Use of Flow-Routing Combination

JANOG36 BoF

maoke@bbix.net
paolo@pmacct.net

Outline　概要

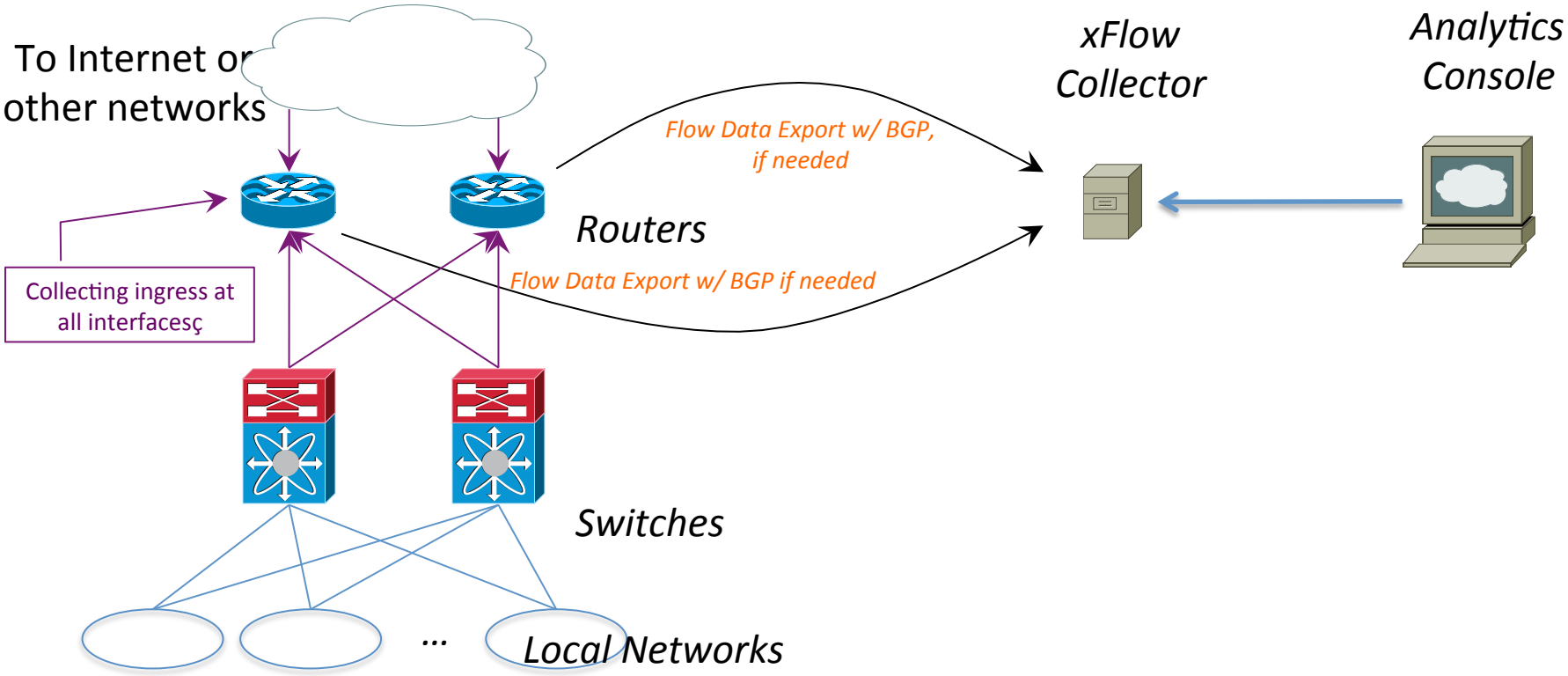
- Purpose of the BoF 目的
  - sharing information on open source flow tools, able to leveraging routing　オープンソースフローツールとローティング連携運用の共有
  - deeply discuss two use-cases　ユーズケースの深読み
    - traffic engineering in Netflix
      Netflixのトラフィックエンジニアリング事例
    - SDN content provision in Spotify
      SpotifyのSDNコンテンツ配信事例
- Contents　内容
  - maoke's introduction on flow collection pattern
    荒井の簡単紹介（10min）
  - Paolo's presentation on the use-cases
    Paoloさんのユーザケース詳細解説（30min x 2）
  - Open Discussion
    自由討論（20min）

Open Source オープンソース

- Why OS flow tool is needed? なぜOSフローツールがいるの
  - Easy deployment 実装の便利
    - Commodity HW is fine　高級な専門設備が要らず
    - Low cost　安い
    - Fast installment　設置まで早い
  - Flexibility　柔軟性
    - Able to customize (hacking)　機能のカストマイズが可能
    - Easy to extent and to combine with other components　他の既存機能と組み合わせて拡張でも可能
    - Easy to accommodate different usage environment 運用環境に適用できるように加減しやすい

