xflow tutorial

JANOG36

Tajima Hirotaka / Genie Networks

agenda

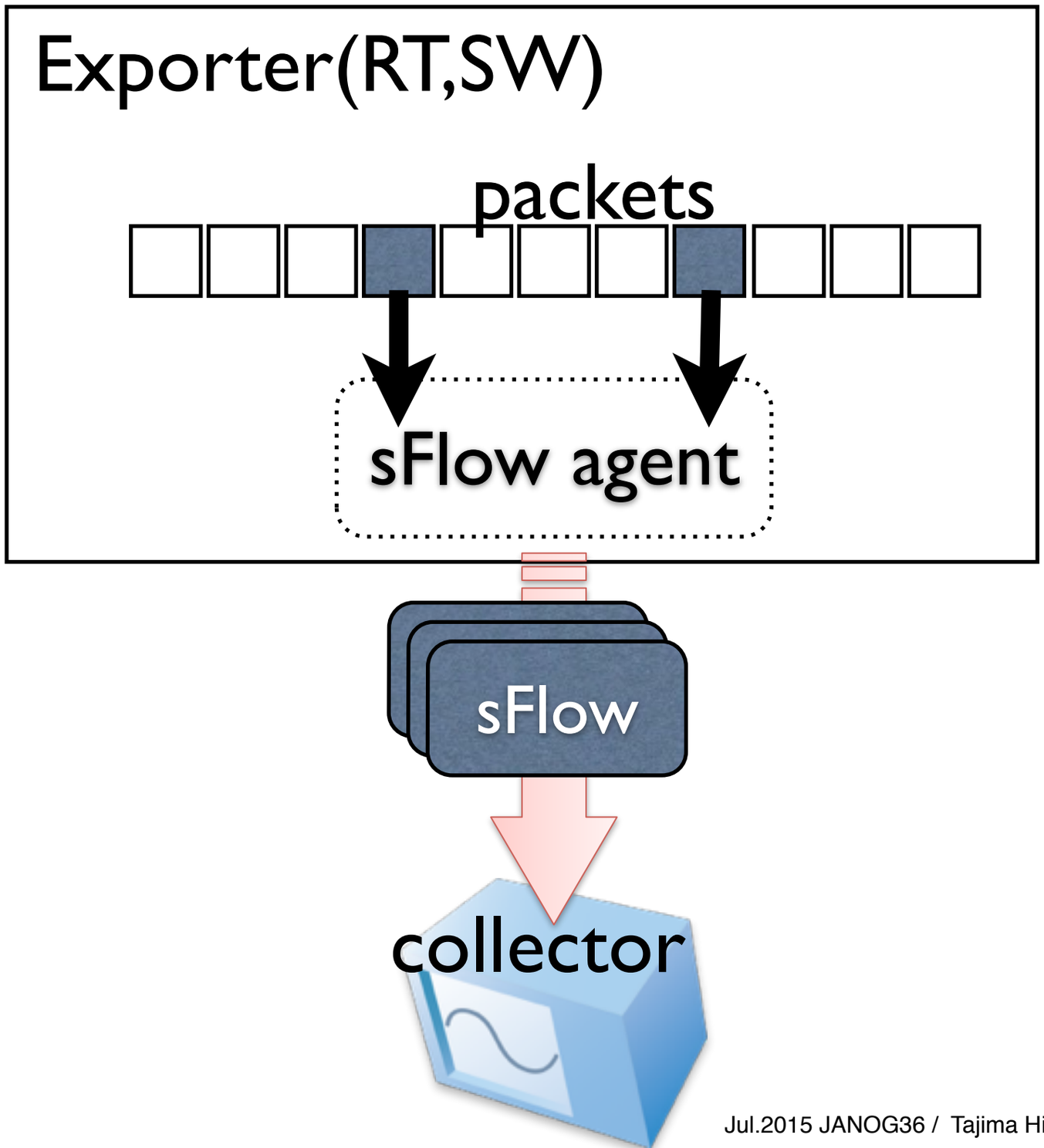
- sFlow
- コレクター
- FAQとtips

sFlow

sFlow

- サンプリングベース
- スイッチで多く実装されている
- 現在はVersion4 か5が多い
  - V4はRFC3176(informational)

