RTG関連 (idr/irs)

Shishio Tsuchiya

shtsuchi@cisco.com

運用/標準化コミュニティと私

Stateless IPv4 over IPv6 report draft-janog-softwire-report-01

Shishio Tsuchiya shtsuchi@cisco.com
Shuichi Ohkubo ohkubo@sakura.ad.jp
Yuya Kawakami kawakami@mfeed.ad.jp

Definitions of Managed Objects for 6rd

draft-cai-softwire-6rd-mib-03

Lei Cai cai.lei3@zte.com.cn

Jacni Qin jacni@jacni.com

Shishio Tsuchiya shtsuchi@cisco.com

Tao of the IETF

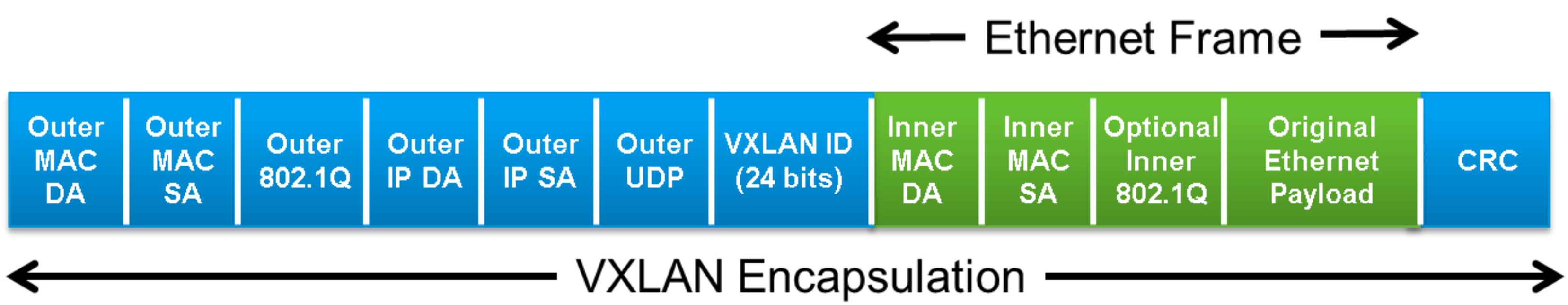
- The Tao of the IETF is now a web page: **http://www.ietf.org/tao.html**
- Five translations are now available: **http://www.ietf.org/tao-translations.html**
- These languages represent the top five non-English languages spoken by people that came to recent IETF meetings
- I hope that these documents help new IETF attendees learn about IETF culture

agenda

- IDR関連
- IRS関連

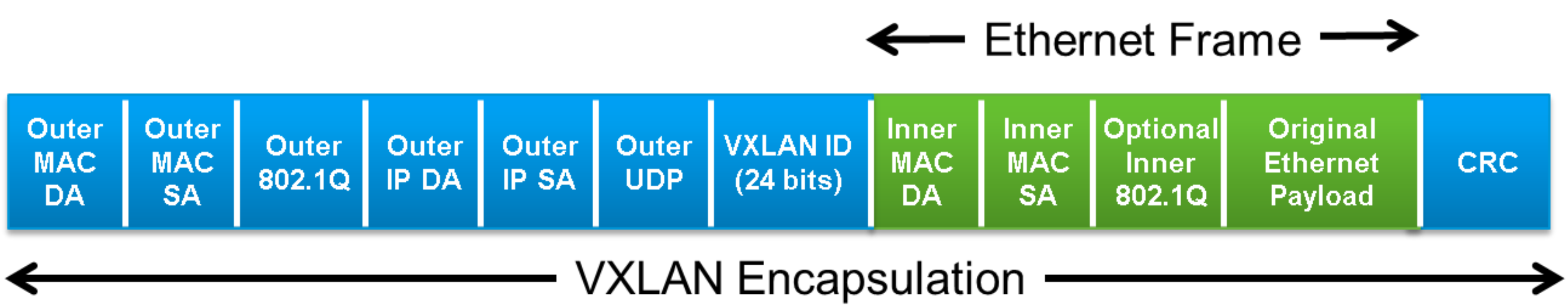
What is VXLAN

- Virtual eXtensible Local Area Network (VXLAN)
- VLANより伸張性(eXtensible)のあるレイヤー2セグメント
- レイヤー2フレームをIP/UDPでカプセル化
  - (50Byteのオーバーヘッド)
- IETF にdraft-mahalingam-dutt-dcops-vxlan をsubmit Cisco,VMWare,Citrix,Redhat,Broadcom,Arista

