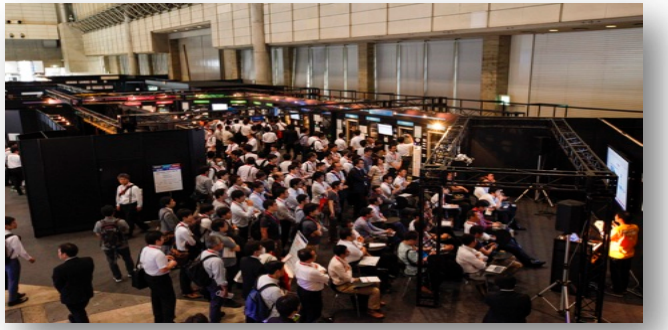
IP経路制御技術でつくる
実践的なサービスチェイニング@ShowNet 2018

Interop Tokyo 2018 ShowNet NOCチームメンバー
中村 遼
ONIC japan 2018

SHOWNET
DIVE INTO THE NEXT

Interop Tokyo

- 世界最大のネットワーク機器と技術の展示会
  - 2018年は25回目の開催
  - 毎年6月に幕張メッセで開催
  - 来場者数約14万人



SHOWNET DIVE INTO THE NEXT

- “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

2018年の検証やデモンストレーション

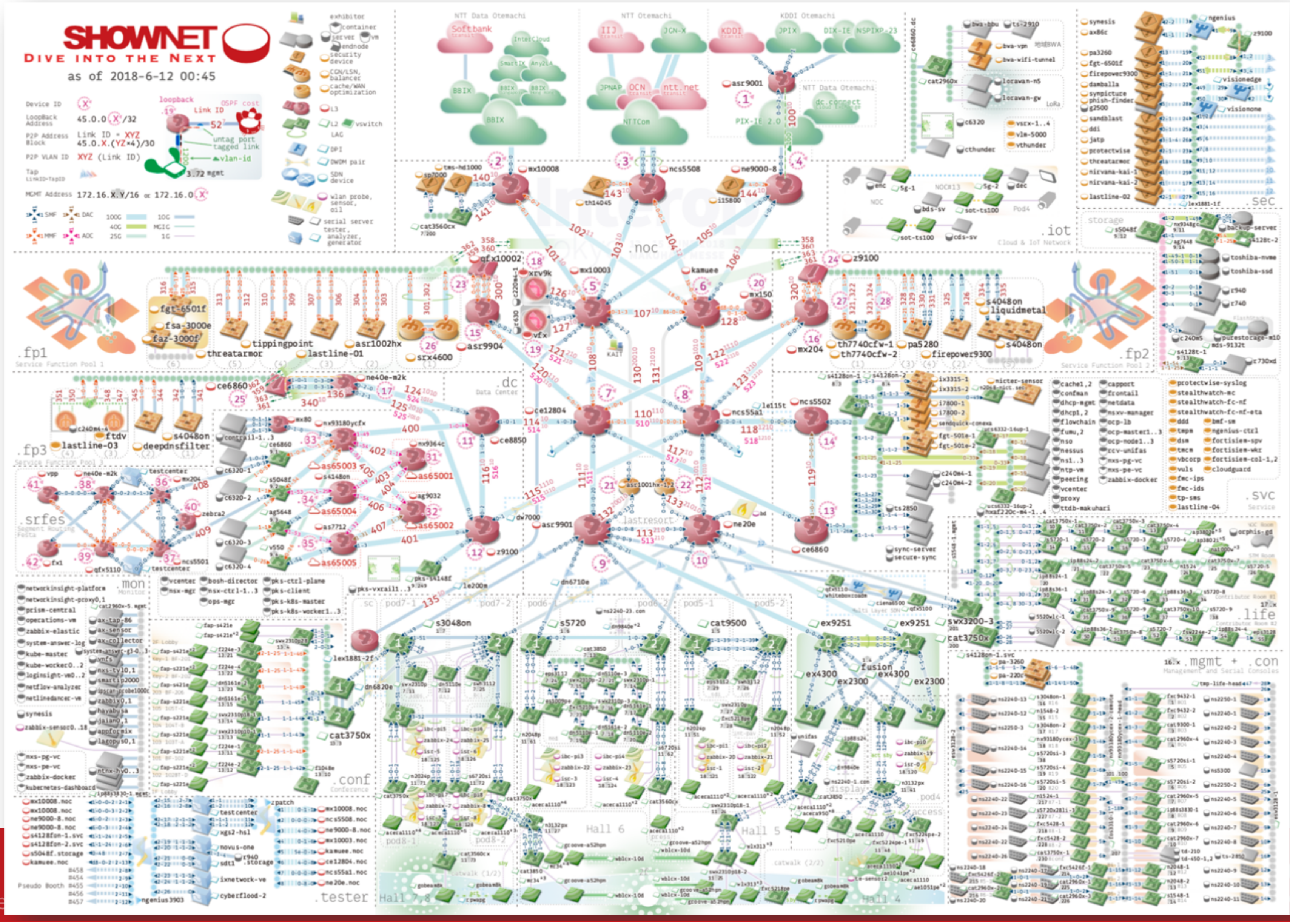
- L2/L3
  - BGP FlowspecとOSPFによるサービスチェイン
  - Segment Routing相互接続性検証
- Wireless
  - 802.11ac wave2 とNBASE-T, MGBASE-Tによる大容量化
  - 無線空間、利用状況のサーベイと評価
- Security
  - サービスチェイニングによる柔軟なセキュリティ機能の提供
  - EDRによる脅威検知後の迅速な対処への備え
  - 多種多様なセキュリティ製品によるモニタリングと相関分析
  - 複数のポリシ施工点における階層的な攻撃防御
- DC/Server
  - ハイパーコンバージド＆コンテナによるスケール可能なインフラ基盤
  - NVMeを活用した高速ストレージ
  - 25G/100GベースのCLOS/EVPN+VXLANネットワーク
  - DCネットワークの構成管理とクイックデプロイメント
- Internet of Things
  - IoTデバイスの低消費電力、長距離通信を支えるLPWA
  - 地域BWA/LTE閉域網によるセキュアで安定した無線通信の実現
  - クラウド上で稼働する仮想モバイルコア
  - 次世代モバイル通信システム(5G)による広帯域・低遅延通信

