次世代のIP経路制御で作る サービスチェイニング@ShowNet 2017

Interop 2017 ShowNet NOCメンバー
/NTTコミュニケーションズ株式会社
上野 幸杜
ueno@interop-tokyo.net

SHOWNET
INHERIT THE INTENTION

アジェンダ

1. ShowNetとは
2. ShowNet 2017における
   サービスチェイニングのコンセプト
3. 設計時の考慮ポイント
4. 設計と実装
5. まとめ

SHOWNET
INHERIT THE INTENTION

ShowNetとは

ShowNetの概要

- Interopの原点
- 産業界、学会、研究機関から集まるトップエンジニアによる世界最大級のライブデモンストレーションプロジェクト
- 2年後、３年後の業界に向けてのメッセージを発信
- 世界、国内で初披露（実稼働）される新製品も実装
- 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

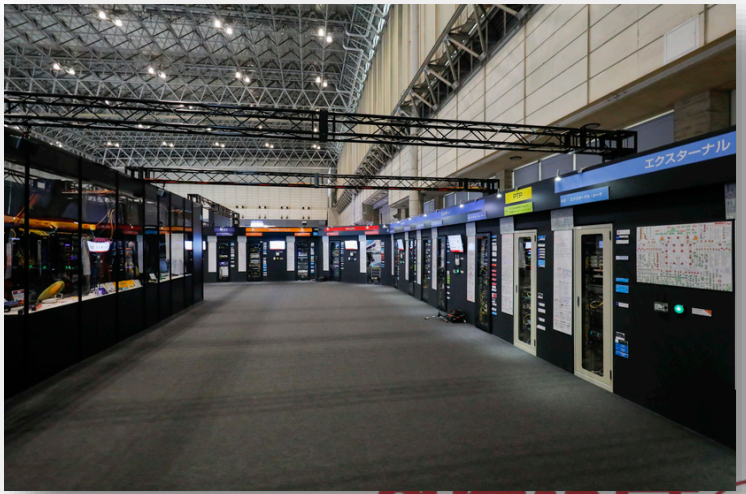
- I know it works because I saw it at Interop -

市場と技術の最前線、未来が見えるネットワーク

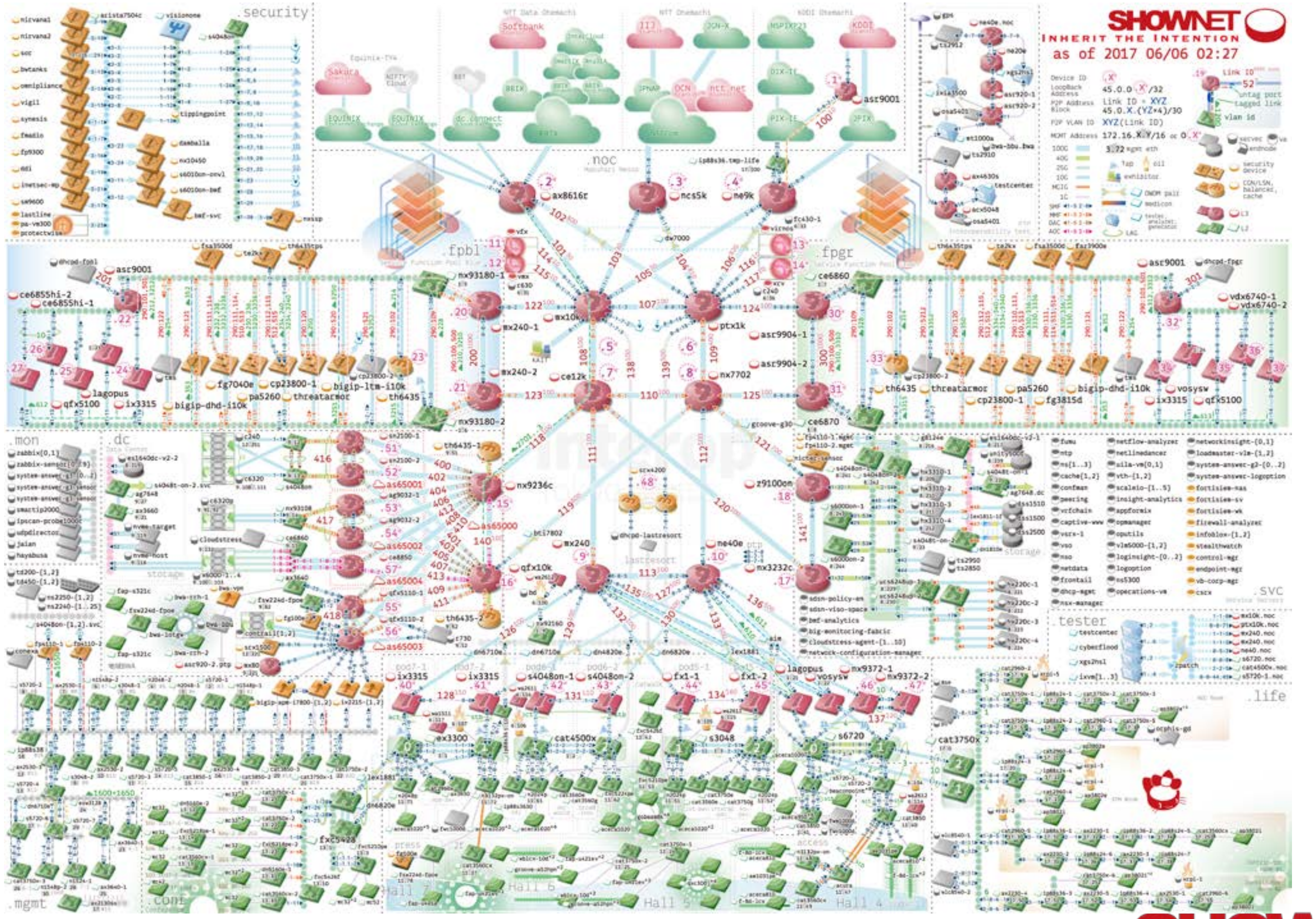
SHOWNET
INHERIT THE INTENTION

ShowNetの目的

- **未来のネットワークのデモンストレーション**
  - 相互接続性検証 = Interopの理念
  - 使っただけ/動かしただけでなく、**どう使うか**を見せる
- **実サービスネットワーク**
  - 出展社、来場者への接続性提供
    - 対外接続性、バックボーン、ディストリビューション、Wifi、CGN、DHCP、DNSなどなど



SHOWNET
INHERIT THE INTENTION

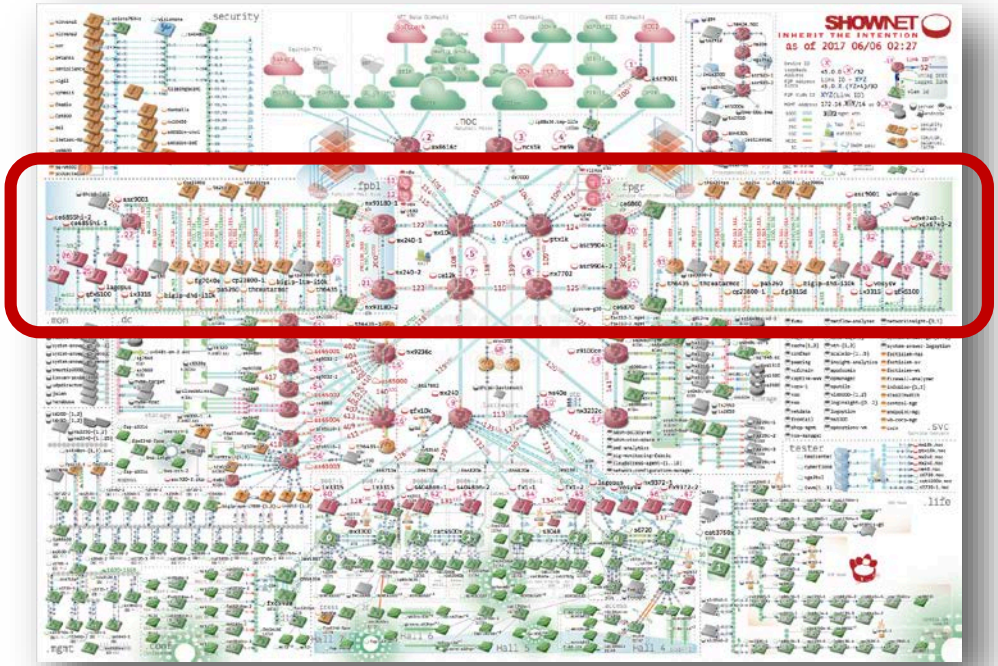


SHOWNET
INHERIT THE INTENTION

ShowNet 2017における
サービスチェイニングの考慮ポイント

サービスチェイニング@ShowNet 2017

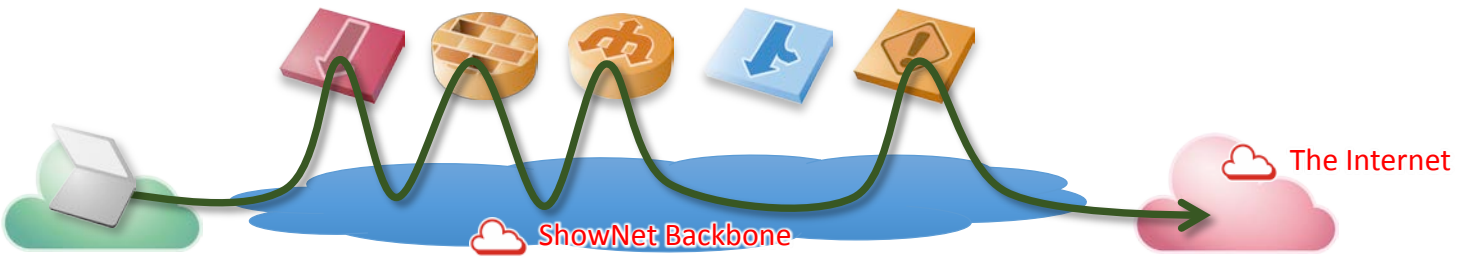
- ShowNet 2017の目玉の一つは"サービスチェイニング"でした
- トポロジ図で俯瞰しても、前年までと大きな違いがあります



