

明日から使える NetFlow・sFlow 設定術

田島弘隆 進藤資訓

htajima@fivefront.com mshindo@fivefront.com

ファイブ・フロント(株)

NetFlow・sFlow、使ってるけど。。。。

- NetFlow,sFlow(以下xFlow)で〇〇したい。
- でも、どう設定すれば正解なのかワカラン。
- なんとなくサンプルコンフィグ通りに設定。
- なんとなくデータが見えてるし、まいっか。

で、ある時に気づく。

- あれ? MRTGのグラフよりズレてるなあ。
- あれ?トラヒックがインパルス状になってるぞ。
- あれ?いきなりグラフが見えなくなったぞ。
- あれ?。。。。etc

で、今日の趣旨

- 現場ですぐ使える xFlow の Know How を語る
 - 日本で一番 xFlow でハマってる(注1)我々が実際に遭遇した実例をもとにしてします。
 - JANOGらしく、技術ネタをどっぷりと☺
 - 仕様だけでは実感しにくいxFlowのパラメータを実践的に考えます。

(注1)「熱中」でなくて、いろんな罫にハマってる

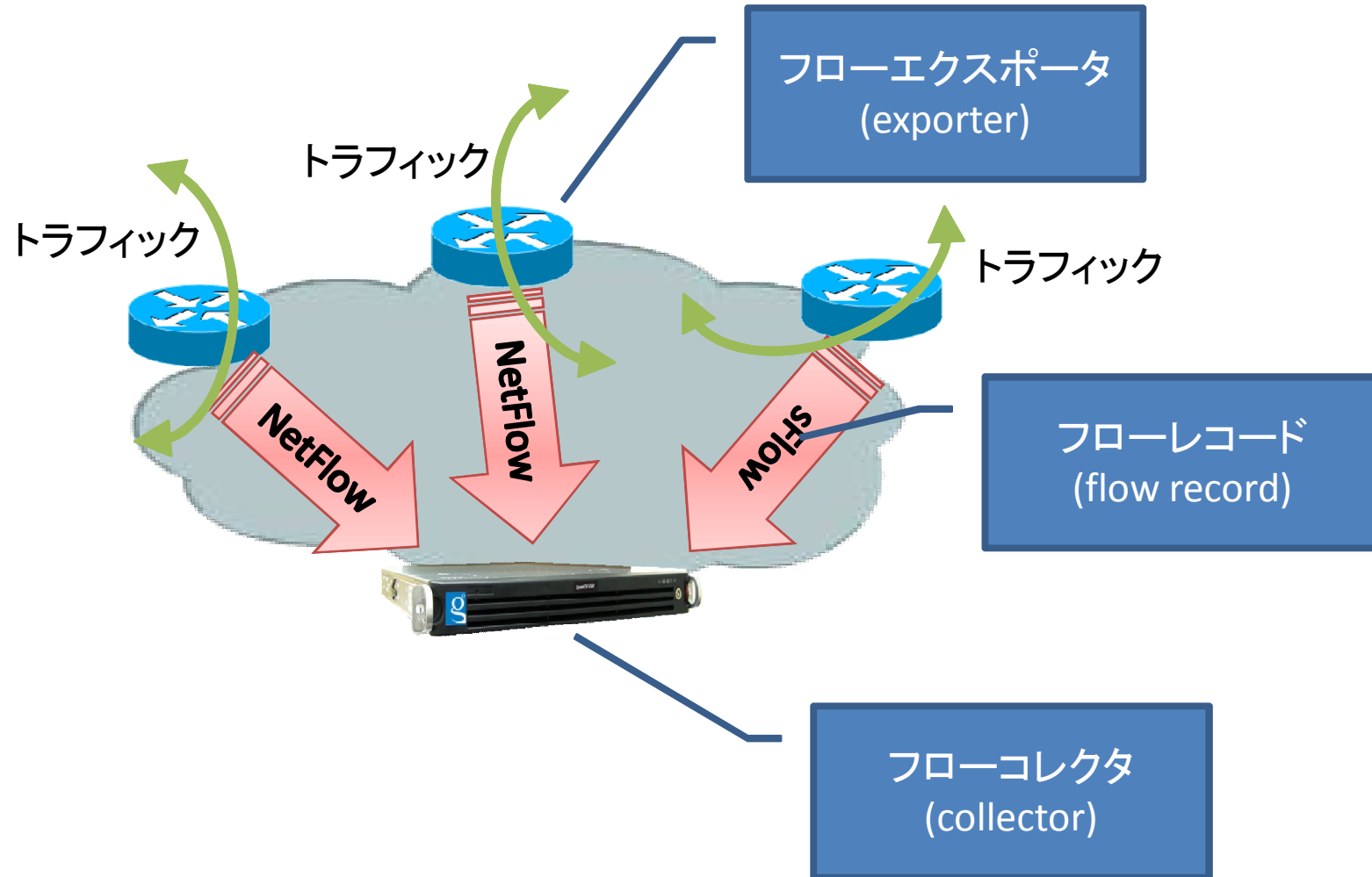
念のためのお約束

- 「明日使えるノウハウ」が趣旨なので、できるだけ実際の機器名やメーカー名を出します。
- でも、リリース時期やバージョン等によって、機器動作は当然違います。
- もちろん、特定機器にダメ出しする意図はまったくありません。
- 本稿を100%信じずに必ず検証してください。

xFlow 基本知識

省略

てのもなんなので、イメージ図



みおとしがちな**基本の基本**

- IFにxFlow設定が無いと xFlowが出ません。
 - ちなみに、必ず全部のIFにxFlow設定を入れる必要はありません。
 - 監視したいトラフィックが通過するIFにのみでOK。

```

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;
      }
    }
  }
}

