

A collection of various medals and a compass are arranged on a wooden surface. The medals include a red ribbon medal with a circular emblem, a white star-shaped medal with a central emblem, and a blue ribbon medal with a circular emblem. A pair of glasses with thin frames and a compass are also visible. The background is a light-colored, textured surface.

広域イーサネット網を利用した 分散IXに関する概況

エスアールエス・さくらインターネット(株)
代表取締役 田中 邦裕

今回はまじめに……

今回の趣旨

- ◆ 広域イーサは良く聞くけれど、とても大きなL2スイッチとして利用できるのか？
- ◆ 地域IXが最近流行っているけれど、中央のコンテンツなしでどのように生き延びるのか？
- ◆ 分散IX (広域IX ?) は成立するのか？

これらを、冗談はほどほどにして解説します。

広域イーサで広域IXを

◆ 分散IXとは……当社の定義

全国どこからでも接続可能なIXサービス

レイヤ2のものとレイヤ3のものの2種類がある？

いっばいつながれば、それが広域IX

◆ レイヤ2かレイヤ3か

当社のものは広域イーサ網を利用したレイヤ2構成で、
加入者同士の協議により直接のピアリングが可能。

というか、そもそもL3のIXというのはどういうものか？

昔MEXさんにて提供されていたスタイルか？

L2網のデメリット

- ◆ WANを構築するにあたりL3よりL2の方がややこしい
(とされている)
このへんは、キャリアに頑張ってもらおう
- ◆ キャリアのL2網を利用した場合には、自社で加入者間のトラフィック交換状況を確認できない
キャリアより各地のポート毎のトラフィックをもらう
ユーザ局舎内に計測器を設置させて頂く
- ◆ 加入者が増えるとブロードキャストが増える
そもそも、同一地域内で加入者が多くなる場合は、
地域IXに頑張っていたかく



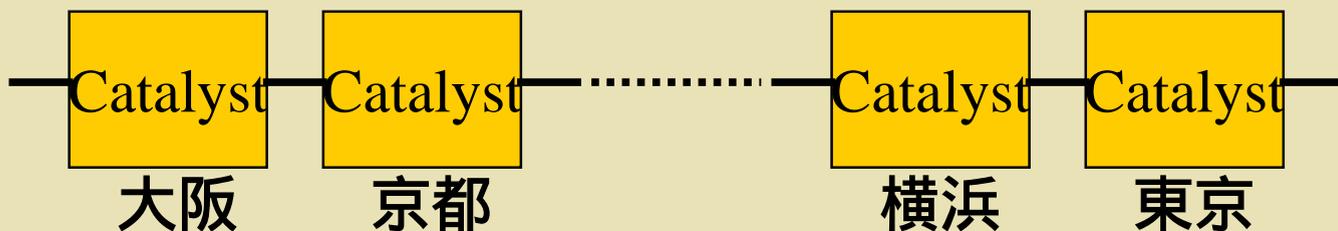
L2網にするメリット

- ◆ 加入者同士のピアリングが容易
- ◆ 「一般的」にL3よりL2の方が高性能
- ◆ ルータの台数を減らすことができる
- ◆ 東京・大阪の両方の境界ルータとピアリングが可能で、平均のRTTが小さくなる
- ◆ HOP数が少ないのでちょっと知ってるエンドユーザが満足する ;)

