愉快なソフトウェアネットワーキング

東京大学 情報基盤センター

関谷 勇司

自己紹介

所属 : 東京大学 情報基盤センター
- 研究分野
  - クラウド要素技術
  - SDN / NFV
  - 次世代ネットワークアーキテクチャ (NSP Consortium)
  - サイバーセキュリティ
- 学内基幹ネットワーク設計・運用

Interop Tokyo ShowNet NOC メンバー
- 2000年より
- 2011年より 2017年まで NOC 統括者

AITAC (一般社団法人 高度ITアーキテクト育成協議会)
- https://aitac.jp/

本日のお話

ソフトウェアネットワーキング
- SDN + NFV = ？

Interop Tokyoでのチャレンジは続いています
- 本日も上野NOCメンバーによるお話があります
- いつから始めてたんだろう？

自身のチャレンジ
- PIX-IE (SDN-IX)
- 学内 SD-WAN

ソフトウェアネットワーキングとは

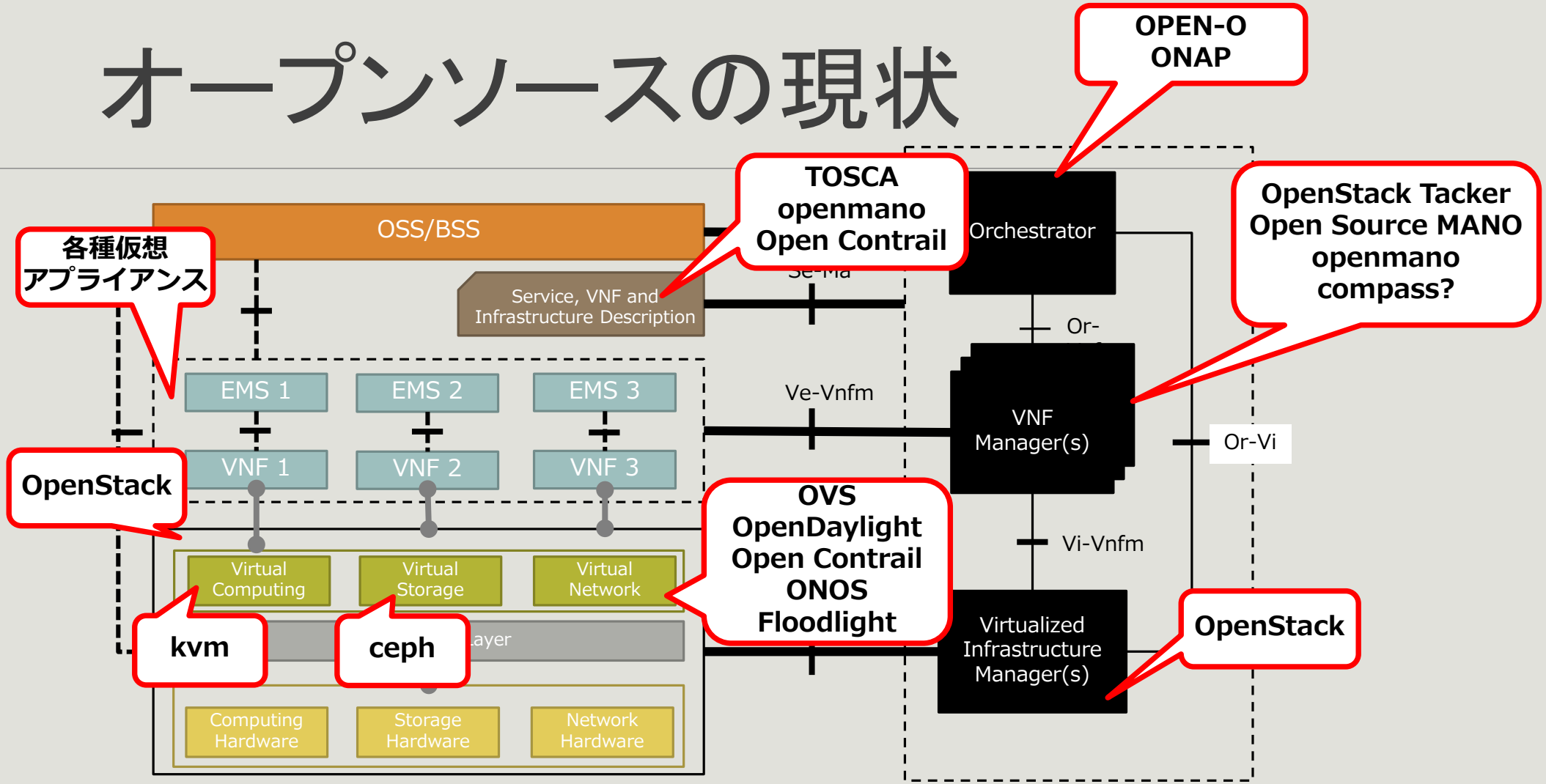
ソフトウェア資源を最大限に活かした
ネットワークシステムを構築すること

- SDN のみならず周辺のシステムも含めソフトウェア資源を有効利用したシステム構築
- Network Softwarization (5G)

ソフトウェア化の意味

- ✕：ハードウェアの処理をソフトウェア化すること
- △：既存のワークフローを自動化すること
- ○：ソフトウェア化による利点を享受すること

オープンソースの現状


OPEN-O
ONAP
TOSCA
openmano
Open Contrail
OpenStack Tacker
Open Source MANO
openmano
compass?
各種仮想
アプライアンス
OSS/BSS
Service, VNF and
Infrastructure Description
Orchestrator
Or-
EMS 1
EMS 2
EMS 3
Ve-Vnfm
VNF 1
VNF 2
VNF 3
VNF
Manager(s)
Or-Vi
OpenStack
OVS
OpenDaylight
Open Contrail
ONOS
Floodlight
Virtual
Computing
Virtual
Storage
Virtual
Network
Vi-Vnfm
kvm
ceph
Virtualized
Infrastructure
Manager(s)
OpenStack
Computing
Hardware
Storage
Hardware
Network
Hardware