SHOWNET
INHERIT THE INTENTION

サービスチェイニング

- ネットワークの機能を鎖のように連携
  - 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
- いままで不可能だったサービス運用を可能に！



2016年までのチャレンジ、そして2017年

- SDNへの挑戦からIP経路制御の次世代への挑戦
  - 昨年までは、SDN/NFV技術を用いた柔軟な制御を志向
  - 2017年は進化してきたIP経路制御の限界にチャレンジ

Interop Tokyo
7 – 9 JUNE 2017 | MAKUHARI MESSE | JAPAN
SHOWNET
INHERIT THE INTENTION

SHOWNET
INHERIT THE INTENTION

2017年のコンセプト

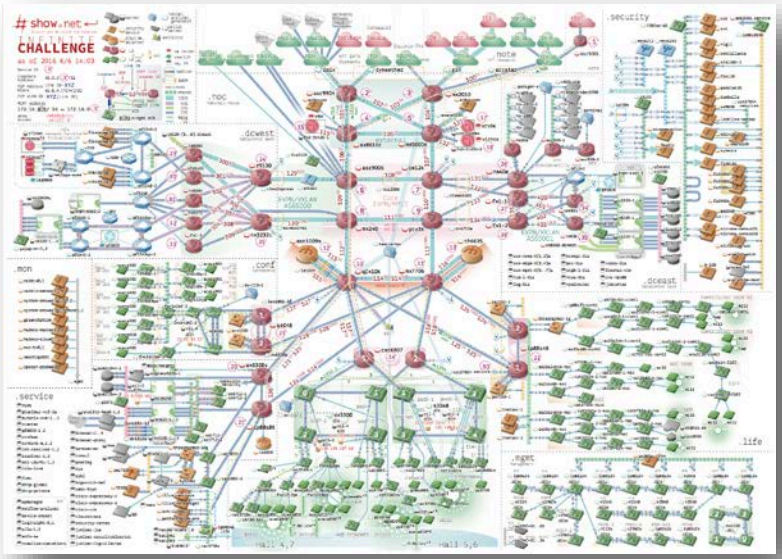
1. 多種多様なサービスを収容可能な汎用性
   - ShowNetにコントリビューションがあった装置のうち、
     ユーザトラフィックを収容するものは基本的に全て収容する
2. 大規模イベントでの本番使用に耐える可用性
   - 特定の出展ブースだけでなく、
     来場者も含めた全員を収容する
3. 既存のIP技術の活用による実現性
   - ベンダを問わず動作し、
     ShowNetでなくとも実現可能なもの

ShowNet 2017における
サービスチェイニングの考慮ポイント

論点1: 何をどこまで収容するか

- そもそも要件は、
  - 「ShowNetにコントリビューションがあった装置のうち、ユーザトラフィックを収容するものは基本的に全て収容する」
- L3で動作するものもあり、「サービスチェイニングのための」最新プロトコルは実装されていないものが多い！

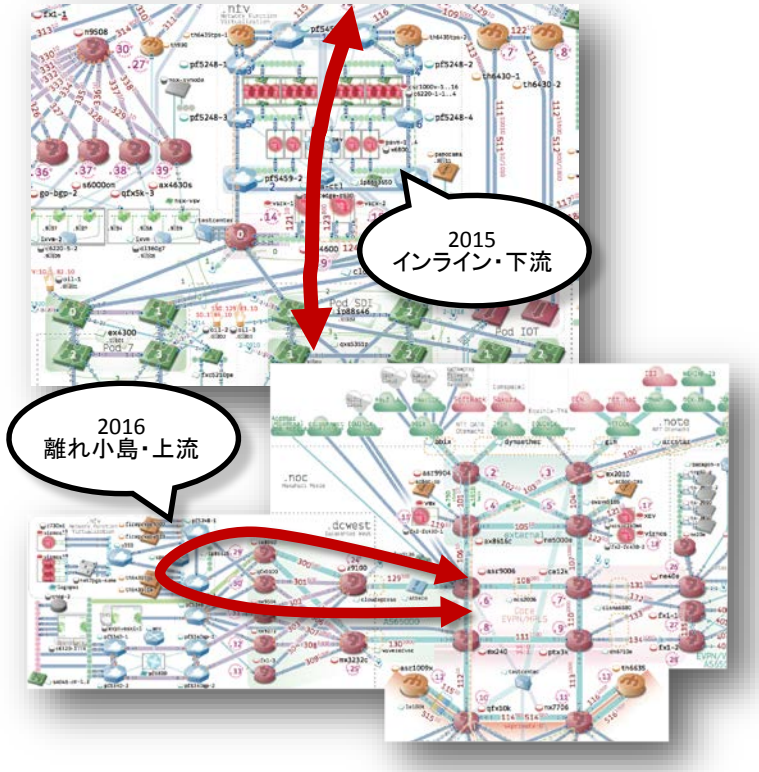
ShowNet 2016より:



SHOWNET
INHERIT THE INTENTION

論点2: バックボーンへの組み込み方

- サービスチェイニングを
  どこで実現するのか？
  - バックボーンにインラインで挟むか、離れ小島までトラフィックを誘導するのか
  - 上流か下流か
- どうやってユーザトラフィックを離れ小島まで誘導するか？



SHOWNET
INHERIT THE INTENTION

論点3: チェイニング手法

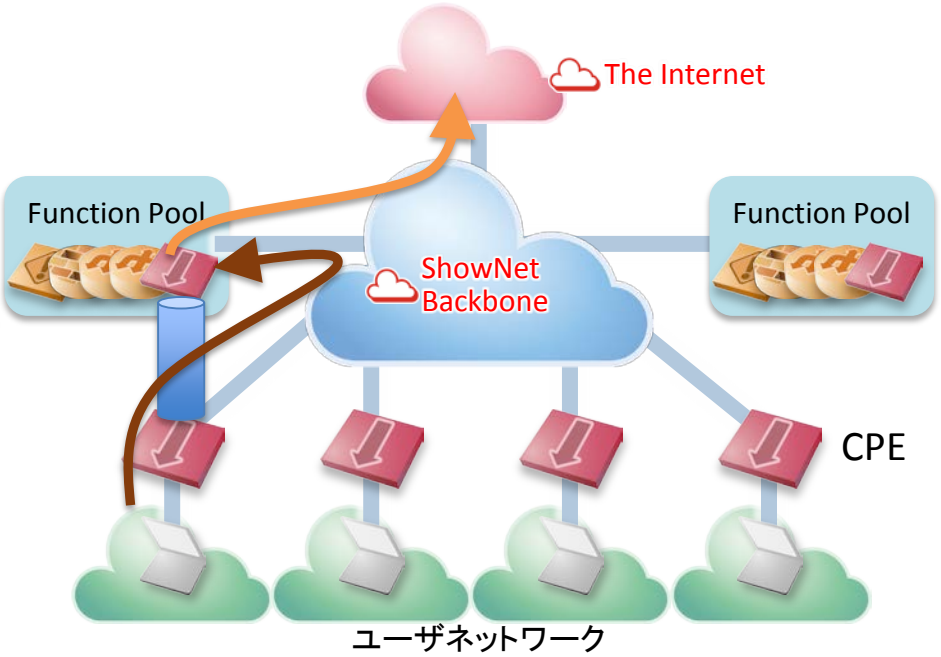
- Service Chaining網の中で、
  どうやってファンクションを連鎖させるか？

	ShowNetに必要なFunctionを全て収容できるか？	ShowNetの本番使用に耐えられるか？	ShowNetの機器で実現できるか？
NSH（トンネル系技術）	○	△（未知数）	△（実装ステータス）
MPLS/Segment Routing	○	○	△（実装ステータス）
Openflow	△（L3の収容は難しい）	△（属人性が高くなりがち）	△（実装機種が限られる）
VRF + PBR	○	△（設定量が非現実的）	○
VRF + BGP flowspec	○	△（未知数）	○

