sFlowで遊ぼう！

さくらインターネット（株）　AS9370, AS9371, AS7684
技術部　大久保修一
ohkubo@sakura.ad.jp

アジェンダ

- 自己紹介
- 弊社での事例紹介
- 使っているツールの説明
- sFlowの憂鬱
- その他

自己紹介

さくらインターネットにてネットワークの仕事を担当

さくらインターネットって？

- 専用サーバ
- 法人向けハウジングサービス
- 個人向けレンタルサーバサービス
- IPトランジット
- オンラインゲーム(MMORPG)も..

SAKURA Internet

AS9370, AS9371, AS7684

一応、上場してます ^^; Mothers

設計？運用？ ⟶ ややˎうちのような小さい会社では
そんなもの、分かれてません；；

- ケーブル配線
- 機材設置
- ネットワーク設計～運用

- スイッチ、ルータのオペレーション
- ピアリング交渉とか・・
- DNSサーバとか・・

何でもやらされます（涙

弊社におけるsFlow利用事例

- sFlowによるトラフィック解析導入のきっかけ
  - ２年ほど前より
  - 主にトラフィックエンジニアリング目的で
    - 何がどこに向かって流れているのかしら？
  - トラフィック調整
    - ちょっとトラフィック、ばらしたいなぁ
  - ピアISPさんとの交渉
    - プライベートピアしませんか？
    - ちょっとDIX-IEからJPN●Pに移してもらえると、うれしいんですけど。。

NetFlow or sFlow?

