運用から見た研究
研究から見た運用

NTT ソフトウェアイノベーションセンタ

石田 渉

Janog35

自己紹介

Innovative R&D by NTT

- 名前
  - 石田 渉
  - Janog：参加3回目，発表2回目
- 所属
  - NTTソフトウェアイノベーションセンタ 2年目
- 仕事
  - nosw：ホワイトボックススイッチOS *Janog34発表内容参照
    - http://github.com/osrg/nosw
  - gobgp：Go言語BGP実装
    - http://github.com/osrg/gobgp
  - Ryu BMP Server：BMPサーバ *Janog34がきっかけ
    - http://osrg.github.io/bmp/

NTT

研究紹介

Innovative R&D by NTT

1. **WAN/DC間ネットワーク** | Google B4
2. **IX事業者** | SDX

NTT

研究紹介

Innovative R&D by NTT

1. **WAN/DC間ネットワーク** | Google B4
2. **IX事業者** | SDX

NTT

Google B4 | 概要

- *B4: Experience with a Globally-Deployed Software Defined WAN*

- sigcomm 2013

- **B4 : GoogleのDC間ネットワーク**
- **目的 : リンク増設の削減, 高価なネットワーク機器からの脱却**
- **結果 : リンク帯域使用率 30% → 70% を独自スイッチで実現**

