

[招待講演] IX の運用管理技術

伊藤 大輔†

†NTT コミュニケーションズ株式会社ブロードバンド IP 事業部 〒100-8019 東京都千代田区内幸町 1-1-6
E-mail: †dai.ito@ntt.com .

あらまし 日本の商用 IX(Internet Exchange)の一つである JPNAP の運用者としての立場より、インターネット上の IX における運用管理技術を紹介する。

キーワード インターネット, IX, 運用管理

[Invited Talk] Operation and maintenance technology of Internet Exchange

Daisuke ITO†

† BBIP Division, NTT Communications Corporation 1-1-6 Uchisaiwai-cho, Chiyoda-ku, 100-8019 Japan
E-mail: †dai.ito@ntt.com

Abstract Introducing operation and maintenance technology of Internet Exchange from operator of JPNAP which is one of the largest IX in JAPAN.

Keyword Internet, IX, Operation, Maintenance

IXの運用管理技術

伊藤大輔
NTTコミュニケーションズ株式会社
(元インターネットマルチフィード株式会社)

Contents

- A. 本発表について
 - 概要
 - 自己紹介
 - IXとは
- B. IXの構成要素
- C. IXの運用管理技術

概要

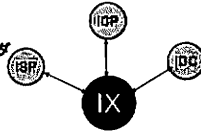
- 目的
 - インターネット上のIXという特殊なネットワークについて、その運用管理技術を紹介する
- 内容
 - 概要
 - IXの構成要素
 - IXの運用管理技術

自己紹介

- 前所属: インターネットマルチフィード株式会社
 - JPNAP開始当初からその設計に携わるとともに、2007年5月まで実務的な運用・管理を6年間担当
- インターネットマルチフィード社について
 - 1997年設立
 - 東京・大手町にてデータセンタを運営
 - 2001年に「JPNAPサービス」を開始
 - 商用インターネット相互接続(IX)サービス
 - 現在MAX120Gbps相当のTraficleを交換

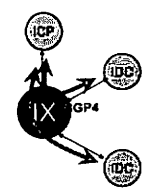
IXの特色 I

- IXとは
 - 複数のインターネットサービスプロバイダなどの事業者ネットワークを相互に接続するインターネット上の相互接続ポイント
- 接続事業者
 - インターネットサービスプロバイダ
 - データセンタ事業者
 - コンテンツ事業者
 - 学術ネットワーク
 - etc...



IXの特色 II

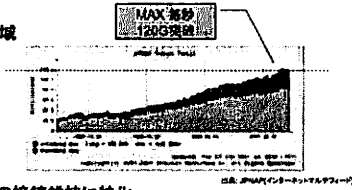
- IXでのネットワーク間接続プロトコル
 - BGP4が一般的
- IXネットワークの種別
 - L3ネットワーク (BGP Multi-Hop)
 - L2ネットワーク
 - MPLSなどの透過L2ネットワーク
 - Ethernet (Transition)



▶▶▶ 本発表ではIXをL2 Ethernetとして想定

IXの特色III

■ 高帯域



■ BGPの接続維持に特化

■ 高い信頼性の要求

- オンラインレードなど社会的インフラとしてのインターネットへの期待
- VOIP/ネットゲームなどTTLにシビアなアプリケーションの台頭

7

Contents

- A. 本発表について
- B. IXの構成要素
 - ネットワーク構成
 - Interface
- C. IXの運用管理技術

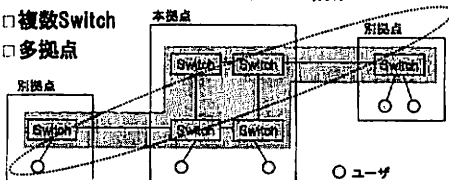
8

ネットワーク構成

■ すべて同一 L2 ネットワークにて構成

□ 複数Switch

□ 多拠点



IXそのものが一つのL2ネットワークとなる

9

Interface

■ Switch間Interface

- 10G Ethernet (IEEE802.3ae)では不足
- Link Aggregation (LAG)が基本

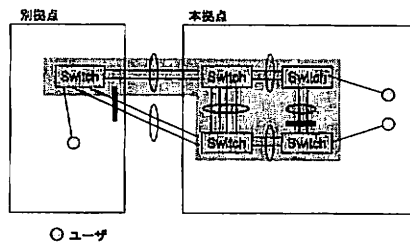


■ ユーザ向けInterface

- Fast/10/10G Ethernet
- 10Gでも不足なユーザはLink Aggregation接続

10

構成図(例)