NetFlowに比較して若干マイナー？？

⟶ NetFlowに対応してないルータが多数・・・

Foundry

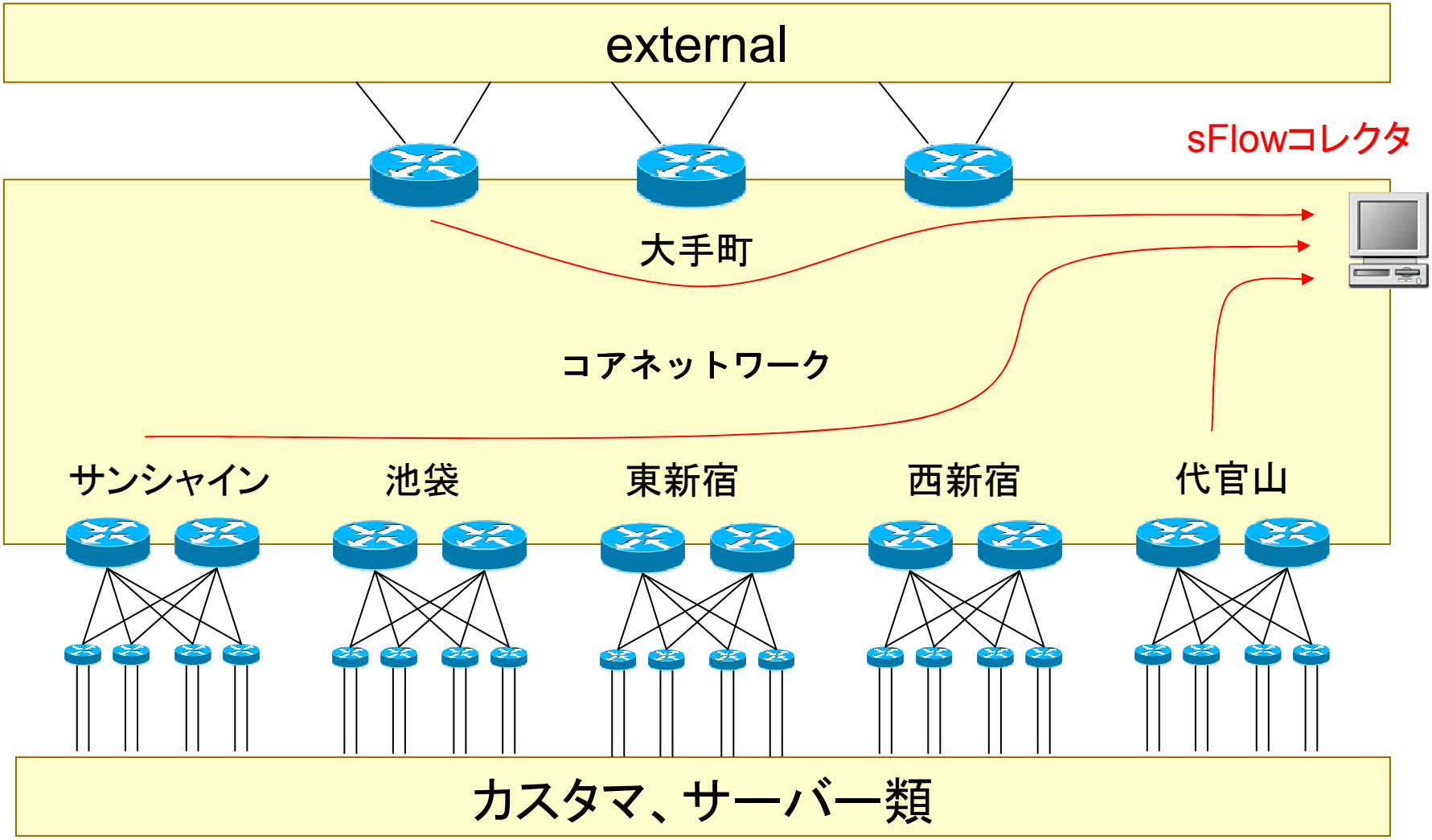
Force10

日立(GRシリーズ)

GRシリーズはNetFlowもOK

⟶ 仕方がなくsFlowへ…
あんまり深い理由はなかったり

ネットワーク構成図


external
sFlowコレクタ
大手町
コアネットワーク
サンシャイン
池袋
東新宿
西新宿
代官山
カスタマ、サーバー類

ルータ設定例

Foundryの例

```
sflow enable
sflow destination xx.xx.xx.xx
sflow polling-interval 30
sflow sample 8192
interface ethernet 1/2
 sflow forwarding
```

コレクタ

サンプリングレート

サンプリングするインターフェイスへ

GR4000の例

ALAXALAも同様

```
sflow yes
  destination xx.xx.xx.xx
  sample 8192
  version 4
  packet-information-type ip
  extended-information-type router gateway
  port 3/0-3,5,7-11
```