ソフトウェア化の利点とは

ユーザや社会の要求に迅速に対応

- **[柔軟性] : 仮想化を用いた抽象化**
- **[即時性] : ソフトウェア展開による迅速なサービス構成**
- **[規模性] : 規模拡張性**

仮想化による抽象化
=> 統一化された管理手法の利用 (OPEX 削減)

ソフトウェアによるサービス構成
=>共通化されたハードウェアの利用と資源の有効利用 (CAPEX 削減)

愉快な。。。の意味

ソフトウェア化にてシステムを内製することによる
利点と欠点

- 利点
  - やっぱり楽しいですよ
  - Requirements に応じた細かなシステムが構築できます
  - アイディアを即時にサービスとして実現することができます
- 欠点
  - 壊れるときはとことん壊れます
  - 構築・管理できる人材が必要です

いくつか私自身が関わった事例を紹介させて頂きます

Interop Tokyo という舞台

2013年 : SDN により「スライス」という概念を構築
- 仮想ルータと OpenFlow を用いて複数ネットワークを構築

2014年 : NaaS の概念を検証
- 接続性提供事業者とネットワークサービス提供者を分割する実験
- SDN-IX (PIX-IE) が登場

2015年 : スケールアウト可能なシステム
- SDN + NFV を用いたサービス提供 (独自コントローラ + ポータル)
- FallFlow モデルの提唱

2016年 : 大規模サービスチェイニングの検証
- BGP FlowSpec + OpenFlow を用いたサービスチェイニングの実践

2017年 : 今日の上野さんの発表におまかせしましょう
- いかに辛かったのか

Interop Tokyo という舞台

2013年 : SDN により「スライス」という概念を構築
- 仮想ルータと OpenFlow を用いて複数ネットワークを構築

2014年 : NaaS の概念を検証
- 接続性提供事業者とネットワークサービス提供者を分割する実験
- SDN-IX (PIX-IE) が登場

2015年 : スケールアウト可能なシステム
- SDN + NFV を用いたサービス提供 (独自コントローラ + ポータル)
- FallFlow モデルの提唱

2016年 : 大規模サービスチェイニングの検証
- BGP FlowSpec + OpenFlow を用いたサービスチェイニングの実践

2017年 : 今日の上野さんの発表におまかせしましょう
- いかに辛かったのか

SDN-IX の実現

2013年の SDN Japan でパネルセッションを開催
- SDN 技術を IX に入れたらどんな利点があるだろう

OpenFlow を使って IX スイッチを作る
- IX スイッチに求められる機能は？
- Ethernet のフル機能は必要ない
- いわばパス交換機能のみ

作って運用してみよう

Programmable Internet Exchange (PIX-IE)

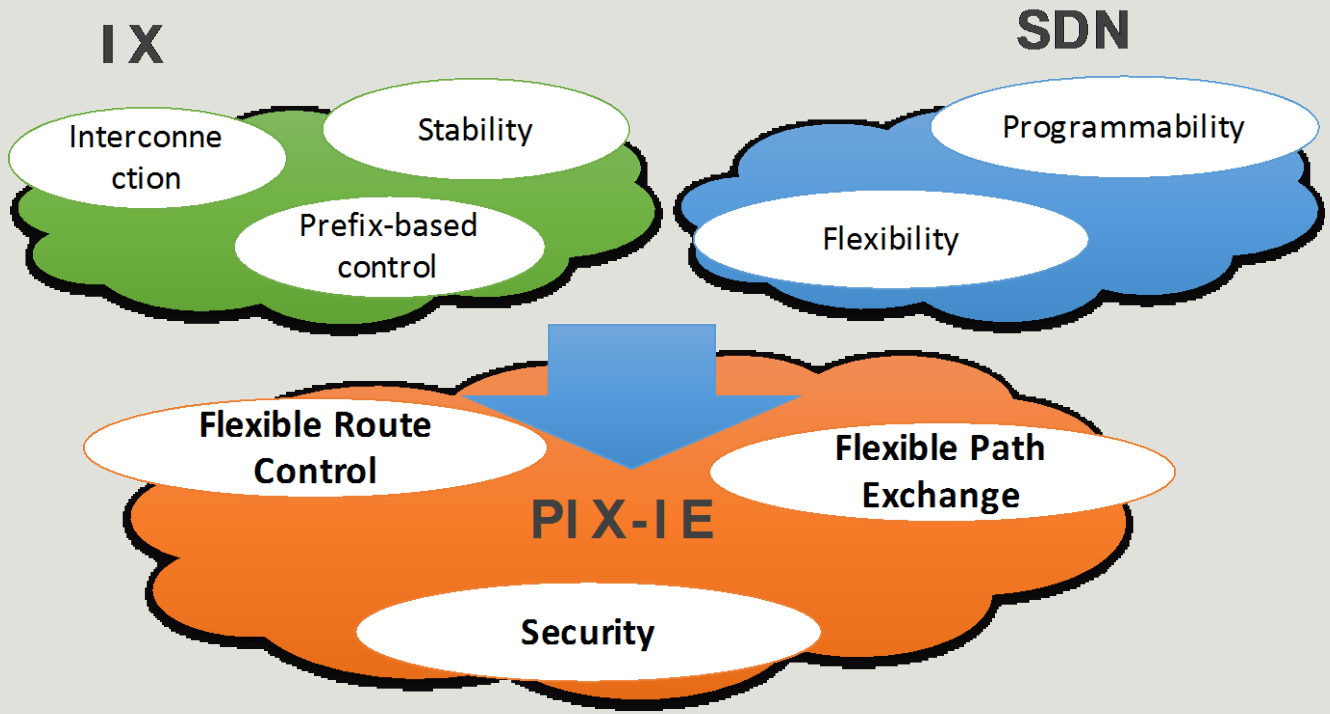
Providing Programmable Functions for Customers on IXP

