

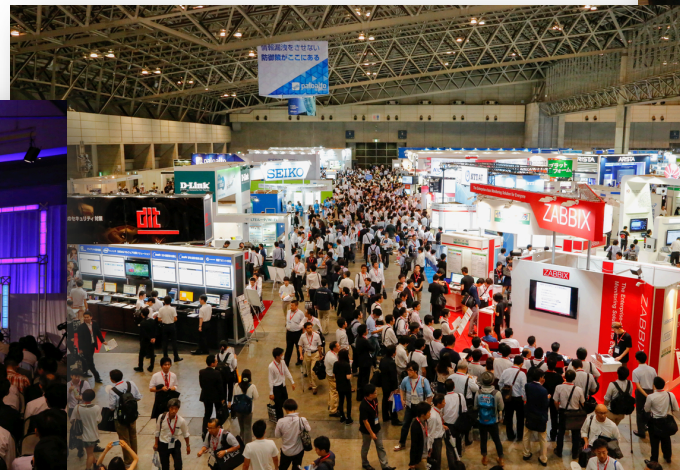
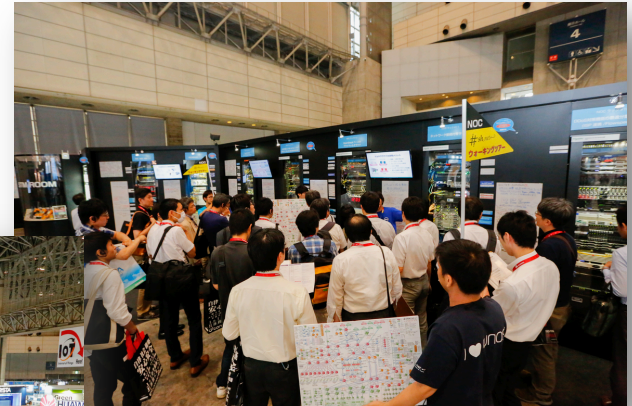
# INTEROP Tokyo ShowNet 2015におけるSDN/NFV

INTEROP Tokyo 2015  
ShowNet NOC Team  
中村 遼



# INTEROP Tokyo

- **世界最大のネットワーク機器と技術の展示会**
  - 2015年で第22回開催
  - 毎年6月に幕張メッセで開催
  - 来場者数約13万人





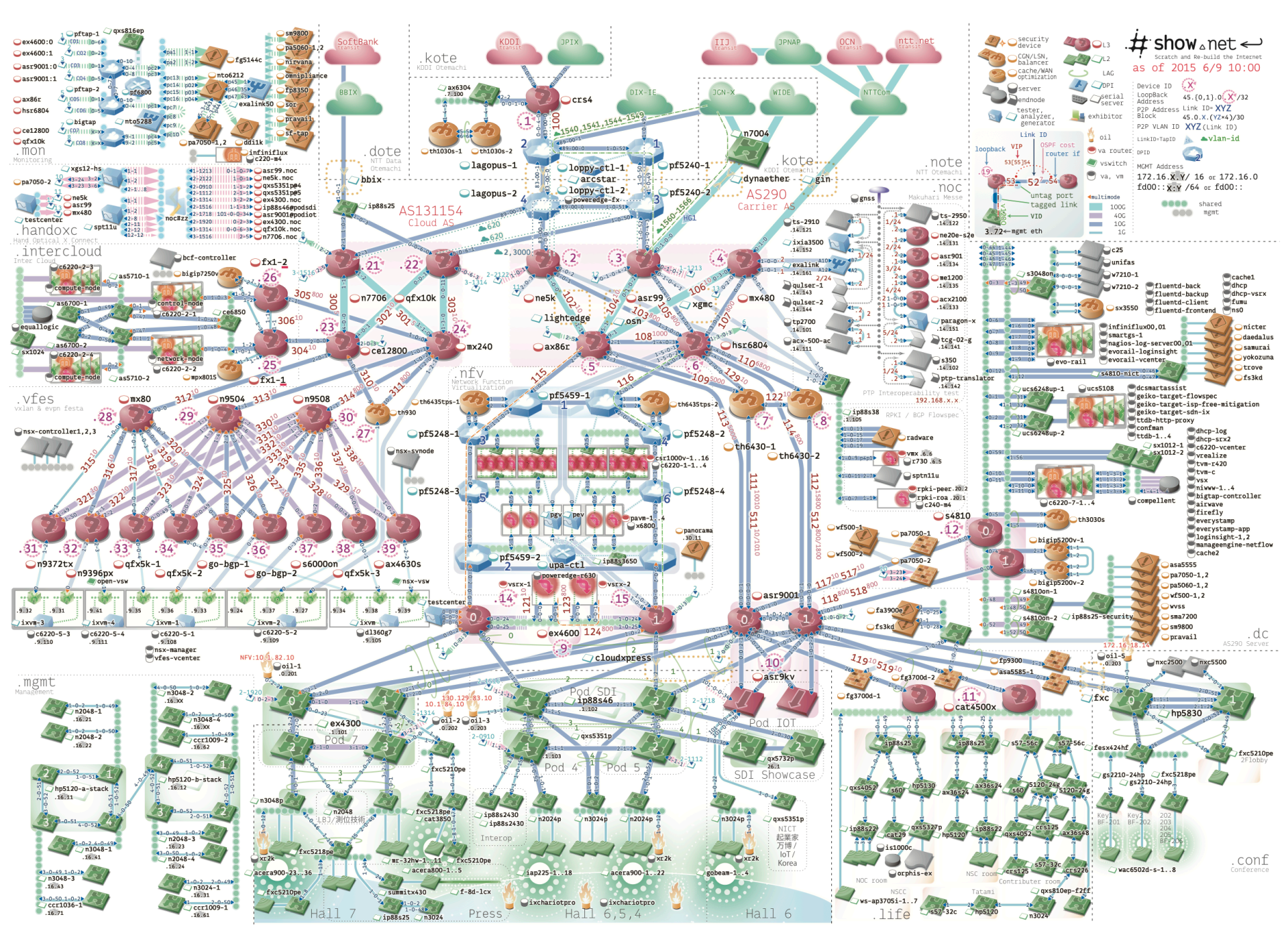
# # show<sub>△</sub>net ←

Scratch and Re-build the Internet

- **“I know it works because I saw it at INTEROP”**
  - 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
  - 最新の技術で10年先のインターネットを構築する
  - 様々な技術の相互接続性検証の場
  - 出展者や来場者へのネットワーク提供









# ShowNetにおけるSDN

- ShowNetは「生きた」ネットワーク
  - 2012年から2015年まで様々なSDNに挑戦
  - ShowNetで求められるのは**動くSDN**
  - そして「**Interoperability**」

### OpenFlow Security

ShowNetのバックボーンネットワークで実際にOpenFlowネットワークを動かす  
今年度はセキュリティ機器との連携によるLIVEデモンストレーション

- OpenFlow auto protection
  - トラフィック解析システム(SAMURAI)やDPIと連携して自動的に特定のフローを制御
  - OpenFlow Switch間の相互接続検証も併せて実施

### SDN Security

ShowNetの10Gbpsリンク17箇所に割り入れた光タップからのキャプチャデータケットを  
OpenFlowスイッチ7台のネットワークでコントロールし解析装置などへ供給

来場者へOpenFlowを体感してもらう  
アクセスコーナでOpenFlowを体感

SDNの代表的な機能であるパス制御を体感  
本要要求と変更後のflow状態を表示  
異なるネットワークへ移行を実際に動画で体感

### SDN Cache 連携

SDN Content Traffic Based Routing  
Cache Applianceと連携して自動的にWebコンテンツトラフィックのフローを制御  
コンテンツトラフィックのフローを効率的に選択・分散処理することでQoSを向上

### SDN 出展社サービス

- SDNによるネットワークの仮想化とプロビジョニング
- 仮想ルータ(Virtual Appliance)インスタンスを出展社ごとに1台ずつ
- OpenFlowによる動的なネットワークの自動構築
- OpenFlowによって2ネットワークの動的なテナントの追加や削除を

全てがソフトウェアで抽象化されたネットワーク  
Software Defined Networkの1つの完成形  
VAのConfig生成とHVへのデプロイ自動化  
NEC ProgrammableFlow ControllerのAPIを用いた  
OpenFlowによる動的なネットワークの自動構築  
NOCお手製ソフトウェアを用いて出展社収容ネットワークを自動生成

### 1出展者1仮想ネットワーク

1つのVirtual Network Function = “ネットワークの機能”は、  
1つのVirtual Appliance(VA)によって実現される

- 仮想ネットワークはNAT, Firewall, DPIの3種類のVAによって構築
- 各仮想ネットワークの設定はWeb画面からオンデマンドに可能

NAT: Firefly  
Firewall: CSR1000V  
DPI: FortiGateVM

Hyper Visor

NOC

出展者ネットワーク

NFV: 仮想ネットワーク

### ASを越えたネットワーク接続の自動化

- AS間ネットワークの自動化
- PIX-IE: VLAN接続のためのAPIを持つInternet eXchange
- AS内ネットワークの自動化
- 各機器の持つ様々なAPIを利用
- 様々なSDN技術を活用し、ネットワークをデプロイ
- ゼロオペレーションでクラウドASのNFVへ接続