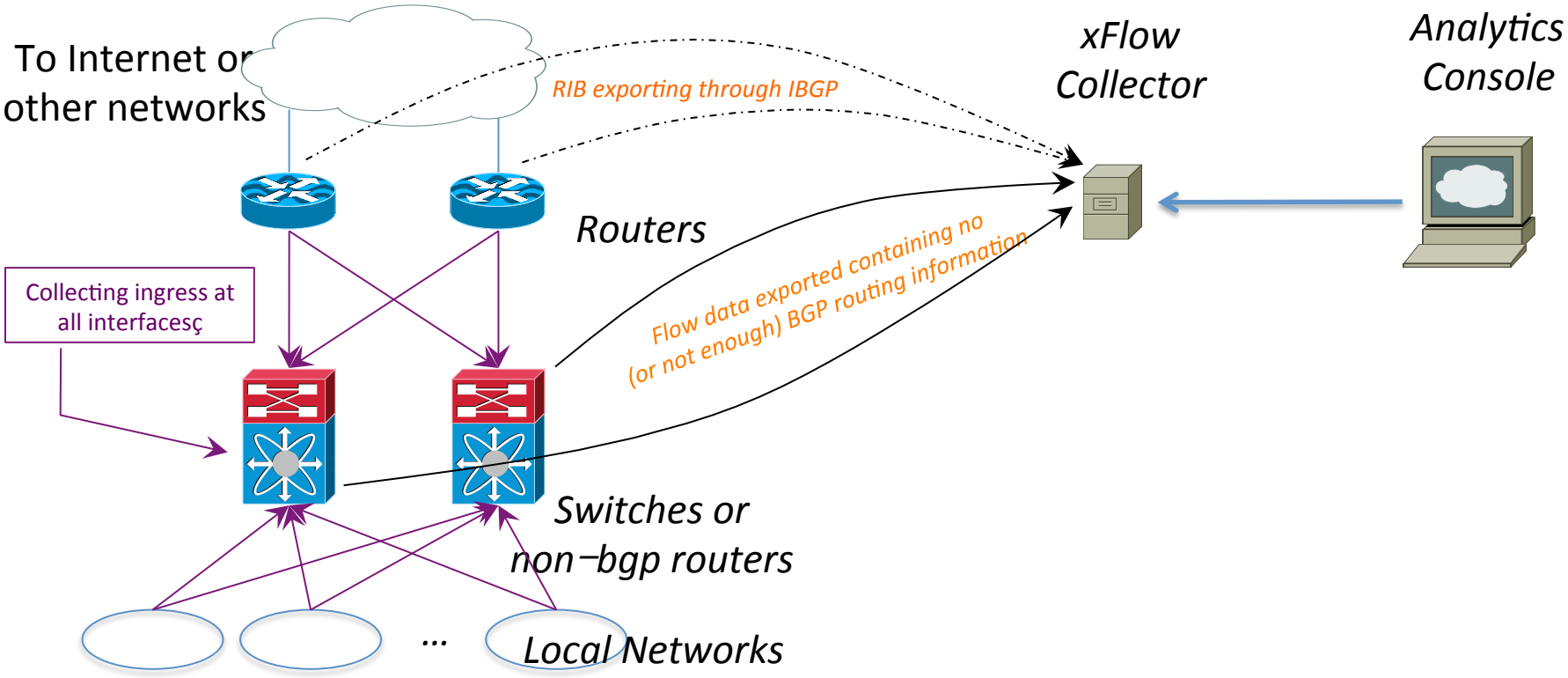
Model 構成モデル

- #1 Generic model of follow collection: router probes, router speaks
  一般的なフローコレクションモデル：ルータがフロー対応



