

# Prefix増大とルータの リソース問題

さくらインターネット(株) 研究所  
大久保 修一 [ohkubo@sakura.ad.jp](mailto:ohkubo@sakura.ad.jp)



# 本日のAgenda

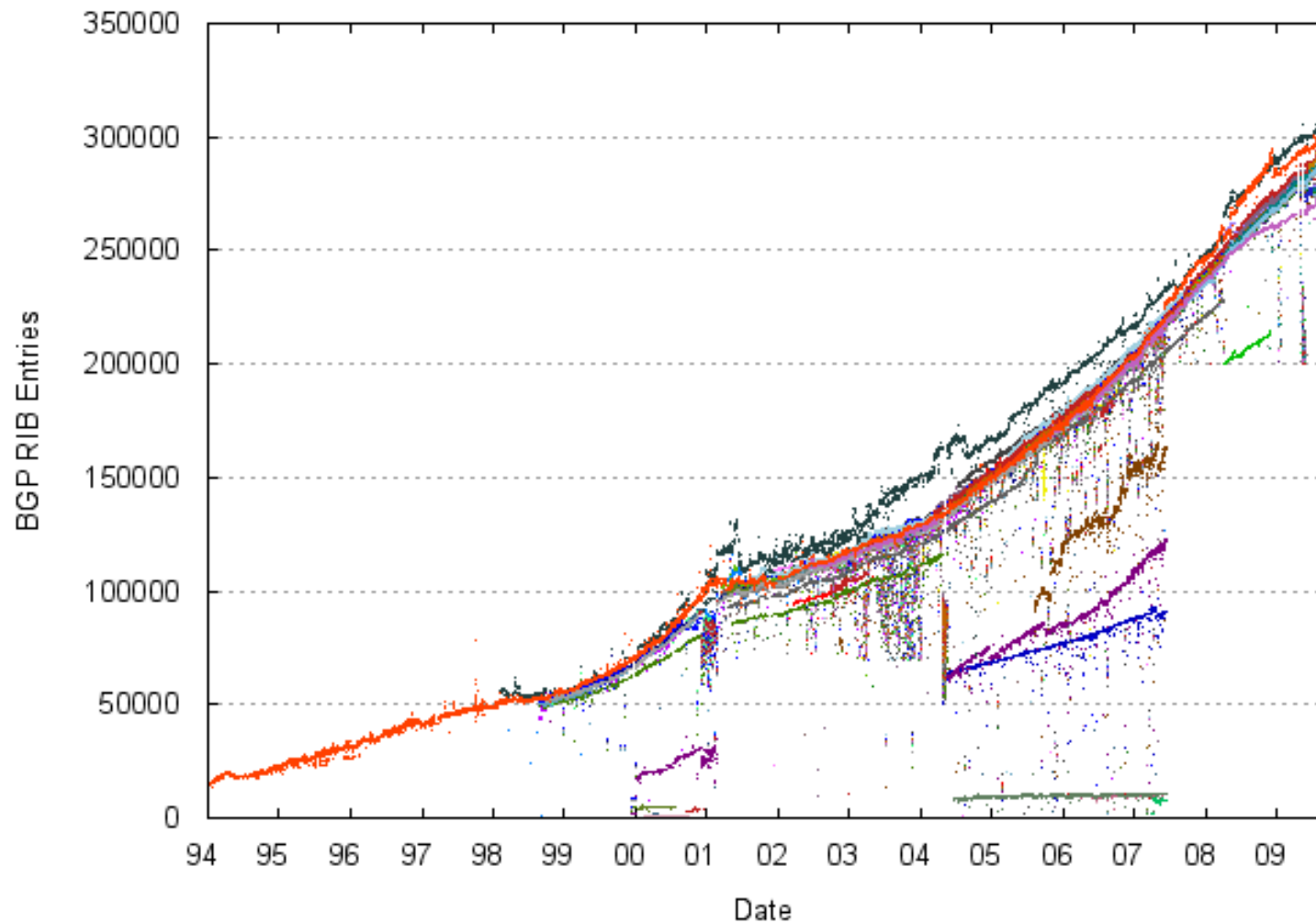
- 自己紹介
- インターネットフルルート増加の状況
- 経路数増大による悪影響
- デフォルトルート活用の検討
- 議論



# 自己紹介

- 2009年7月より「さくらインターネット研究所」が発足、異動になりました。  
(前部署:技術部ネットワークチーム)
- 研究所の目的
  - 弊社Webページより引用  
[http://www.sakura.ad.jp/press/2009/0630\\_sakura\\_randd.html](http://www.sakura.ad.jp/press/2009/0630_sakura_randd.html)
  - 一般に普及することを目的としたインターネット技術の基礎研究と情報発信
  - 新規事業への展開や既存事業の更なる発展を目的とした応用研究
  - 先端技術の発展や次世代エンジニアの育成を目指す産学共同研究
- 今後ともよろしく申し上げます。

# インターネットフルルート数の伸び



2009年9月7日現在

<http://bgp.potaroo.net/> より

# インターネットフルルート数の伸び

## ■ 2009年9月7日現在、約29万経路

```
# show ip bgp sum
```

```
BGP router identifier xx.xx.xx.xx, local AS number 9370
```

```
RIB entries 539569, using 33 MiB of memory
```


```
Peers 5, using 12 KiB of memory
```

```
Peer groups 1, using 16 bytes of memory
```

Neighbor	V	AS	MsgRcvd	MsgSent	TbIVer	InQ	OutQ	Up/Down	State/PfxRcd
xx.xx.xx.xx	4	9370	7195651	129010	0	0	0	06w2d20h	290399
xx.xx.xx.xx	4	9370	7194603	129010	0	0	0	06w2d20h	290399