```

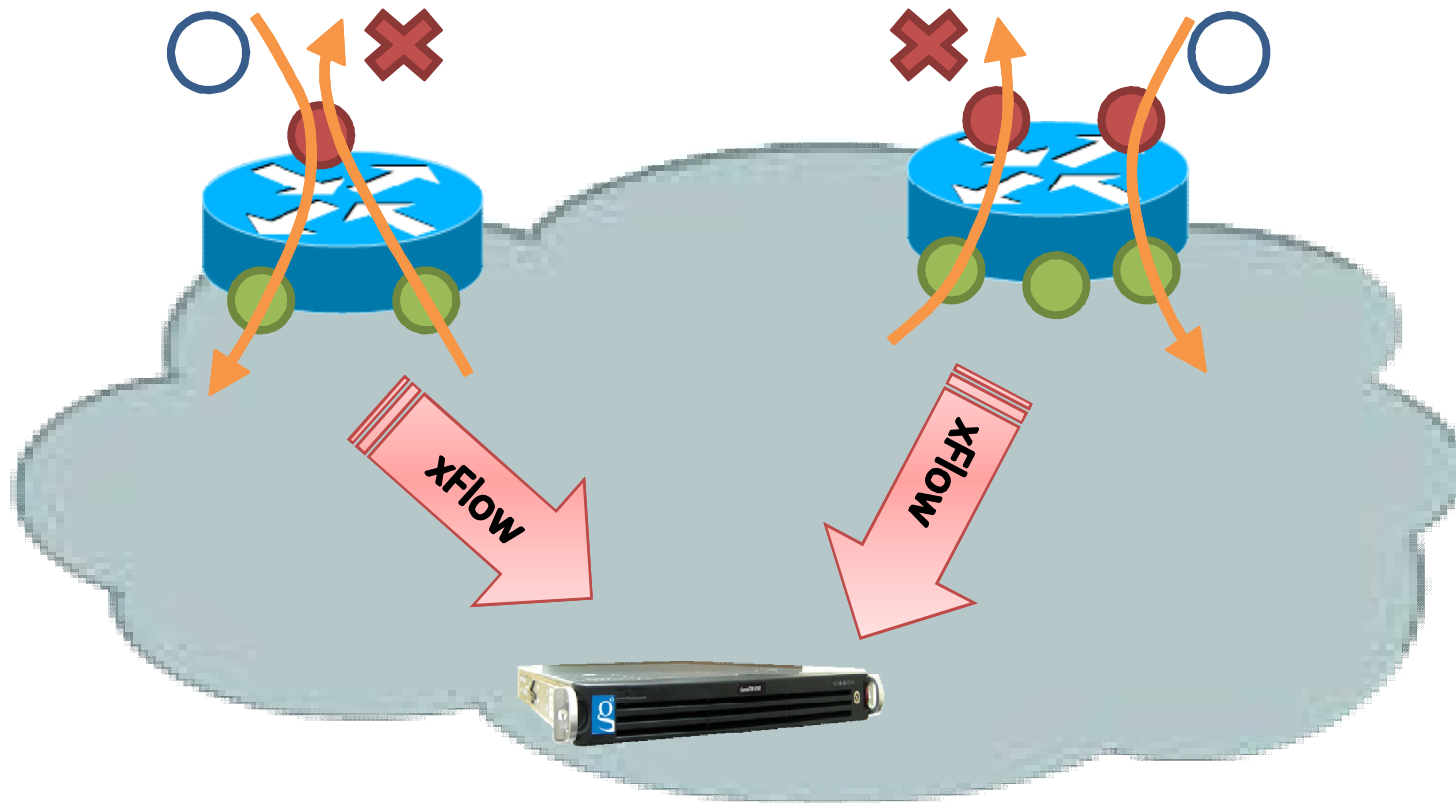

だからといって、例えば・・・

- うちにはボーダーのトラフィックしか興味ない
- 見たいトラフィックはすべてボーダーのIFを通過する
- よって、ボーダーのみでxFlowをenableにすればいい

フローレコード生成のトリガー

- 原則的に “Ingress” で効きます。
- まれに
 - “Egress” で効く奴
 - 設定で “Ingress” or “Egress” を指定できる奴がいたりします。

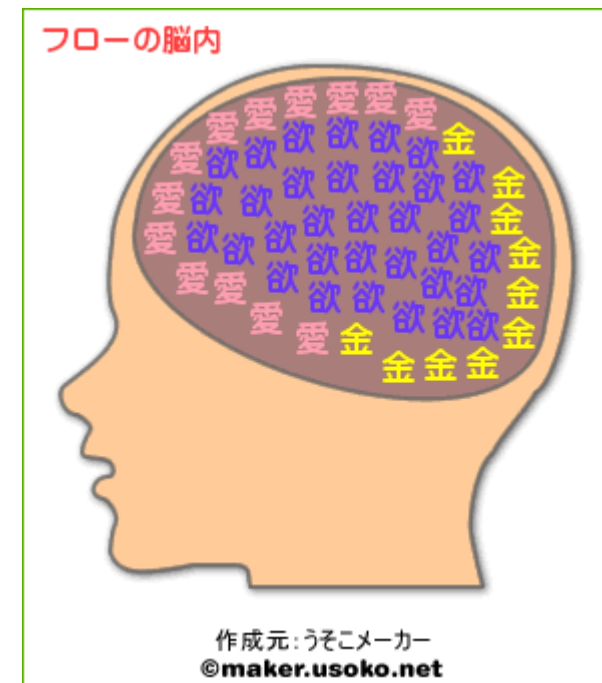
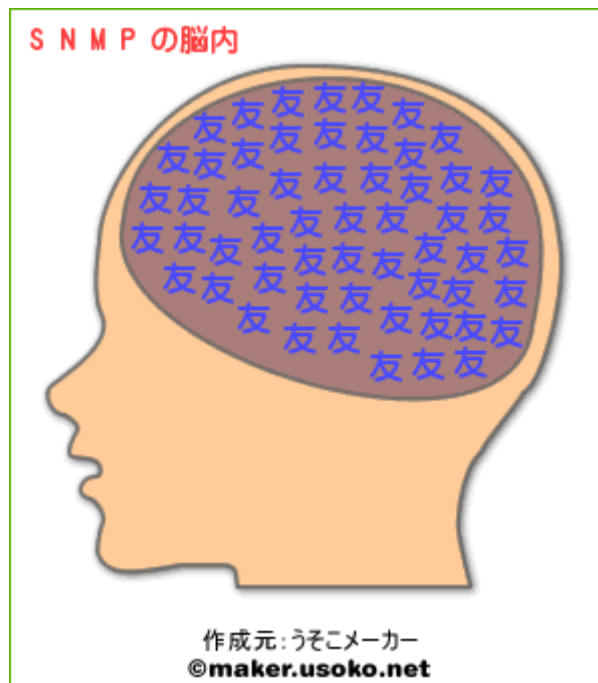
なので、こんなことになります



