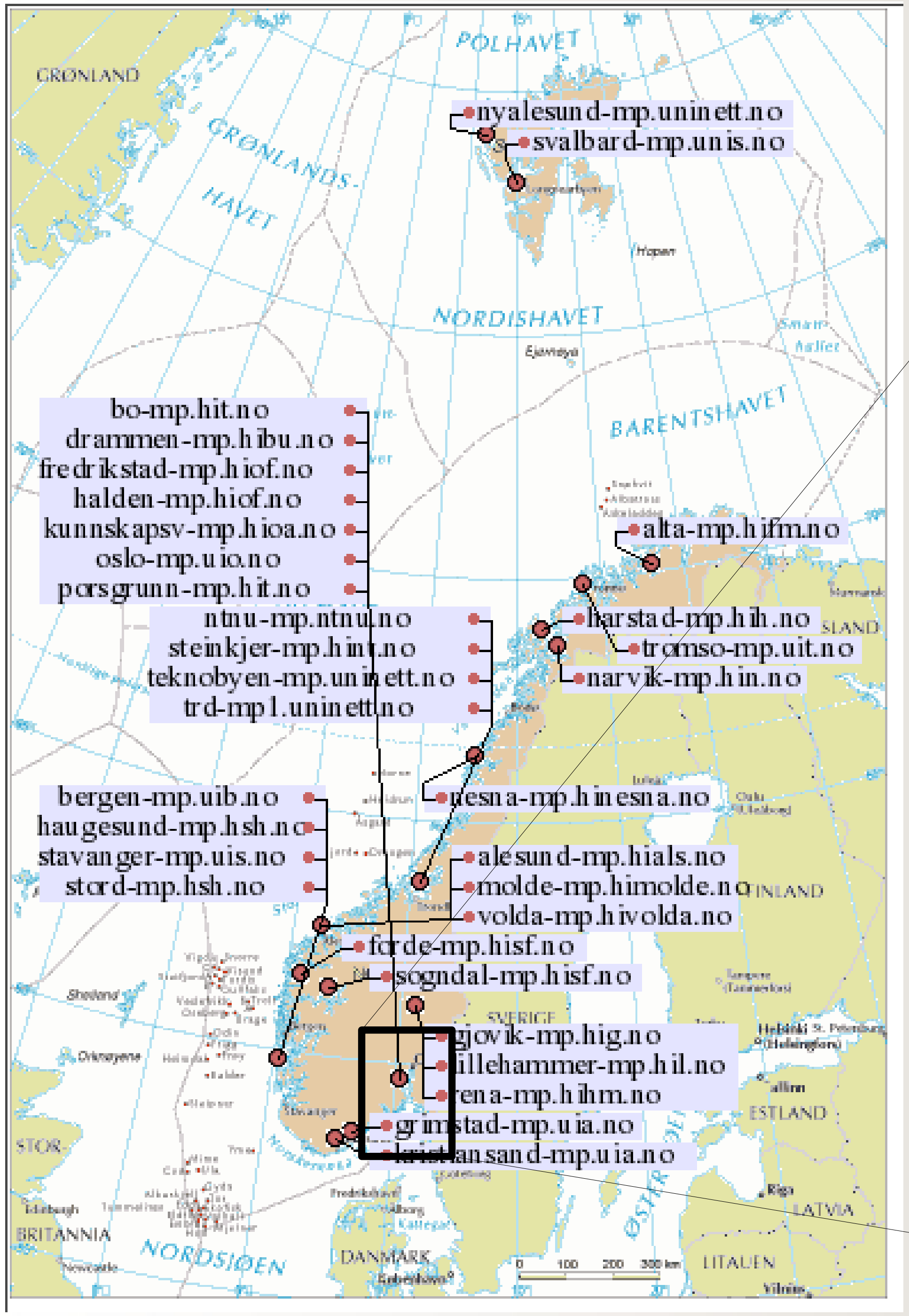
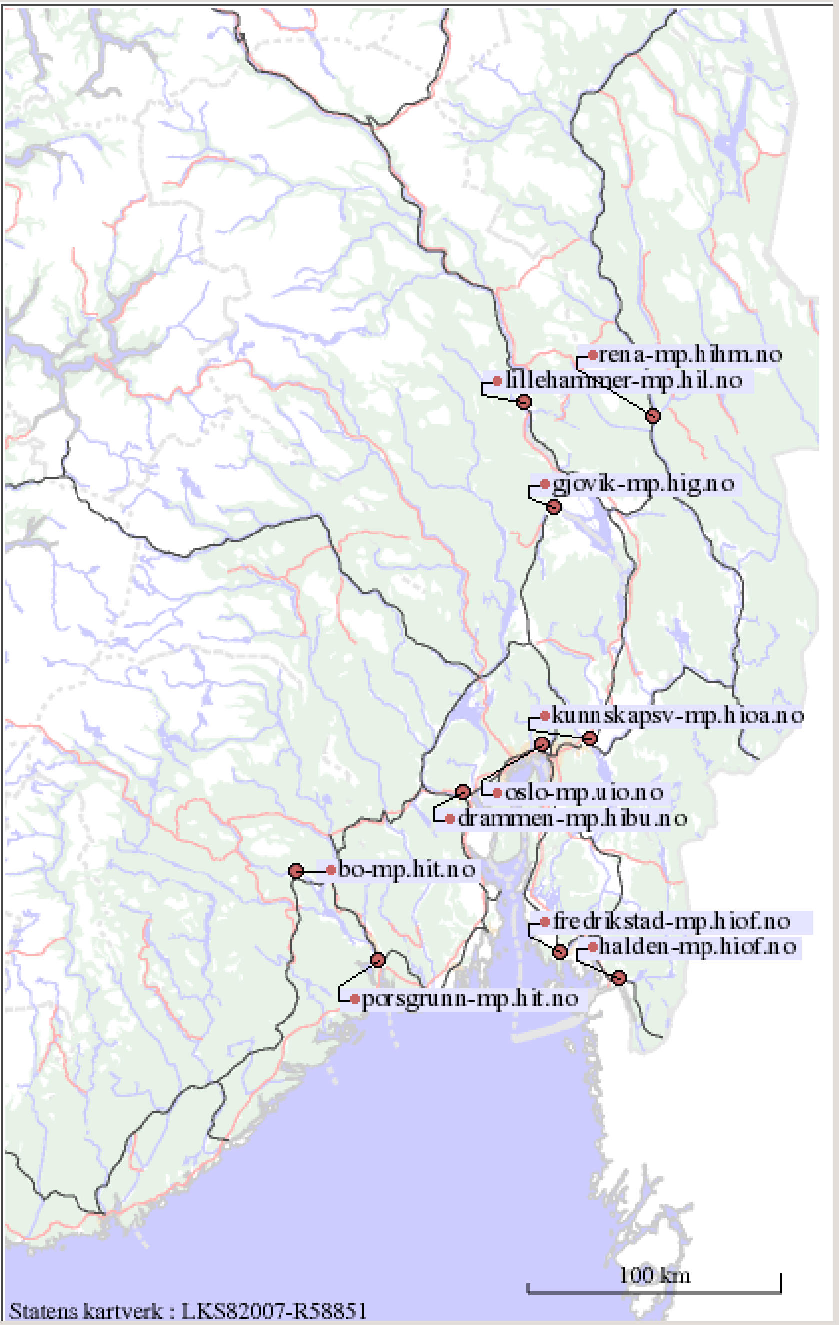
# Large scale passive monitoring at 10Gbps on commodity hardware

Campus network monitoring and security workshop
April 24, 2014
Arne Øslebø, arne.oslebo@uninett.no