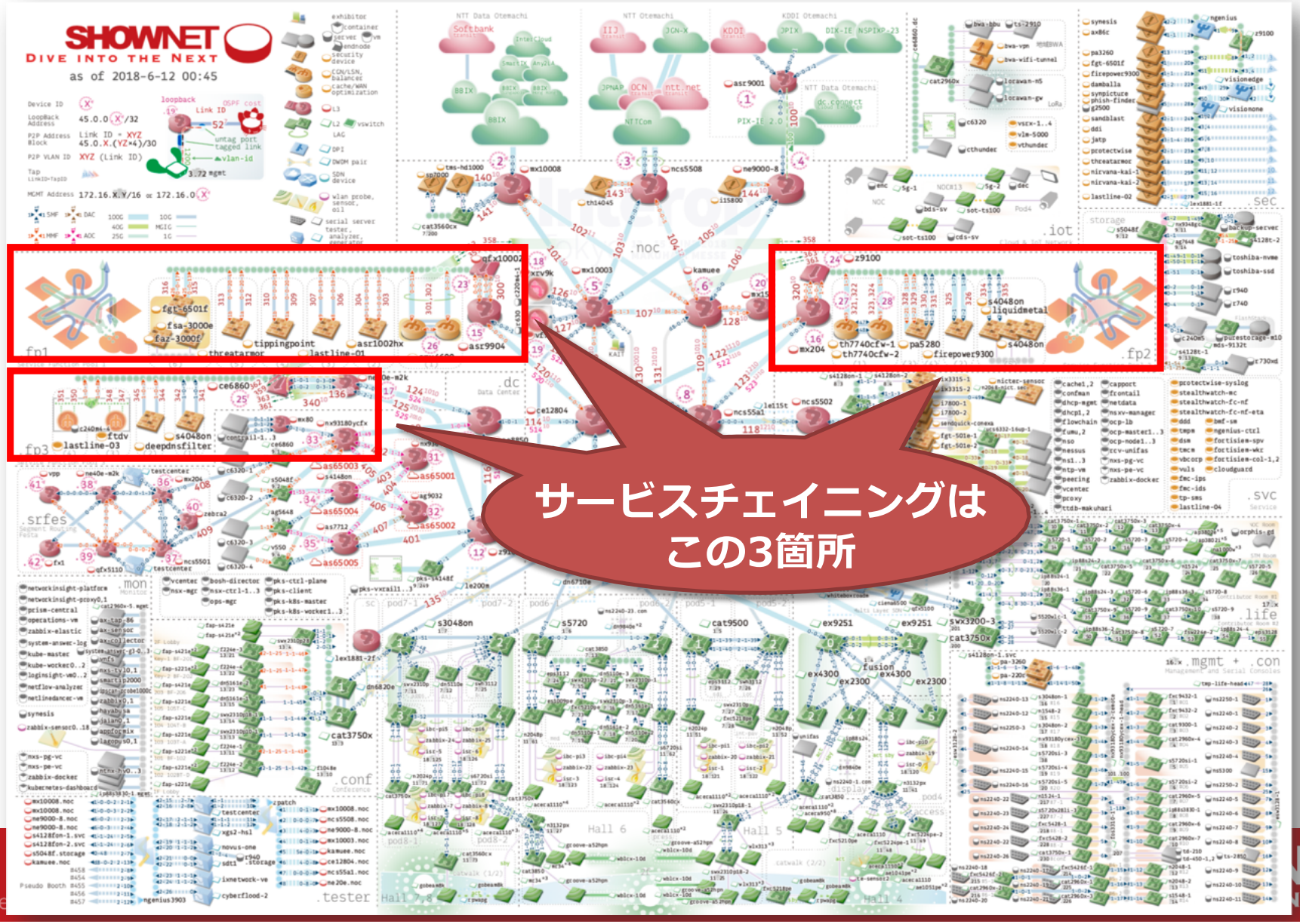
他にもL1からL7まで様々な取り組みを実施

2018年の検証やデモンストレーション

- L2/L3
  - **BGP FlowspecとOSPFによるサービスチェイン**
  - Segment Routing相互接続性検証
- Wireless
  - 802.11ac wave2 とNBASE-T, MGBASE-Tによる大容量化
  - 無線空間、利用状況のサーベイと評価
- Security
  - サービスチェイニングによる柔軟なセキュリティ機能の提供
  - EDRによる脅威検知後の迅速な対処への備え
  - 多種多様なセキュリティ製品によるモニタリングと相関分析
  - 複数のポリシ施工点における階層的な攻撃防御
- DC/Server
  - ハイパーコンバージド＆コンテナによるスケール可能なインフラ基盤
  - NVMeを活用した高速ストレージ