コレクタ

サンプリングレート

BGPの情報を利用する場合

サンプリングするインターフェイスへ

設定の注意点

コレクタの指定について

sflow destination xx.xx.xx.xx
sflow destination xx.xx.xx.xx

→ 複数定義した場合、同じデータが送信される
※バックアップコレクタを立てる場合など

サンプリングレートについて

sflow sample 8192

→ 任意の値が指定できるわけではない

たとえば、2×4^n等、装置によって異なる
1,2,8,32,128,512,2048,8192,32768・・・

※小さな値を指定する場合、CPU負荷に注意！！

sFlowに対応したコレクタは？

商用のものはいろいろあるが・・

ARBOR NETWORKS　GenieNRM　InMon internet monitoring

フリーソフトでお手軽に使えるツールは？

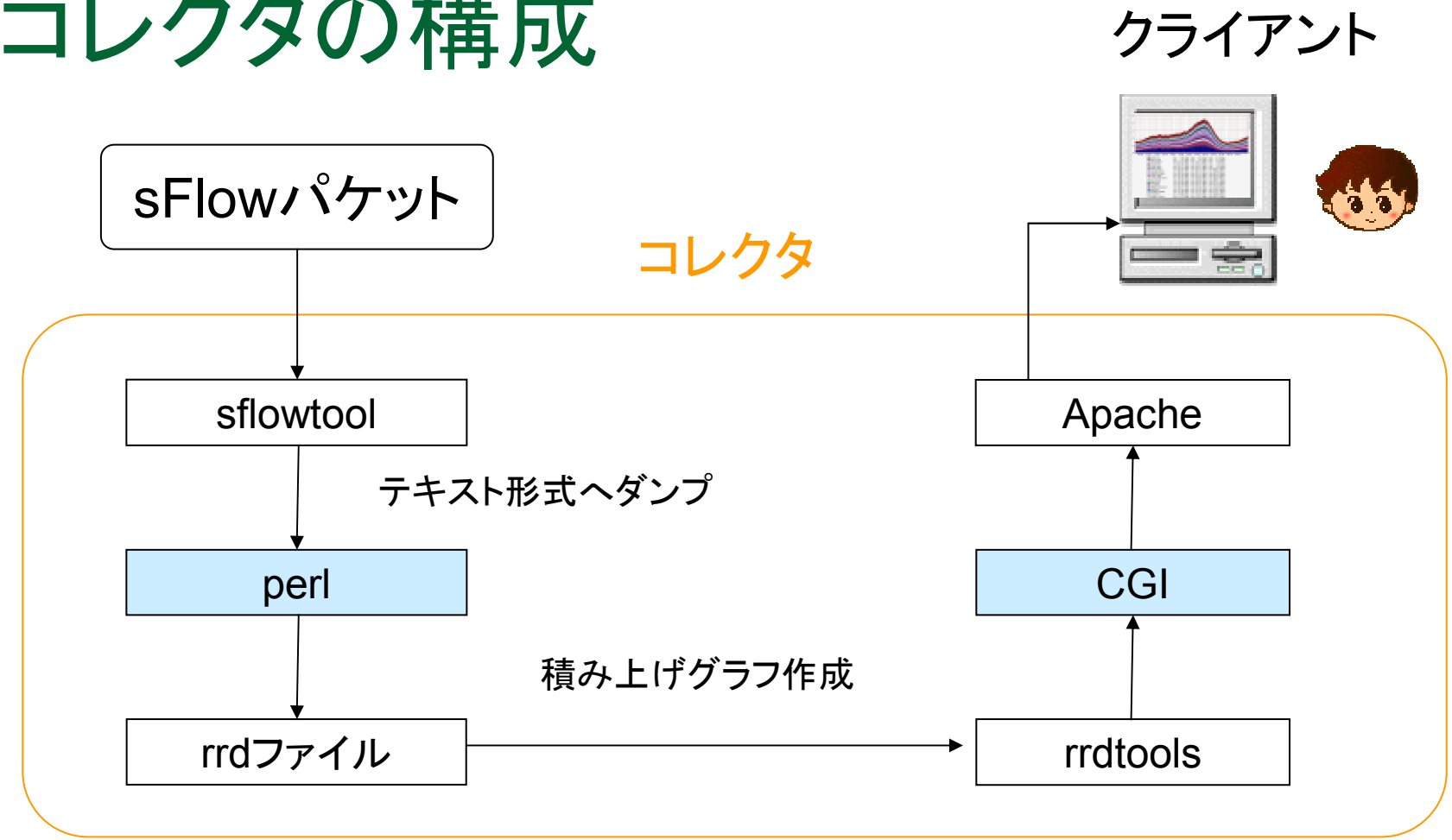
↓

NetFlowに比べると情報が少ない

↓

結局、自作の解析ツールを開発することに。。

コレクタの構成



下記URLで公開中
http://micho.mimora.com/sflow-sgraph.html

sflowtoolの使い方

- ダウンロード
  - http://www.inmon.com/technology/sflowTools.php
- コンパイル
  - % ./configure
  - % make
- libpcap(tcpdump)形式へ
  - % sflowtool –t | tcpdump –r -
- NetFlow形式へ
  - % sflowtool –c hostname –d port
- テキストへダンプ (引数なし)
  - % sflowtool

どんな情報がとれるのか？

テキストへダンプした出力結果

meanSkipCount 8192	サンプリングレート
inputPort 1 outputPort 144	ポート番号(MIBの値)
sampledPacketSize 1498	フレームサイズ(要注意！)
dstMAC 000b45b7XXXX srcMAC 00008798XXXX	MACアドレス
srcIP 61.211.XXX.XXX dstIP 222.91.XXX.XXX	IPアドレス
IPProtocol 6	IPプロトコル(6=TCP)
IPTTL 60	TTL