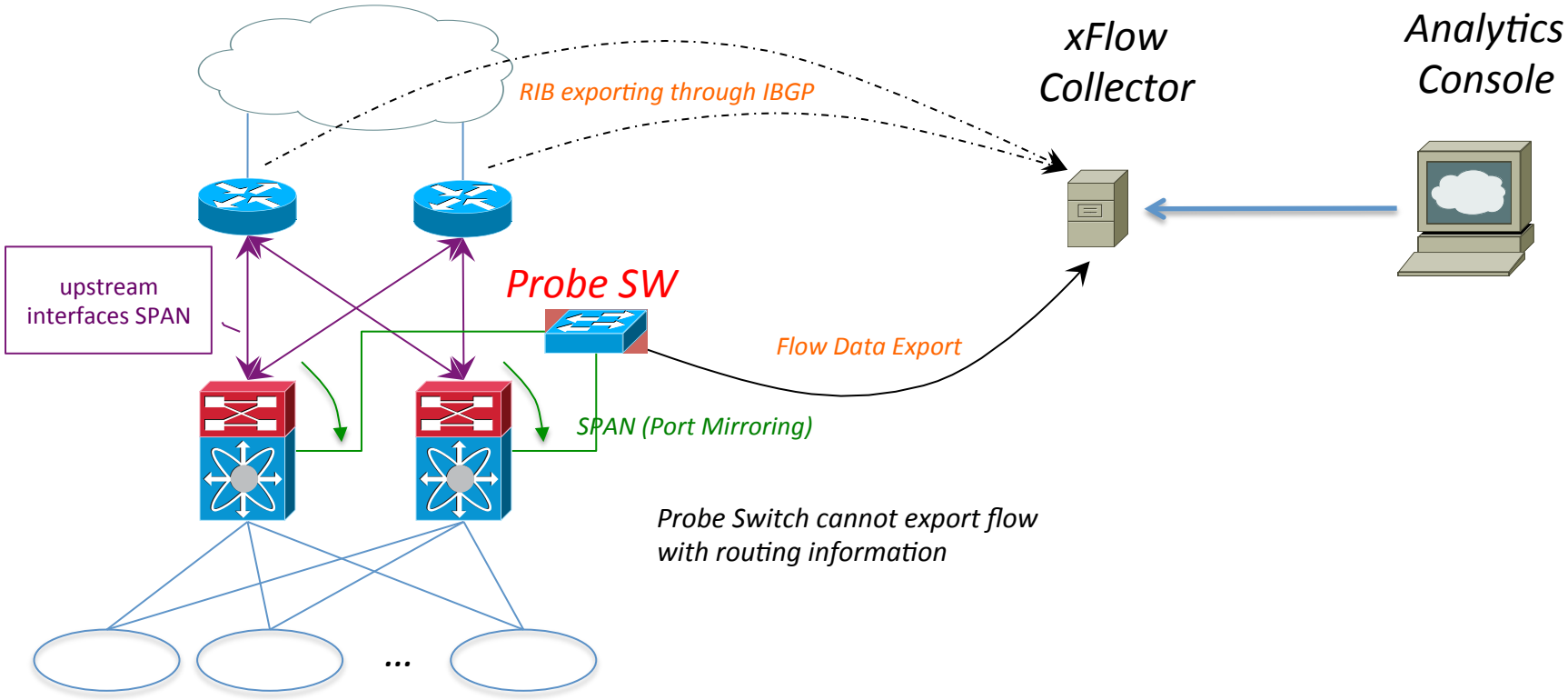
Model 構成モデル

- #2 Probe cannot speak BGP (or not fully speak) but it is needed
  フローデータにはBGP情報がないか少ないが、欲しい



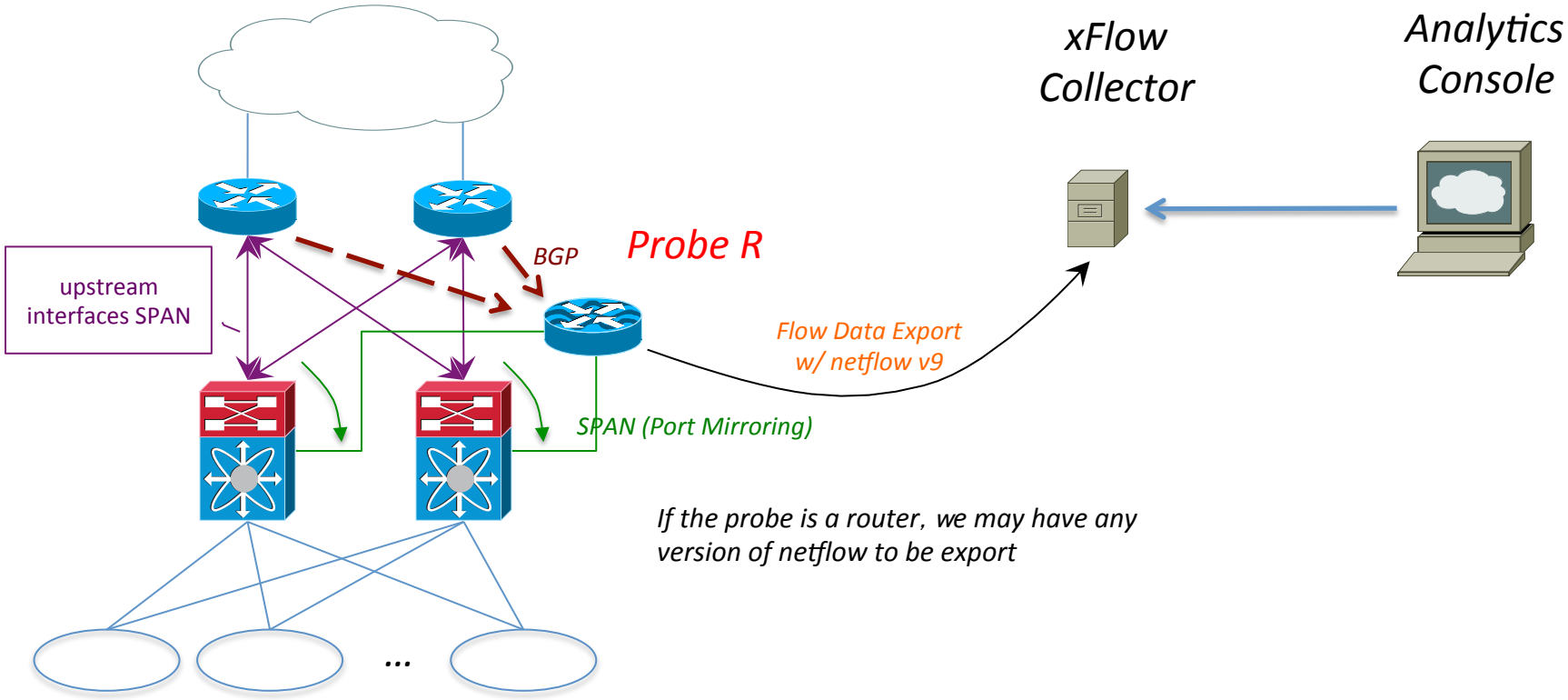
Model　構成モデル

- #3 Router or SW cannot report (wanted version's) xFlow
  RouterやSWはxFlowもしくは欲しいxFlowができない