sFlowはサンプリング



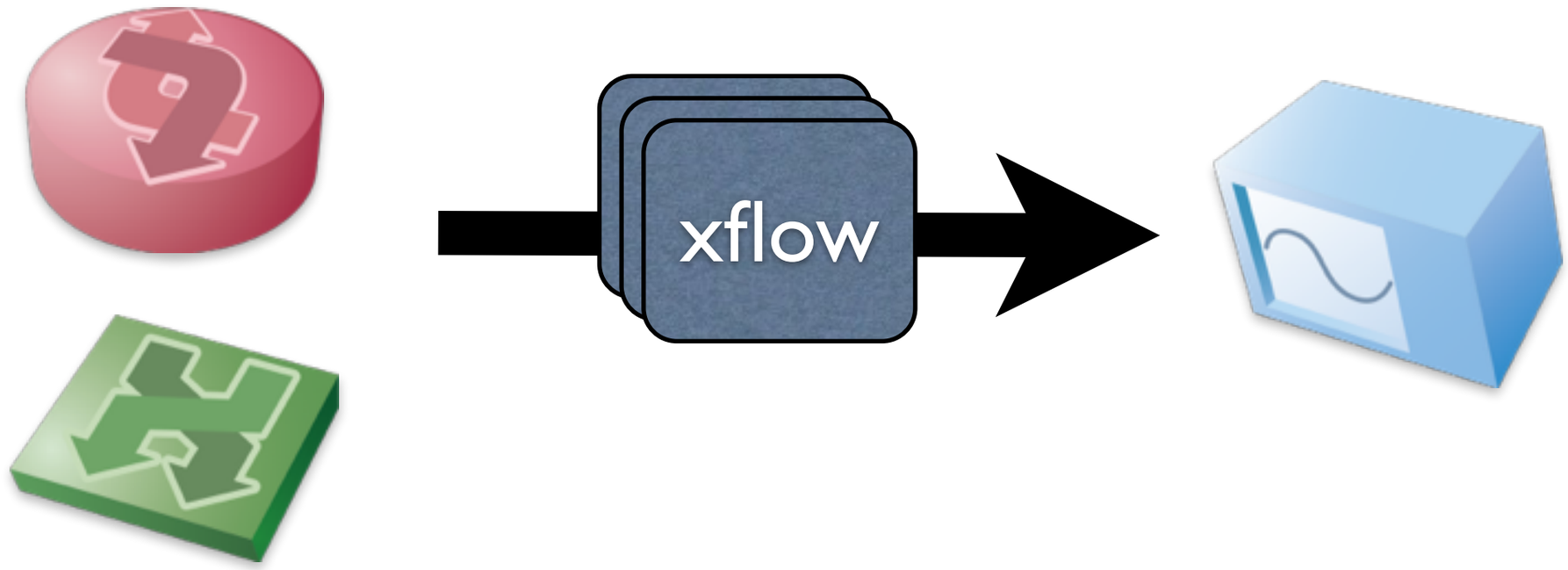
- パケットを拾い出してsFlowとして送る
- 複数パケットのアグリゲーションはしない

sFlowの特徴

- サンプリング・レートが内包される
  - コレクタ側で自動設定が容易
- 小型NW箱やUTM,Firewallなど**sFlow**出力可能な機器が多い
- SNMPライクなカウンターのサポート

コレクター

コレクタって?



エクスポータ

コレクタ

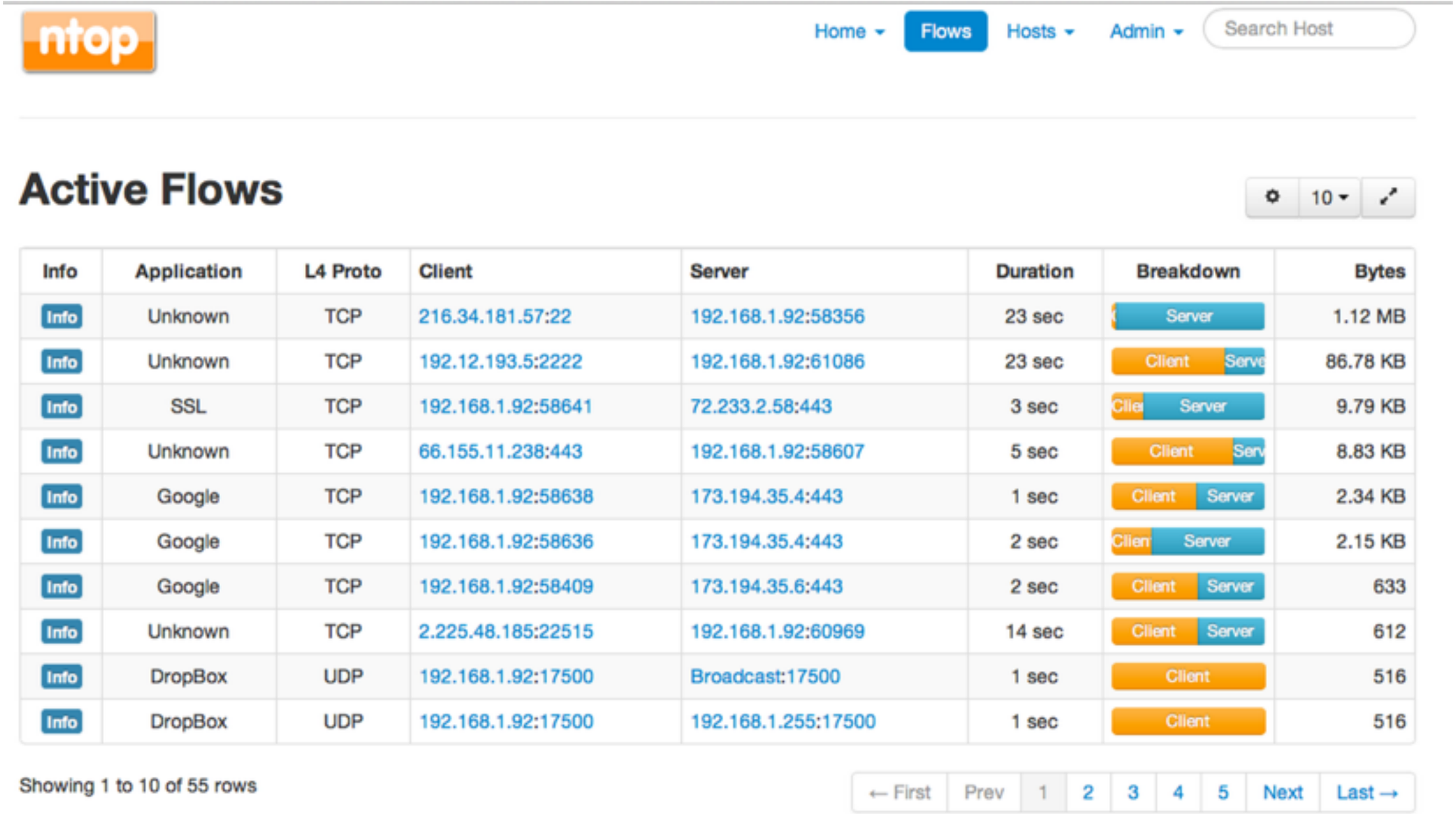
*ここではアナライザや描画ツールも含めてコレクタとします。

コレクタあれこれ

- OpenSourceなもの
  - FlowViewer,nfdump,nfsen,ntop,pmacct,sflow-tools, etc.
  - お手軽に試せる
- 商用なもの
  - ノウハウと時間を買いたい人向け