## ■ 直近6ヶ月の伸び

□ 約2,000経路/月、約24,000経路/年



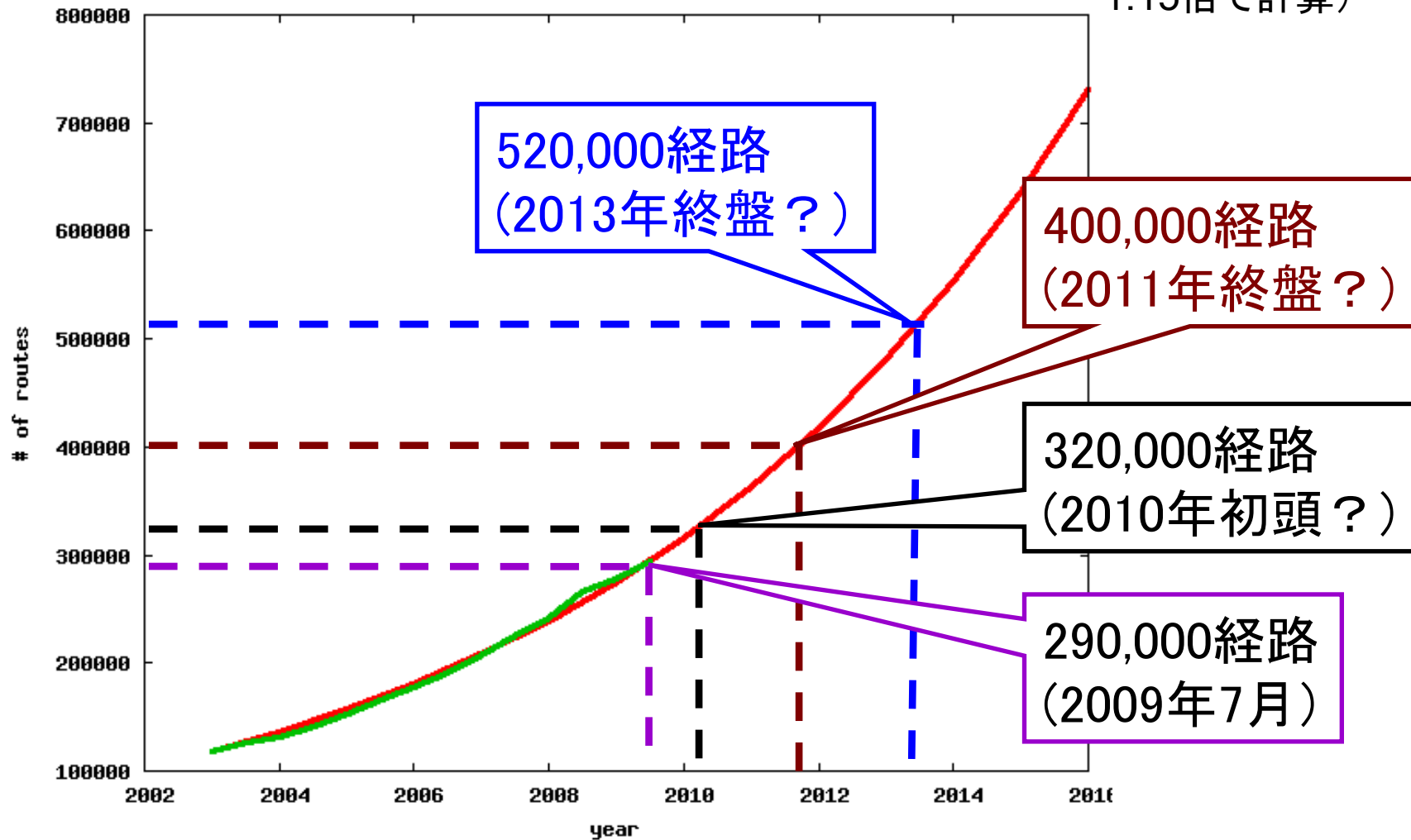
# 経路数増大による悪影響

## ■ ルータFIB枯渇

- IPv4,IPv6デュアルスタック
- IPアドレス移転
- ますますルータのリソース不足が懸念

# 今後の経路数伸び予測

— 実績値  
— 予測値(伸び年率  
1.15倍で計算)



## 最大収容経路数の例(IPv4のみ)

製品名	経路数	枯渇時期
Force10 EEシリーズ	26万経路	2008年中盤
ALAXALA AX2000Rシリーズ		
Force10 EFシリーズ	32万経路	2010年初頭？
Brocade IENI,BIシリーズ	40万経路	2011年終盤？
Brocade RX,MLXシリーズ	52万経路	2013年終盤？
Force10 EGシリーズ		
Brocade XMRシリーズ	100万経路	ずっと先？
ALAXALA AX7800Rシリーズ		