- Joint Research with NICT/JGN

Granular Route Control

Flexible Path Exchange

Security

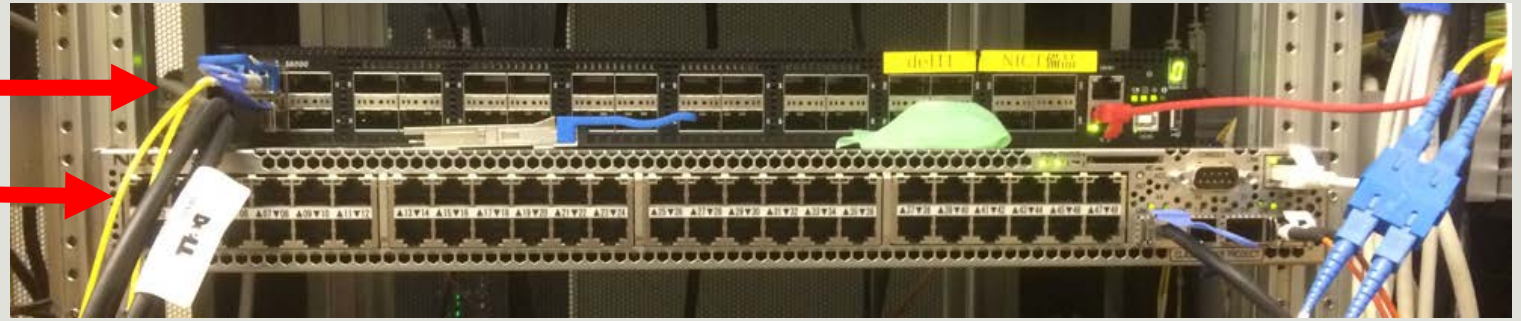


Hardware – Phase 1

Otemachi-1 Site

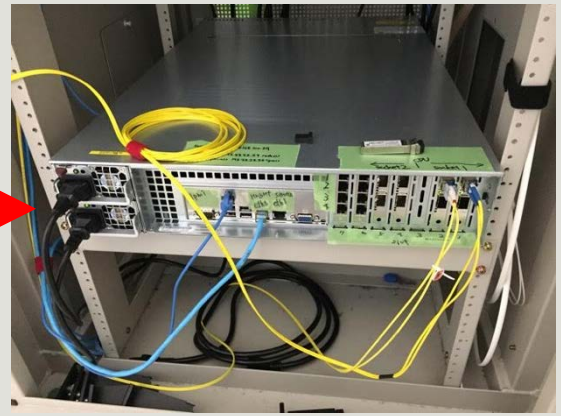
S6000-ON

PF5240



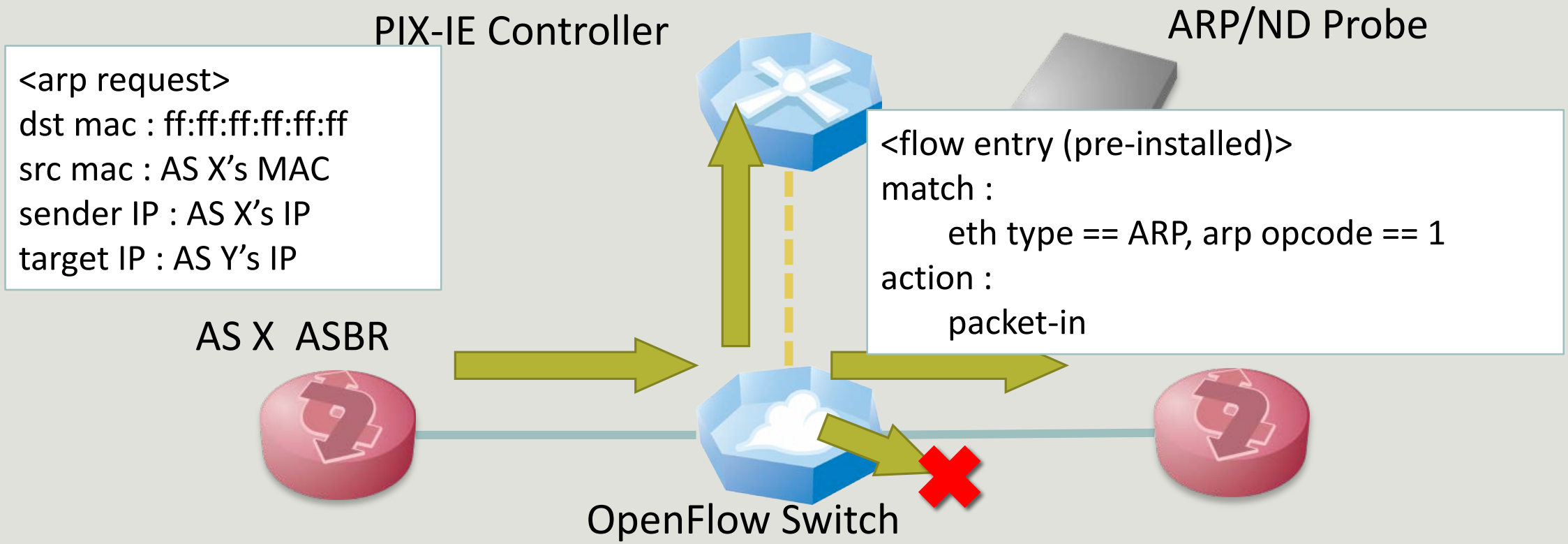
Otemachi-2 Site

Lagopus Software Switch



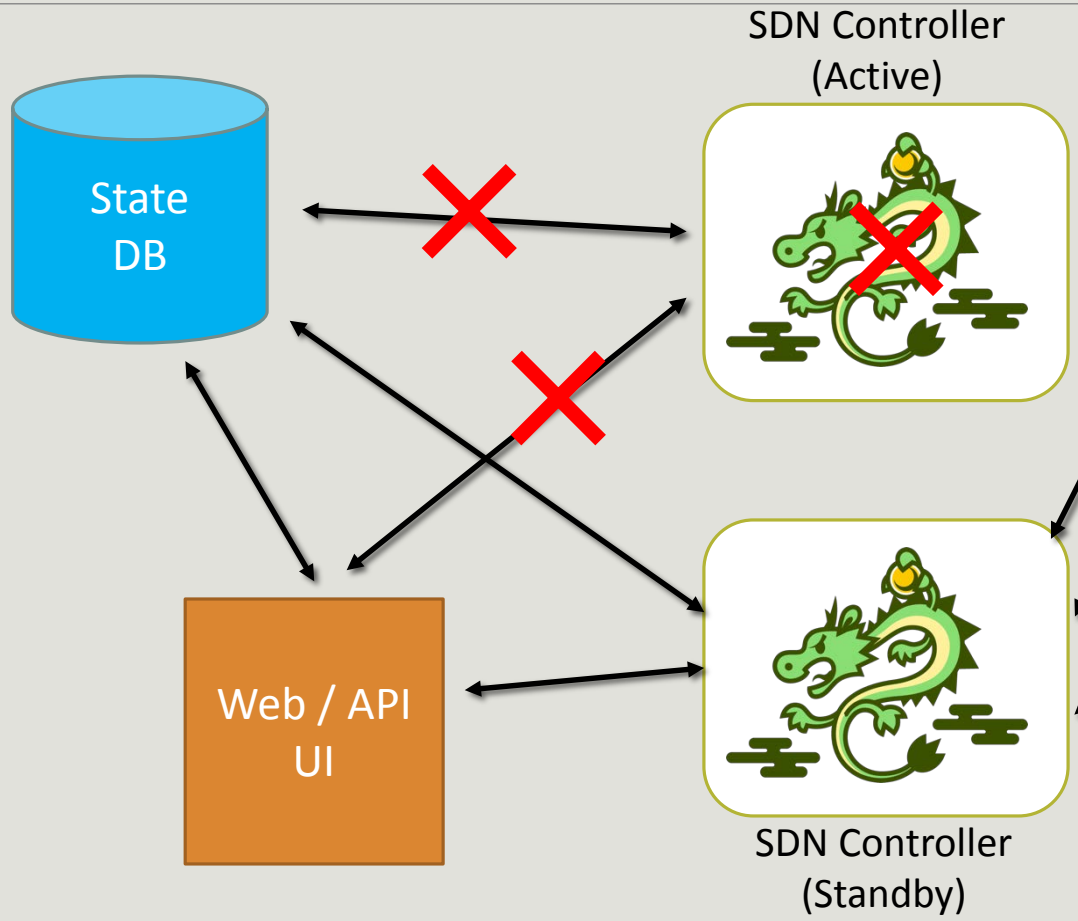
When make a peering between AS X and AS Y

BUM traffic localization



Forward the ARP request packet to the controller following the flow entry

Controller Architecture (HA)



コントローラとのコネクションが切れても
ルールを保持するよう設定

パス交換に必要なルール情報や ARP/NDP に必要なエントリ情報はすべてバックエンドの DB に格納し SDN コントローラはその Proxy (インタプリタ) として動作するよう構築

Phase-2：FAUCET の利用

同じく OpenFlow で IX を運用している組織

- Toulouse IX by TOUSIX Project

この IX と協力体制

- FAUCET SDN Controller (http://faucet.nz/)
- 共通のプラットフォームでアーキテクチャ構成

FAUCET SDN Controller

OpenFlow スイッチでファブリックを構成できる
オープンソースソフトウェア

- 異機種の OpenFlow スイッチ同士でも可能
- OpenFlow 1.3 + MultiTable のサポートが必要

Faucet Conference and Plugfest

- October 2017, Berkeley, CA, USA

FAUCET 検証済み OpenFlow スイッチ

https://github.com/faucetsdn/faucet/tree/master/docs/vendors より

- Allied-Telesis
- HPE
- Lagopus
- NorthBound Networks
- noviflow
- ovs
- 某社さんのスイッチも検証済み