- xFlowがenabledなIF
- xFlowがdisabledなIF

パラダイムシフト

- みなさん、**SNMP脳**から**フロー脳**に切り替えましょう！



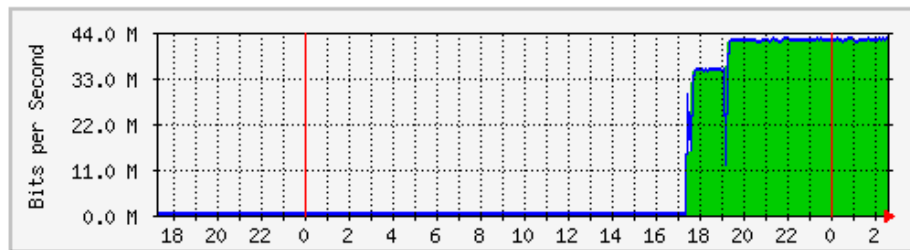
では、deep
な世界へ
どうぞ。

ケース1

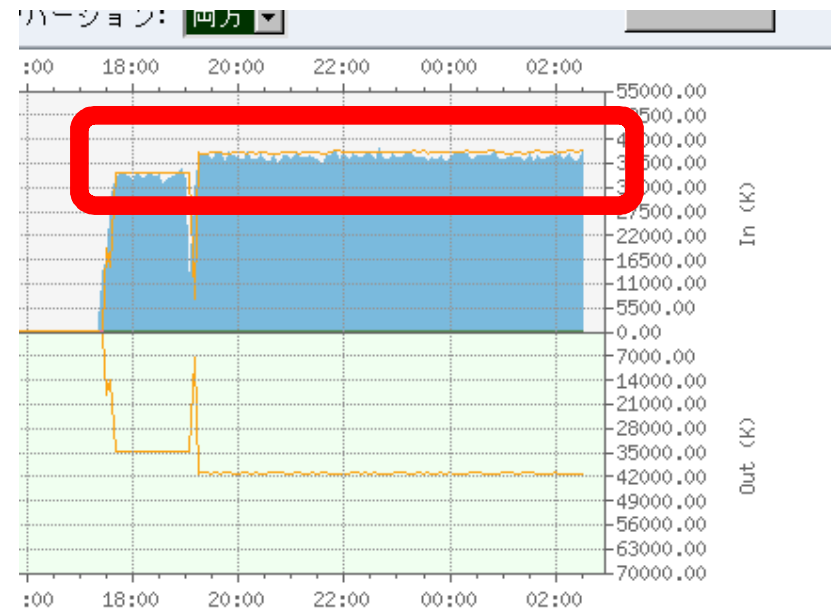
- MRTG(SNMP)のグラフと違うんですけど。。。。
 -トラフィック量がちょっとズレてる場合

更新日時 2009年1月19日(月) 2:35,
 'ax24'の稼働時間 175 days, 10:40:29.

日グラフ(5分間 平均)



	最大	平均	最新
受信	42.5 Mb/s (42.5%)	11.0 Mb/s (11.0%)	42.0 Mb/s (42.0%)
送信	42.5 Mb/s (42.5%)	11.1 Mb/s (11.1%)	42.1 Mb/s (42.1%)



ケース1:原因

1. サンプリングレートが低すぎる
2. SNMPとxFlowはL2?L3?問題
3. ルータで終端するトラヒックとマルチキャスト
4. その他の明日つかえるトリビア