自社で作るかキャリアに頼るか

- ◆ 自社でつくるのは大変

Catalystを延々接続するの？



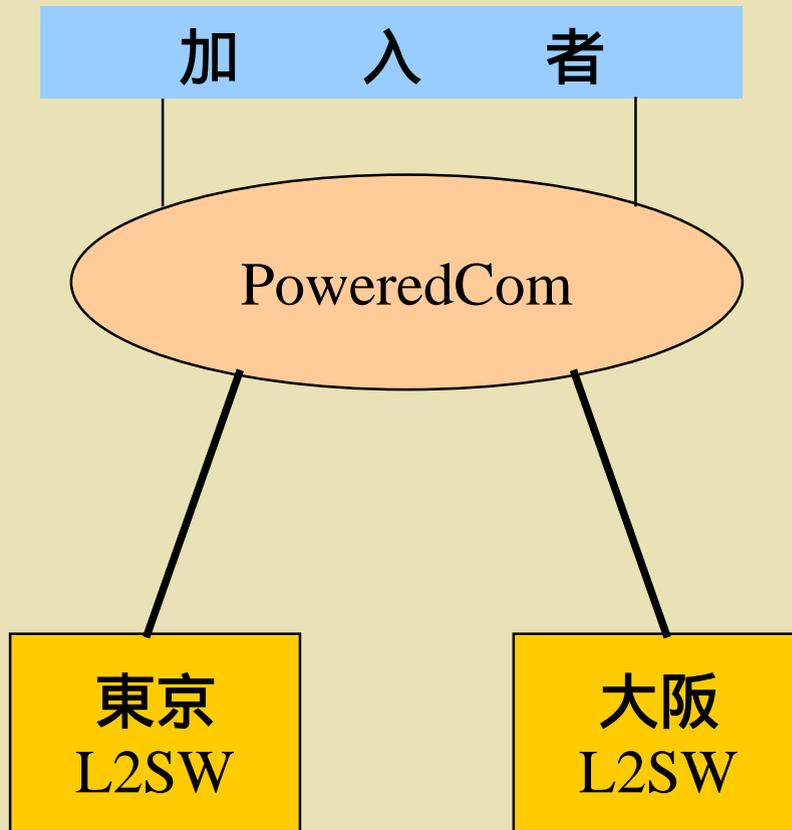
とてもスケールしない

L2 over L3でやるの？



なら、L3だけでいいんじゃない？

でも、やっぱりキャリアに頼ろう