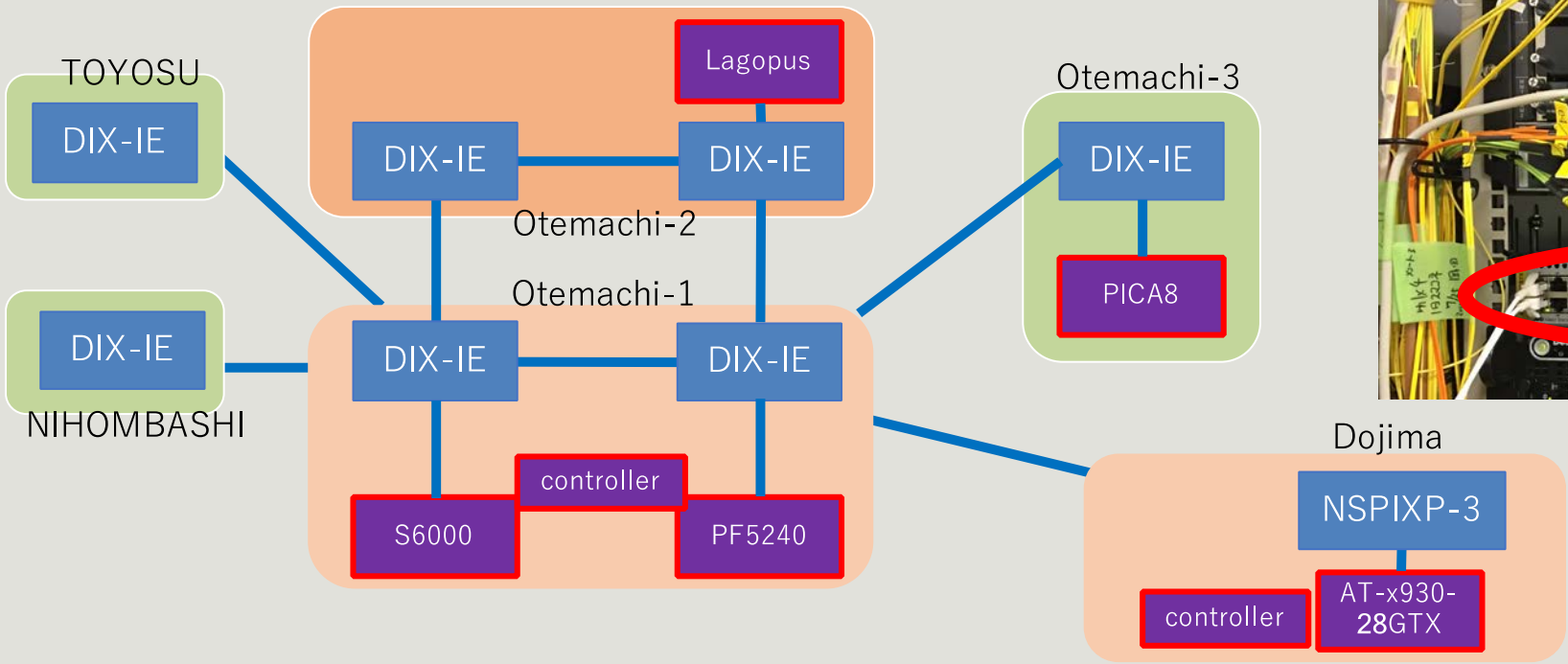
機能

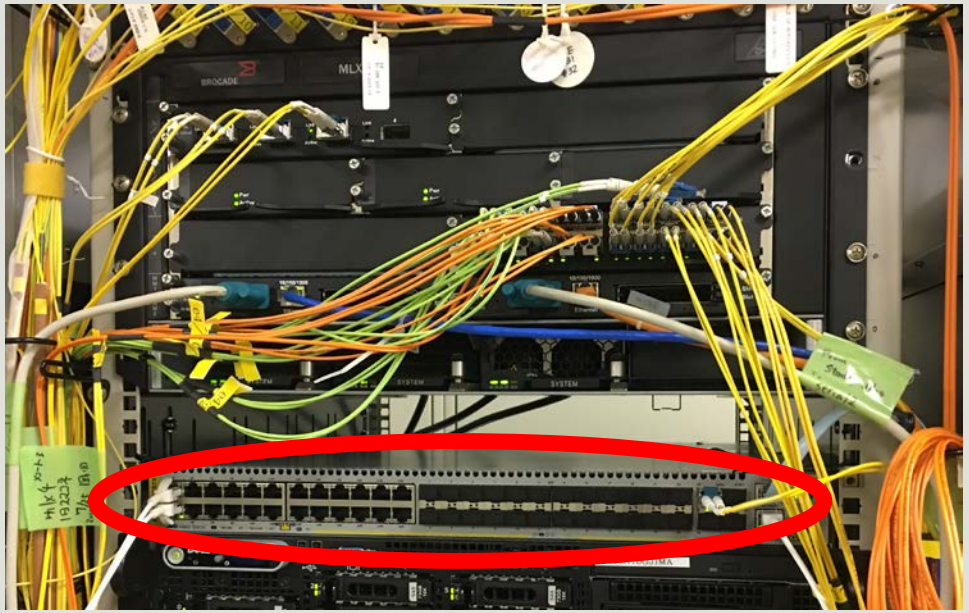
Features

- VLANs
- IPv4 and IPv6 support
- IPv6 neighbor solicitation and router advertisement support
- Static and BGP routing
- Flexible port and VLAN based Access Control Lists
- Port mirroring
- Fast configuration reloads
- Vendor neutral stacking of Openflow switches
- Policy based forwarding to offload processing to external systems (Eg 802.1x via hostapd)
- Configurable learning: Control unicast flooding by port and by VLAN
- Dataplane for NFV - Offload functions such as DHCP, NTP, Firewall, and IDS
- CouchDB support for storing flows from switches to enable north bound applications
- Influx support for time-series OpenFlow port statistics
- Prometheus integration for monitoring and instrumentation of FAUCET
- Grafana based dashboards for monitoring
- Comprehensive test suite – tests for all features that can be run against mininet (software switching) and on hardware
- Code: Python based, easy readability (PEP8 style)