ShowNet 2017 での決め手

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
- 出展社に提供する機能は両系の**ファンクションプール**に集約
- **L2オーバーレイ技術**で出展社収容からファンクションプールまでL2延伸
- ファンクションプール内では**VRF**と**BGP Flowspec**を用いてサービスチェイニングを実現

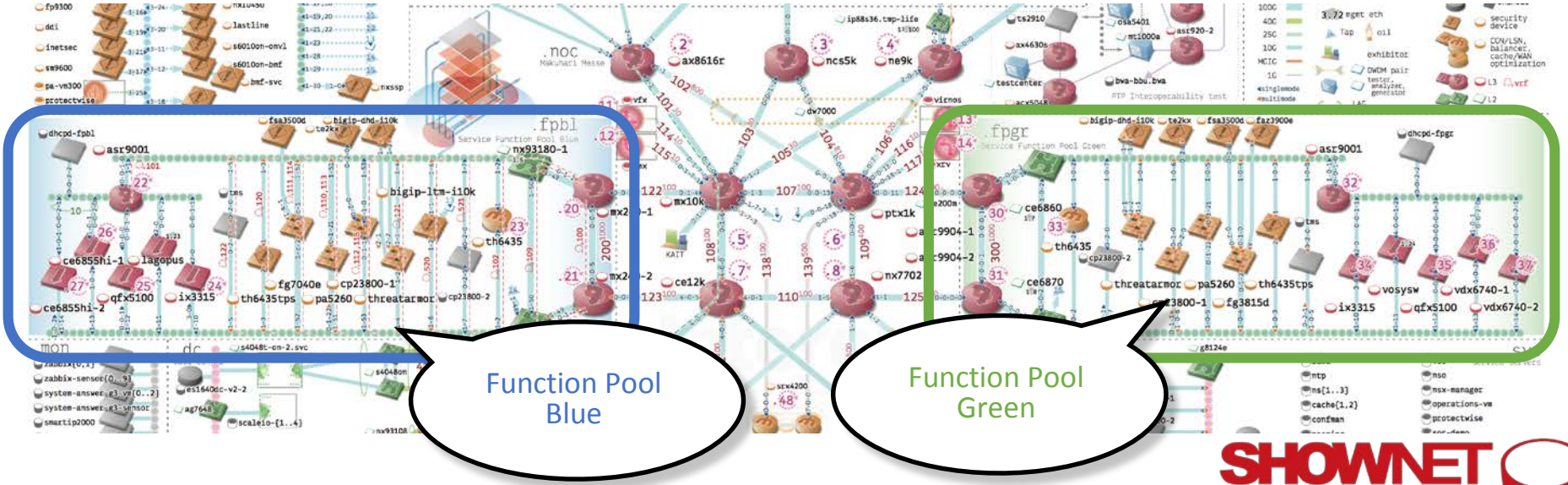


SHOWNET
INHERIT THE INTENTION

ShowNet 2017における
サービスチェイニングの設計と実装

ファンクションプールとは

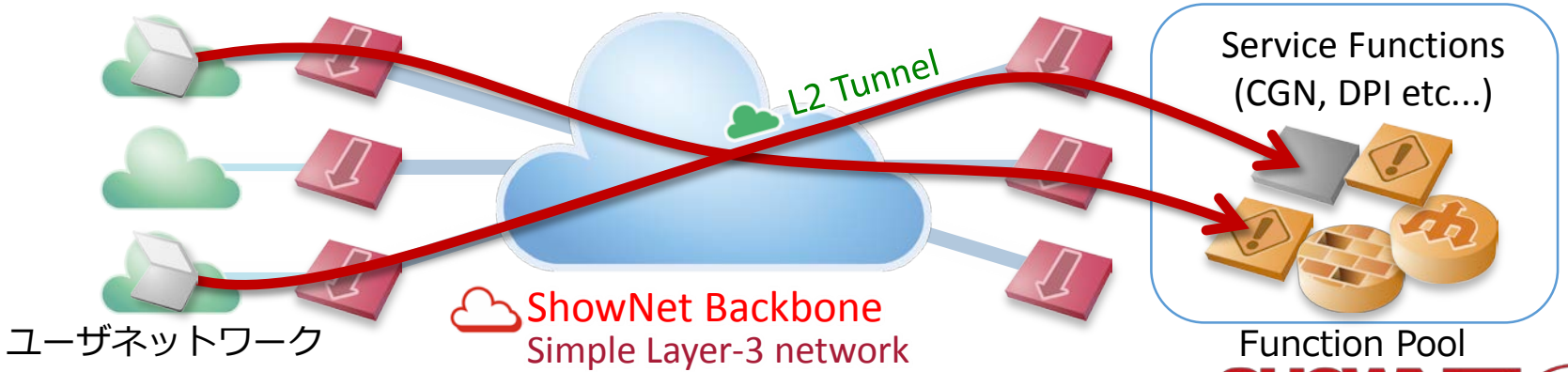
- ネットワーク機能を集約した"貯水槽"
  - ShowNetでは冗長のため２つのファンクションプールを構築
    左右のファンクションプールをそれぞれBlue/Greenとし切り替え



SHOWNET
INHERIT THE INTENTION

トラフィック誘導

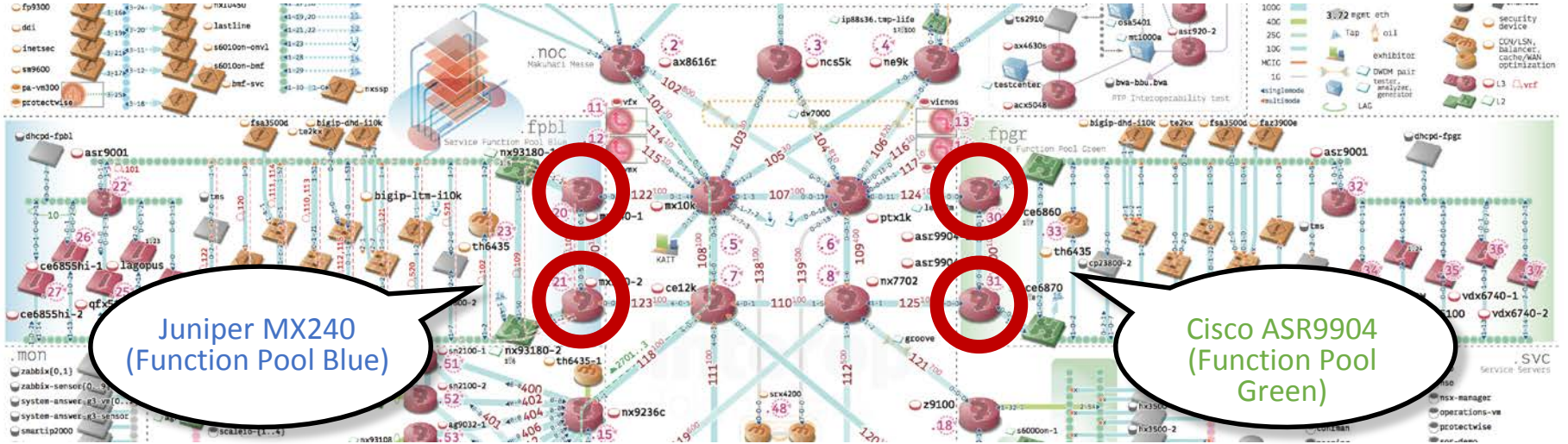
- 異なる複数のオーバレイ技術を利用
  - ShowNet 2017ではL2オーバーレイ技術をトラフィック誘導に使用
  - それぞれ別の技術を用いた4つのL2オーバーレイ面を構築
  - 全てのユーザトラフィックは一旦トンネルされ、
    ファンクションプールまでは通常のL3転送で到達する



SHOWNET
INHERIT THE INTENTION

サービスチェイニング

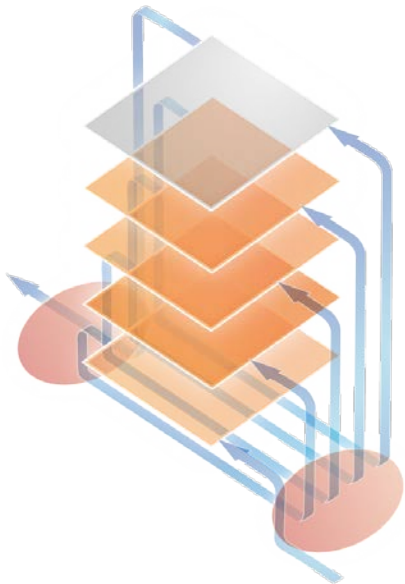
- チェイニングに使用するルータは2機種を2台ずつ使用
  - ファンクションプールBlue: Juniper MX240
  - ファンクションプールGreen: Cisco ASR9904



