



Flow measurements from the packet-switched NREN PIONIER: technology and experience

Szymon Trocha

A.Binczewski, W.Procyk, M.Stroi_ski

Poznan Supercomputing and Networking Center
Poland

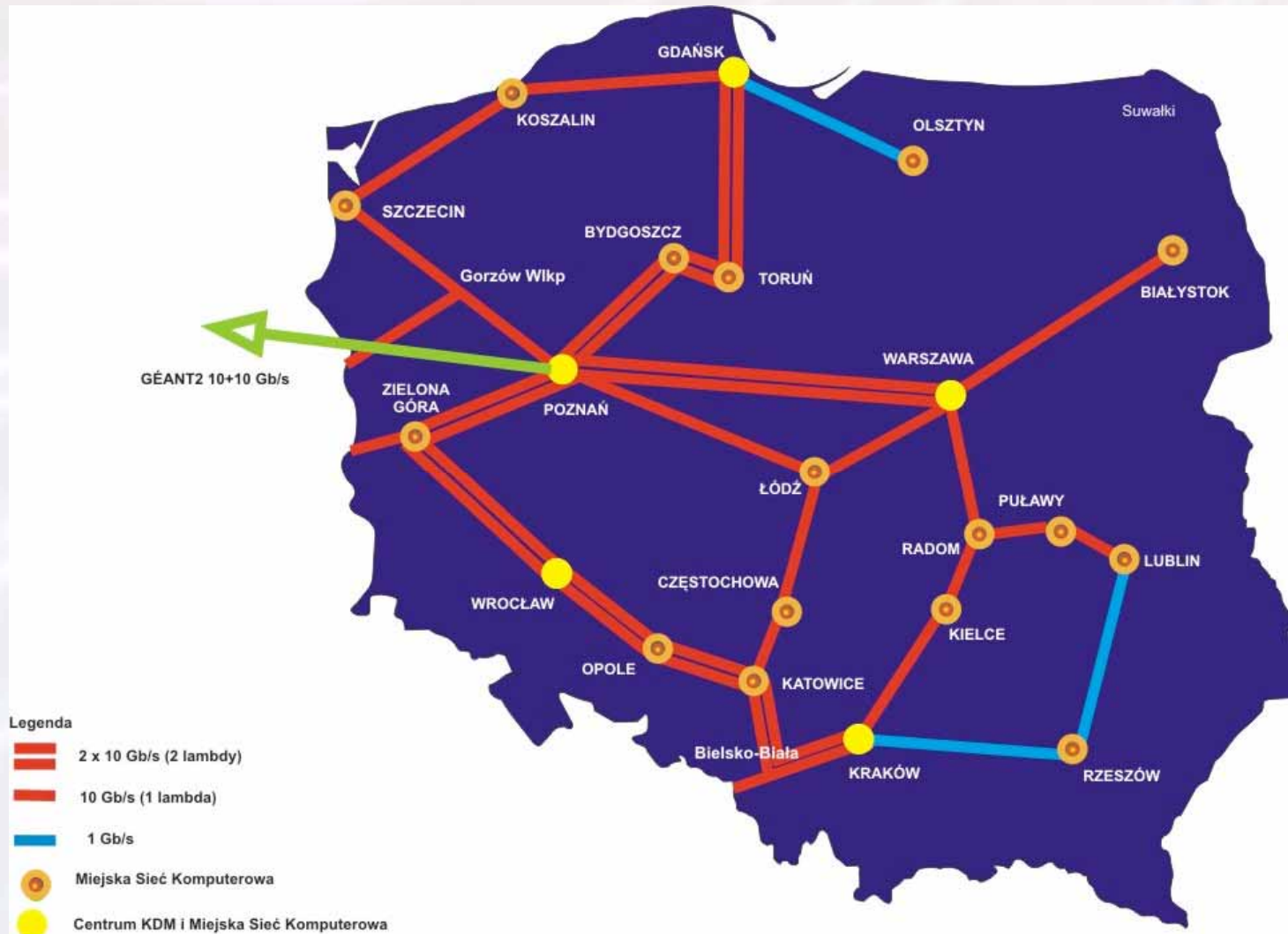
SANOG7, Mumbai, January 2006



Quick overview

- R&D institution
 - 10 years of experience
- PIONIER – Polish National Research & Education Network operator
 - 3400 km of our own fibres
- Metropolitan Area Network operator
- 24 / 7 / 365 NOC
- High Performance Computing Centre
- Active international projects participant