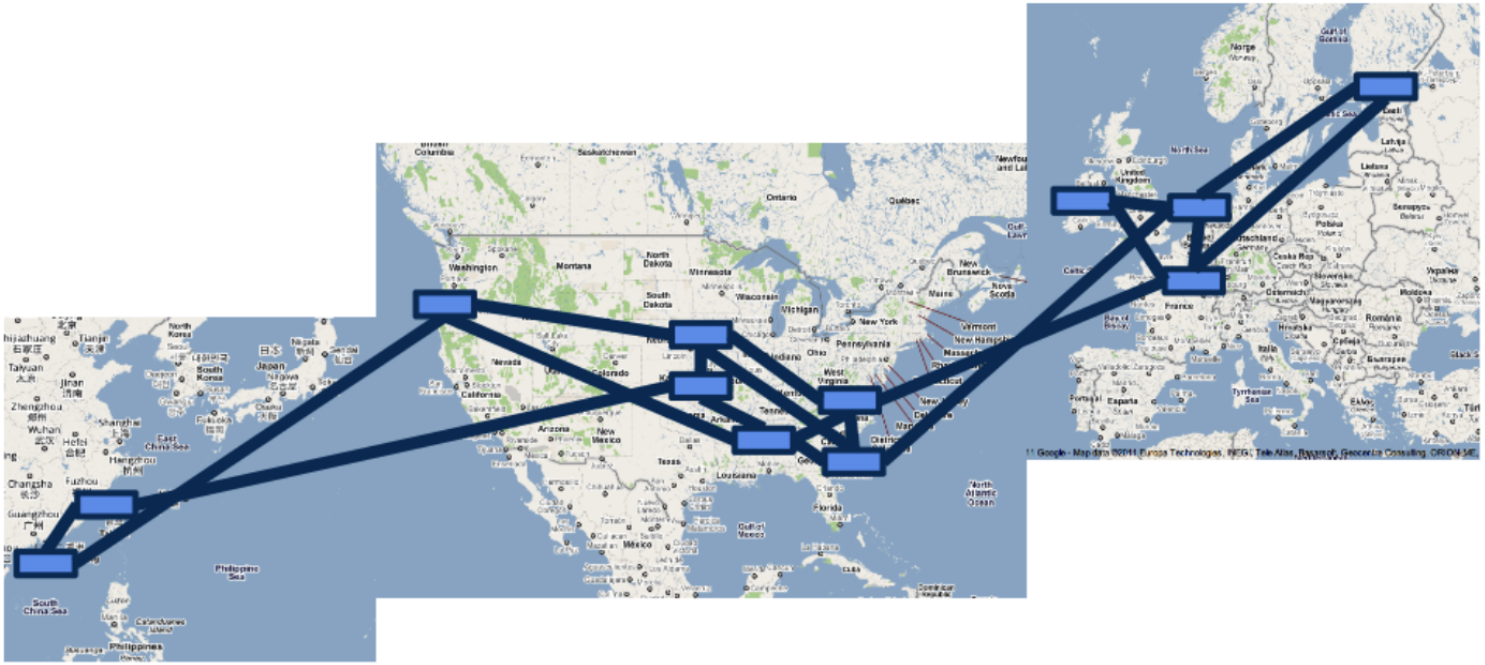


Figure 1: B4 worldwide deployment (2011).

Google B4 | 前提

Innovative R&D by NTT

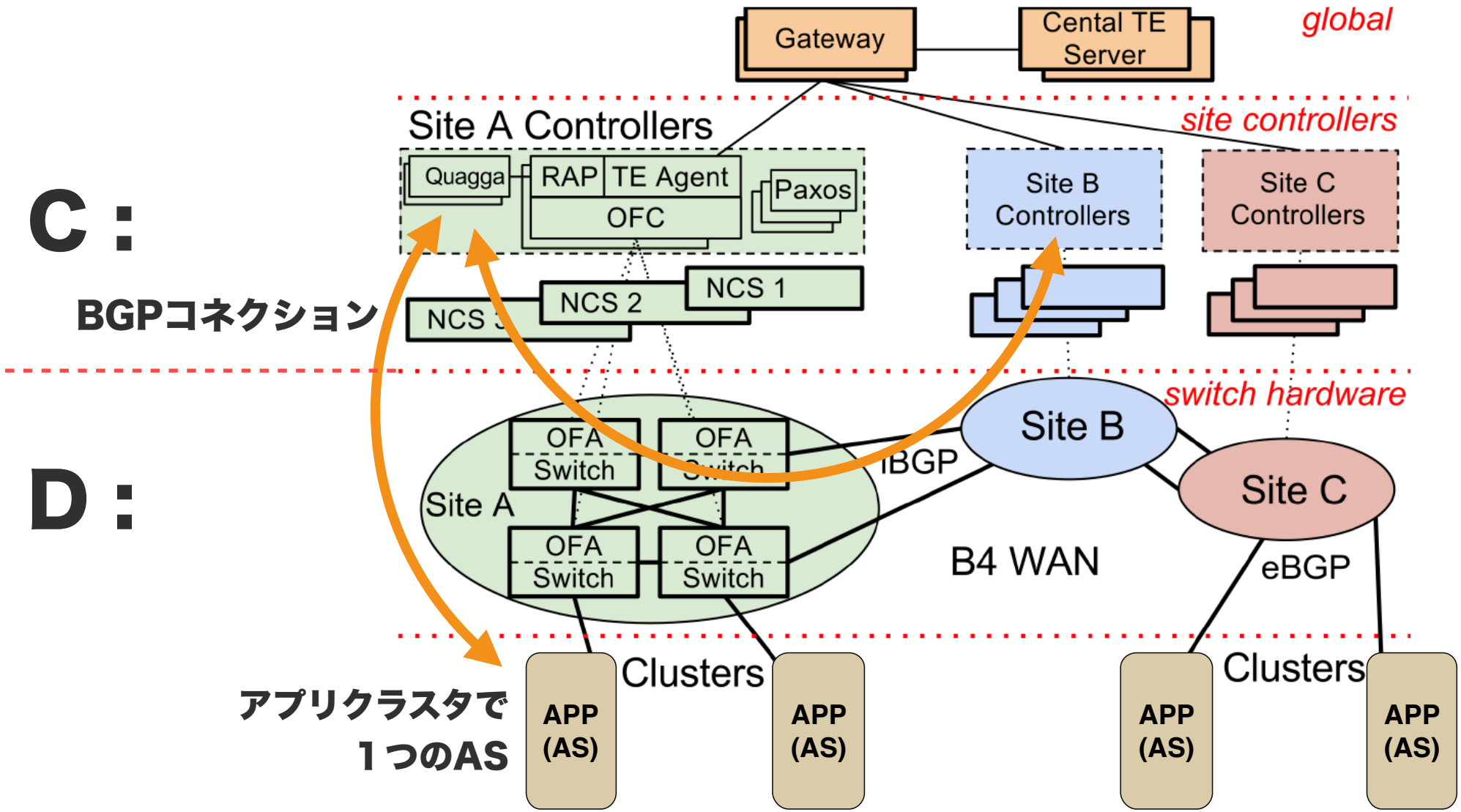
- **(特殊な)前提**
  - **アプリケーションの転送レートを制御できる**
  - **アプリケーションの優先度、使用帯域を知ることができる**
    - 優先度の低いものはリスケ、ドロップしてもよい
  - **拠点は数十程度**
    - 中央集権的に現実的な時間で経路計算しきれる