SHOWNET
INHERIT THE INTENTION

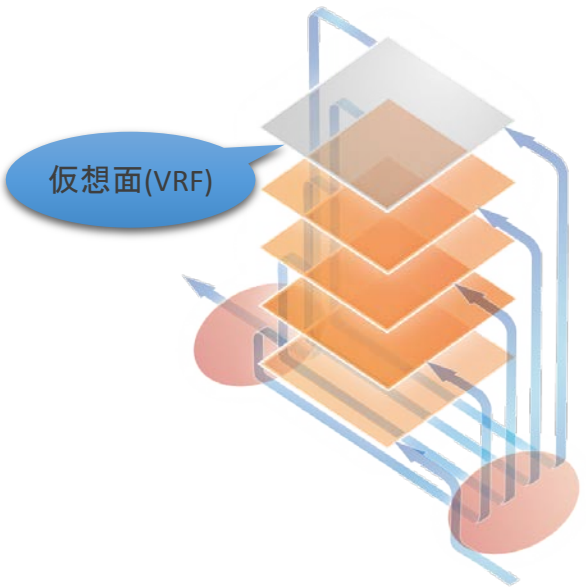
サービスチェイニング

- 設計上の要件
  - 対称性 -> 双方向で機能群を正しい順序で経由させなければならない（e.g. ステートフルファイアウォール、CGN）
  - 冗長性 -> 特定のFunctionに障害が起きた際、サービスを継続できる必要
  - 柔軟性 -> Functionの順番を設定変更なしに任意に入れ替えられる必要

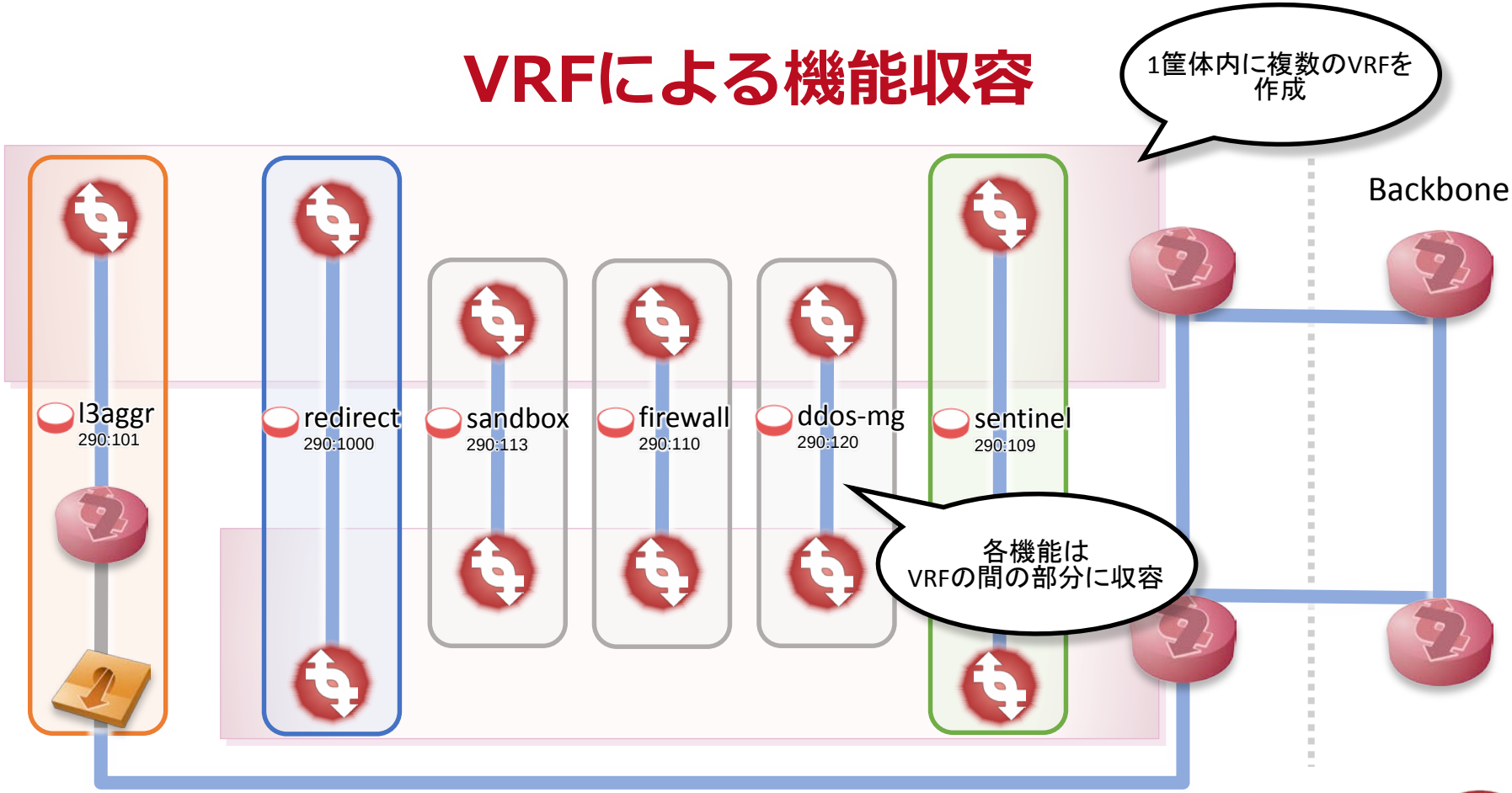


VRFによる機能収容

- 各ファンクションを独立したルーティング面として構築
  - VRF: Virtual Routing and Forwarding ルータの筐体内で経路表を分離する技術
  - L3のFunctionを収容できる
- 各ルーティング面の中ではOSPFによる経路交換を行う
  - 全てのルーティング面で静的経路を書くのが労力的に辛いため