AS131154 クラウドAS

AS290 キャリアAS

NOC自作 Ansible Netconf OpenFlow py-junos-etc Expect py-junos-etc

NFV Server /S4800N /ASR9000 /PF248 /MX440 /HP1250 /EX4550

ShowNet NFV Controller



Device ID:  L3  
LoopBack Address: 45.0.1.0/32  
P2P Address: Line ID: XYZ  
Block: 45.0.X.(Y2+4)/30  
P2P VLAN ID: XYZ (Link ID)  
Link ID:   
LinkID-TapID:   
vlan-id:

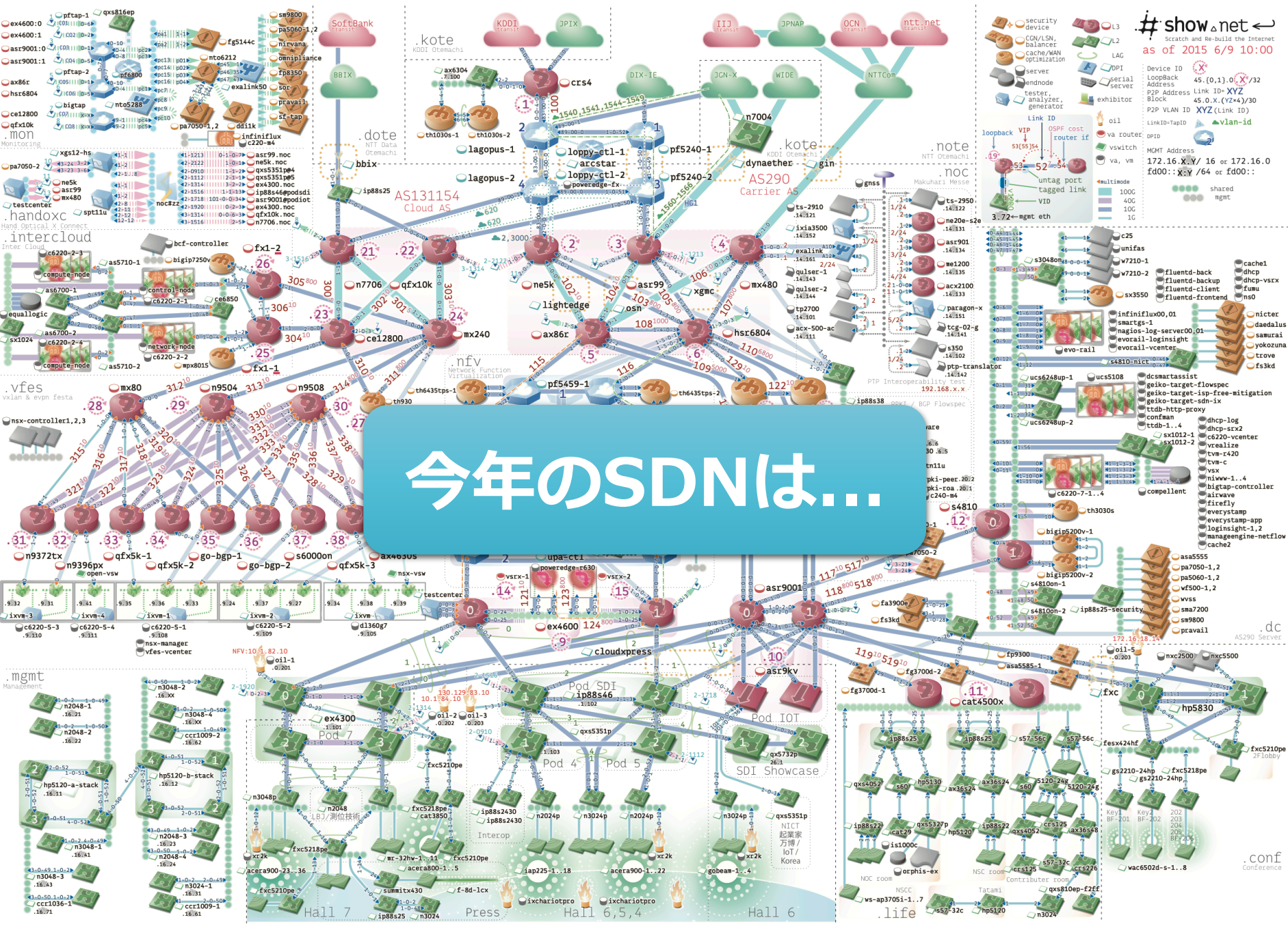
security device  
CON/LSN, balancer  
cache/WAN optimization  
LAG  
server  
endnode  
tester  
analyzer, generator  
exhibitor

oil  
va router: p2p  
vs switch  
va, vm  
MMT Address: 172.16.X.Y / 16 or 172.16.0  
fd00: X.Y / 64 or fd00:..

multimode  
100G  
40G  
1G

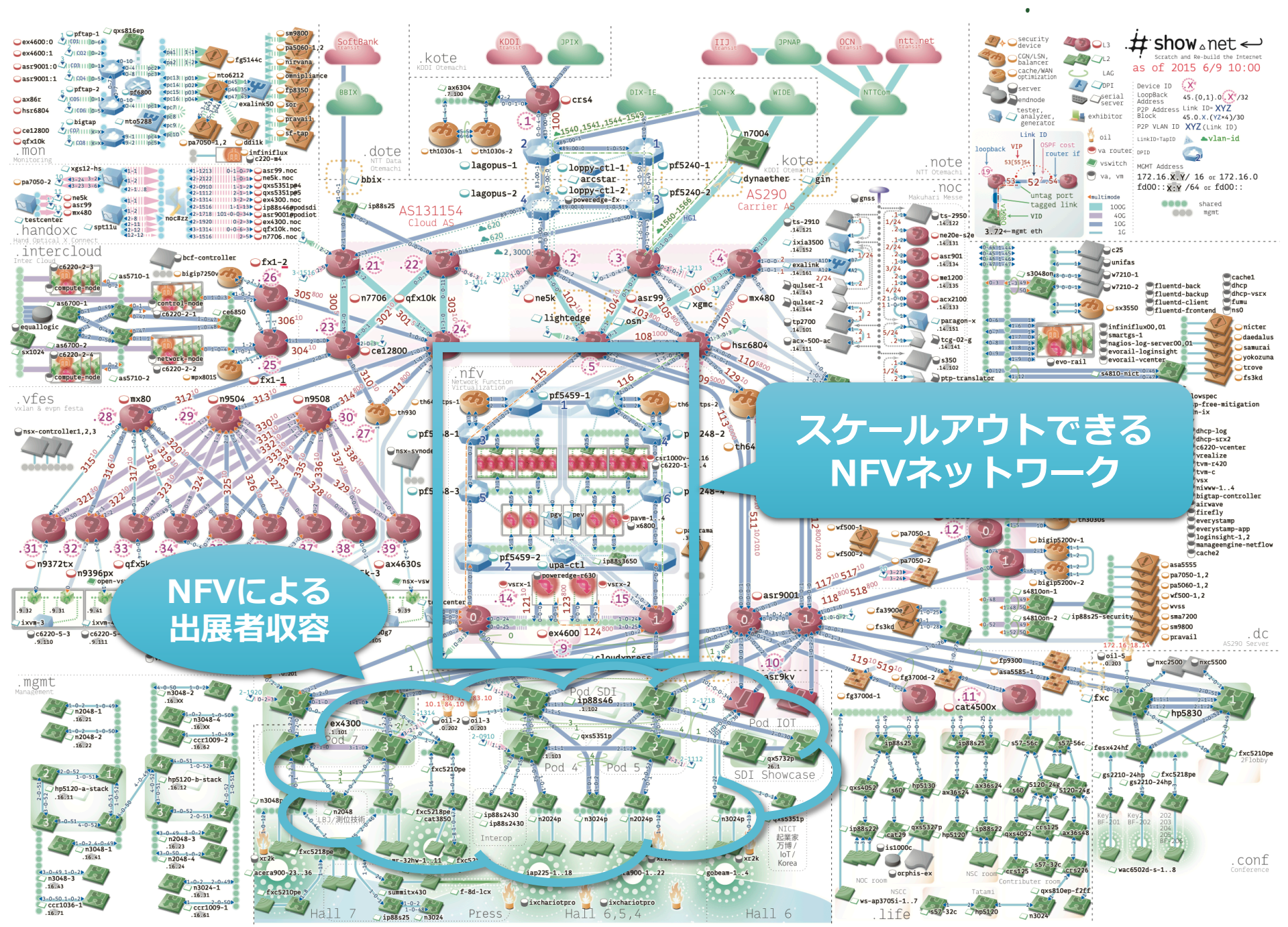
shared  
ngmt

loopback VIP  
OSPF cost  
router if  
tag port  
tagged link  
3.72~ngmt eth



# 今年のSDNは...

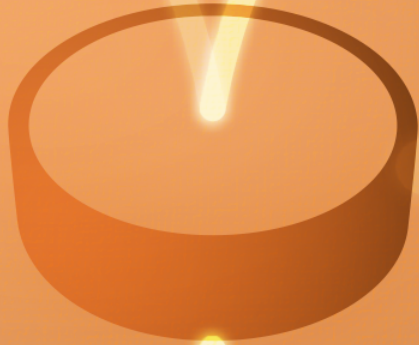




NFVによる  
出展者収容

スケールアウトできる  
NFWネットワーク





2015年の取り組みの前に。。



# SDN@ShowNetの歩み

---

- **ShowNetのSDNにおける仮想ネットワーク**
  - 2012年にOpenFlowの実機が登場
  - 2013年から出展者収容に導入
  - 2014年はNFVを構築
  - 毎年アップデート
    - その中で「できること」と「できないこと」
    - 課題の洗い出しと解決の積み重ね