とりあえず絵が見たい人は

ntop



http://www.ntop.org/wp-content/uploads/2013/06/ActiveFlows.png?w=809

加工したい人向け : nfdump

- xFlowをキャプチャ

```
% mkdir work/nf
% nfcapd -w -t 60 -D -l work/nf -p 2055
```

- xFlowを見る

```
% nfdump -R work/nf
```

Date flow start	Duration	Proto	Src IP Addr:Port		Dst IP Addr:Port	Packets	Bytes	Flows
2015-07-14 00:43:21.341	1014.755	UDP	100.0.0.3:3	->	10.0.0.1:18	1738	2522	1
2015-07-14 00:31:16.821	1739.275	UDP	100.0.0.3:4	->	10.0.0.1:19	1328	4951	1
2015-07-14 00:30:05.421	1810.675	UDP	100.0.0.3:5	->	10.0.0.1:20	1218	3650	1
2015-07-14 00:29:39.098	1836.998	UDP	100.0.0.3:6	->	10.0.0.1:0	1871	9765	1
2015-07-14 00:38:27.524	1309.572	UDP	100.0.0.3:7	->	10.0.0.1:1	1572	6852	1
2015-07-14 00:32:13.685	1683.411	UDP	100.0.0.3:8	->	10.0.0.1:2	1478	8956	1

加工したい人向け : nfdump

- xFlowをキャプチャ

```
% mkdir work/nf
% nfcapd -w -t 60 -D -l work/nf -p 2055
```

- xFlowを見る

```
% nfdump -R work/nf
```

Date flow start	Duration	Proto	Src IP Addr:Port		Dst IP Addr:Port	Packets	Bytes	Flows
2015-07-14 00:43:21.341	1014.755	UDP	100.0.0.3:3	->	10.0.0.1:18	1738	2522	1
2015-07-14 00:31:16.821	1739.275	UDP	100.0.0.3:4	->	10.0.0.1:19	1328	4951	1
2015-07-14 00:30:05.421	1810.675	UDP	100.0.0.3:5	->	10.0.0.1:20	1218	3650	1
2015-07-14 00:29:39.098	1836.998	UDP	100.0.0.3:6	->	10.0.0.1:0	1871	9765	1
2015-07-14 00:38:27.524	1309.572	UDP	100.0.0.3:7	->	10.0.0.1:1	1572	6852	1
2015-07-14 00:32:13.685	1683.411	UDP	100.0.0.3:8	->	10.0.0.1:2	1478	8956	1

加工したい人向け:nfdump (cont.)

- 暴れてる子Top10を見つける

```
% nfdump -s srcip/bytes -n 10 -R work/nf/

Top 10 Src IP Addr ordered by bytes:

Date first seen          Duration Proto      Src IP Addr   Flows(%)      Packets(%)       Bytes(%)         pps      bps   bpp
2015-07-14 00:26:58.133  2465.453 any        100.0.0.3     89115(19.7)   133.7 M(19.7)    488.5 M(19.6)    54212    1.6 M    3
2015-07-14 00:35:58.527  3358.072 any        10.0.0.6      17468( 3.9)   26.2 M( 3.9)     96.5 M( 3.9)     7809     229897   3
2015-07-14 00:35:49.543  3367.056 any        10.0.0.19     17436( 3.8)   26.1 M( 3.8)     96.1 M( 3.9)     7752     228428   3
2015-07-14 00:36:31.065  3325.534 any        10.0.0.15     17543( 3.9)   26.3 M( 3.9)     96.1 M( 3.9)     7920     231070   3
2015-07-14 00:34:12.124  3464.475 any        10.0.0.3      17335( 3.8)   26.0 M( 3.8)     95.2 M( 3.8)     7494     219849   3
°°°°°
```

加工したい人向け:nfdump (cont.)

- 暴れてる子 Top10を CSVで出す

% nfdump -s srcip -n 10 -R work/nf/ -o csv

ts,te,td,pr,val,fl,flP,ipkt,ipktP,ibyt,ibytP,pps,pbs,bpp

2015-07-14 00:26:58,2015-07-14 01:08:03,2465.453,any,100.0.0.3,89115,16.7,133657656,16.7,488520741,16.7,54212,1585171,3

2015-07-14 00:35:26,2015-07-14 01:36:56,3690.504,any,10.0.0.9,21784,4.1,32632006,4.1,119204149,4.1,8842,258401,3

2015-07-14 00:36:31,2015-07-14 01:36:56,3625.547,any,10.0.0.15,21712,4.1,32556340,4.1,118899952,4.1,8979,262360,3

2015-07-14 00:35:58,2015-07-14 01:36:56,3658.085,any,10.0.0.6,21656,4.1,32513734,4.1,119911530,4.1,8888,262238,3

2015-07-14 00:35:49,2015-07-14 01:36:56,3667.069,any,10.0.0.19,21606,4.1,32335801,4.1,118764627,4.1,8817,259094,3

2015-07-14 00:36:07,2015-07-14 01:36:56,3649.070,any,10.0.0.1,21545,4.0,32325530,4.1,118425920,4.1,8858,259629,3

ooooo

pmacct もイケてる

% pmacct -s -T bytes -p /tmp/a.pipe

TAG	SRC_IP	DST_IP	SRC_PORT	DST_PORT	PACKETS	BYTES
110	100.0.0.2	10.0.0.12	12347	54	148748831	5480618878
110	100.0.0.2	10.0.0.12	12346	54	148679571	5462237366
110	100.0.0.1	10.0.0.11	12345	53	148772815	5461617413
110	100.0.0.2	10.0.0.12	12345	54	148814637	5454732403

°°°°°

商用なコレクタ

- ノウハウと時間を買いたい人向け

FAQとtips

(Q) 対外トラヒック見たいから対外IFだけxflowを有効にすればいいんだよね?