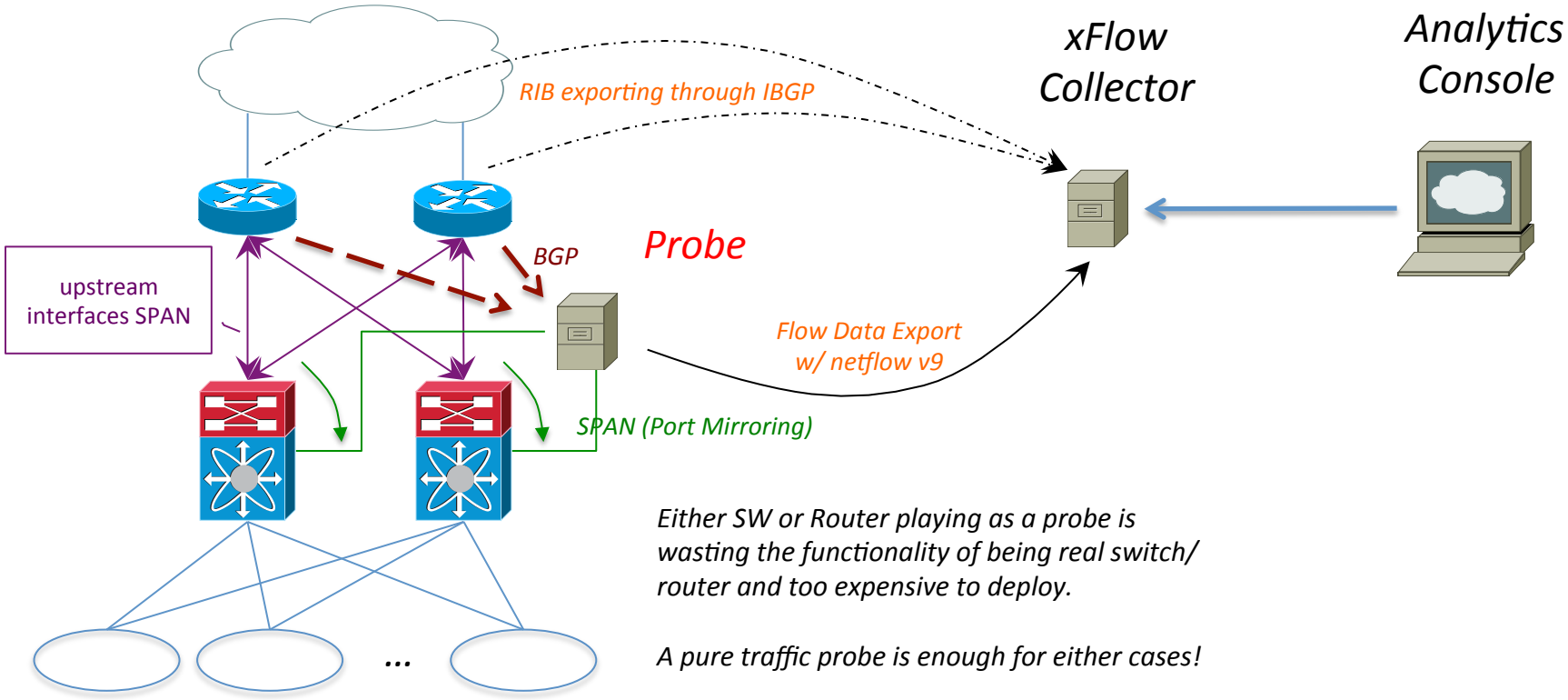
Model　構成モデル

- #3 Router or SW cannot report (wanted version's) xFlow (cont.)
  RouterやSWはxFlowもしくは欲しいxFlowができない



Model　構成モデル

- #3 Router or SW cannot report (wanted version's) xFlow (cont.)
  RouterやSWはxFlowもしくは欲しいxFlowができない



Either SW or Router playing as a probe is wasting the functionality of being real switch/ router and too expensive to deploy.

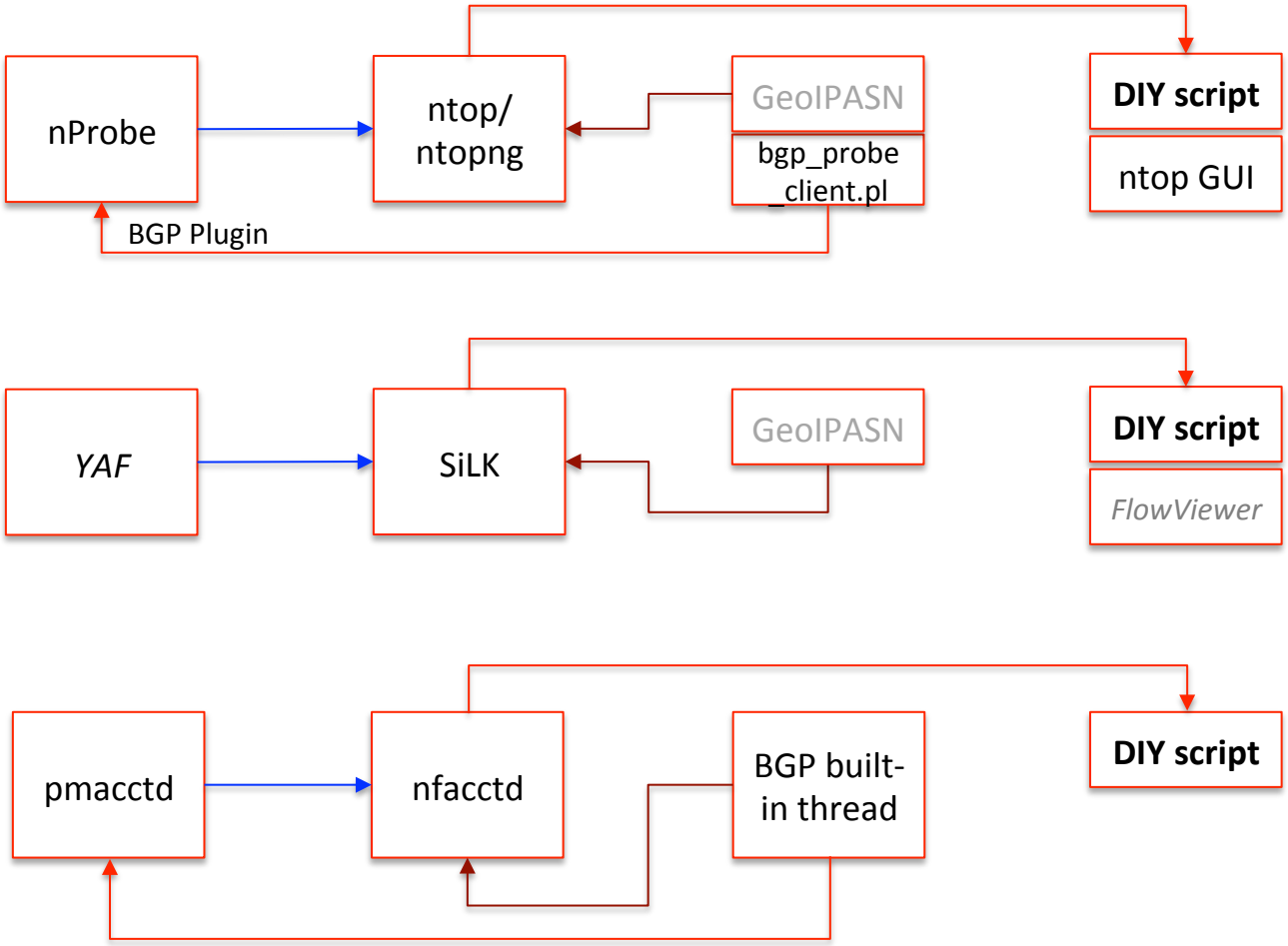
A pure traffic probe is enough for either cases!