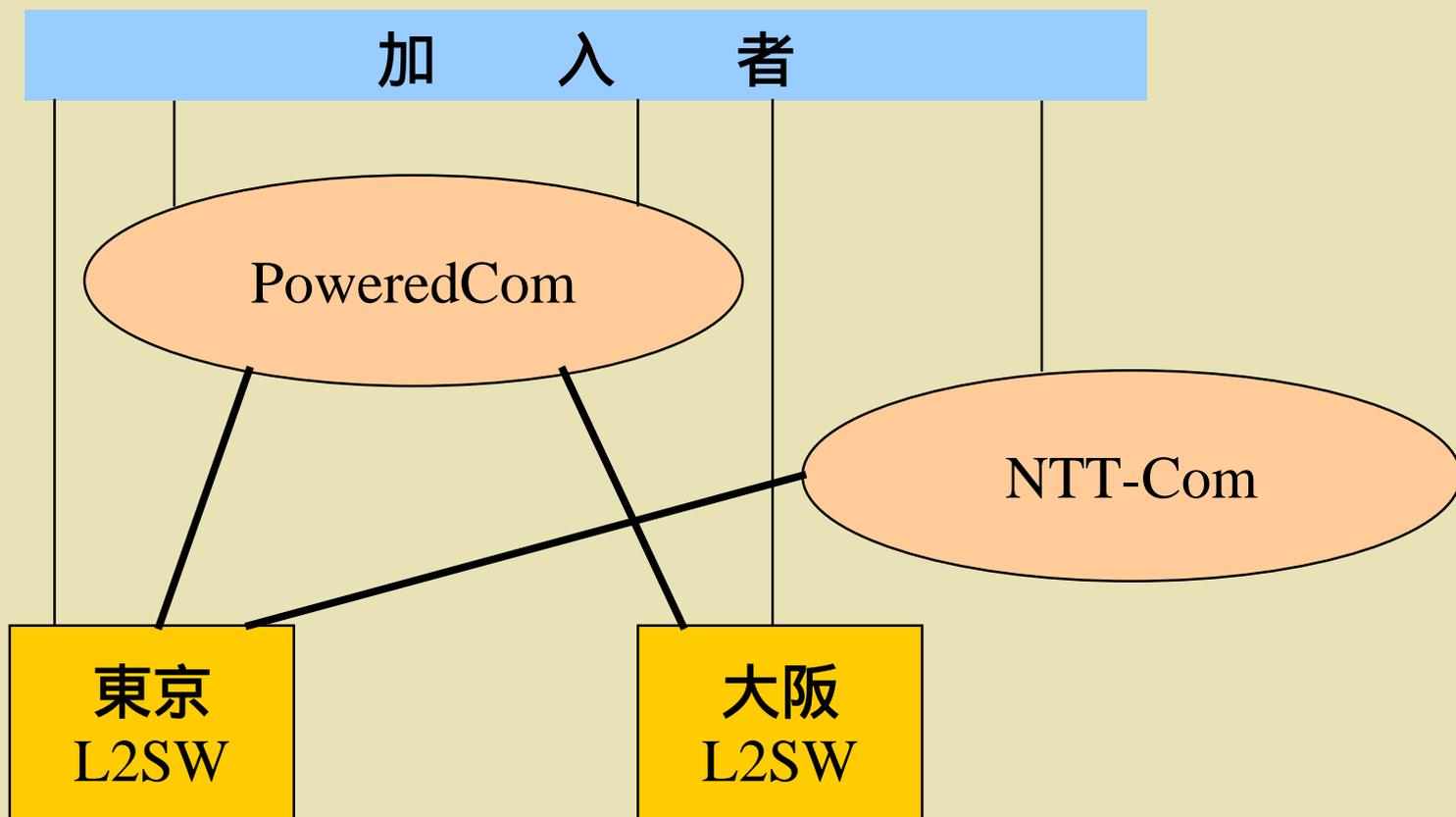
でも、大きな問題が…

PoweredComの回線が
NTT-POI局に入らない
問題発生！

最初はいける
ゆうたやないか…

オマケにTT/CTC/OMP
以外はすごく弱かった

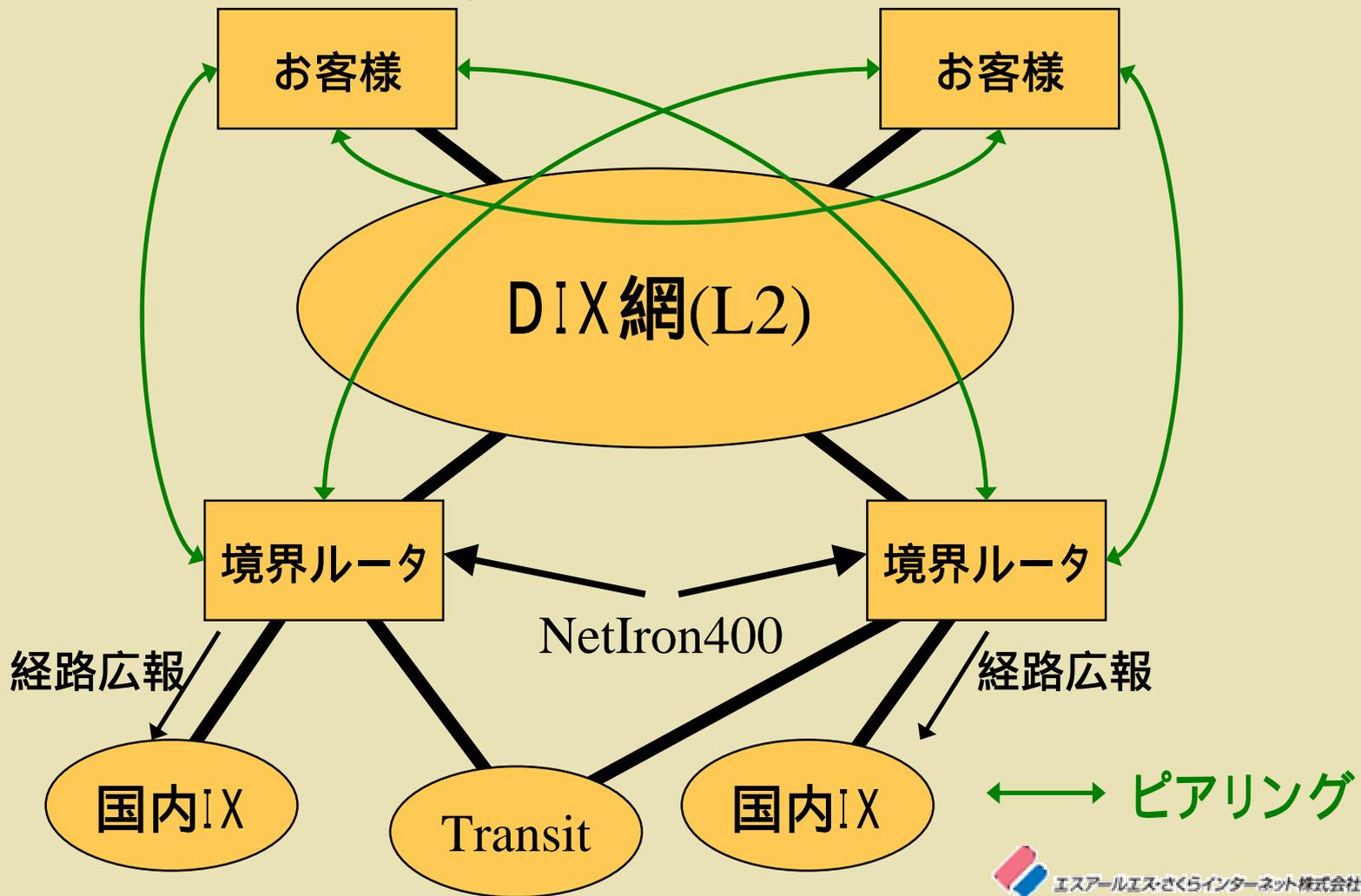
