

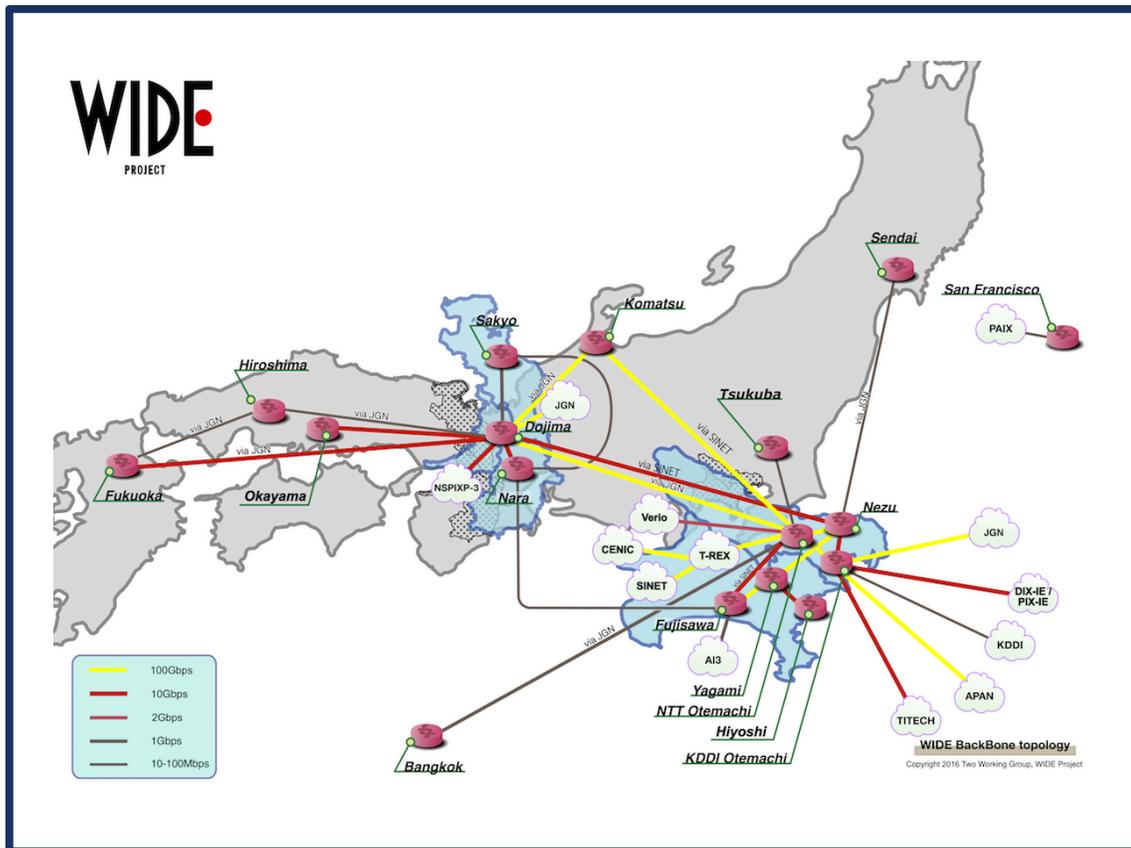
東京大学
情報基盤センター

WIDE PROJECT

関谷 勇司

WIDE PROJECT UPDATE 2019

WIDE PROJECT ネットワークバックボーン

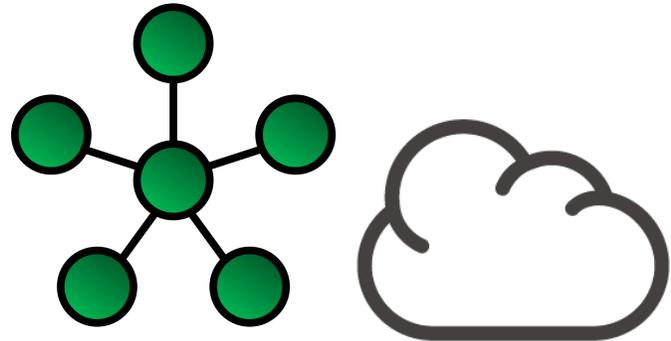


- 日本国内を中心としたネットワーク構成
- SINET / JGN の L2 網を利用して構築
- 拠点となる NOC
 - 大手町 DC、堂島 DC、慶應義塾大学、東京大学、奈良先端科学技術大学院大学、北陸先端科学技術大学院大学、筑波大学、広島大学 等
- 主に 100Gbps の回線を利用