NTT

Google B4 | 設計方針

- **前提をもとに..**
  - **全てのアプリケーションはトラフィックを流す前にリンクの使用を予約**
  - **アプリケーションの優先度と使用帯域をもとに中央集権的に経路計算し、スイッチに経路注入**
    - 計算手法に関してはカレンダリングという研究分野

Google B4 | アーキテクチャ



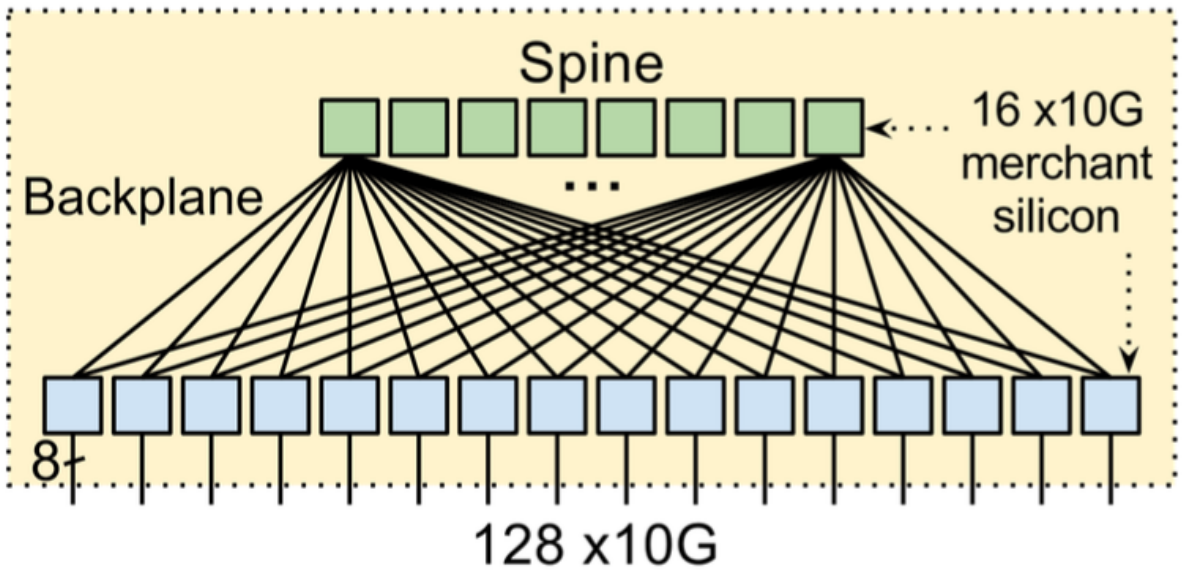
Google B4 | アーキテクチャの特性

Innovative R&D by NTT

- **アーキテクチャの特性**
  - **C/D分離**
    - C/D間プロトコルはOpenFlow
  - **制御プレーンはさらにレイヤリング**
    - BGP(quagga) / TE
    - IPの疎通性は担保したままIPinIPでTEを提供
  - **制御プレーンは分散ソフトウェア分野の知見を生かす**
    - Paxosによる冗長化
  - **データプレーンは汎用ASICを利用し、自社開発**
    - 16個のASICをスタッキングし、10Gbps ×128ポート