どんな情報がとれるのか？

テキストへダンプした出力結果

TCPSrcPort 80
TCPDstPort 3726
nextHop 210.138.XXX.XXX
srcSubnetMask 30
dstSubnetMask 16
my_as 9370
src_as 0
dst_as_path 2497-4134
dst_as 4134
dst_peer_as 2497

ポート番号
ネクストホップ
マッチするPrefixのlength
AS情報

例:ポート番号別統計

サンプルコード

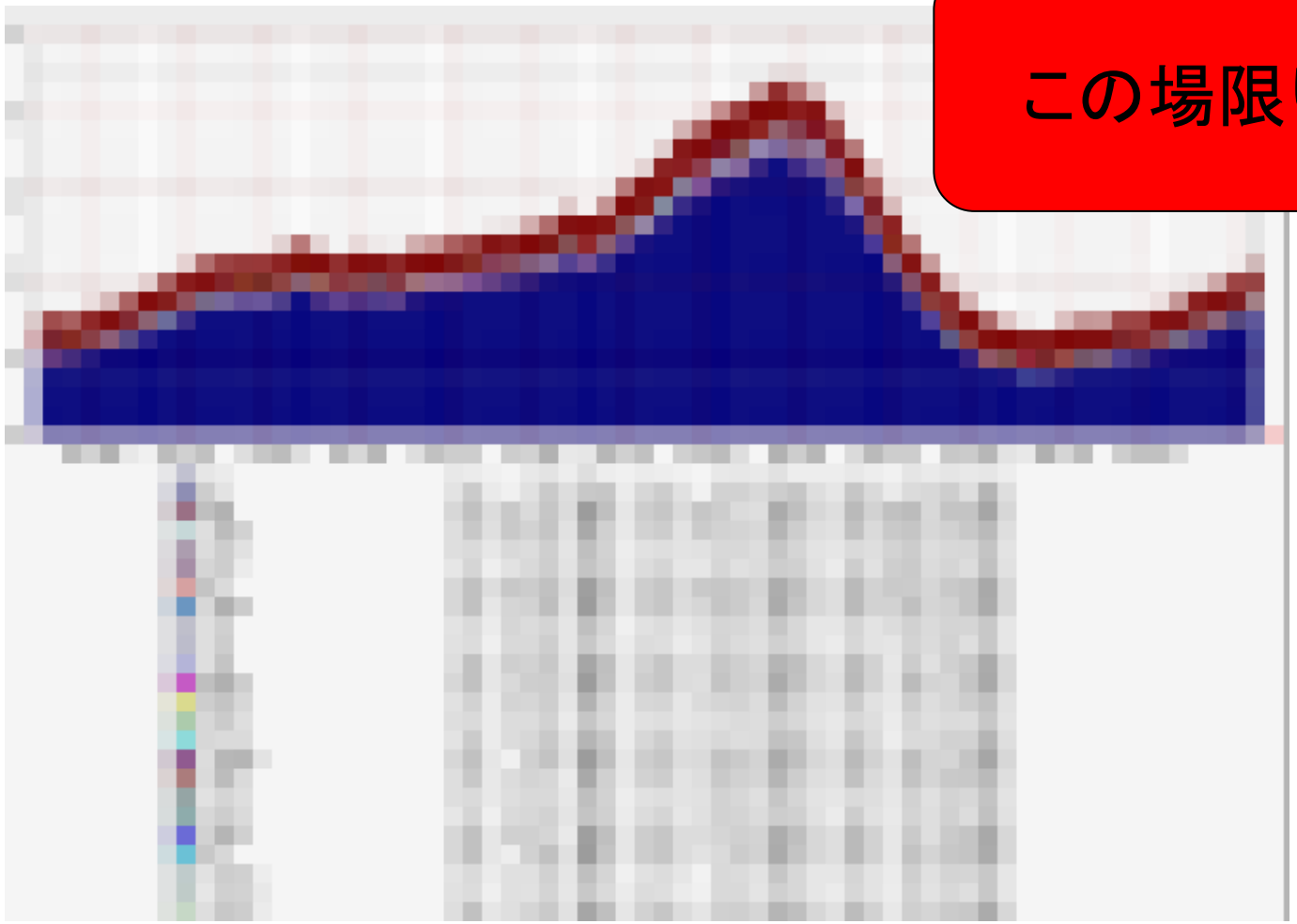
```
$agent = $header{'agent'};
$inputport = $sflow{'inputPort'};
$length = $sflow{'sampledPacketSize'};
$rate = $sflow{'meanSkipCount'};
$tcpport = $sflow{'TCPSrcPort'};

if ($agent eq '192.168.100.200') {
    if ($inputport == 50) {
        $count{$tcpport} += $length * $rate;
    }
}
```