  - 25G/100GベースのCLOS/EVPN+VXLANネットワーク
  - DCネットワークの構成管理とクイックデプロイメント
- Internet of Things
  - IoTデバイスの低消費電力、長距離通信を支えるLPWA
  - 地域BWA/LTE閉域網によるセキュアで安定した無線通信の実現
  - クラウド上で稼働する仮想モバイルコア
  - 次世代モバイル通信システム(5G)による広帯域・低遅延通信

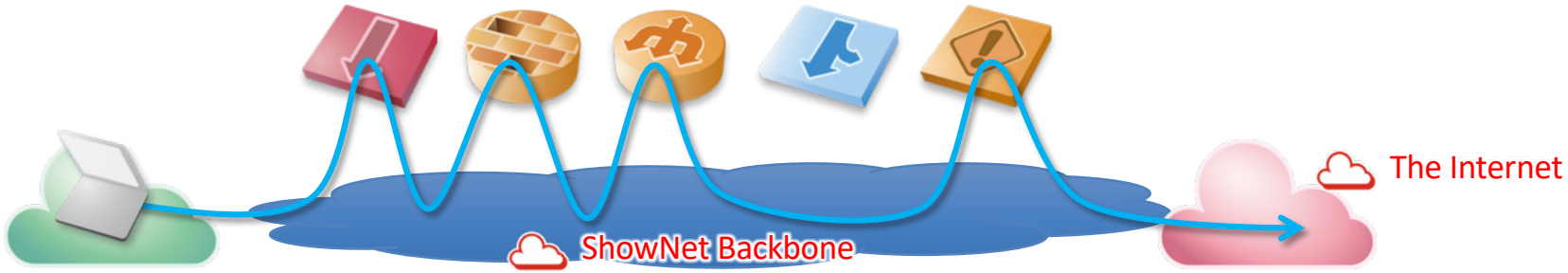
他にもL1からL7まで様々な取り組みを実施





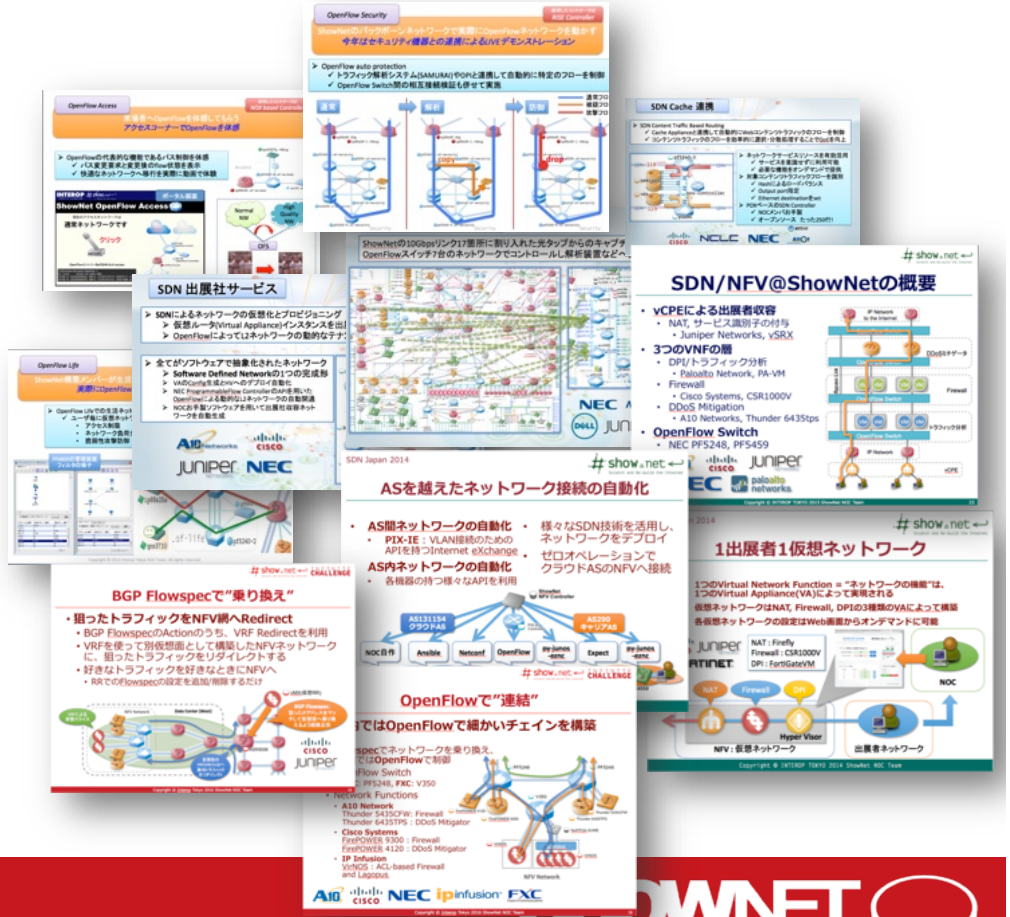
Service Chaining

- ネットワークの機能(Middlebox)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
  - 障害部分の自動迂回
- 最短経路配送のIPには難しい=>SDN/NFV技術の利用



