Extending the IPFIX protocol for better QoS monitoring

22.05.07

Olav Kvittem, Arne Øslebø
End to end QoS measurements

- Goal: Knowing the QoS by measurements
- Traditional SNMP/netflow measure volumes of packets and errors on network components - not quality of service
- End to end active measurement does not scale
- Both engineering and customer oriented statistics
- User deserve end-to-end inter domain view
- Passive probes see the flows quality
Scampi/Lobster Software

- **MAPI - Measurement API – in C**
  - Interfaces passive measurement cards (DAG, Combo6)
  - Abstraction, sharing, branching, anonymization, efficiency (0-copy)
  - Functions – filtering hw/sw, counters, flow analysis

- **Applications**
  - Service detection - appmon
  - Polymorphic attack detector
  - Extended flow analysis – Stager
    - SubSecond Bandwidth measurement (SSB)
Flow characterization

- MAPI with passive monitoring cards
  - Investigate flows per packet with microsec clock
- Extended IPFIX flow records with statistics for intensity, intervals, sizes
  - Will allow us to assess quality by service and location (AS or IP prefix).
  - Research actual quality of services
IPFIX flow

- IPFIX is an IETF standardization inspired by netflow v9 (Cisco)
- Extensible: custom defined record types and parameters
- Flow is defined by a flow key:
  - std. 5-tuple: IP-addresses, transport protocol and ports
  - end by FIN, time gap or max
  -Parms: bytes, octets, time, AS, protocol, ..
Extended flow parms.

- Compute cheap numbers: count, sum, sum of squares => standard deviation
- Packet size distribution
- Interarrival time distribution
- Bit rate vs time – 1, 10, 100, 1000 ms (max/min)
- RTP time-stamps/payload-type
- TCP properties: windows, retransmissions/out-of-seq
- direction - initiator
- Service classification
Framework

- Flow collector based on NERD
- Stager backend

- Passive monitoring card
- MAPI

Stager user interface

http://www.ist-lobster.org
### Destination IP report

![Destination IP report](image)

<table>
<thead>
<tr>
<th>Select</th>
<th>IP address</th>
<th>Total</th>
<th>Max</th>
<th>Min</th>
<th>Max</th>
<th>Min</th>
<th>Max</th>
<th>Min</th>
<th>Max</th>
<th>Min</th>
</tr>
</thead>
<tbody>
<tr>
<td>✔️</td>
<td>w.x.y.z</td>
<td>18.8M</td>
<td>10.5M</td>
<td>780k</td>
<td>15.8M</td>
<td>0</td>
<td>53.8M</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>✔️</td>
<td>w.x.y.z</td>
<td>8.91M</td>
<td>5.21M</td>
<td>583k</td>
<td>5.37M</td>
<td>0</td>
<td>32.4M</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>✔️</td>
<td>w.x.y.z</td>
<td>7.75M</td>
<td>195k</td>
<td>75.3k</td>
<td>1.33M</td>
<td>163k</td>
<td>4.72M</td>
<td>5.78k</td>
<td></td>
<td></td>
</tr>
<tr>
<td>✔️</td>
<td>w.x.y.z</td>
<td>8.79M</td>
<td>143k</td>
<td>140k</td>
<td>937k</td>
<td>731k</td>
<td>7.62M</td>
<td>315k</td>
<td></td>
<td></td>
</tr>
<tr>
<td>✔️</td>
<td>w.x.y.z</td>
<td>7.48M</td>
<td>45.9k</td>
<td>2.38k</td>
<td>143k</td>
<td>1.88k</td>
<td>1.22M</td>
<td>7.34k</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Mon, 18. September 2006, 15:00

???(in, 1/1)

---

*Stager, 2004-2006 © UNINETT AS*

*Processing the report took 131.8ms*
SSB
(Sub-Second Bandwidth)

Bytes and packets

Wednesday 17. January 2007

Bytes and packets for trd-oslo

- trd-oslo (c_bps)
- trd-oslo (c_maxbps1s)
- trd-oslo (c_maxbps100ms)
Flow length

Flow duration distribution

- 0ms: 50
- 10ms: 1
- 100ms: 1
- 1s: 10
- 10s: 10
- 100s: 5
- 241s: 1
Web performance

- Flows > 100 ms to measure 100ms bursts
- TCP Src port 80 = web servers
- who initiated the flow – TCP SYN
  - request 9.1 mflow 63.2 GB
  - response 6.2 mflow 213.3 GB
  - uncertain 7.8 mflow 729.9 GB
**AS - Web server performance**

- 1ms 10ms 100ms eff win
- Kbps Kbps Kbps Byte
- 52535 7330 1582 5743
- 37056 4793 1729 14456
- 56092 8710 2147 2762
- 41778 5975 2193 13628
- 54324 7476 2197 14513
- average flows per AS, > 100 ms src port 80
Out of sequence TCP

- Count of TCP sequence < previous
- does not distinguish:
  - retransmission of lost packets
  - reordered by network
    - cost CPU for sequencing out of fast path
- Some networks reorder
- distinguish by retransmission time >> reordered packet distance?
## Most out of sequence dst-AS

- out 10ms 100ms eff win
- % Kbps Kbps Byte
- 0.9 8449 3482 7105
- 1.4 4584 502 5210
- 1.5 5903 1007 8814
- ..................
- 11.9 1917 204 3081
- 13.6 1054 112 1650
- 15.1 1445 156 2467
- 21.4 1973 208 3120
Scaling MAPI

- MAPI/ipfixlib copes on 2.5Gbps
- 10Gbps demands
  - parallellism - splitting captured data on more cpus for processing
  - offloading processing to hardware - flow generation too complex?
- Faster report-generator – in parallel - ipfix/netflow v6?
Summary

- Passive flow measurements can
  - measure performance and indicate quality
  - show bursts in flows – 1ms too low
  - TCP effective window and out of sequence indicates bad performance
  - RTP measurements for further study
- Application recognition demanding
- More work to be done

- Thanks to Scampi and Lobster partners

- Our software: http://software.uninett.no
- Lobster: http://www.ist-lobster.org