The first-class optical network



Evolving networks

- Changes in the computer networks during recent years
 - Increased capacity and popularity
 - New services – grids, videoconferencing, VoIP requiring bandwidth and quality
 - Growing meaning of quality assurance and SLA
- Focus on network monitoring and engineering

Motivations

- Need for monitoring Layer 2 for demanding users and dedicated connections
- For monitoring Layer 2 sFlow multivendor sampling technology can be used
- GPL based tools are used in many networks for monitoring purposes because of their low cost and flexibility
- sFlow is poorly supported by nearly all GPL tools

The environment

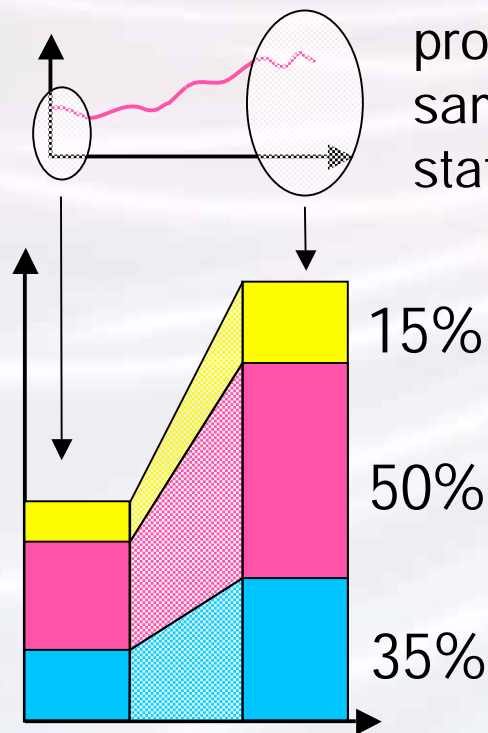
- Network devices
 - Extreme Networks BlackDiamond 6800 switch
 - sFlow enabled per interface
 - sampling rate set per whole device
- Applications
 - RRD tool
 - sflowtool
 - Home-made perl tool

Common problems

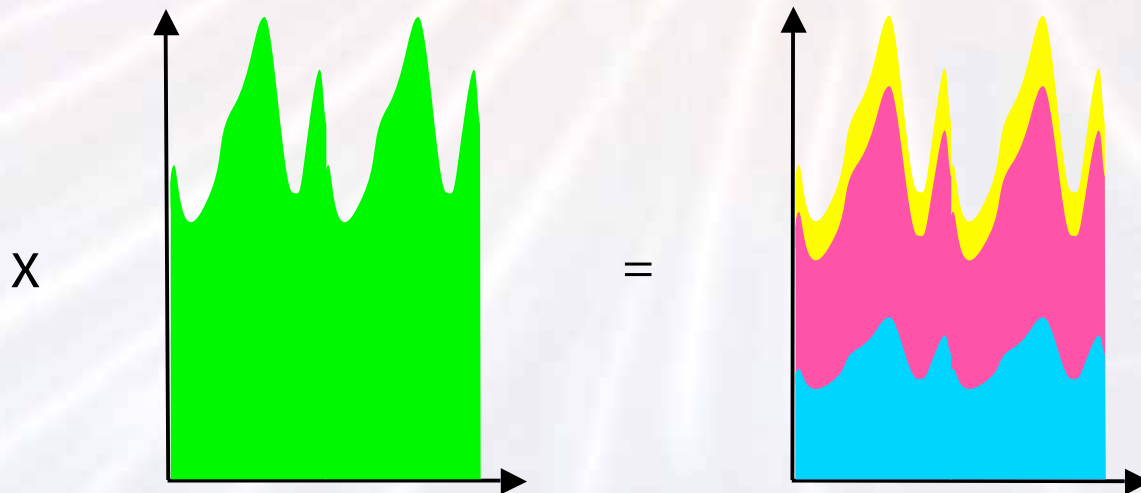
- Sampling rate set per whole device
 - Changing sFlow configuration needs measurement tool reconfiguration
 - We count every packet crossing the switch but sampled packets are taken only from sFlow-enabled interface
- Increasing traffic on non-sampled interfaces increases number of samples from sFlow-enabled interface
- Enabling sFlow on more interfaces decreases number of samples per interface (total number of samples is constant)
- Data loss for underused VLANs
 - Commonly used RRD tool introduce error while aggregating data
 - Misleading difference between MRTG and sFlow

Proposed solution

•It's observed that if traffic did not change, the proportions between different VLANs remained at the same level, even if we changed sampling rate or sFlow status on other interfaces.