# 2012年: OpenFlow実機の登場

- デモンストレーション
  - オンデマンドなパスの切り替え
  - 攻撃トラフィックの吸い込み
  - トラフィックの分離

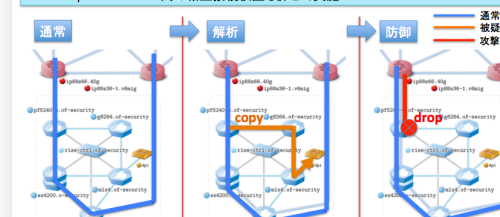


> Reborn to the Future

**OpenFlow Security** 最新: 本コントローラは RISE Controller

ShowNetのバックボーンネットワークで実際にOpenFlowネットワークを動かす  
今年セキュリティ機器との連携によるLIVEデモンストレーション

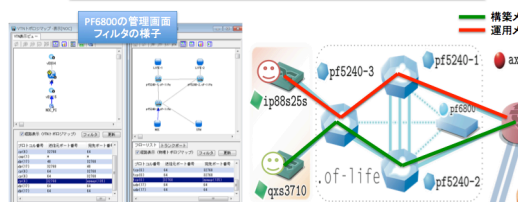
- OpenFlow auto protection
  - ✓ トラフィック解析システム(SAMURA)やDPIと連携して自動的に特定のフローを制御
  - ✓ OpenFlow Switch間の相互接続検証も併せて実施



**OpenFlow Life** 最新: 本コントローラは プログラムフローコントローラ

ShowNet構築メンバーが生活するスペース(LIFEネットワーク)での活用  
実際にOpenFlowネットワークで生活してみる

- OpenFlow Lifeでの生活ネットワークを提供
  - ✓ ユーザー毎に仮想ネットワークを構築し複数のポリシーを効果的に適用
    - アクセス制限
    - ネットワーク負荷分散
    - 脆弱性攻撃防御 等



構築メンバー (緑線)  
運用メンバー (赤線)


PF8000の管理画面  
フィルタの様子

Copyright © 2012 Interop Tokyo NOC Team. All rights reserved.

**OpenFlow Access** 最新: 本コントローラは NOX based Controller

来場者へOpenFlowを体感してもらおう  
アクセスコナーでOpenFlowを体感

- OpenFlowの代表的な機能であるパス制御を体感
  - ✓ パス変更要求と変更後のflow状態を表示
  - ✓ 快適なネットワークへ移行を実際に動画で体感



INTEROP #show.net **ポータル画面**

ShowNet OpenFlow Access

現在のアクセスコナーは  
通常ネットワークです

クリック

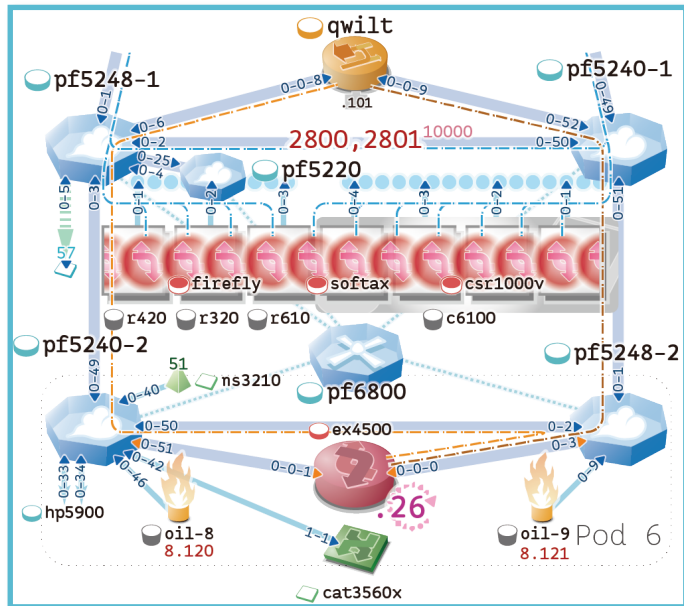
OpenFlowコントローラ-NOXbased Controller

# 2013年: 1出展者1仮想ルータ

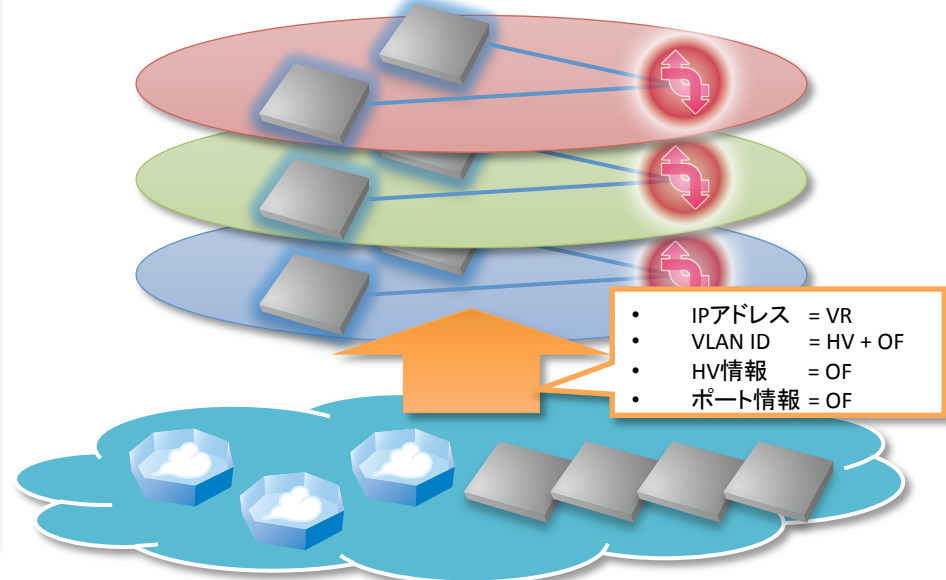
## • 初めて仮想ルータを導入

- 出展者ネットワークをテンプレート化
- 仮想ルータとOpenFlowの組み合わせでネットワークのプロビジョニングを実現

# show<sub>Δ</sub>net ←  
> Go to the next decade



ユーザの仮想ネットワークインスタンス群

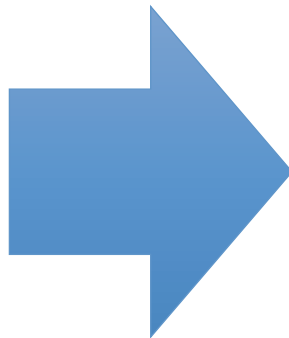




# 2013年：1出展者1仮想ルータ

## • 得られた知見と課題

- SDNによるプロビジョニングの可能性
  - ソフトウェアでL2/L3各パーツを構成する技術
- 論理構成の多重化と複雑化
  - ネットワークとサーバと仮想化の組み合わせ
  - 各技術への習熟が必要

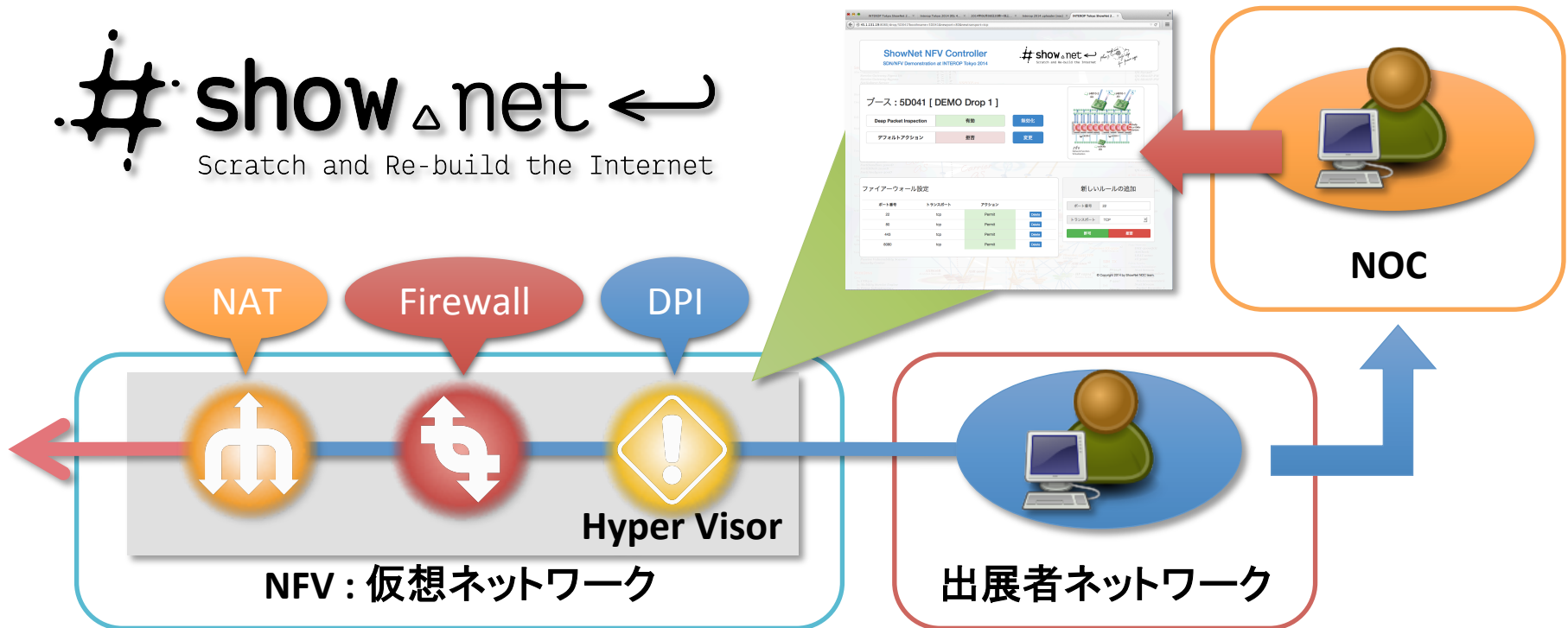


- 2013年以降、NFV/サービスチェーンニングが本格化
- 2013年1月 ETSI NFV
  - 2013年9月 IETF SFC Working Group
  - 2014年10月 OPNFV