調べたいルータ
インターフェイス

フレーム長とサンプリング
レートの積をカウント

例：ポート番号別統計

解析結果



この場限り。

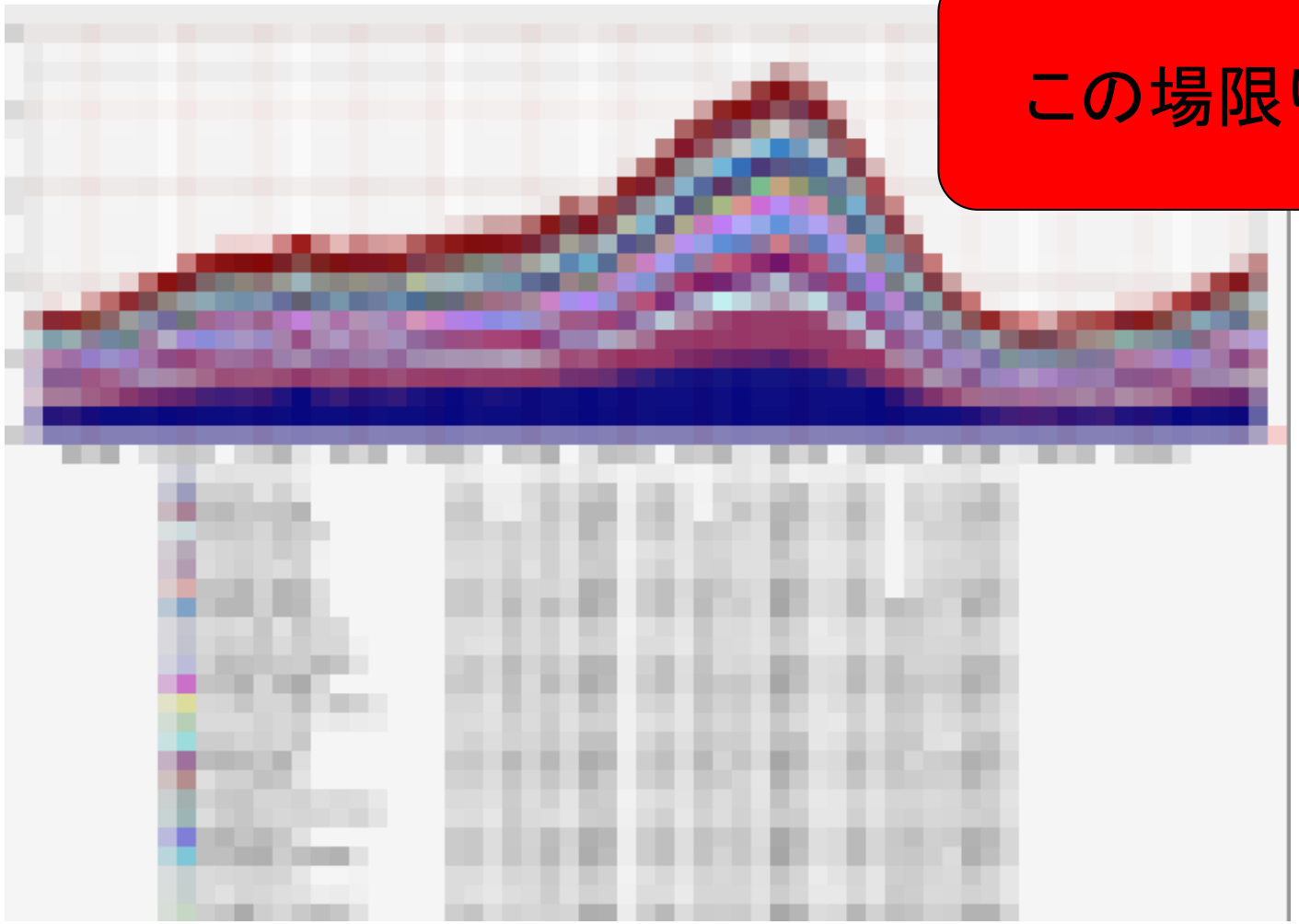
例：AS番号別統計

サンプルコード

```
$agent = $header{'agent'};
$inputport = $sflow{'inputPort'};
$length = $sflow{'sampledPacketSize'};
$rate = $sflow{'meanSkipCount'};
$dstas = $sflow{'dst_as'};

if ($agent eq '192.168.100.200') {
    if ($inputport == 50) {
        $count{$dstas} += $length * $rate;
    }
}
```