UNINETT

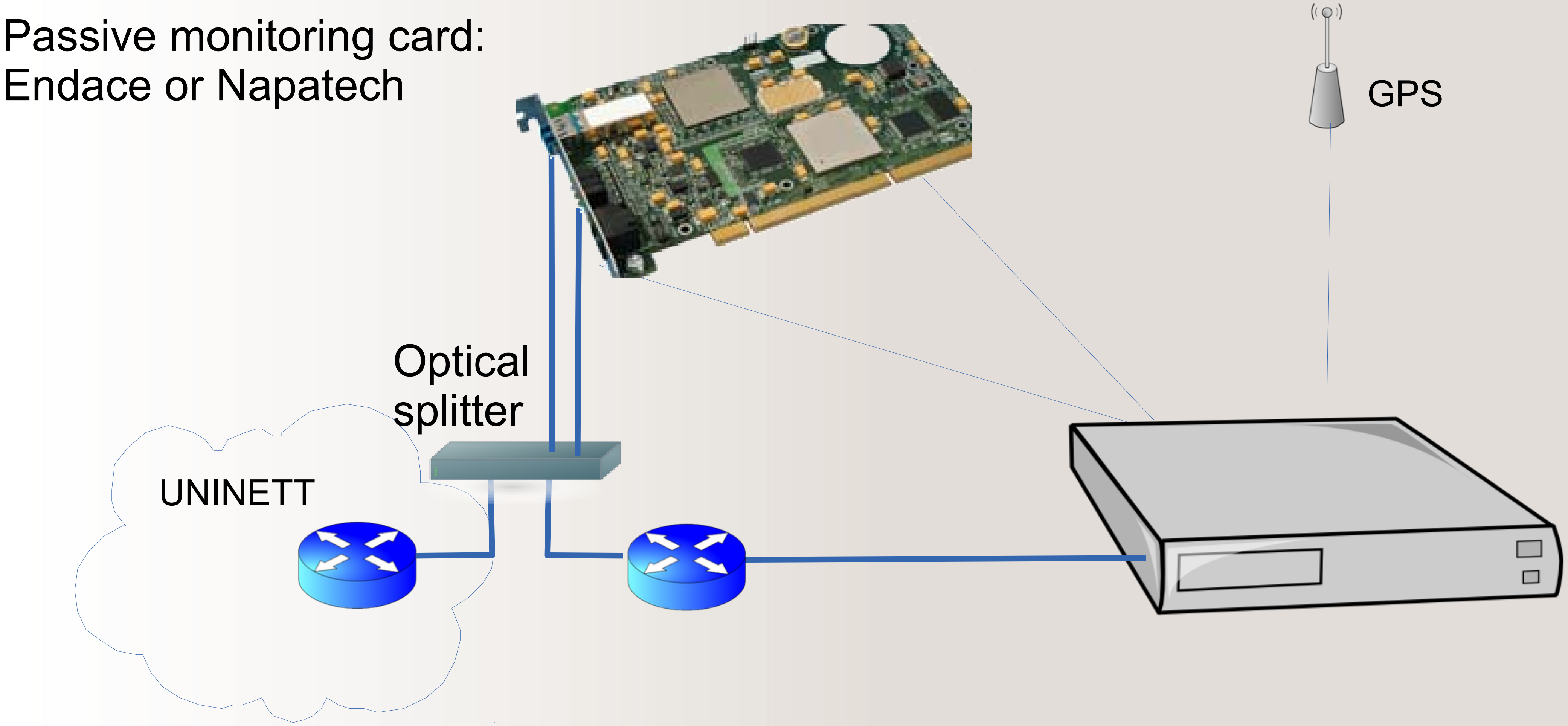
# UNINETT monitoring infrastructure





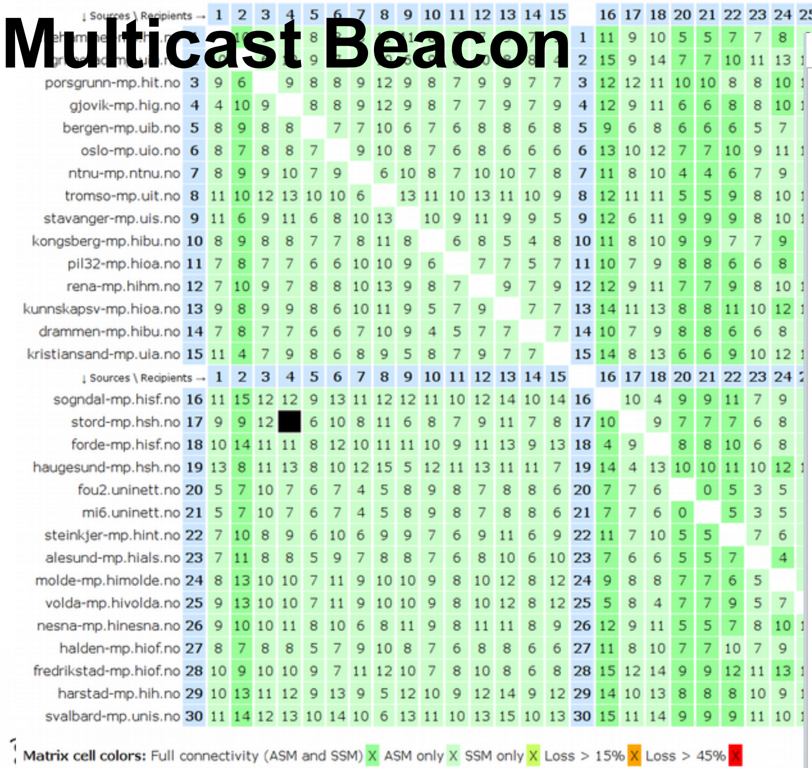
Statens kartverk : LKS82007-R58851

# Original hardware setup



# Active and passive monitoring

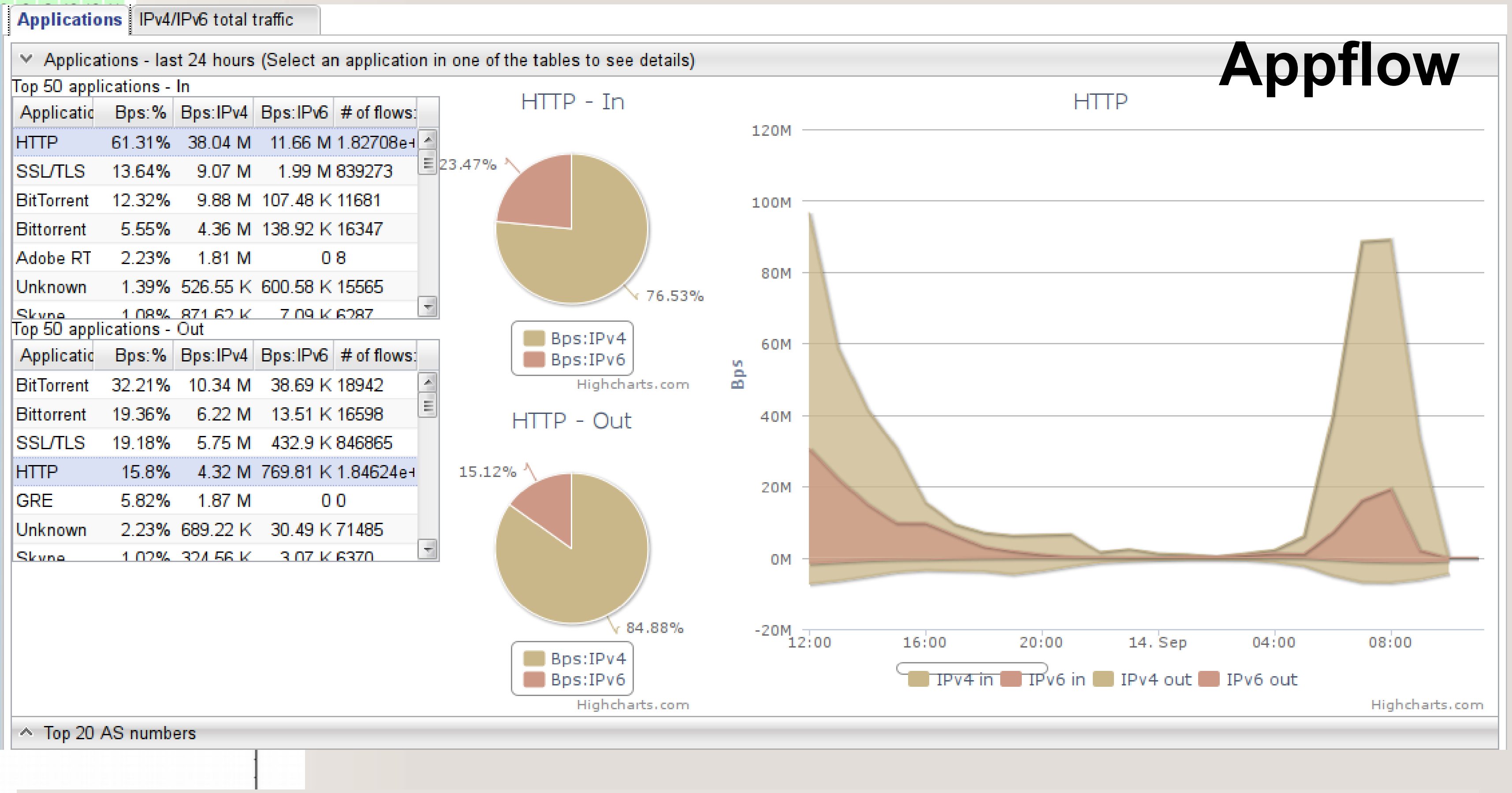
## Multicast Beacon



| ↓ Sources \ Recipients → | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | | 16 | 17 | 18 | 20 | 21 | 22 | 23 | 24 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | 8 | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | 1 | 11 | 9 | 10 | 5 | 5 | 7 | 7 | 8 |
| [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | 9 | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | 4 | 2 | 15 | 9 | 14 | 7 | 7 | 10 | 11 | 13 |
| porsgrunn-mp.hit.no 3 | 9 | 6 | | 9 | 8 | 8 | 9 | 12 | 9 | 8 | 7 | 9 | 9 | 7 | 7 | 3 | 12 | 12 | 11 | 10 | 10 | 8 | 8 | 10 |
| gjovik-mp.hig.no 4 | 4 | 10 | 9 | | 8 | 8 | 9 | 12 | 9 | 8 | 7 | 7 | 9 | 7 | 9 | 4 | 12 | 9 | 11 | 6 | 6 | 8 | 8 | 10 |
| bergen-mp.uib.no 5 | 8 | 9 | 8 | 8 | | 7 | 7 | 10 | 6 | 7 | 6 | 8 | 8 | 6 | 8 | 5 | 9 | 6 | 8 | 6 | 6 | 6 | 5 | 7 |
| oslo-mp.uio.no 6 | 8 | 7 | 8 | 8 | 7 | | 9 | 10 | 8 | 7 | 6 | 8 | 6 | 6 | 6 | 6 | 13 | 10 | 12 | 7 | 7 | 10 | 9 | 11 |
| ntnu-mp.ntnu.no 7 | 8 | 9 | 9 | 10 | 7 | 9 | | 6 | 10 | 8 | 7 | 10 | 10 | 7 | 8 | 7 | 11 | 8 | 10 | 4 | 4 | 6 | 7 | 9 |
| tromso-mp.uit.no 8 | 11 | 10 | 12 | 13 | 10 | 10 | 6 | | 13 | 11 | 10 | 13 | 11 | 10 | 9 | 8 | 12 | 11 | 11 | 5 | 5 | 9 | 8 | 10 |
| stavanger-mp.uis.no 9 | 11 | 6 | 9 | 11 | 6 | 8 | 10 | 13 | | 10 | 9 | 11 | 9 | 9 | 5 | 9 | 12 | 6 | 11 | 9 | 9 | 9 | 8 | 10 |
| kongsberg-mp.hibu.no 10 | 8 | 9 | 8 | 8 | 7 | 7 | 8 | 11 | 8 | | 6 | 8 | 5 | 4 | 8 | 10 | 11 | 8 | 10 | 9 | 9 | 7 | 7 | 9 |
| pil32-mp.hioa.no 11 | 7 | 8 | 7 | 7 | 6 | 6 | 10 | 10 | 9 | 6 | | 7 | 7 | 5 | 7 | 11 | 10 | 7 | 9 | 8 | 8 | 6 | 6 | 8 |
| rena-mp.hihm.no 12 | 7 | 10 | 9 | 7 | 8 | 8 | 10 | 13 | 9 | 8 | 7 | | 9 | 7 | 9 | 12 | 12 | 9 | 11 | 7 | 7 | 9 | 8 | 10 |
| kunnskapsv-mp.hioa.no 13 | 9 | 8 | 9 | 9 | 8 | 6 | 10 | 11 | 9 | 5 | 7 | 9 | | 7 | 7 | 13 | 14 | 11 | 13 | 8 | 8 | 11 | 10 | 12 |
| drammen-mp.hibu.no 14 | 7 | 8 | 7 | 7 | 6 | 6 | 7 | 10 | 9 | 4 | 5 | 7 | 7 | | 7 | 14 | 10 | 7 | 9 | 8 | 8 | 6 | 6 | 8 |
| kristiansand-mp.uia.no 15 | 11 | 4 | 7 | 9 | 8 | 6 | 8 | 9 | 5 | 8 | 7 | 9 | 7 | 7 | | 15 | 14 | 8 | 13 | 6 | 6 | 9 | 10 | 12 |
| ↓ Sources \ Recipients → | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | | 16 | 17 | 18 | 20 | 21 | 22 | 23 | 24 |
| sogndal-mp.hisf.no 16 | 11 | 15 | 12 | 12 | 9 | 13 | 11 | 12 | 12 | 11 | 10 | 12 | 14 | 10 | 14 | 16 | | 10 | 4 | 9 | 9 | 11 | 7 | 9 |
| stord-mp.hsh.no 17 | 9 | 9 | 12 | ■ | 6 | 10 | 8 | 11 | 6 | 8 | 7 | 9 | 11 | 7 | 8 | 17 | 10 | | 9 | 7 | 7 | 7 | 6 | 8 |
| forde-mp.hisf.no 18 | 10 | 14 | 11 | 11 | 8 | 12 | 10 | 11 | 11 | 10 | 9 | 11 | 13 | 9 | 13 | 18 | 4 | 9 | | 8 | 8 | 10 | 6 | 8 |
| haugesund-mp.hsh.no 19 | 13 | 8 | 11 | 13 | 8 | 10 | 12 | 15 | 5 | 12 | 11 | 13 | 11 | 11 | 7 | 19 | 14 | 4 | 13 | 10 | 10 | 11 | 10 | 12 |
| fou2.uninett.no 20 | 5 | 7 | 10 | 7 | 6 | 7 | 4 | 5 | 8 | 9 | 8 | 7 | 8 | 8 | 6 | 20 | 7 | 7 | 6 | | 0 | 5 | 3 | 5 |
| mi6.uninett.no 21 | 5 | 7 | 10 | 7 | 6 | 7 | 4 | 5 | 8 | 9 | 8 | 7 | 8 | 8 | 6 | 21 | 7 | 7 | 6 | 0 | | 5 | 3 | 5 |
| steinkjer-mp.hint.no 22 | 7 | 10 | 8 | 9 | 6 | 10 | 6 | 9 | 9 | 7 | 6 | 9 | 11 | 6 | 9 | 22 | 11 | 7 | 10 | 5 | 5 | | 7 | 6 |
| alesund-mp.hials.no 23 | 7 | 11 | 8 | 8 | 5 | 9 | 7 | 8 | 8 | 7 | 6 | 8 | 10 | 6 | 10 | 23 | 7 | 6 | 6 | 5 | 5 | 7 | | 4 |
| molde-mp.himolde.no 24 | 8 | 13 | 10 | 10 | 7 | 11 | 9 | 10 | 10 | 9 | 8 | 10 | 12 | 8 | 12 | 24 | 9 | 8 | 8 | 7 | 7 | 6 | 5 | |
| volda-mp.hivolda.no 25 | 9 | 13 | 10 | 10 | 7 | 11 | 9 | 10 | 10 | 9 | 8 | 10 | 12 | 8 | 12 | 25 | 5 | 8 | 4 | 7 | 7 | 9 | 5 | 7 |
| nesna-mp.hinesna.no 26 | 9 | 10 | 10 | 11 | 8 | 10 | 6 | 8 | 11 | 9 | 8 | 11 | 11 | 8 | 9 | 26 | 12 | 9 | 11 | 5 | 5 | 7 | 8 | 10 |
| halden-mp.hiof.no 27 | 8 | 7 | 8 | 8 | 5 | 7 | 9 | 10 | 8 | 7 | 6 | 8 | 8 | 6 | 6 | 27 | 11 | 8 | 10 | 7 | 7 | 10 | 7 | 9 |
| fredrikstad-mp.hiof.no 28 | 10 | 9 | 10 | 10 | 9 | 7 | 11 | 12 | 10 | 7 | 8 | 10 | 8 | 6 | 8 | 28 | 15 | 12 | 14 | 9 | 9 | 12 | 11 | 13 |
| harstad-mp.hih.no 29 | 10 | 13 | 11 | 12 | 9 | 13 | 9 | 5 | 12 | 10 | 9 | 12 | 14 | 9 | 12 | 29 | 14 | 10 | 13 | 8 | 8 | 8 | 10 | 9 |
| svalbard-mp.unis.no 30 | 11 | 14 | 12 | 13 | 10 | 14 | 10 | 6 | 13 | 11 | 10 | 13 | 15 | 10 | 13 | 30 | 15 | 11 | 14 | 9 | 9 | 9 | 11 | 10 |