# 2014年: 1出展者1仮想ネットワーク

## • Network Function Virtualization

- サービスチェーンの提供
- ASを越えた仮想ネットワーク接続の自動化



# 2014年: 1出展者1仮想ネットワーク

- **得られた知見と課題**

- 2度目のネットワークプロビジョニング
  - 30出展者ネットワークの構築に4,50分
- サーバ運用とネットワーク運用の融合
  - プログラミングしよう！
    - 様々なライブラリやAPIの登場
    - サーバ運用の知識や経験をネットワークへ

- **性能面の課題**

- 単純なVMとソフトウェアパッケージ転送の限界
- 規模性
  - 本当にスケールアウトするには



# 2014年: 1出展者1仮想ネットワーク

## • 得られた知見と課題

- 2度目のネットワークプロビジョニング
  - 30出展者ネットワークの構築に4,50分
- サーバ運用とネットワーク運用の融合
  - プログラミングしよう！
    - 様々なライブラリやAPIの登場
    - サーバ運用の知識や経験をネットワークへ

## • 性能面の課題

- 単純なVMとソフトウェアパッケージ転送の限界
- 規模性
  - 本当にスケールアウトするには



# SDN/NFV@ShowNet 2015

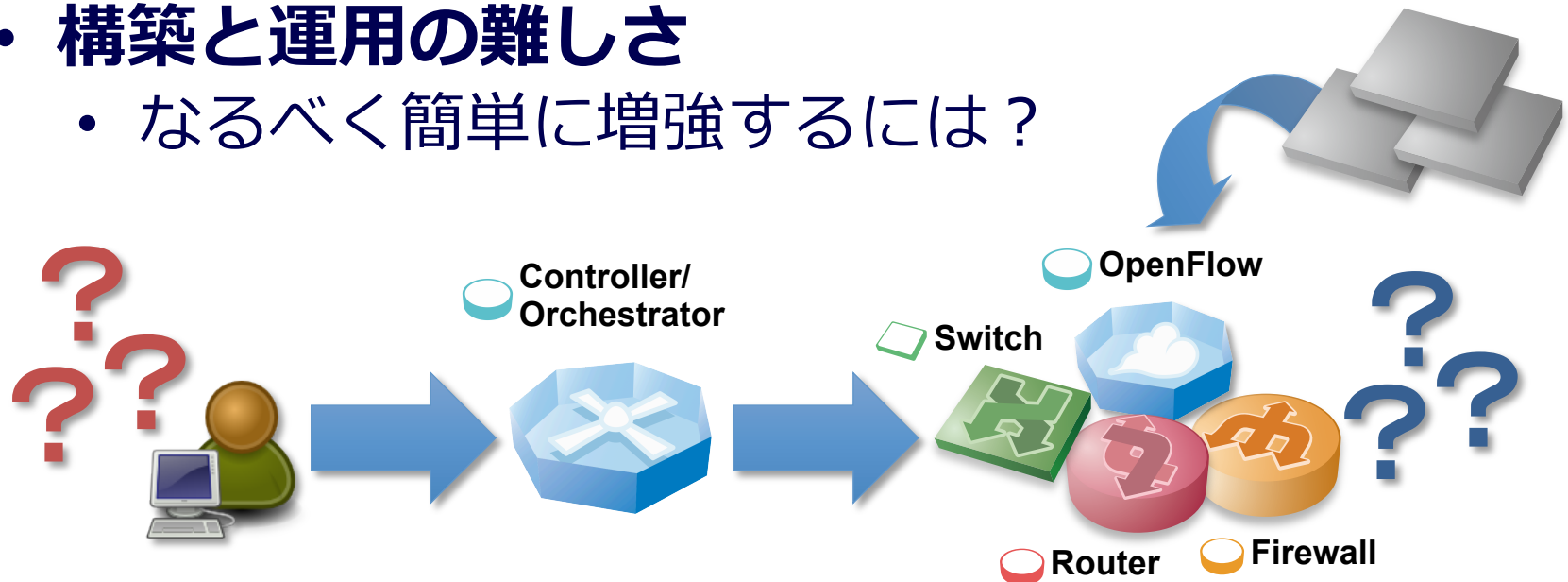
# 2015年のフォーカス

- **性能面の課題**

- ソフトウェアパケット転送のボトルネック
- どうやってサーバの追加で帯域や性能を強化していけばいいのか？

- **構築と運用の難しさ**

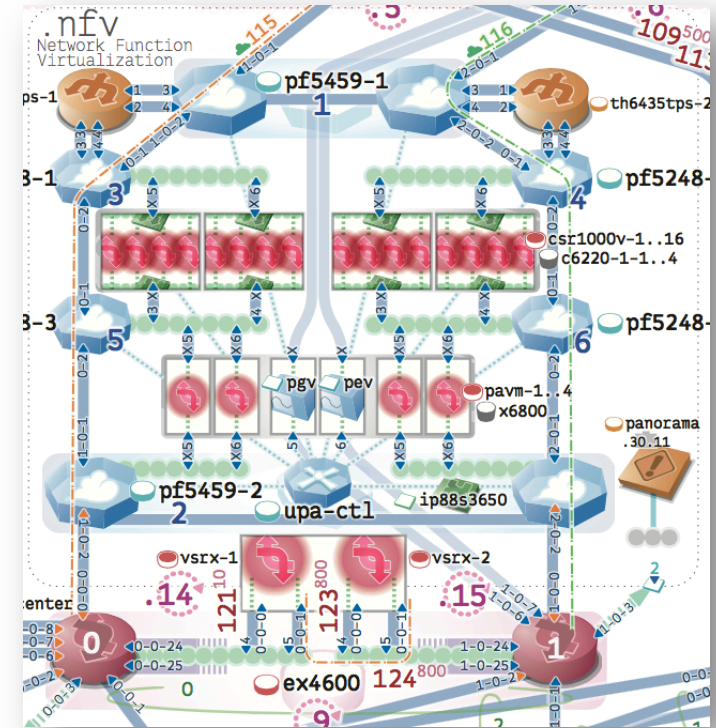
- なるべく簡単に増強するには？





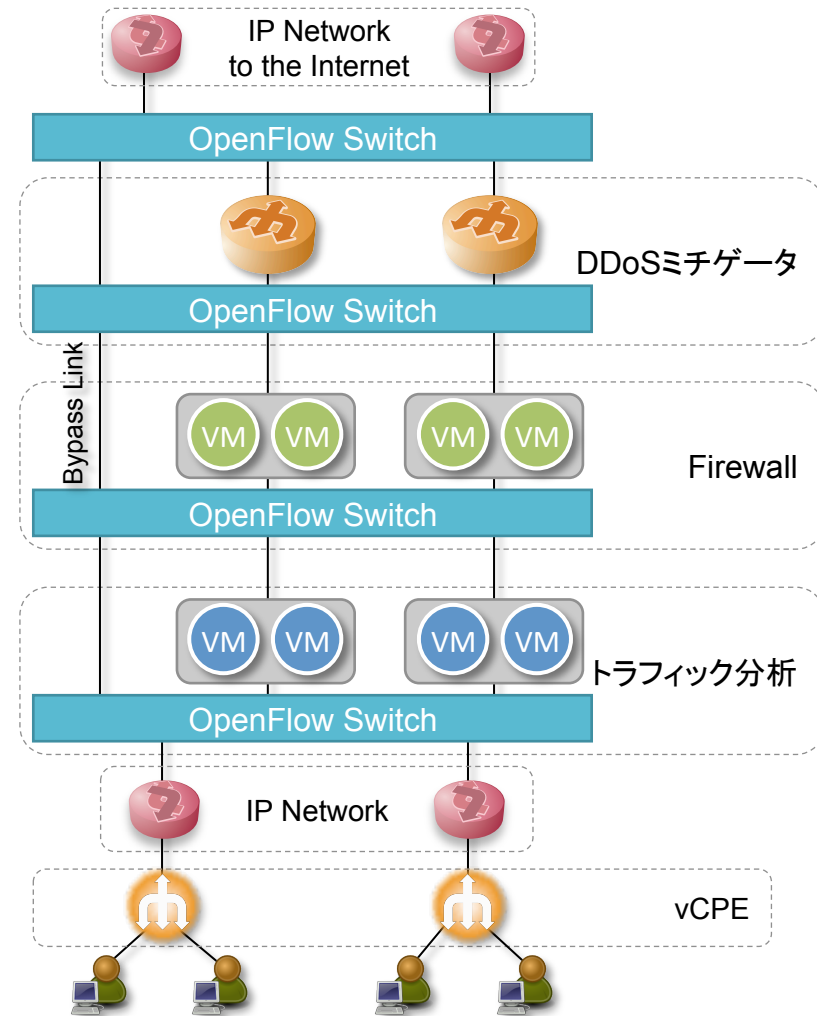
# SDN/NFV@ShowNet 2015

- スケールアウトするNFV
  - バックボーンにおいて、
  - サーバを追加すればするほど、ネットワーク全体の転送量が向上するNFV
- ポイント
  - トポロジに制約
  - 1. 同一機能を複数のバーチャルアプライアンスで構成
  - 2. OpenFlowを用いた複数VMへのロードバランス
  - 3. 様々な高速化手法の利用による性能の向上



