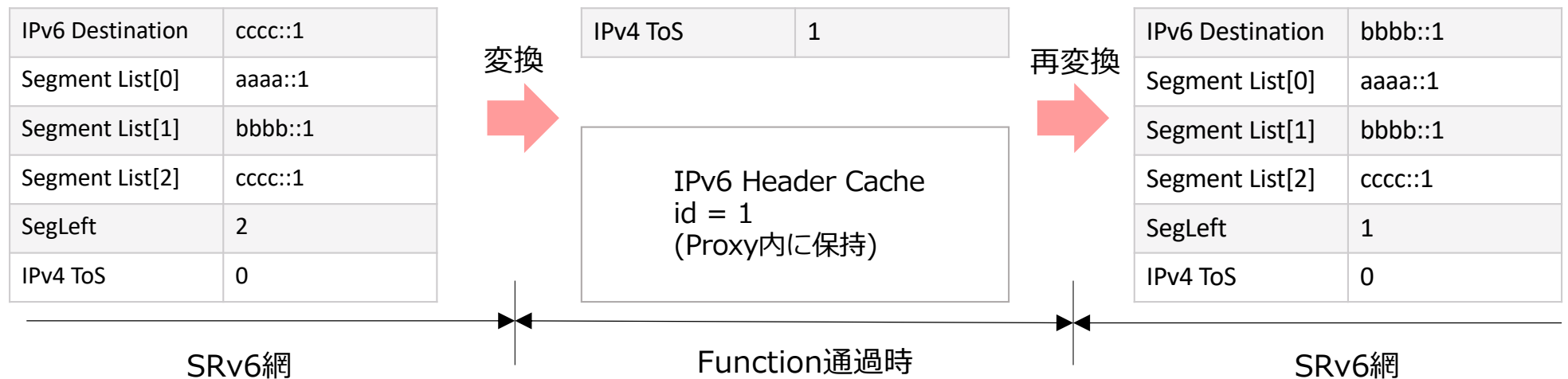
IPv6 over SRv6のProxy方式: End.AM

- Function通過前にSIDの最後をDestination Addressに上書き
- Function通過後、次のSIDをDestination Addressに上書き



IPv4 over SRv6のProxy方式: End.AC

- Function通過前にIPv6ヘッダをキャッシュし、IPv4のToSフィールドにSRv6 Argument(SIDの内ファンクションに渡す値として使える領域)を埋め込む
- Function通過後、ToS値に基づいてIPv6ヘッダをキャッシュから復元



End.ACにまつわるあれこれ

- End.ACは標準化過程にある技術ではなく、ShowNet 2019での要件を満たすためNOCチームが考案したもの
- 現時点では1つのSRv6 Proxyで複数のIPv4チェーンを収容可能な標準技術がない
 - 大規模環境では必要性が高いため、標準化提案等のフィードバックを検討中
- ShowNet 2019ではAF_XDPベース/Linuxカーネルベースの2実装を作成し、AF_XDPベースの実装を実運用
 - <https://github.com/edenden/end.ac>



本日の発表内容

1. ShowNet 2019における
SRv6 Service Chainingのアーキテクチャ
2. **SRv6最新動向と標準化の場へのフィードバック**
3. ShowNet 2020 STMプログラム

標準化活動へのフィードバック

- IETF
(Internet Engineering Task Force)
 - IPをはじめインターネット関連技術を標準化する団体
 - 会議を年3回程度開催
- IETF 105(2019/7)にてSRv6標準化関係者にフィードバックを実施



出典: <https://www.ietf.org/blog/ietf-105-highlights/>

SRv6関連仕様の標準化状況

- 主にSPRING WGにて仕様策定が進行中
 - 主要部分はRFC、WG Draftなどある程度固まっている状況
- SR（MPLS含む）の基本的な仕様
 - RFC8402
- SRv6のヘッダフォーマット
 - draft-ietf-6man-segment-routing-header
- Service Programming(SR Proxyなど)
 - draft-ietf-spring-sr-service-programming
 - IETF 105後にWG draft化

