A multi-task adaptive monitoring system combining different sampling primitives.

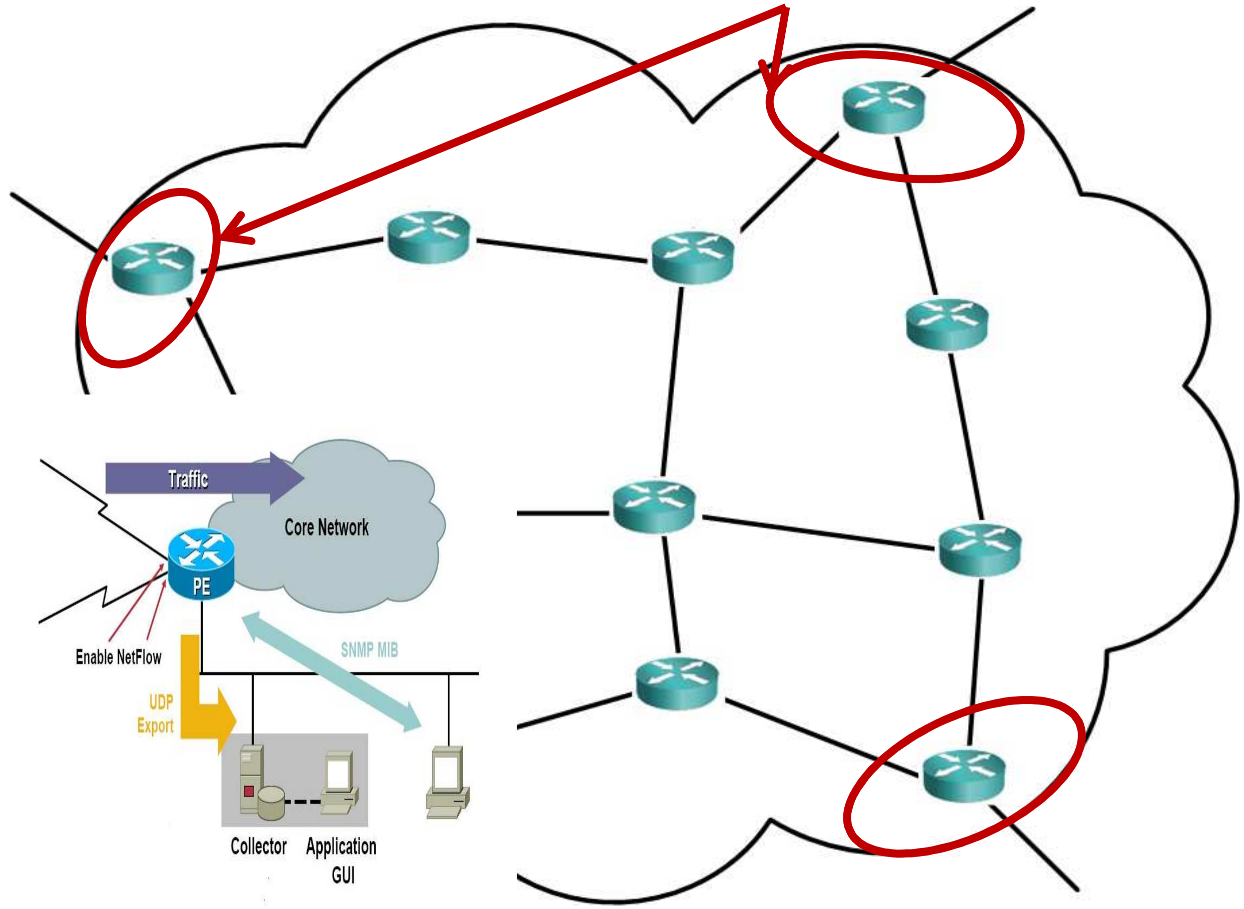
Imed LASSOUED (Imed.Lassoued@.inria.fr)
Chadi BARAKAT (Chadi.Barakat@.inria.fr)

Introduction
General context

Existing systems use sampling primitives **<u>separately</u>** and configure them **<u>statically</u>** to achieve some performance objective

Several monitoring techniques are deployed in network routers

- Low sampling rates (between 0.01 and 0.001)
- Reduce accuracy
- Entails loss of information
- Bias against small flows
- Problem in configuring sampling rates
- Resources usage problem.
- Sample and export unnecessary data