# SDN/NFV@ShowNetの概要

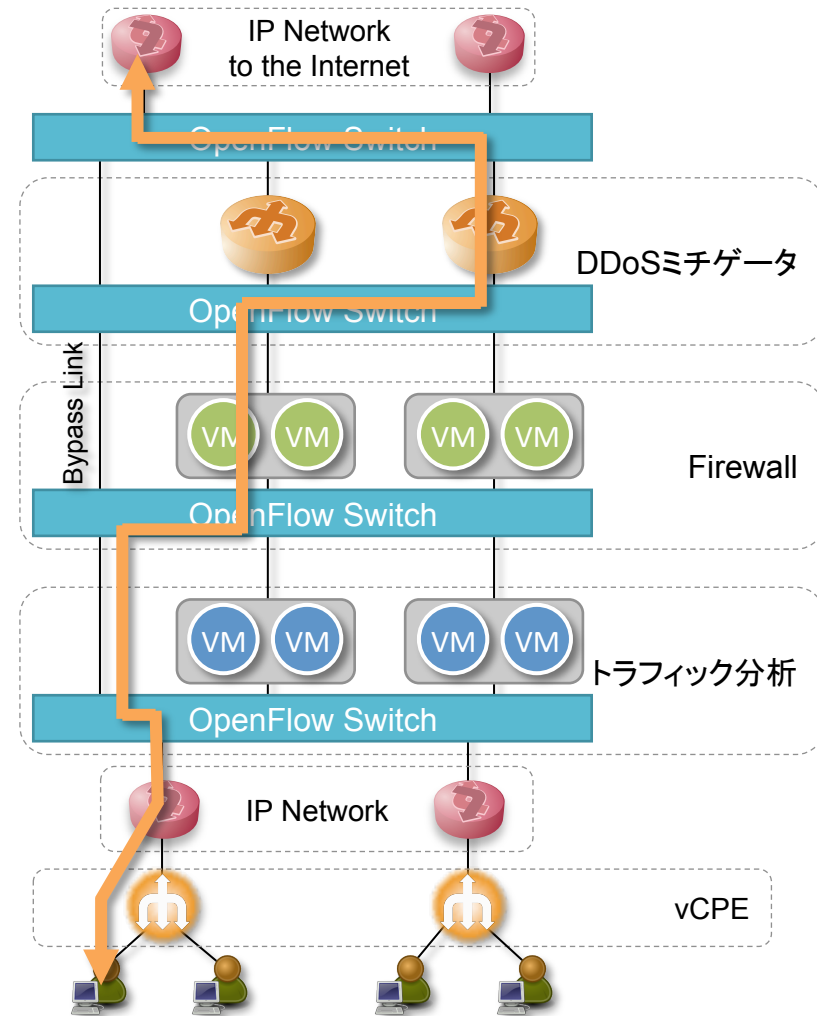
- **VNF Layering Model**
  - 1つのVNFを複数のVMで構成
  - 各VNFを層として積んで構成
- **CPEでパケットにマーキング**
  - 各サービスをToSのbitに埋め込む
- **OpenFlowでサービスの適用を判断**
  - 各パケットのToSの特定bitが1ならVNFへ、0ならバイパス
- **1つのサービスを構成する複数VMへトラフィックを分散**
  - 送信元アドレスをハッシュして転送するVMを決定
  - OpenFlowで決定したVMへOutput





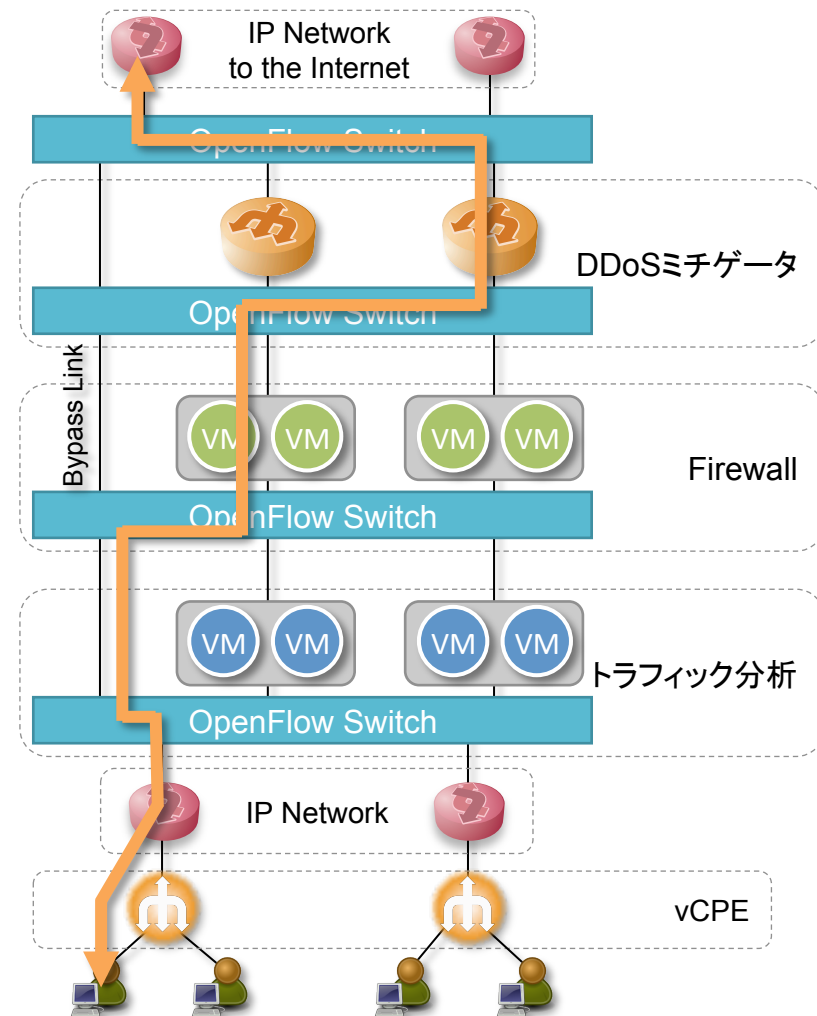
# SDN/NFV@ShowNetの概要

- **VNF Layering Model**
  - 1つのVNFを複数のVMで構成
  - 各VNFを層として積んで構成
- **CPEでパケットにマーキング**
  - 各サービスをToSのbitに埋め込む
- **OpenFlowでサービスの適用を判断**
  - 各パケットのToSの特定bitが1ならVNFへ、0ならバイパス
- **1つのサービスを構成する複数VMへトラフィックを分散**
  - 送信元アドレスをハッシュして転送するVMを決定
  - OpenFlowで決定したVMへOutput



# SDN/NFV@ShowNetの概要

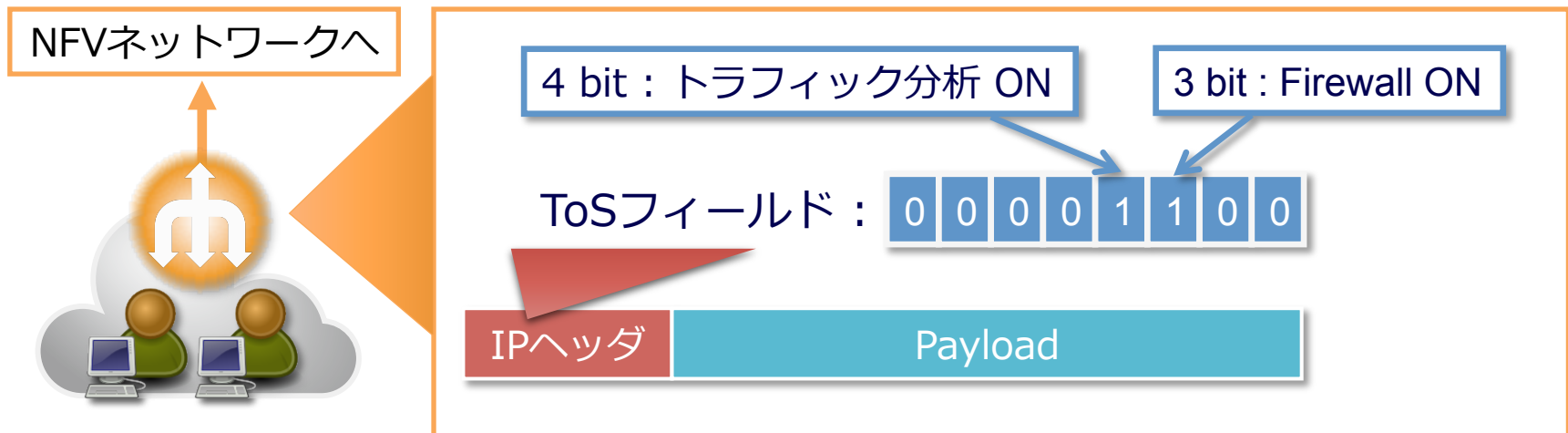
- **vCPEによる出展者收容**
  - NAT, サービス識別子の付与
  - Juniper Networks, vSRX
- **3つのVNFの層**
  - DPI/トラフィック分析
    - Paloalto Network, PA-VM
  - Firewall
    - Cisco Systems, CSR1000V
  - DDoS Mitigation
    - A10 Networks, Thunder 6435tps
- **OpenFlow Switch**
  - NEC PF5248, PF5459