**Matrix cell colors:** Full connectivity (ASM and SSM) X ASM only X SSM only X Loss > 15% X Loss > 45% X

## Appflow



Applications | IPv4/IPv6 total traffic

Applications - last 24 hours (Select an application in one of the tables to see details)

Top 50 applications - In

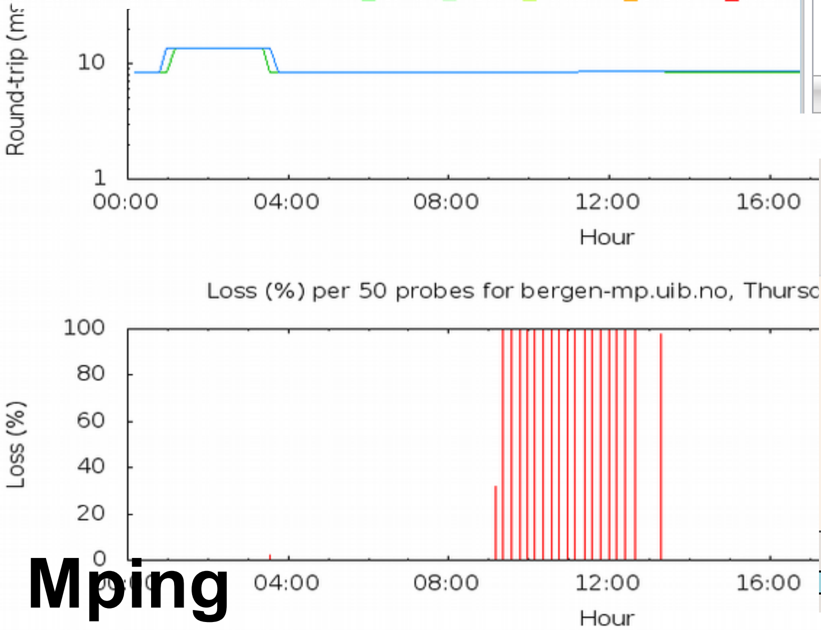
| Applicatic | Bps:% | Bps:IPv4 | Bps:IPv6 | # of flows: |
|---|---|---|---|---|
| HTTP | 61.31% | 38.04 M | 11.66 M | 1.82708e+ |
| SSL/TLS | 13.64% | 9.07 M | 1.99 M | 839273 |
| BitTorrent | 12.32% | 9.88 M | 107.48 K | 11681 |
| Bittorrent | 5.55% | 4.36 M | 138.92 K | 16347 |
| Adobe RT | 2.23% | 1.81 M | 0 | 8 |
| Unknown | 1.39% | 526.55 K | 600.58 K | 15565 |
| Skype | 1.08% | 871.62 K | 7.09 K | 6287 |

Top 50 applications - Out

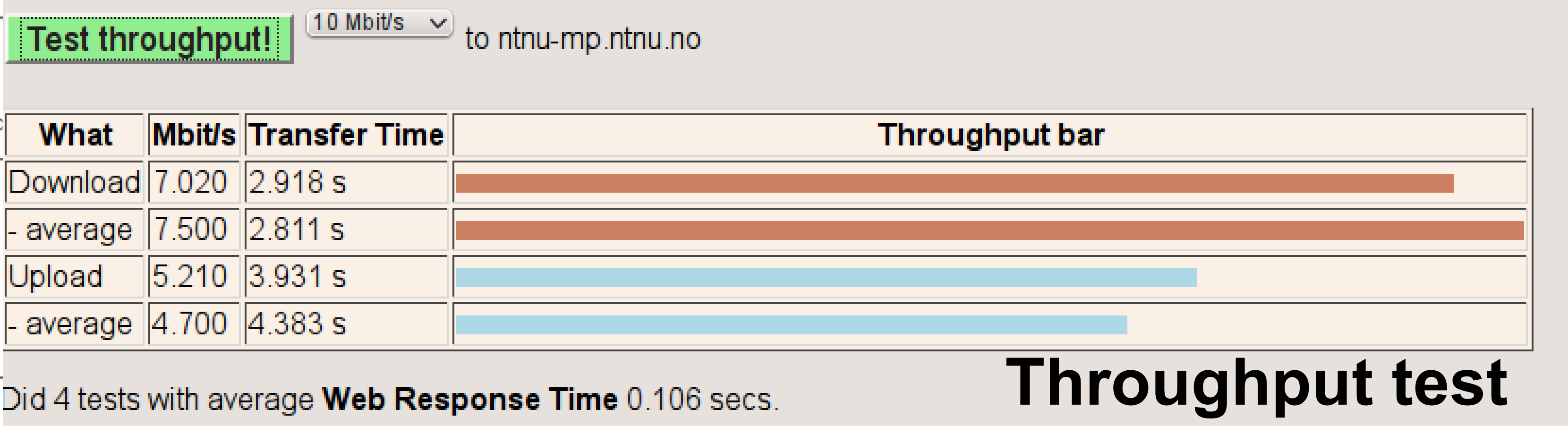
| Applicatic | Bps:% | Bps:IPv4 | Bps:IPv6 | # of flows: |
|---|---|---|---|---|
| BitTorrent | 32.21% | 10.34 M | 38.69 K | 18942 |
| Bittorrent | 19.36% | 6.22 M | 13.51 K | 16598 |
| SSL/TLS | 19.18% | 5.75 M | 432.9 K | 846865 |
| HTTP | 15.8% | 4.32 M | 769.81 K | 1.84624e+ |
| GRE | 5.82% | 1.87 M | 0 | 0 |
| Unknown | 2.23% | 689.22 K | 30.49 K | 71485 |
| Skype | 1.02% | 324.56 K | 3.07 K | 6370 |

