Generalized Router Configuration in IX

日本インターネットエクスチェンジ株式会社
平尾　利崇
hirao@jpix.ad.jp

2005年4月8日

前置き

具体的な設定についての前に・・・

IXの定義

- JPNICによるIXの定義
  - インターネットエクスチェンジポイント(IXP)は、物理的なネットワークインフラストラクチャーであり、独立したISP間でのインターネットトラフィックの交換を円滑化するために運用される。
  - IXPに接続されるISP数は最低でも3つあるべきで、他のISPが参加するための明確でオープンなポリシーがなければならない。
- この資料でいうIXとは・・・・
  - スイッチネットワークにより拡張された単一のブロードキャストドメイン
  - いわゆる「レイヤ2 IX」をさす。

IXの提供するサービス

- もともとL2での接続性を保証するサービス
  - IX上ではIP/BGPで経路およびトラフィックを交換するとしているもののその上位レイヤに関しては特に制約を設けていなかった。

→機器固有の設定だけでなく、各社同士のポリシーにも依存するため、関与するのが難しかった。

IX接続に関する推奨設定について（その1）

- 時代は変わって・・・・
  - IXとしての、さまざまな経験の蓄積から問題点が整理されてきた。
  - IXに対する注目度、依存度も上がってきた。

  → 意識の変化

- IX接続の際のISP側機器に推奨設定を提示するようになった
  - 必要のないトラフィックは流さない
  - トラブル発生時の影響を小さくする

IX接続に関する推奨設定について（その2）

- 欧州のIXでは、「許可するトラフィック」を規定している例も・・・・。
  - AMS-IX（オランダ）
    - 「Allowed Traffic Types on Unicast Peering LANs」(http://www.ams-ix.net/technical/allowed.html)
    - 条件を満たさない場合は、「Quarantine VLAN」という別VLANに収容
  - LINX（イギリス）
    - 「Memorandum of Understanding(APPENDIX 1)」(http://www.linx.net/joining/mou.thtml#_Toc15876470)

本題

IX接続に関する設定について

MACアドレス・IPアドレス

- 1ポートにつき
  - 1つのMACアドレス
    - IXによってはSrc MACアドレスによりフィルタを行う場合も。
  - 1つのIPアドレス
    - IX事業者の管理するIPアドレスから割当て
- ただし、最近では Link Aggregation によるサービス提供もあるので・・。

イーサフレーム

- イーサタイプ
  - 0x0806 ARP
  - 0x0800 IPv4
  - 0x86DD IPv6
- 宛先MACアドレス
  - BroadcastパケットはARP/NDのみ許可
  - Multicastパケットは禁止
    - インタフェースのエラーカウンタがアップするケースがある。
      - Juniper のDiscard パケットカウンタ
      - Catalyst6500のInputエラーカウント etc…

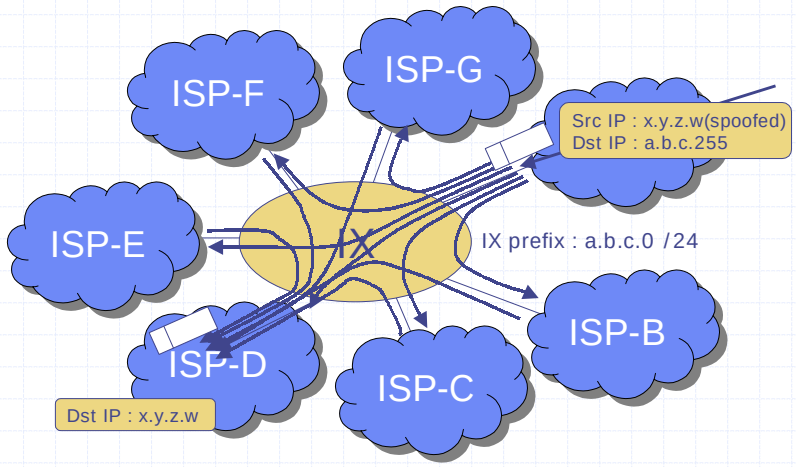
リンクローカルトラフィックの転送禁止

- 以下に限らず、不要なものは停止
  - IRDP
  - IEEE 802 Spanning Tree
  - Vendor proprietary discovery protocols (e.g. CDP, EDP)
  - Interior routing protocol broadcasts (e.g. OSPF, ISIS, IGRP, EIGRP)
  - BOOTP/DHCP
  - PIM-SM/PIM-DM/DVMRP
  - Keepalive
  - VTP
  - DEC MOP etc.….