インフラ

■ IX

- DIX-IE
- PIX-IE
- NSPIXP-3
- T-LEX (T-REX)



■ クラウド

- WIDE Project 内で利用する広域クラウド
- 大手町 DC、堂島 DC、東京大学、慶應義塾大学、奈良先端科学技術大学院大学、北陸先端科学技術大学院大学 を拠点として構成

■ DNS

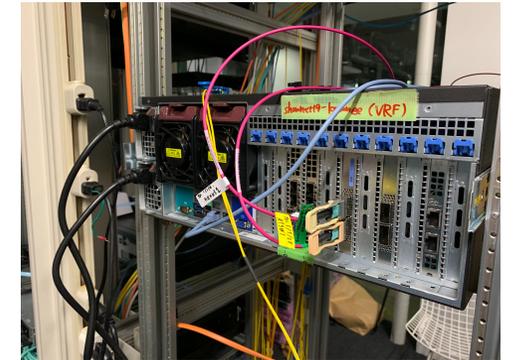
- JPRS との共同運用
- M-ROOT DNS
 - 日本 (2拠点)、フランス (2拠点)、US 西海岸 (2拠点)、韓国
 - 来年度にかけて拠点増を目指す
- JP DNS (E.DNS.JP)
 - 日本、フランス、US 西海岸



WIDE BACKBONE における WHITEBOX SWITCH の活用

Kamuee (SW) developed by NTT Com.

- DPDK-based router
- Poptrie (look up for large amount of FIB entries)
- OSS: Ubuntu, FRR (zebra), ...
- Protocol: Static, OSPFv2/v3, BGP4+, ...



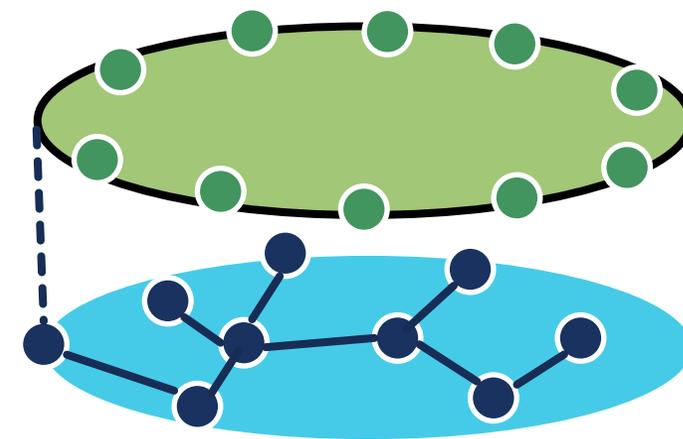
ThalarctOS (WBS) developed by KDDI

- WBS (Tomahawk) router
- Dev. Target: Switch Fabric
- OSS: ONL, FRR (zebra), ...
Protocol: Static, OSPFv2/v3, BGP4+, ...