NTT

Google B4 | 自社開発スイッチ



Google B4 | 歴史

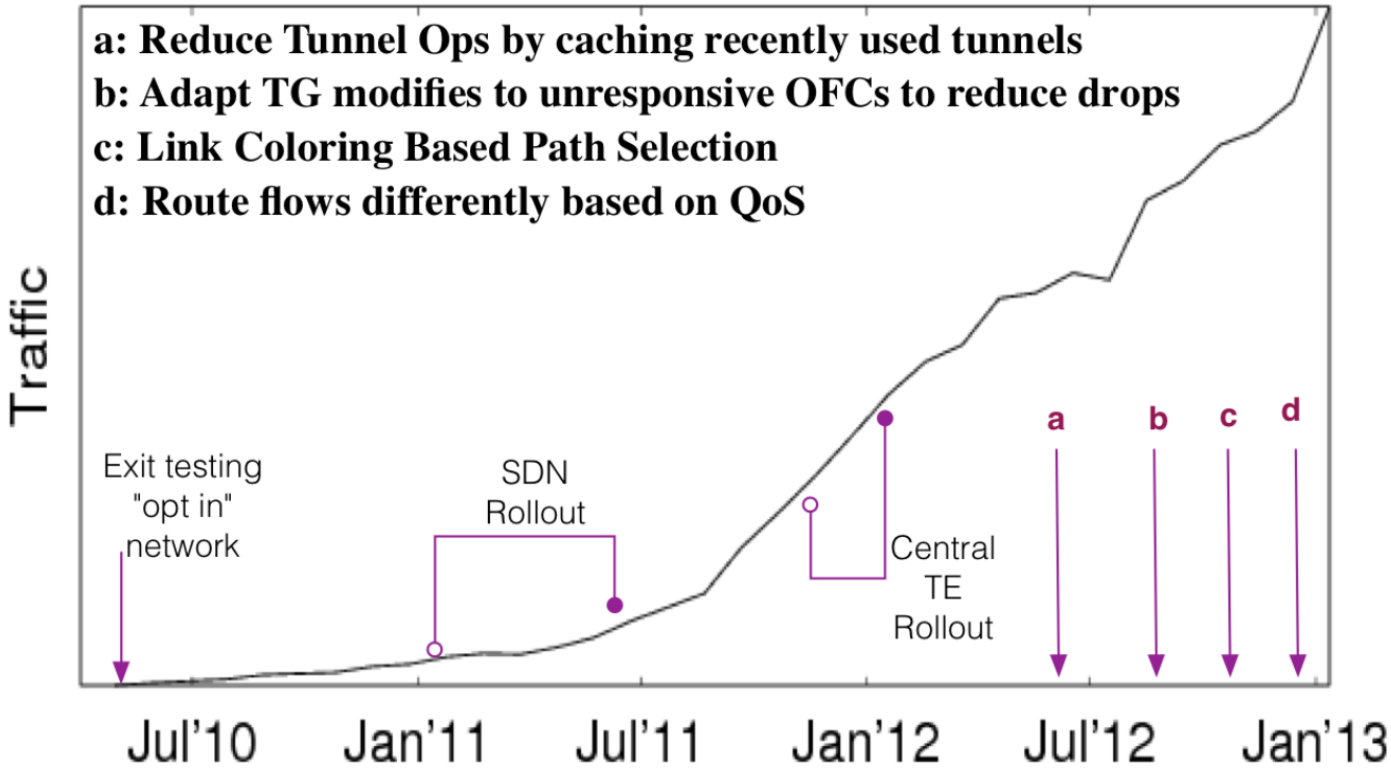


Figure 11: Evolution of B4 features and traffic.