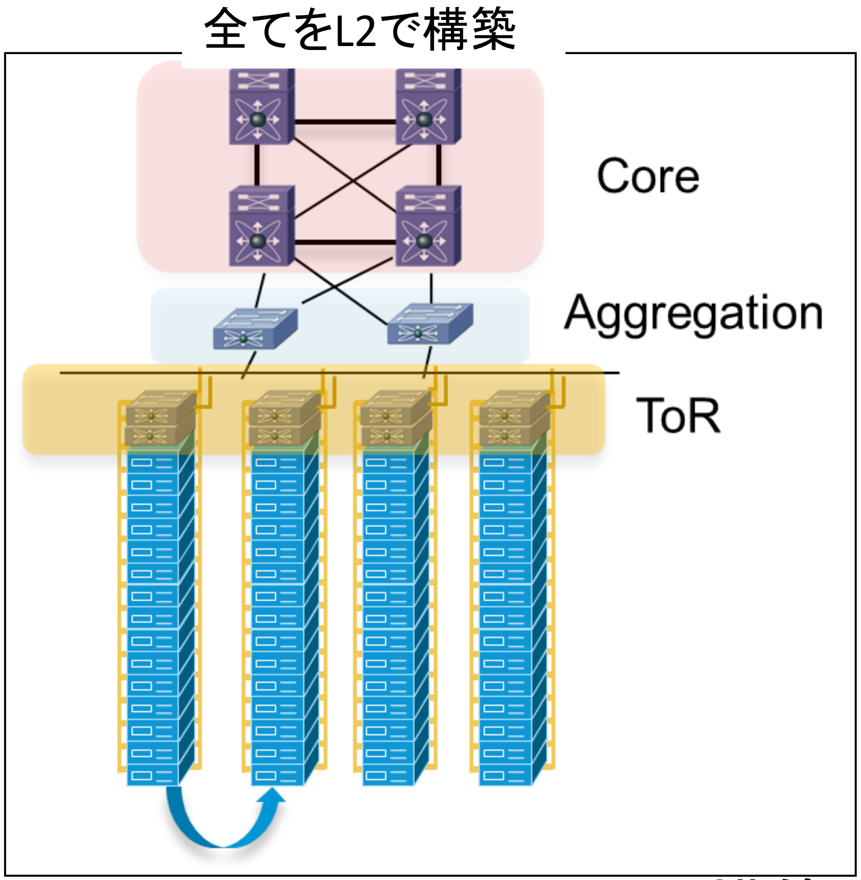


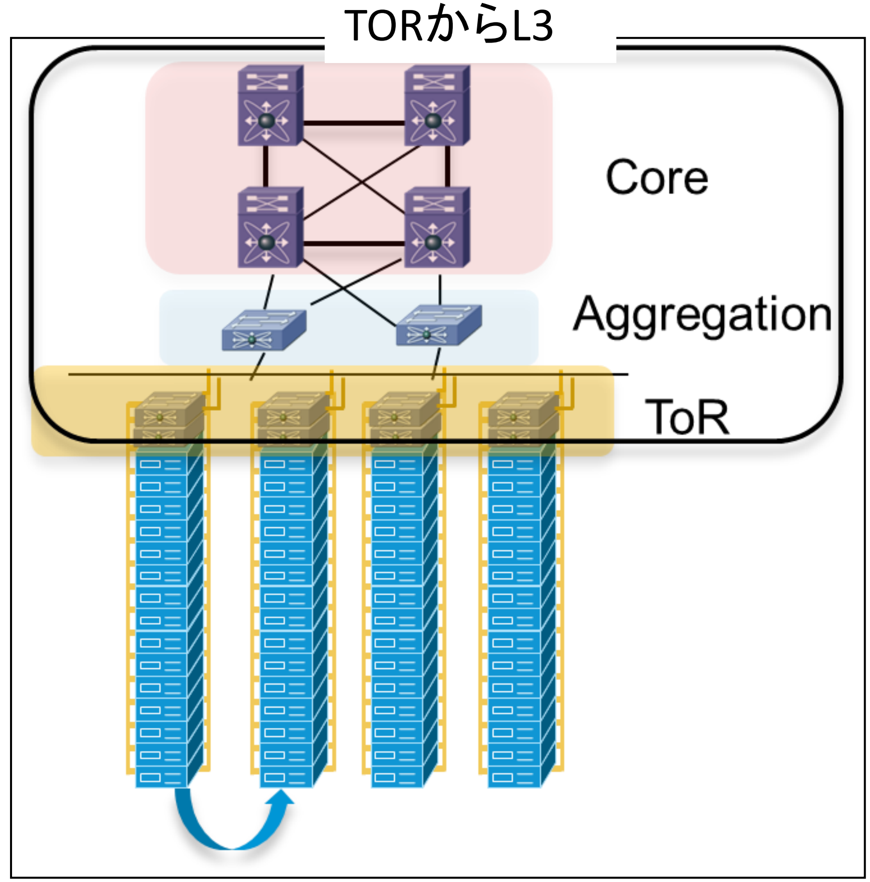
What’s is VXLAN *cont’d*

- 24ビットのVXLAN ID 2^24=16M
- VTEP(VXLAN Tunnel End Point)間でカプセル化
  - VMはVXLAN IDをケアしない。
- ブロードキャスト/マルチキャスト/Unknown ユニキャストにはIPマルチキャストを使用