溢れたら上位機種へのリプレイス等が必要





# 経路数増大による悪影響

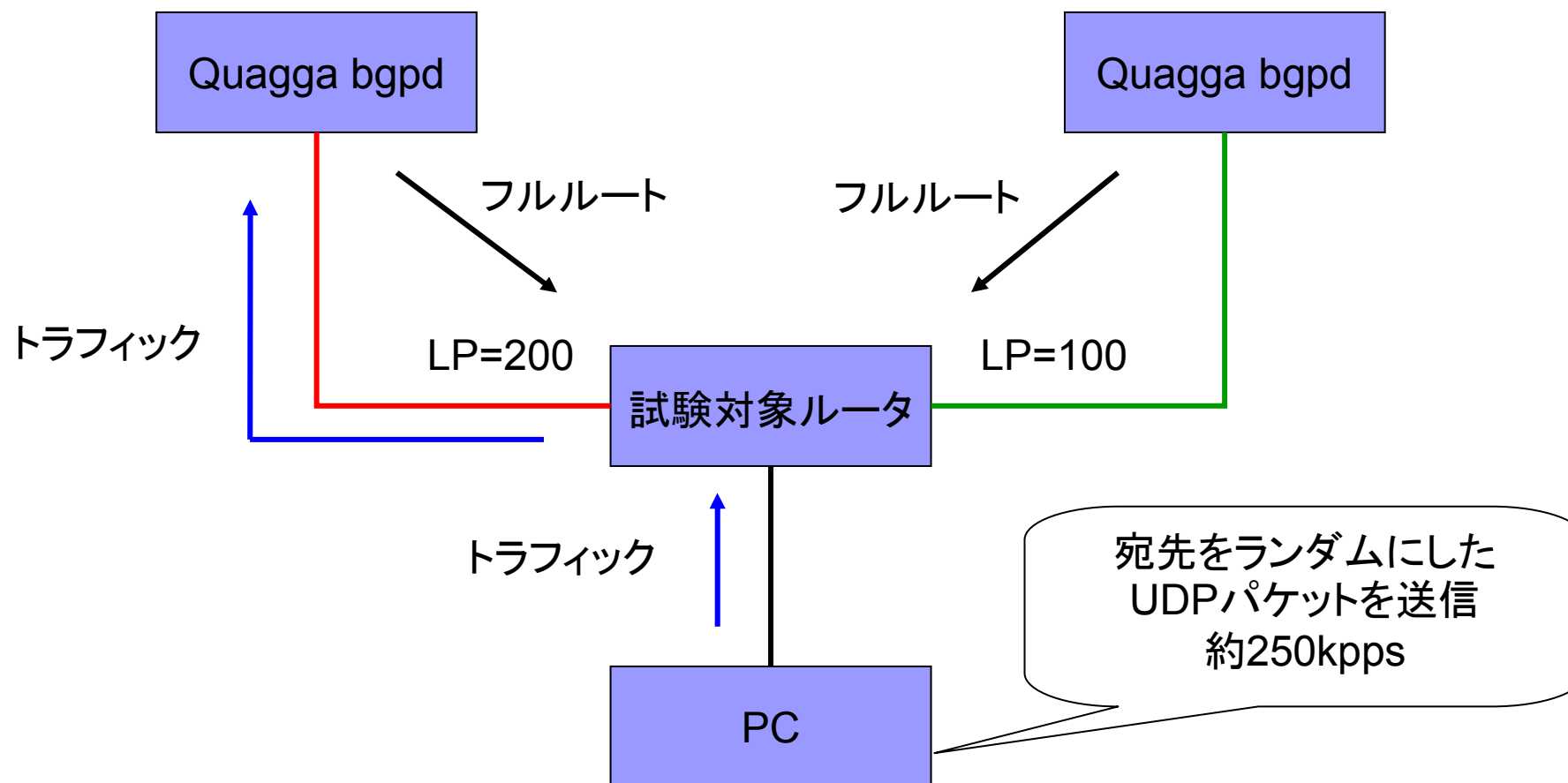
## ■ 経路収束の遅延

- BGPテーブルの更新
- ルーティングテーブルの更新
- FIBの更新
- 他のルータへUpdate

## ■ 29万経路の迂回にはどれくらいの時間を要するのか？

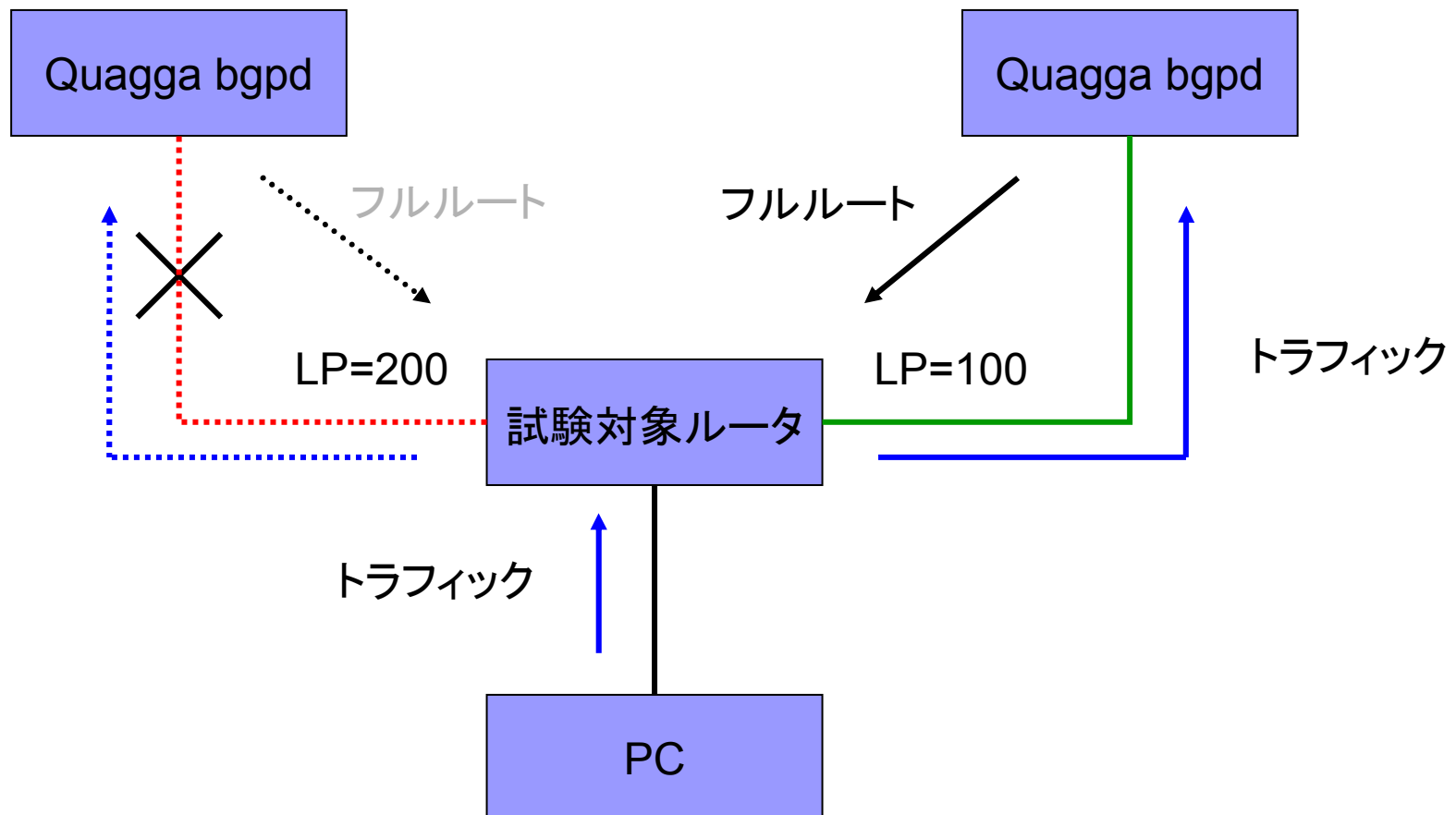
# 経路収束の遅延

## ■ 実験例



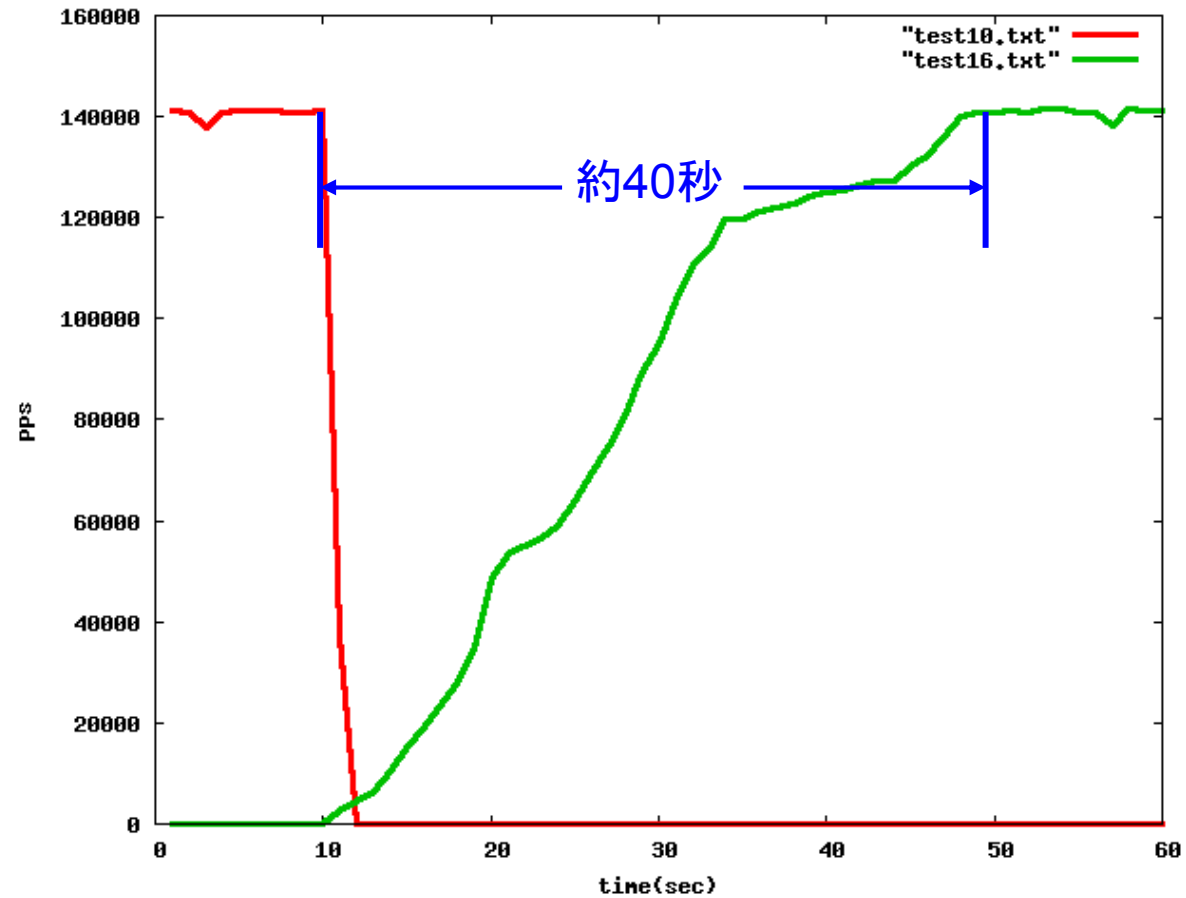
# 経路収束の遅延

## ■ 実験例



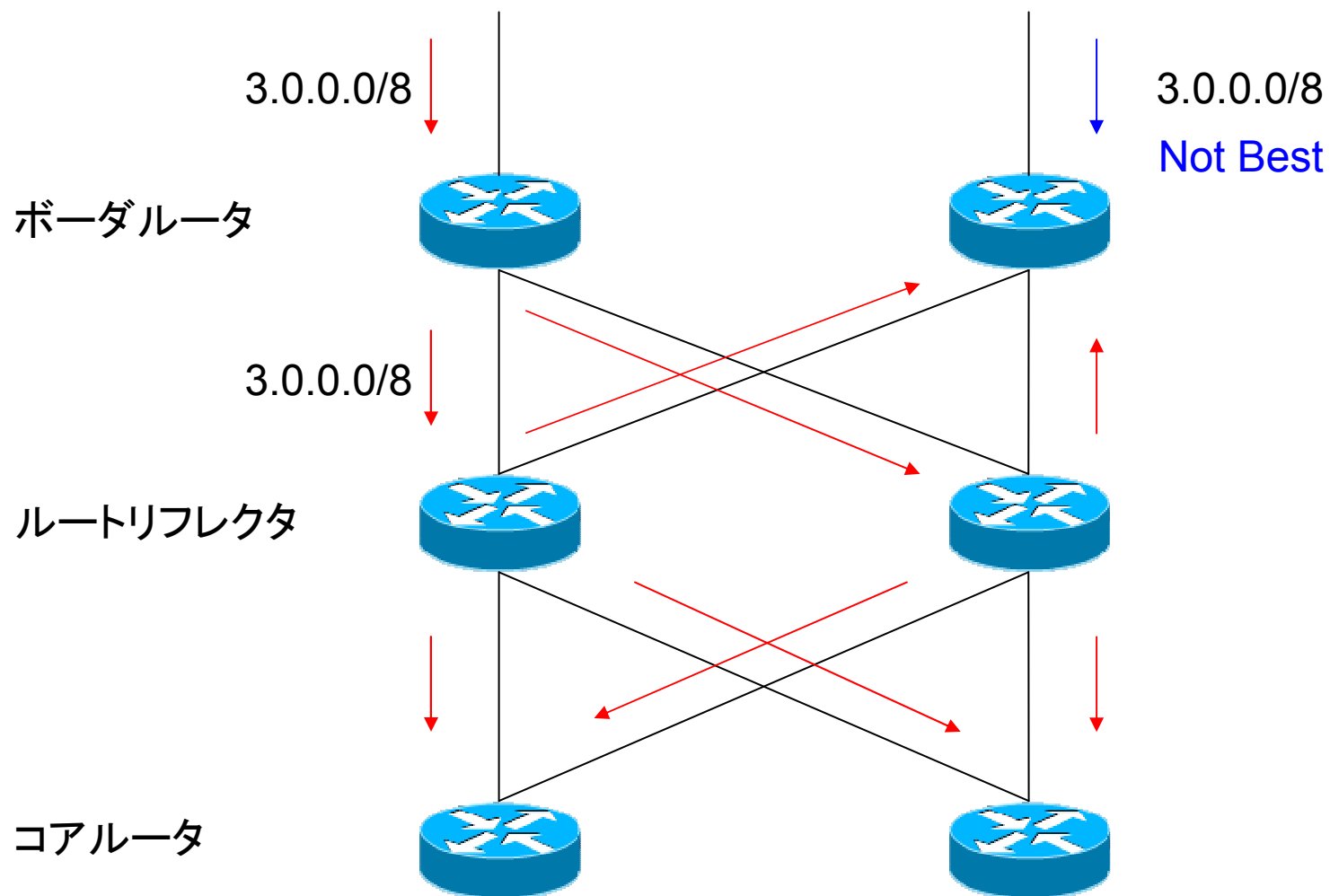