- (A) だめです(例外もあり)

内側IFも含めて、見たいトラヒックが通るIFすべてでxflowを有効にしてください。

忘れがちな基本

- IFにxFlow設定が無いと xFlowが出ません。
  - 必ず全部のIFにxFlow設定を入れる必要は
ありません。

管理系には
たいてい不要

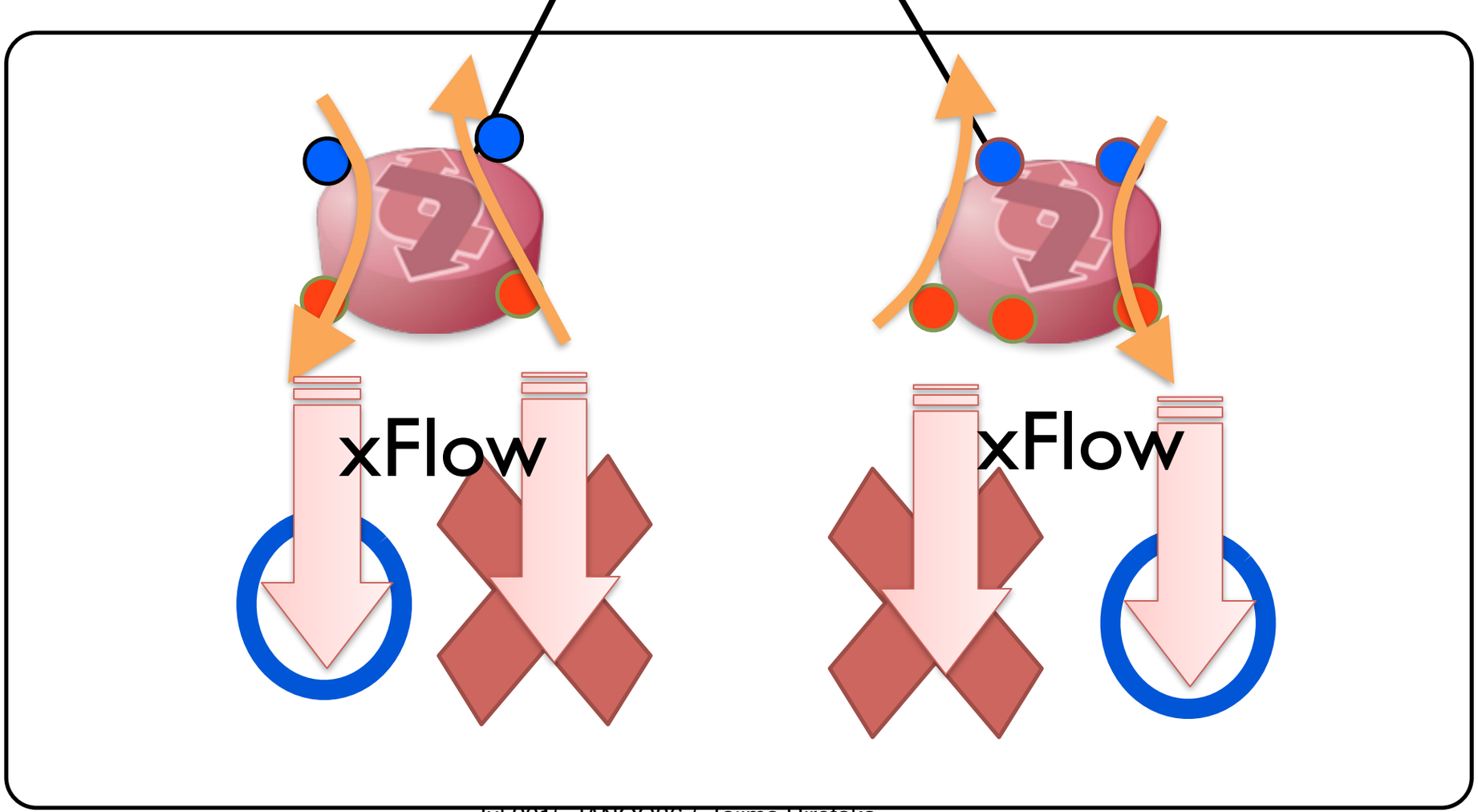
```
interfaces {
    xe-1/0/0 {
        unit 0 {
            family inet {
                address 192.168.1.1/24;
                filter {
                    input cflowd;
                }}}
    fxp0 {
        unit 0 {
            family inet {
                address 10.0.0.1/24;
    }}}}
```

なのでこういうのはダメ(*1)

