23rd Chaos Communication Congress
23C3
Who can you trust?

sFlow
I can feel your traffic

23rd Chaos Communication Congress
Elisa Jasinska
06/12/30

Agenda

- What is sFlow?
- What is AMS-IX?
- Existing Software
- Performance Issues
- AMS-IX Software
- Privacy
- Results

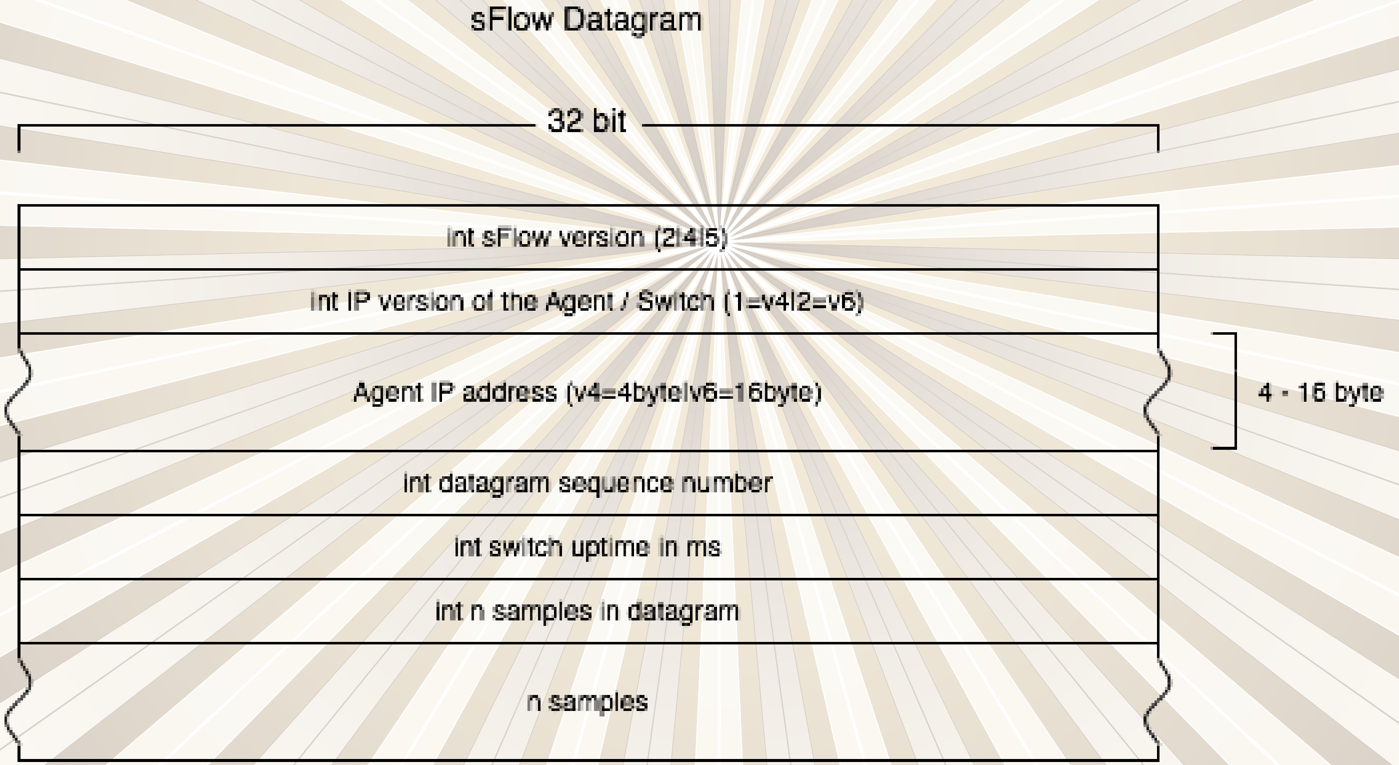
What is sFlow?

- Sampling mechanism
  (not “touching” every packet)
- Monitoring switched or routed networks
- Cisco IOS - NetFlow
- Applicable to high speed networks
  (>= 1GE)

What is sFlow?

- sFlow datagrams sent via UDP
- Datagram format standard defined in RFC 3176
- Implemented on a wide range of devices (Foundry, Force10, Extreme...)

What is sFlow?



sFlow Datagram

32 bit

int sFlow version (2|4|5)

int IP version of the Agent / Switch (1=v4|2=v6)

Agent IP address (v4=4byte|v6=16byte)

4 - 16 byte

int datagram sequence number

int switch uptime in ms

int n samples in datagram

n samples

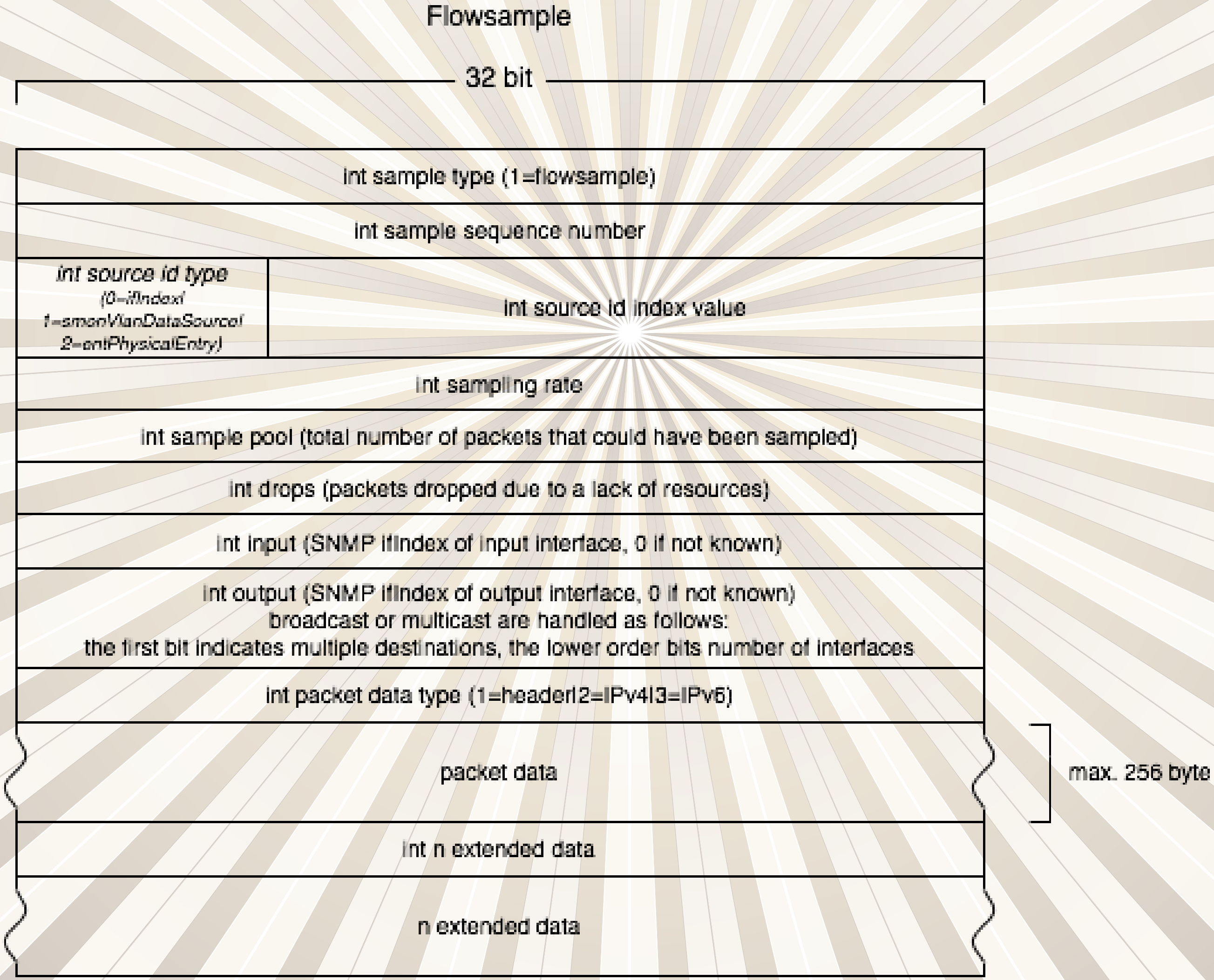
What is sFlow?