すべてのユーザはL2で同一セグメントに存在する
⇒ 同じIPセグメントにBGP接続先がある

11

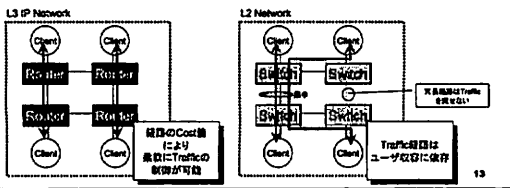
Contents

- A. 本発表について
- B. IXの構成要素
- C. IXの運用管理技術
 - Traffic Control
 - IXの監視
 - LAGの運用
 - 信頼性の確保
 - sFlow

12

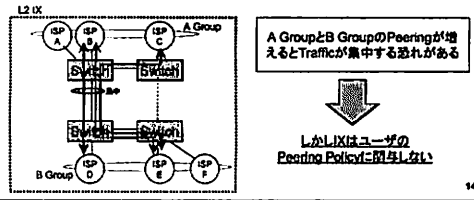
Traffic Control I

- L2ネットワークでは"cost"による経路選択が難しい
 - 特定の接続にTrafficが集中する可能性がある
 - ユーザ収容に大きく依存



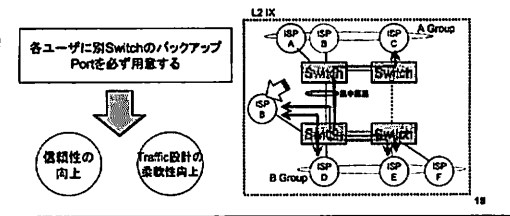
Traffic Control II

- TrafficはユーザのPeering Policyによって増減する



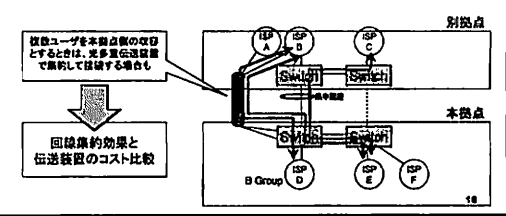
Traffic Control III

- ユーザ収容によるコントロール①
 - Portの二重化



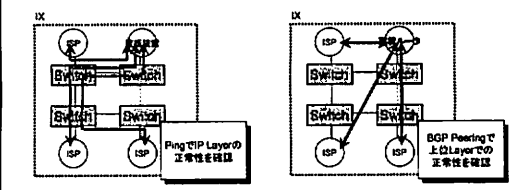
Traffic Control IV

- ユーザ収容によるコントロール②
 - 光多重伝送装置による収容



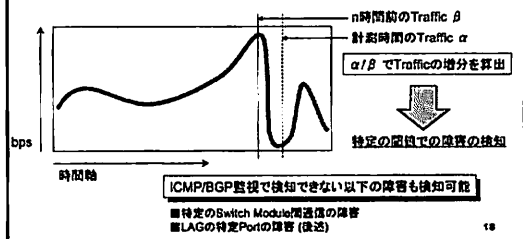
IXの監視 I

- ICMP(Ping)による監視
- BGP Peeringによる監視



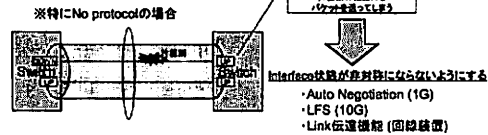
IXの監視 II

- Port Traffic計測による監視



LAGの運用 I

- プロトコルの選定
 - LACP (Link Aggregation Control Protocol)
 - No protocol
- 片線断への対処



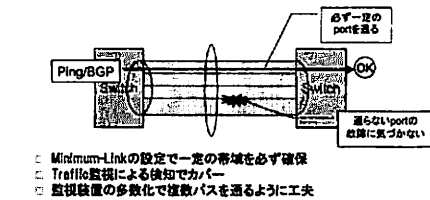
LAGの運用 II

- LAGのTraffic Control
 - Load Balancing 方式
 - L2 (Mac Addressベース)
 - L3以上 (IP, TCP/UDP Portなど)
 - 何経路に収束されるか
 - 8経路に収束される場合