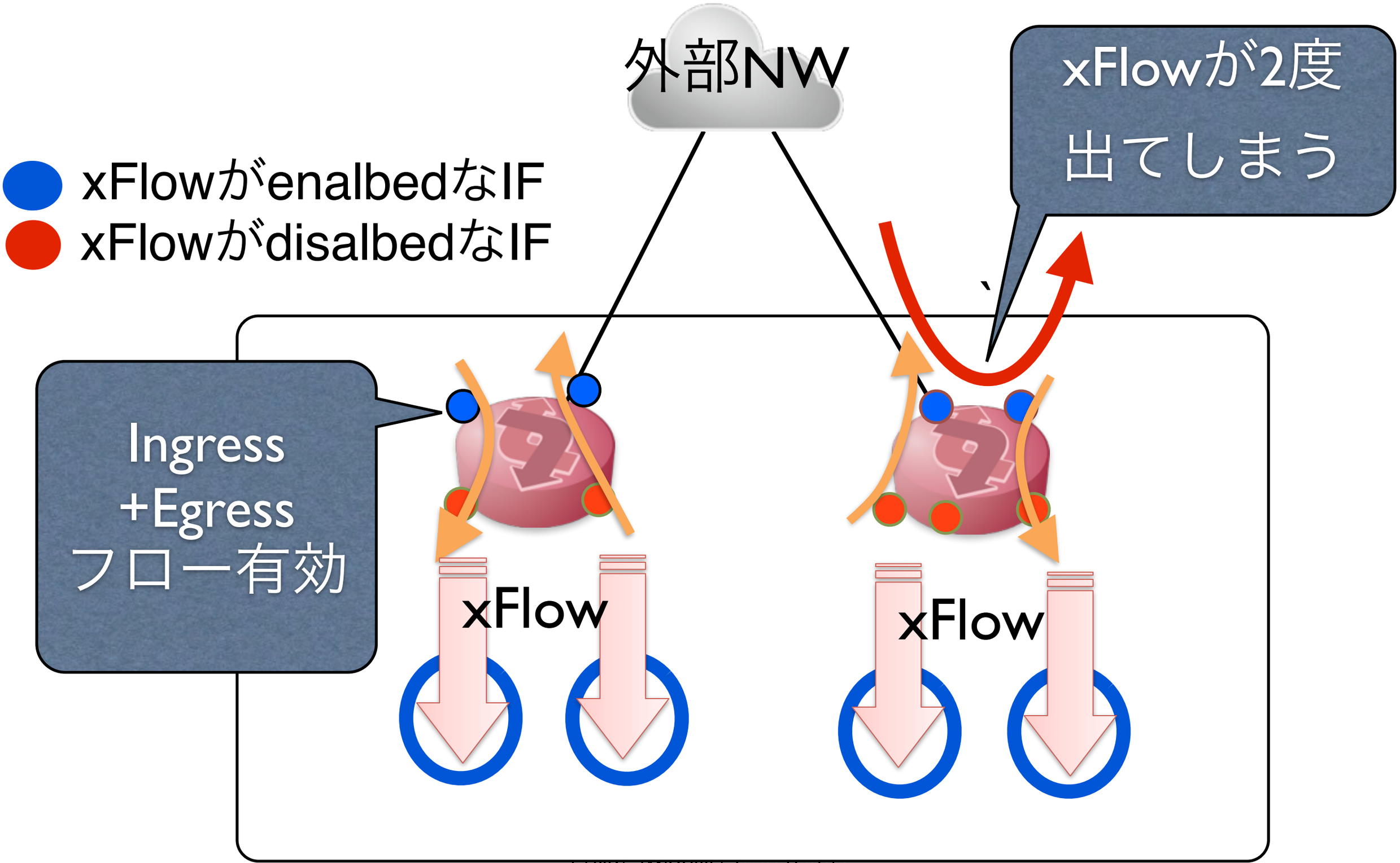
外部NW

(*1)ingressでflow有効の場合

- xFlowがenalbedなIF
- xFlowがdisalbedなIF



例外もあるけど、おすすめしない



(Q)サンプリングレートはどれくらいがよい?