- Not everything is sampled information
- Two different types provided by the datagram format:
  - Flow samples
  - Counter samples

What is sFlow?

- Flow samples
  - Defined sampling rate (e.g. one out of 8192)
  - Up to 256 bytes of captured packet (L2-L7)

What is sFlow?



Flowsample

32 bit

int sample type (1=flowsample)

int sample sequence number

int source id type
(0=ifIndex|
1=smonVlanDataSource|
2=entPhysicalEntry)

int source id index value

int sampling rate

int sample pool (total number of packets that could have been sampled)

int drops (packets dropped due to a lack of resources)

int input (SNMP ifIndex of input interface, 0 if not known)

int output (SNMP ifIndex of output interface, 0 if not known)
broadcast or multicast are handled as follows:
the first bit indicates multiple destinations, the lower order bits number of interfaces

int packet data type (1=header|2=IPv4|3=IPv6)

packet data

max. 256 byte

int n extended data

n extended data

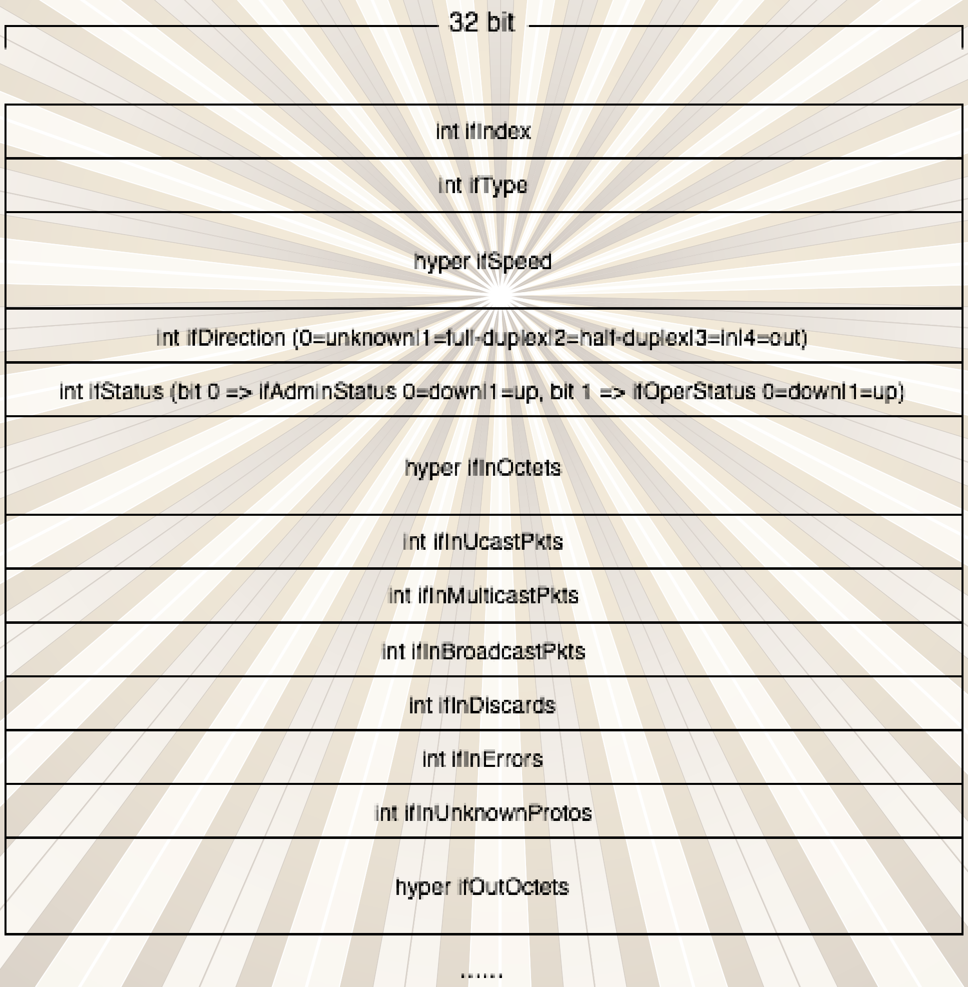
What is sFlow?

- Counter samples
  - Polling interval (e.g. 30 seconds)
  - Interface counters (octets/packets/errors)

What is sFlow?

Counterstype - Generic

see RFC2233



32 bit

int ifIndex

int ifType

hyper ifSpeed

int ifDirection (0=unknown|1=full-duplex|2=half-duplex|3=in|4=out)

int ifStatus (bit 0 => ifAdminStatus 0=down|1=up, bit 1 => ifOperStatus 0=down|1=up)

hyper ifInOctets

int ifInUcastPkts

int ifInMulticastPkts

int ifInBroadcastPkts

int ifInDiscards

int ifInErrors

int ifInUnknownProtos

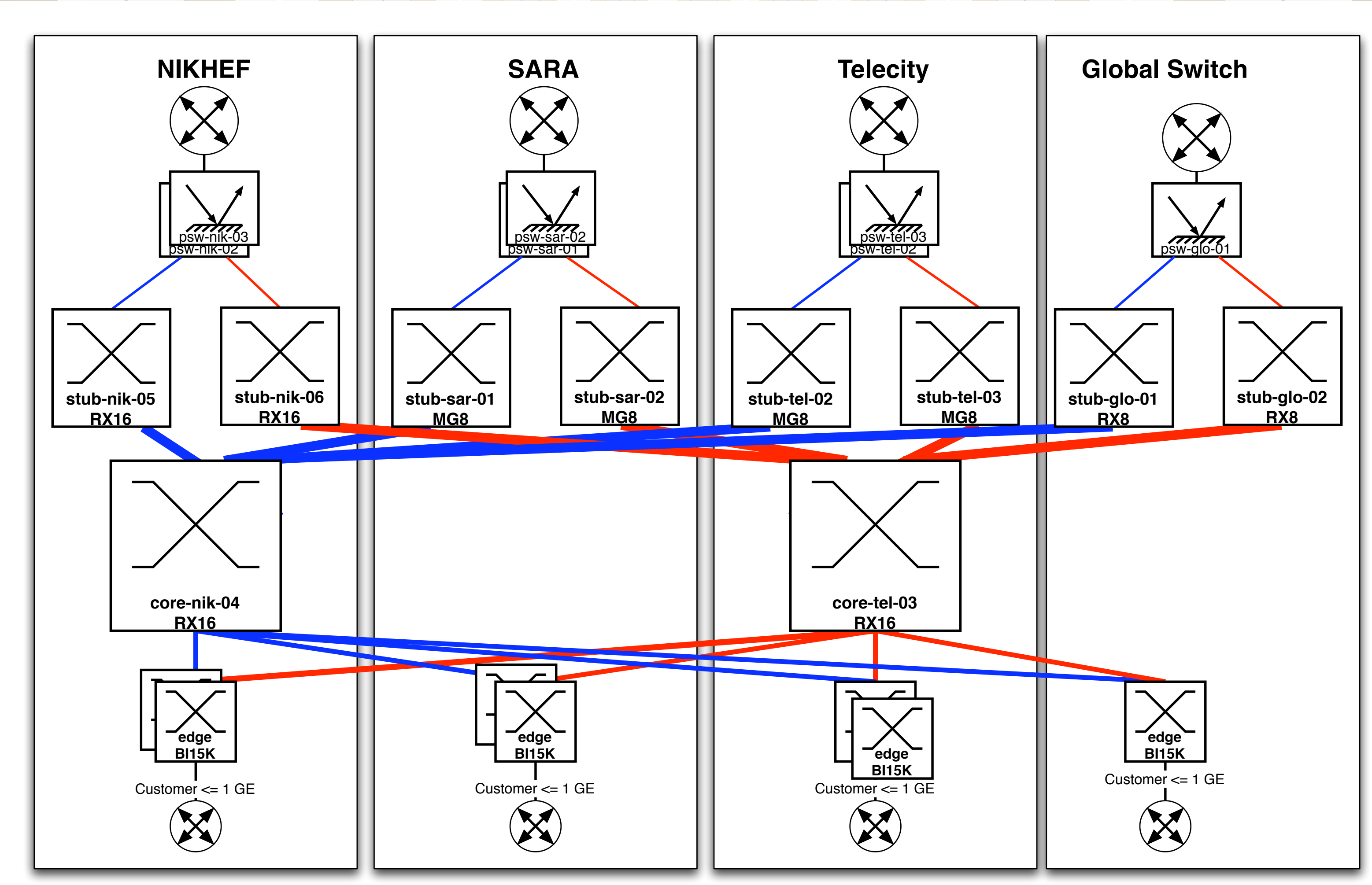
hyper ifOutOctets

......