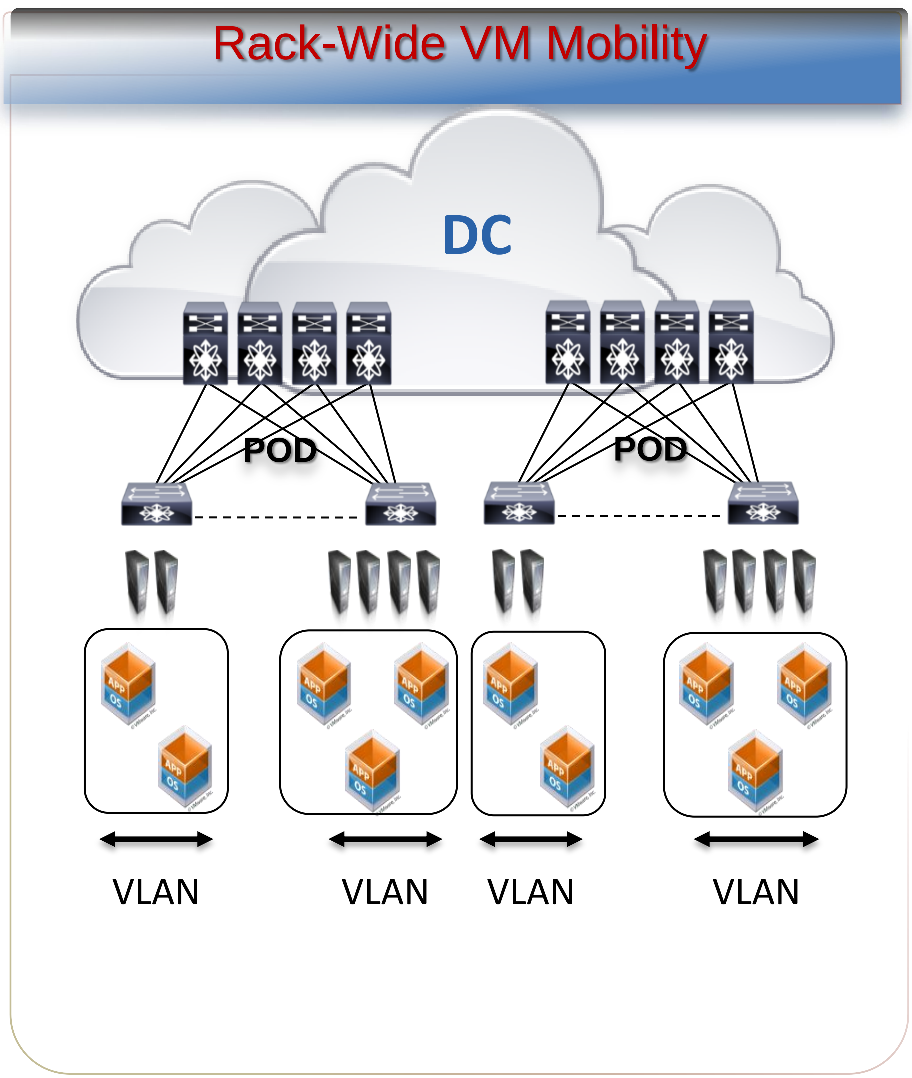
データセンターのスケーラビリティの問題を解決

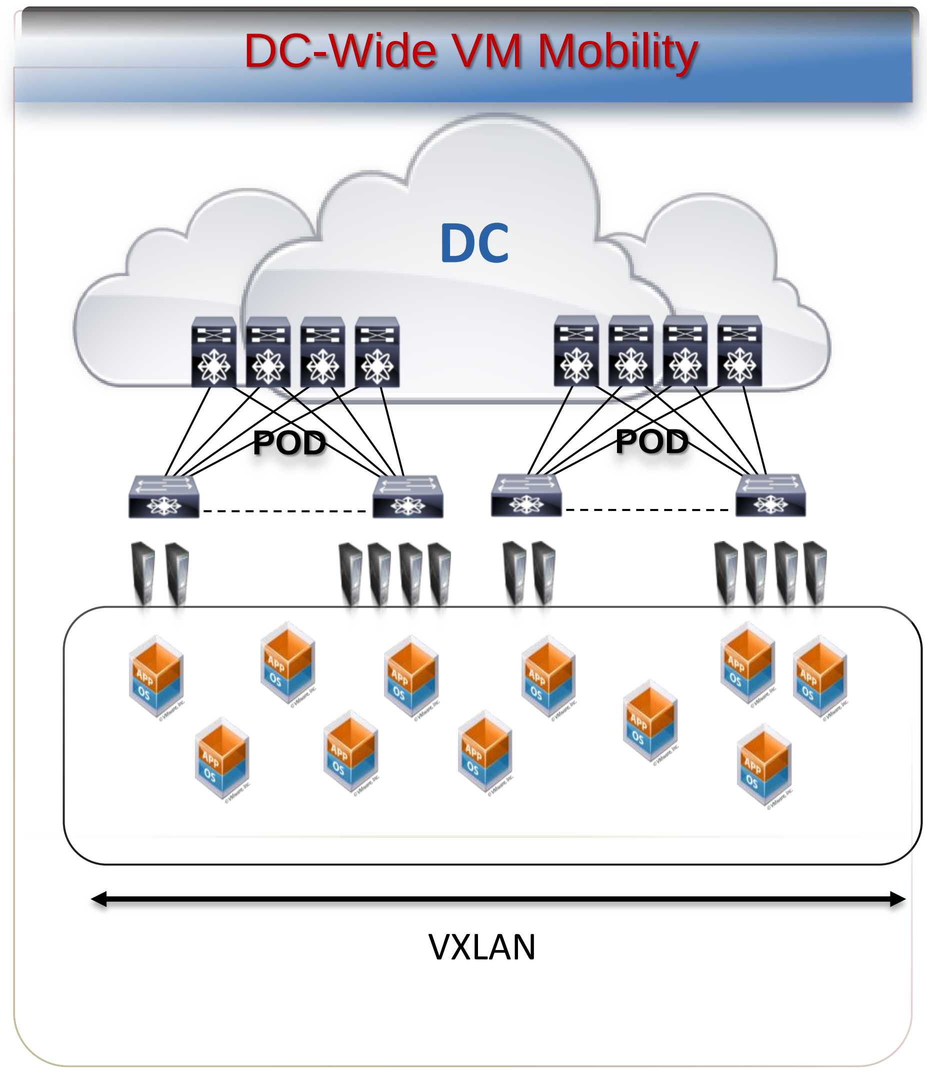




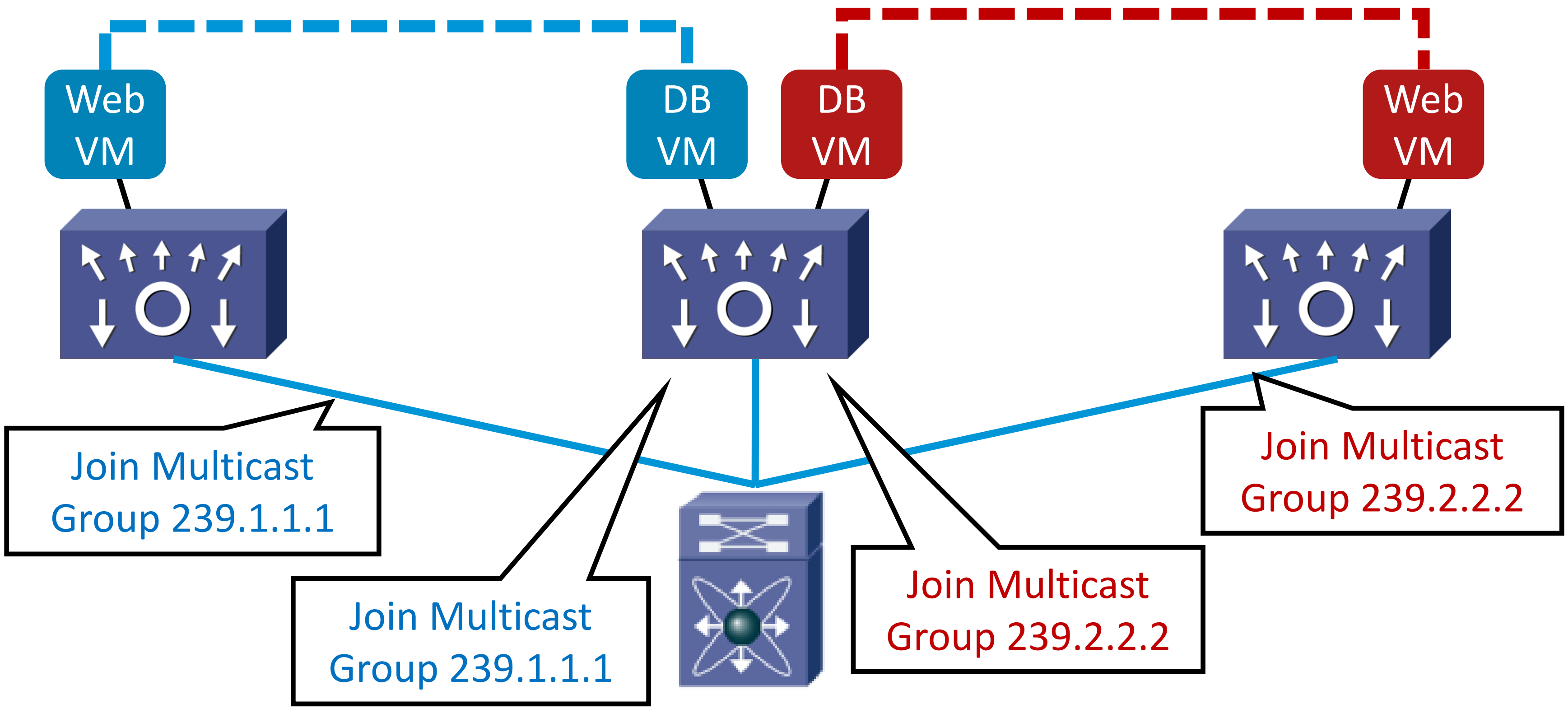
- ToRスイッチからL3で構築
- ARP/VLANのスケールの問題が解決し、ネットワークは安定に
- VMのモビリティは？？