# 経路収束の遅延

## ■ 実験結果

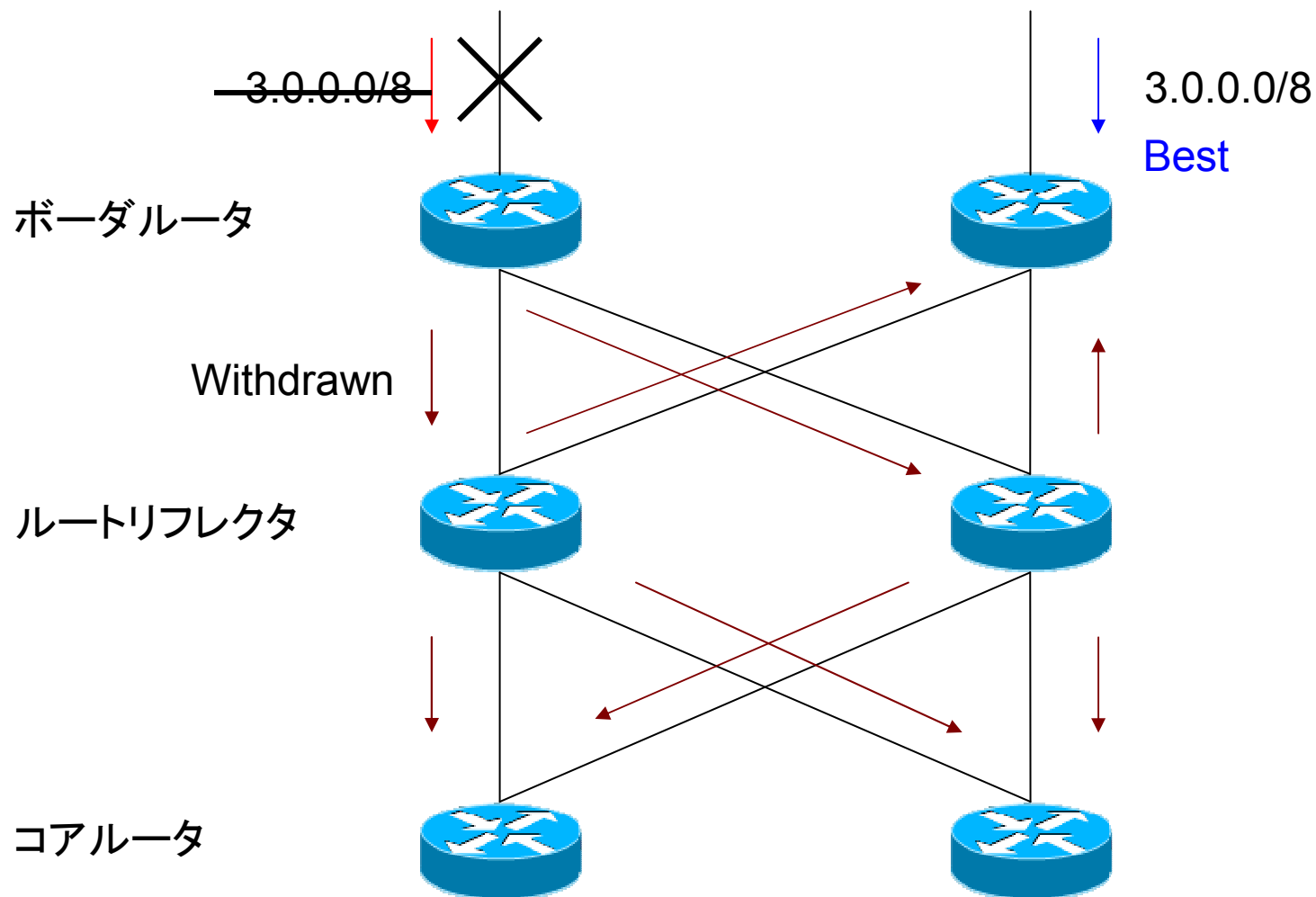


29万経路の切り替りに約40秒

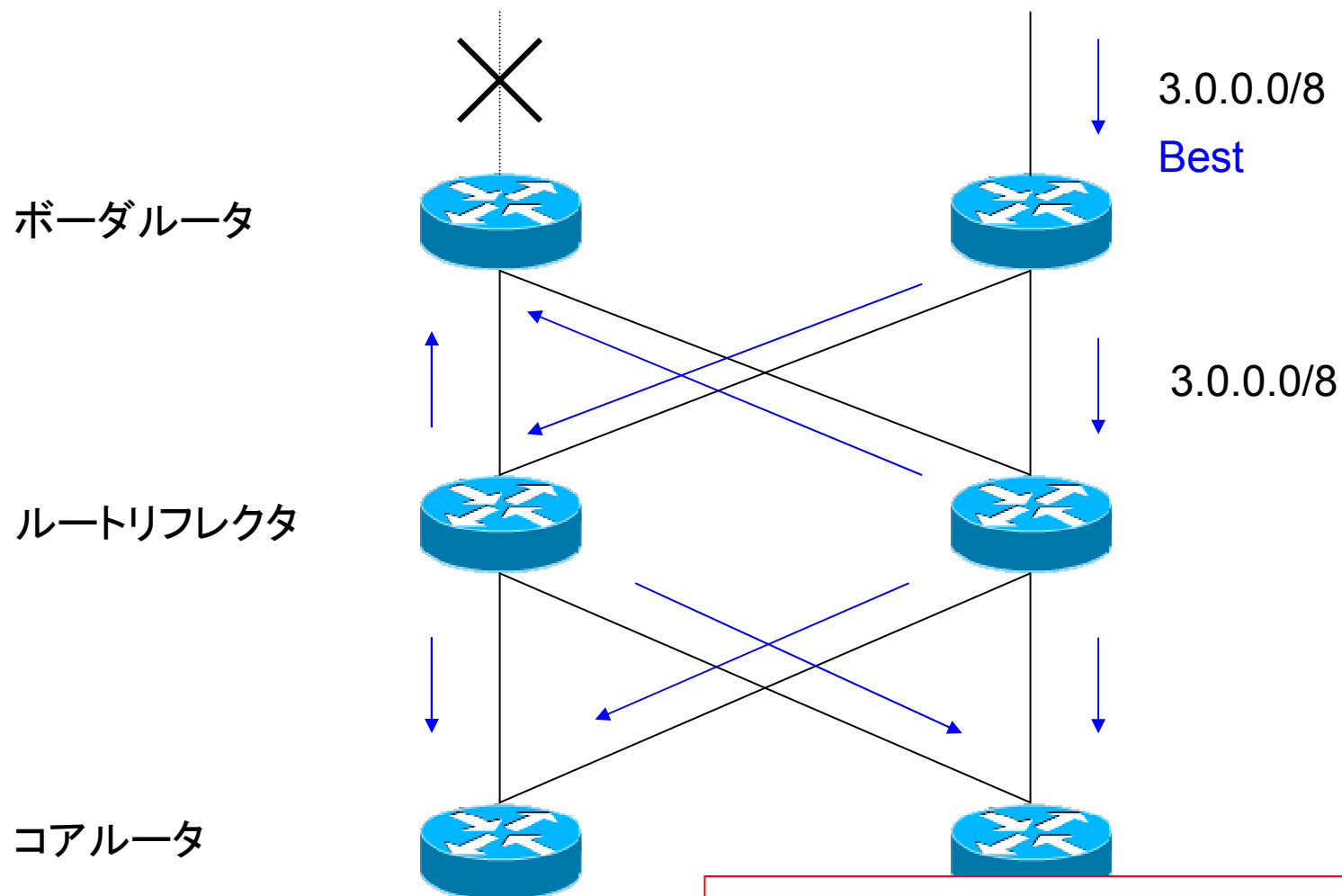
# 実際のネットワークはもっと複雑



# 実際のネットワークはもっと複雑



# 実際のネットワークはもっと複雑



網全体の収束には時間がかかる。



## フルルート利用率調査

- そもそも29万経路存在するフルルートは、どれくらい使われているのか？
  - トラフィックの多い宛先もあれば、少ない宛先もある。
  - Eg. トラフィックの多い上位XX経路で、全体のトラフィックのXX%を運んでいる。
- 弊社バックボーンで調べてみました。