Some Flow Tools

- nProbe
  - nProbe Pro (with BGP plug-in included)
    - €299.95 for product and one-year remote support
  - ntopng
    - free download for UNIX-platform
- YAF/SiLK
  - free, open source
  - making graphics with FlowViewer is complicated
    - self-scripting needed

  ※ no BGP plug-in, no sampling (らしい)
- pmacct
  - free, open source
  - fully functional as exporter/collector (w/ BGP), while graphics needs scripting
    - good community support

How we use them? どう使う



nProbe

- nProbe as flow collector
  - Flow dump format
    - binary, text, etc…

```
$ sudo nprobe -n none -D t -i docker0 -V 9 -P /home/maoke/dumped -T
"%IPV6_SRC_ADDR %IPV6_DST_ADDR %IPV6_NEXT_HOP %INPUT_SNMP %OUTPUT_SNMP %IN_PKTS
%IN_BYTES %FIRST_SWITCHED %LAST_SWITCHED %L4_SRC_PORT %L4_DST_PORT %TCP_FLAGS
%PROTOCOL %SRC_TOS %SRC_AS %DST_AS"
```

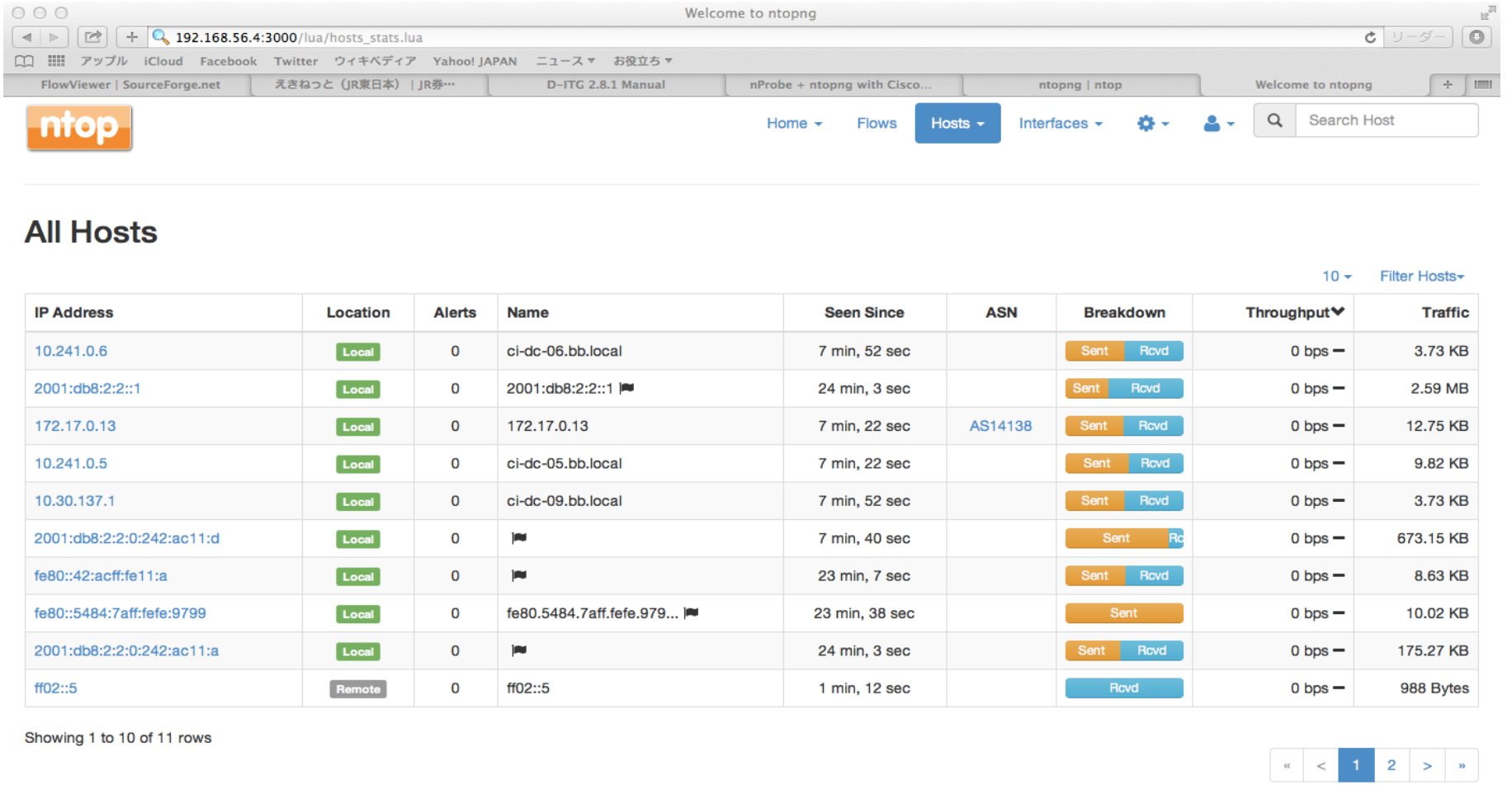
- nProbe exports to ntopng
  - Start ntopng as collector (w/o pcap locally)

```
$ sudo ntopng –i “tcp://127.0.0.1:5556” –F
```