結局、L2サービスの併用に



この上でSTPを使ったらどうなるのかしらねえ(笑)
大きなスイッチとして使えるのか否か...

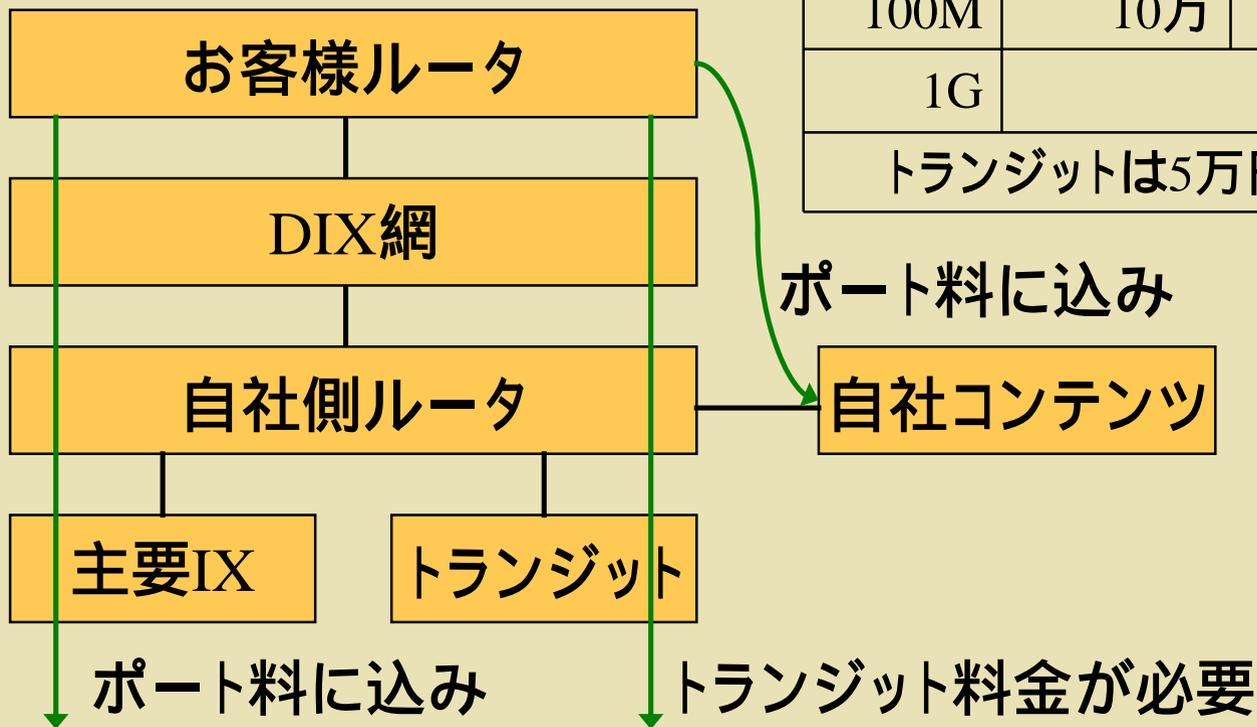
DIXの仕組み

CISCO/Foundry/Riverstone/Hitachi/Zebra/Gated



DIXの仕組み(2)