# フルルート利用率調査

## ■ 調査対象

- 弊社バックボーンからインターネット向けout方向  
トラフィック

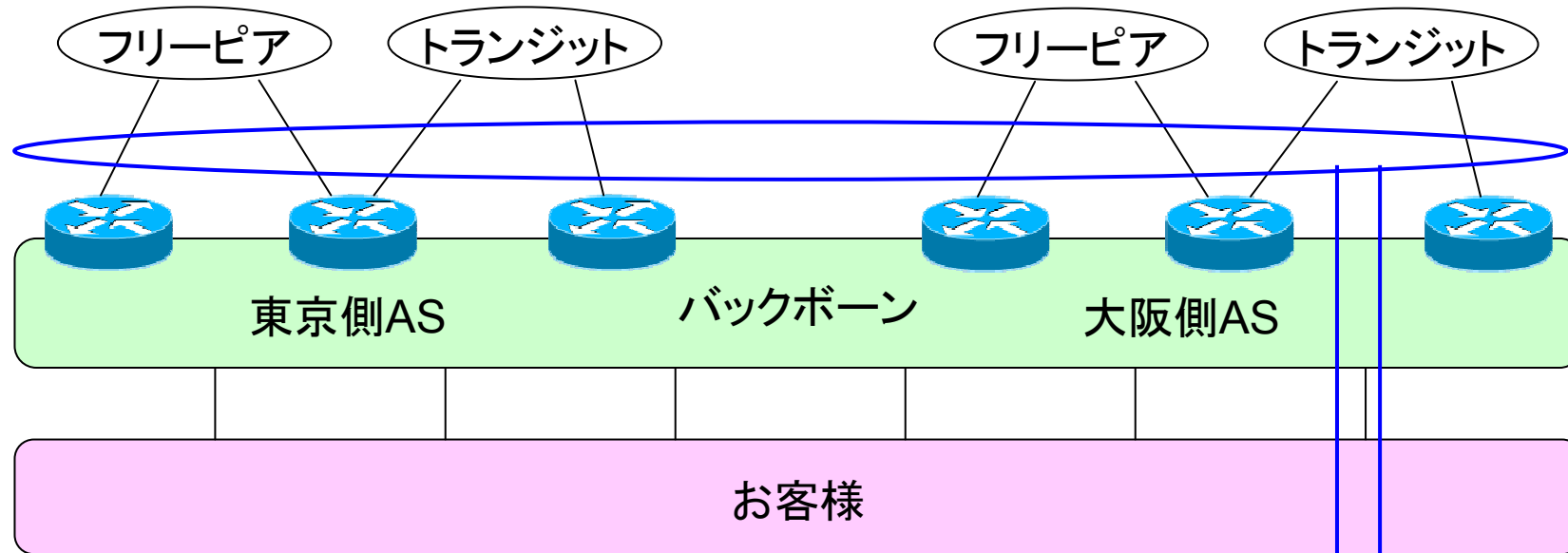
## ■ 調査方法

- sFlowサンプルパケットより、各経路にマッチする  
トラフィックの量を計測する。

## ■ 調査期間

- 2009/8/31 19:00～2009/9/1 10:00(15時間)
- 該当時間帯に165TBのトラフィックが流れた  
(平均24Gbps)

# sFlowを使ってPrefixごとにカウント



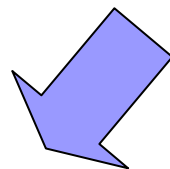
経路テーブル	カウンタ(bytes)	sFlowコレクタ
3.0.0.0/8	23232512	+
4.0.0.0/9	13182566	
4.21.103.0/24	24870912	+
4.23.88.0/24	18720961	
⋮	⋮	
29万エントリー		

sFlowサンプル

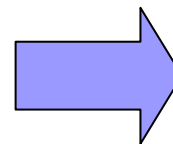
# トラフィックが多い順にPrefixをソート

経路テーブル	カウンタ(bytes)
3.0.0.0/8	23232512
4.0.0.0/9	131825664
4.21.103.0/24	24870912
4.23.88.0/24	187209613
⋮	⋮

ソート



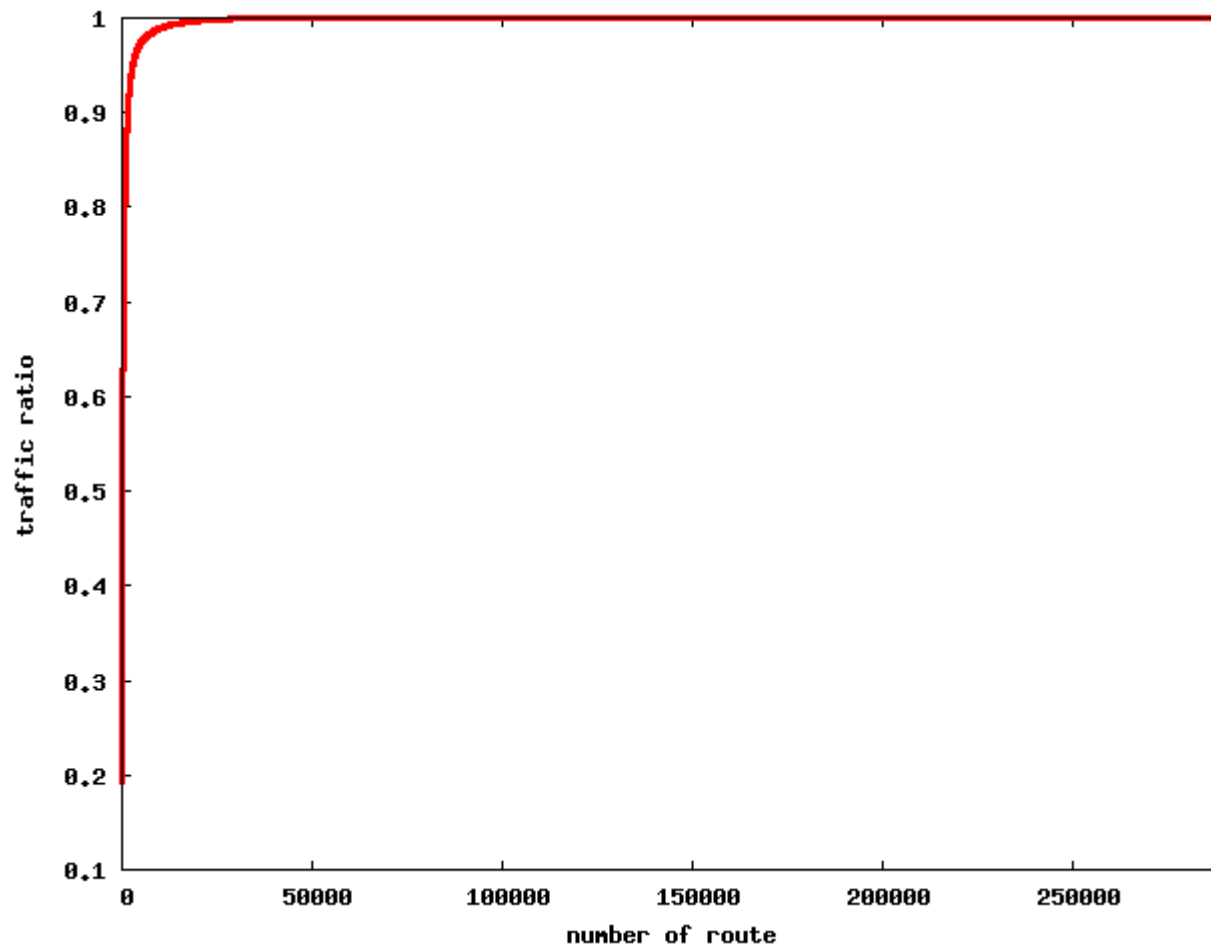
経路テーブル	カウンタ(bytes)
210.XX.0.0/17	8558762033152
122.XX.0.0/12	3086620721152
114.XX.0.0/11	3035447263232
118.XX.0.0/12	2987496521728
⋮	⋮