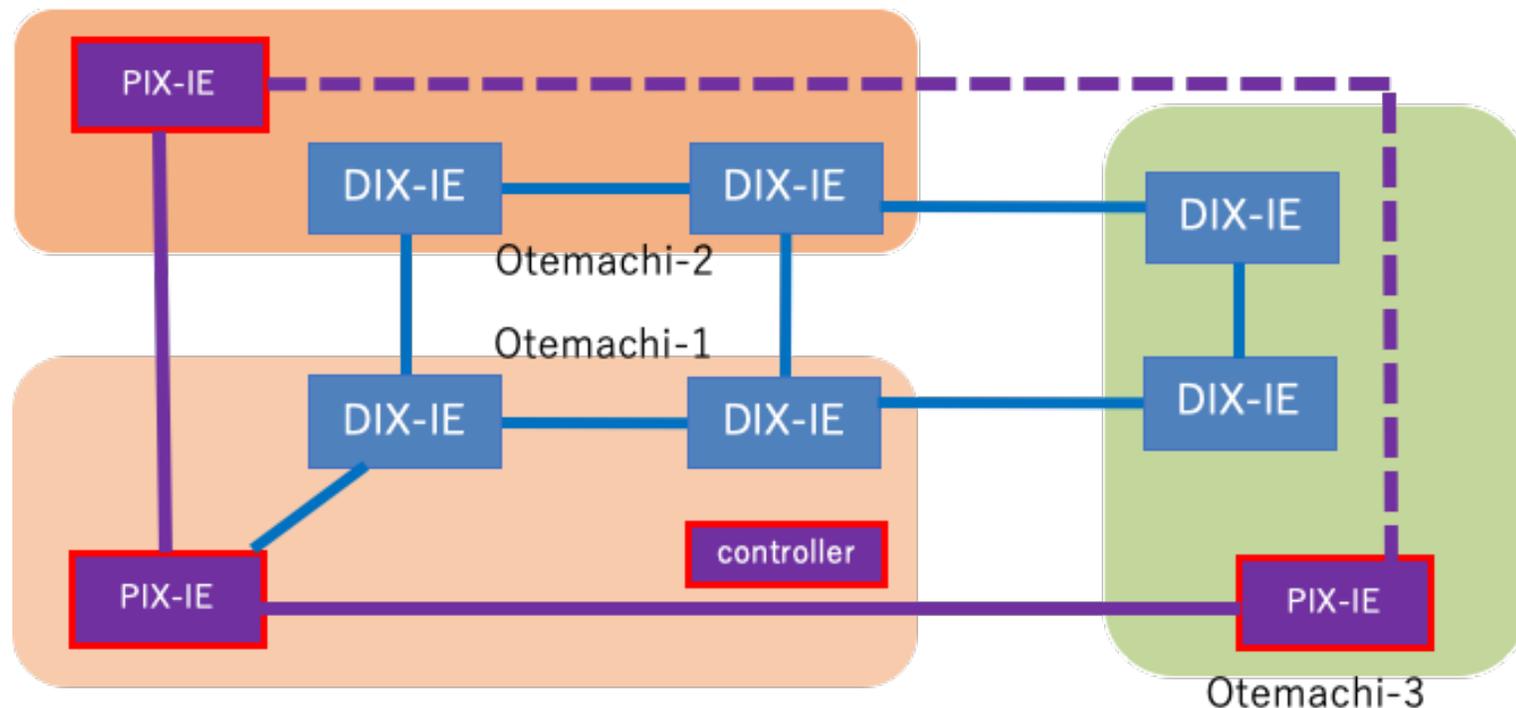
L2 OVER L3

- L2 を構成するにあたって迂回技術が問題に
 - 自身で構成できる網を持っていると L2 パスが多くなる
- 最近のネットワーク機器は優秀でそれほどトラブル発生しませんが
 - それでもメンテナンス等のために迂回技術は必要
 - でもここ 3週間くらいお騒がせしており申し訳ありません
 - 大手町コアルータが調子悪くてですね。。
- VLAN + STP はとうの昔に限界に達している
- WIDE バックボーン内で試みられているのは VXLAN
 - PIM-SM + VXLAN という VXLAN 本来 (?) の構成を実現
 - でもやはり EVPN + VXLAN の方がいいかな。。
- アンダーレイ (ハードウェアルータ) + VTEP (PC ルータ) という構成