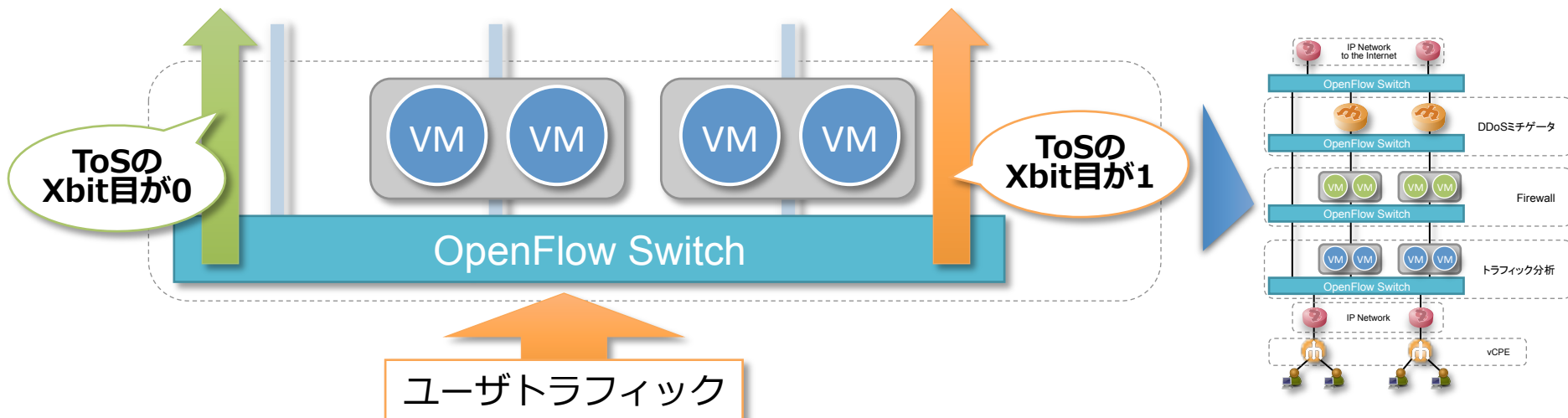
# サービスの識別

- **ToSフィールドを利用したサービス識別**
  - Type of Serviceフィールドの各bitをVNFで実現される各サービスにマッピング
  - vCPEでユーザ毎に付与するサービスのbitを1にする
  - その他のフィールドでも代用可能



# OpenFlowによるパスの切り替え

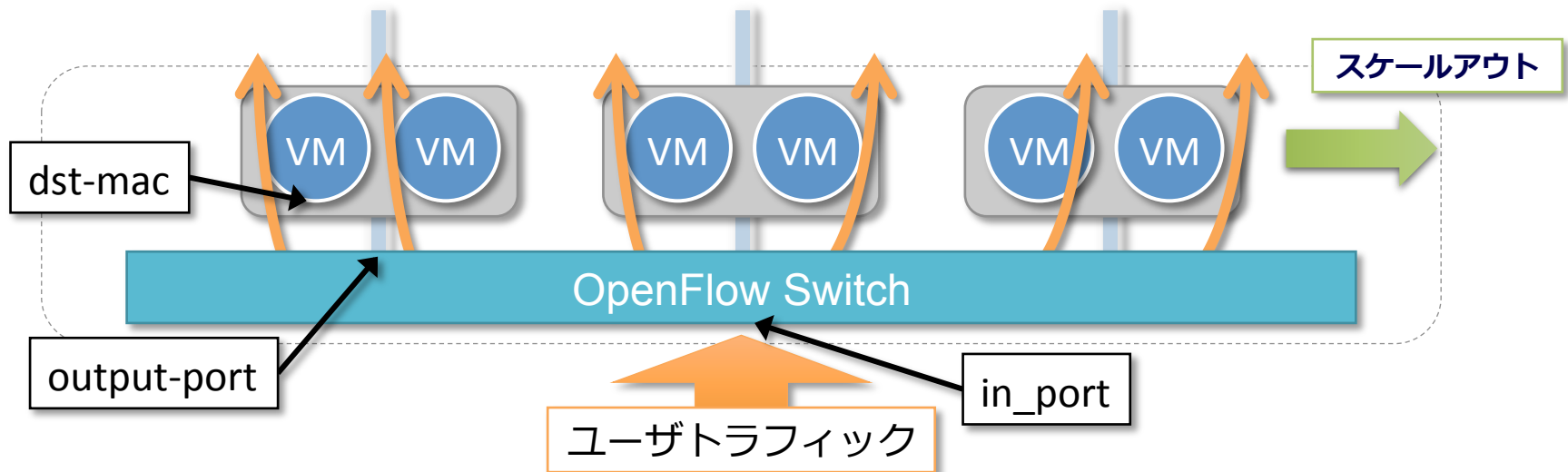
- 転送時にパケットをVNFに通すか判断
  - ToSと送信元アドレスをマッチ
  - ユーザ(アドレス)とそのサービス(ToS)をフローとして、フローごとにVNFを通すか制御
  - 各VNFの層は全く同じ手順で動作





# 複数VMへのロードバランス

- OpenFlowを使ってトラフィックを複数のVMに分散
  - 同じ設定のVM/サーバをコピー、追加投入、トラフィックをロードバランスしてスケールアウト
    - OF Match : [in\_port, ip\_src, ip\_tos]
    - OF Action : [set-dl-dst, output-port] ← portはハッシュで決定
  - どうしても変更が必要な部分には様々な自動化技術で対応



# Incremental Deployment

- **VM→HV→OpenFlow Switch**

- VMの追加: VMごとの要件、CPU、メモリ、NIC
  - APIを利用したオーケストレーション

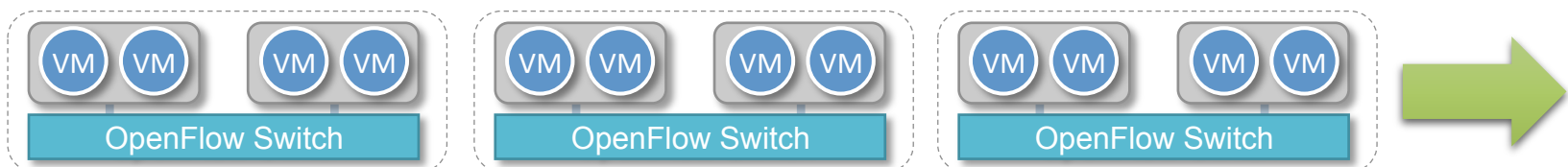


VM単位、サーバ単位、  
ネットワーク単位での、  
水平スケールアウト

- HVの追加: スイッチのポート数
  - 様々なサーバ設定自動化手法



- OpenFlow Switch単位で追加






# ShowNet NFV Control Panel ver 2015

## ShowNet NFV Controller

SDN/NFV Demonstration at INTEROP Tokyo 2015

ShowNet 2015 NFVコントロールパネル。ユーザ出展者ごとに、適用するVNFを選択することができます。



Scratch and Re-build the Internet

*Balance.*

### ユーザごとのサービスチェーン状態

ユーザ	DDoSミティゲーション	標準アクセスコントロール	トラフィック分析
NFV User 1 (Hall 5)	OFF	ON	OFF
NFV User 3 (Hall 7)	OFF	ON	OFF
NFV User 4 (Hall 5)	OFF	ON	ON
NFV User 5 (Hall 4)	OFF	ON	ON
NFV User 6 (Hall 6)	OFF	ON	OFF
NFV User 7 (Hall 5)	OFF	ON	ON