VXLANが可能にするもの





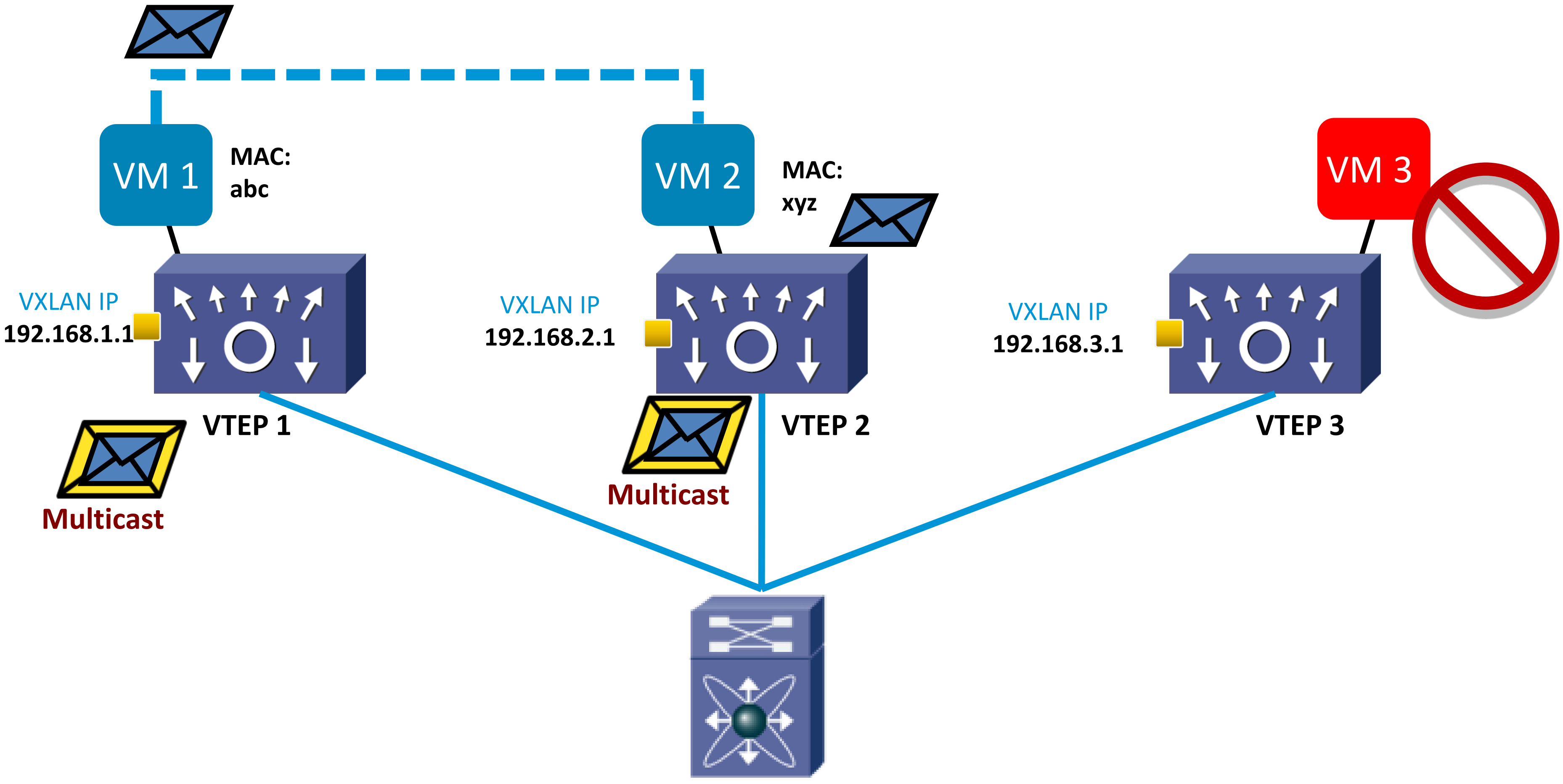
VXLANコントロールプレーン



- VMが接続されているポートとVNI(VXLAN Network Identifier)をひもづけ、VNI毎にマルチキャストグループを設定
- Muticast Join/Leaveにより最適化