Top 20 AS numbers

## Mping



## Throughput test



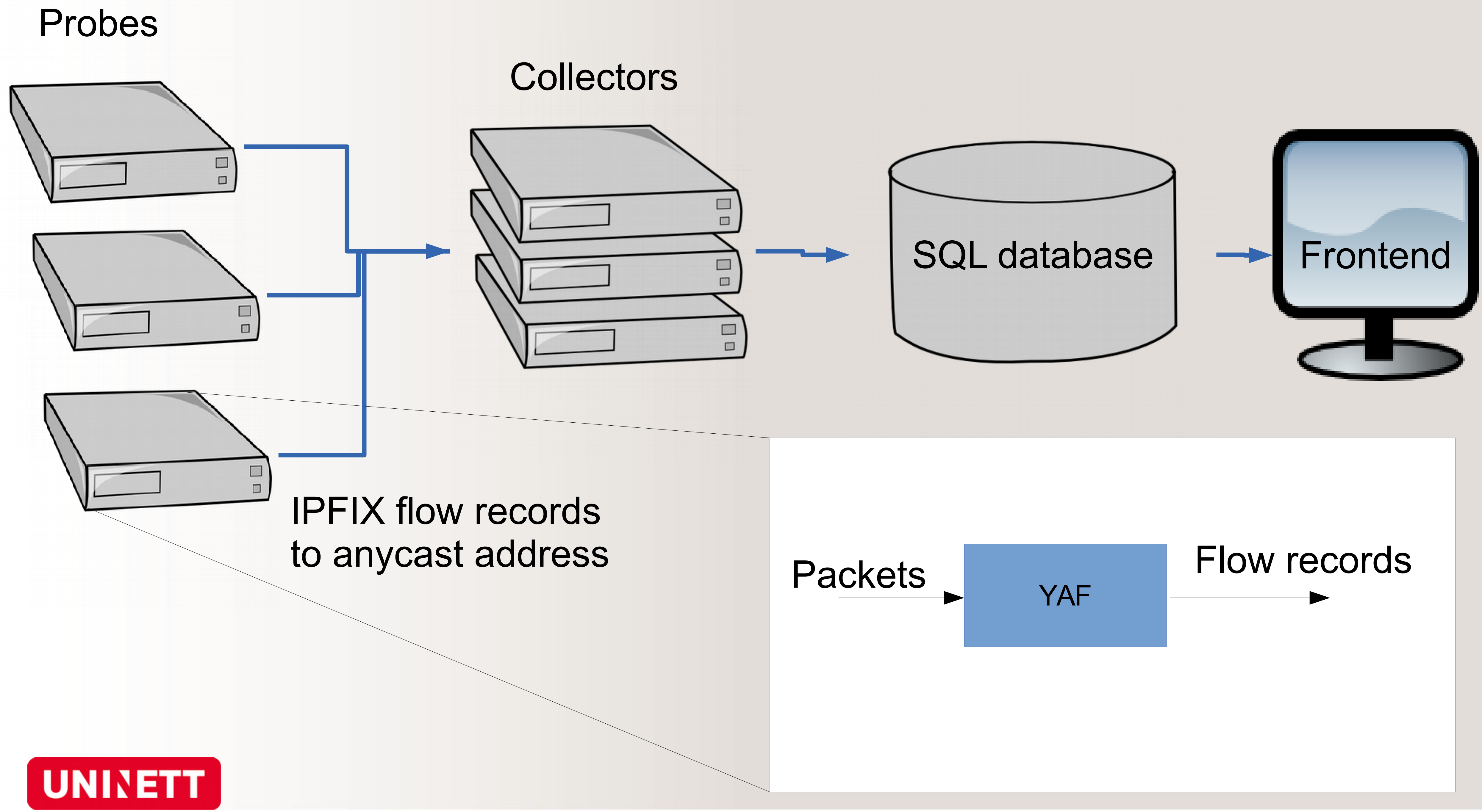
Test throughput! | 10 Mbit/s | to ntnu-mp.ntnu.no

| What | Mbit/s | Transfer Time | Throughput bar |
|---|---|---|---|
| Download | 7.020 | 2.918 s | |
| - average | 7.500 | 2.811 s | |
| Upload | 5.210 | 3.931 s | |
| - average | 4.700 | 4.383 s | |

Did 4 tests with average **Web Response Time** 0.106 secs.