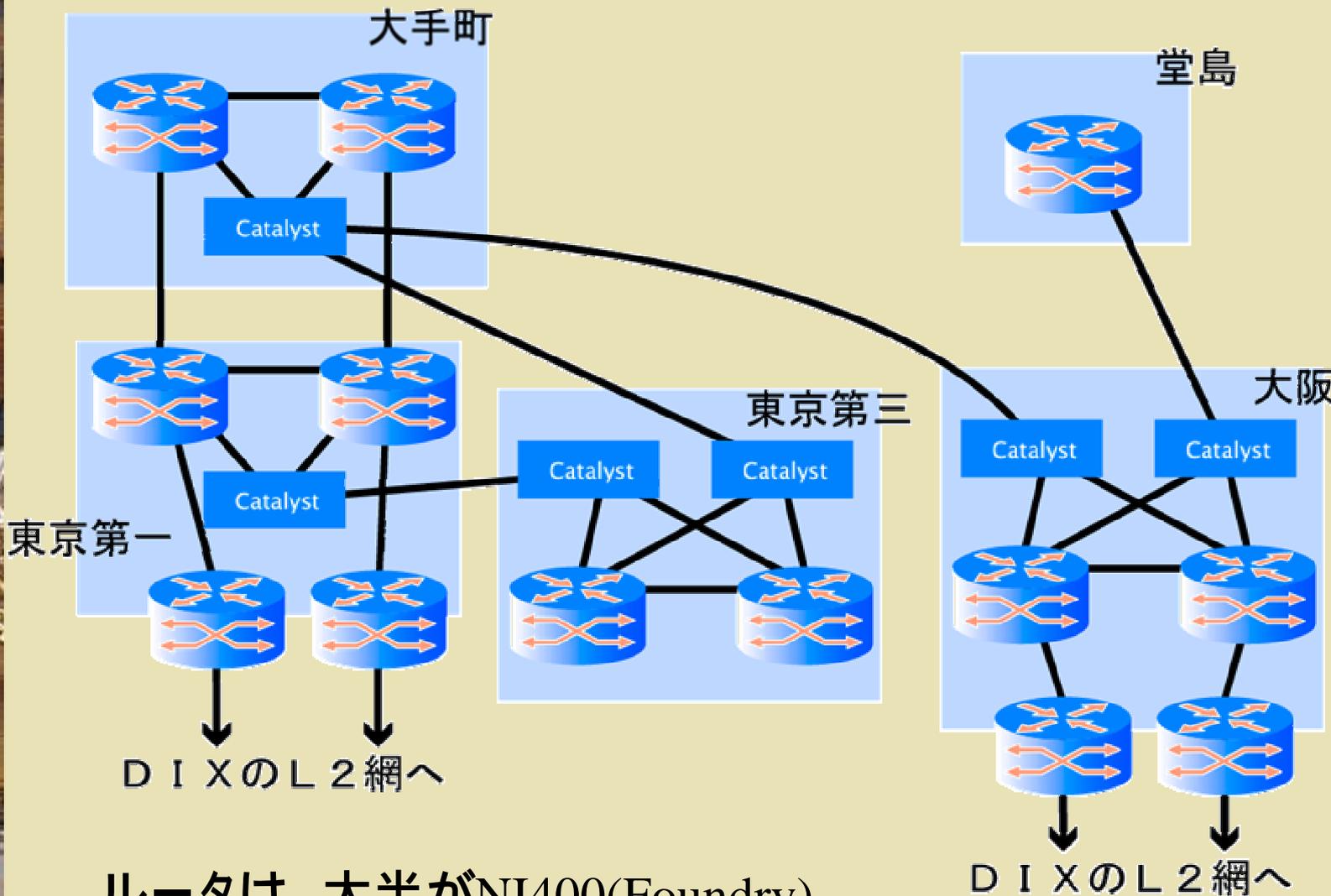
トラフィックの流れとコスト



速度	ポート料	回線料金
10M	3万	20万
100M	10万	70万
1G	要見積り	
トランジットは5万円/1M(定価)		

←→ トラフィック

バックボーンの構成



ルータは、大半がNI400(Foundry)
一部が、RS8000(Riverstone)