- VLAN1
- VLAN2
- VLAN3



Algorithm

- Take relative values and apply 4-steps algorithm
 1. Collect data
 2. Check time
 3. Calculate ratio and update files
 4. Write data to temporary tables

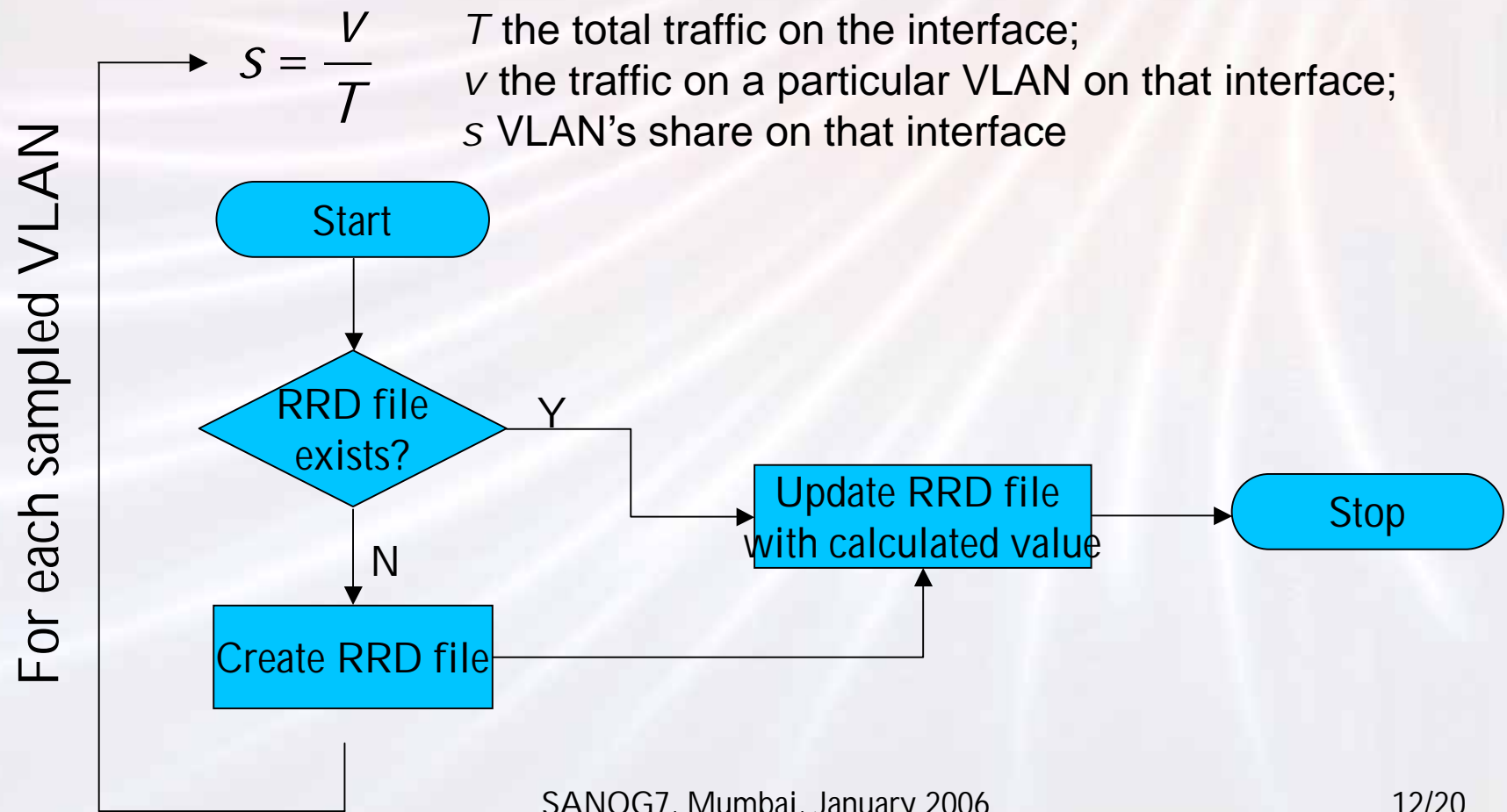
Step 1 – Collecting data

- Collect data samples and classify based on
 - Agent IP address
 - Input interface
 - Output interface
 - Sampled packet size
 - Decoded VLAN