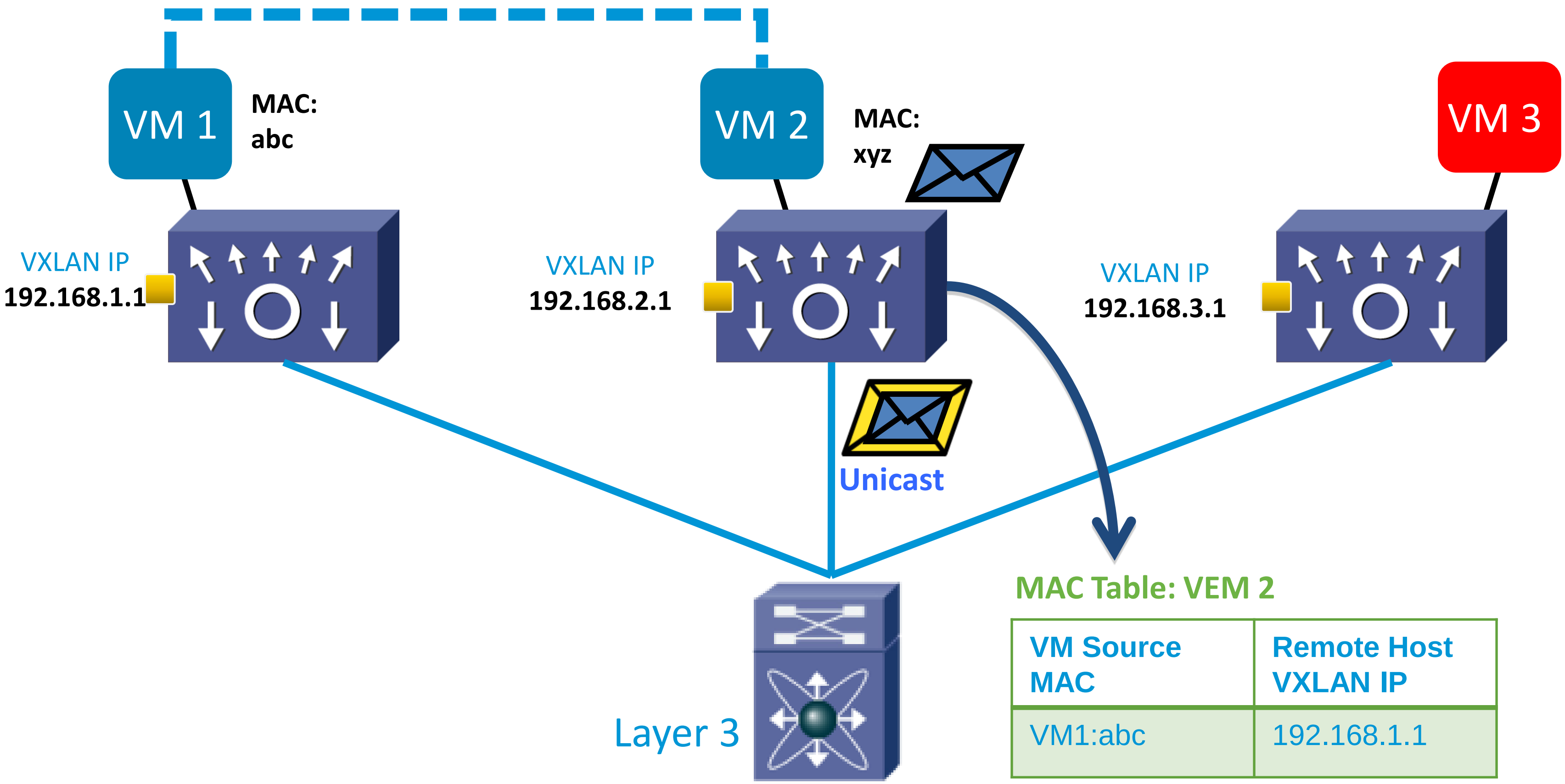
VXLANデータプレーン
Unknown Unicast

VM1 は VM2 と VXLAN 内で通信



VXLANデータプレーン
ラーニング

VM1 は VM2 と VXLAN 内で通信

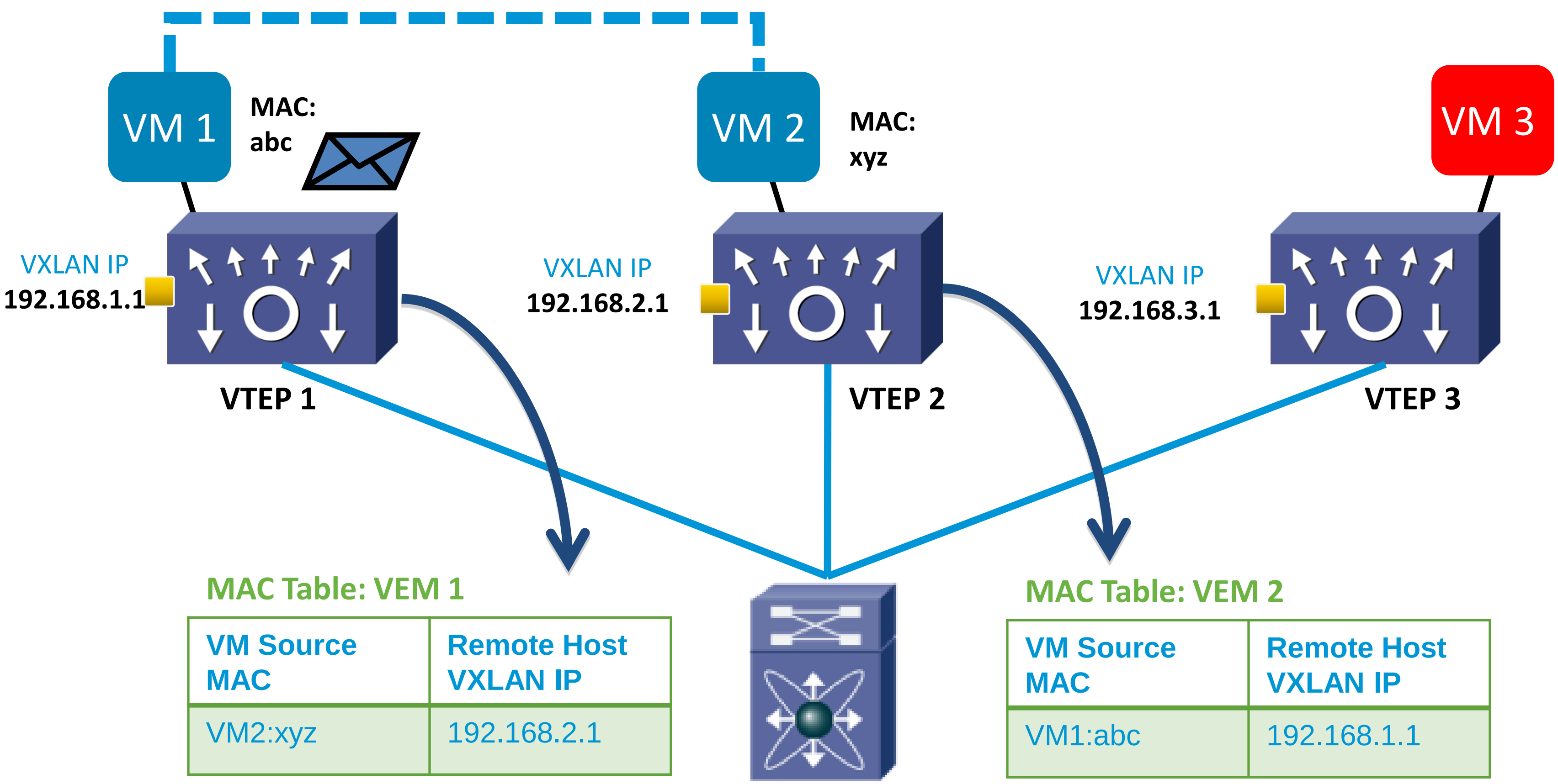


MAC Table: VEM 2

VM Source MAC	Remote Host VXLAN IP
VM1:abc	192.168.1.1

VXLANデータプレーン ラーニング *cont'd*

VM1 は VM2 と VXLAN 内で通信



MAC Table: VEM 1

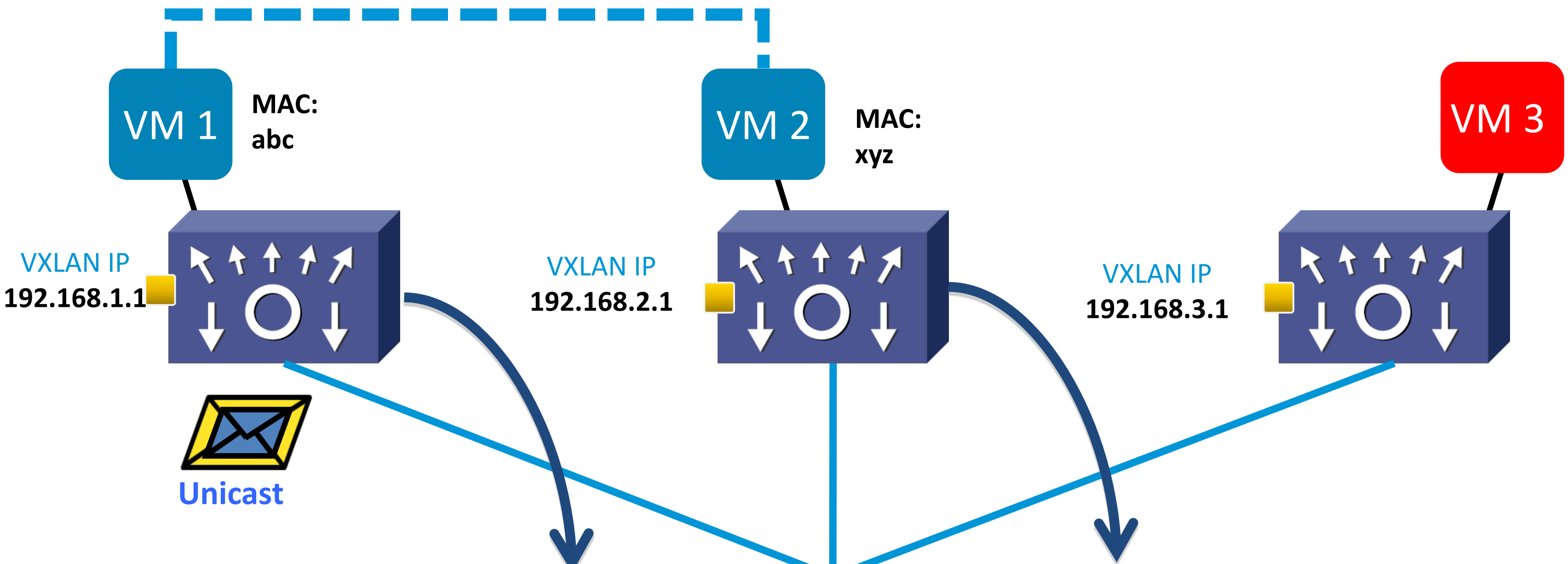
VM Source MAC	Remote Host VXLAN IP
VM2:xyz	192.168.2.1

MAC Table: VEM 2

VM Source MAC	Remote Host VXLAN IP
VM1:abc	192.168.1.1

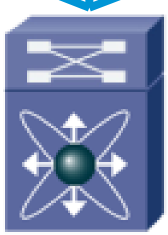
VXLANデータプレーン
Known MACアドレス

VM1 は VM2 と VXLAN 内で通信



MAC Table: VEM 1

VM Source MAC	Remote Host VXLAN IP
VM2:xyz	192.168.2.1



MAC Table: VEM 2

VM Source MAC	Remote Host VXLAN IP
VM1:abc	192.168.1.1

VXLAN vs NVGRE

類似点

- IPトランスポート
  - アクセススイッチで終端する
- IP マルチキャスト
  - ブロードキャスト/マルチキャストに関して
- 24ビットのセグメントID

相違点

- エンキャプスレーション
  - VXLAN: UDP 50バイト
  - NVGRE: GRE 42 バイト
- ポートチャネル負荷分散
  - UDP: 5 Tupleハッシュ(UDP,s/d ip,s/d port)
  - GREヘッダーでのハッシュは難しい
- ファイヤーウォールやACL、NATはVXLANのUDPポートで動作可能
  - GREプロトコルフィールドの実装は難しい
- フォワーディング手法
  - VXLAN: フラッディング/ラーニング
  - NVGRE: 定義無し？