新たなアーキテクチャによる SDN-IX

PIX-IE 大阪





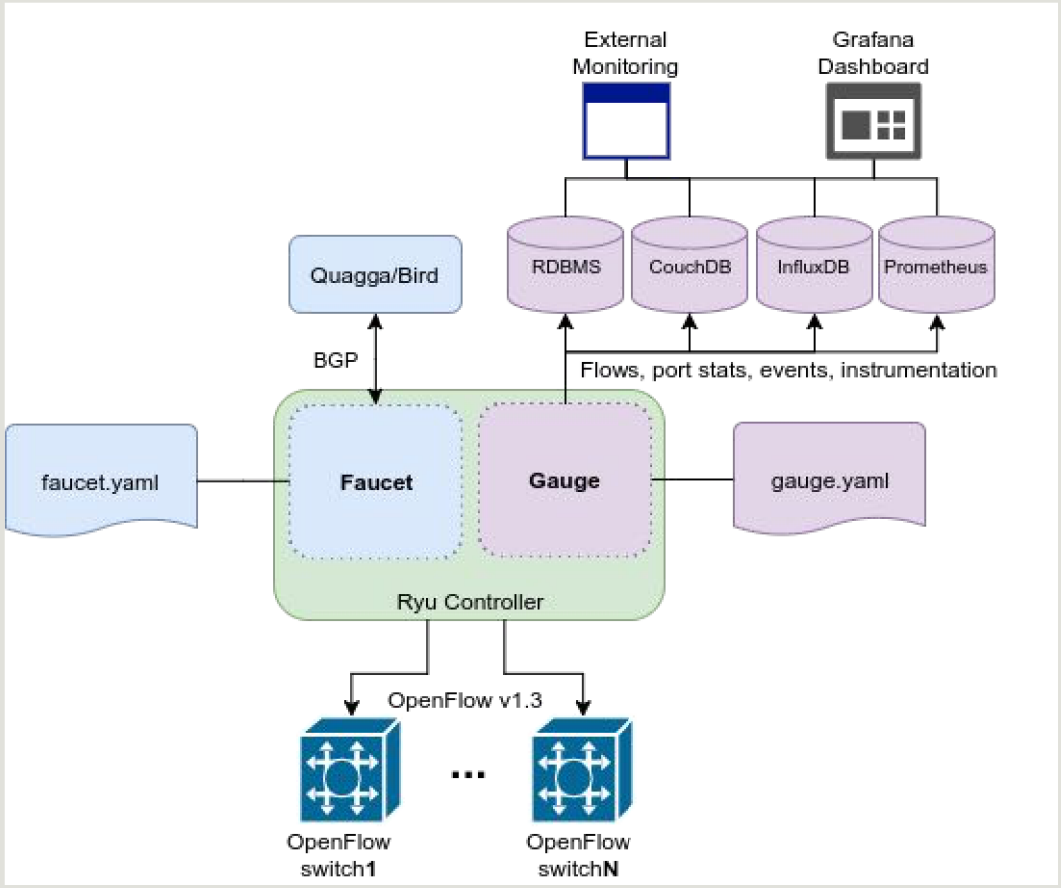
導入したスイッチ

Allied Telesis x930-28GTX

導入前テスト

- ARP unicasted test
- ICMPv6 ND unicasted test
- Filtering undesirable traffic test
- IPv4 and IPv6 traffic test
- Basic performance test
  - Iperf and HW Tester > 1G～10Gbps fully in HW
  - 1000 hosts config test
  - TCPReplay stress test with ARP / ICMPv6ND

FAUCETを利用したシステム概要



実際の config は YAML

```
version: 2
vlans:
    100:
        name: "clock"
        unicast_flood: True
        max_hosts: 3
    2001:
        name: "trusted network"
        unicast_flood: True
    2002:
        name: "untrusted network"
        unicast_flood: False
    2003:
        name: "roof network"
        unicast_flood: True
        max_hosts: 10
acls:
    101:
        - rule:
            dl_src: "ae:ad:61:7d:02:2f"
            actions:
                allow: 1
        - rule:
            actions:
                allow: 0
```

```
dps:
    zodiac-fx-1:
        dp_id: 0x1
        hardware: "ZodiacFX"
        interfaces:
            1:
                native_vlan: 100
                name: "clock"
            2:
                native_vlan: 100
                name: "VLAN 2001"
                acl_in: 100
    windscale-faucet-1:
        dp_id: 0x2
        description: "Josh's experimental AT-
X930"
        hardware: "Allied-Telesis"
        interfaces:
            1:
                tagged_vlans: [2001,2002,2003]
                name: "port1.0.1"
                description: "windscale"
            2:
                native_vlan: 2001
                name: "port1.0.2"
```

config 生成

人間が書くものではない

- YANG model に基づいて自動生成

何かしらの DB / ネットワークコントローラと連携する必要がある

- PIX-IE の場合は自作 DB + コントローラ (Umbrella)
- ixpmanager (https://www.ixpmanager.org/) との連携