例：AS番号別統計

解析結果



この場限り。

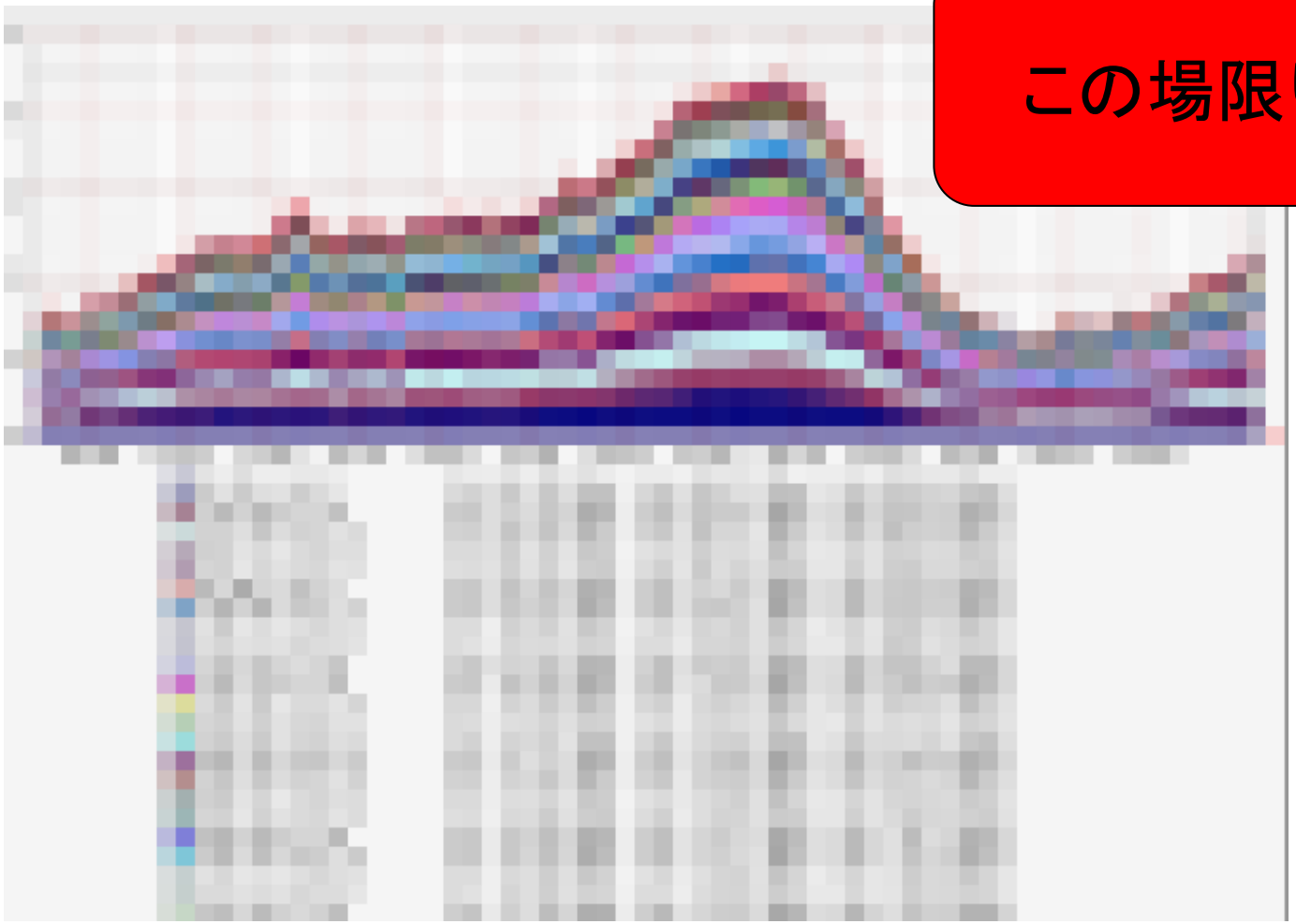
例：Prefix別統計

サンプルコード

```
$agent = $header{'agent'};
$inputport = $sflow{'inputPort'};
$length = $sflow{'sampledPacketSize'};
$rate = $sflow{'meanSkipCount'};

$dstip = $sflow{'dstIP'}
$dstmask = $sflow{'dstSubnetMask'}
$prefix = &prefix($dstip, $dstmask);

if ($agent eq '192.168.100.200') {
    if ($inputport == 50) {
        $count{$prefix} += $length * $rate;
    }
}
```