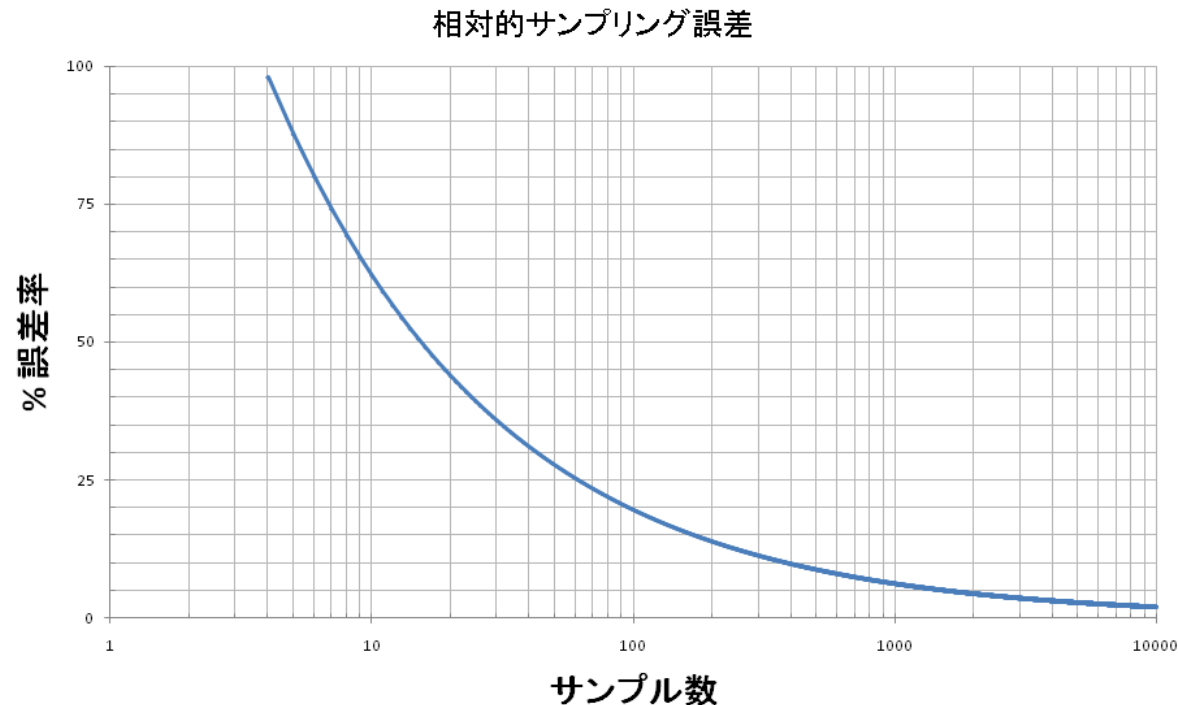
原因1:サンプリングレートが低すぎる

- サンプリングレート選定の大原則
 - サンプルをいくつ集めればよいか、で決める。
(レート値は必要サンプル数から導かれる結果論である)
- すごく簡単にいうと
 - サンプルをたくさん集めるほど精度が高くなる
 - 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CPUのむだつかい。

サンプリング理論

- 誤差率 = $196 \times \sqrt{1/c}$

※c: サンプル数
(=集めたパケット数)
※詳細は参考文献にて。



注: 信頼区間
95%の場合

注目すべきは、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
 $1=196 \times \sqrt{1/c} \quad \rightarrow \quad c=196^2=38416$ **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、
 $PPS = 1Gbps / (500Byte \times 8) = 250$ kpps

観測周期が5分の場合
 5分間に流れるパケット数 = $250 \text{ kpps} \times 300\text{sec} = 75$ **Mパケット**

(STEP3)必要なサンプリングレートを求める
 $75M \text{パケット} / 38416 \text{パケット} \doteq 1952$

解: $1/1952$ 以上にすればよい。

ただし、あくまでも理論値。

		サンプリングレートの理論値						
誤差率% ＼パケットサイズ	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			39046	31237	26031	22312	
5	244039			61010	48808	40673	34863	

1/7809にしても
誤差はわずか2%

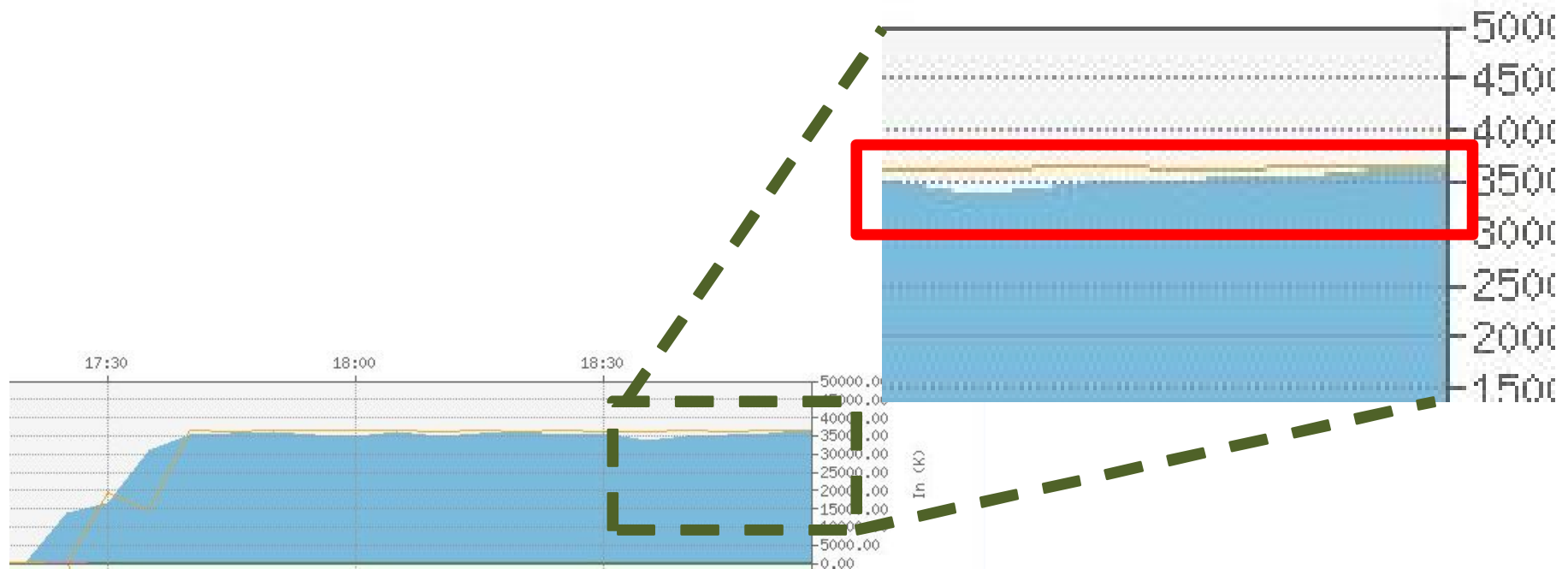
エクスポートの負荷と求める精度の
バランスを考えましょう。

原因2:SNMPとxFlowはL2?L3?問題

- 取得できるトラフィック量の違い
 - SNMP はL2 (フレーム長)
 - xFlow はL3 (パケット長)
- 一般的に、xFlowによるトラフィックはMRTGより少なめに表示される。

原因2:SNMPとxFlowはL2?L3?問題

(cont.)