SRv6の最新動向

- SRv6 uSIDとC-SRH
 - とともにSRv6のヘッダサイズを圧縮する目的で提案
 - 既存のSRv6にそのまま適用可能
- SRm6(SRv6+ until 2019/10)
 - IPv6仕様の厳密な遵守とASICへの最適化を目的に提案
 - SRv6の構成要素を新たに定義し直し、SIDもIPv6形式ではなく16-32bitの別形式を採用
(結果的にヘッダサイズが小さくなる)
 - 既存のSRv6と互換性がない

ShowNet発の標準化提案

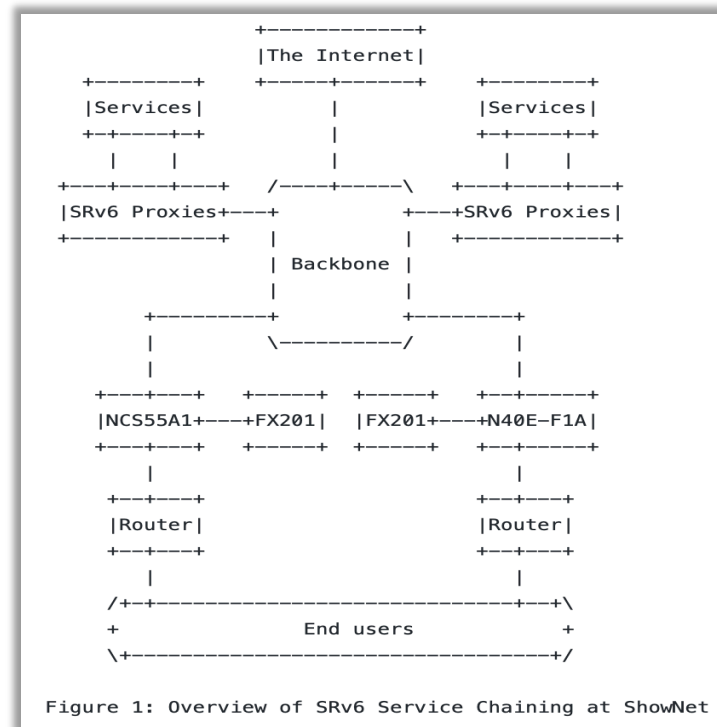
- 下記2本のInternet-Draftを提案
 - ShowNet 2019のSRv6によるサービスチェイニング
 - SRv6 Proxyの新方式 (End.AT)
- IETF 106にて提案活動を実施予定

```
SPRING                                Y. Ueno, Ed.  
Internet-Draft                        NTT Communications Corporation  
Intended status: Standards Track      R. Nakamura  
Expires: April 29, 2020                The University of Tokyo  
                                        T. Kamata  
                                        Cisco Systems, Inc.  
                                        October 27, 2019  
  
SRv6 Tagging proxy  
draft-eden-srv6-tagging-proxy-00  
  
Abstract
```

```
SPRING                                R. Nakamura, Ed.  
Internet-Draft                        The University of Tokyo  
Intended status: Informational        Y. Ueno  
Expires: May 3, 2020                 NTT Communications Corporation  
                                        T. Kamata  
                                        Cisco Systems, Inc.  
                                        October 31, 2019  
  
An Experiment of SRv6 Service Chaining at Interop Tokyo 2019 ShowNet  
draft-upa-srv6-service-chaining-exp-00  
  
Abstract
```

draft-upa-srv6-service-chaining-exp

- ShowNet 2019のSRv6によるサービスチェイニング
 - Informational draft
 - 実運用から得られた知見をまとめたもの
 - 今後のSRv6標準化過程で考慮すべき点などを記述・主張
 - ついにShowNetのトポロジがあのアスキーアートに！



draft-eden-srv6-tagging-proxy

- SRv6 Proxyの新方式 (End.AT)
 - Standards-track draft
 - IETF有識者との検討の結果、End.AT(SRv6 Tagging Proxy)に名称を変更
 - End.ATの仕様が公開されることで、ShowNetのサービスチェイニングのような大規模・多数チェインの構成を誰もが実現できるように

```
1. IF NH=SRH & SL > 0 & ENH == 4 THEN
2.     Cache IPv6 Header and SRH into CACHE[ARG]
3.     Remove the (outer) IPv6 header and its extension headers
4.     Embed ARG into the ToS field of the (inner) IPv4 header
5.     Forward the exposed packet on IFACE-OUT towards NH-ADDR
6. ELSE
7.     Drop the packet
```