What is AMS-IX?

- Non-profit Internet Exchange
- Based in Amsterdam
- 4 independent colocation facilities
- Operates only on Layer 2
- Interconnects parties to exchange IP traffic (e.g. ISP's, web hosters, content providers)

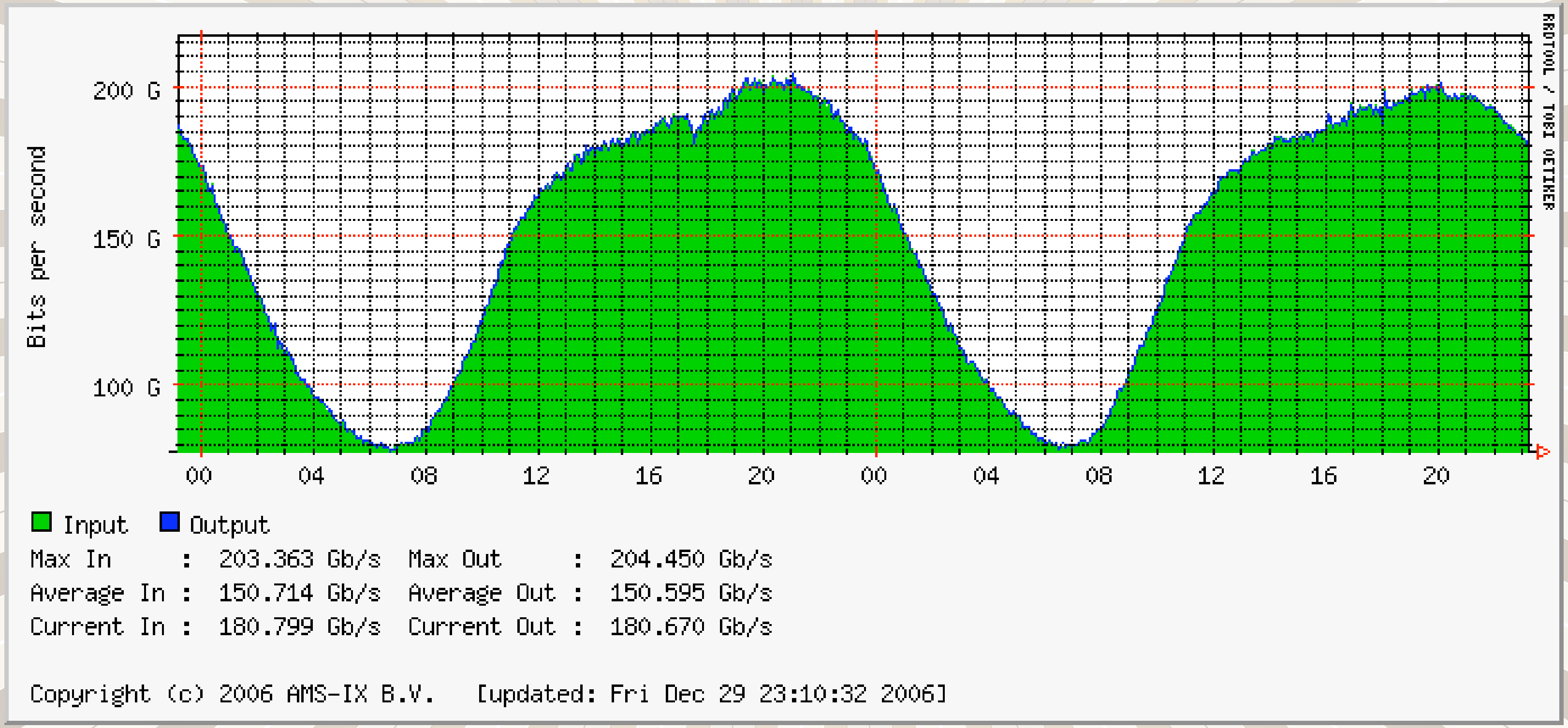
What is AMS-IX?



What is AMS-IX?

- Statistics
- Total traffic statistics
- Interface counter (octets/packets/errors)
- Polled via SNMP
- MRTG
- .... no sFlow so far ...

What is AMS-IX?



```
🟩 Input  🟦 Output
Max In      :  203.363 Gb/s  Max Out      :  204.450 Gb/s
Average In  :  150.714 Gb/s  Average Out  :  150.595 Gb/s
Current In  :  180.799 Gb/s  Current Out  :  180.670 Gb/s

Copyright (c) 2006 AMS-IX B.V.   [updated: Fri Dec 29 23:10:32 2006]
```

What is AMS-IX?

Use flow samples to...

- Provide member-to-member traffic information
- See growth of (or lack of) IPv6
- Due to high throughput a very efficient system is required

Existing Software

- Free software:
  - InMon – sflowtool
  - Pmacct
  - sFlow2MySQL
- Commercial:
  - InMon – Traffic Sentinel

Performance Issues

- Issues with existing software
    - Saves each sample to DB
    - No caching or preprocessing possible
    - Graphing with RRDtool
        - overhead due to data export to RRD
        - same data saved twice

Performance Issues

- Traffic up to 220 Gb/s (35 Mpps)
- ca. 3500 samples per second
- Cannot store each sample in a DB

AMS-IX Software

Net::sFlow

- Decodes sFlow datagrams
- Supports sFlow version 2/4 and 5
- Single (exportable) function, decode()
- Available on CPAN

AMS-IX Software

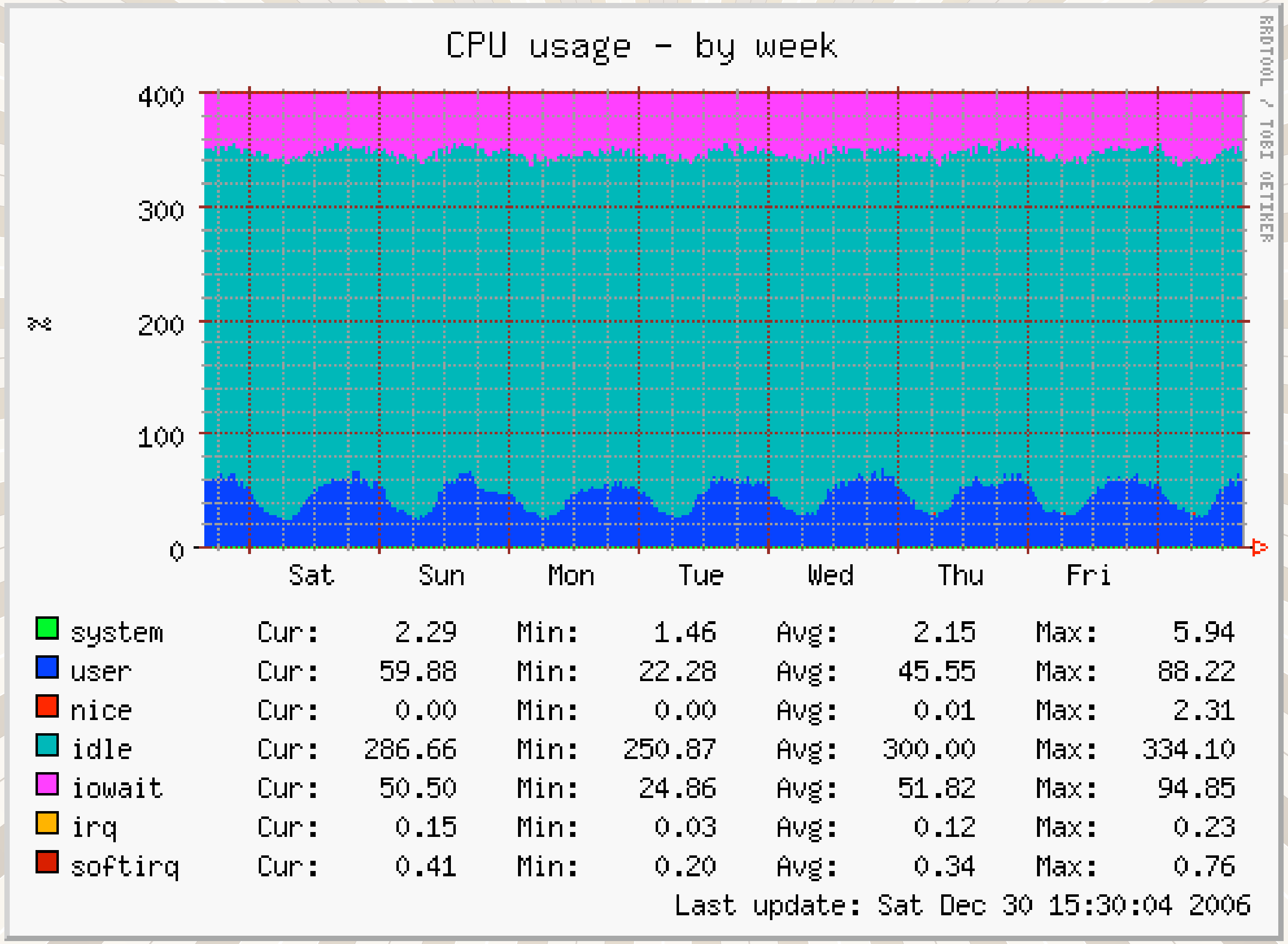
sFlow daemon

- Based on module Net::sFlow
- Receives UDP datagrams
- Analyses the information
- Stores data in RRD files