例：Prefix別統計

解析結果



この場限り。

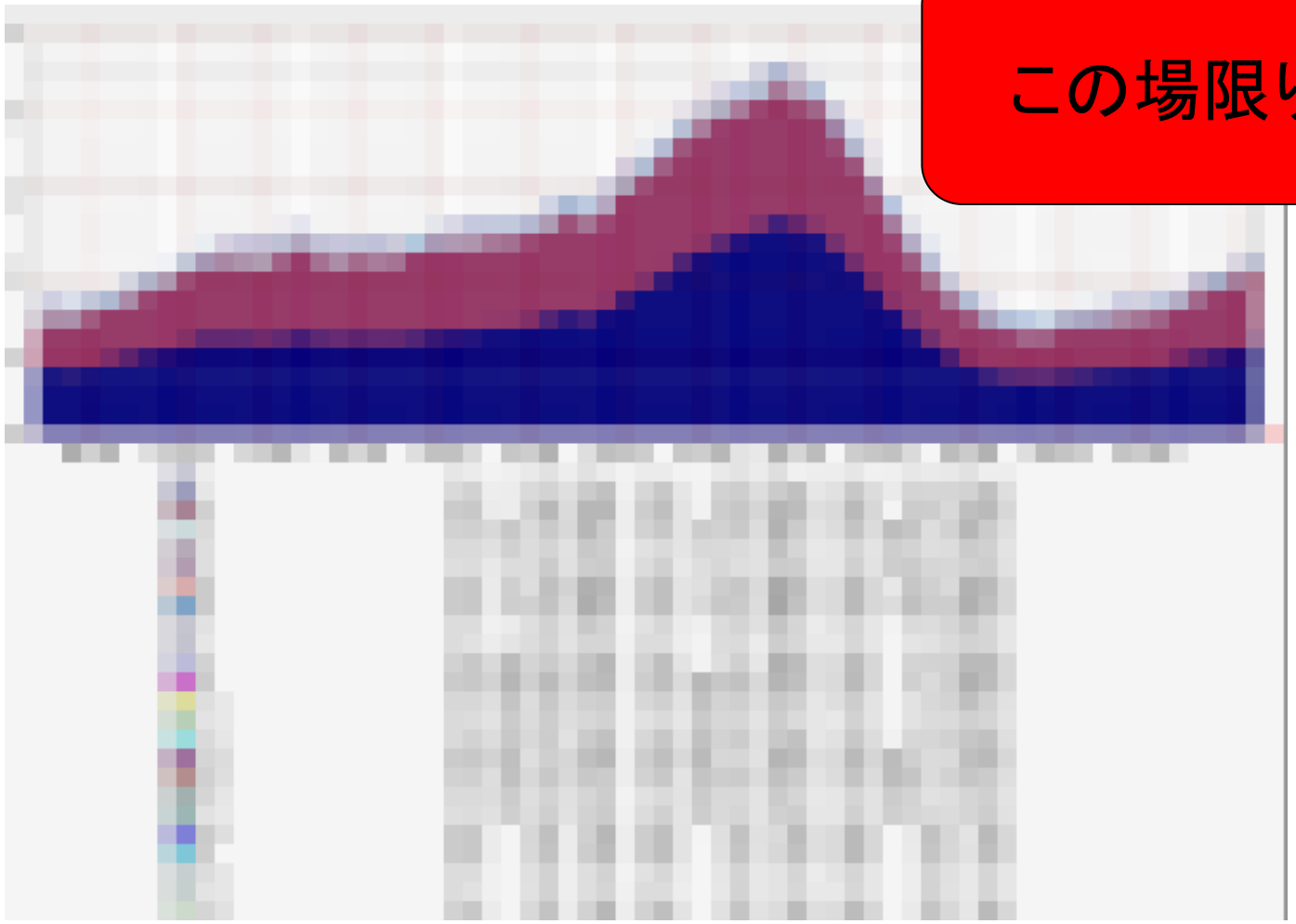
例：AS Hop数統計

サンプルコード

```
$agent = $header{'agent'};
$inputport = $sflow{'inputPort'};
$length = $sflow{'sampledPacketSize'};
$rate = $sflow{'meanSkipCount'};

$ashop = scalar(split(/¥-/, $sample{'dst_as_path'}));

if ($agent eq '192.168.100.200') {
    if ($inputport == 50) {
        $count{$ashop} += $length * $rate;
    }
}
```

例：AS Hop数統計

解析結果



この場限り。

sampledPacketSizeの微妙な違い

最長フレームの場合

Foundry BigIron RX4	1518	→ OK!
Foundry NetIron 400	1514	→ FCSが抜けてる。。
HITACHI GR4000	1500	→ ペイロードだけ？

最短フレームの場合

Foundry BigIron RX4	64	→ OK!
Foundry NetIron 400	60	→ FCSが抜けてる。。
HITACHI GR4000	40	→ ペイロードだけ？

※プロトコルによりさらに短いこともある

コレクタ側で補正して、カウントする必要がある（涙

AS番号フィールドの微妙な違い

正しいルーター

dst_as_path 2914-1239-19782-1767-20452
dst_as 20452
dst_peer_as 2914

HITACHI GR4000

dst_as 2914
dst_peer_as 20452

意味が逆になっている