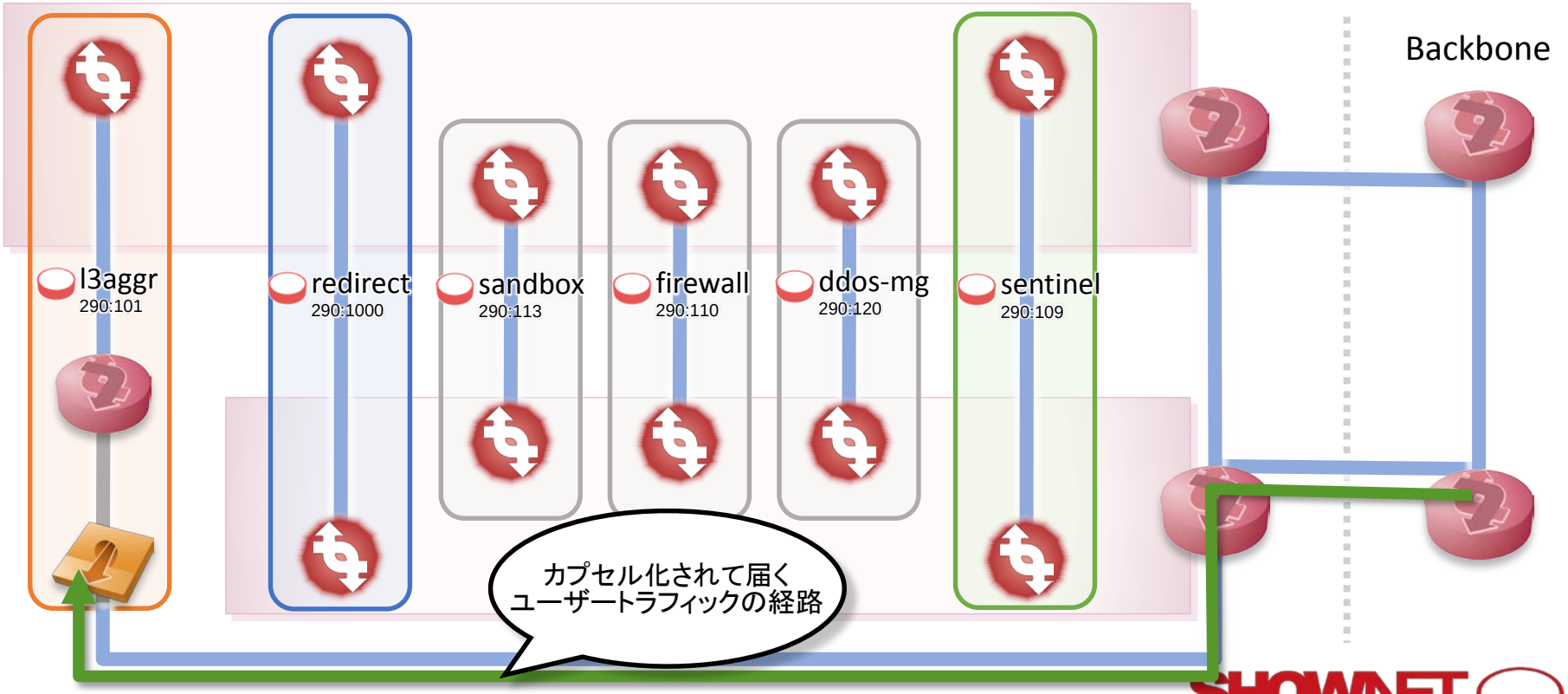


VRFによる機能収容



SHOWNET
INHERIT THE INTENTION

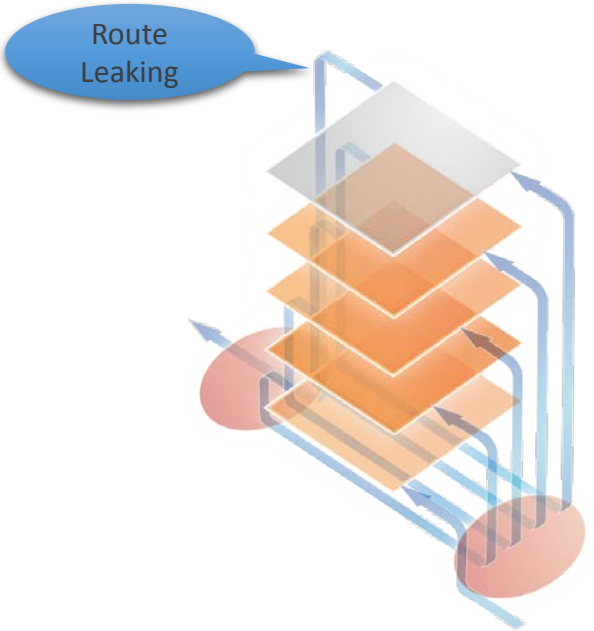
VRFによる機能収容



SHOWNET
INHERIT THE INTENTION

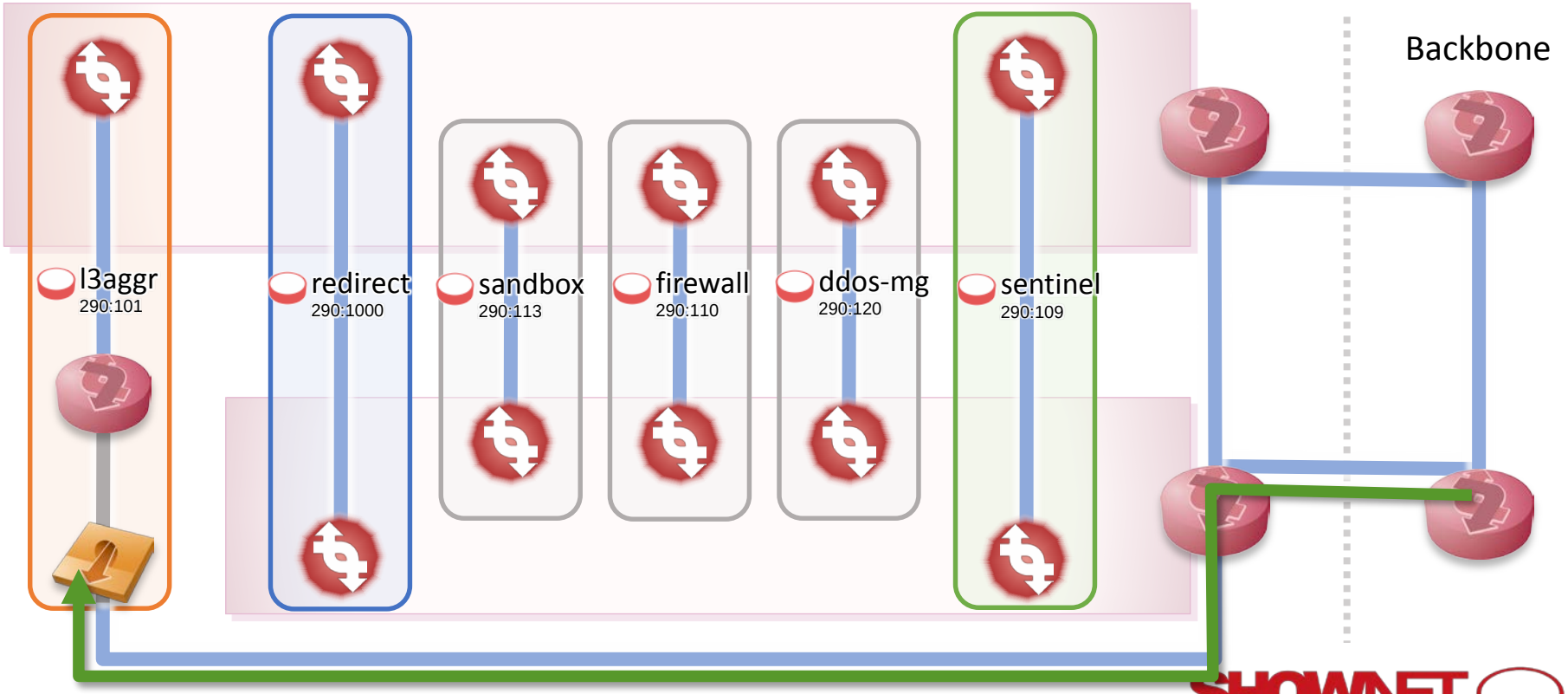
Route Leakingを用いたデフォルトチェイン構築

- VRF間のRoute Leakingを用いて必須ファンクションのみのチェインを構築
  - Route Leaking: VRF間で経路情報を漏洩させる手法（VRF間ルーティングを実現可能）
  - デフォルトで全てのトラフィックがこのチェインを通る
    - BGP Flowspecが止まった場合でも、疎通性が確保できる