AMS-IX Software

- Performance reasonable
  - Less I/O usage due to preprocessing
  - PERL unpack() slower than decoder written in C

AMS-IX Software

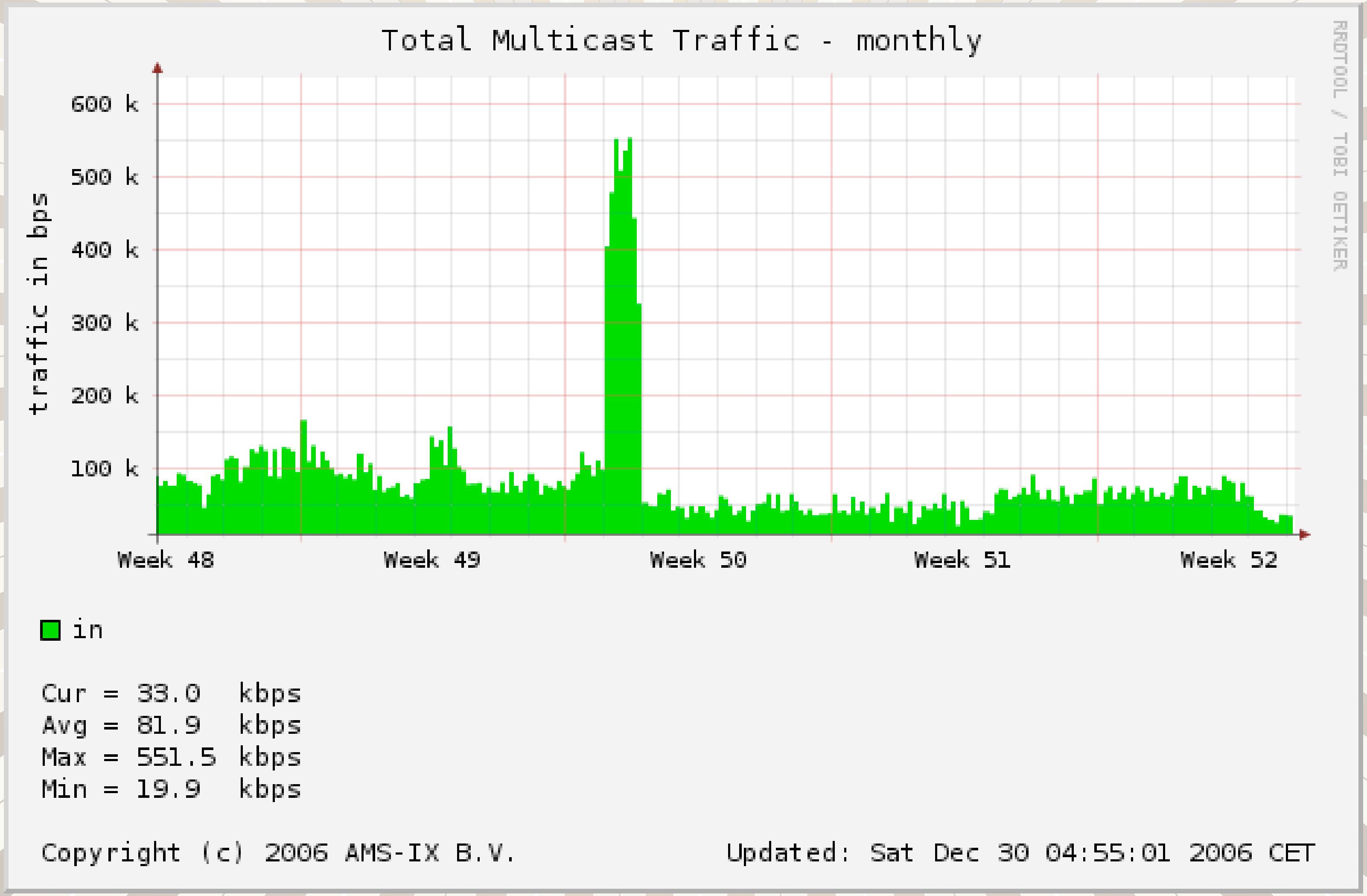


```
🟩 system     Cur:   2.29   Min:   1.46   Avg:   2.15   Max:   5.94
🟦 user       Cur:  59.88   Min:  22.28   Avg:  45.55   Max:  88.22
🟥 nice       Cur:   0.00   Min:   0.00   Avg:   0.01   Max:   2.31
■ idle        Cur: 286.66   Min: 250.87   Avg: 300.00   Max: 334.10
🟪 iowait     Cur:  50.50   Min:  24.86   Avg:  51.82   Max:  94.85
🟨 irq        Cur:   0.15   Min:   0.03   Avg:   0.12   Max:   0.23
🟥 softirq    Cur:   0.41   Min:   0.20   Avg:   0.34   Max:   0.76
                                 Last update: Sat Dec 30 15:30:04 2006
```

Privacy

- Whole packet header (up to 256 Byte)
- Statistical analysis
- Samples not saved after decoding
- Decoding only up to L2 (ethernet)