2016年まではSDN/NFVを志向

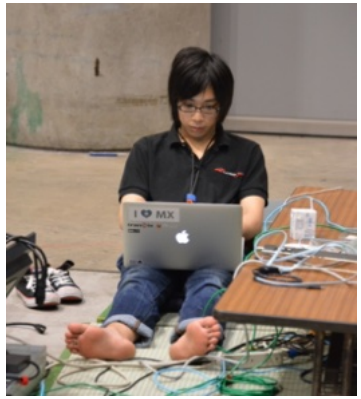
- ShowNetにおける変遷
  - 2012年: OpenFlowによるデモ
  - 2013年: 仮想ルータによるユーザ収容
  - 2014年: 仮想ネットワーク(NAT, FW, DPI)によるユーザ収容
  - 2015年: OpenFlowによるService Chaining (IEEE CoolSDN WS, IRTF nfvrgで発表)
  - 2016年: OpenFlowとBGP Flowspecを併用したService Chaining



5年間で得られた知見

- やはりIPと比べると、SDN/NFVは構築・運用が難しい
  - 開発コスト: データプレーン、コントロールプレーン
  - 運用コスト: 専用のパケット転送/制御ロジックへの精通
  - とりわけShowNetでは、2週間で構築、検証を行う

日中



明け方

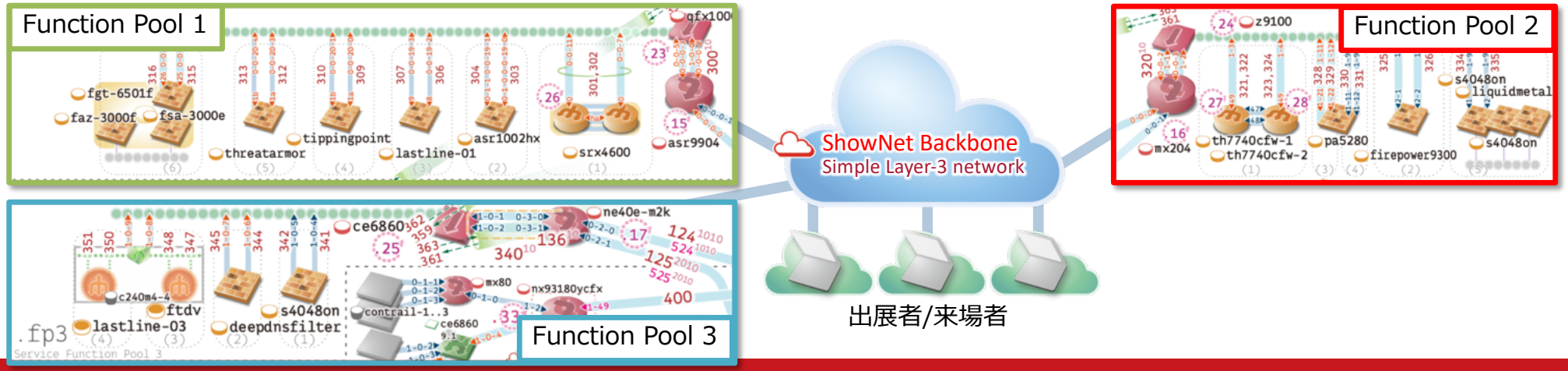
2017年からのチャレンジ

- IP経路制御技術でどこまでやれるか
  - 2016年まではSDN/NFVを志向: OpenFlow, Virtual Appliance
  - 2017年からはNetwork Programmabilityが普及する前の、IP経路制御技術の進化と限界を見極めるチャレンジを実施

SHOWNET
DIVE INTO THE NEXT
Interop
Tokyo 13 - 15 JUNE 2018
MAKUHARI MESSE

IP経路制御によるService Chaining@ShowNet 2018

- 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
- ユーザに提供する機能は**Function Pool**に集約
- Pool内では**BGP Flowspec**を用いてService chainingを実現
- Pool間は**EVPN/VXLAN**で接続、拠点をまたいだチェインを実現