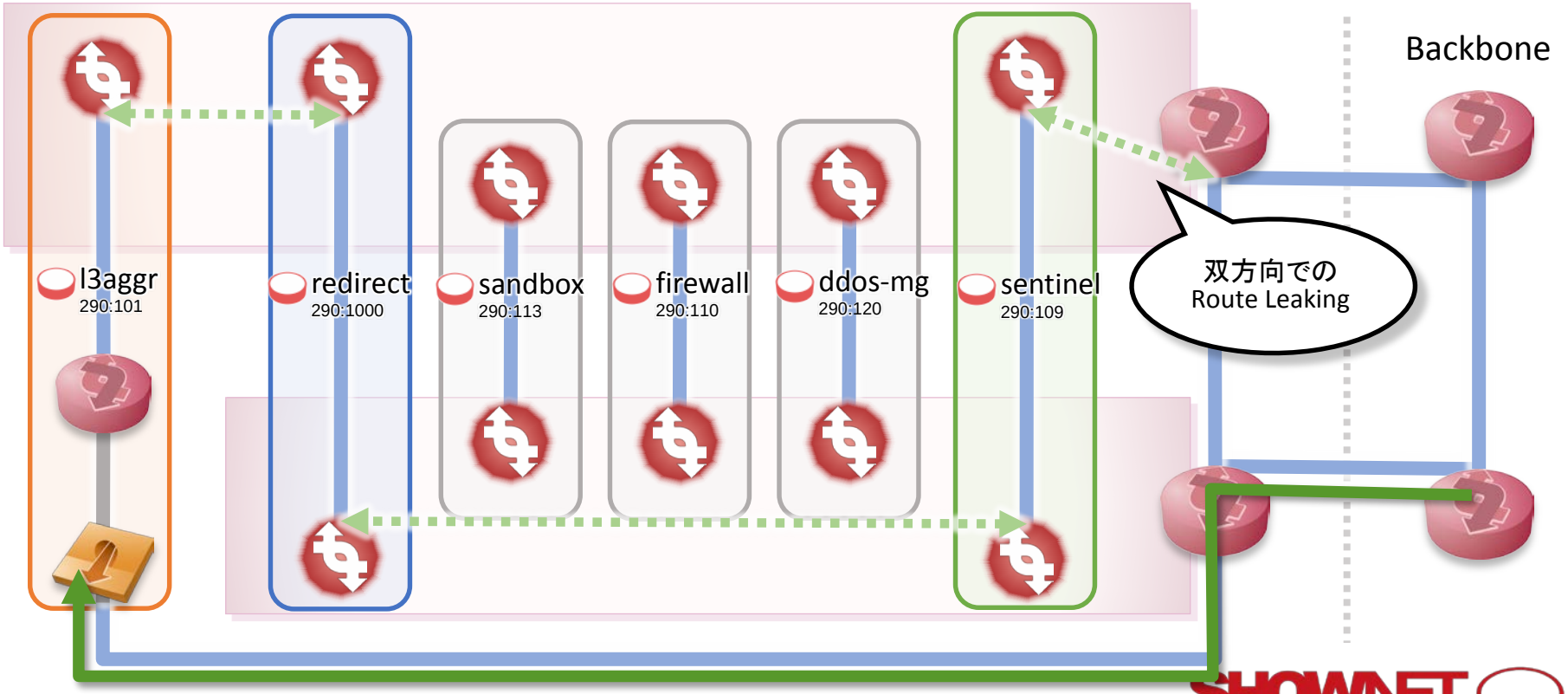
SHOWNET
INHERIT THE INTENTION

Route Leakingを用いたデフォルトチェイン構築

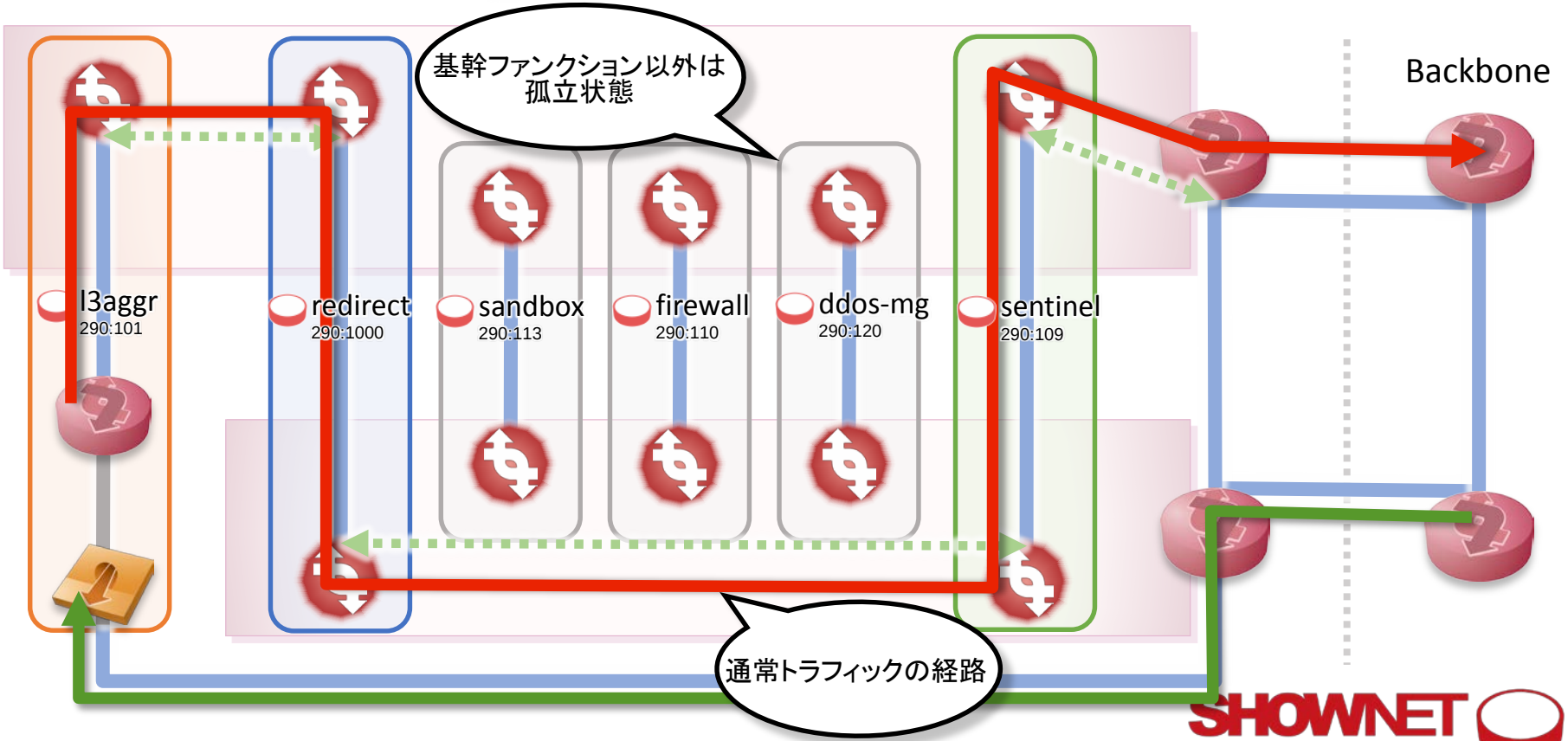


SHOWNET
INHERIT THE INTENTION

Route Leakingを用いたデフォルトチェイン構築



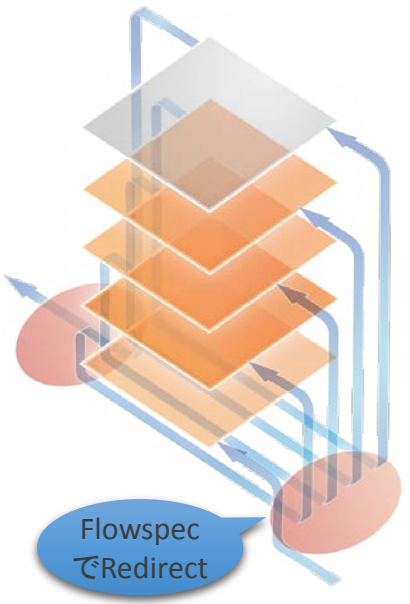
Route Leakingを用いたデフォルトチェイン構築



SHOWNET
INHERIT THE INTENTION

BGP Flowspecによる経路の"上書き"

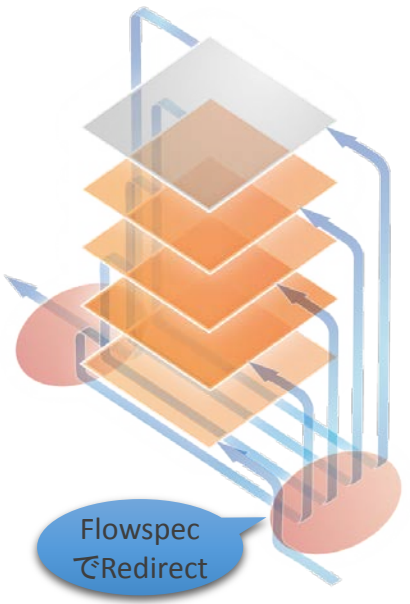
- （パケットがデフォルトチェインを通る過程で）BGP Flowspecにより、チェインを上書きすることでサービスチェイニングを実現
  - BGP Flowspec: BGP Flow specification
    BGPでフィルタールールを伝搬する技術
    RFC5575
  - BGPのNLRIを使って細かなパケットのタイプを指定し、適用するアクションを伝搬する



SHOWNET
INHERIT THE INTENTION

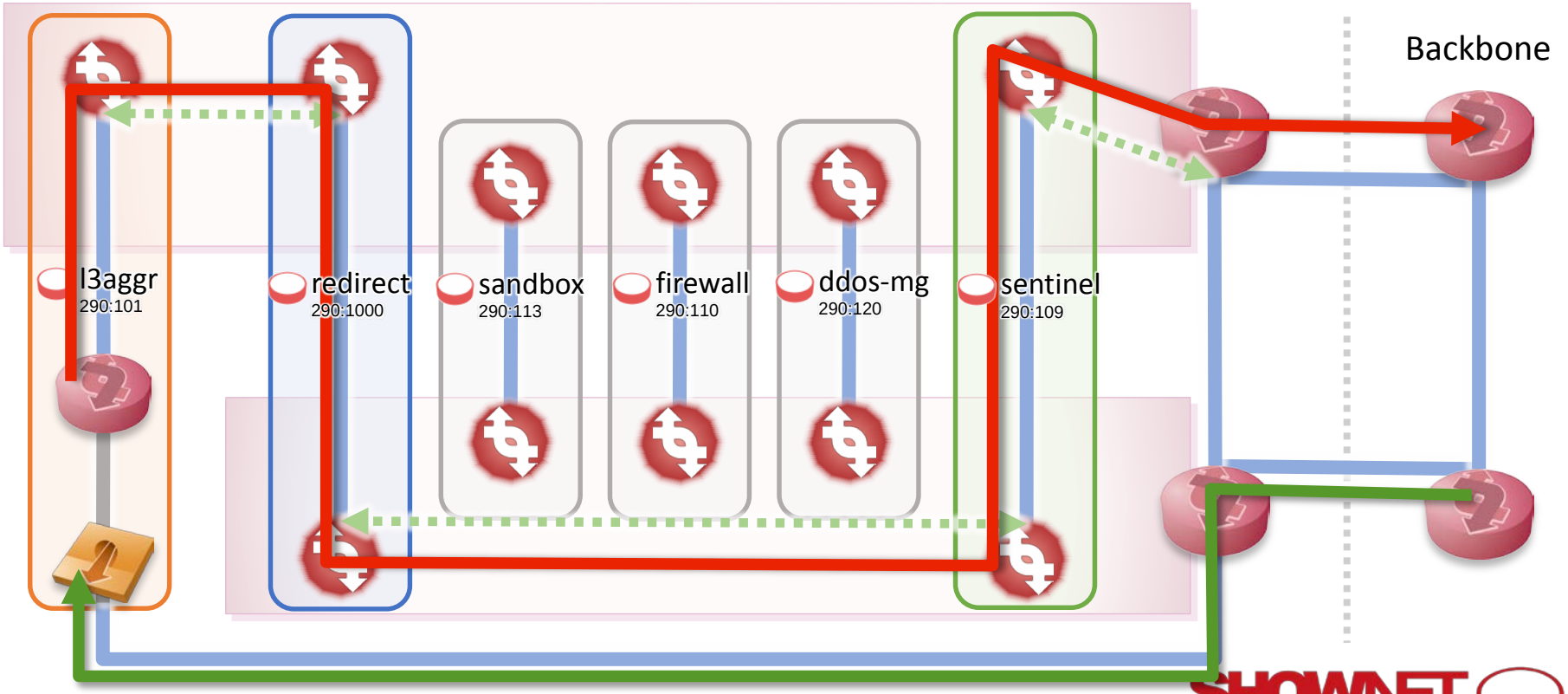
BGP Flowspecによる経路の"上書き"