  - Start nProbe as exporter
    - nProbe-ntopng communication is based on ZMQ JSON but –V 9 should be applied for IPv6 support

```
$ sudo nprobe –n none --zmq “tcp://*:5556” –i docker0 eth1 –V 9
```

ntopng



IP Address	Location	Alerts	Name	Seen Since	ASN	Breakdown	Throughput	Traffic
10.241.0.6	Local	0	ci-dc-06.bb.local	7 min, 52 sec		Sent Rcvd	0 bps	3.73 KB
2001:db8:2:2::1	Local	0	2001:db8:2:2::1	24 min, 3 sec		Sent Rcvd	0 bps	2.59 MB
172.17.0.13	Local	0	172.17.0.13	7 min, 22 sec	AS14138	Sent Rcvd	0 bps	12.75 KB
10.241.0.5	Local	0	ci-dc-05.bb.local	7 min, 22 sec		Sent Rcvd	0 bps	9.82 KB
10.30.137.1	Local	0	ci-dc-09.bb.local	7 min, 52 sec		Sent Rcvd	0 bps	3.73 KB
2001:db8:2:2:0:242:ac11:d	Local	0		7 min, 40 sec		Sent Rc	0 bps	673.15 KB
fe80::42:acff:fe11:a	Local	0		23 min, 7 sec		Sent Rcvd	0 bps	8.63 KB
fe80::5484:7aff:fefe:9799	Local	0	fe80.5484.7aff.fefe.979...	23 min, 38 sec		Sent	0 bps	10.02 KB
2001:db8:2:2:0:242:ac11:a	Local	0		24 min, 3 sec		Sent Rcvd	0 bps	175.27 KB
ff02::5	Remote	0	ff02::5	1 min, 12 sec		Rcvd	0 bps	988 Bytes

« < 1 2 > »

pmacct

- Probe & collector
  - pmacctd: using libpcap or PF_RING for traffic collection
    - works independently or exports to collectors including nfacctd
  - nfacctd: receiving netflow v5, v9, IPFIX
  - sfacctd: receiving sFlow
- BGP thread
  - Daemons are working in multi-thread mode, where IBGP session is established for getting full routing table from a peer
    - BGP peer should be specified and configured!
- Data output
  - Pipe (able to be processed with RRDtool)
  - Flat file like .csv
  - SQL database

pmacctd as nf probe

```
!
daemonize: true
imt_path[inbound]:     /tmp/collect.pipe-eth0-in
imt_path[outbound]:    /tmp/collect.pipe-eth0-
out
imt_path[debug]: /tmp/collect.pipe-debug
pidfile:    /var/run/pmacctd.pid
logfile:    /var/log/pmacctd.log
interface: eth0
!
pmacctd_net:      bgp
bgp_peer_src_as_type:  bgp
bgp_src_as_path_type:  bgp
aggregate[inbound]:    src_host, dst_host,
src_as, peer_src_as, peer_src_ip, src_as_path
aggregate[outbound]:   src_host, dst_host,
dst_as, peer_dst_as, peer_dst_ip, as_path
aggregate_filter[inbound]:   dst net
192.0.128.0/24
aggregate_filter[outbound]:  src net
192.0.128.0/24
aggregate[collect]:    src_host, dst_host,
src_port, dst_port, proto, tos
aggregate[debug]:src_host, dst_host, src_port,
dst_port, proto, tos
!
```

```
!
plugins:    memory[inbound], memory[outbound],
memory[debug], nfprobe[collect]
!
nfprobe_receiver:172.17.0.2:2100
nfprobe_source_ip:     172.17.0.2
nfprobe_version: 9
!
pmacctd_as: bgp
bgp_daemon: true
bgp_daemon_ip:    192.0.128.2
bgp_daemon_id:    192.0.128.2
bgp_agent_map:    /home/maoke/pmacct_work/maps/
agent_to_peer.map-v4-eth0
!
plugin_pipe_size:2000000
plugin_buffer_size:    10000
imt_mem_pools_number:  0
!
bgp_table_dump_file:   /tmp/bgp-$peer_src_ip.txt
bgp_table_dump_refresh_time: 300
!
! for debug only
nfprobe_timeouts[collect]:
     maxlife=60:expint=60
```