- (A)理論的な計算式はあります。

  理論値と実測値で調整しましょう。

サンプリングレートの考え方

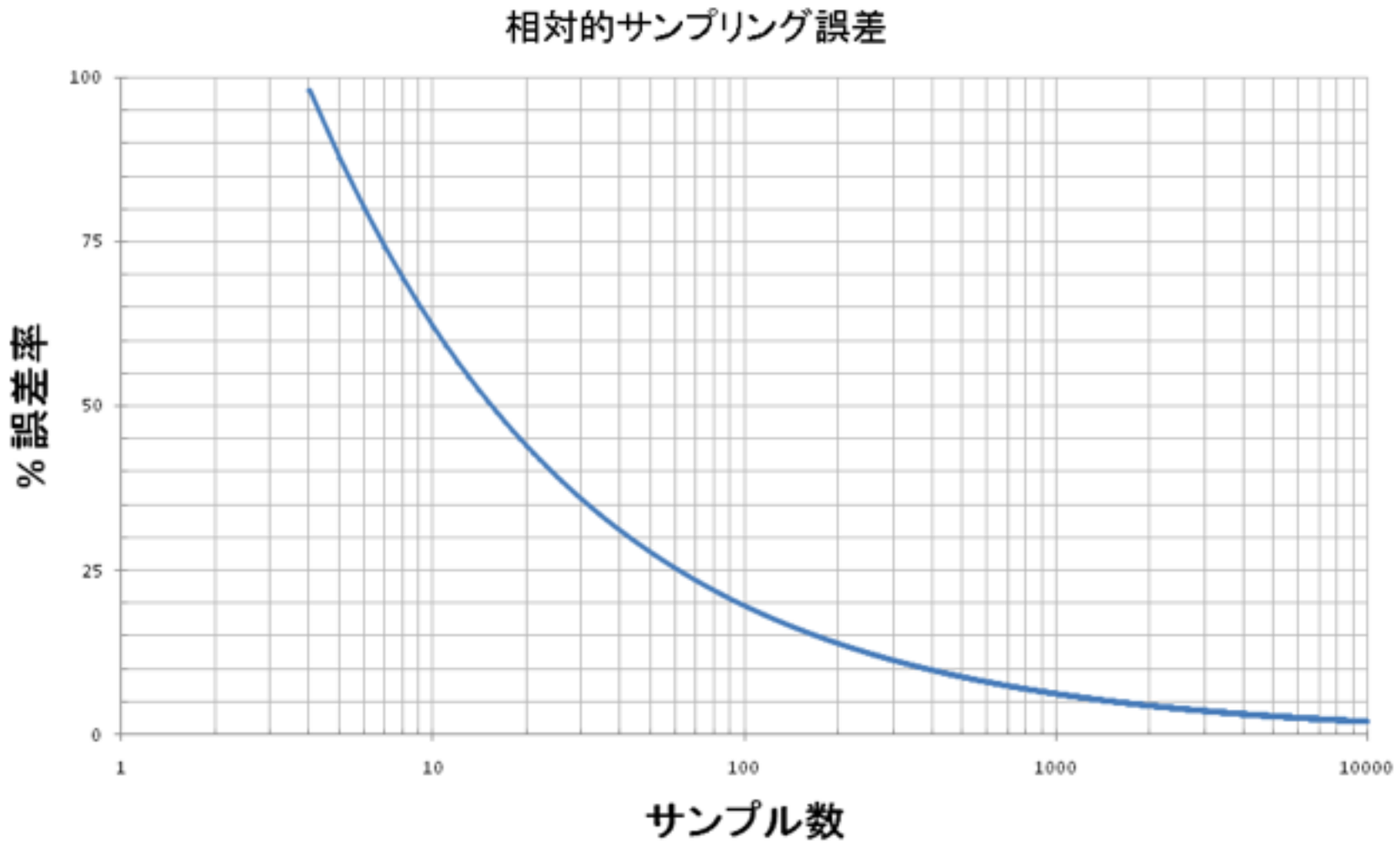
–サンプル数が多いほど精度が上がります

- e.g. 1分あたりサンプル数10個より20個が高精度
- 低トラヒック(～数Mbps)では高レートが必要
- でも必要以上な高レートは無駄になることも

サンプリング理論

- 誤差率= 196×sqrt(1/c)

※ c: サンプル数
(=集めたパケット数)



注: 信頼区間95%の場合

誤差率はレートでなくサンプル数に依存すること。

計算例

- 1Gbpsが流れてるIFを誤差1%で見たい。

(STEP1)必要なサンプル数(パケット数)を求める

誤差率1%にしたいので、最低必要なパケット数は

l=196×sqrt(1/c)　　→　　c=196^2=38416パケット

(STEP2)観測する周期毎に流れるパケット数を求める

パケットサイズが平均500Byteとすると、

PPS = 1Gbps/(500Byte×8)=250 kpps

観測周期が5分の場合

5分間に流れるパケット数= 250 kpps ×300sec=75 Mパケット

(STEP3)必要なサンプリングレートを求める

75Mパケット/35416パケット ≒ 1952

解:　1/1952 以上にすればよい。

でも理論値をよくみると。。。

	サンプリングレートの理論値						
誤差率%	100	200	300	400	500	600	700
0.1	98	49	33	24	20	16	14
0.5	2440	1220	813	610	488	407	349
1	9762	4881	3254	2440	1952	1627	1395
2	39046	19523	13015	9762	7809	6508	5578
3	87854	43927	29285	21964	17571	14642	12551
4	156185	78092	52062	39046	31237	26031	22312
5	244039	122019	81346	61010	48808	40673	34863

経験的には1000~10000が多いです

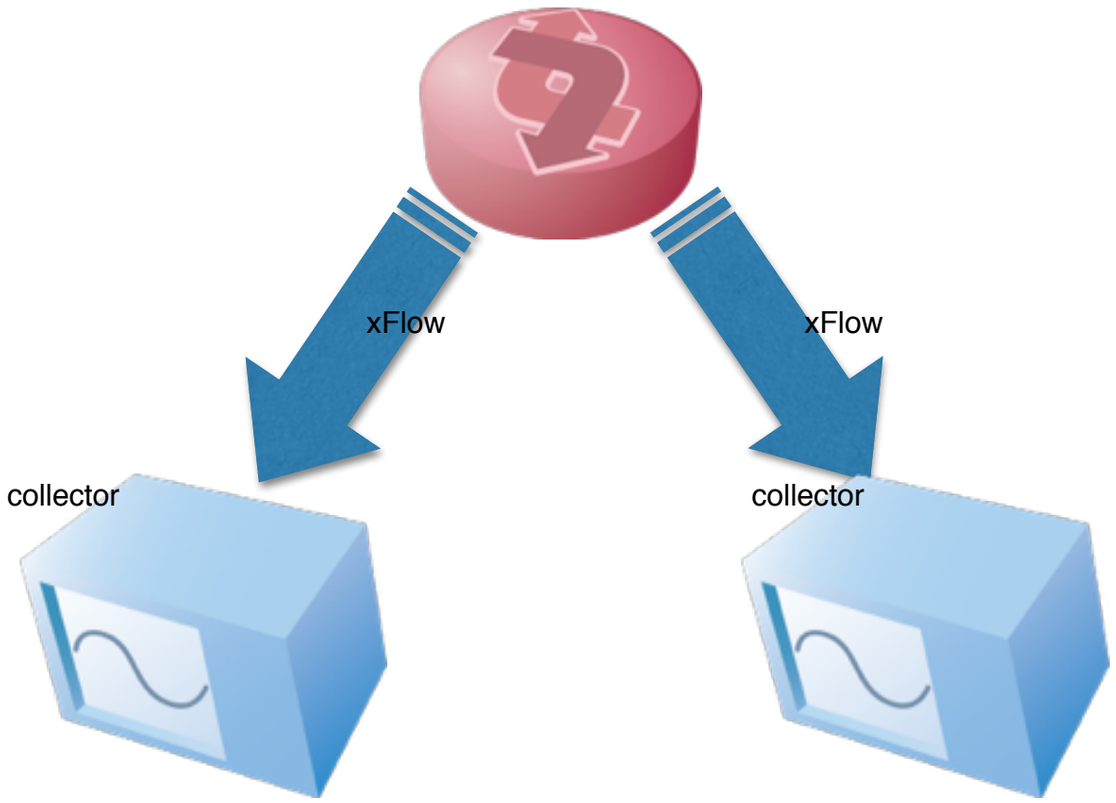
(Q) コレクタが複数あるので

それぞれにxFlowを食わせたい

- (A) いくつか方法があります
  - –方法1:エクスポーターが複数出す
  - –方法2:分岐ツールを使う
  - –方法3:コレクタでリレーする
  - –方法4:SWやtapで分岐する