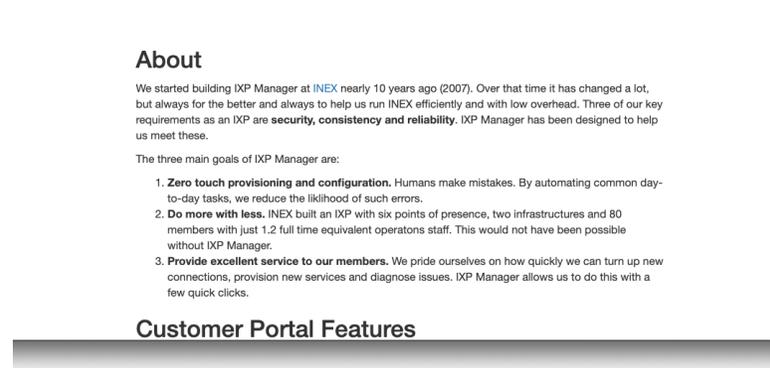
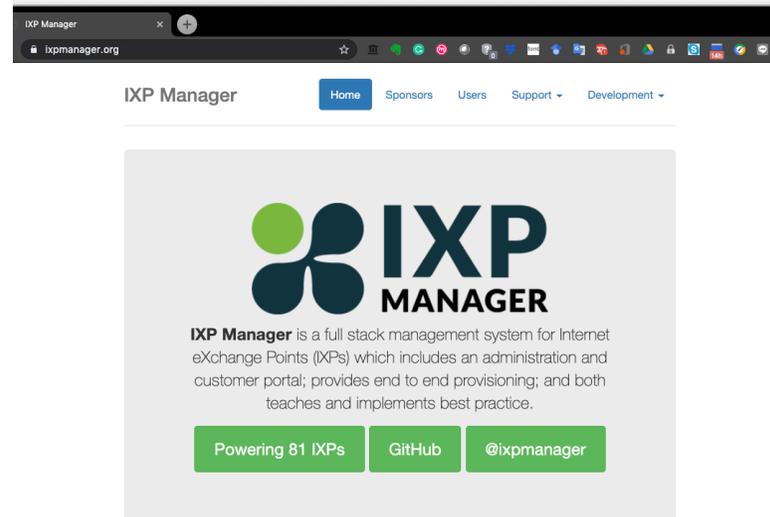
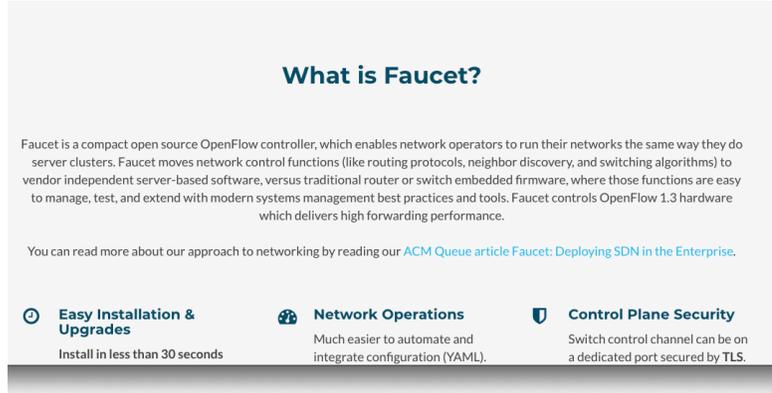
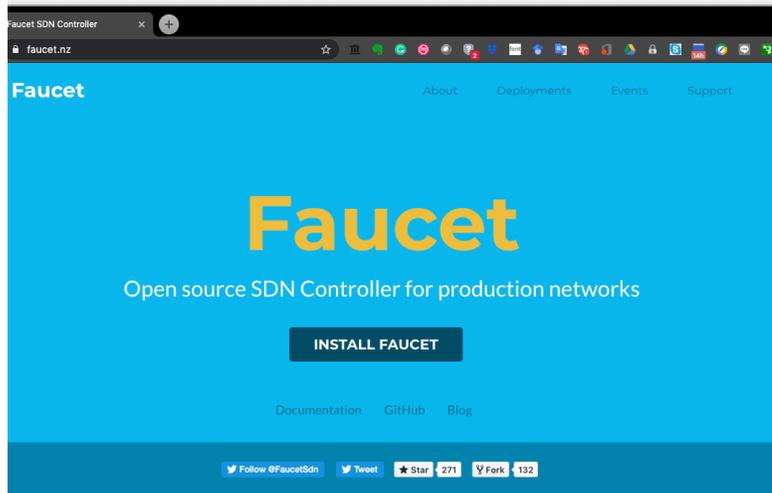


PIX-IE (PROGRAMMABLE INTERNET EXCHANGE)