ダウンロードExcel-

		▼ In	▼ Out	▼ 総和	
<input checked="" type="checkbox"/>	—	Flow Traffic(bps)	25.63M	0.00	25.63M
<input checked="" type="checkbox"/>	—	SNMP Traffic(bps)	25.03M	23.95M	48.98M
<input checked="" type="checkbox"/>	—	Flow Traffic(pps)	6,128.21	0.00	6,128.21
<input checked="" type="checkbox"/>	—	SNMP Traffic(pps)	5,831.32	5.46	5,836.78

2009/1/23

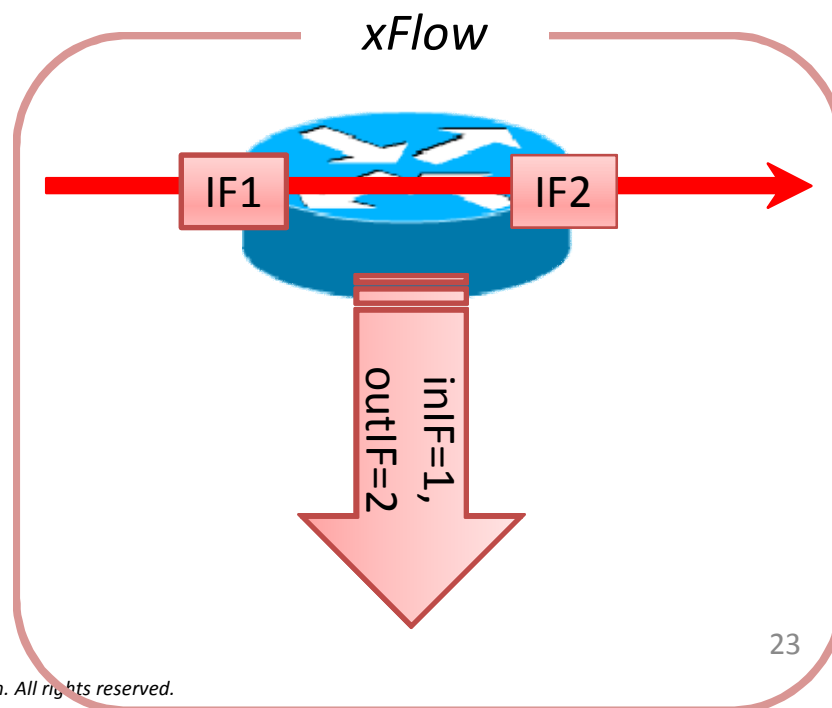
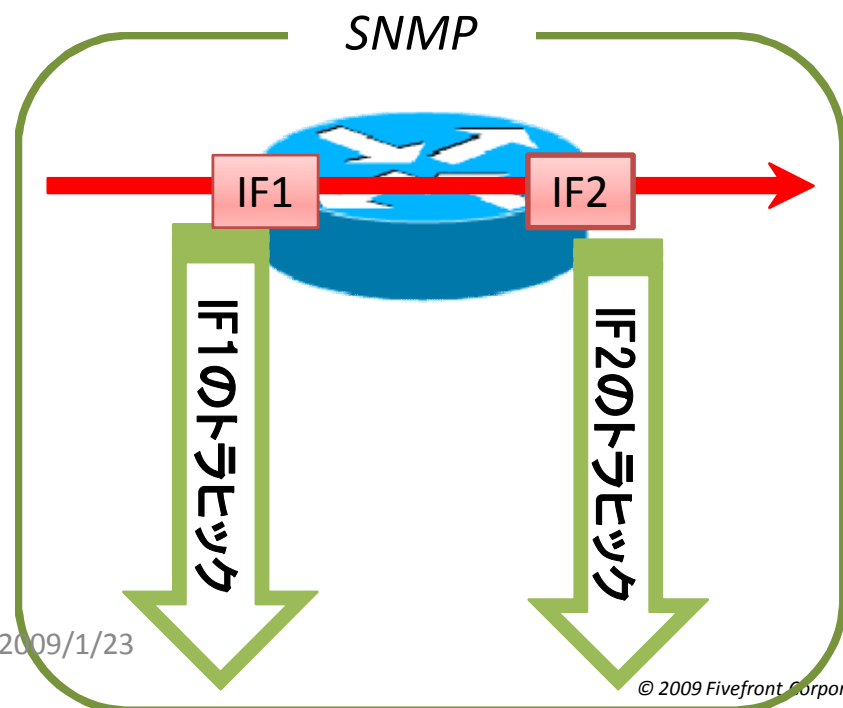
原因2:SNMPとxFlowはL2?L3?問題

(.cont)