SRv6を用いたサービスチェイニング @ShowNet 2019

- SRv6によってIPv6のシンプルさ・スケーラビリティを享受しつつ、
順番・場所自由なチェーンを実現
- 本番3日間の故障はゼロ
- 構築を通して得られた知見を標準化の場へフィードバック



本日の発表内容

1. ShowNet 2019における
SRv6 Service Chainingのアーキテクチャ
2. SRv6最新動向と標準化の場へのフィードバック
3. **ShowNet 2020 STMプログラム**

ShowNetの構成要素

3つの要素が、協力、調和することで、ShowNetは構築、運営されている

産学界から様々な人材が参加

東京大学、情報通信研究機構、レピダム、
NTTコミュニケーションズ、慶應義塾大学、
ジュニパーネットワークス、シスコシステムズ、
and more

NOC
Team
Member

SHOWNET
EVOLVE INTO
THE NEXT GENERATION



Contributor

STM

一般公募によるボランティアスタッフ

大学，専門学校，高専，企業(ISP, Sier...)など多種多様

ShowNet機材提供企業

各種ネットワーク機器，負荷試験装置，
ソフトウェア，そして人材の提供

STMプログラムとは？

- ShowNetへの参加を通して
大規模かつ最先端のネットワークの構築に触れる
ボランティアプログラム
- NOCチームやコントリビュータと協力して、
ShowNetを実際に構築、運用する作業に携わる

STMプログラムで体験できること

- 大規模ネットワークを"0"から作り"0"に戻す経験
 - ネットワーク構築に必要な作業を短期間で一通り行う
 - ただし、分担作業のため各人が全ての作業を経験できるわけではない
- 組織に囚われない人と人との繋がり
 - 大規模ネットワークの構築を通じて、組織を越えたヨコの繋がりを作ることができる

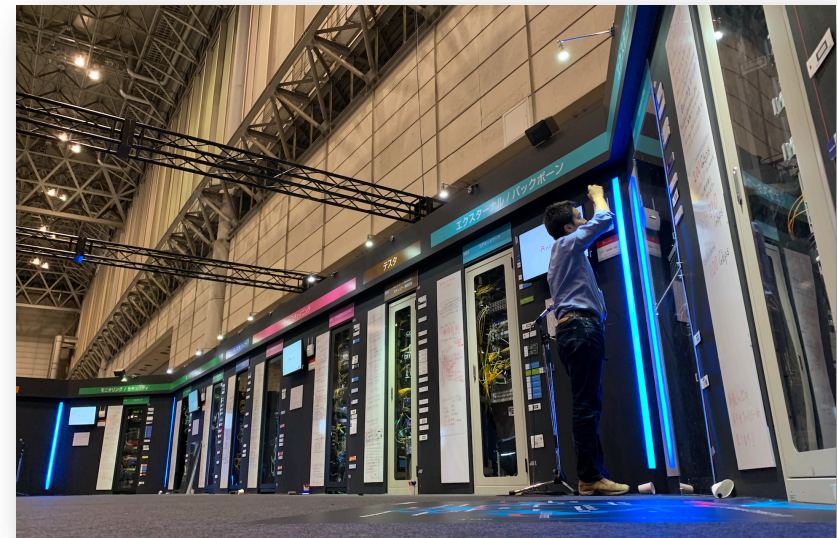
2020年度の募集・選考

- 応募情報はInterop Tokyoのウェブサイト (<https://www.interop.jp/>) にて公開中
- STMの採用は、書類による一次選考と NOCチームメンバーによる面接を経て決定される

- 10月17日 STM募集ページオープン
- 11月9日 / 11月30日 説明会実施（両日とも東京、テレコン有）
- 12月6日 事前アンケート締め切り
- 1月18日 / 1月25日 東京（テレコン有）で面接
- 1月末以降 結果通知

Interop Tokyo 2019 ShowNet

- 未来のネットワークの1つのカタチ
 - 10年先のインターネットをつくる
 - そのモデルを示すデモと検証
 - 相互接続性
 - ShowNetは異種ベンダー、異種機器間の相互接続で成り立つ
 - オープンな技術の上に成り立つ組み合わせと設計の自由度
 - 検証とフィードバック



SHOWNET

EVOLVE INTO
THE NEXT GENERATION