- NICT / JGN との共同研究
- SDN を利用した IX の構築
- なぜ SDN を使うのか
 - そもそも限られたエリアにて用いるネットワーク技術
 - IX には Ethernet のフル機能は必要ない (むしろ邪魔になる)
 - IX に適した Ethernet の機能とスイッチ群の集中管理
- IX 運用の自動化



PIX-IE を構成する技術



2019年10月4日

- 光部材も PC パーツのように購入できるように
 - 低コストかつ運用自動化の IX が作れないか
- 現状は FAUCET + IXP Manager
 - IXP Manager との連携による運用自動化
 - Umbrella という方式を FAUCET にて実現
- Whitebox Switch の利用
 - OS が問題
 - IX のファブリック構成に適した機能とは？

IX 運用の自動化

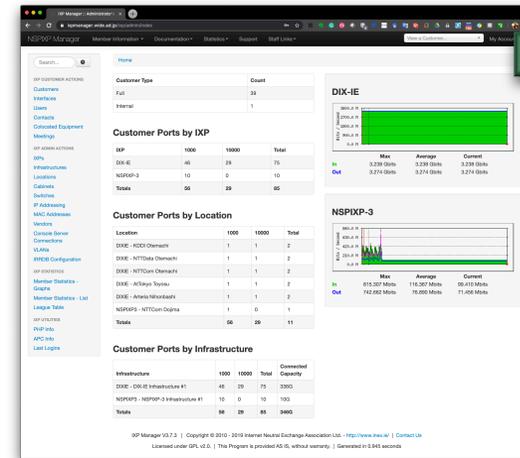
- IXP Manager による接続組織管理



- FAUCET コンフィグの自動生成



- IX 各拠点スイッチへの自動反映
(SDN Controller によるルール更新)

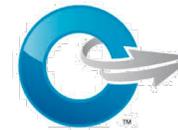


```
actions:
  - rule:
    dl_dst: "00:00:00:00:02"
    actions:
      output:
        port: 2

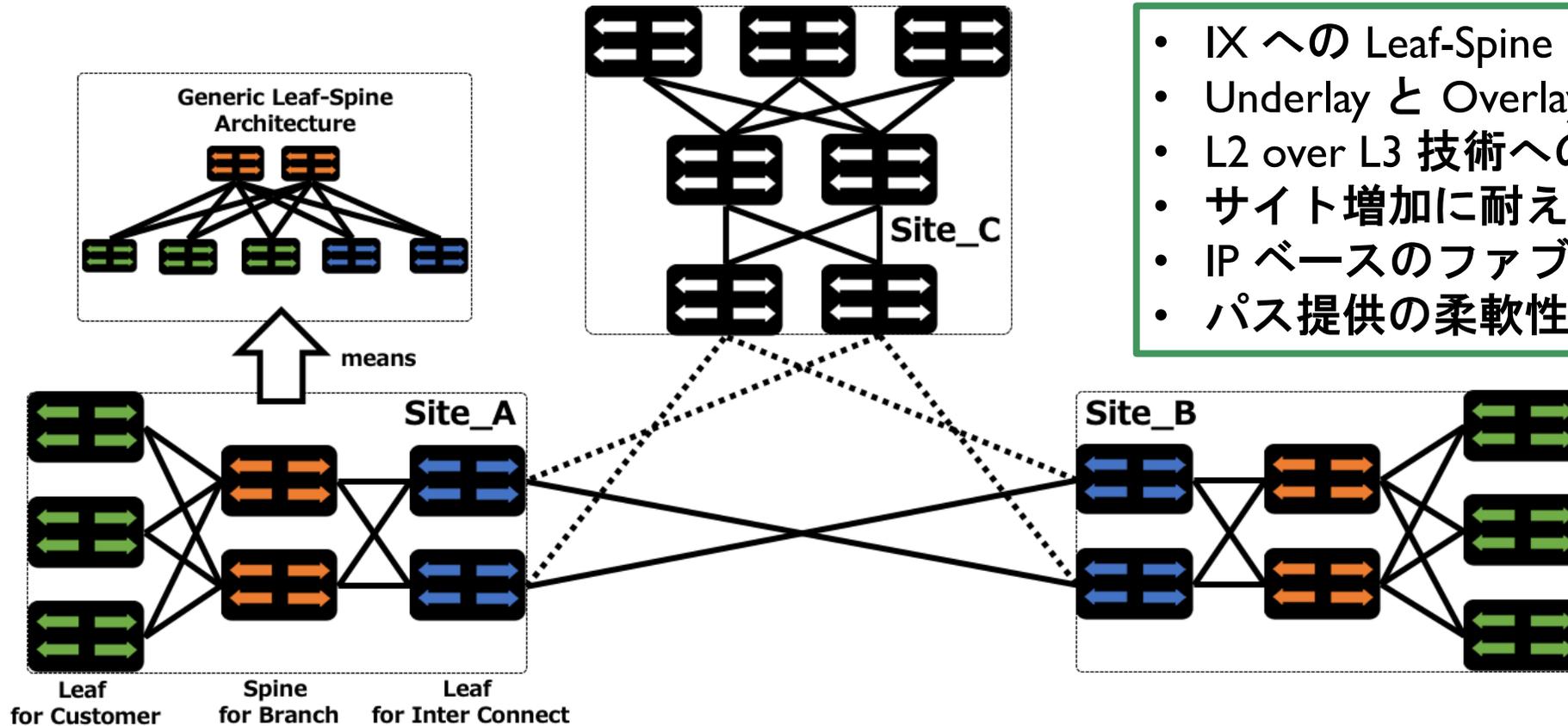
  - rule:
    dl_type: 0x0806
    dl_dst: "ff:ff:ff:ff:ff:ff"
    arp_tpa: "10.0.0.2"
    actions:
      output:
        port: 2

  - rule:
    dl_type: 0x86dd
    ip_proto: 58
    icmpv6_type: 135
    ipv6_nd_target: "2001::2/128"
    actions:
      output:
        port: 2

  - rule:
    actions:
      allow: 0
```