SHOWNET DIVE INTO THE NEXT

Key technologies

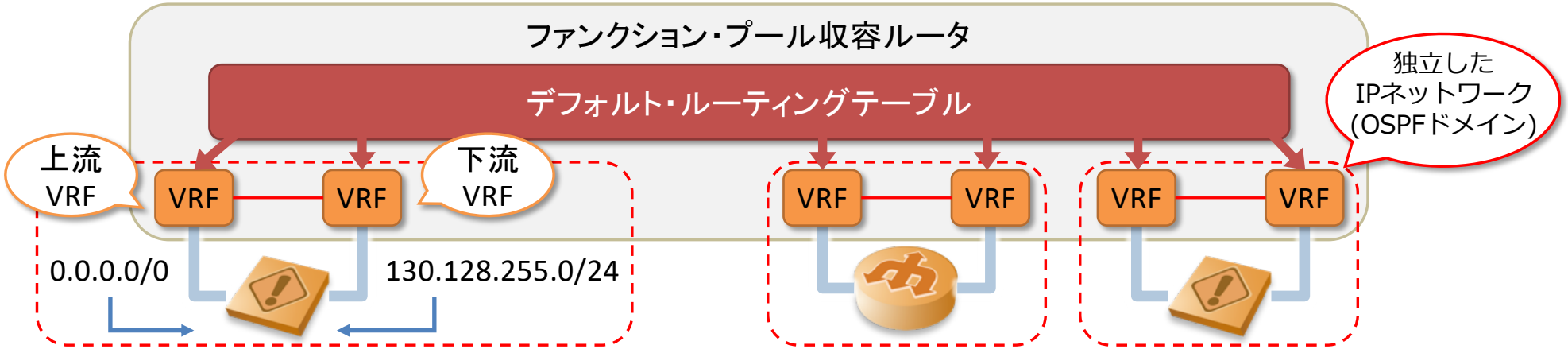
- BGP Flow Specification (RFC5575)
  - NLRIでAccess Control Listを伝搬するBGP拡張
    - Match: アドレス、ポート、プトロコル番号、ICMP typeなど
    - Action: Rate-control, Drop, Redirect (VRF)など
    - 本来の用途はDDoSミチゲーションなど
- Virtual Routing and Forwarding (VRF)
  - VPN経路を保持するための独立した経路表(BGP/MPLS VPN)
  - (ひとつのルータ内に持てる複数の経路表)

Flowspecを使ったファンクションの連結

- IP経路制御によるファンクションチェイン
  - VRFを用いてファンクションごとに個別のネットワークを構築
  - BGP Flowspecでユーザごとに適切なファンクションへと次々乗り換え
- 各ファンクションのネットワークは個別のOSPFドメイン
  - 様々な形態(L2, L3, over L4)の機器をファンクションとして収容可能

個別のネットワークとチェインの構築

- VRFを用いてFunctionごとに個別のIPネットワークを構成
  - 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

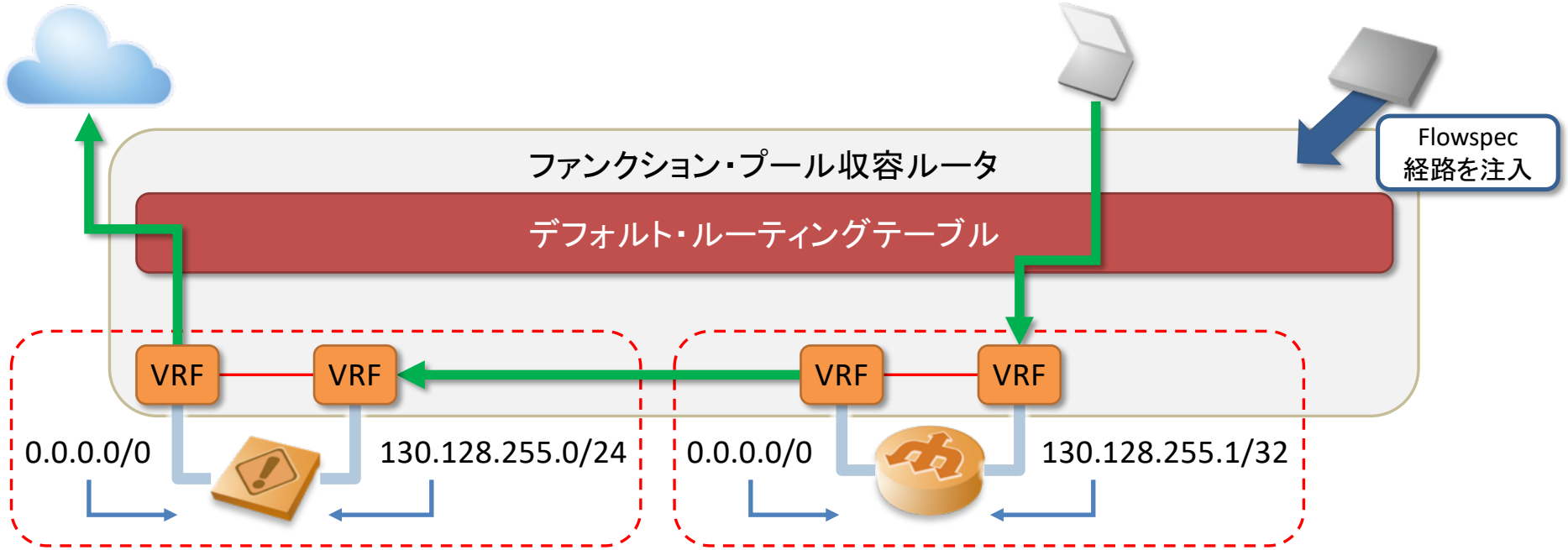


個別のネットワークとチェインの構築 (cont'd)

- Flowspec VRF Redirectによるネットワークの乗り換え
  - 上流行きはSrc Prefix Match, Action Redirect, Target 下側VRF
  - 下流行きはDst Prefix Match, Action Redirect, Target 上側VRF