その他にも。。。

最近 SD-WAN に手を出しています

学内のネットワーク管理に活かせないかと。。。

- ネットワーク利用・管理権限の委譲
- 位置に拘束されないネットワーク利用
- 構成変更の即時性

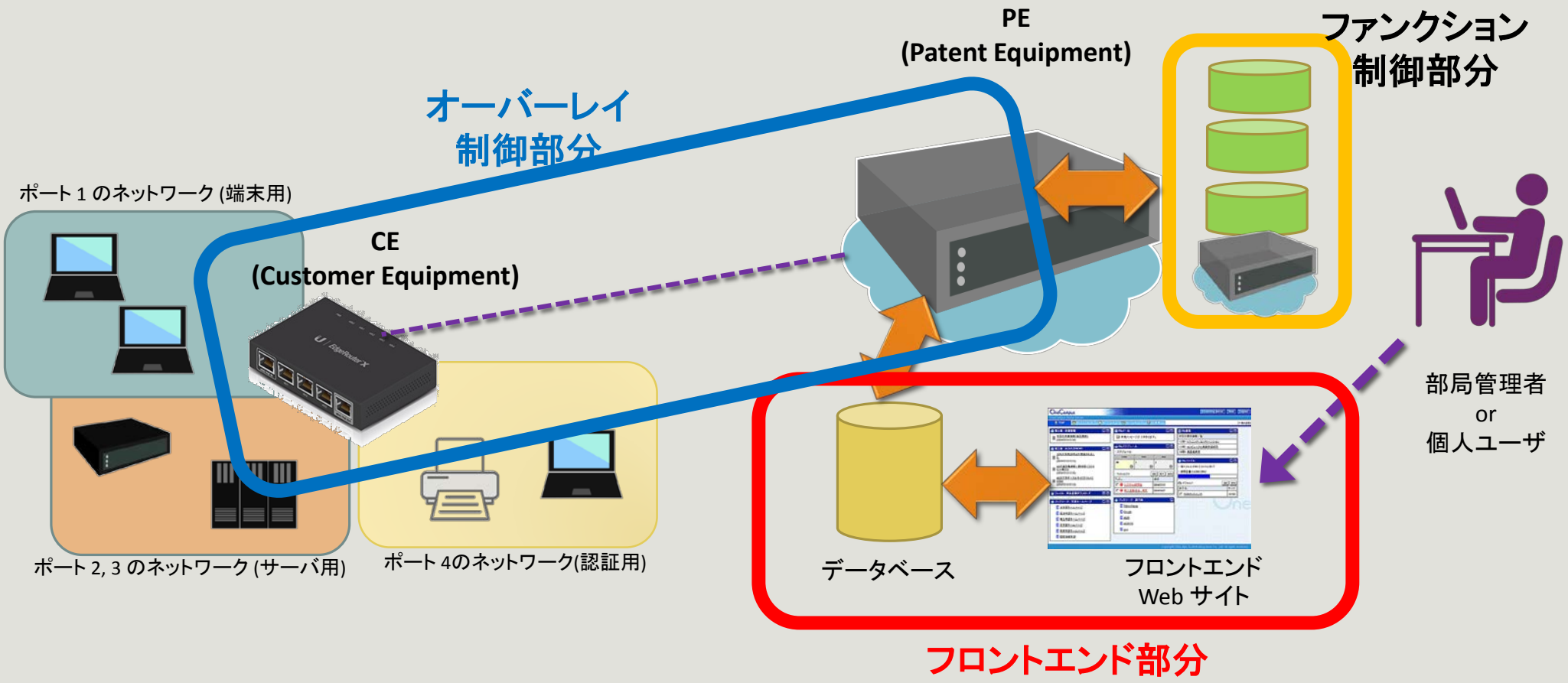
学内 SD-WAN + NFV を作ればいい？

一つの解法ではあるかもしれない

- 管理者・利用者に対して利用できるネットワーク権限を委譲
- 利用できるネットワークに対して利用できる機能の権限を委譲
- 学内・学外のどこにでも必要なネットワークを提供

ポータルを通じて全てを制御できるようにする

システム概要図


PE
(Patent Equipment)
ファンクション
制御部分
オーバーレイ
制御部分
ポート 1 のネットワーク (端末用)
CE
(Customer Equipment)
部局管理者
or
個人ユーザ
ポート 2, 3 のネットワーク (サーバ用)
ポート 4のネットワーク(認証用)
データベース
フロントエンド
Web サイト
フロントエンド部分

実装するにあたって

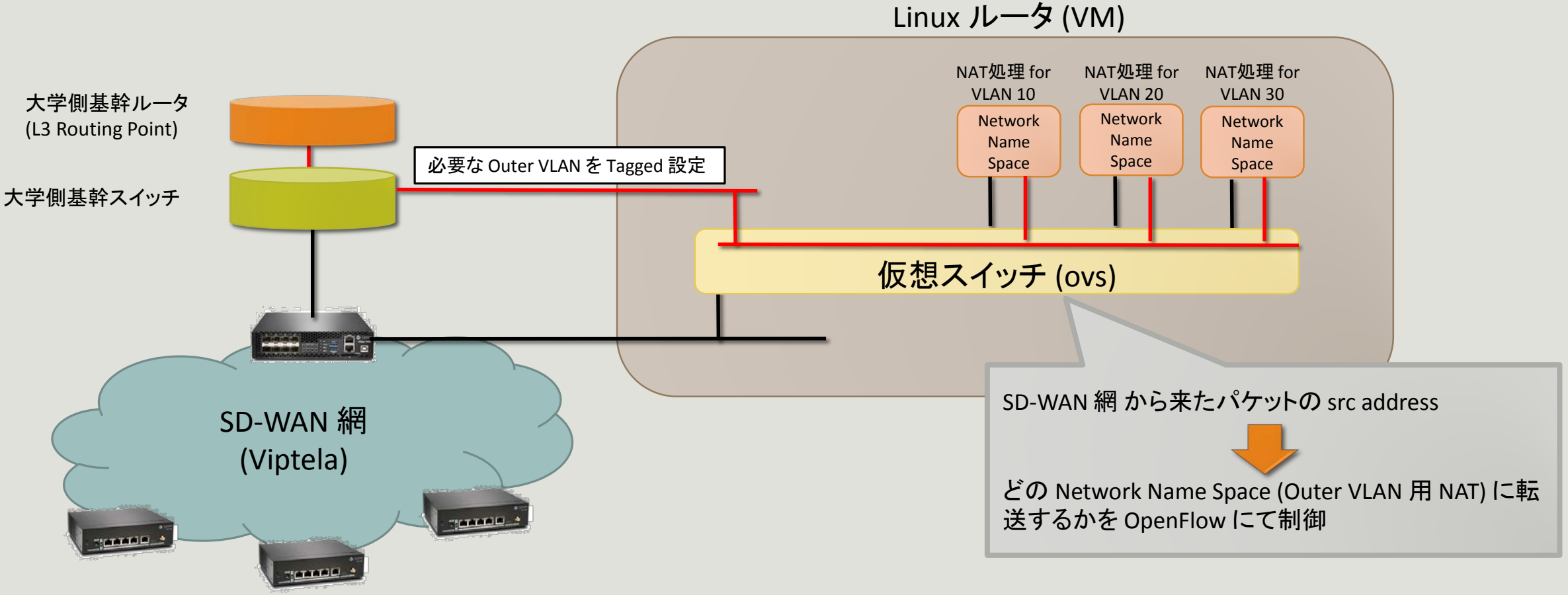
仮想ルータ
- VyOS
- Linux kernel (Network Namespace)