UNINETT

# Original Appflow architecture

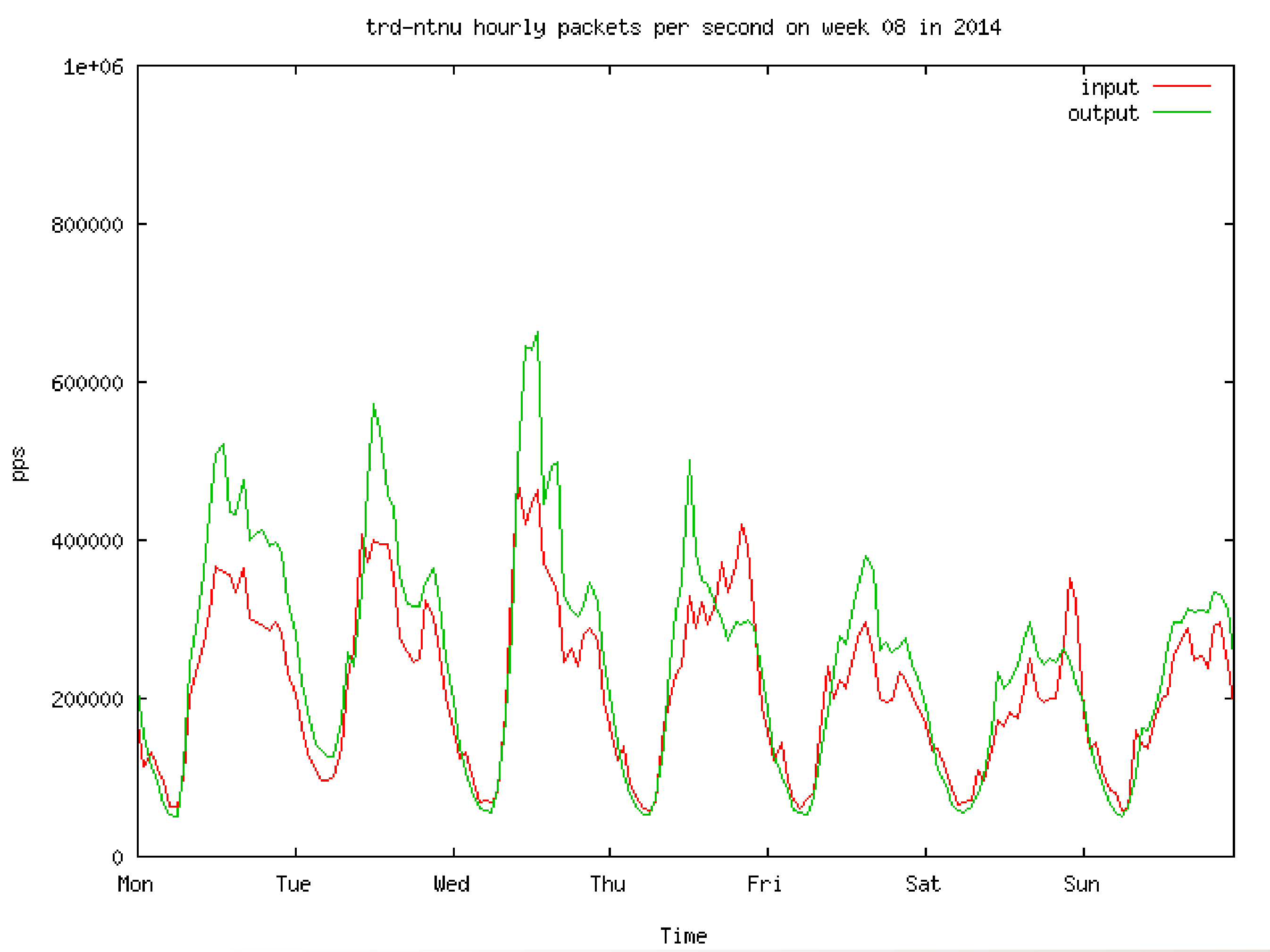


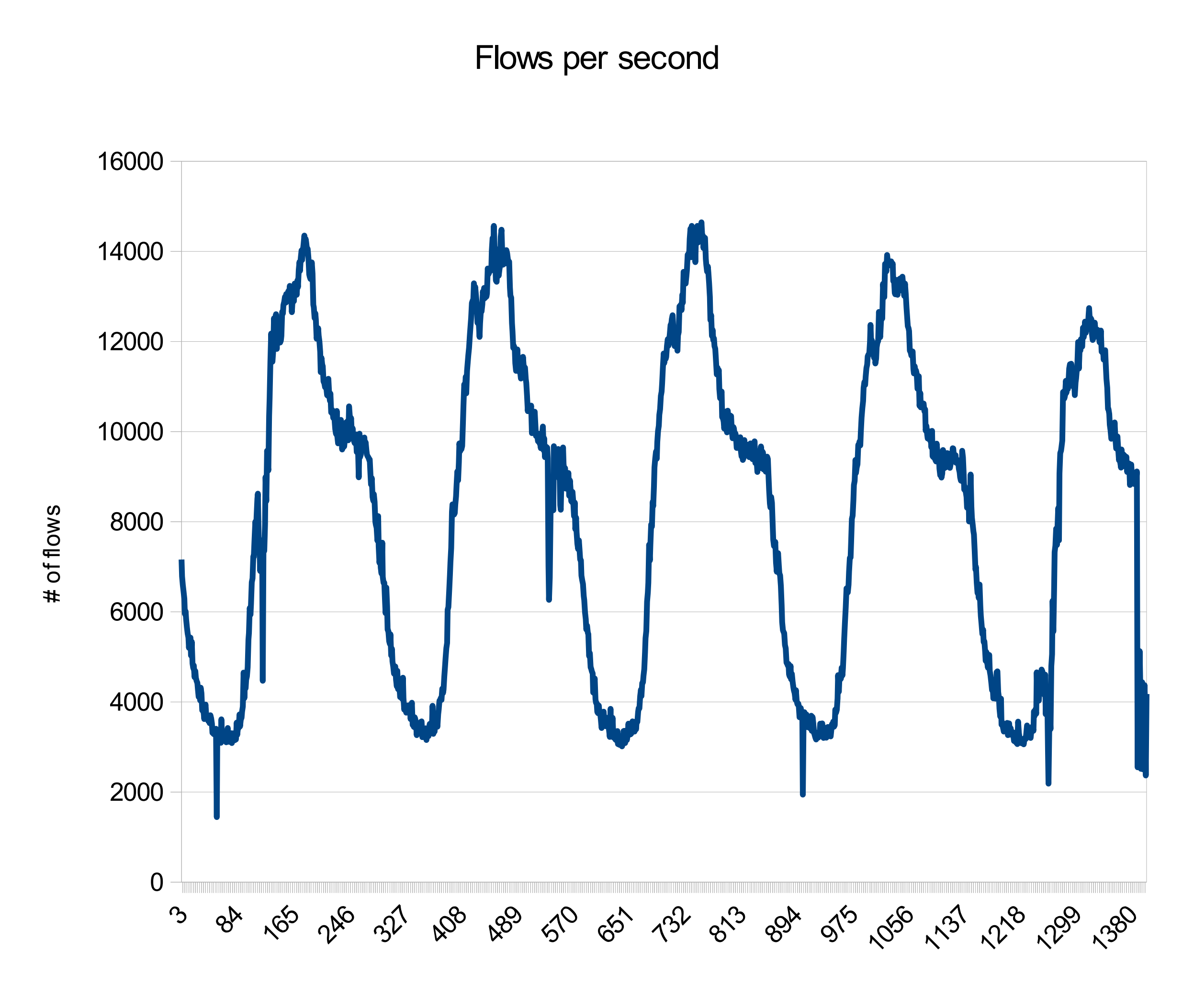
UNINETT

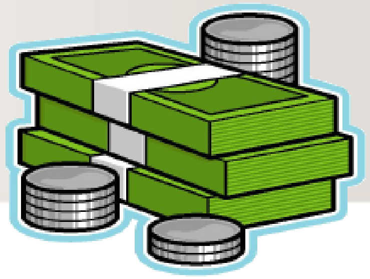
# 10Gbps challenges

Theoretical packet rate: 14.88 million pps

Number of flows

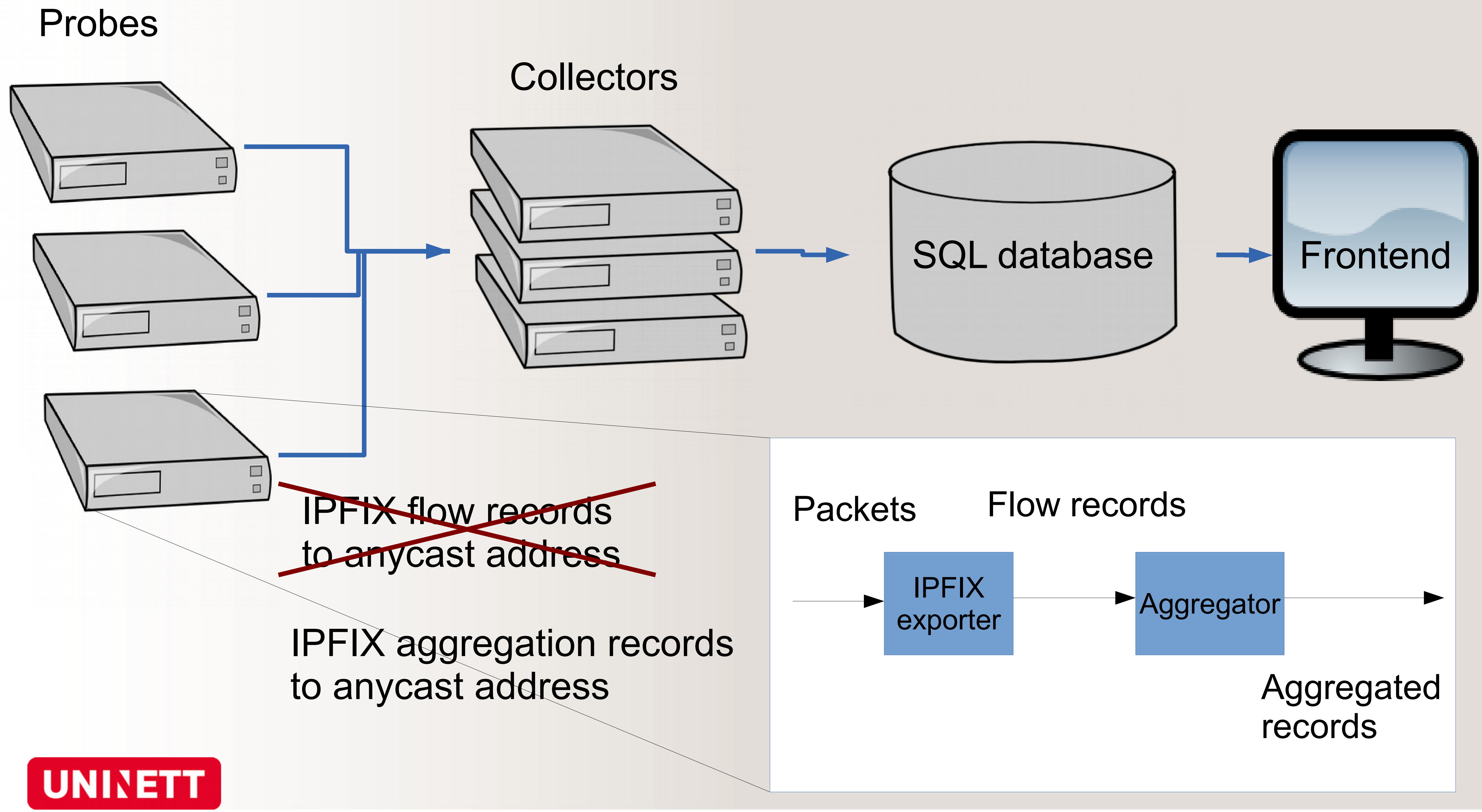






UNINETT

# New Appflow architecture



Probes

Collectors

SQL database

Frontend

~~IPFIX flow records to anycast address~~

IPFIX aggregation records to anycast address

Packets

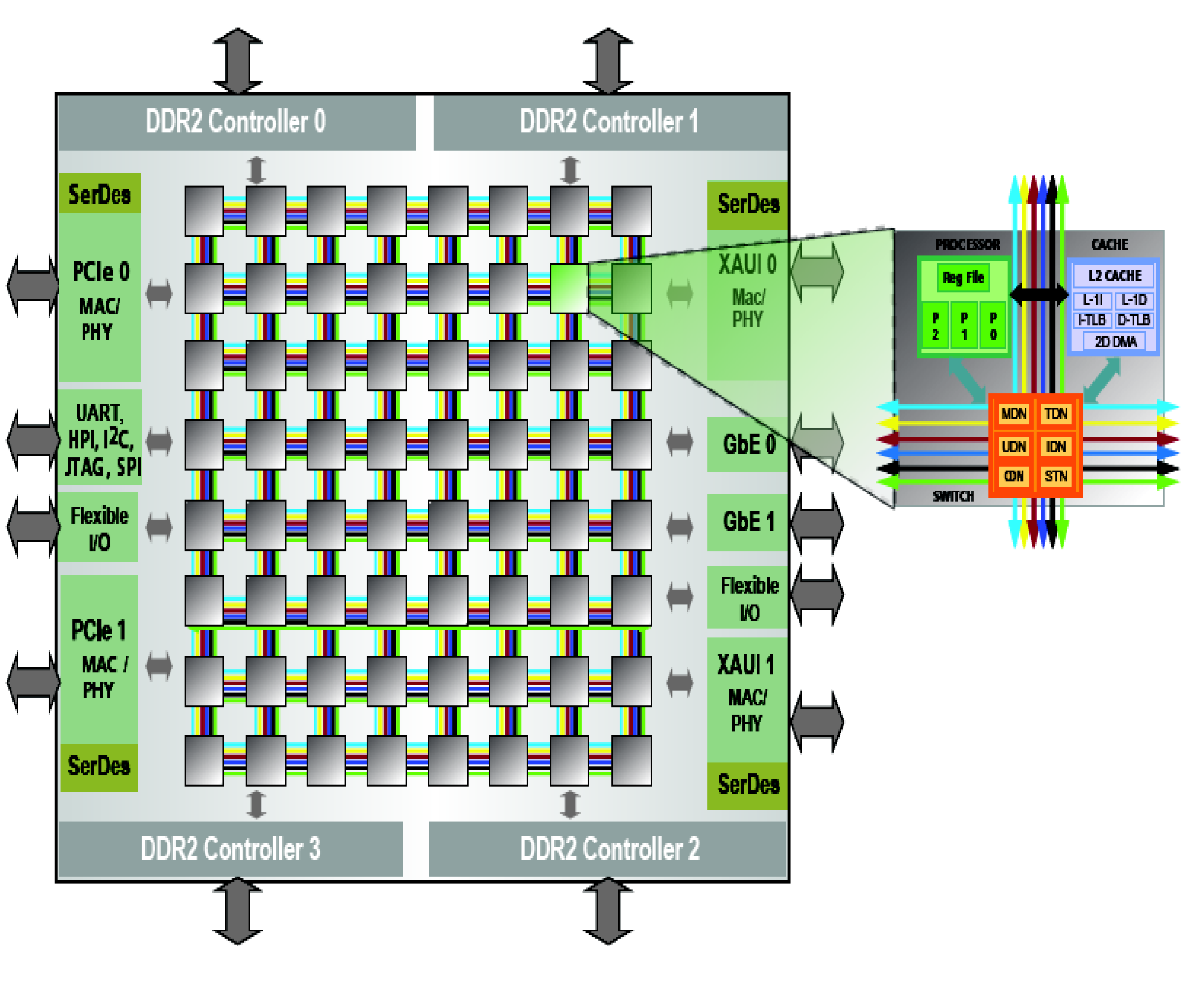
Flow records

IPFIX exporter

Aggregator

Aggregated records

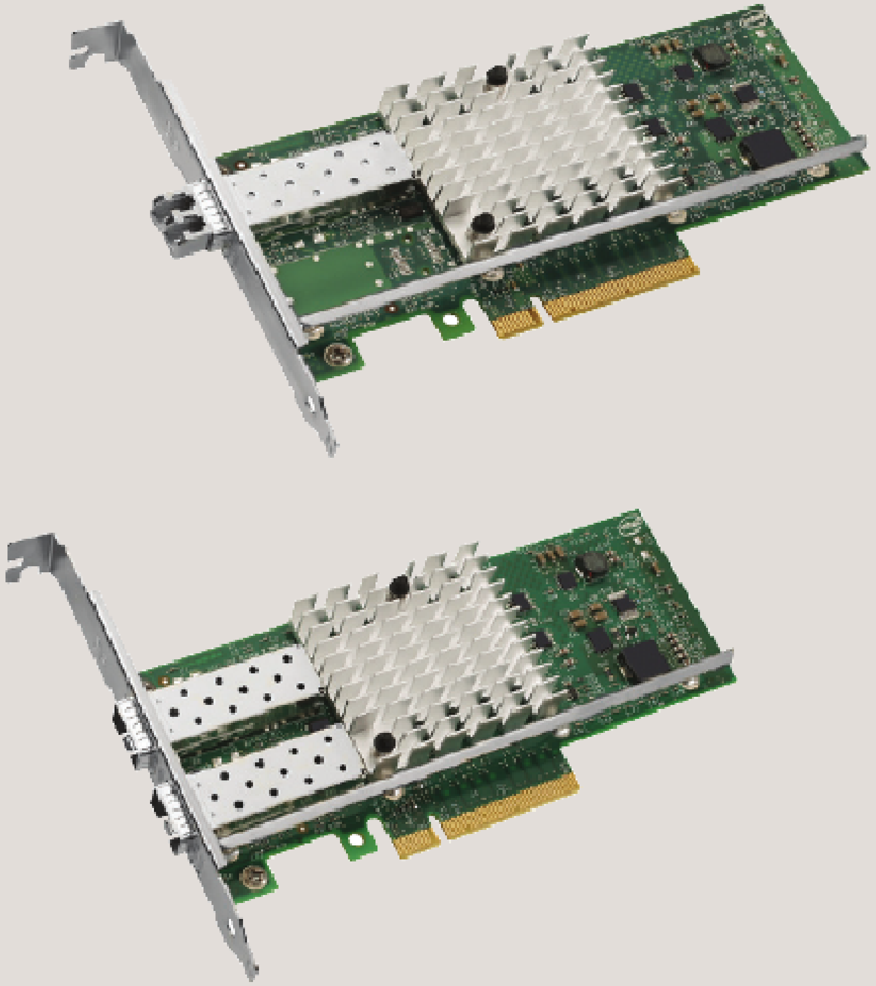
# TILEempower



- Based on TILERA cpu
  - Up to 72 cores
- Pros
  - Good performance
  - Special instructions for packet processing
  - Very good documentation
  - DPI library
- Cons
  - Difficult to program
  - Price