### サービスチェーンの切り替え

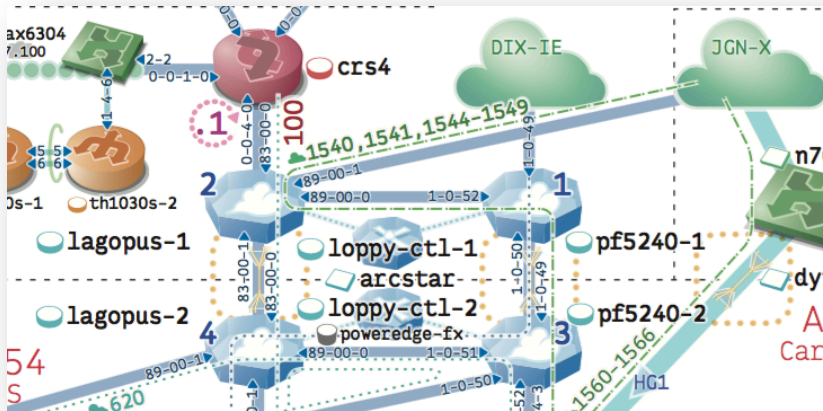
ユーザ	DDoSミティゲーション	標準アクセスコントロール	トラフィック分析
NFV User 1 (Hall 5)	<input checked="" type="checkbox"/> ON <input type="checkbox"/> OFF	<input type="checkbox"/> ON <input checked="" type="checkbox"/> OFF	<input checked="" type="checkbox"/> ON <input type="checkbox"/> OFF
NFV User 3 (Hall 7)	<input checked="" type="checkbox"/> ON <input type="checkbox"/> OFF	<input type="checkbox"/> ON <input checked="" type="checkbox"/> OFF	<input checked="" type="checkbox"/> ON <input type="checkbox"/> OFF
NFV User 4 (Hall 5)	<input checked="" type="checkbox"/> ON <input type="checkbox"/> OFF	<input type="checkbox"/> ON <input checked="" type="checkbox"/> OFF	<input type="checkbox"/> ON <input checked="" type="checkbox"/> OFF
NFV User 5 (Hall 4)	<input checked="" type="checkbox"/> ON <input type="checkbox"/> OFF	<input type="checkbox"/> ON <input checked="" type="checkbox"/> OFF	<input type="checkbox"/> ON <input checked="" type="checkbox"/> OFF
NFV User 6 (Hall 6)	<input checked="" type="checkbox"/> ON <input type="checkbox"/> OFF	<input type="checkbox"/> ON <input checked="" type="checkbox"/> OFF	<input checked="" type="checkbox"/> ON <input type="checkbox"/> OFF
NFV User 7 (Hall 5)	<input checked="" type="checkbox"/> ON <input type="checkbox"/> OFF	<input type="checkbox"/> ON <input checked="" type="checkbox"/> OFF	<input type="checkbox"/> ON <input checked="" type="checkbox"/> OFF

# パケット転送性能の挑戦

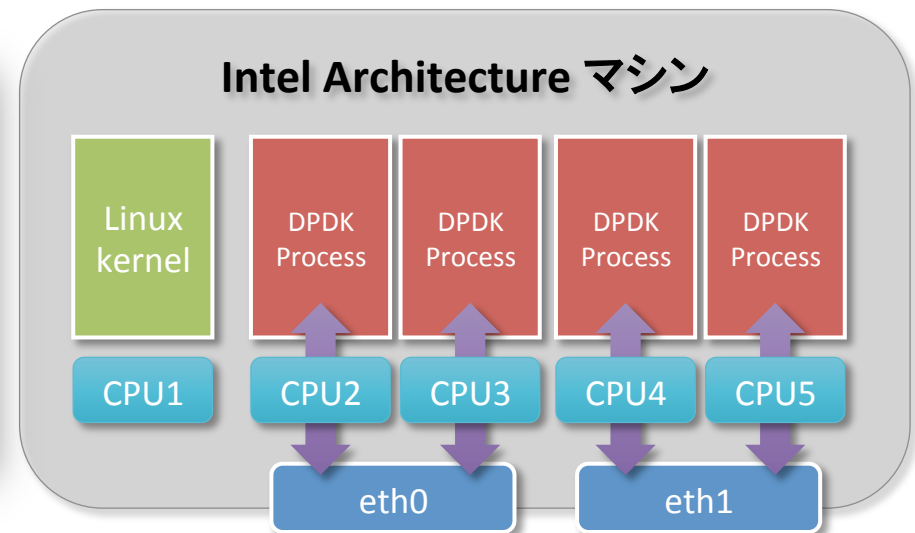
- できるだけ、  
**“ソフトウェアによるパケット処理は減らす”**
  - 現状ではハードウェア処理の方が当然早い
  - なるべく仮想スイッチなどは使わない
  - または、高速パケットI/O技術の利用
- **Intel DPDK**
- **PCI Pass-through**
- **Single Root I/O Virtualization**

# Intel DPKK

- **Intel Data Plane Development Kit**
  - Intelによる高速パケットI/O技術の1つ
  - DPKK自体は専用のドライバとライブラリ(SDK)
  - データプレーンの処理に最適化
  - Open Source Software, BSDライセンス



SDN-IXではLagopusも

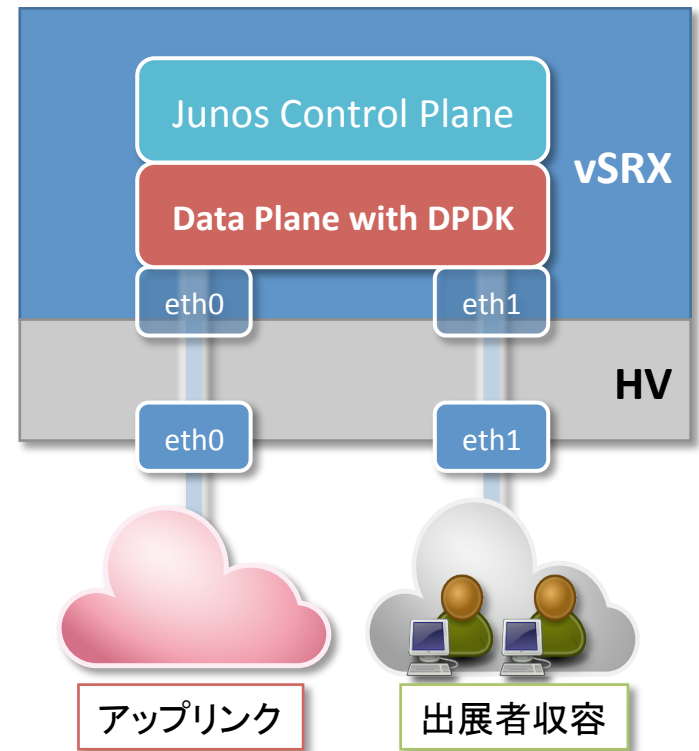




# DPDKを採用した仮想ルータ

## • Data PlaneをDPDKで実装した仮想ルータ

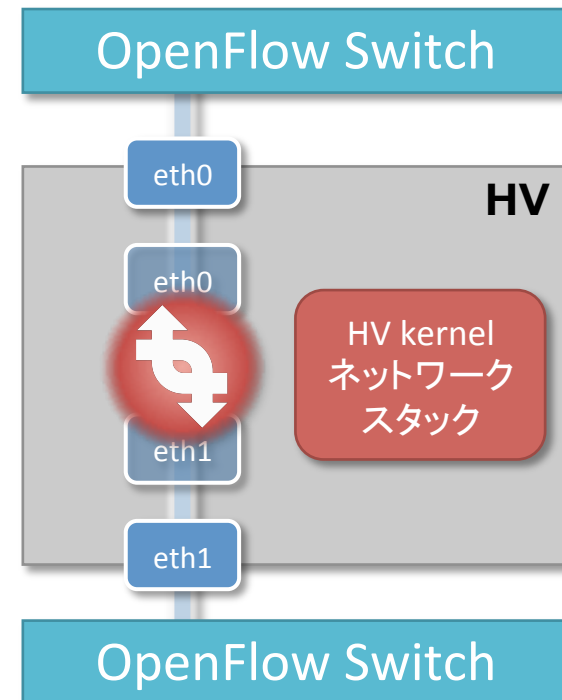
- Juniper Networks, vSRX
  - Control PlaneはJunos, Data PlaneにDPDK利用
- vCPEとして出展者セグメントを直接収容
- サービスメニューに応じ、ユーザごとにToS書き換え
  - API: py-junos-eznc



JUNIPER  
NETWORKS

# PCI Pass-through

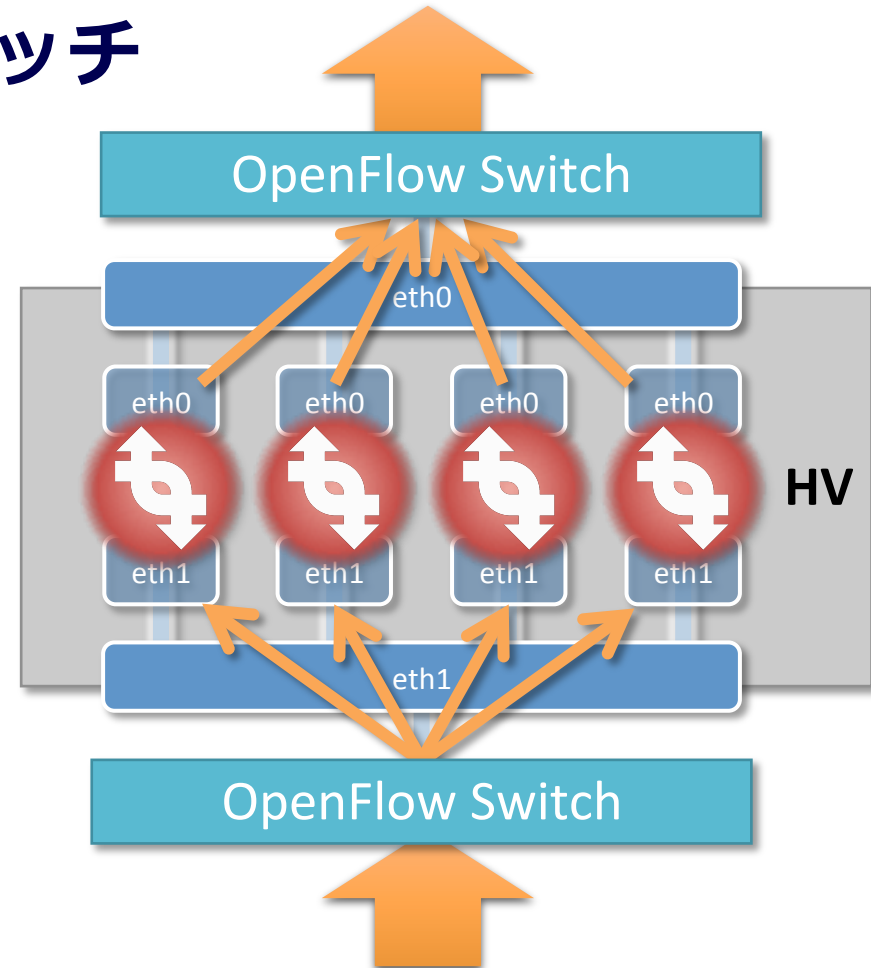
- **HVのNICをVMにアタッチ**
  - 最もレガシーな手法
  - HVのネットワークスタックを経由しない
  - ソフトウェアパケット処理はVA内のみ
  - Paloalto Networks, VM-300にて利用
    - Panoramaによる集中管理