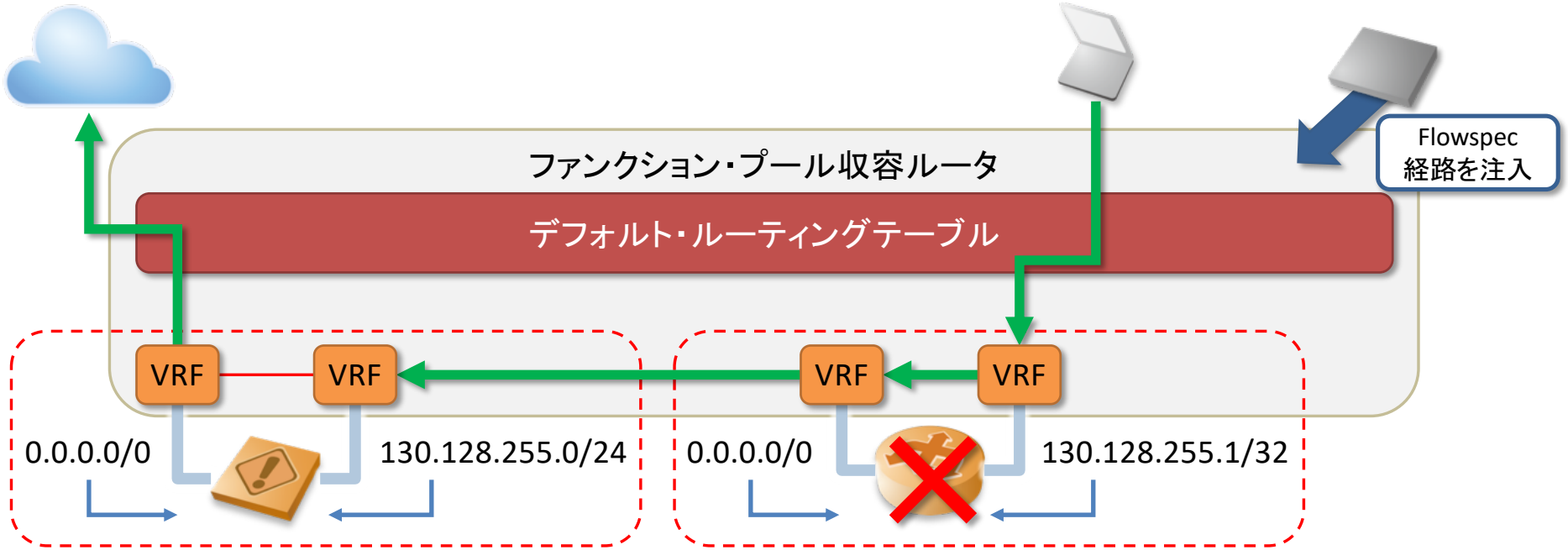
個別のネットワークとチェインの構築 (cont'd)

- Functionの連結も同様にVRF Redirectで実施



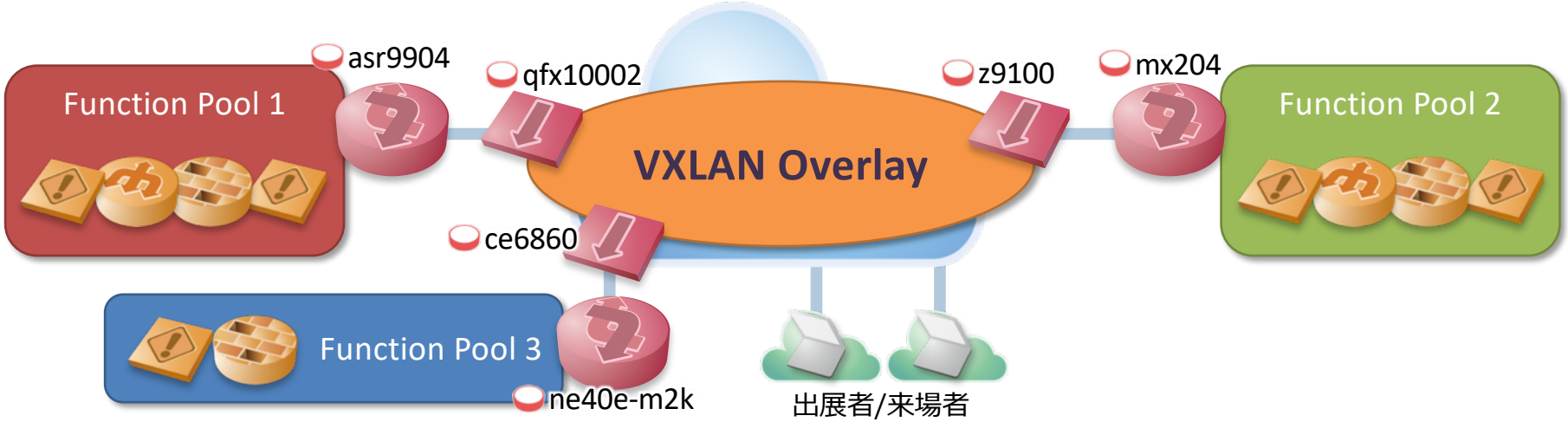
個別のネットワークとチェインの構築 (cont'd)