IETF Network Virtualization Overlays (NVO3) WG

- NVO3 (Network Virtualization Over Layer 3) WG
  - nvo3@ietf.org
- Framework for DC Network Virtualization
  - draft-ietf-nvo3-framework-01
- Problem Statement: Overlays for Network Virtualization
  - draft-ietf-nvo3-overlay-problem-statement-01
- A Control Plane for Network Virtualized Overlays
  - draft-drake-nvo3-evpn-control-plane-00
- Network Virtualization Hypervisor-to-NVE Overlay Control Protocol Requirements
  - draft-kreeger-nvo3-hypervisor-nve-cp-00
- NVO3 Data Plane Requirements
  - draft-bl-nvo3-dataplane-requirements-03

BGP/E-VPN Extension for VXLAN/NVGRE

- A Network Virtualization Overlay Solution using E-VPN
  - draft-sajassi-nvo3-evpn-overlay

  データセンター間で使われるE-VPNテクノロジーをNVOでも使用出来るような提案。inter-asやマルチテナントのVXLAN/NVGREの拡張として使用出来る可能性

- BGP Remote-Next-Hop
  - draft-vandevelde-idr-remote-next-hop

What is BGP remote nexthop?

- new optional transitive attribute
- AS間で自動的にTunnelをセットアップする為の拡張
- Tunnel Encapsulationの情報はattributeの中に含む
- RFC5512 Tunnel Encapsulation Attributeでは任意のIPアドレスをTunnelのエンドポイントとしてセットが出来無い