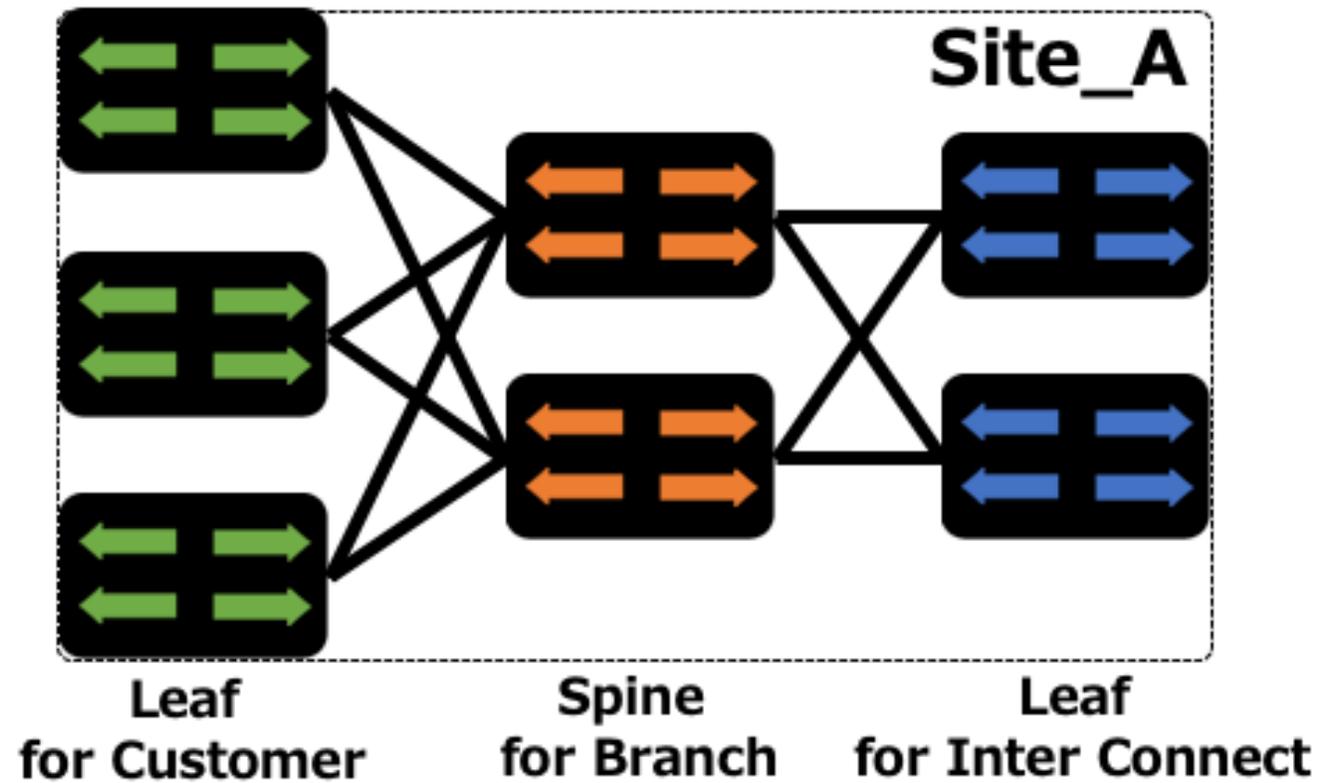
IX も次なるアーキテクチャへ (PIX-IE ++)