- Functionの障害時はOSPFでフェイルオーバ



3つのFunction Poolを構築

- VXLAN/EVPNによる拠点間トラフィック転送
  - VXLANオーバーレイでファンクションプール間を接続
  - 拠点をまたいでファンクションをつなぐサービスチェインも構成可能
  - ExaBGPを用いたFlowspecスピーカを実装



実際の設計や構成はもう少し(だいぶ？)複雑

- 論文として11月にPublish予定
  - Ryo Nakamura, Yukito Ueno, Teppei Kamata, Kazuki Shimizu, Takashi Tomine, Takayuki Watanabe, “Practical Service Chaining based on IP Routing”, Asian Internet Engineering Conference, Nov 2018 (will be appeared)

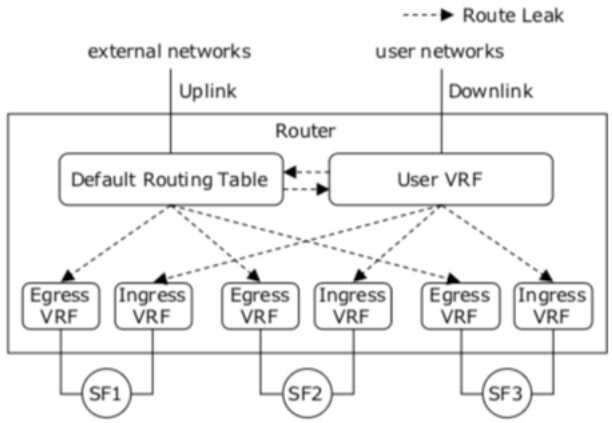


Figure 2: The basic setup of VRFs in a router accommodating service functions.

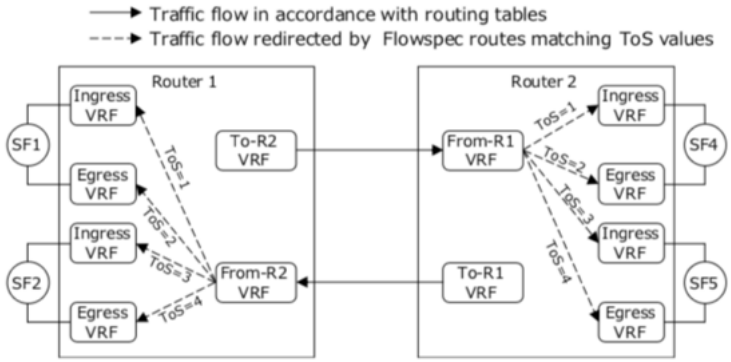


Figure 4: Constructing service chaining with two routers. Two unidirectional point-to-point links bring packets to each router.

SHOWNET DIVE INTO THE NEXT

稼働実績

- Flowspec経路数: 約2700経路/ルータ
- VRF数: 約20/ルータ
- サービスチェイン数: 492
- Function数: 15

Network Function一覧

- Function Pool 1
  - SRX4600, ASR1002HX, Tipping Point, ThreatARMOR, Fortigate, Lastline
- Function Pool 2
  - Thunder 7740 CFW, PA-5280 (Firewall), PA-5280 (Captive Portal), Firepower 9300, S4048ON + Liquid Metal
- Function Pool 3
  - S4048ON, Firepower Thread Defense Virtual, Lastline, deep-dns-filter

冗談のようなTraceroute

- ShowNet内で28hop
  - Function Poolをまたぎつつ、
  - すべてのファンクションを通過するサービスチェイン

```
C:¥>tracert -d -h 64 8.8.8.8

Tracing route to 8.8.8.8 over a maximum of 64 hops

  1     1 ms     1 ms     1 ms  45.0.248.2
  2    23 ms    52 ms    34 ms  45.0.2.45
  3     1 ms     1 ms    <1 ms  45.0.2.82
  4    <1 ms    <1 ms    <1 ms  45.0.3.17
  5     1 ms     1 ms     1 ms  45.0.3.13
  6     1 ms     1 ms     1 ms  45.0.3.234
  7     1 ms     1 ms     1 ms  45.0.3.101
  8     1 ms     1 ms     1 ms  45.0.3.246
  9     2 ms     1 ms     1 ms  45.0.3.165
 10     1 ms     1 ms     1 ms  45.0.3.249
 11    12 ms    11 ms    11 ms  45.0.3.25
 12    12 ms    12 ms    12 ms  45.0.3.234
 13    12 ms    12 ms    13 ms  45.0.3.113
 14     2 ms     2 ms     2 ms  45.0.3.246
 15    13 ms    13 ms    12 ms  45.0.3.177
 16     2 ms     2 ms     2 ms  45.0.3.249
 17    13 ms    13 ms    13 ms  45.0.3.37
 18    13 ms    13 ms    13 ms  45.0.3.234
 19    53 ms    13 ms    13 ms  45.0.3.137
 20     2 ms     2 ms     2 ms  45.0.3.241
 21    13 ms    13 ms    13 ms  45.0.3.61
 22     2 ms     2 ms     2 ms  45.0.3.238
 23    26 ms    24 ms    24 ms  45.0.3.201
 24     2 ms     2 ms     2 ms  45.0.3.249
 25    24 ms    24 ms    24 ms  45.0.3.49
 26    60 ms    76 ms    40 ms  45.0.1.81
 27    24 ms    24 ms    24 ms  45.0.1.33
 28    39 ms    26 ms    24 ms  45.0.1.5
 29    25 ms    25 ms    25 ms  218.100.6.173
 30    27 ms    26 ms    27 ms  108.170.242.193
 31    25 ms    25 ms    25 ms  209.85.250.23
 32    25 ms    25 ms    25 ms  8.8.8.8

Trace complete.
```