pmacct nfprobeのoutput例

```
~/pmacct_work$ pmacct -s -p /tmp/collect.pipe-eth0-in
SRC_AS      SRC_AS_PATH                PEER_SRC_AS PEER_SRC_IP
        SRC_IP                                             DST_IP
                    PACKETS                BYTES
0           ^$                         0           0
        192.0.128.1                                        192.0.128.65
                    375                    25008
0           ^$                         0           0
        192.0.128.2                                        192.0.128.1
                    523                    85010
0           ^$                         0           0
        192.0.128.1                                        192.0.128.2
                    815                    52575
0           ^$                         0           0
        192.168.56.2                                       192.0.128.65
                    21                     1404
0           ^$                         0           0
        192.0.128.65                                       192.0.128.1
                    245                    25532
0           ^$                         0           0
        192.32.0.2                                         192.0.128.65
                    37                     4983
65533       65530_65533                65530       0
        192.32.0.2                                         192.0.128.65
                    214                    23992

For a total of: 7 entries
```

pmacct nfprobeのoutput例

```
~/pmacct_work$ pmacct -s -p /tmp/collect.pipe-eth0-out
DST_AS      AS_PATH                  PEER_DST_AS PEER_DST_IP                                SRC_IP
      DST_IP                                            PACKETS          BYTES
0           ^$                       0           0                                          192.0.128.65
      192.32.0.2                                        30               4841
0           ^$                       0           0                                          192.0.128.2
      224.0.0.5                                         690              47000
0           ^$                       0           0                                          192.0.128.1
      192.0.128.65                                      72               5044
0           ^$                       0           0                                          192.0.128.2
      192.0.128.1                                       8                526
0           ^$                       0           0                                          192.0.128.1
      192.0.128.2                                       10               823
0           ^$                       0           192.0.128.1                                192.0.128.1
      192.0.128.2                                       842              54216
0           ^$                       0           192.0.128.1                                192.0.128.65
      192.0.128.1                                       202              21268
0           ^$                       0           0                                          192.0.128.65
      192.0.128.1                                       43               4264
0           ^$                       0           0                                          192.0.128.2
      192.168.56.2                                      9                828
0           ^$                       0           192.0.128.1                                192.0.128.2
      192.0.128.1                                       536              88264
65533       65530_65533              65530       192.168.56.2                               192.0.128.65
      192.32.0.2                                        323              22124
0           ^$                       0           192.0.128.1                                192.0.128.1
      192.0.128.65                                      303              19964
0           ^$                       0           0                                          192.0.128.1
      224.0.0.5                                         690              47064
0           ^$                       0           0                                          192.0.128.65
      192.168.56.2                                      12               828

For a total of: 14 entries
```

bgp_nexthop

nfacctd as collector

```
!
daemonize:     true
logfile:  /var/log/nfacctd.log
nfacctd_ip:    172.17.0.2
nfacctd_port:  2100
plugins:  memory[display]
!
nfacctd_net:    bgp
bgp_peer_src_as_type:   bgp
bgp_src_as_path_type:   bgp
!
aggregate[display]:      src_host,
dst_host, src_as, dst_as, peer_src_as,
peer_dst_as, as_path, src_as_path,
peer_src_ip, peer_dst_ip
!
!classifiers:  /home/maoke/pmacct_work/
maps/pretag.map-eth0
!
```

```
!
nfacctd_as_new:     bgp
bgp_daemon:    true
bgp_daemon_ip: 172.17.0.2
bgp_daemon_id: 172.17.0.2
bgp_agent_map: /home/maoke/pmacct_work/
maps/agent_to_peer.map-v4-eth0
!
!plugin_pipe_size:  2000000
!plugin_buffer_size:     10000
!imt_mem_pools_number:   0
!
bgp_table_dump_file:     /tmp/bgp-nfacctd-
$peer_src_ip.txt
bgp_table_dump_refresh_time:  300
!
```

