Adding QoS Attributes to
the IPFIX Protocol

NORDUnet conference 2006
Sept 28

Arne Øslebo
arine.oslebo@uninett.no
Outline

• Brief NetFlow/IPFIX introduction
• Motivation
• Implementation
  – Exporter
  – Collector
  – Presentation
• Performance
• Future work
NetFlow

- Originally Cisco technology - 1996
- IETF is standardizing NetFlow in the IPFIX working group
- IPFIX definition of an IP network flow:
  - A set of IP packets passing an observation point in a network during a certain time interval. All packets belonging to a particular flow have a set of common properties.
- Flow Key
  - Each of the properties that are used for defining a flow
- Flow record
  - Measured properties of a flow
Flow records

- **NetFlow v5**
  - Src&dst IP address
  - Next hop router's IP address
  - In&out interface index
  - Pkts and bytes in the flow
  - sysUptime at start and end of flow
  - TCP/UDP src and dst port number
  - Type of service
  - TCP flags
  - IP protocol
  - Src&dst AS number
  - Src&dst address prefix mask bits

- **NetFlow v9**
  - Configurable

- **IPFIX**
  - Based on NetFlow v9
  - Possible to add enterprise specific attributes
Motivation

- **Standard NetFlow** can tell you things like:
  - which AS number you send the most traffic to
  - which IP address sends/receives the most traffic
  - ++++

- **It can NOT tell you**
  - who initiated the traffic
  - the quality of the traffic
    - retransmissions
    - jitter
    - burstiness
  - the type of application that generates traffic

- **We want to extend IPFIX with enterprise specific attributes** so that this information is available
Framework

- Passive monitoring card
- MAPI
- Flow collector based on NERD
- Stager backend

Stager user interface

Stager DB

Collector

Exporter

http://www.ist-lobster.org

lobster

UNINETT
Exporter

- Supports: NetFlow v5, v9 and IPFIX
- Part of Monitoring API (MAPI)
  - [http://mapi.uninett.no](http://mapi.uninett.no)
- New attributes:
  - pktLenHistogram, pktDistHistogram
  - pktPayload
  - pkt[Dist/Length][Var/Sum/SumQ]
  - direction
  - reordered
  - rtpJitter, rtpLostFraction, rtpLostPackets, rtpSequenceCycles
  - maxRate[1sec/100ms/10ms/1ms]
  - minRate[1sec/100ms/10ms/1ms]
  - service
Measurement probes

Part of the GigaCampus project
Collector

- Modified version of NERD
  - [http://www.nerdd.org](http://www.nerdd.org)
- Stores flow records to file
- FlowStat generates high level reports that are inserted into the Stager database.

FlowStat -o "2 desc" -f "packets>500" -s "dst_as sum(octets) avg(hpktdst)"
# dst_as sum(octets) avg(hpktdst)
0, 800721047, 28_0c_08_0c_4c_27_08_04_0c_d3_,
15659, 719072910, 1a_06_03_04_20_07_04_02_08_f2_,
8642, 615117228, 25_0c_07_0c_47_16_07_0a_10_dd_,
3301, 592188599, 12_04_05_0a_3f_0e_02_01_0b_ef_,
2119, 446154823, 1d_06_03_06_2a_07_02_01_04_f0_,
Stager – user interface

- A web-based tool for aggregating and presenting most types of network statistics
- GPL
- 140+ people on a public mailing list
- Easy to use web frontend
  - Text based reports and graphs
- Easy to add new reports
  - Templates and plugins
- Statistical functions
- Access control
  - Observation points and reports
- Scalable
  - Handle large volumes of data and store years of statistics
### Destination AS report