- ただし例外もある!
 - Juniperの論理I/FのSNMPはL3で答えてくれる。
 - よってJuniperの論理I/Fはほぼ同じ値になる。

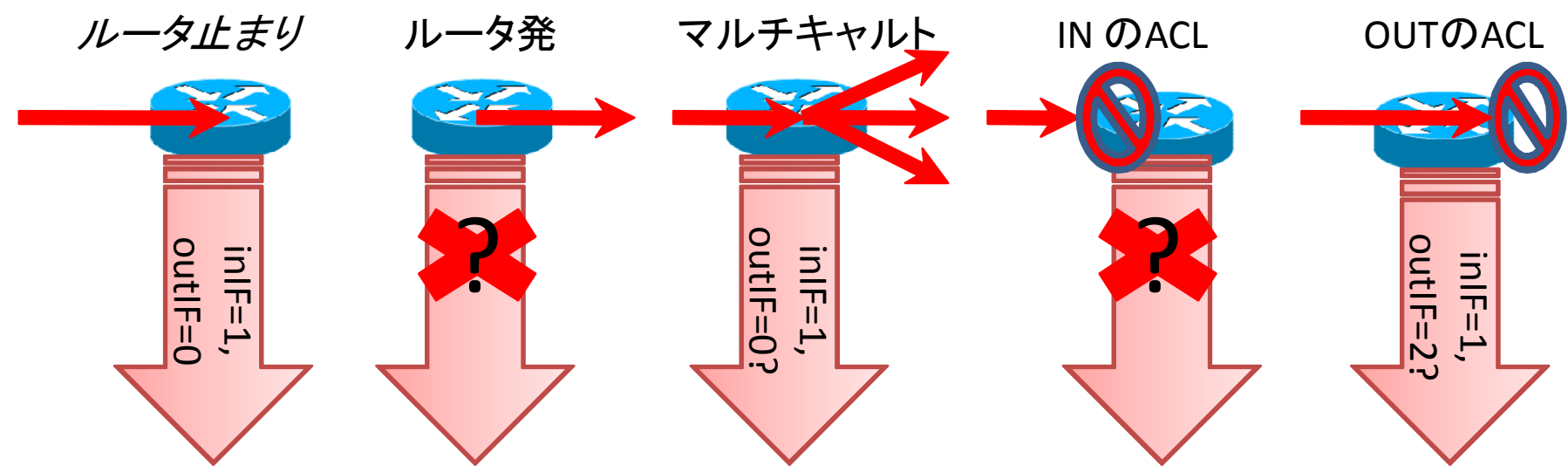
原因3:ルータで終端するトラヒックと マルチキャスト

- 提供されるトラヒック情報の**素性**が違う
 - SNMPは各IF毎のトラヒック量を提供
 - xFlowはIn&Out IF番号とトラヒック量を提供



原因3:ルータで終端するトラフィックとマルチキャスト(cont.)

xFlowのイメージ(Ingressの場合)



- SNMPではいずれもIF毎に計上される(と思う)。

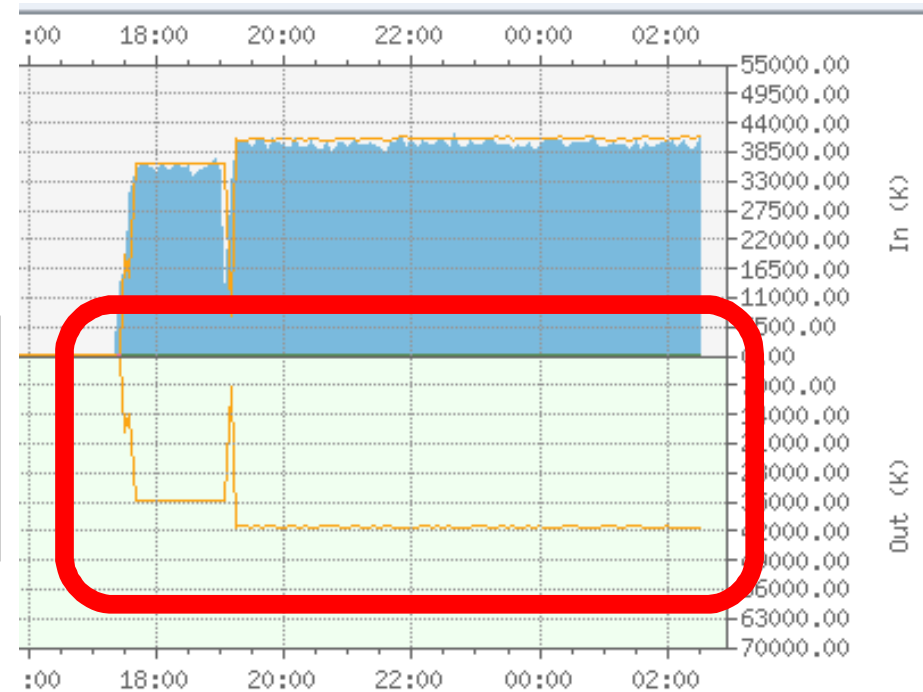
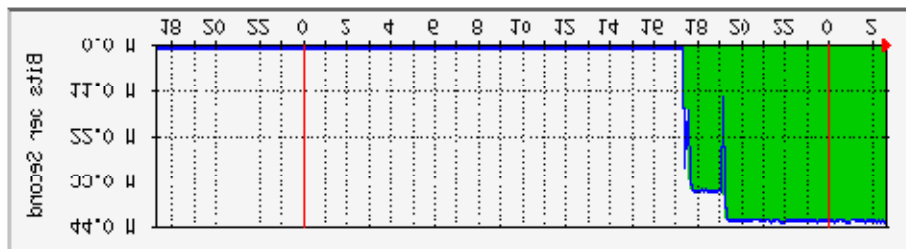
他にもこんなトリビアが。

- ほかにもある要素
 - IP以外のトラヒック (AppleTalk等)
 - V6がのらないxFlow を使用している
 - xFlow実装の機器による違い
 - etc

サンプリングレートが原因と考えがちですが、他の要素も疑ってみてください。

ケース2

- MRTG(SNMP)のグラフと違うんです part2。
 - 描画されないグラフがある。



ケース2:原因

1. 論理I/Fと物理I/Fの違い
2. 出力のIF番号がゼロになるエクスポート
3. その他 飲み会で使えるスベらない話

原因1:物理I/Fと論理I/F

- 論理I/Fのフロー情報が出ない実装もある。
 - 「SNMP getできる = flow情報が採れる」、**ではない!**
 - VLAN I/F、バンドルI/F は特に注意。