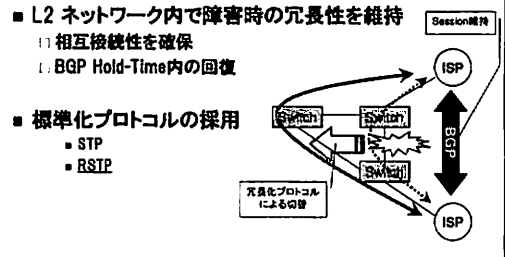
LAGの運用 III

- 特にno protocolの場合
- 監視用Ping/BGP
 - LB方式に基づき、LAGのうちのどこか1portを固定で通る
 - 他のLAG構成Portは監視できない



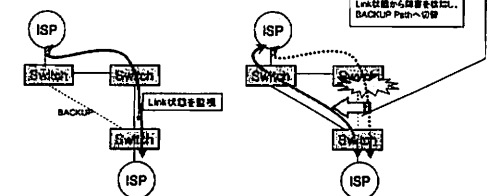
信頼性の確保 I

- L2 ネットワーク内で障害時の冗長性を維持
 - 相互接続性を確保
 - BGP Hold-Time内の回復
- 標準化プロトコルの採用
 - STP
 - RSTP



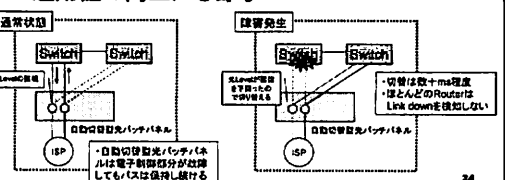
信頼性の確保 II

- Port Redundantの利用



信頼性の確保 III

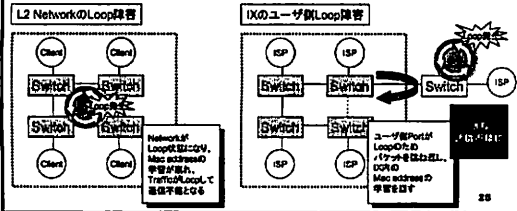
- 自動切替型光パッチパネルの導入
 - 信頼性の向上
- 運用性の向上にも寄与



信頼性の確保Ⅳ

Loop障害

- L2 接続のためユーザ側でのLoopにより全体に影響を受ける

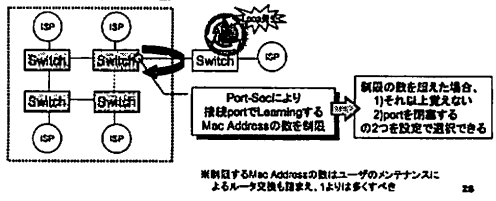


26

信頼性の確保Ⅴ

Port Security / Mac address Limitの導入

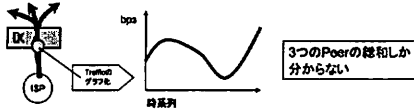
IXのユーザ側Loop障害の防衛



28

sFlow I

- Port InterfaceではユーザのTotalのTrafficしか計測できない



- sFlowを利用することにより各Peer毎のTrafficが取得可能



27

sFlow II

sFlow概要

- RFC3176
- Switchやルータにて通過パケットをsampling抽出し、sFlow-collectorにexportする
- sFlow-collectorにて統計処理を行う

sFlow使用例

- MACアドレスベースでのTraffic流量の可視化
- 商用IXではPeer毎のTraffic可視化サービスとして提供しているところも

28

おわりに

今後、IXの運用管理で解決すべきこと

- 障害時の切り分け能力の向上
 - Broadcast-domainの広域化に伴った、より精度の高い障害切り分け手法
 - L3以上のプロトコル(OAM/BGP)に頼らない監視方法
 - Traffic監視は一つの解
- Traffic Control (輻輳回避)をより容易に

一つの解

- sIx
- EtherOEM
- MPLS-IX

29