# Intel X520 family of NICs

- Designed for virtualization
- Support multi-core processors
  - Hardware based load balancing
- DMA transfer of captured packets
- Hardware counters
- Supports both 1 and 10 Gbps

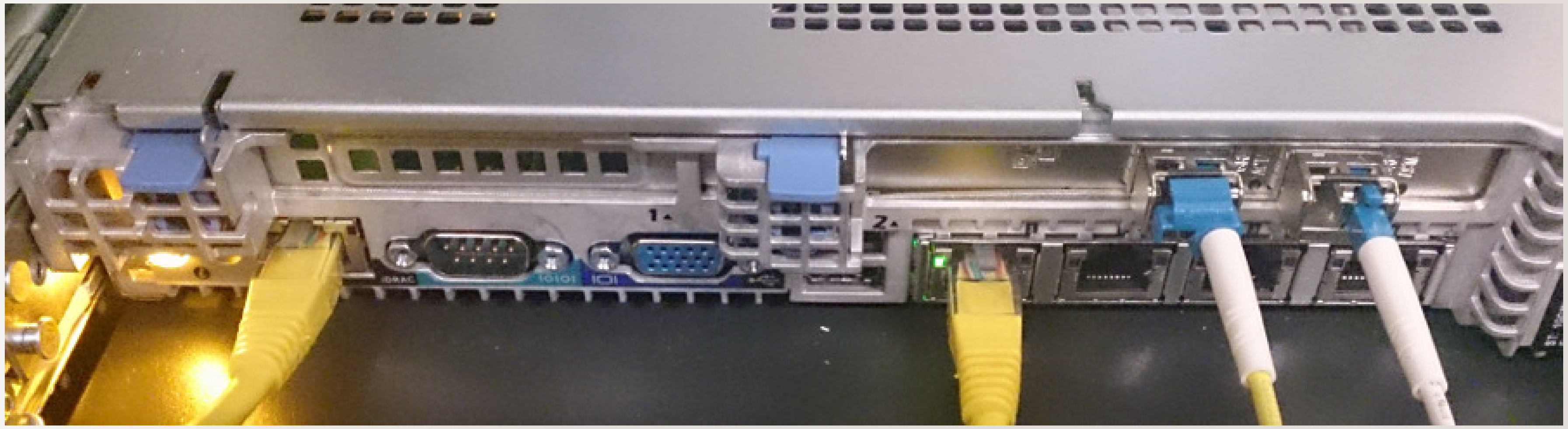


# Drivers for Intel X520

- Standard drivers not very good for passive monitoring
  - Too many interrupts per second
- Packet I/O Engine
  - No longer maintained
  - http://shader.kaist.edu/packetshader/io_engine/
- netmap - a novel framework for fast packet I/O
  - Originally developed for FreeBSD
  - Unstable port to Linux
  - http://info.iet.unipi.it/~luigi/netmap/
- PF_RING with DNA
  - Stable and well maintained
  - Multiple applications can access same buffer
  - Not GPL, but free for academic use
  - http://www.ntop.org/products/pf_ring/dna/

# Server hardware

- Dell PowerEdge R620
- CPU: Intel Xeon E5-2690, 2.9GHz, 8 cores , hyper-threading
  - Support for second CPU
- 32GB 1600MHz RDIMM
- Intel X520DP
  - Two ports with pluggable SFP+



# Packet capture performance

64 bytes packet size, two ports, one core

| Gbps | Mpps | Cpu load (%) | Packet drop (%) |
|---|---|---|---|
| 0.7 | 1 | 1 | 0 |
| 3.3 | 5 | 4 | 0 |
| 6.7 | 10 | 7 | 0 |
| 10.2 | 15 | 13 | 0 |
| 13.9 | 20 | 18 | 0 |
| 16.8 | 25 | 23 | 0 |
| 20 | 29.8 | 31 | 3.2 |

Realistic packet size distribution, two ports, 8 cores for each port

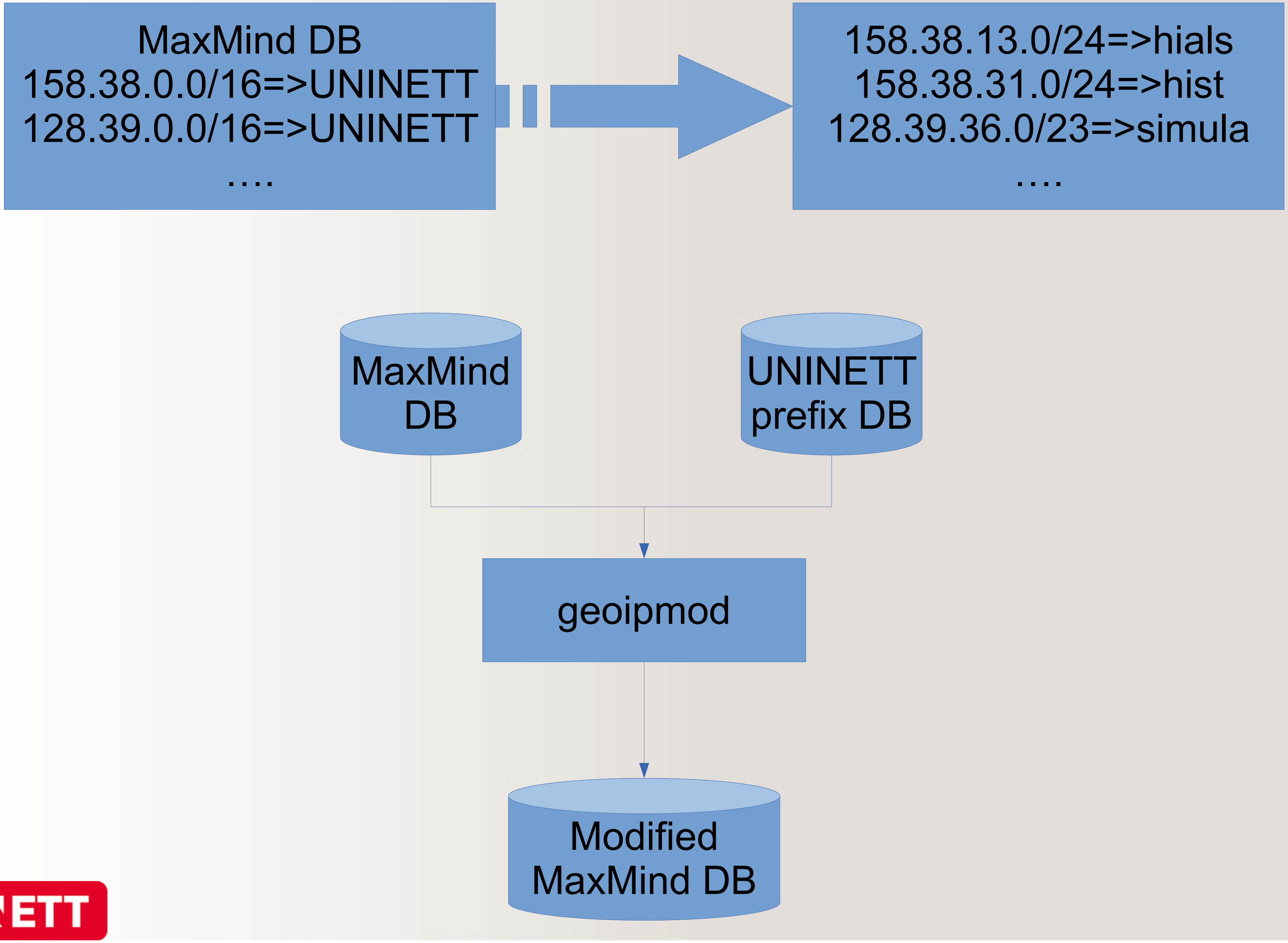
| Gbps | Mpps | Cpu Load(%) | Packet drop (%) |
|---|---|---|---|
| 17.3 | 5 | 7 | 0 |
| 20 | 6.5 | 9 | 0 |

UNINETT

# nProbe

- An Extensible NetFlow v5/v9/IPFIX GPL Probe for IPv4/v6
- http://www.ntop.org/products/nprobe/
- Good performance
- Well maintained
- Large user base
- Multi-threaded
  - But recommends running multiple single-thread instances
- IP tagging
  - AS numbers, countries
  - MaxMind: http://dev.maxmind.com/geoip/legacy/geolite/
- Support plugins
  - HTTP, DNS, BGP, SIP/RTP

# Adding IP prefix information



MaxMind DB
158.38.0.0/16=>UNINETT
128.39.0.0/16=>UNINETT
....

158.38.13.0/24=>hials
158.38.31.0/24=>hist
128.39.36.0/23=>simula
....