多いほうから積算していく

# 調査結果

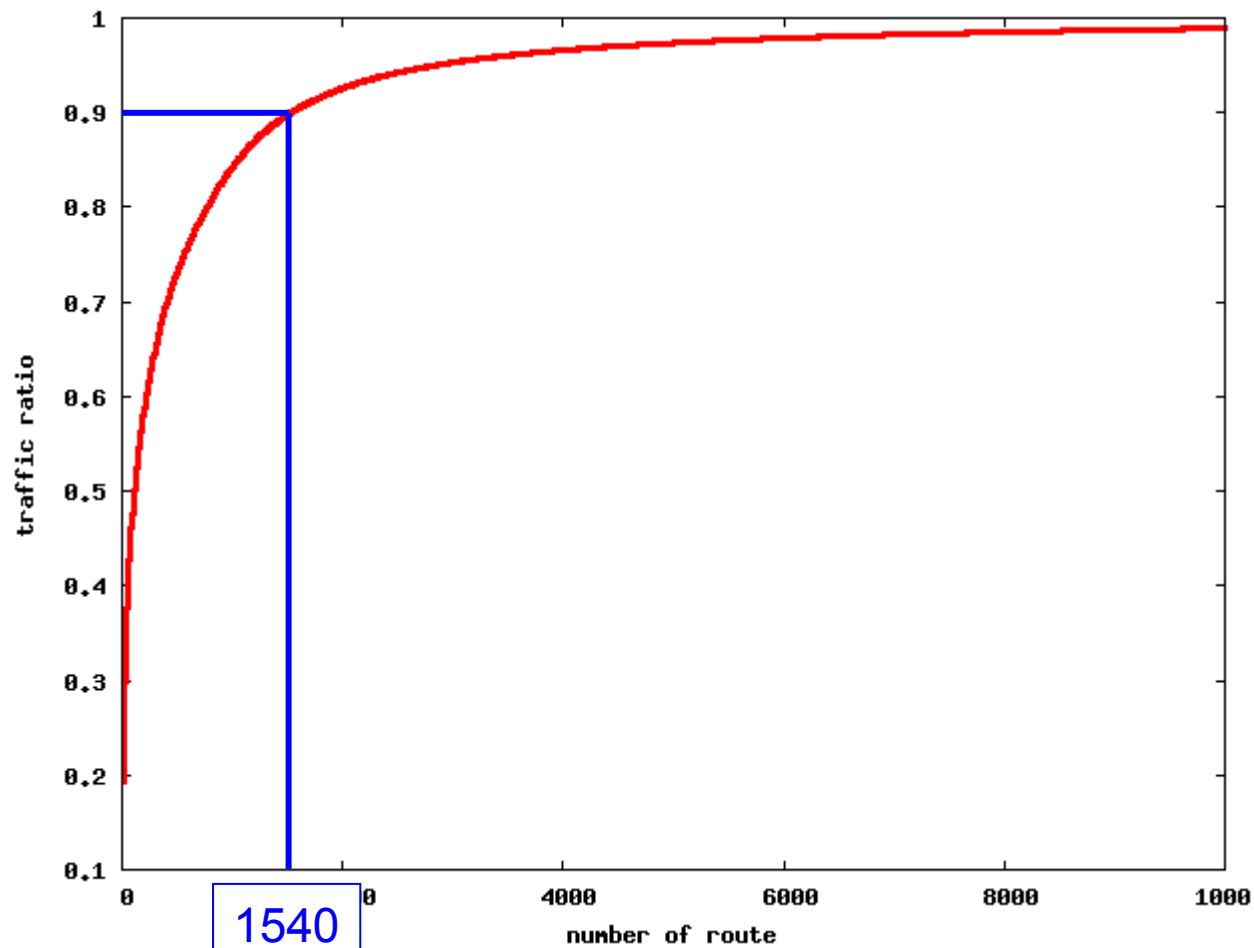
積算トラフィック割合



経路数

# 調査結果(1万経路までを拡大)

積算トラフィック割合



経路数



# フルルート利用率調査


トラフィック(%)	経路数	経路数(%)
20%	10	0.003%
50%	120	0.041%
80%	780	0.269%
90%	1,540	0.530%
99%	10,640	3.67%

約29万経路のうち、約28万経路は「ほぼ使われていない！」



## 理不尽な話

- フルルートは今後も増加する(だろう)。
  - 経路数増大に伴い、収束時間も長くなる。
  - FIBの枯渇により、ルータのリプレイスが必要。
  - コストがかかる。
- 
- その割には、使われている経路は微量。



# デフォルトルート活用の検討

- フルルートを捨てて、デフォルトルートを使おう！
- トラフィックエンジニアリング、マルチホームによる冗長化のため、BGPは引き続き使う。
- 利用率の多い経路のみ内部に伝播させ、その他の雑多な経路は上位ISPに向けたデフォルトルートでルーティング。
- 雑多な宛先は、近い上位ISPへ適当に流れてくれればよい。