広域IXとJPIX/NSPIXP

- ◆ 東京・大阪以外のISP/CATV事業者が、JPIXやNSPIXPに接続する例も少なくない

ブロードバンド化に伴って、専用線を中央のIXに引っ張ってきた

しかし専用線代が安くないし、ピアリングがOKとも限らない

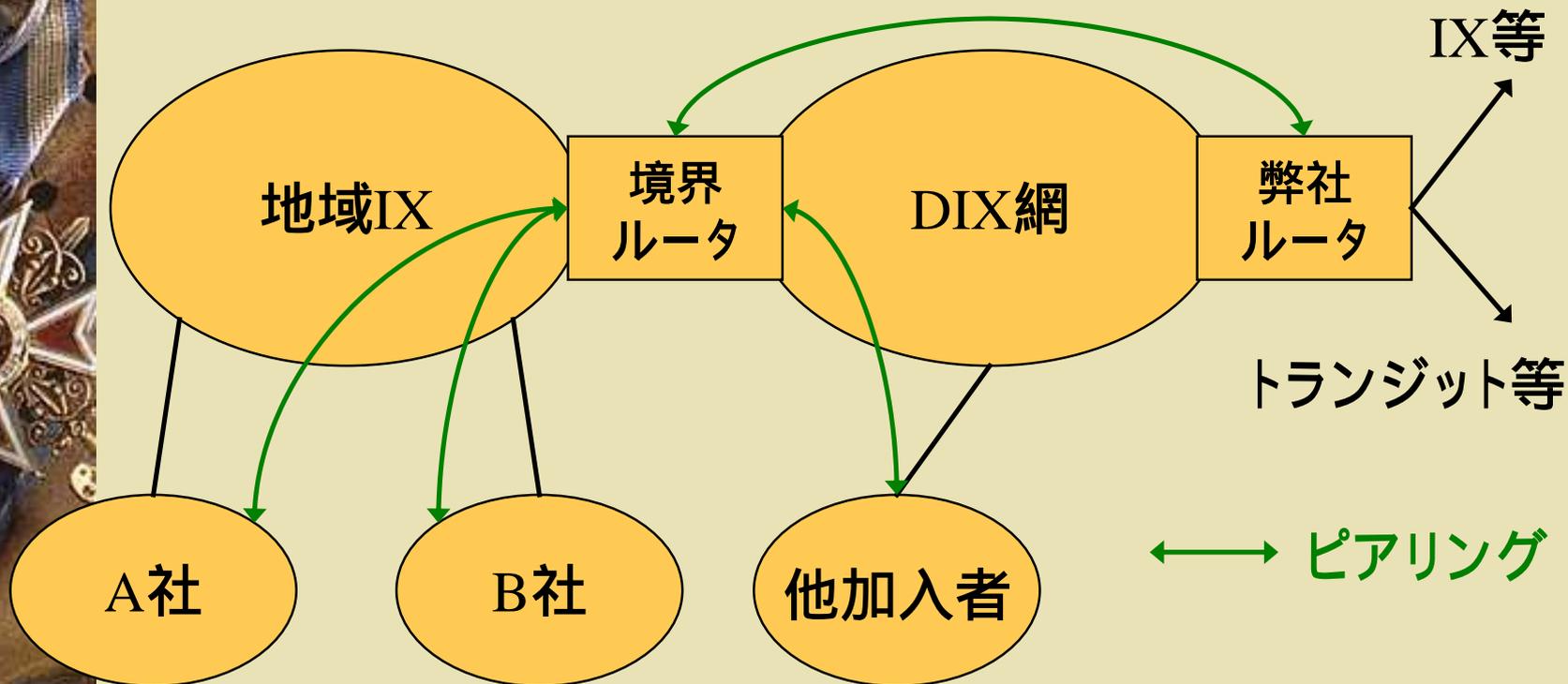
データセンターとしても、増え続けるIXでのピアリングの本数に困っている

広域IXとアプリケーション

- ◆ 例えば某匿名掲示板のトラフィック等を配信
大手一次プロバイダでも提供できないコンテンツを
安価に配信。
- ◆ コンテンツプロバイダを無償でハウジング
大手コンテンツ会社のハウジングを無償で受託。
トラフィックはDIX加入ISPがポート料として「購入」する。
- ◆ 錬金術でトラフィック自体をお金に
今まではお金でトラフィックを流してもらう構図。
これを、トラフィックをお金で売る構図に変化できるかも