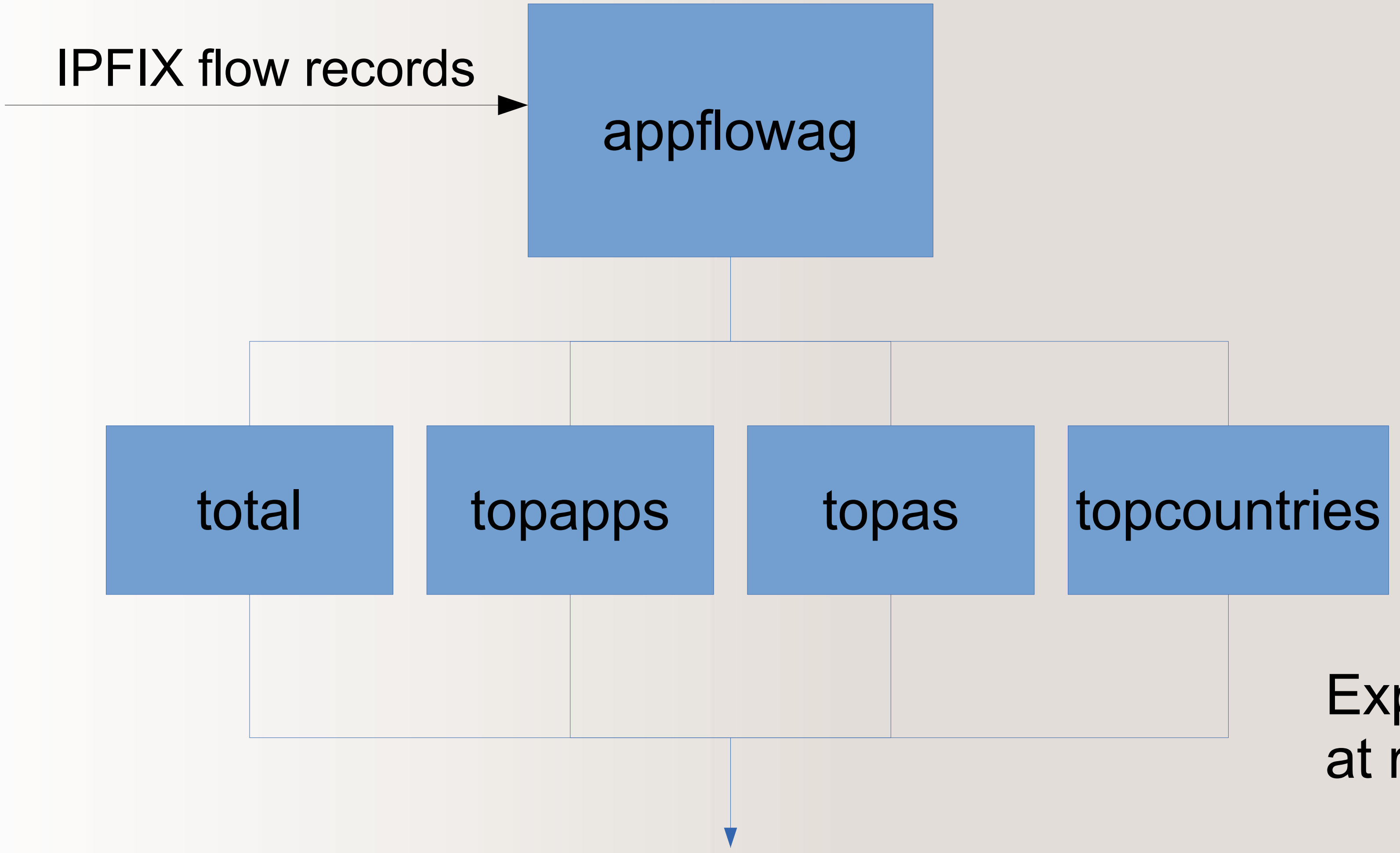
MaxMind DB

UNINETT prefix DB

geoipmod

Modified MaxMind DB

# Appflowag



IPFIX flow records

appflowag

total

topapps

topas

topcountries

Export records
at regular intervals

IPFIX aggregation records

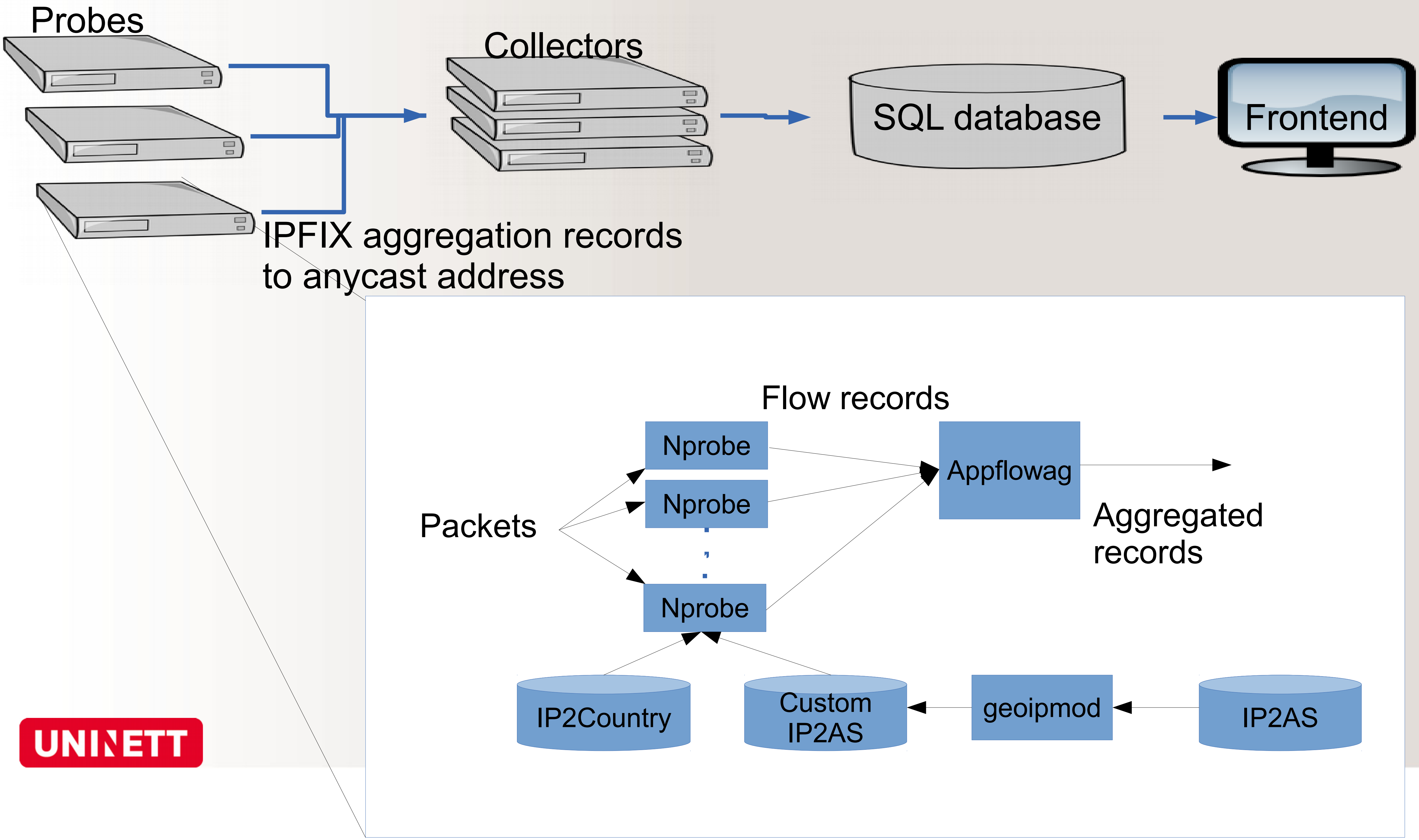
# Appflowag IPFIX records

- Total traffic
  - flowStartMilliseconds
  - flowEndMilliseconds
  - octetDeltaCount
  - packetDeltaCount
  - deltaFlowCount
  - ipVersion

- Top source AS number
  - flowStartMilliseconds
  - flowEndMilliseconds
  - octetTotalCount
  - packetTotalCount
  - deltaFlowCount
  - bgpSourceAsNumber
  - l7_proto
  - ipVersion

```
261, 1378883700000, 1378883999999, 55430087856, 51440429, 792359, 4
261, 1378883700000, 1378883999999, 3666166884, 3127366, 73943, 6
259, 1378883700000, 1378883999999, 11979801504, 9847245, 29923, 224, 0, 4
259, 1378883700000, 1378883999999, 3600748945, 2413758, 9, 42307, 0, 4
```