TLVフォーマット

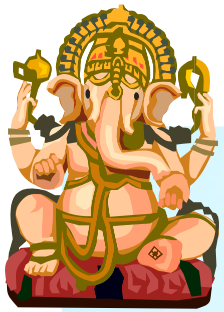
```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|    Tunnel Type (2 Octets)     |            Length             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Addr len      |      Tunnel Address (IPv4 or IPv6)            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                          AS Number                            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                      Tunnel Parameters                        |
~                                                               ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

- Tunnel Type
  - L2TPv3 over IP [RFC3931]: Tunnel Type = 1
  - GRE [RFC2784]: Tunnel Type = 2
  - IP in IP [RFC2003] [RFC4213]: Tunnel Type = 7
  - VXLAN: Tunnel Type = TBD by IANA
  - NVGRE: Tunnel Type = TBD by IANA
- TunnelパラメーターはRFC5512のものを使用する

agenda

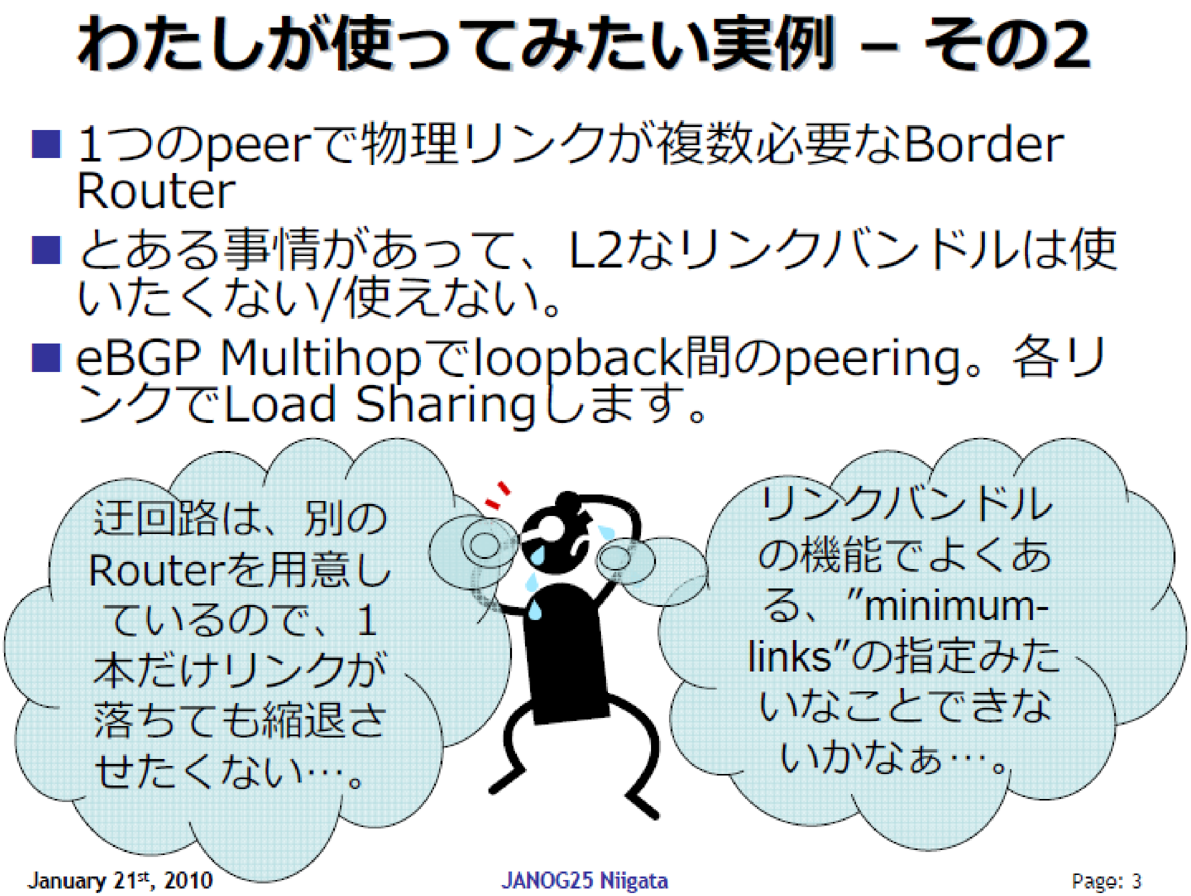
- IDR関連
- IRS関連

Openflow/SDNは？



- Openflow/SDNに期待する事は以下の様なものがある
  - ✓ 新しいデータプレーンの創造
  - ✓ **サービスとルーティングシステムの連動・自動化**
  - ✓ 共通のマネージメントプレーンの創造
  - ✓ サービスとDPIによるセキュリティ製品との連携

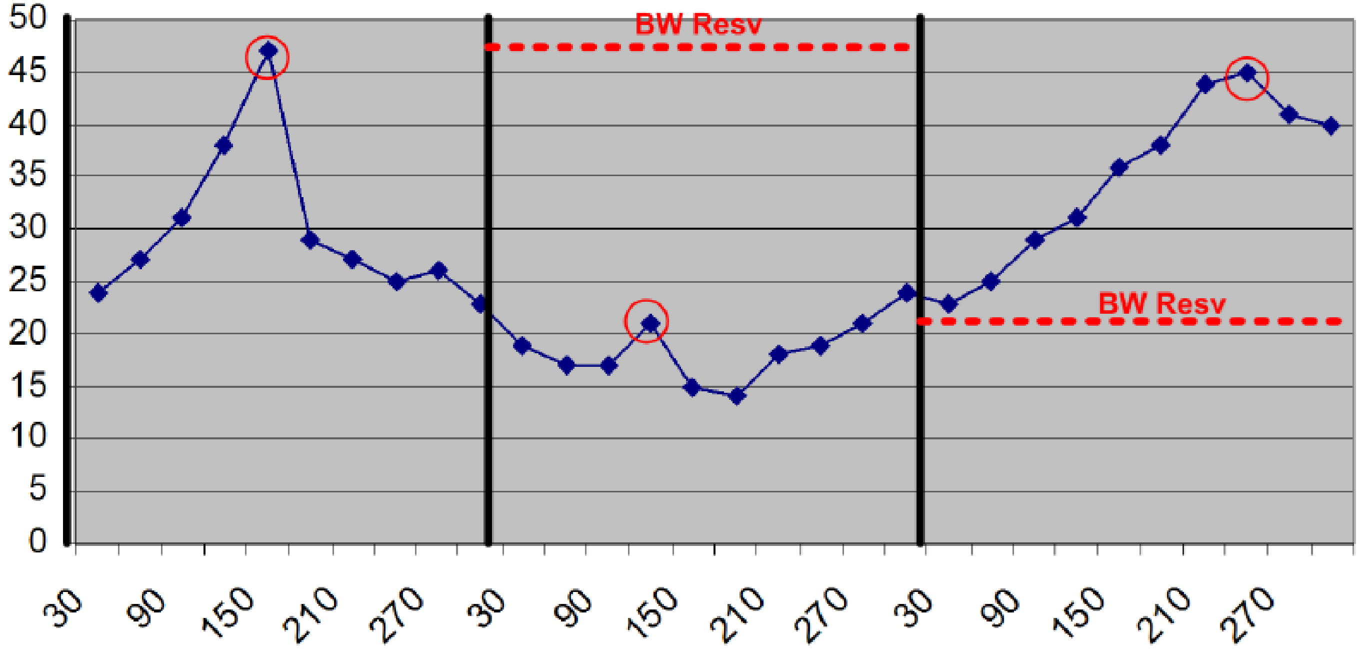
サービスとルーティングシステムの連携・自動化①



- ソフトバンク印南さん　JANOG25新潟
- http://www.janog.gr.jp/meeting/janog25/doc/janog25-rscript-00-intro-tinnami.pdf

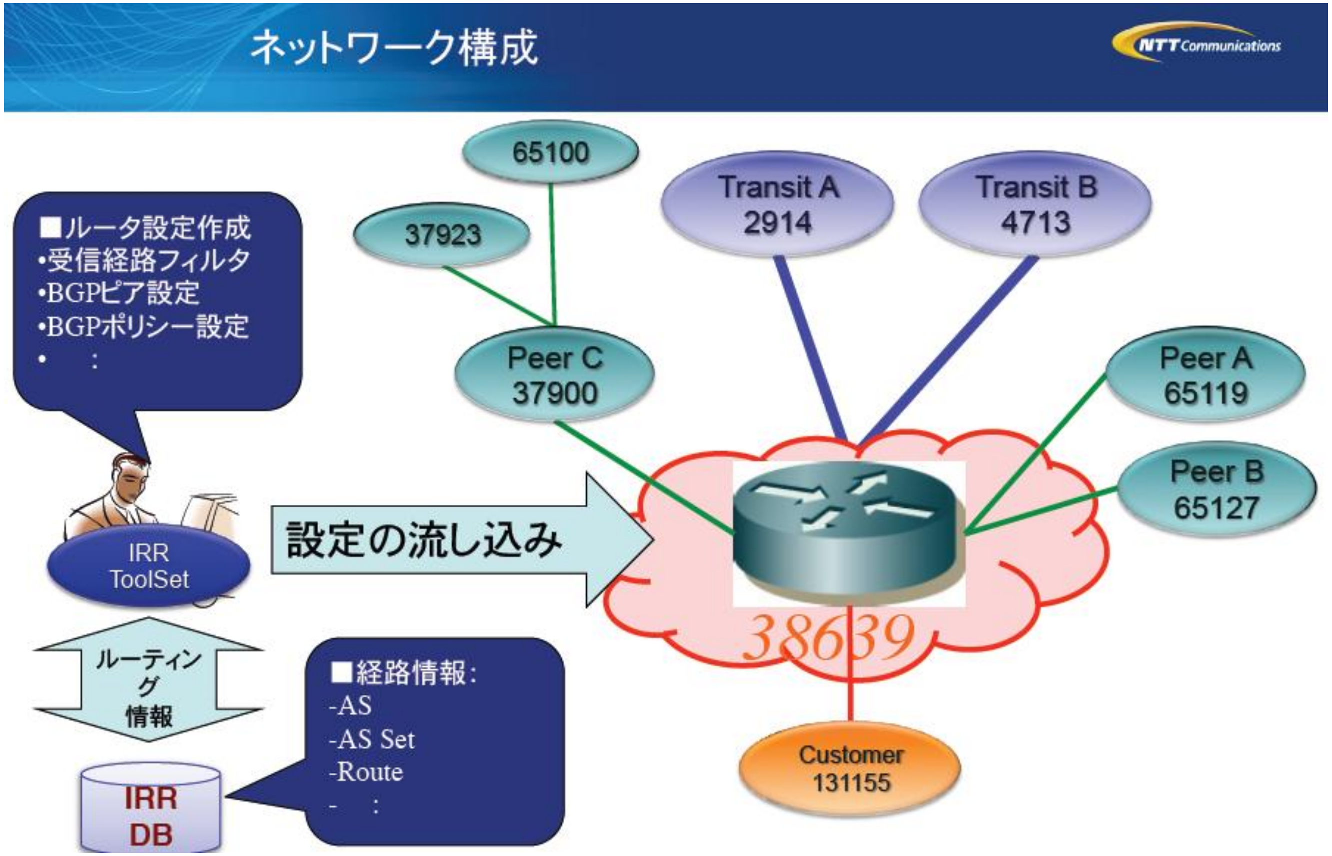
サービスとルーティングシステムの連携・自動化②

When Auto-Bandwidth Doesn't Work Well



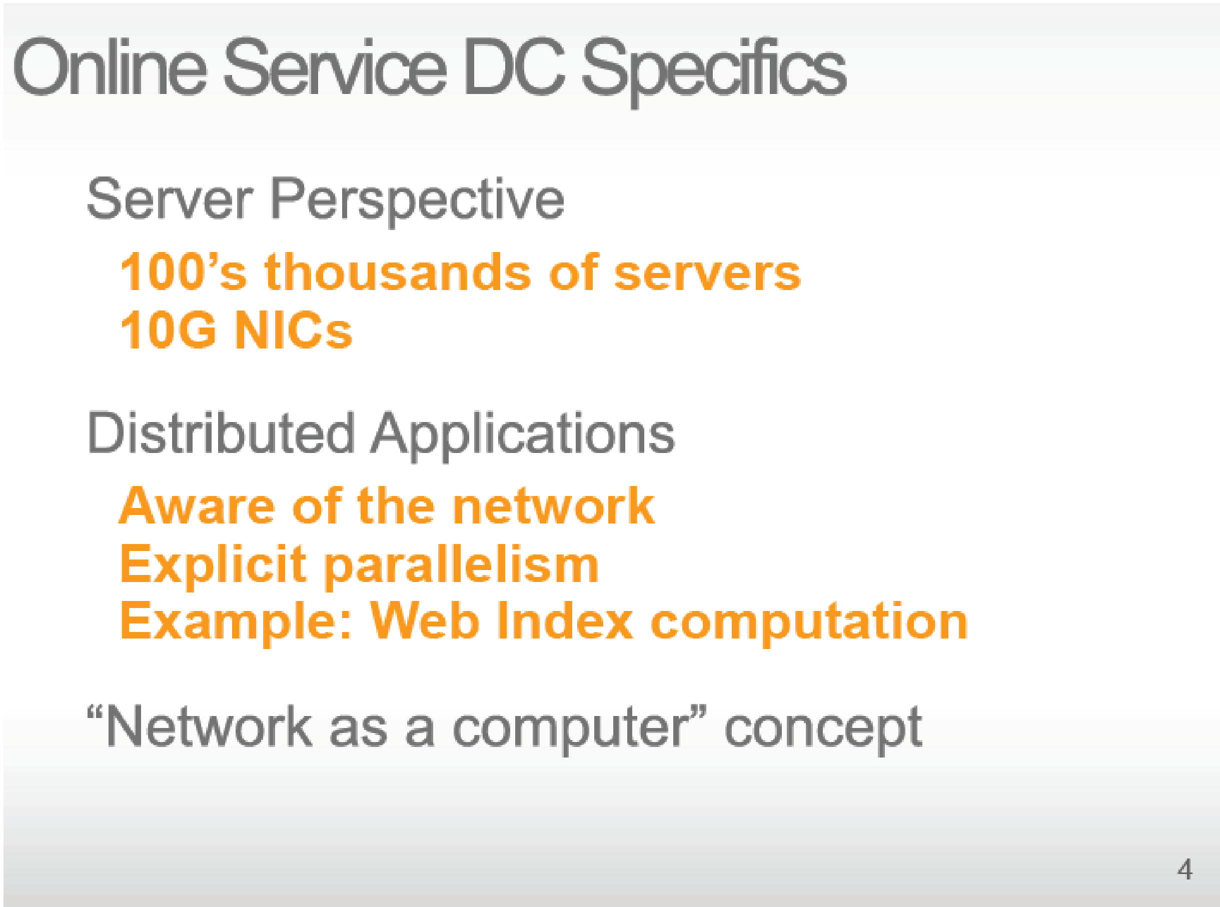
- Richard A Steenbergen,nLayer Communications,APNIC33
- http://meetings.apnic.net/_data/assets/pdf_file/0011/45596/MPLS-autobandwidth-ras.pdf
-

サービスとルーティングシステムの連携・自動化③



- NTTコミュニケーションズ水口さん
- http://www.janog.gr.jp/meeting/janog29/_downloads/janog29-auto-after-mizuguchi-01.pdf

サービスとルーティングシステムの連携・自動化④