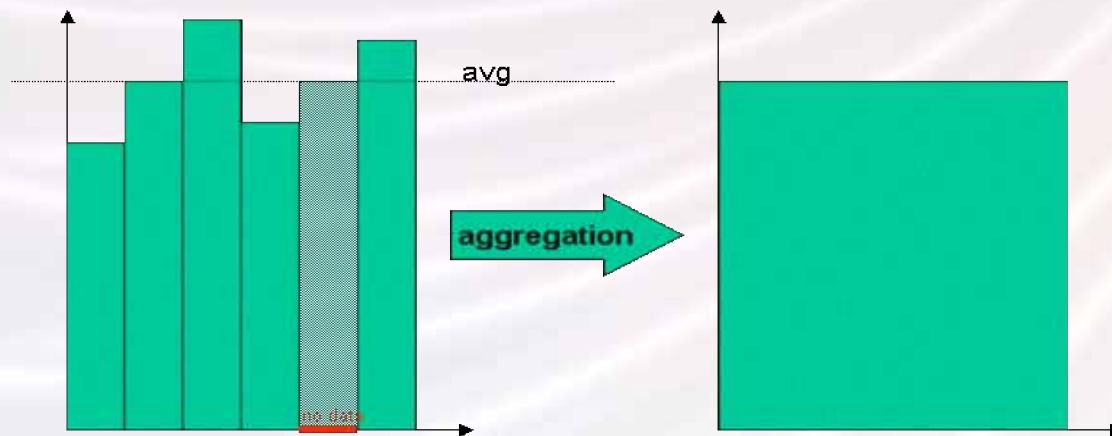
Step 2 – Timing is important

- Check current time
- Compare last RRD file update time with the measurement period
 - Measurement data is kept in RRD files
 - `<agentIP>-<interface>-<VLANnumber>.rrd`
 - Measurement period as configured in RRD files

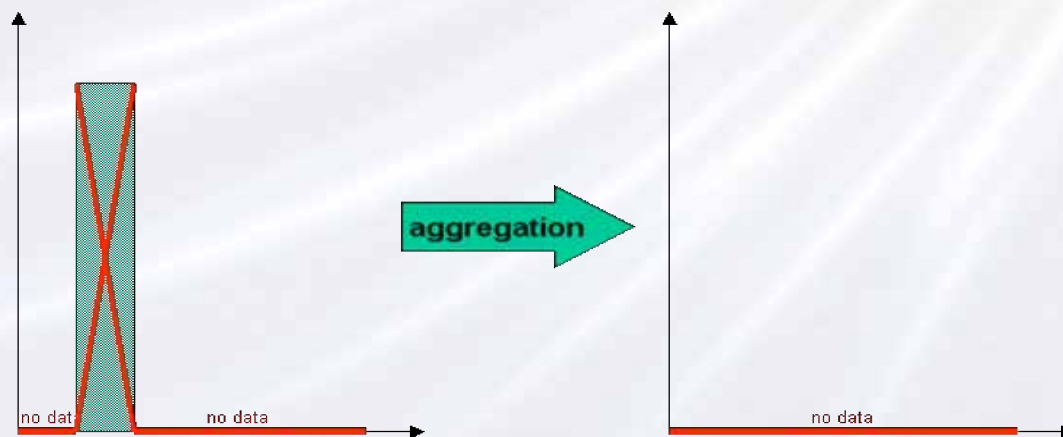
Step 3 – Updating files



The rules for RRD aggregation



Sum of all shares on the interface exceeds 100%



Sum of all shares on the interface is not 100%

Step 3 – Updating files (mod.)