dst_as_path 20452-2914

ひっくり返って、途中が抜けている

Foundry BigIron RX4（昔のバージョン）

dst_as_path 2914-0-0-0-20452

途中が0に・・

コレクタ側で補正して、カウントする必要がある（涙

その他注意点

サンプリングレートの設定

あまり頻度を上げると、CPU負荷が上昇・・・
GRシリーズでは、推奨値がマニュアルに掲載

http://www.hitachi.co.jp/Prod/comp/network/manual/switch/gs4k/1002/HTML/CFREFHY2/0049.HTM

Extended Gateway(BGPの情報)の対応状況

Force10・・・未対応　→　コレクタ側で頑張れるか？？

サンプリングのタイミングの違い

Ingress・・・Foundry、ALAXALA、日立GRシリーズ
Egress・・・Force10

QoSやパケットフィルタによって、結果が異なる場合も・・

Ingress VS Egress

Ingress

- ルータ自身から発せられるパケットがサンプリングできない。
- Egressでフィルタされるパケットがサンプリングされてしまう。
- Egress側でQoSが掛かっていると、実際のパケット送出順序とサンプルの到着順序が異なる場合がある。

Egress

- ルータ自身に向かうパケットがサンプリングできない。
- どちらかというと、Ingressよりも実状に近いサンプリングが可能。

あまり気にする必要はないかもしれませんが。。。

Discussion!

- Exporterの微妙な差異を吸収するノウハウ！？
  - あの箱は〇〇とか、この箱は××とか・・
- サンプリングレート、どれくらいあれば十分？
  - 課金に使うのであれば、誤差が気になるところ・・
  - 森さんの発表で聞けるでしょうか？
- DoSアタックの検出もできればいいんだけど。。
  - うーん、やっぱり商用の箱を使うしかないのか！？