# Single Root I/O Virtualization

## 各仮想NICを各VMにアタッチ

- NICの仮想化
- OpenFlowでトラフィックを各VMへ分散(dl-dst)
- VMを各CPUに割り当てることで複数のコアを有効に使ってパケット転送が可能
- CPUのコア数/メモリ量/NICのキュー数だけVM設置
- Cisco Systems, CSR1000Vで利用

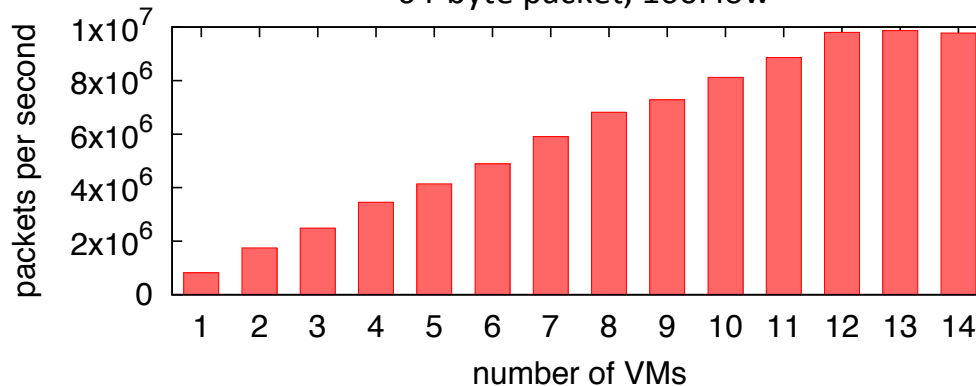




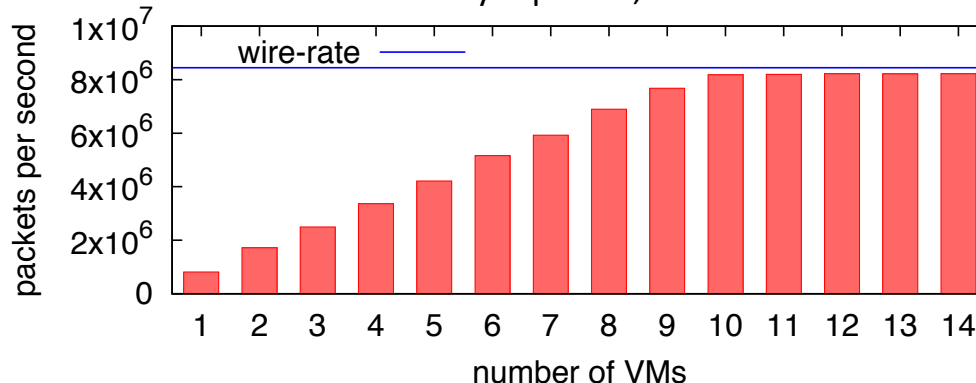
# SR-IOV+OpenFlow

## • 1HV内複数VMへのトラフィック分散(予備)実験

64-byte packet, 100Flow



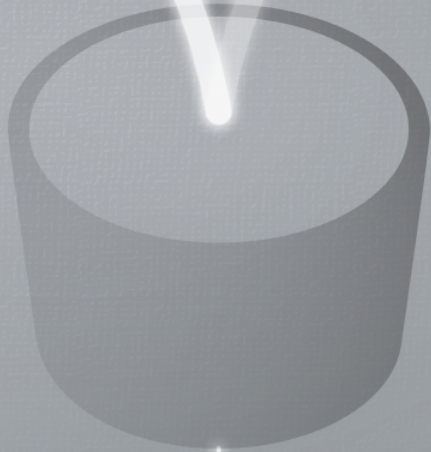
128-byte packet, 100Flow



- VMを足せば足すだけ性能が向上
  - VMにはVyOSを利用
  - L3ルーティング
  - HVはKVM(Default)
- 高速パケットI/O無し
- 1台の汎用サーバでも10VMあれば128-byteパケットでwire-rate

# やってみて解ったこと

- **仮想ネットワークのプロビジョニング**
  - 3回目、もう普通にできることです。APIの整備も進みました
- **型にはめよう！**
  - 自由度の高さはもちろん重要
  - 制約を加え、要所要所を定型化、モジュール化することで、構築と運用を楽にできる(もちろんSDNに限らず)
- **スケールアウトの大事さ**
  - サーバと違い、ネットワークは足せばいいってものではない
  - SDN的なアプローチは有効
- **そして作りこみも大切**
  - 監視手法のアイデアもあったが、実装が間に合わず。。
  - ネットワークのプログラミング/プログラマ重要！

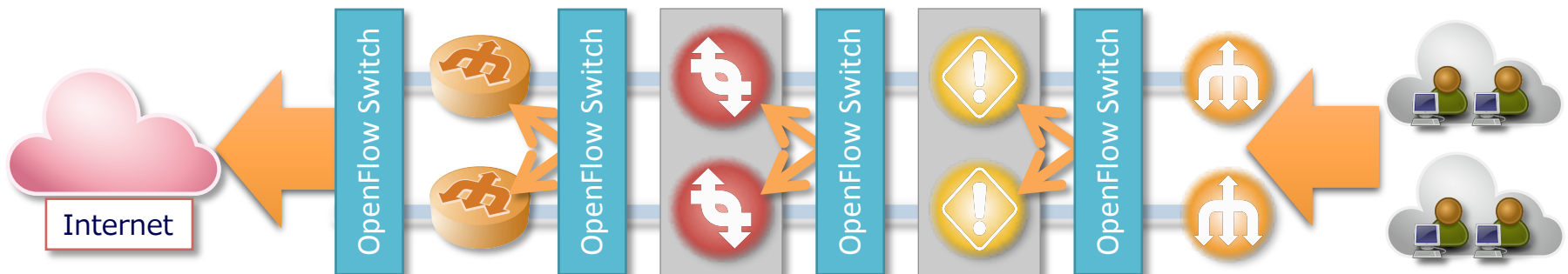


# まとめ



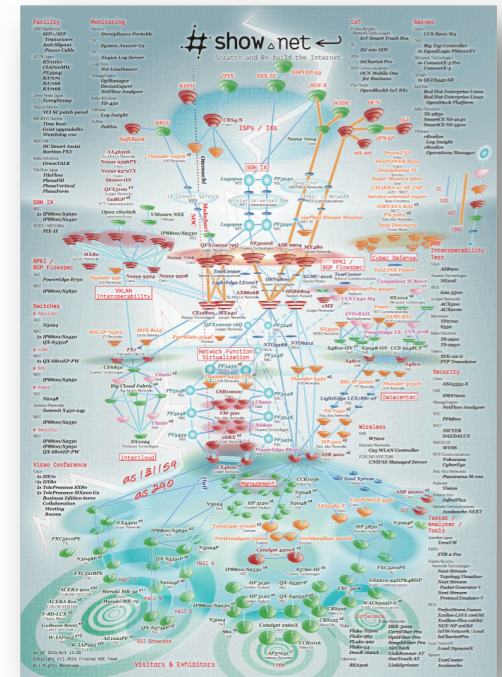
# ShowNet 2015におけるSDN/NFV

- **スケールアウトするNFV**
  - 仮想ルータを横に並べてスケールアウト
  - OpenFlowを用いたトラフィックの分散
  - サーバを追加するだけでスループットを向上
- **ソフトウェアによるパケット転送**
  - ハードウェアには数で対抗
  - Intel DPDKをはじめ、ソフトウェアも高速化



# INTEROP Tokyo 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1つのモデル、実現手法としてのSDN/NFV
- **Live Network**
  - ShowNetは生きたネットワーク
  - 実際に動いて使えるSDNとNFV
- **相互接続性**
  - 多種多様な機器と様々な技術
  - それぞれの特徴と活用手法
  - そしてフィードバック





# show net ←

Scratch and Re-build the Internet

*simplifying* *reliability*  
*flexibility* *phased:*  
Ultimate  
Balance.