**Monday 18. September 2006, 15:00**  
??? (in, 1/1)

#### Pie chart

#### Plot graph

<table>
<thead>
<tr>
<th>Dst AS</th>
<th>Octets</th>
<th>Packets</th>
<th>Flows</th>
<th>Out of sequence</th>
<th>Octets</th>
<th>Packets</th>
<th>Flows</th>
<th>% of flows</th>
<th>Octets</th>
<th>Packets</th>
<th>Flows</th>
<th>% of flows</th>
</tr>
</thead>
<tbody>
<tr>
<td>SCV-AS-AP SCV Broadband Access Provider</td>
<td>633k</td>
<td>118</td>
<td>6.41</td>
<td>35.4-10³</td>
<td>2.29-10³</td>
<td>9.91%</td>
<td>15</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TELEFONICA-DATA-ESPAÑA Internet Access Network of TDE</td>
<td>6.45M</td>
<td>1.16-10³</td>
<td>1.18%</td>
<td>192-10³</td>
<td>13.0-10³</td>
<td>5.89%</td>
<td>15</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PROXAD AS for Proxad ISP</td>
<td>3.93M</td>
<td>676</td>
<td>20.9</td>
<td>43.2-10³</td>
<td>4.24-10³</td>
<td>5.63%</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ASS413 PIPEX Communications</td>
<td>887k</td>
<td>185</td>
<td>5.57</td>
<td>34.4-10³</td>
<td>1.67-10³</td>
<td>5.33%</td>
<td>32</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AS-INET iiNet Limited</td>
<td>1.11M</td>
<td>194</td>
<td>5.10</td>
<td>14.0-10³</td>
<td>928</td>
<td>4.99%</td>
<td>15</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TMNET-AS-AP TM Net. Internet Service Provider</td>
<td>990k</td>
<td>170</td>
<td>16.8</td>
<td>51.0-10³</td>
<td>2.88-10³</td>
<td>4.76%</td>
<td>18</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MFX-AS Microplex PTY LTD</td>
<td>1.77M</td>
<td>292</td>
<td>6.04</td>
<td>15.9-10³</td>
<td>1.00-10³</td>
<td>4.62%</td>
<td>16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Net Sentry Corp</td>
<td>22.2k</td>
<td>27.7</td>
<td>7.95</td>
<td>1.31-10³</td>
<td>1.28-10³</td>
<td>4.47%</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TTNET Tntnet Autonomous System</td>
<td>1.23M</td>
<td>182</td>
<td>11.2</td>
<td>26.8-10³</td>
<td>1.61-10³</td>
<td>4.01%</td>
<td>17</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Telecomunicacoes de Santa Catarina SA</td>
<td>605k</td>
<td>110</td>
<td>7.28</td>
<td>7.21-10³</td>
<td>926</td>
<td>3.53%</td>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
## Destination AS report

![Image of network interface destination AS report](image.png)

**Monday 18. September 2006, 15:00**

### Destination AS Report

<table>
<thead>
<tr>
<th>Select</th>
<th>Dst AS</th>
<th>Name</th>
<th>Traffic</th>
<th>Requests</th>
<th>Responses</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
<td>0</td>
<td>15.4G</td>
<td>480·10^3</td>
<td>26.98%</td>
</tr>
<tr>
<td></td>
<td>15659</td>
<td>NEXTGENTEL Autonomous System</td>
<td>15.1G</td>
<td>480·10^3</td>
<td>13.88%</td>
</tr>
<tr>
<td></td>
<td>8642</td>
<td>B2 Bredband AB (pub)</td>
<td>11.2G</td>
<td>73.4·10^3</td>
<td>12.12%</td>
</tr>
<tr>
<td></td>
<td>2119</td>
<td>Telenor-NEXTEL Teleon Internet Access</td>
<td>8.43G</td>
<td>135·10^3</td>
<td>15.05%</td>
</tr>
<tr>
<td></td>
<td>3301</td>
<td>TeliaNet-Sweden TeliaNet Sweden</td>
<td>8.05G</td>
<td>130·10^3</td>
<td>30.64%</td>
</tr>
<tr>
<td></td>
<td>3320</td>
<td>DTAG Deutsche Telekom AG</td>
<td>5.01G</td>
<td>158·10^3</td>
<td>4.31%</td>
</tr>
<tr>
<td></td>
<td>6830</td>
<td>UPC UPC Broadband</td>
<td>4.64G</td>
<td>92.7·10^3</td>
<td>6.77%</td>
</tr>
<tr>
<td></td>
<td>7132</td>
<td>SBC Internet Services</td>
<td>4.24G</td>
<td>110·10^3</td>
<td>6.82%</td>
</tr>
<tr>
<td></td>
<td>8394</td>
<td>Alfanet Autonomous System</td>
<td>3.61G</td>
<td>28.8·10^3</td>
<td>7.67%</td>
</tr>
<tr>
<td></td>
<td>3215</td>
<td>AS3215 France Telecom Transpac</td>
<td>3.55G</td>
<td>93.6·10^3</td>
<td>4.80%</td>
</tr>
</tbody>
</table>