- ある日突然、すべての機能が動き出したわけではない
- まずは拡張性のある形(SDNな形)で既存NWと同等なものを

Google B4 | 所感

Innovative R&D by NTT

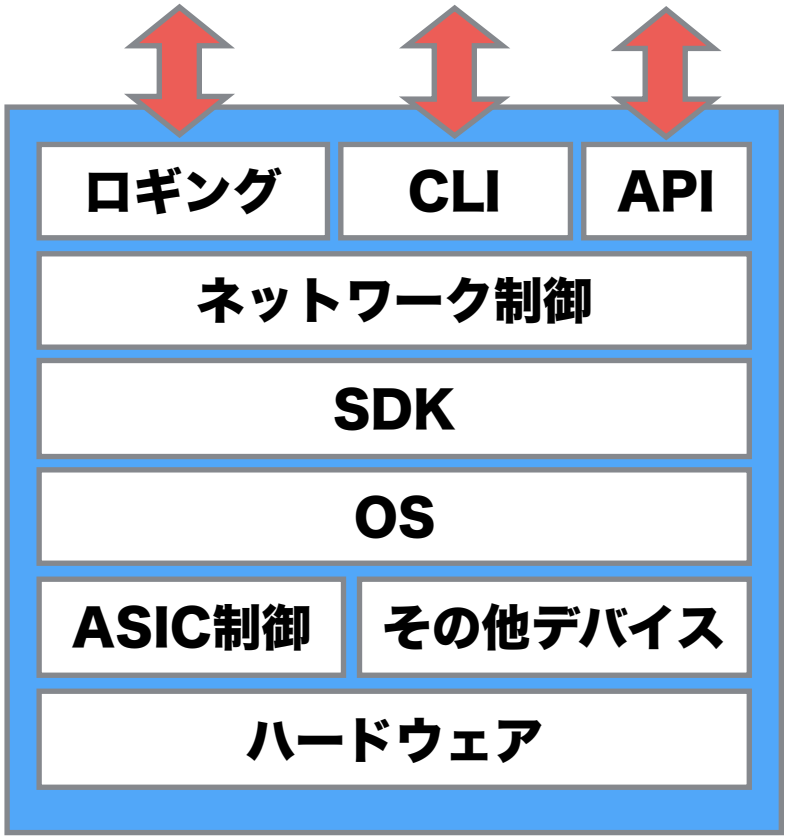
- OpenFlowが、SDNが、C/D分離が、独自スイッチがGoogleで動いている!!
- 前提は特殊。しかしアーキテクチャとしては参考になるのでは
- より柔軟、かつ堅牢なネットワークのためにはネットワーク機器のレイヤリング、ビルディングブロック化が必要