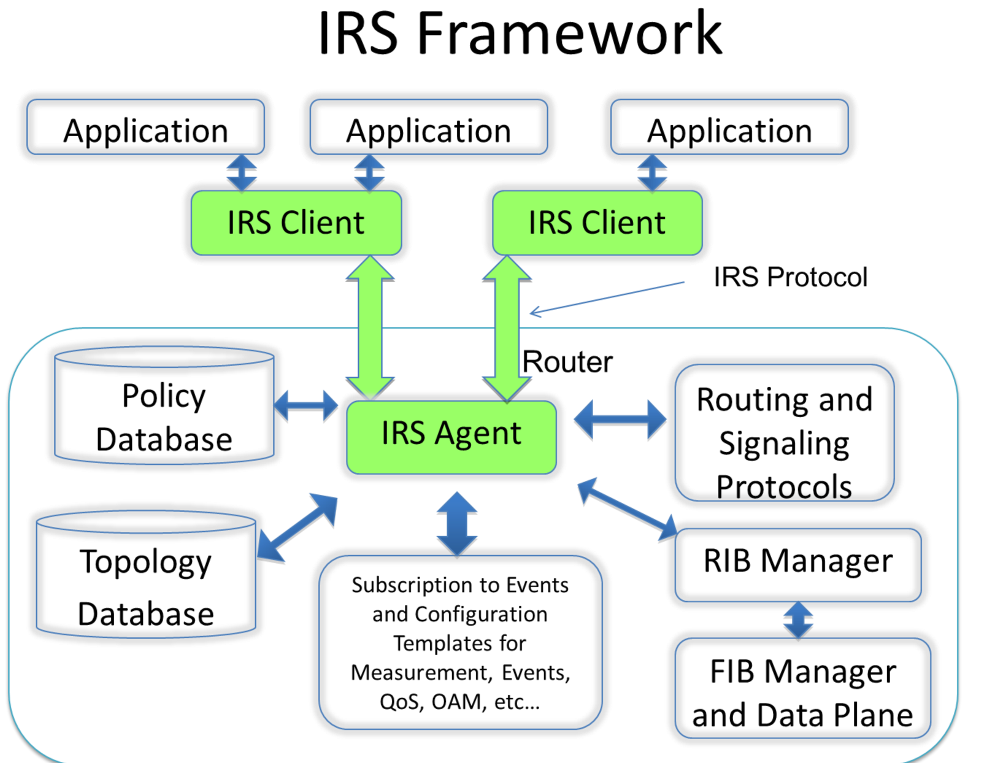
- NANOG55/IETF84 Petr Lapukhov, Microsoft
- http://www.ietf.org/proceedings/84/slides/slides-84-idr-1.pdf
- http://www.nanog.org/meetings/nanog55/presentations/Monday/Lapukhov.pdf

ルーティングシステムの連携のまとめ

- コンディションに応じたダイナミックなルーティング
- 使用帯域に応じたルーティングコントロール
- IRRデータベースとの連携
- アプリケーションに応じたトラフィックエンジニアリング

既存のルーティングプロトコルと外部データベース
や条件の連携により、目的を実現

IETF IRS



- IETFではIRS(Interface to the Routing System)でフレームワーク・Problem Statementを定義し議論中
  - http://tools.ietf.org/html/draft-ward-irs-framework
  - http://tools.ietf.org/html/draft-atlas-irs-problem-statement

IRSは

- ルーターが常に必要とする唯一無二の設定手法**では無い**
- 現在のルーティング/シグナリングプロトコルを直接的に置き換えるもの**ではない**
- 常に必要となるルータのデータとトポロジーを読む唯一無二の方法**ではない**
- 一つのルーティングデバイスに制限された物**ではない**

case study

- **draft-amante-irs-topology-use-cases-00**
  - Traffic Engineering, VPN Services Provisioning, Rapid IP & ASN Renumbering
- **draft-keyupdate-irs-bgp-usecases-01**
  - ASN migration scenarios; RT (RD) changes, BGP Route Manipulation, Notify applications when changes occur to “important routes”
- **draft-white-irs-use-case-00**
  - Optimized Exit Control, Reacting to DDoS attacks, Dynamically optimize traffic flows in a hub & spoke network, Inside DataCenter Routing
- **draft-medved-irs-topology-requirements-00**
  - Define standards-based data models with common vocabulary to describe various network components, MUST support publish/subscribe capability
- **draft-rfernando-irs-framework-requirement-00**
  - Publish-Subscribe for Asynchronous Notifications of Events

name?

- IRS Internal Revenue Service
- IRS Inter-Domain Routing Security for Japan

Interface to the Routing System (i2rs)

http://datatracker.ietf.org/wg/i2rs/charter/

(proposed WG)