*Stager, 2004-2008 © UNINETT AS*

*Processing the report took 70.70ms*
## Destination IP report

**Destination IP**

**Monday 18. September 2006, 15:00**

<table>
<thead>
<tr>
<th>Dst IP</th>
<th>Dcctets</th>
<th>Rate 1 second</th>
<th>Rate 100 milliseconds</th>
<th>Rate 10 milliseconds</th>
</tr>
</thead>
<tbody>
<tr>
<td>wx.y.z</td>
<td>18.8M</td>
<td>10.5M</td>
<td>780k</td>
<td>15.8M</td>
</tr>
<tr>
<td>wx.y.z</td>
<td>9.01M</td>
<td>6.21M</td>
<td>603k</td>
<td>6.37M</td>
</tr>
<tr>
<td>wx.y.z</td>
<td>7.75M</td>
<td>195k</td>
<td>75.3k</td>
<td>1.33M</td>
</tr>
<tr>
<td>wx.y.z</td>
<td>8.79M</td>
<td>143k</td>
<td>140k</td>
<td>937k</td>
</tr>
<tr>
<td>wx.y.z</td>
<td>7.48M</td>
<td>45.9k</td>
<td>2.38k</td>
<td>143k</td>
</tr>
</tbody>
</table>

**Processing the report took 131.8ms**
Performance - exporter

- Depends on included attributes in the flow records
- Most common attributes
  - Tested on OC48 without problems

<table>
<thead>
<tr>
<th>Element</th>
<th>Instructions</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAPI</td>
<td>200</td>
</tr>
<tr>
<td>Normal attributes</td>
<td>680</td>
</tr>
<tr>
<td>HIST_PKT_LEN</td>
<td>30</td>
</tr>
<tr>
<td>HIST_PKT_DIST</td>
<td>495</td>
</tr>
<tr>
<td>VAR of DIST/LEN</td>
<td>100</td>
</tr>
<tr>
<td>MAX/MIN RATE</td>
<td>300</td>
</tr>
<tr>
<td>SERVICE</td>
<td>1760</td>
</tr>
</tbody>
</table>
Performance - collector

- Collecting and storing to disk is no problem
- Current bottleneck:
  - Processing data to generate high level reports
  - >1 hour to process 1 hour of data
- Main problem
  - Not possible to generate multiple reports in one pass
  - Some reports demands multiple passes
Performance - Stager

- Highly scalable
- Our NetFlow setup:
  - 27 routers
  - 207 interfaces
  - >30Gb of raw Netflow data every day
  - 400,000 new entries in the db every hour
  - >450 millions entries in a single table
  - >700Gb database size

<table>
<thead>
<tr>
<th></th>
<th>PC1</th>
<th>PC2</th>
<th>PC3</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Netflow size</td>
<td>537MB</td>
<td>161MB</td>
<td>399MB</td>
<td>1097MB</td>
</tr>
<tr>
<td>Sequentially</td>
<td>9min 17s</td>
<td>3min 8s</td>
<td>5min 51s</td>
<td>18min 16s</td>
</tr>
<tr>
<td>No insert in DB</td>
<td>7min 11s</td>
<td>1min 48s</td>
<td>4min 52s</td>
<td>13min 51s</td>
</tr>
<tr>
<td>Simultaneously</td>
<td>9min 21s</td>
<td>3min 7s</td>
<td>5min 56s</td>
<td>18min 24s</td>
</tr>
<tr>
<td># of new DB entries</td>
<td>164702</td>
<td>69549</td>
<td>184706</td>
<td>418957</td>
</tr>
<tr>
<td># of entries/second</td>
<td>295.69</td>
<td>369.94</td>
<td>526.23</td>
<td>382.26</td>
</tr>
</tbody>
</table>
Future work

• Verify correctness of attributes
• Improve performance of FlowStat
  – Generate multiple reports on one pass over raw NetFlow files
• Full deployment on all measurement probes
  – Gain more experience to see which attributes are the most useful
Software availability

• http://software.uninett.no
• E-mail contact:
  – arne.oslebo@uninett.no