- BGP Flowspecを用いると、マッチしたパケットに対して様々なアクションが可能
  - Permit: 通常通り転送
  - Drop: 破棄
  - VRF Redirect: 別のVRFにパケットを乗せ替える
  - Marking: IPヘッダのToSフィールドに任意の値を埋め込む



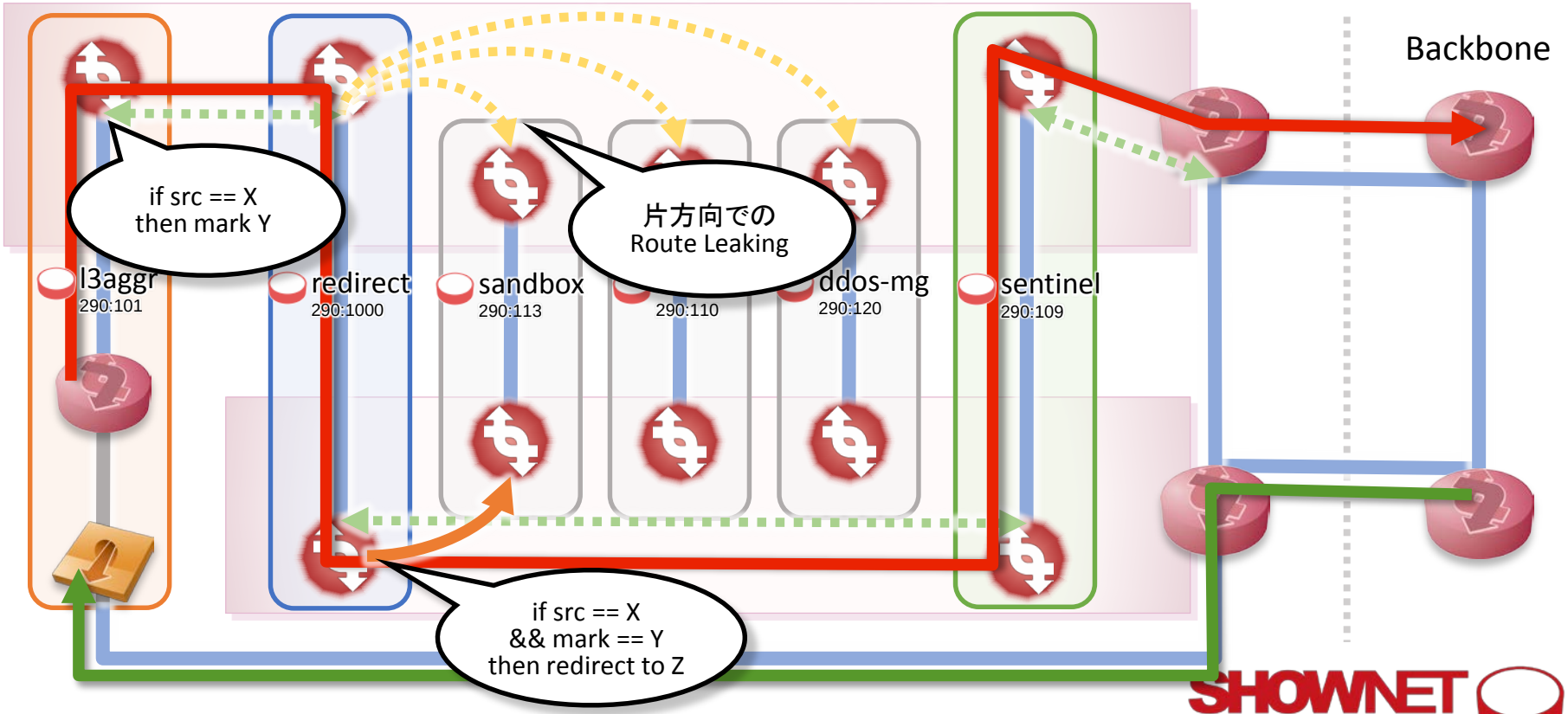
SHOWNET
INHERIT THE INTENTION

BGP Flowspecによる経路の"上書き"（上りの場合）



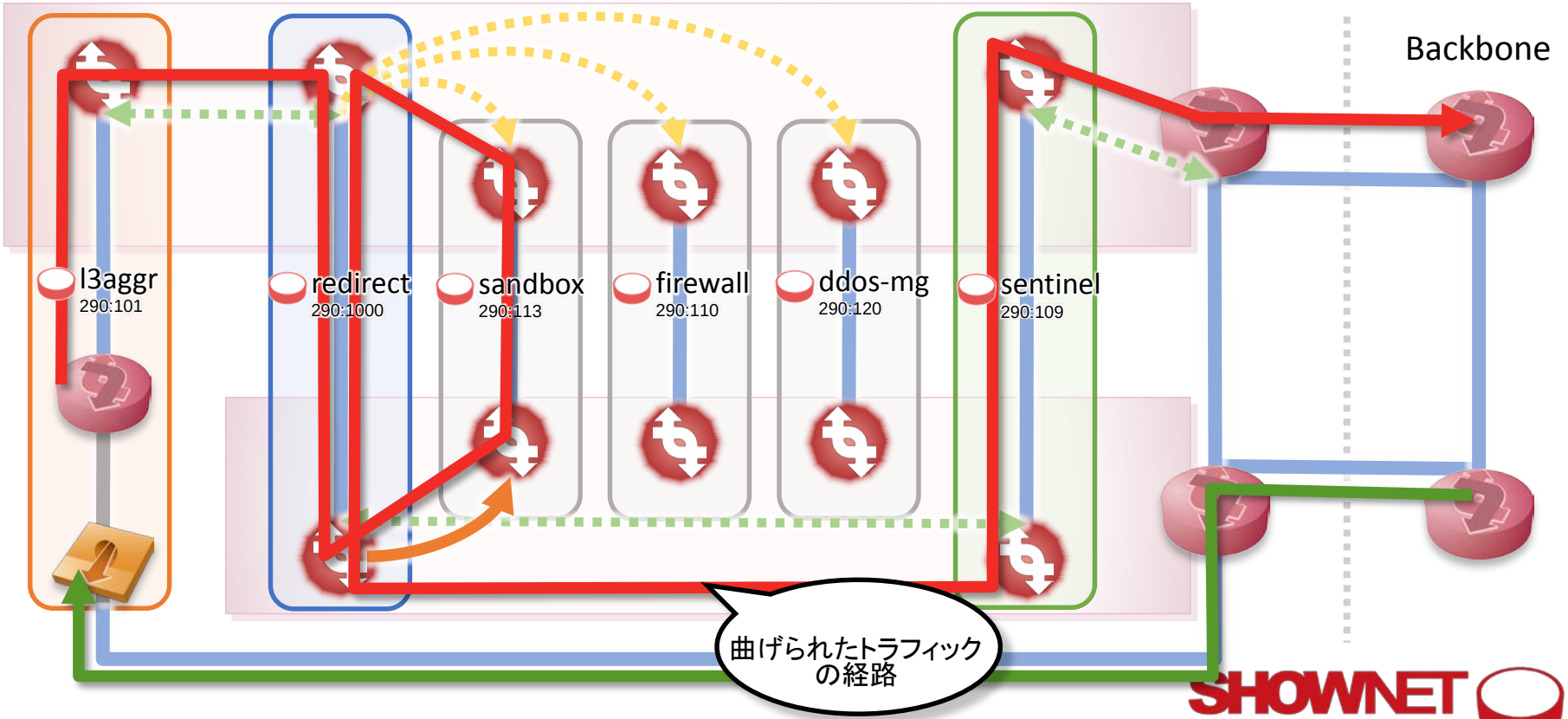
SHOWNET
INHERIT THE INTENTION

BGP Flowspecによる経路の"上書き"（上りの場合）



SHOWNET
INHERIT THE INTENTION

BGP Flowspecによる経路の"上書き"（上りの場合）



SHOWNET
INHERIT THE INTENTION

トレーサビリティ

- サービスチェイニングの運用にはトレーサビリティが必須
  - 各VRFのIPアドレスの逆引きに機能名を記載

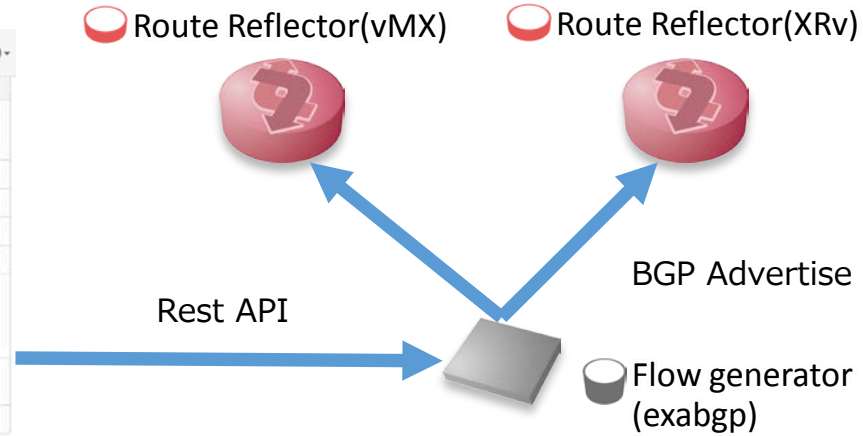
```
 My traceroute  [v0.86]
 Host                                                    Loss%   Snt   Last   Avg  Best  Wrst StDev
 //snip
 4. 10.0.2.137                                           0.0%    10    1.2   1.0   0.8   1.2   0.0
 5. 10.0.2.42                                            0.0%    10    0.9   0.9   0.7   1.1   0.0
 6. 10.0.2.161                                           0.0%    10    2.0   1.8   1.5   2.2   0.0
 7. 10.0.2.42                                            0.0%    10    1.1   1.2   1.1   1.4   0.0
 8. 10.0.2.61                                            0.0%    10    1.8   1.9   1.7   2.1   0.0
 9. vrf-cgn.class0.mx240-1.fpbl.interop-tokyo.net        0.0%    10    2.1   2.0   1.7   2.3   0.0
10. vrf-redirect.class0.mx240-2.fpbl.interop-tokyo.net   0.0%    10    2.0   5.9   1.9  37.9  11.3
11. vrf-ddosmg-a10.class0.mx240-1.fpbl.interop-tokyo.net 0.0%    10    3.2   2.3   1.8   3.2   0.0
12. vrf-redirect.class0.mx240-2.fpbl.interop-tokyo.net   0.0%    10    2.2   8.5   2.0  40.5  13.8
13. vrf-sentinel.class0.mx240-1.fpbl.interop-tokyo.net   0.0%    10    3.2   2.4   2.0   3.2   0.0
14. 45.0.1.89                                            0.0%    10    2.4   4.6   2.3  16.8   4.8
15. 45.0.1.5                                             0.0%    10    3.6   3.6   3.1   4.2   0.0
16. 218.100.6.173                                        0.0%    10    3.2   3.2   3.0   3.6   0.0