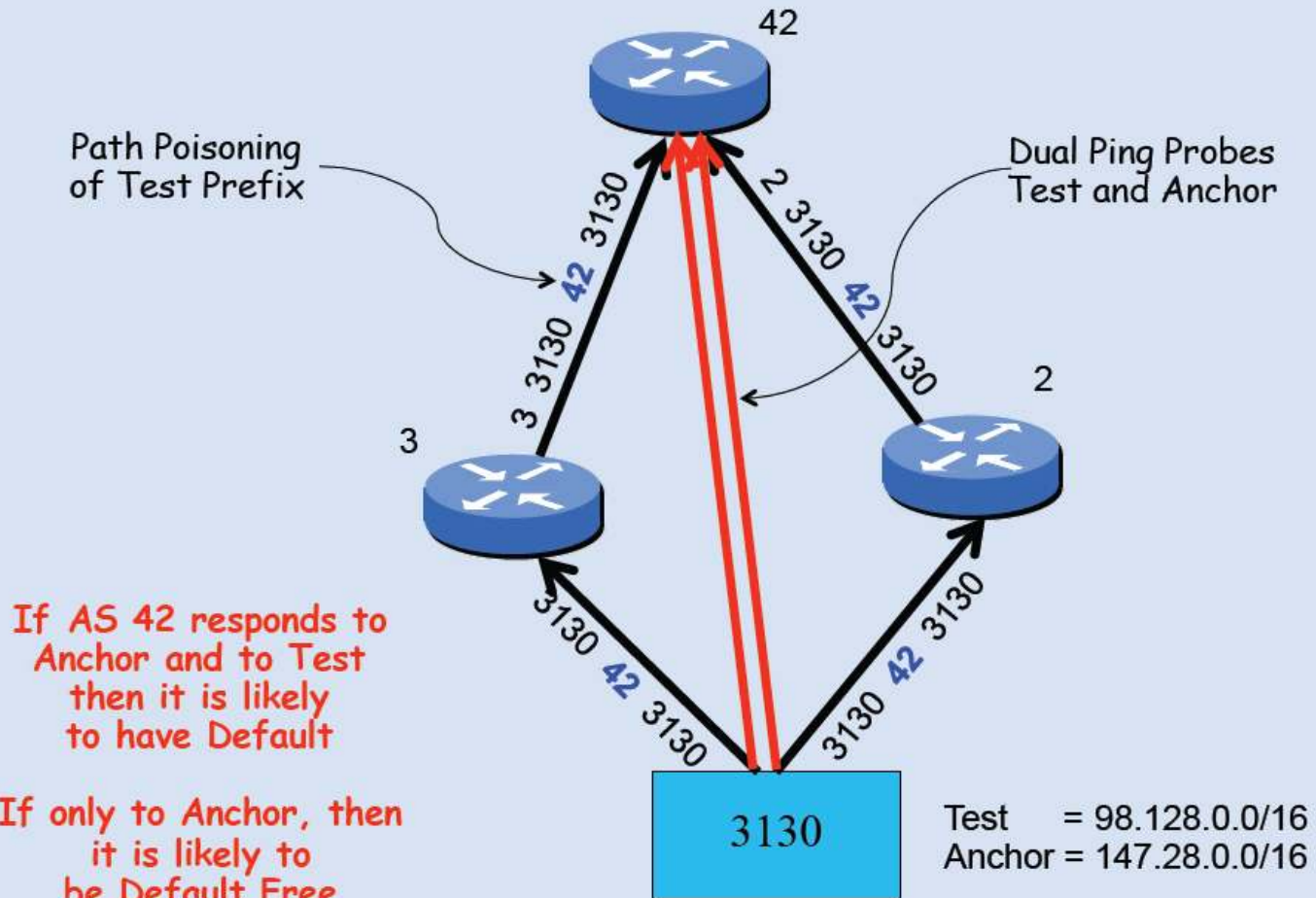




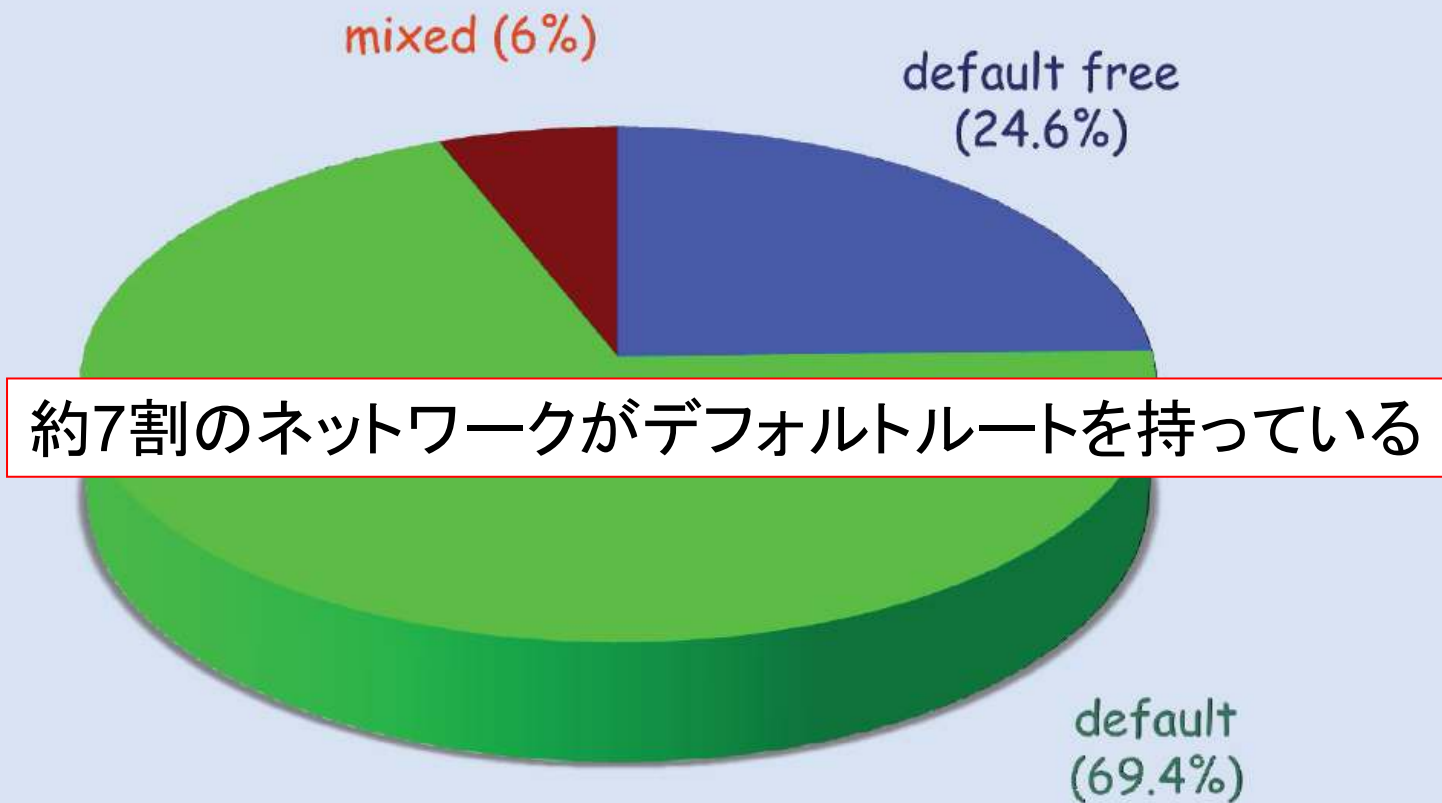
# JANOG24 Randyさんのお話

- Internet上でのDefault Route観測
- <http://www.janog.gr.jp/meeting/janog24/program/st3.html>