冗談のようなTraceroute

- ShowNet内で28hop
  - Function Poolをまたぎつつ、
  - すべてのファンクションを通過するサービスチェイン
  - 同じアドレスが何度も登場しますが、ループではありません

```
C:¥>tracert -d -h 64 8.8.8.8

Tracing route to 8.8.8.8 over a maximum of 64 hops

  1     1 ms     1 ms     1 ms  45.0.248.2
  2    23 ms    52 ms    34 ms  45.0.2.45
  3     1 ms     1 ms    <1 ms  45.0.2.82
  4    <1 ms    <1 ms    <1 ms  45.0.3.17
  5     1 ms     1 ms     1 ms  45.0.3.13
  6     1 ms     1 ms     1 ms  45.0.3.234
  7     1 ms     1 ms     1 ms  45.0.3.101
  8     1 ms     1 ms     1 ms  45.0.3.246
  9     2 ms     1 ms     1 ms  45.0.3.165
 10     1 ms     1 ms     1 ms  45.0.3.249
 11    12 ms    11 ms    11 ms  45.0.3.25
 12    12 ms    12 ms    12 ms  45.0.3.234
 13    12 ms    12 ms    13 ms  45.0.3.113
 14     2 ms     2 ms     2 ms  45.0.3.246
 15    13 ms    13 ms    12 ms  45.0.3.177
 16     2 ms     2 ms     2 ms  45.0.3.249
 17    13 ms    13 ms    13 ms  45.0.3.37
 18    13 ms    13 ms    13 ms  45.0.3.234
 19    53 ms    13 ms    13 ms  45.0.3.137
 20     2 ms     2 ms     2 ms  45.0.3.241
 21    13 ms    13 ms    13 ms  45.0.3.61
 22     2 ms     2 ms     2 ms  45.0.3.238
 23    26 ms    24 ms    24 ms  45.0.3.201
 24     2 ms     2 ms     2 ms  45.0.3.249
 25    24 ms    24 ms    24 ms  45.0.3.49
 26    60 ms    76 ms    40 ms  45.0.1.81
 27    24 ms    24 ms    24 ms  45.0.1.33
 28    39 ms    26 ms    24 ms  45.0.1.5
 29    25 ms    25 ms    25 ms  218.100.6.173
 30    27 ms    26 ms    27 ms  108.170.242.193
 31    25 ms    25 ms    25 ms  209.85.250.23
 32    25 ms    25 ms    25 ms  8.8.8.8

Trace complete.
```

BGP FlowspecによるService Chainingの利点

- 開発コストの小ささ
  - データプレーンは既存のIPルータをそのまま利用可能
    - VRFはMPLS/VPNで、BGP FlowspecはDDoSミチゲーションのため、バックボーンルータであれば商用で実装が進んでいる
  - コントロールプレーンはBGP
    - Route ReflectorにCLIで設定をいれるだけでもService Chainを制御可能
    - ExaBGPやGoBGPなどで実装することも容易
- 積み重ねてきたIPネットワークの技術・スキルの転用
  - OSPFをはじめとする枯れた動的経路制御技術による可用性の実現
  - ネットワークオペレータの慣れ親しんだCLIでもそのまま構築/デバッグ/運用できる
    - 一方で簡単なオーケストレータ(BGP Flowspec speaker)による制御も同時に可能

BGP FlowspecによるService Chainingの欠点

- トラフィック制御の粒度
  - データプレーンは結局IPなので、TEやQoSはIPに引きずられる
  - BGP Flowspecはstatelessな経路なので、セッション単位の制御は難しい
- Pre-configuredなトポロジ
  - VRFの構成やRoute leakなど、ある程度Fixedなトポロジ構成
- 規模性
  - BGP Flowspec routeの数
    - Cisco ASR9000シリーズで約3000経路

今後検討したい技術

- BGP/MPLS VPN
  - Service-topology-RTによる柔軟なトポロジ設計
    - draft-ietf-bess-service-chaining-04
- Segment Routing
  - Data PlaneはSR(MPLS or IPv6)
  - Control PlaneがPCEPで打ち込むだけなら比較的容易か
- Network Service Header
  - 実装が進むかどうか
- P4
  - 実装が進むかどうか、OpenFlowとはまた違った難しさがあるか

まとめ

- IP経路制御技術を用いたサービスチェイニング
  - VRFを用いてFunctionごとに個別のネットワークを構築
  - BGP FlowspecのVRF Redirectでチェインを構築
  - 15 Functions, 492 Service Chainsで実稼働@ShowNet 2018

いわゆるSDN/NFV技術でなくても、
実践的なService Chainingを構築することはできる

Interop Tokyo 2018 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の自由度
    - そして検証とフィードバック


SHOWNET
DIVE INTO THE NEXT