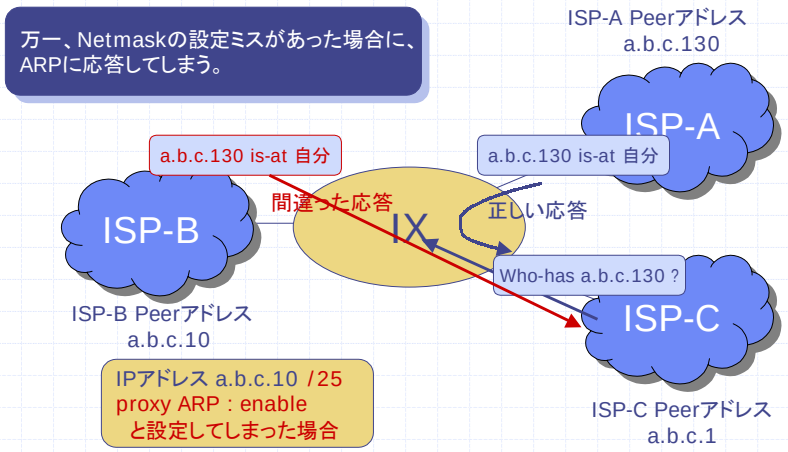
インタフェースの設定

- IXセグメントのブロードキャスト宛の Directed Broadcast の停止
  - Smurf攻撃の防止
- ICMP redirect の転送停止
- Proxy ARP機能の停止
  - ネットマスクの設定ミスをすると他人宛のARP リクエストに応答。

Directed Broadcast によるSmurf攻撃



Proxy ARP設定による問題



ルーティング関連（その1）

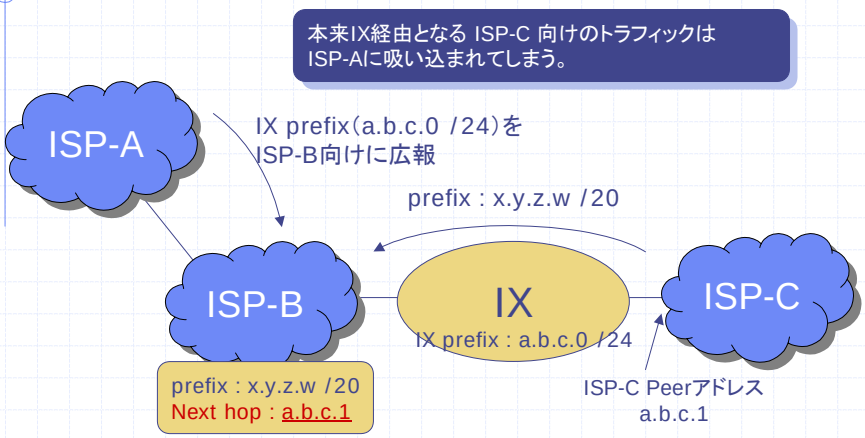
- IXセグメントのアドレス空間は他のASに広報しない
  - JPNICでは、IXセグメント用IPアドレスの割当てを行っているが、その割当てに関して、以下の制限を設けている。

    「この条件下で行われたすべての割り当てにおける特別な制約として、そのIXPはグローバルなインターネット経路表に、そのアドレスを広告してはならない。」
- 自衛手段としては、AS間ボーダの inbound において、IX prefix を filter する。
  - （ただし、IX運営者自身がIXセグメントのアドレス空間「を含む」prefixを広報していることが多いので注意）

ルーティング関連（その2）

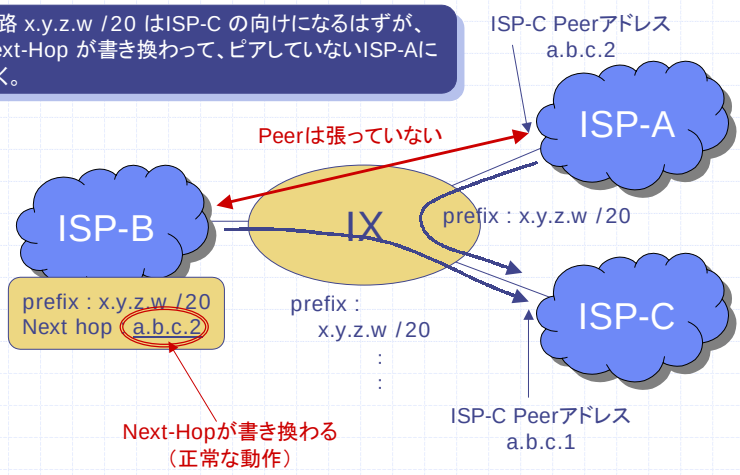
- Next-Hop-Selfの利用
  - 隣接ASからのルーティング情報をIBGPピアに広報する場合には、Next-Hop-Self を使用する。
  - AS同士のEBGPピアについてもNext-Hop-Selfを使用する（要検討？？）
    - MA（多重アクセス）メディアの動作によるNext-Hop書き換え問題の防止策

ルーティング問題（その1）



ルーティング問題（その2）

経路 x.y.z.w /20 はISP-C の向けになるはずが、Next-Hop が書き換わって、ピアしていないISP-Aに向く。



BGP脆弱性対策

- 顧客収容ルータでingress フィルタ
- IX接続ルータでegressフィルタ
  - Src IP : any
  - Dst IP : [IX prefix]
  - Protocol : TCP
  - Port : BGP

- フィルターなので、馬渡さんネタ？？

その他

- 不要なPeerについては（こまめに）削除
  - BGPピアの確立のため、周期的にARPのパケットを送出してしまう。まったくトラフィックを流していない機器でも数kbpsを受信(JPIXの場合)。
  - トラフィックをキャプチャし、ARPリクエストを解析することにより、IX側では確認が可能。ISPに対してレポートしているケースも。