AMS-IX - Results

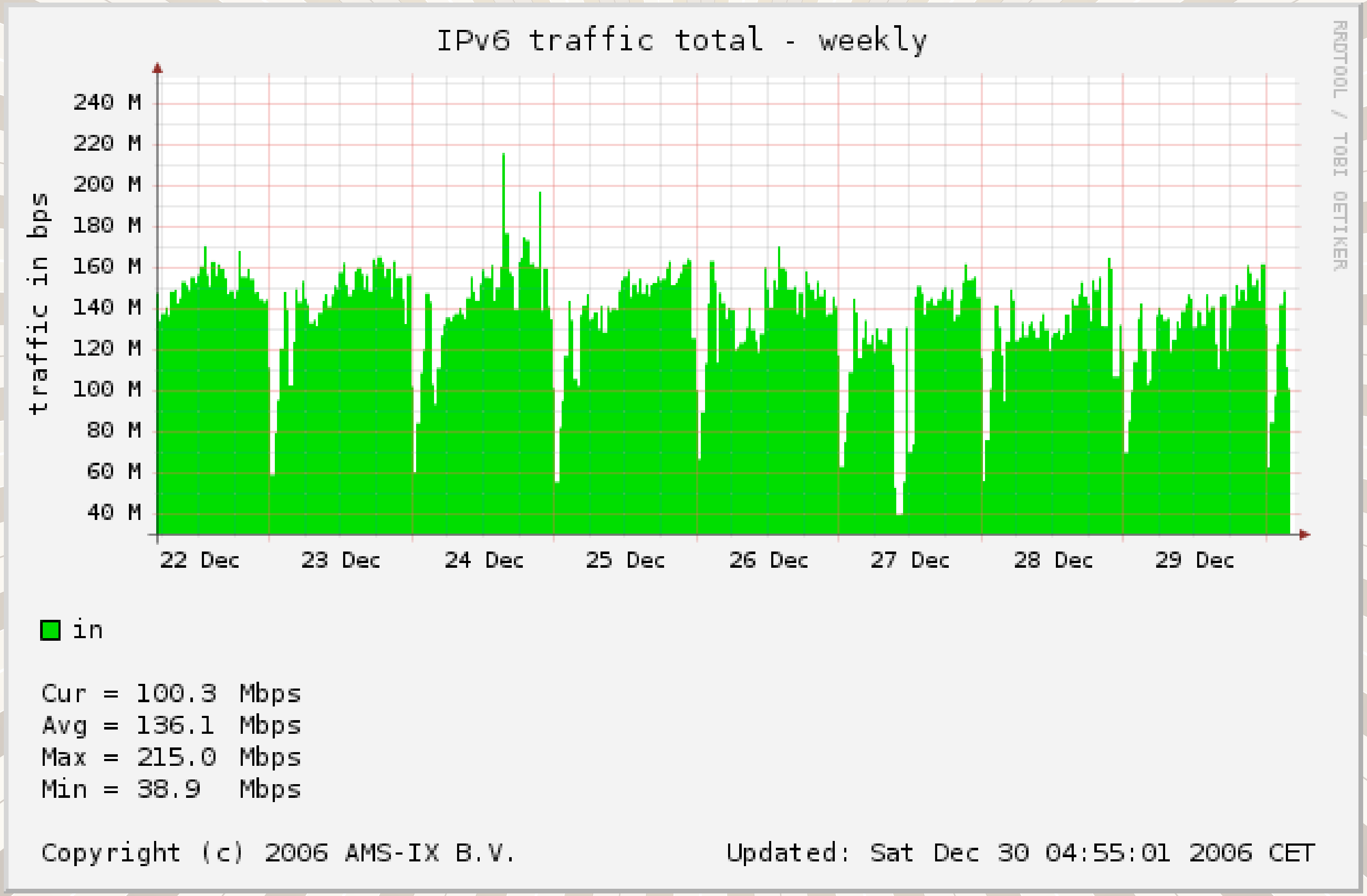


🟩 in

```
Cur = 33.0  kbps
Avg = 81.9  kbps
Max = 551.5 kbps
Min = 19.9  kbps
```

Updated: Sat Dec 30 04:55:01 2006 CET

AMS-IX - Results

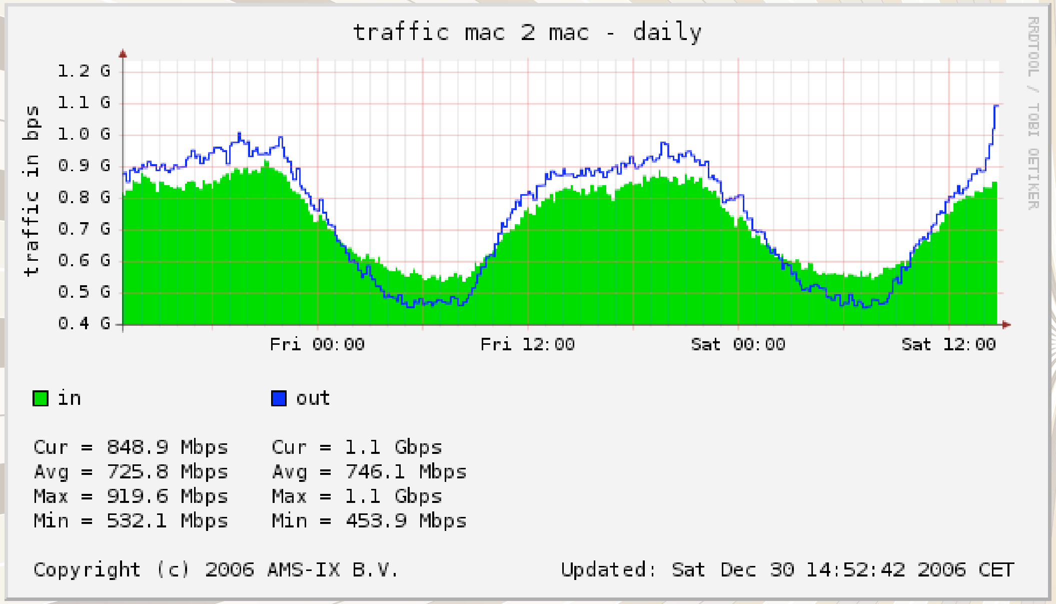


🟩 in

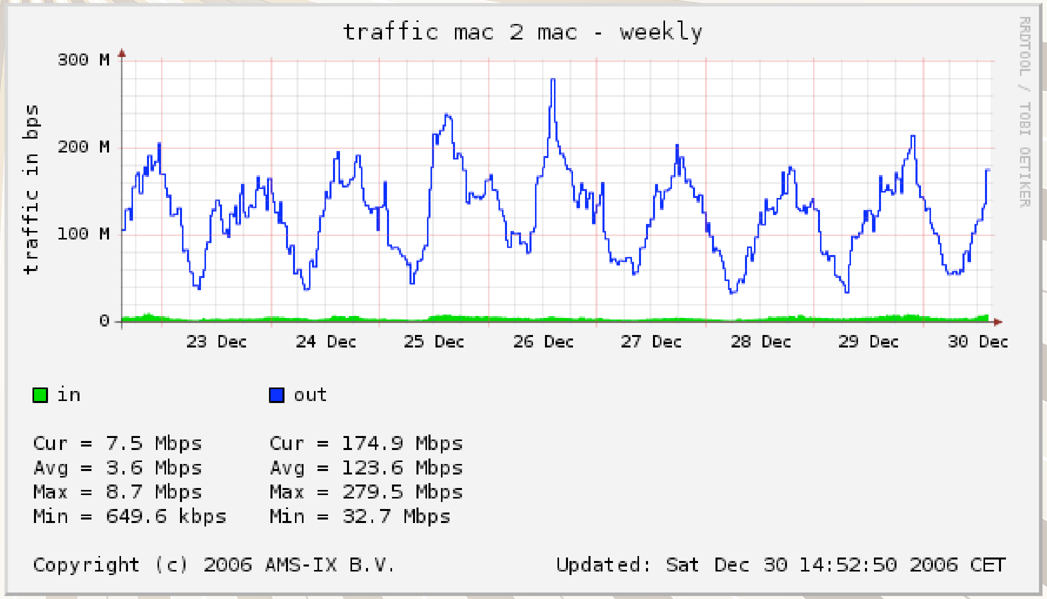
```
Cur = 100.3  Mbps
Avg = 136.1  Mbps
Max = 215.0  Mbps
Min = 38.9   Mbps
```

Copyright (c) 2006 AMS-IX B.V.   Updated: Sat Dec 30 04:55:01 2006 CET

AMS-IX - Results

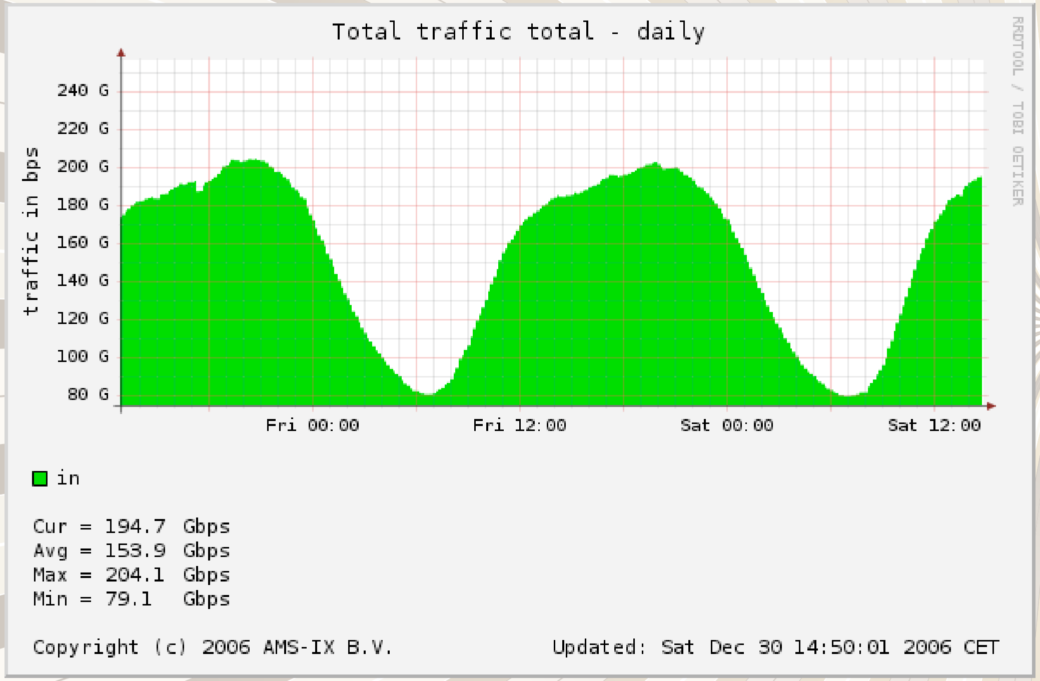


```
🟩 in                 🟦 out

Cur = 848.9 Mbps      Cur = 1.1 Gbps
Avg = 725.8 Mbps      Avg = 746.1 Mbps
Max = 919.6 Mbps      Max = 1.1 Gbps
Min = 532.1 Mbps      Min = 453.9 Mbps

Copyright (c) 2006 AMS-IX B.V.          Updated: Sat Dec 30 14:52:42 2006 CET
```