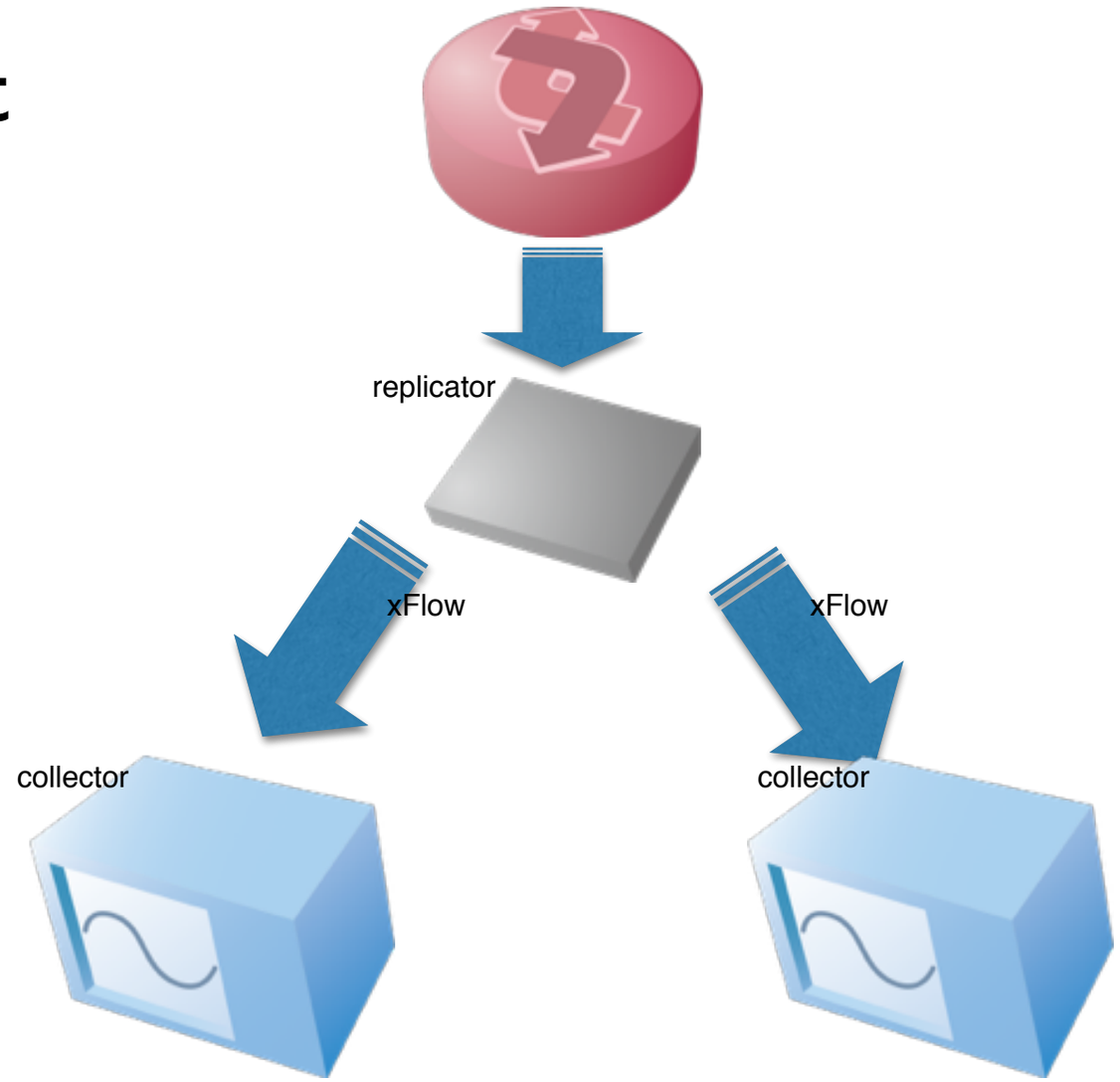
方法1:エクスポータが複数出す

- 多くのルータ/スイッチが複数の宛先にxFlowを出せます。
- ルータは上限2つが多いので3以上は。。。



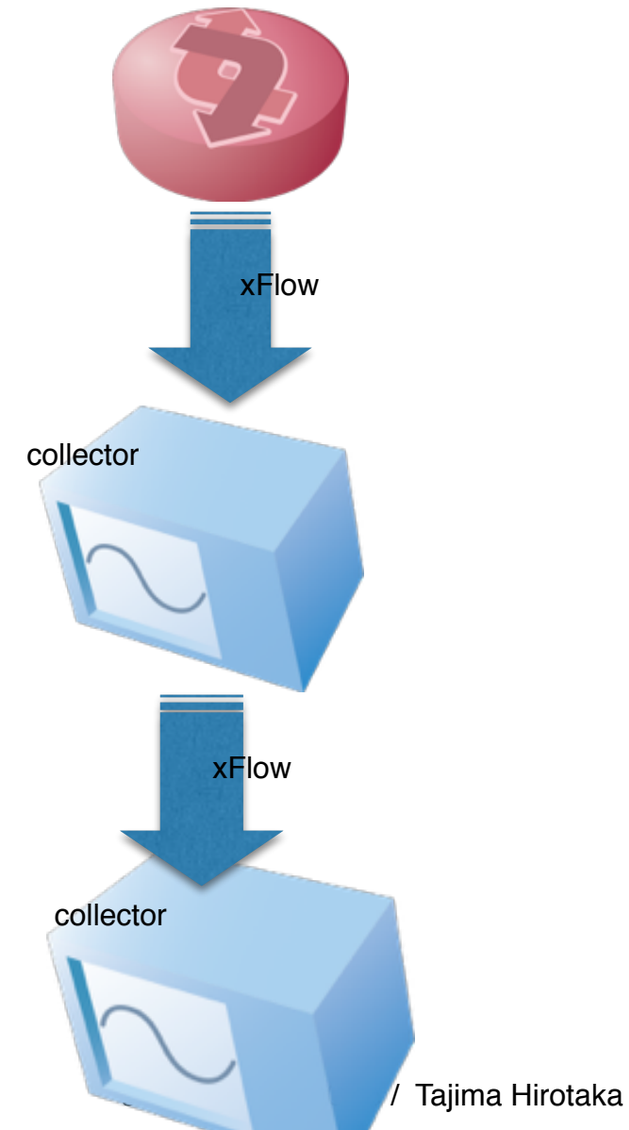
方法2:分岐ツールを使う

- xFlowパケットを分岐するツール(replicator)
  - flow-fanout
  - pmacct



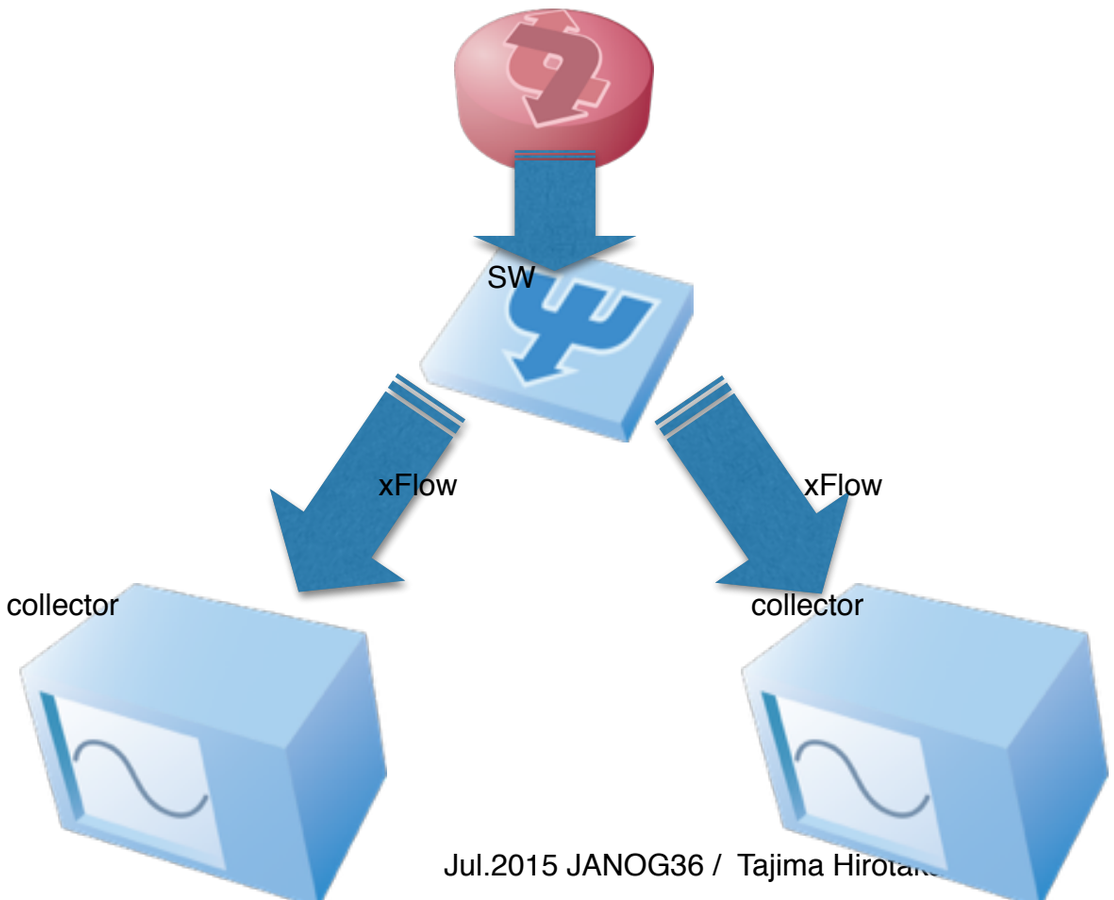
方法3:コレクタでリレーする

- xFlowパケットをコレクタでうけ、それを別コレクタにリレーする



方法4:SWやtapで分岐する

- ミラーやtap、OpenFlow SW等で分岐する
  - ミラー/tapはIPアドレスに注意
  - OF SWでUDPパケットを分岐する



others

- ifIndexはSNMPとxFlowで同じ?
- NAT越しのxFlow
- ICMPが全部echo replyに見えるのはなぜ?

- ルータ再起動したらトラヒックが消えた?

またどこかの機会で。

まとめ

- SNMPとは違う世界が見えます。
- ルータ1つ、Open Sourceなコレクタ1つから始めてみてはいかがでしょう。
- 習うより慣れよの世界です。

Questions?