地域IXとはどう関係を持つか

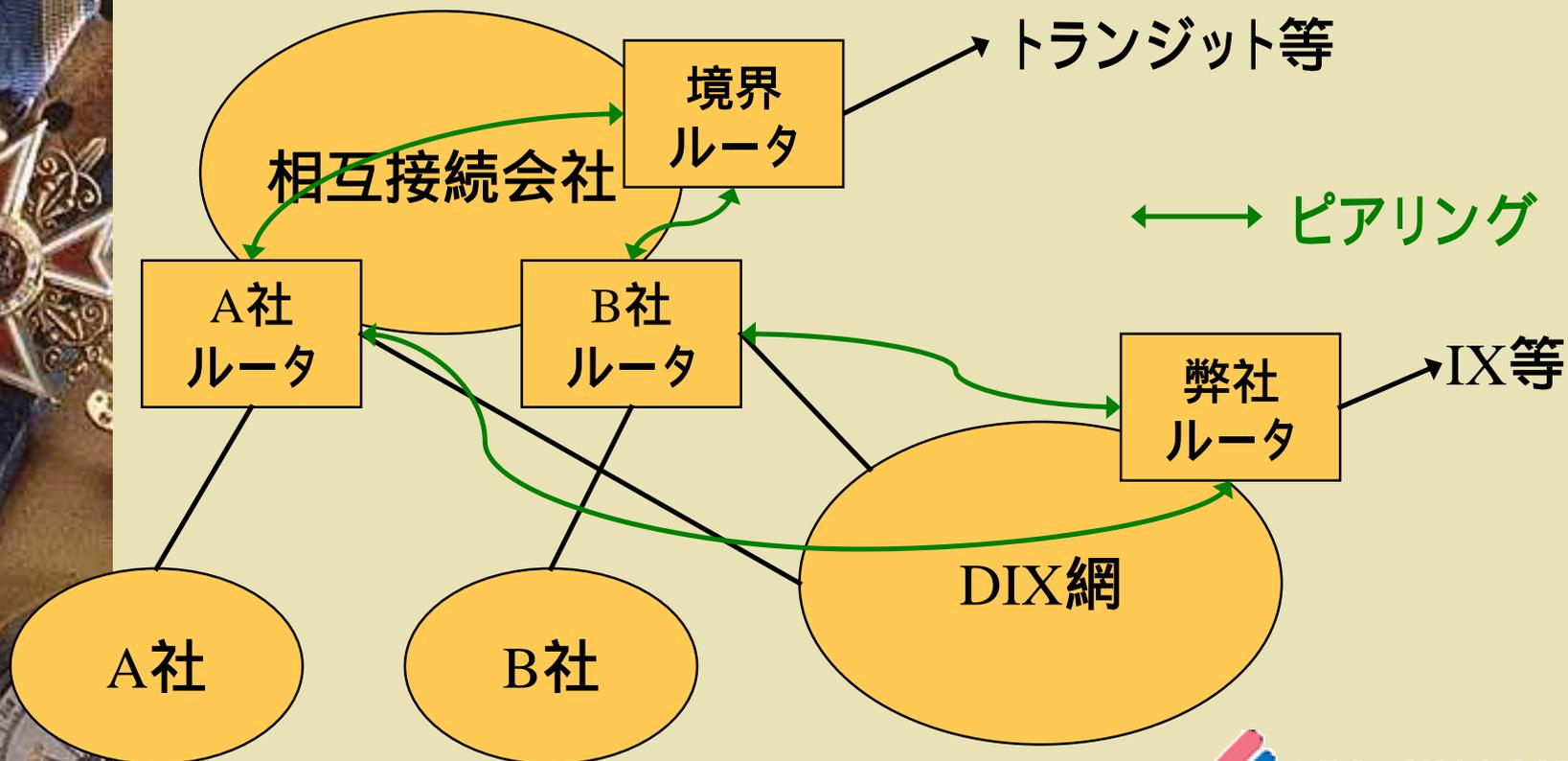
- ◆ DIXの加入者である地域IX様の実例



BGPをしゃべる事業者とそうでない事業者がごった煮なので、/24とかの経路が流れることも…

CATV網との相互接続

CATV事業者では、複数の事業者でバックボーンを運営する会社を持っていることが多いため、そこへ接続を行う場合も、下記の例では、IXのトラフィックの効率化がメインとなる。