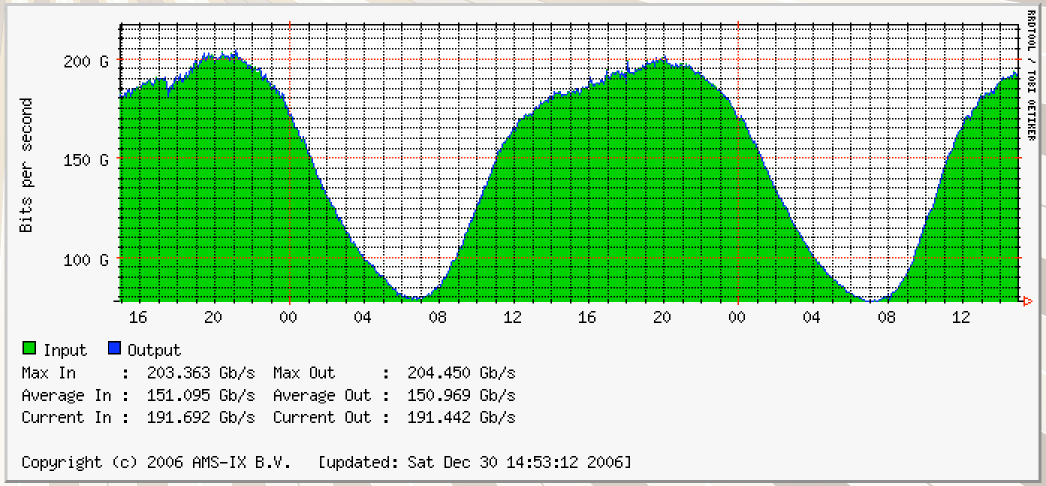


```
🟩 in                 🟦 out

Cur = 7.5 Mbps        Cur = 174.9 Mbps
Avg = 3.6 Mbps        Avg = 123.6 Mbps
Max = 8.7 Mbps        Max = 279.5 Mbps
Min = 649.6 kbps      Min = 32.7 Mbps

Copyright (c) 2006 AMS-IX B.V.          Updated: Sat Dec 30 14:52:50 2006 CET
```

AMS-IX - Results

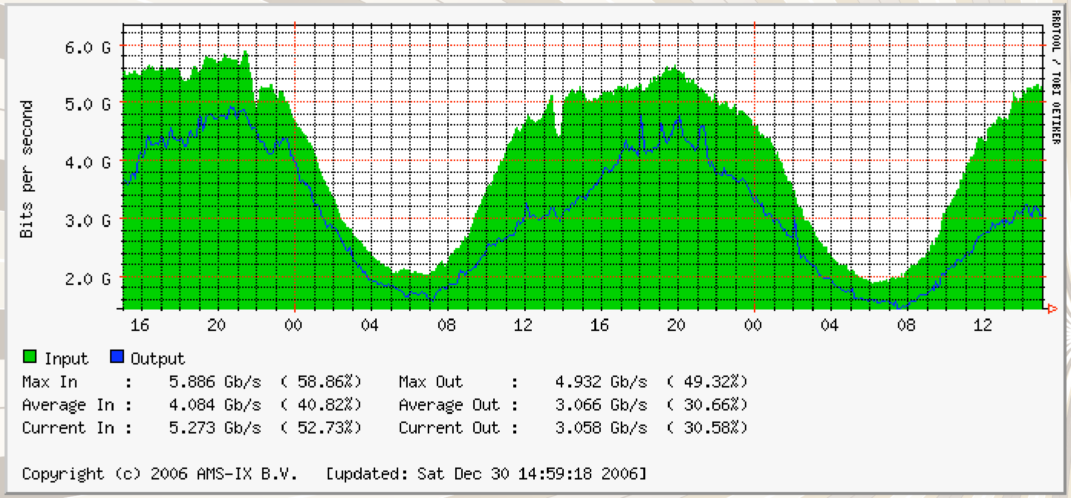


```
Cur = 194.7  Gbps
Avg = 153.9  Gbps
Max = 204.1  Gbps
Min = 79.1   Gbps

Copyright (c) 2006 AMS-IX B.V.              Updated: Sat Dec 30 14:50:01 2006 CET
```

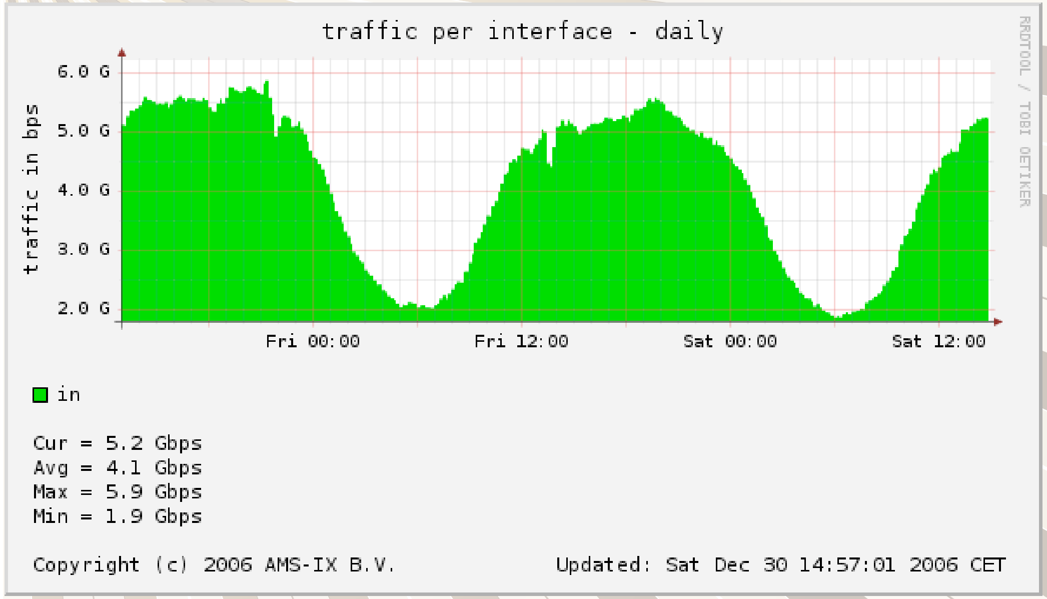


```
Max In     :  203.363 Gb/s  Max Out      :  204.450 Gb/s
Average In :  151.095 Gb/s  Average Out  :  150.969 Gb/s
Current In :  191.692 Gb/s  Current Out  :  191.442 Gb/s

Copyright (c) 2006 AMS-IX B.V.   [updated: Sat Dec 30 14:53:12 2006]
```