17. 108.170.242.193                                      0.0%     9    3.6   3.7   3.5   4.2   0.0
18. 216.239.54.21                                        0.0%     9    3.6   3.6   3.4   3.7   0.0
19. google-public-dns-a.google.com                       0.0%     9    3.6   3.2   3.0   3.6   0.0
```

BGPコントローラ

- BGP FlowspecのNLRIはNOC手製のBGPコントローラから配信
  - NLRI: BGP Updateでやりとりされる経路情報
  - オープンソースのBGP実装であるexaBGPを使用
    - https://github.com/Exa-Networks/exabgp

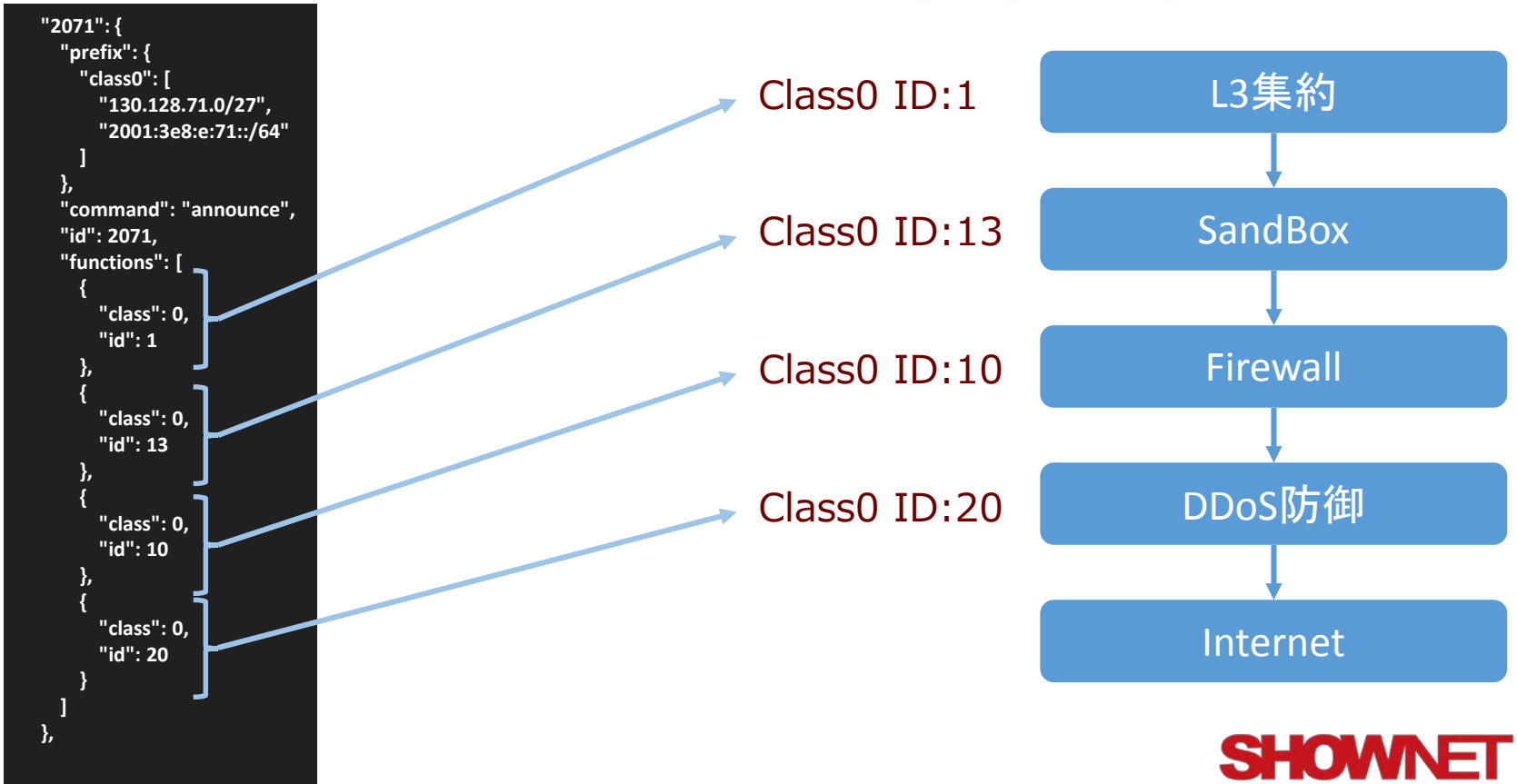
チェイン構成の自動化

- ShowNetの機器情報、出展社向け回線情報などを一元管理するデータベース(TTDB)とREST APIにより連携



SHOWNET INHERIT THE INTENTION

REST APIリクエストボディの例



SHOWNET
INHERIT THE INTENTION

運用結果

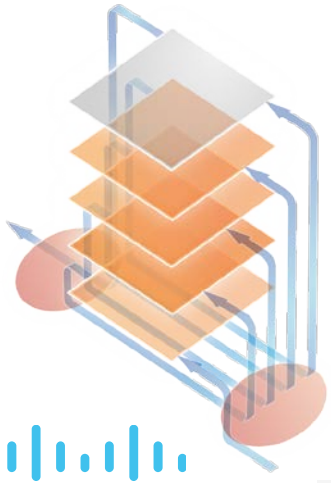
- VRF数
  - 21個 / Device
- フロー数
  - IPv4: 3240経路 / Device
  - IPv6: 2564経路 / Device
  - 合計: 5804経路 / Device
- ユーザー数
  - 200弱
    - 出展社、来場者無線等の
      ユーザトラフィックを全て収容

SHOWNET
INHERIT THE INTENTION

まとめ

まとめ

- 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
  1. 多種多様なサービスを収容可能な汎用性
  2. 大規模イベントでの本番使用に耐える可用性
  3. 既存のIP技術の活用による実現性



CISCO

JUNIPER NETWORKS

exa networks

SHOWNET INHERIT THE INTENTION

関係コントリビュータ・機器一覧

- Router(VRF + BGP Flowspec)
  - **Juniper** MX240
  - **Cisco** ASR9904
- Switches
  - **Cisco** Nexus 93180
  - **Huawei** Cloud Engine 6860, Cloud Engine 6870
- Tunnel End Points
  - **Huawei** Cloud Engine 6855hi
  - **Brocade** VDX6740
  - **Juniper** QFX5100
  - **NEC** IX3315
  - **日本電信電話** Lagopus
  - **Virtual Open Systems** ARMv8 CPE
- Tunnel End Points(cont.)
  - **古河電気工業株式会社** FX1
  - **Cisco** Nexus 9372
  - **Dell EMC** S4048-ON
  - **IP Infusion** OCNOS
- Network Functions
  - **A10 Networks** Thunder 6435
  - **NEC** Traffic Management System
  - **Cisco** ASR9001
- Security Functions
  - **Fortinet** FortiGate-7040e, 3815d, FortiSandbox3500d
  - **PaloAlto Networks** PA5260, WildFire
  - **Checkpoint** 23800, TE2000X
  - **IXIA** ThreatARMOR
  - **NTT-AT** Herculon DHD i10800, BIG-IP LTM i10800
  - **A10 Networks** Thunder 6435 TPS

SHOWNET
INHERIT THE INTENTION