どれだけ流れるか

- ◆ CATVや地域IXの場合にはマルチホームが多い

無調整の場合には、IXからのトラフィックもトランジットISPを経由する例が少なくない 対処が大変！

対策1:ピア先に頼んで優先度を上げてもらう

対策2:加入者から契約トランジットISPにゴルフ電してもらう

対策3:トランジット側にプリペンドを行う

最終的には、1:1くらいでトランジットとDIX側のトラフィックが分散する

一部のお客様すいません…。

トラフィック課金をどうするか

- ◆ DIXの場合、トランジットとIXのトラフィックを切り分けるため、別個の課金サーバが必要

ケン: S S Rには、LFAPがあるんじゃない

ボブ: まさか、また良く落ちるネットワークにするのかい？

ケン: なら、FundryのsFlowにしたらどう

ボブ: いやーきみの意見はいつも傑作だね。あれは、JetCoreの搭載されたNetIronしか無理なんだよ。

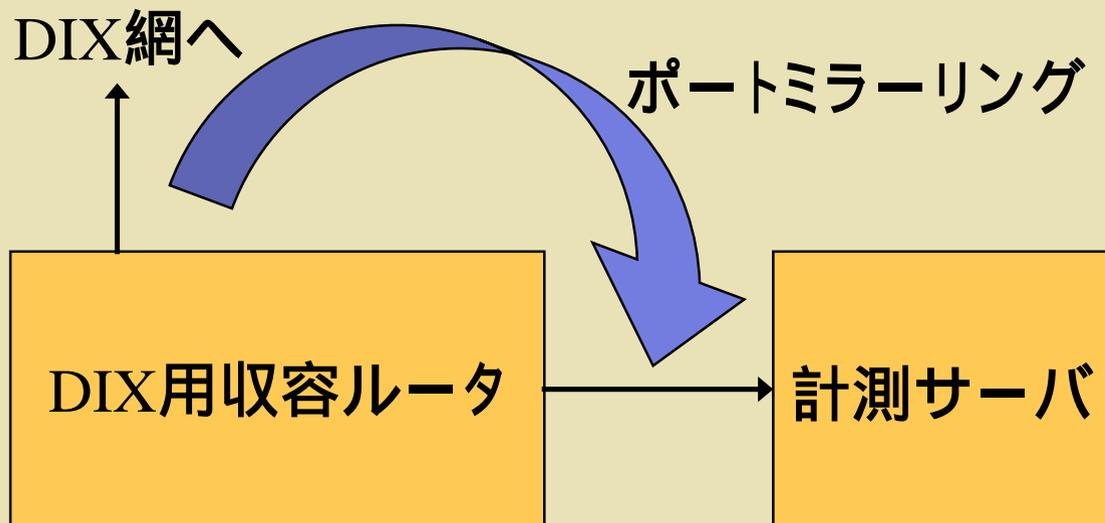
誰かが失敗した頃に始めるのが吉だろうね。

ケン: じゃあ、CISCOのGSRを買ってNetFlowにしよう

ボブ: ……

ボブ: それは無理だよ。

トラフィック課金の方法



全てのIPパケットのSRC/DEST/Sizeのみを抽出しリアルタイムに解析する。
そして、5分に1回RRDを利用してトラフィックを保存

かなりの荒技。でも超高性能な計測サーバを入れても、ソフトを買うより安い；)

トラフィック課金の方法(2)

ちょっと技術的な話

トラフィックはIPヘッダを解析してます

Ver	LEN	TOS	パケット長
識別番号		フラグメントオフセット	
TTL	プロトコル	ヘッダチェックサム	
送信元IPアドレス			
送信先IPアドレス			
オプション			

この部分でどの加入者なのかを判別

この部分でトランジットか否かを判別

パケット長はIPヘッダのLENを使ってはならない
理由は簡単 ;)。今回はbpfのlenを流用

そろそろまとめ

- ◆ 広域イーサは良く聞くけれど、とても大きなL2スイッチとして利用できるのか？

自社でやるのは大変だけれど、キャリアに頼って

- ◆ 地域IXが最近流行っているけれど、中央のコンテンツなしでどのように生き延びるのか？

地域IX同士を広域IXで接続して、且つJPIX等への経路があれば生き残れるかも…

- ◆ 広域IXは成立するのか？

一次プロバイダの接続料が下がらない以上、広域IXは重要なソリューションになるでしょう ;)

ありがとうございました