NTT

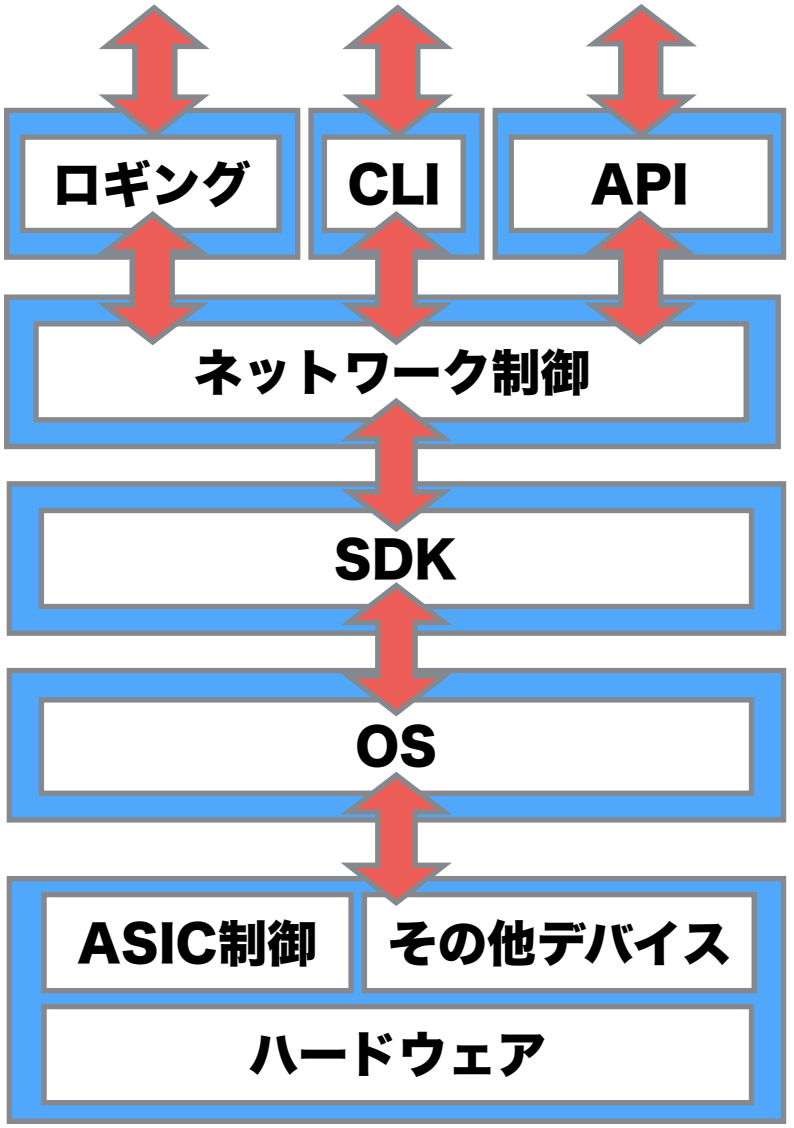
ビルディングブロック化

Innovative R&D by NTT

従来ネットワーク機器



これからのネットワーク機器



：公開されたインターフェイス

NTT

研究紹介

Innovative R&D by NTT

1. **WAN/DC間ネットワーク** | Google B4
2. **IX事業者** | SDX

NTT

SDX | 概要

Innovative R&D by NTT

- ***SDX: A Software Defined Internet Exchange*** *- sigcomm 2014*

- **IXへのSDNの適用の提案**
  - 実運用はされていない
- **類似のプロジェクト/講演は多数存在**
  - **Project Cardegan** (Janog33 - Joe Stringer)
  - **SDN-IXの可能性** (SDN Japan 2013 - MF 吉田さん)
  - **マルチラテラルピアリングの可能性** (Janog29 - MF 吉田さん, 渡辺さん)
- **論文の中心：ポリシ記述言語のOpenFlowへの効率的なコンパイル**
  - 今回はユースケースにフォーカス