# Default Detection

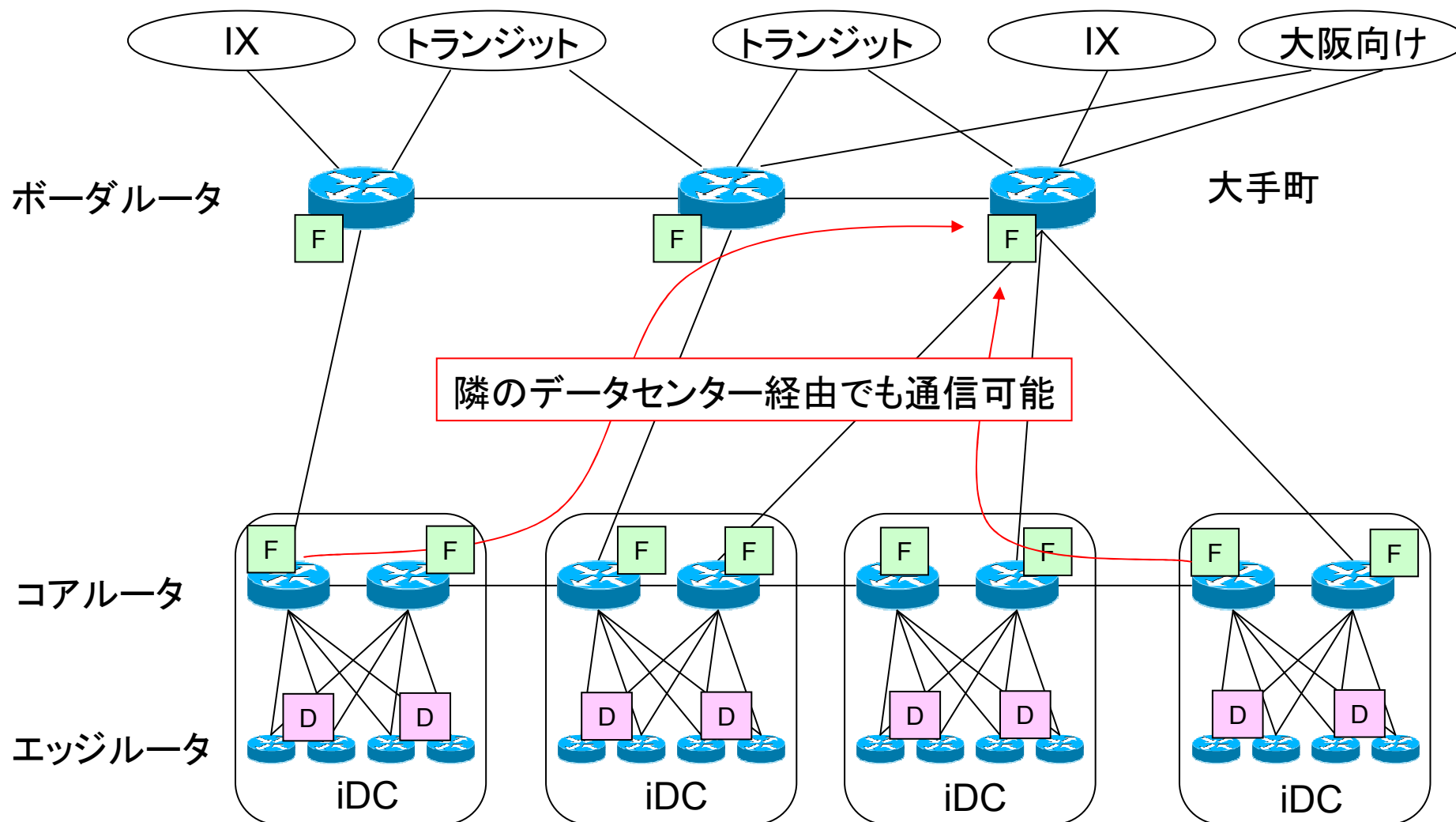


# Defaults in /25-Experiment

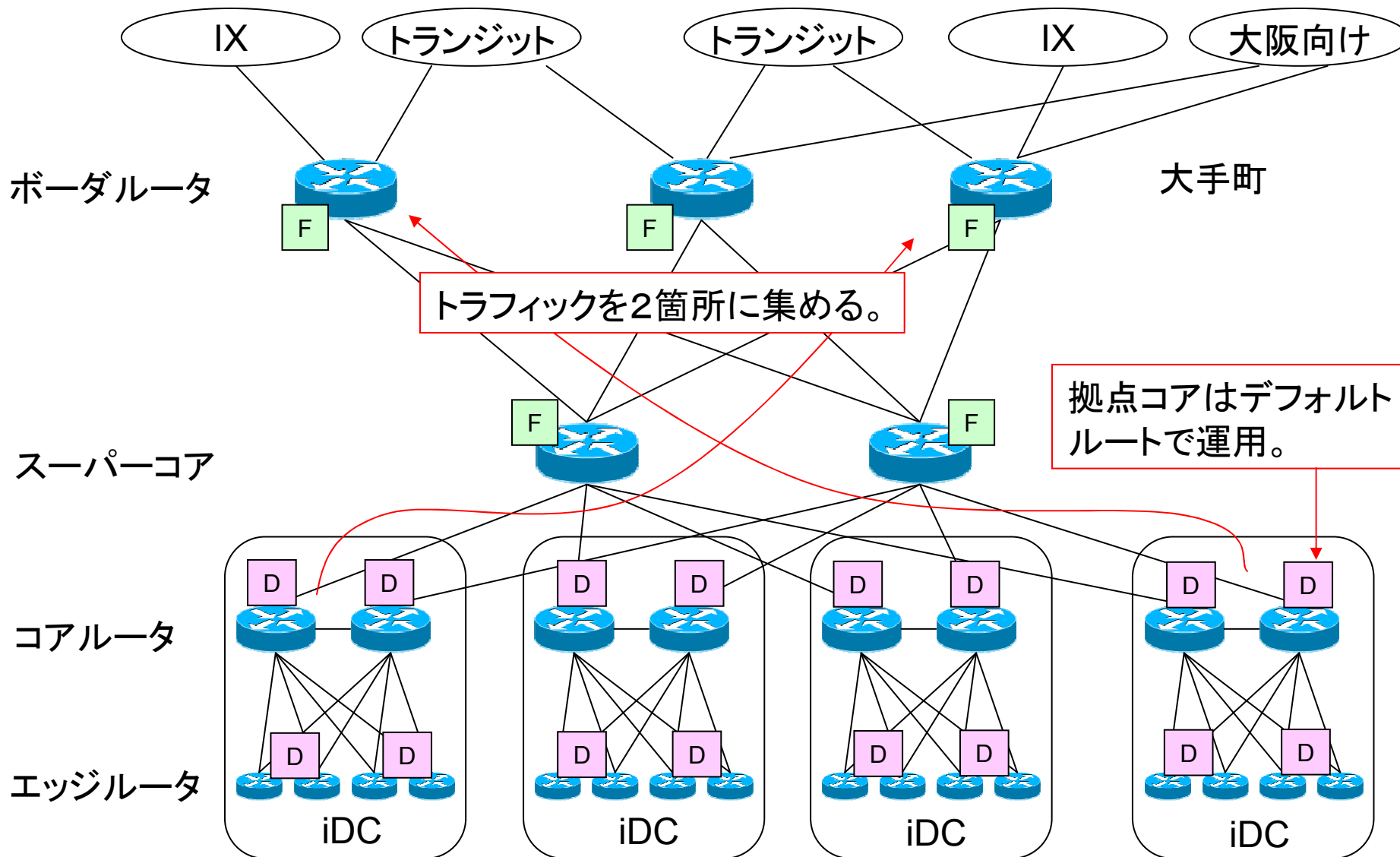


# 以前のネットワーク構成

F フルルート  
D デフォルトルート

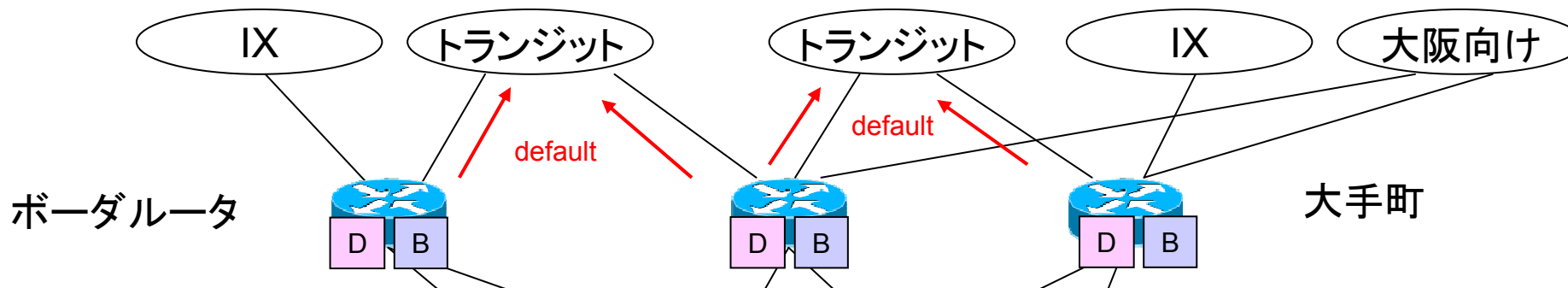


# コアのアグリゲーション



# デフォルトルート化

B BGP一部経路

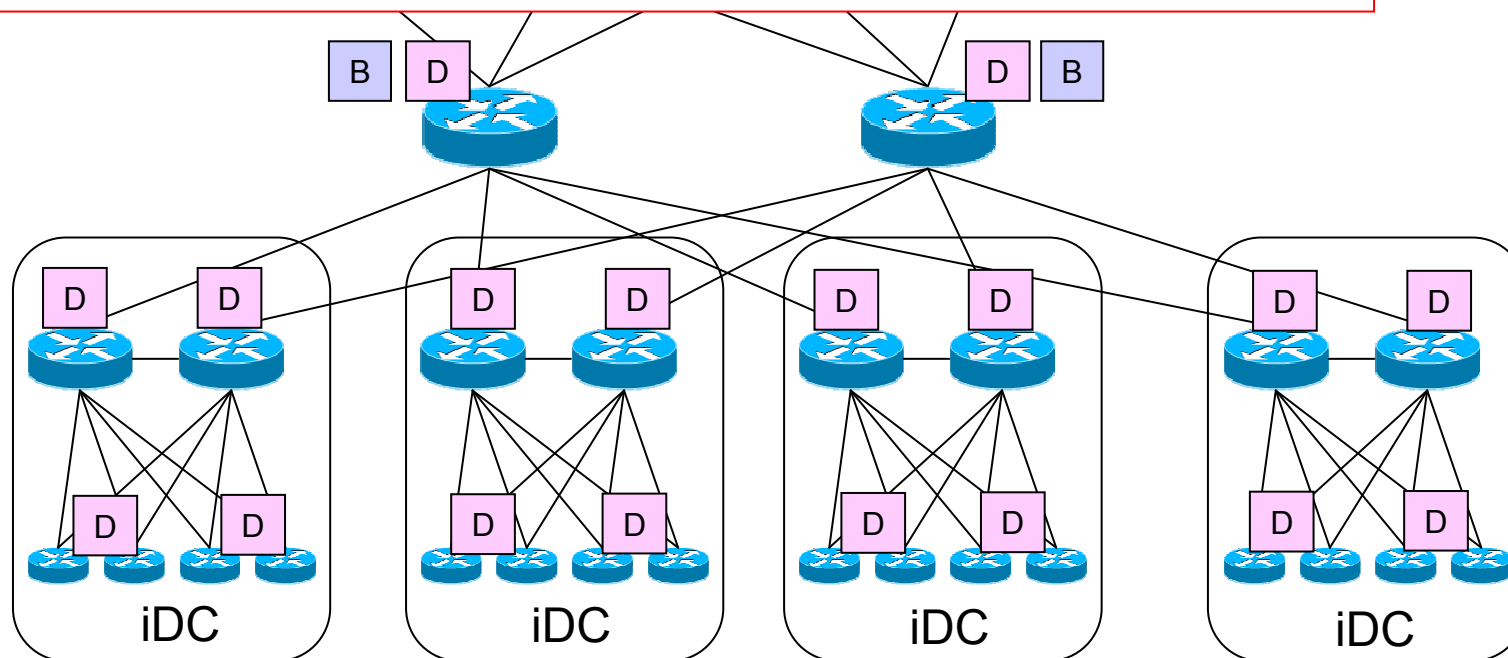


BGP一部経路のみ内部へ伝播。その他はデフォルトルートでまかなう。  
BGP顧客は別ASへ收容変更。

スーパーコア

コアルータ

エッジルータ





# 議論

## ■ ISPの皆さんへ

- フルルートってどこまでもっていています？
- フルルートをFIBにいれるのってつらくないですか？
- フルルート捨てて、デフォルトルートにしませんか？

## ■ トランジットISPの皆さんへ

- トランジットISPさん向けにデフォルトルートを書いてもいいでしょうか？
- トランジットISPさんからBGPでdefault originateしていただけるとうれしいんですが。

## ■ ルーターメーカーの皆さんへ

- FIBを節約する良い方法ないですか？