AMS-IX - Results

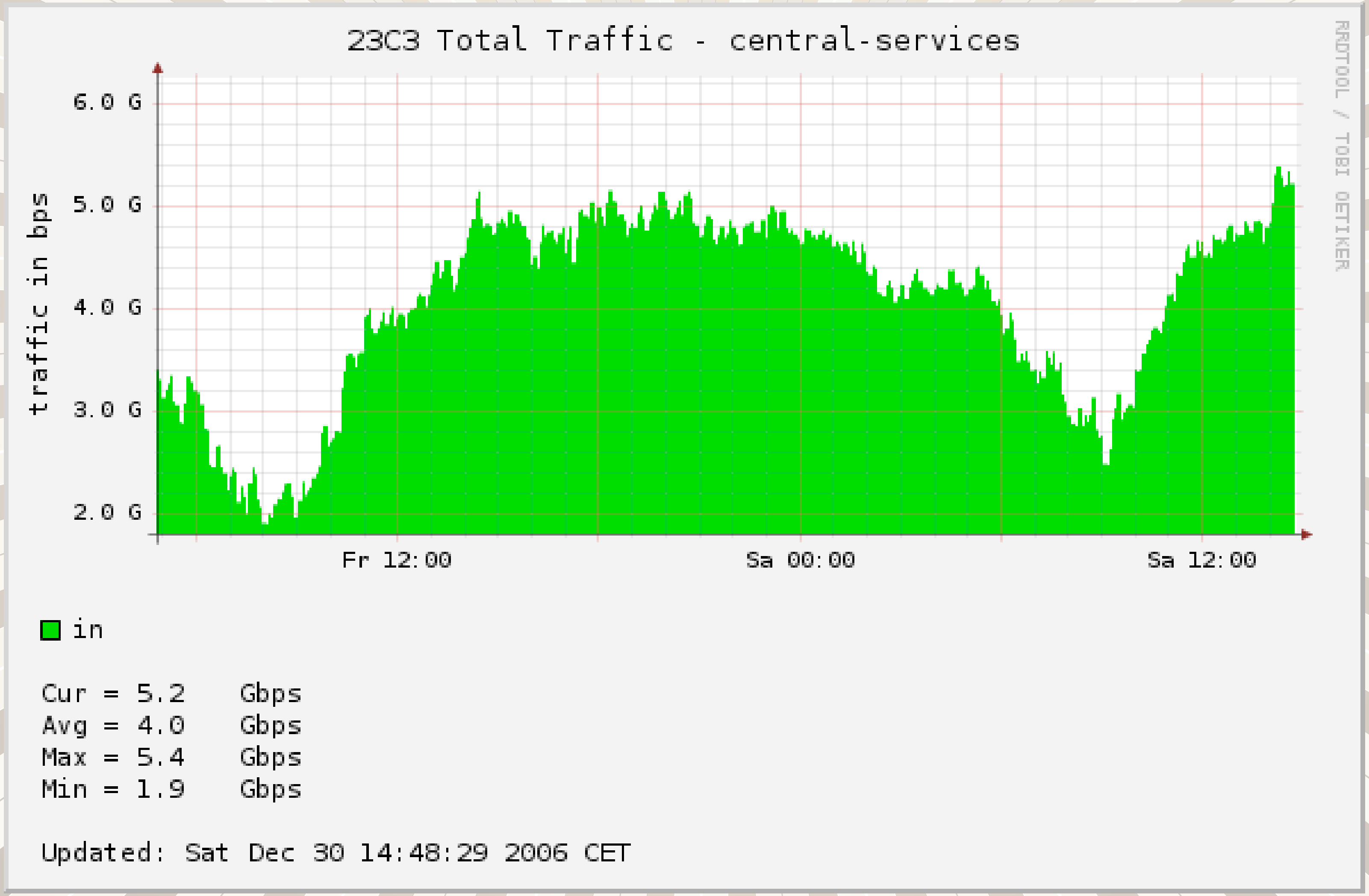


```
🟩 Input  🟦 Output
Max In     :    5.886 Gb/s  ( 58.86%)   Max Out     :    4.932 Gb/s  ( 49.32%)
Average In :    4.084 Gb/s  ( 40.82%)   Average Out :    3.066 Gb/s  ( 30.66%)
Current In :    5.273 Gb/s  ( 52.73%)   Current Out :    3.058 Gb/s  ( 30.58%)

Copyright (c) 2006 AMS-IX B.V.   [Updated: Sat Dec 30 14:59:18 2006]
```



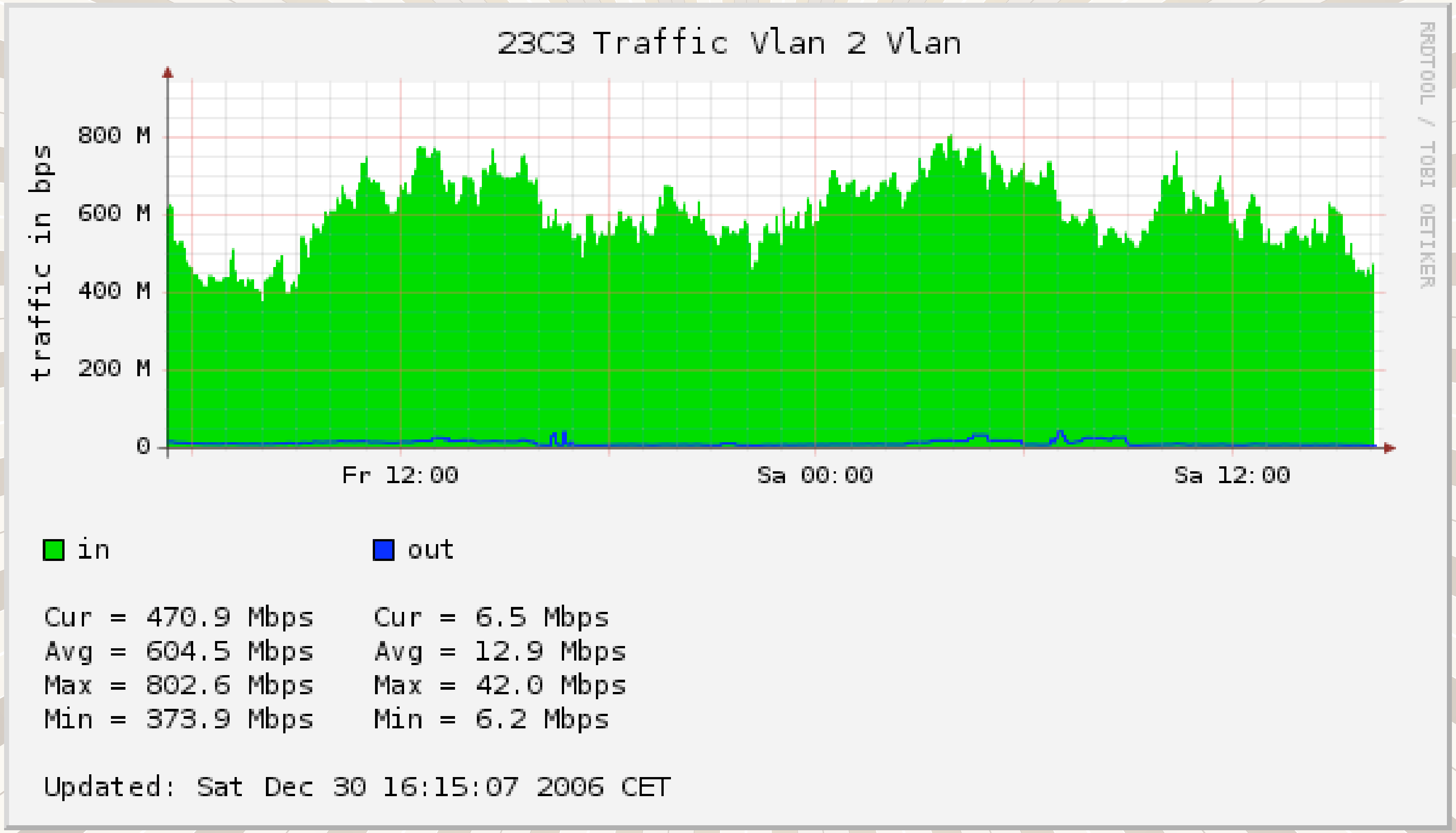
```
🟩 in

Cur = 5.2 Gbps
Avg = 4.1 Gbps
Max = 5.9 Gbps
Min = 1.9 Gbps

Copyright (c) 2006 AMS-IX B.V.          Updated: Sat Dec 30 14:57:01 2006 CET
```