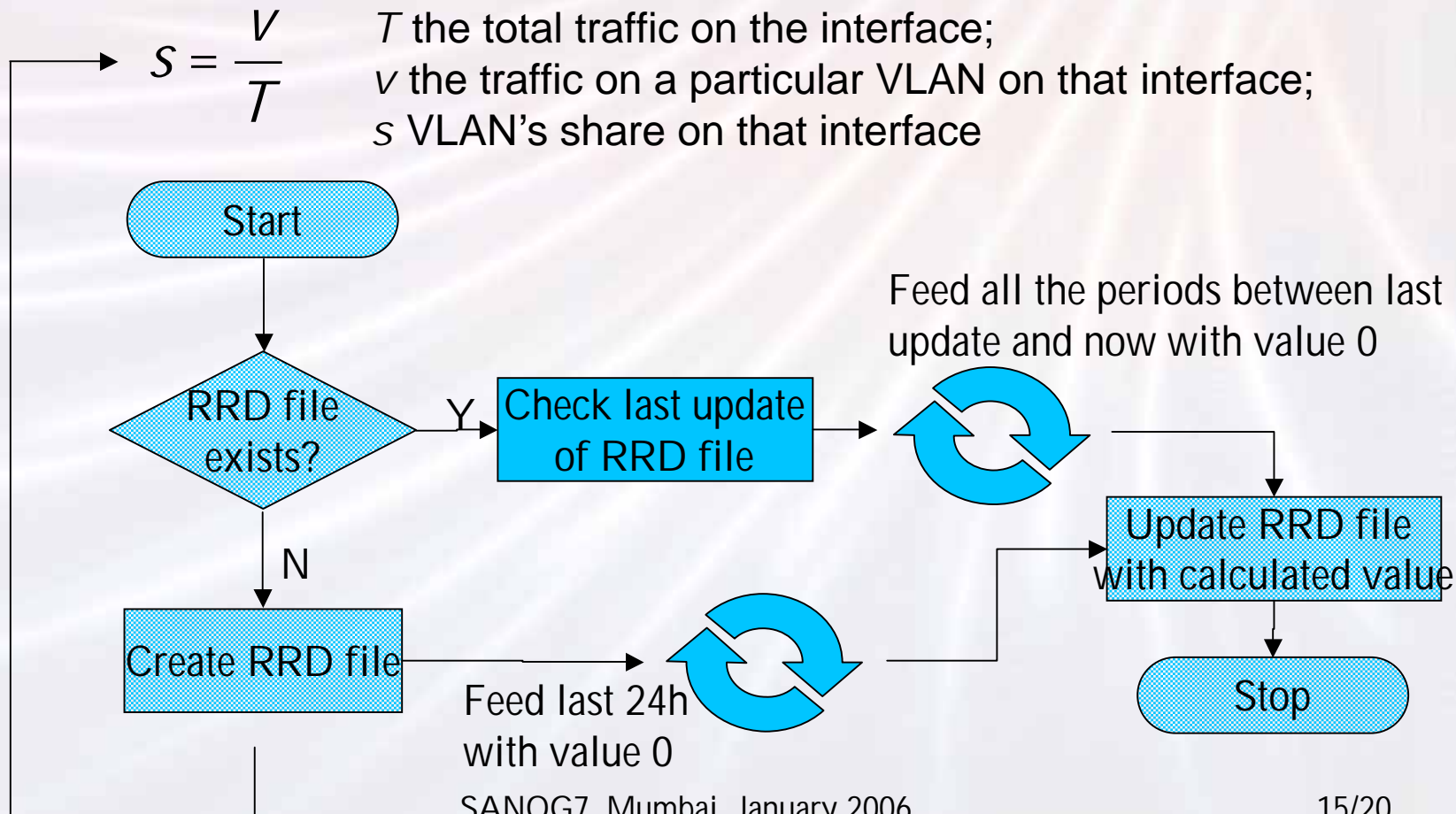
- If time elapsed
 - Update appropriate RRD file with VLAN's share
 - If last update of file is older than the end of previous period
 - Complete missing values with 0
 - Helps to keep 100% interface volume
- If not
 - Go to step 4

Step 3 (block schema)

$$s = \frac{v}{T}$$

T the total traffic on the interface;
 v the traffic on a particular VLAN on that interface;
 s VLAN's share on that interface