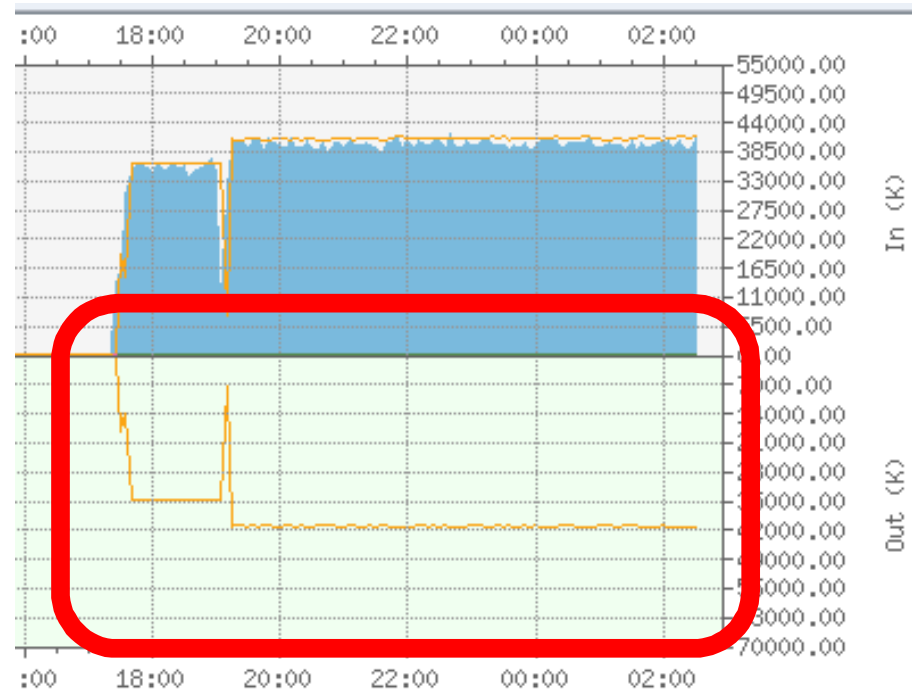
ダウンロードExcel-XM

All	インターフェース名	インターフェース説明	Flw In	SNMP In
<input checked="" type="checkbox"/>	TenGigE0/4/0/0	TenGigE0/4/0/0	9,797.89	51.63K
<input checked="" type="checkbox"/>	Bundle-Ether1.61	Bundle-Ether1.61	2,502.54	3,544.56
<input checked="" type="checkbox"/>	Dundle-Ether1.49	Dundle-Ether1.49	4,072.76	3,812.12
<input checked="" type="checkbox"/>	TenGigE0/4/0/6	TenGigE0/4/0/6	2,535.42	263.64K
<input checked="" type="checkbox"/>	Bundle-Ether1.65	Bundle-Ether1.65	1,988.00	2,902.83
<input type="checkbox"/>	Bundle-Ether1.63	Bundle-Ether1.63	1,520.86	1,908.36
<input type="checkbox"/>	Dundle-Ether1.166	Dundle-Ether1.166	0.00	14.79K
<input type="checkbox"/>	TenGigE0/4/0/4	TenGigE0/4/0/4	0.00	722.54K
<input type="checkbox"/>	Bundle-Ether1.164	Bundle-Ether1.164	0.00	143.33K
<input type="checkbox"/>	Bundle-Ether1	Bundle-Ether1	0.00	839.03K
<input type="checkbox"/>	Bundle-Ether1.162	Bundle-Ether1.162	0.00	663.57K

Sample

原因2:出力IFがゼロになるエクスポート

- L2製品は一般的に出力IF番号がゼロ。
- Catalyst系で flow mask 設定がないとゼロ。
 - mls flow ip interface-full
- SNMPは出力も当然カウントされる。
 - IPフローの入力と出力は関係ないから。

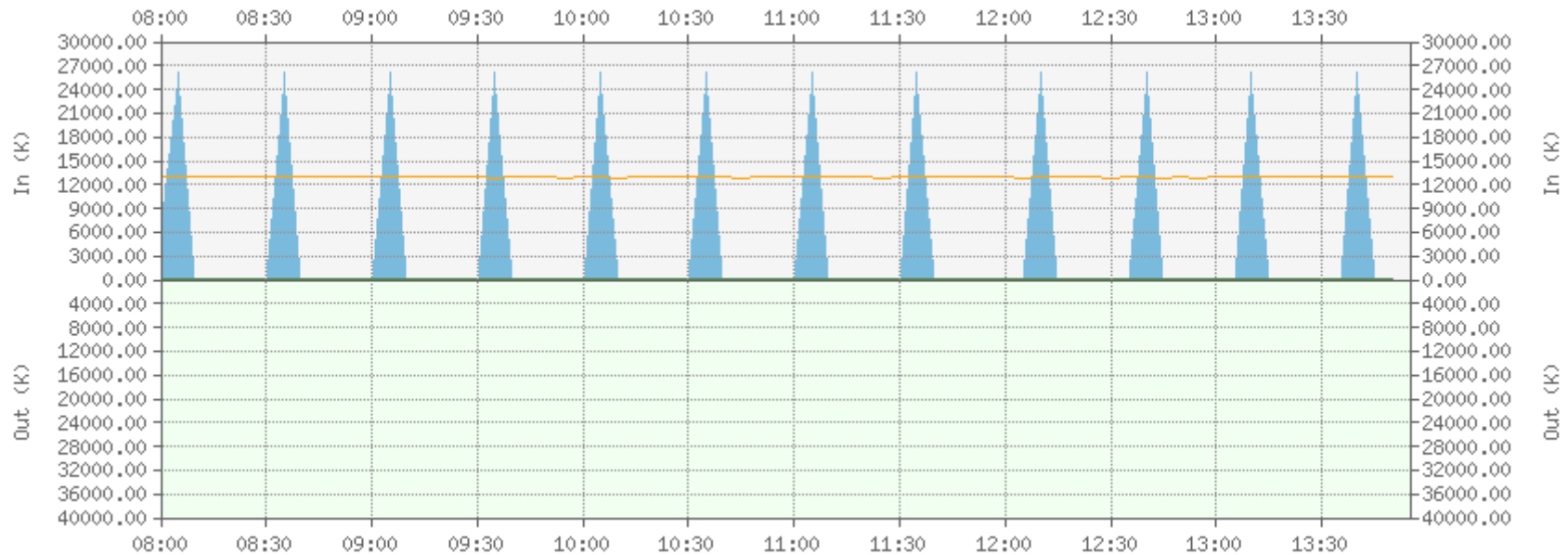


他にもこんな飲み会ネタが。

- IFにxFlow設定が入っていない。
- ルータの再起動で ifIndexが変ってしまった。
- SNMP と xFlow のifIndexが異なっていた。
- etc