# Final Appflow architecture



Probes

Collectors

SQL database

Frontend

IPFIX aggregation records
to anycast address

Flow records

Packets

Nprobe

Nprobe

Nprobe

Appflowag

Aggregated
records

IP2Country

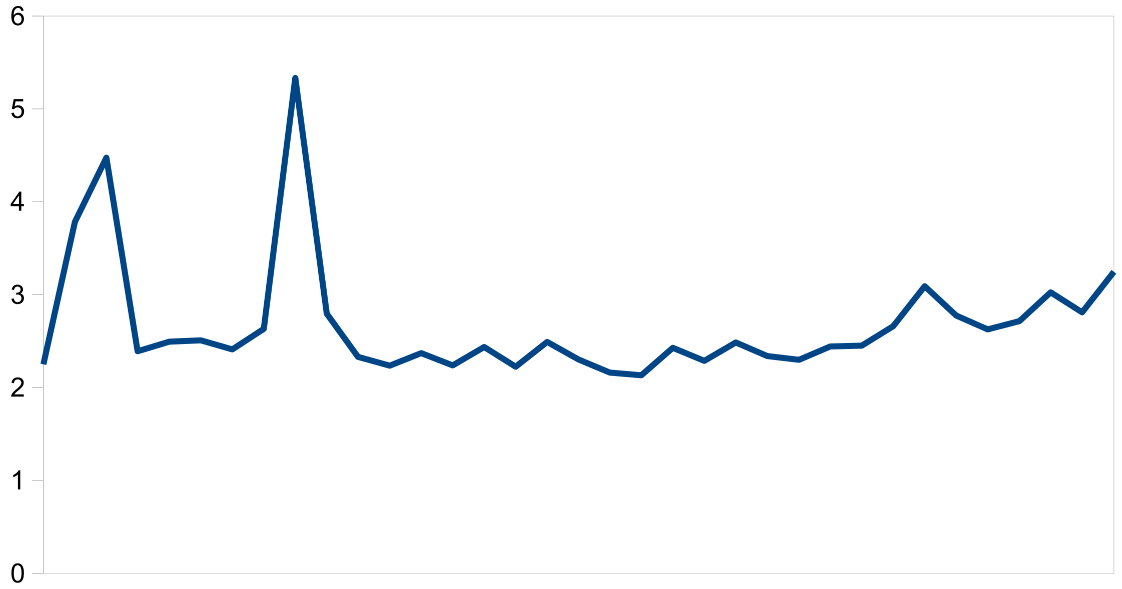
Custom
IP2AS

geoipmod

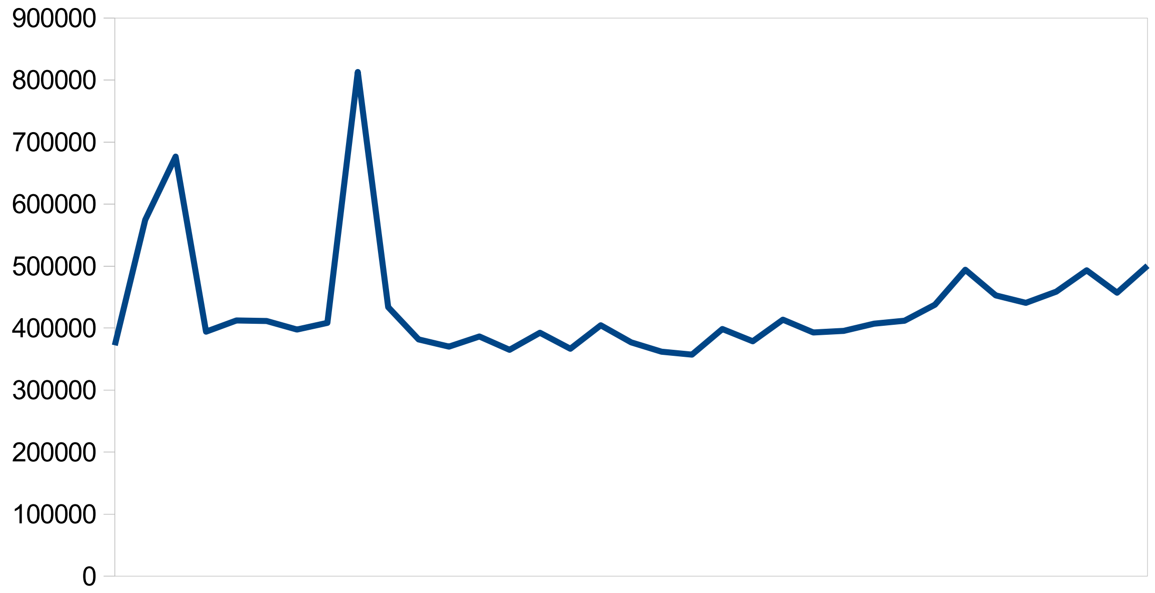
IP2AS

# Nprobe and Appflowag performance

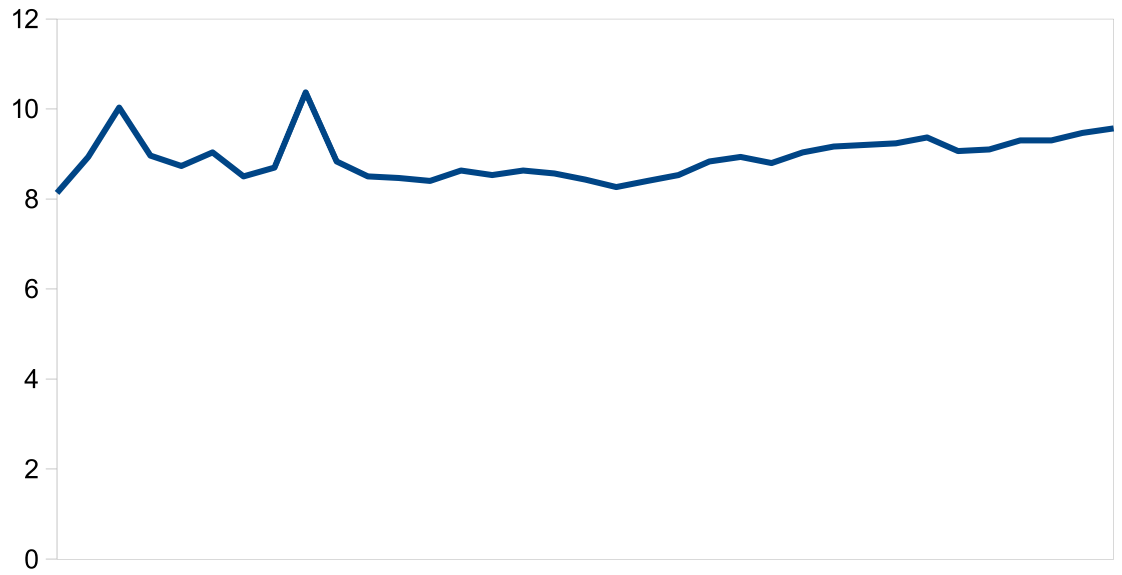
Gigabit per second



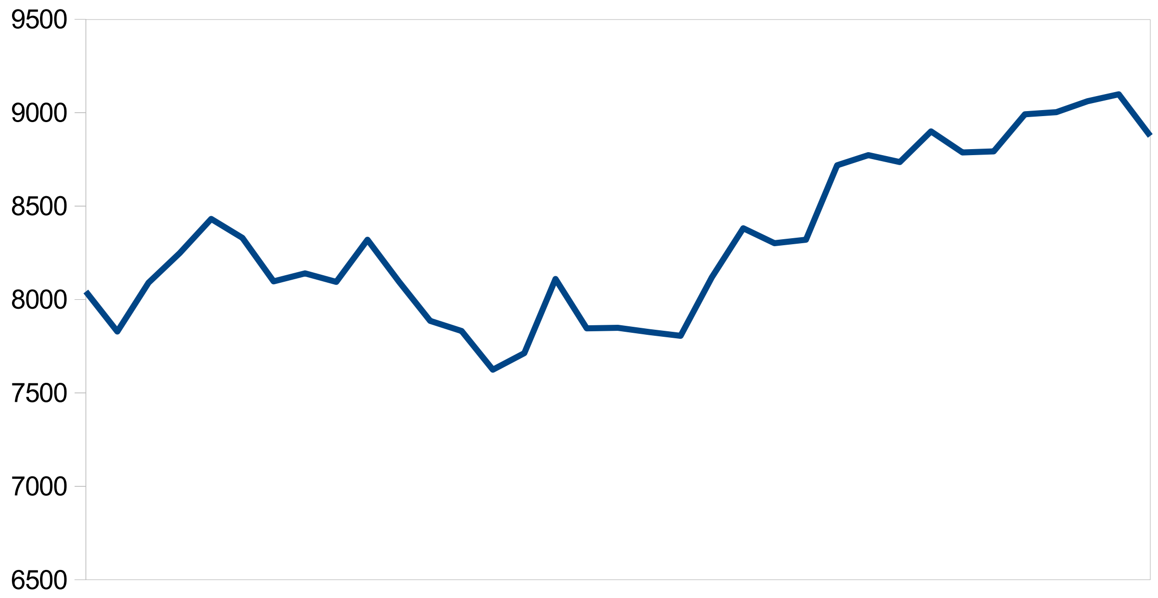
Packets per second



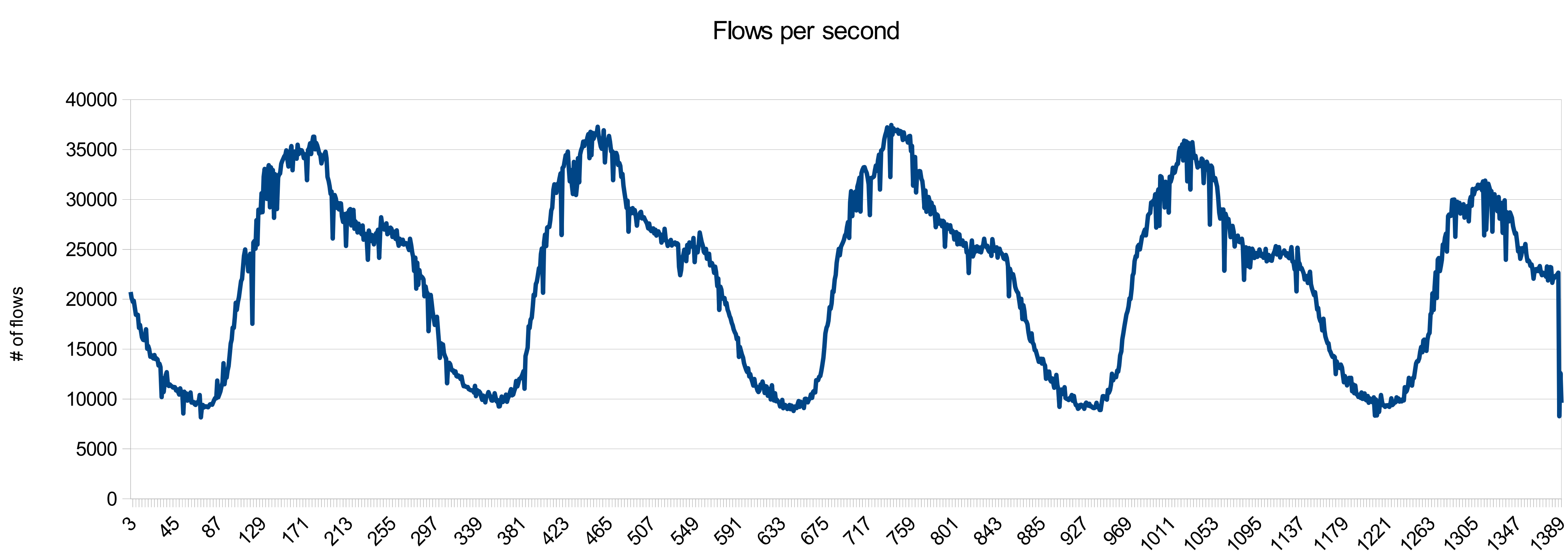
Total CPU usage
8 cores for nProbe, 1 for appflowag

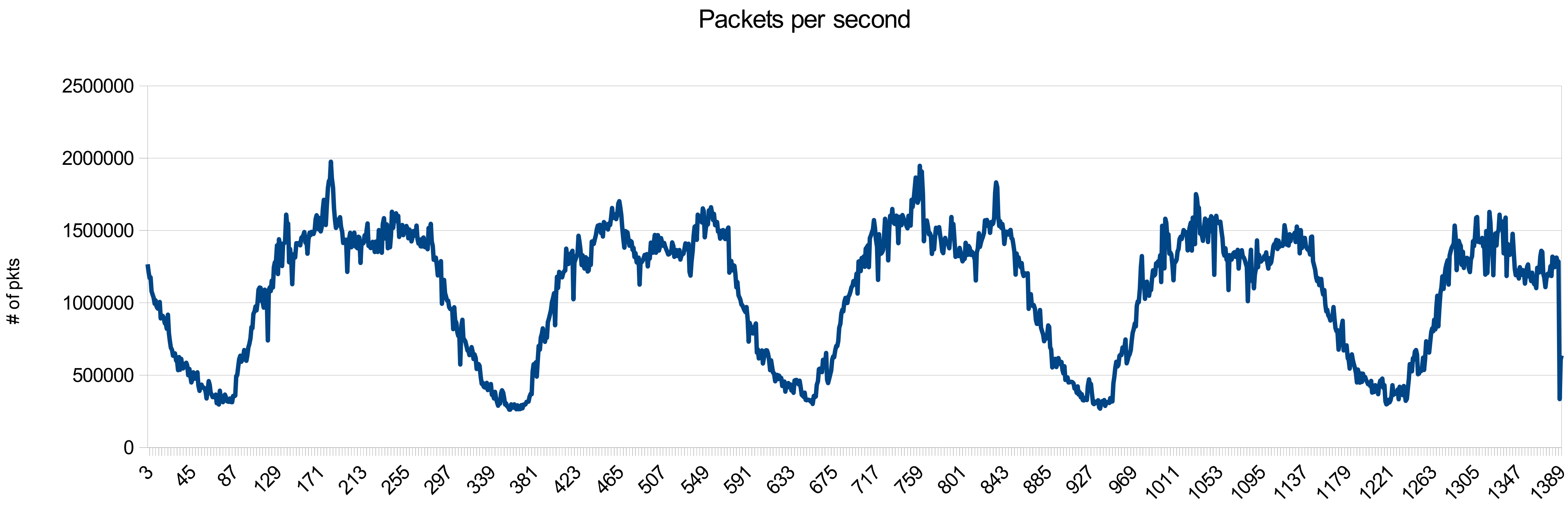


Flows per second



# Total processing (19 probes)





# Current status and future work

- 30 new monitoring probes being deployed
  - 19 in full production
- Appflow in full production
  - Want to improve unknown traffic
  - Customers wants to add their own prefixes to classify traffic
- Activate nProbe plugins
  - SIP/RTP, DNS
- Other QoS measurements
  - Packet reordering, jitter …
- Software will be released
  - http://software.uninett.no/