nfacctdのoutput例

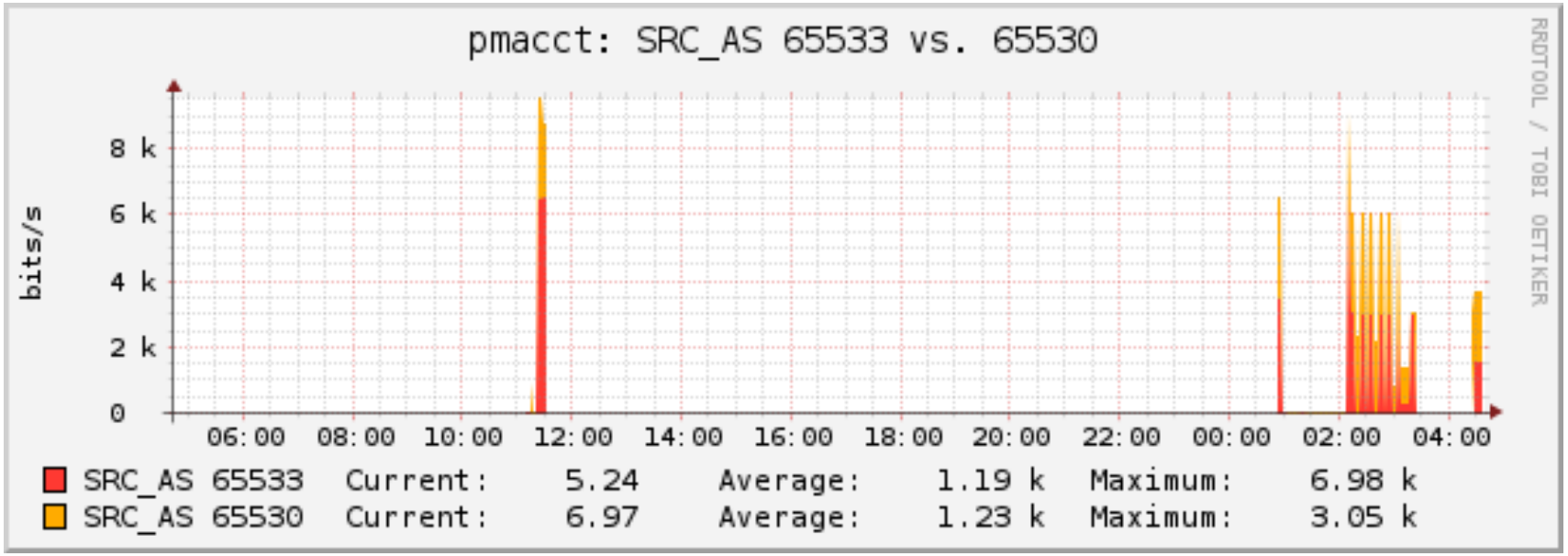
```
~/pmacct_work$ pmacct -s -p /tmp/collect.pipe
SRC_AS      DST_AS      AS_PATH                 SRC_AS_PATH             PEER_SRC_AS PEER_DST_AS PEER_SRC_IP                 PEER_DST_IP
SRC_IP                                          DST_IP                                          PACKETS     BYTES
0           65533       65530_65533               ^$                    0           65530       172.17.0.2                  192.168.56.2
172.17.0.2                                      192.32.0.2                                      12          1008
0           0           ^$                        ^$                    0           0           172.17.0.2                  0
fe80::5484:7aff:fefe:9799                       ff02::5                                         209         15884
0           0           ^$                        ^$                    0           0           172.17.0.2                  172.17.42.1
192.0.128.65                                    192.0.128.1                                     34          3536
0           0           ^$                        ^$                    0           0           172.17.0.2                  172.17.42.1
192.0.128.2                                     192.0.128.1                                     330         59865
0           65530       65530                     ^$                    0           65530       172.17.0.2                  192.168.56.2
192.0.128.65                                    192.16.0.2                                      2980        158870
0           0           ^$                        ^$                    0           0           172.17.0.2                  0
192.0.128.2                                     192.0.128.1                                     225         22520
0           0           ^$                        ^$                    0           0           172.17.0.2                  0
192.0.128.65                                    192.32.0.2                                      5           420
0           0           ^$                        ^$                    0           0           172.17.0.2                  0
192.0.128.1                                     192.0.128.2                                     424         29740
0           0           ^$                        ^$                    0           0           172.17.0.2                  0
192.0.128.2                                     192.168.56.3                                    2004        2231292
65530       0           ^$                        65530                 65530       0           172.17.0.2                  172.17.42.1
192.16.0.2                                      192.0.128.65                                    11          924
65533       0           ^$                        65530_65533           65530       0           172.17.0.2                  0
192.32.0.2                                      172.17.0.2                                      4           336
0           0           ^$                        ^$                    0           0           172.17.0.2                  0
192.0.128.1                                     224.0.0.5                                       197         13432
0           0           ^$                        ^$                    0           0           172.17.0.2                  0
172.17.42.1                                     172.17.0.2                                      64          4134
0           65533       65530_65533               ^$                    0           65530       172.17.0.2                  192.168.56.2
192.0.128.65                                    192.32.0.2                                      149         11622
0           0           ^$                        ^$                    0           0           172.17.0.2                  0
192.0.128.2                                     224.0.0.5                                       196         13348
0           0           ^$                        ^$                    0           0           172.17.0.2                  0
192.32.0.2                                      192.0.128.65                                    6           504
0           0           ^$                        ^$                    0           0           172.17.0.2                  172.17.42.1
192.0.128.1                                     192.0.128.65                                    34          1768
65533       0           ^$                        65530_65533           65530       0           172.17.0.2                  172.17.42.1
192.32.0.2                                      192.0.128.65                                    964         114610
0           0           ^$                        ^$                    0           0           172.17.0.2                  0
192.0.128.65                                    192.16.0.2                                      6           504
0           0           ^$                        ^$                    0           0           172.17.0.2                  172.17.42.1
192.168.56.3                                    192.0.128.2                                     1494        314102
0           0           ^$                        ^$                    0           0           172.17.0.2                  172.17.42.1
192.0.128.1                                     192.0.128.2                                     494         32077
0           0           ^$                        ^$                    0           0           172.17.0.2                  0
192.168.56.3                                    192.0.128.2                                     8           648
0           0           ^$                        ^$                    0           0           172.17.0.2                  0
192.16.0.2                                      192.0.128.65                                    6           504
0           0           ^$                        ^$                    0           0           172.17.0.2                  0
172.17.0.2                                      172.17.42.1                                     35          2462

For a total of: 24 entries
```

Cactiでグラフ

- pmacctでpipeから呼び出す
- RRDtoolで処理
- MySQLでリソース定義を記録され

- PNGグラフィック
- csv出力可能



デスクトップ検証の環境