ネットワークファンクション
- Firewall (Filtering) ファンクション
- DHCP ファンクション
- NAT ファンクション
- DPI (URL filtering やサンドボックス) ファンクション

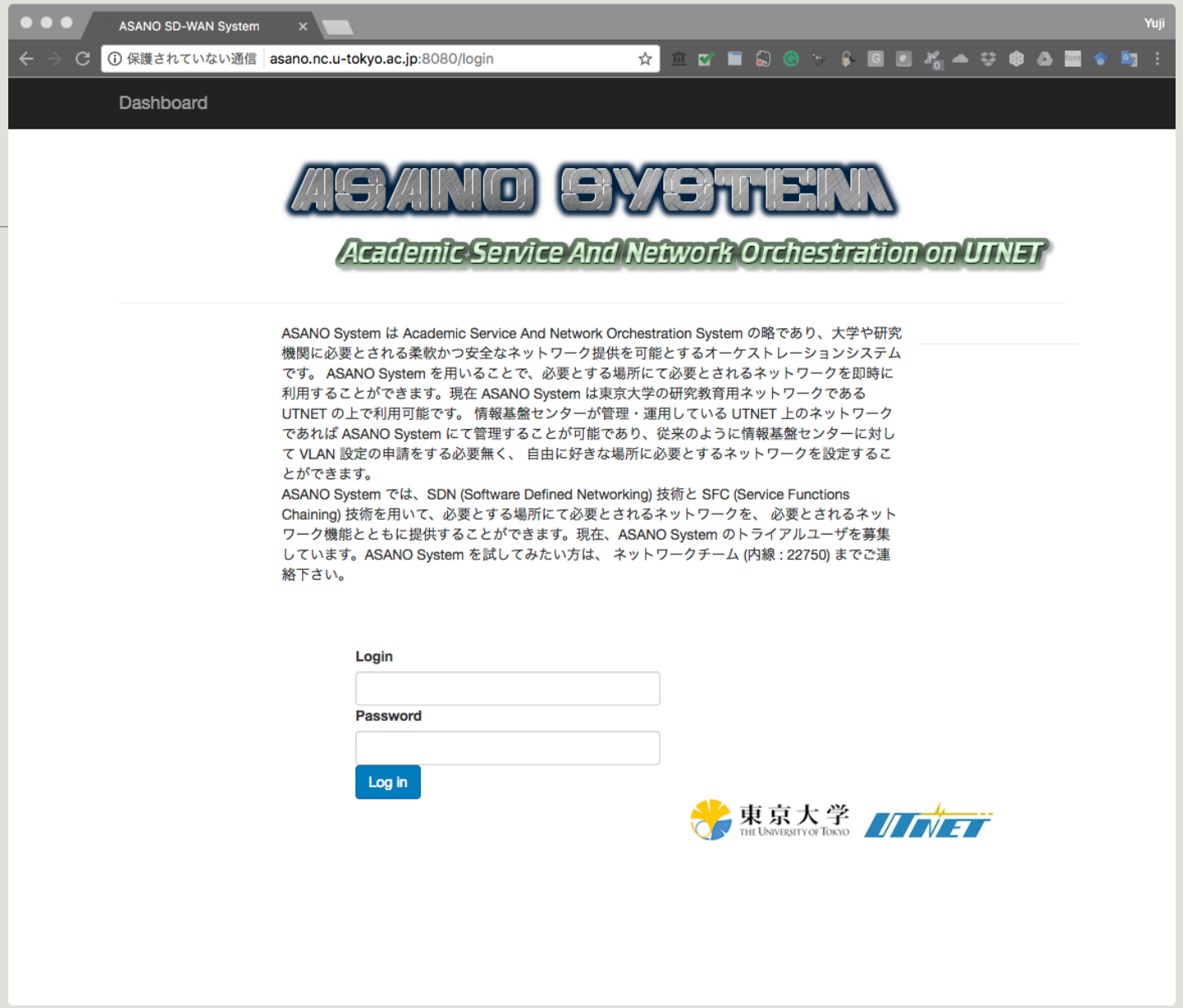
オーバレイネットワーク制御
- L2TPv3
- OpenVPN

P2P 型の VPN ではなく網として使えるオーバレイ型の VPN である必要がある
- DMVPN が必要
- VXLAN + EVPN ?

NAT と仮想ルータ


Linux ルータ (VM)
NAT処理 for VLAN 10
NAT処理 for VLAN 20
NAT処理 for VLAN 30
Network Name Space
Network Name Space
Network Name Space
大学側基幹ルータ
(L3 Routing Point)
大学側基幹スイッチ
必要な Outer VLAN を Tagged 設定
仮想スイッチ (ovs)
SD-WAN 網
(Viptela)
SD-WAN 網 から来たパケットの src address
どの Network Name Space (Outer VLAN 用 NAT) に転送するかを OpenFlow にて制御

とりあえず
サービス開始



OpenSource だけで SD-WAN

現在プロトタイプが動いています
- XCY のベアボーン PC の上で試験中

構成としては
- IPsec + DMVPN + VXLAN + EVPN
- 詳細は次の機会に。。。
- コンテナとしてデプロイ可能

さらなる構想として。。。

SD-WAN ブローカーみたいなものできないかな

- OpenSource による SD-WAN
- 商用の SD-WAN

異なる種類の SD-WAN 同士でネットワーク交換を行ったり、ファンクションを提供する

- SD-WAN ブローカー IX ・・・？

PIX-IE で SD-WAN ブローカーやってみよう

ソフトウェアをうまく使える人材を

一般社団法人 高度ITアーキテクト育成協議会（AITAC）

- https://aitac.jp/