ケース2:トラヒックがドカンと出てきた。

- しばらく観測されなかったトラヒックが、一定時間後にドカンとでてきた。



ケース2:原因

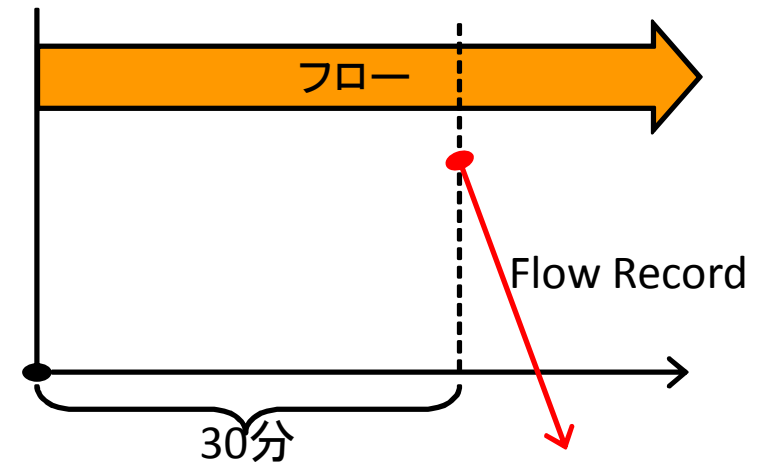
1. active-timeoutパラメータ
2. first-seenとlast-seenの利用方法

長寿命フロー

- 原則NetFlowはフローが終了した際にフローレコードが出る
 - TCPならFIN or RST
 - 一定時間当該フローが観測されなかった場合
- 長寿命なフローは (e.g.ストリーミング、P2P、検証時のgeneratorトラフィック、等)はどうなんねん？
 - さすがに1日後にフローレコードが出ても・・・。

active-timeoutパラメータ

- Active-timeoutで指定された時間以上継続したフロー情報をフラッシュする。
 – Ciscoのデフォルトは30分。



- スパイク問題解決策
 →active-timeoutを短かい時間(たとえば1分)にすればよい。

例: ip flow-cache timeout active 1

でも、ちょっと心配??

- 補足: active-timeout を短かくしたら、CPUが痛くなるんじゃない?

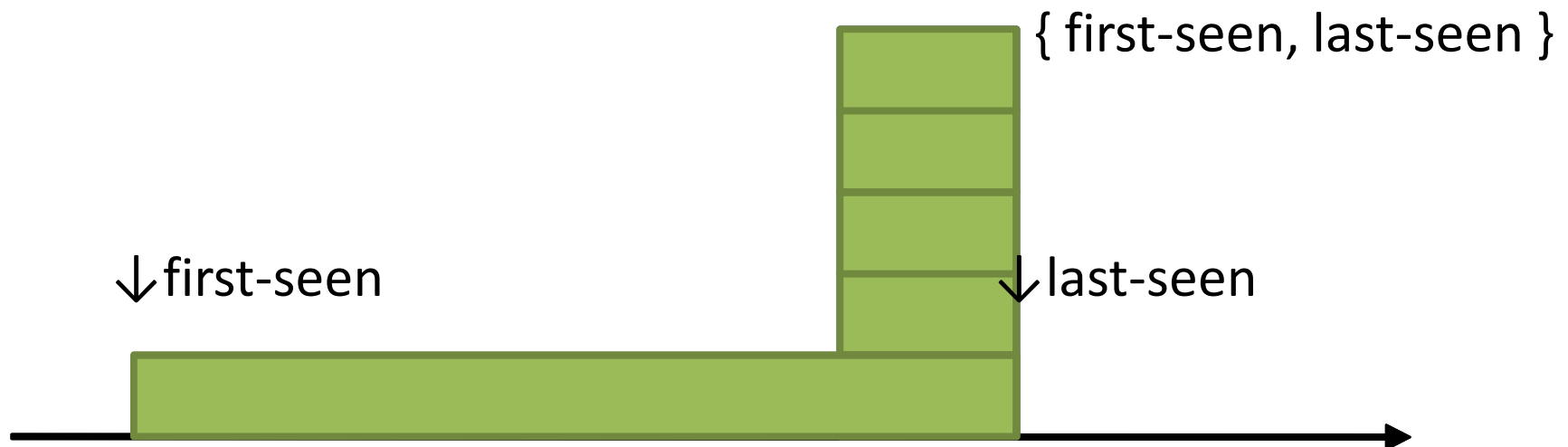
→ (大抵は)問題ない。

- 長寿命フローはトラフィックボリュームには寄与しているが、フロー数には寄与していない!

first-seenとlast-seenを使えば？

ムニアック

- フィールドの意味
 - first-seen: IPフローが初めて観測された時間
 - last-seen: IPフローが最後に観測された時間
- 理論的にはActive-timeoutが30分のときでもfirst-seenでグラフ補正が可能。



でも、難しいのよ(涙)

- でも、ほとんどのコレクタは{first,last}-seenは見ない。
 - 理由① 統計処理がえらい複雑になる。
 - 理由② 30分前のグラフが書き変わるのは現実的でない。
- active-timeoutを1分にするのが現実解。

まとめ

- xFlowのパラメータを理解して使いましょう。
 - サンプルコンフィグの鵜呑みは危険。
- エクスポートだけでなくコレクタの実装にも注意が必要です。
 - xFlowの仕様だけ読んでも不十分です。
- 脳みそ切り替えてください ☺

参考文献

- *Packet Sampling Basics*
<http://www.sflow.org/packetSamplingBasics/index.htm>
※勝手な日本語訳と補足説明:
http://www.fivefront.com/technology/sampling_theory/index.html
- フローベースのトラフィック計測と解析
<http://www.soi.wide.ad.jp/class/20060031/slides/51/>
- Flow最新情報 (注:もう古いです)
http://www.bugest.net/irs/docs_20060922/IRS10-flow-tajima-kokai.pdf