NTT

SDX | ユースケース

Innovative R&D by NTT

1. Application-specific peering

- L4単位での経路交換

2. Inbound traffic engineering

- AS Path prepending, コミュニティ, route-mapは難しい

3. Wide-area server load balancing

- ラウンドロビンDNSはキャッシュに起因する問題点がある

4. Redirection through middle-boxes

- サービスチェイニングのIXでの適用

NTT

SDX | ポリシ記述の例

Innovative R&D by NTT

1. Application-specific peering

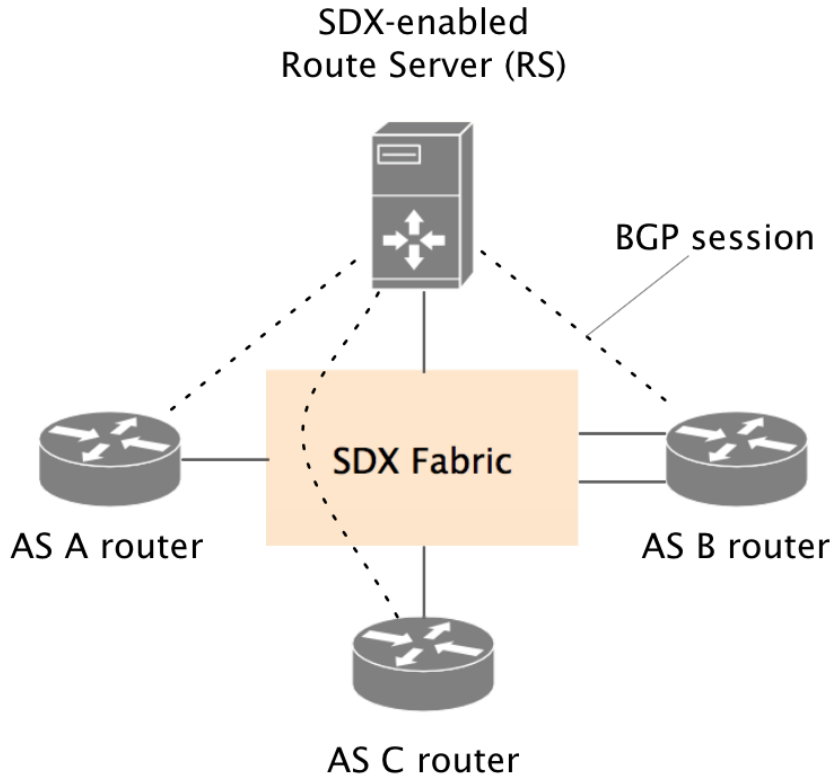
```
(match(dstport = 80) >> fwd(B)) +
(match(dstport = 443) >> fwd(C))
```

3. Wide-area server load balancing

```
match(dstip=74.125.1.1) >>
    (match(srcip=96.25.160.0/24) >>
        mod(dstip=74.125.224.161)) +
    (match(srcip=128.125.163.0/24) >>
        mod(dstip=74.125.137.139))
```

NTT

SDX | アーキテクチャ



- マルチラテラルピアリングをベースにポリシ機能の実装を提案

SDX | 所感

Innovative R&D by NTT

- **取り上げられているユースケースは本当に求められているのか**
- **BGP本体に機能追加するアプローチとのPros/Cons**
  - 今後BGPを運用したくないIXの顧客が出現するなら、BGPと共存でき、かつ単独でも利用できるシステムが望まれる
- **ルートサーバ実装は商用サービスでもOSSが利用されている**
  - OSS / プロプライエタリソフトウェアの違いは運用者にとってどのようなインパクトがあるのか

NTT

議論 | 進め方

Innovative R&D by NTT

1. 個々の研究事例に関して(各論)：15分

2. 運用と研究のサイクルに関して(総論)：15分

NTT

議論1 | 個々の研究事例

Innovative R&D by NTT

1. DC : Statesman

FWアップデートの安全な自動実行基盤

Q: NW運用自動化システム導入への障壁は？

2. Internet/ISP : Hybrid Networking Model

BGP C/D分離によるエッジルータレスAS

Q: BGP C/D分離の是非、導入への課題

3. WAN/DC間ネットワーク : B4

アプリケーションドリブンなWAN

Q: 通信に対する要求のレベル分けの是非

4. IX事業者 : SDX

IXでのポリシー制御

Q: モチベーションは正しい？ 需要はある？

NTT

議論2 | 運用と研究のサイクル

Innovative R&D by NTT

1. 研究者の感じるHyperGiantとの差

- 世界では研究段階の技術がどんどん実用化

2. なぜHyperGiantにこのような技術革新が可能か

- 圧倒的な規模 (人, 物, 金)
- アプリケーションからインフラまで所有

3. HyperGiantなき日本はどうすべきか

- コミュニティベースのサイクルを回したい e.g. OSS
- SIGCOMM Updateの開催

NTT