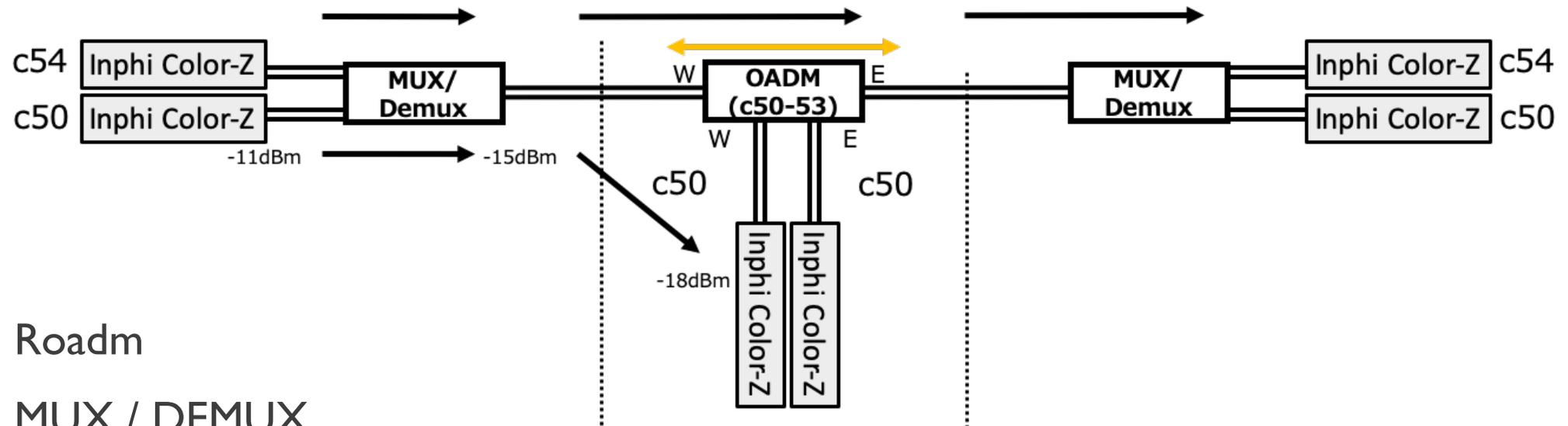
- IX への Leaf-Spine モデルの導入
- Underlay と Overlay の分離
- L2 over L3 技術への移行
- サイト増加に耐えうる規模性
- IP ベースのファブリック
- パス提供の柔軟性

3拠点で実験を開始 (予定)

- 実験拠点
 - KDDI 大手町拠点
 - NTT-C 大手町拠点
 - BBTower 大手町拠点
- L2 over L3 技術
 - VXLAN ?
- この場合経路サーバ
どうしよう
- Inter-Connect
 - 多波長 Optics の活用による
ファイバ本数削減



サイト間伝送路も安価な OOLS によって構成



- Roadm
- MUX / DEMUX
- 多波長光トランシーバ
- サイト間 400Gbps の実現に向けて



ありがとうございました