For each sampled VLAN



Step 4 – Data processing

- Data is written to temporary tables
 - Interface
 - VLANs

Advantages

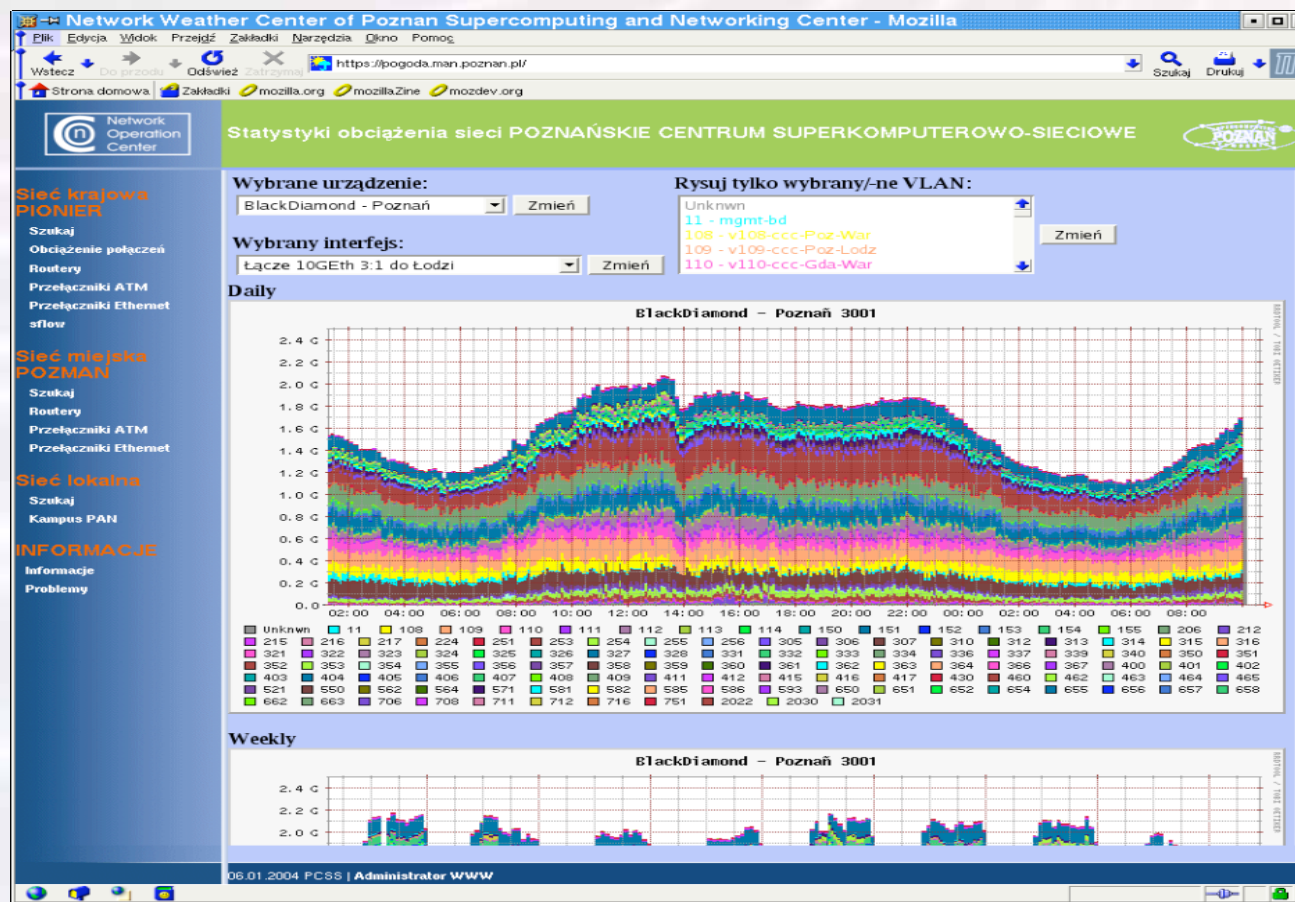
- Solves sFlow measurement distortion problems
- Makes sure one has exact statistics
- Can modify the sampling rate without changes in measurement application
- Uses open-source tool

Implementation

- Network
 - 16 BlackDiamonds
 - 278 VLANs
- Workstation
 - P III 500 (CPU usage ~5%)
 - 512 MB RAM (~10MB for scripts)
- Data
 - 5424 files (inc. 2349 updated during <24h)
 - 350 MB



Friendly frontend



User interface

- Written entirely in PHP
- Visualisation of all or only selected VLANs on particular port
- Dependencies between RRD files from MRTG and sFlow measurements stored in the database
- Used mostly by PIONIER NOC