23C3 - Results



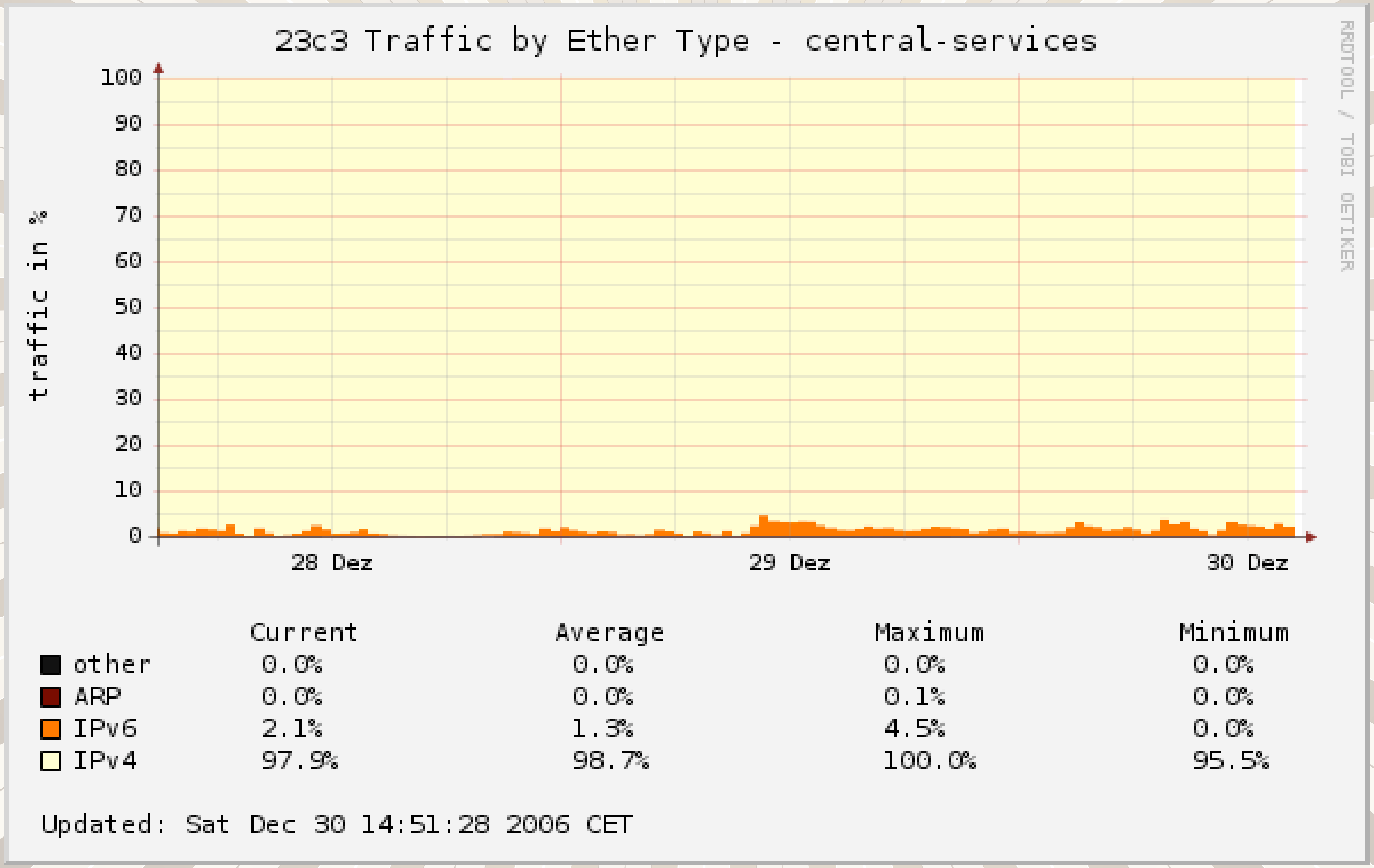
```
Cur = 5.2   Gbps
Avg = 4.0   Gbps
Max = 5.4   Gbps
Min = 1.9   Gbps

Updated: Sat Dec 30 14:48:29 2006 CET
```

23C3 - Results



```
🟩 in                🟦 out

Cur = 470.9 Mbps    Cur = 6.5 Mbps
Avg = 604.5 Mbps    Avg = 12.9 Mbps
Max = 802.6 Mbps    Max = 42.0 Mbps
Min = 373.9 Mbps    Min = 6.2 Mbps

Updated: Sat Dec 30 16:15:07 2006 CET
```

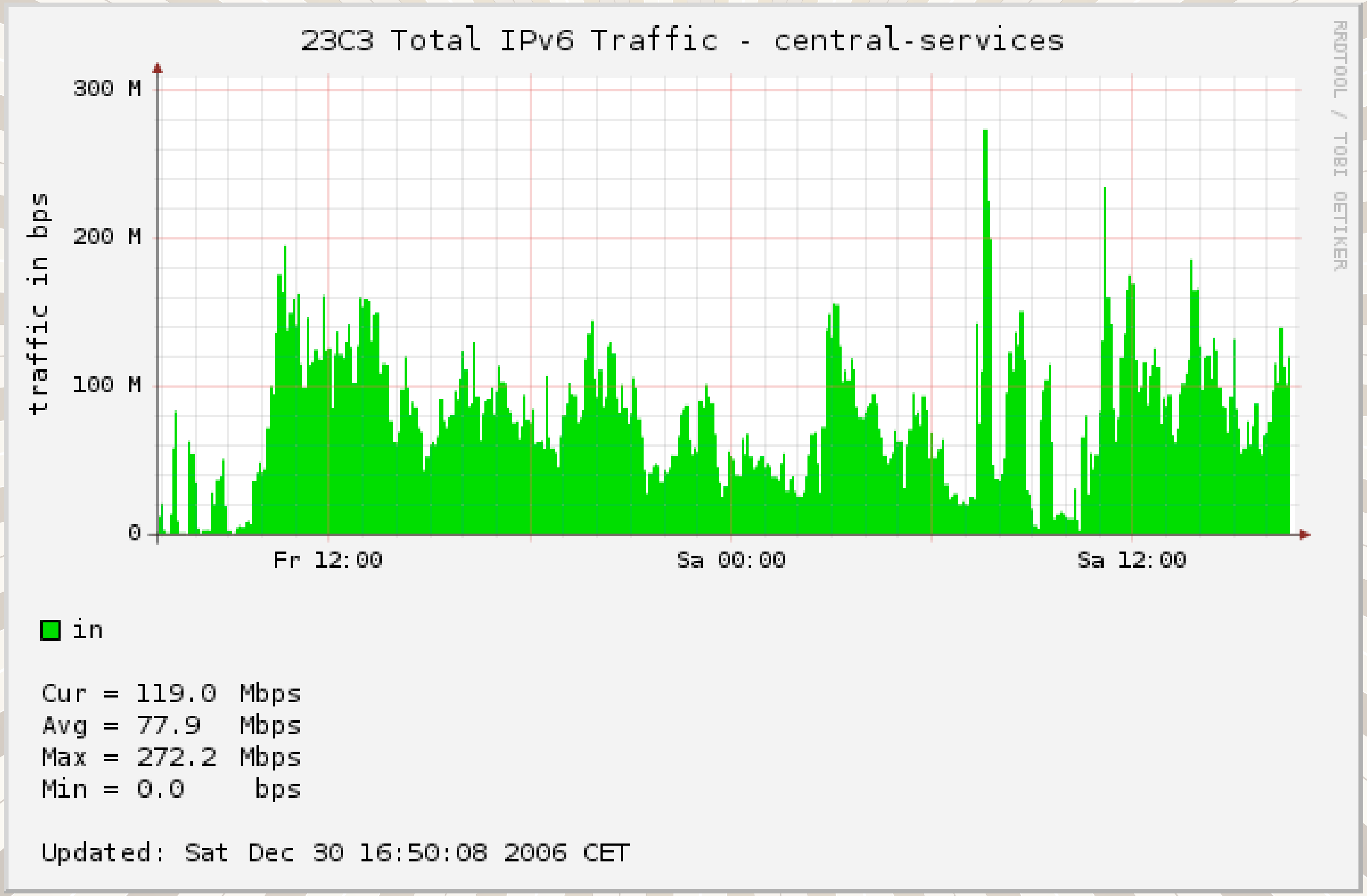
23C3 - Results



	Current	Average	Maximum	Minimum
■ other	0.0%	0.0%	0.0%	0.0%
🟥 ARP	0.0%	0.0%	0.1%	0.0%
🟧 IPv6	2.1%	1.3%	4.5%	0.0%
□ IPv4	97.9%	98.7%	100.0%	95.5%

Updated: Sat Dec 30 14:51:28 2006 CET

23C3 - Results



23C3 Total IPv6 Traffic - central-services

🟩 in

```
Cur = 119.0  Mbps
Avg = 77.9   Mbps
Max = 272.2  Mbps
Min = 0.0     bps
```

Updated: Sat Dec 30 16:50:08 2006 CET

Thank You!

- Questions?