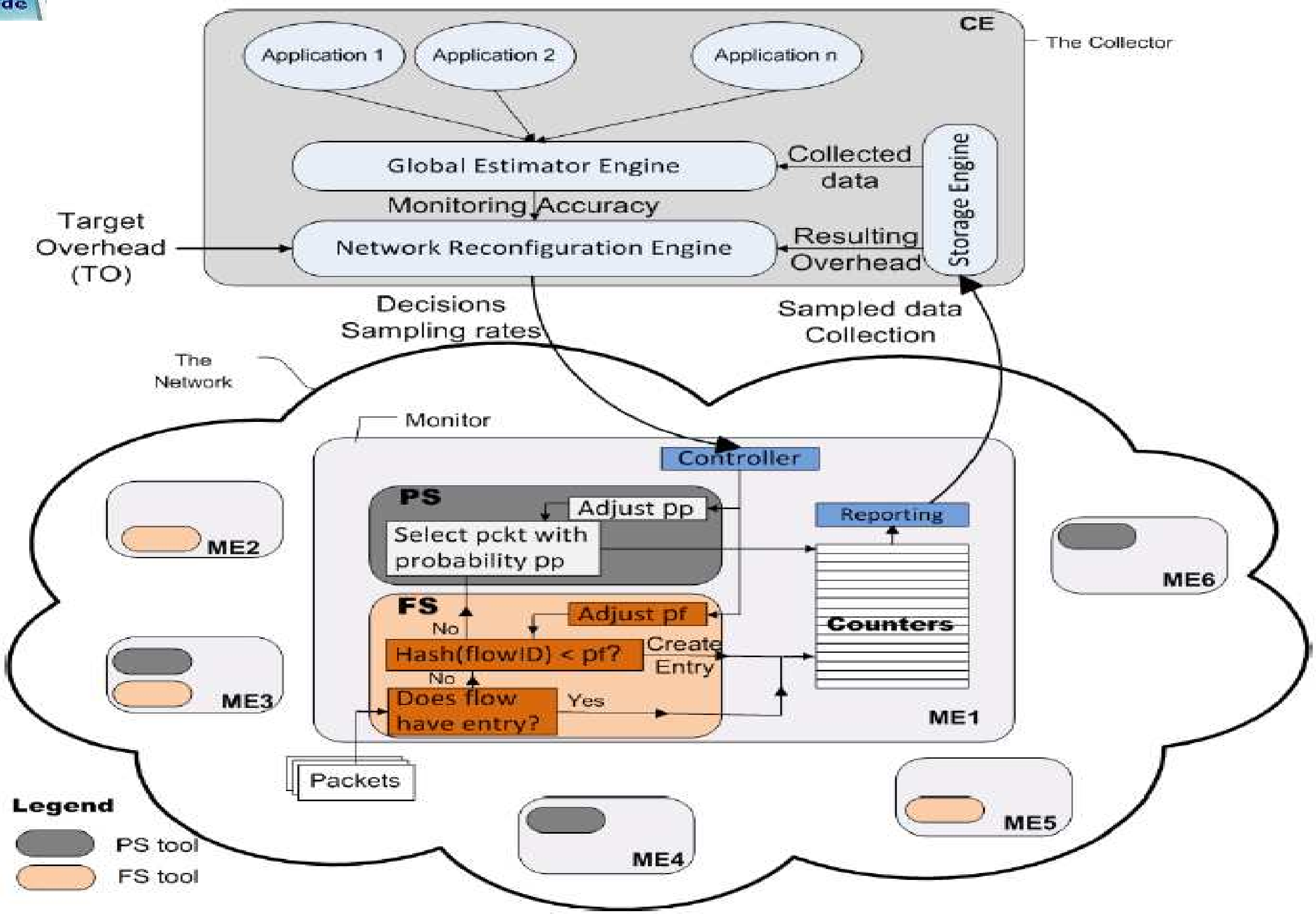
Introduction
Motivation

- Multiplying the monitoring points inside the network.
- Coupling their observations in order to improve the global accuracy
- Sharing responsibilities between the different monitors
- Automatically and periodically reconfiguring sampling rates as a function of monitoring task requirement and network conditions

Challenges

- Instead of placing monitors, How to coordinate responsibilities across the different monitors and how to share resources between the different sampling primitives in order to improve the global accuracy while respecting resource consumption constraints.
- How to deal with multiple monitoring objectives and how to combine independent measurements collected across the network using different sampling primitives and different monitoring tools.
- How to adapt to variations in the monitored traffic and in network conditions

System Architecture



Problem formulation

- Given a set of measurement task T and a monitoring constraint TO (Maximum exported flow records in the entire network during a period t), find a method that:
  - sets the sampling rate of the different sampling primitives on all interfaces (some can be turned off).
  - guarantees optimal use of resources
    (in terms of processed packets and volume of collected data)
  - can adapt to changes in the traffic
- Find the optimal sampling rate vector that minimizes the global estimation error under the following

$$\mathcal{O} \leq \mathcal{TO}$$

$$p_k \leq SR_{max} \quad ; \quad \forall k \in \mathcal{R}$$

$$p_k \geq SR_{min} \quad ; \quad \forall k \in \mathcal{R}$$

Cognitive Engine
Global estimation module:

Objective:

- Investigate the different local and noisy estimations, of a given task Ti, to build a global and more reliable estimation.
- Combine measurements of the different deployed sampling primitives
- Minimize the amplitude of measurement error.

$$\hat{T}_i = \sum_{s \in \mathcal{S}} \lambda_s \hat{T}_i^s \quad \text{with} \quad \lambda_s = \frac{\frac{1}{Var(\hat{T}_i^s)}}{\sum_{l \in \mathcal{S}} \frac{1}{Var(\hat{T}_i^l)}}.$$

Ti can be the flow size estimation task, flow counting or heavy hitter detection

Cognitive Engine

Network configuration method

Optimization Problem $U = \sum_i \gamma_i Var(\hat{T}_i),$

Under the constraints:

$$\mathcal{O} \leq \mathcal{TO}$$

$$p_k \leq SR_{max} \quad ; \quad \forall k \in \mathcal{R}$$

$$p_k \geq SR_{min} \quad ; \quad \forall k \in \mathcal{R}$$

Lagrange problem:

$$L = U + \delta(O - TO) + \sum_k a_k(p_k - SR_{max}) + \sum_k b_k(SR_{min} - p_k).$$

Optimization method

- **Newton Method**
  - **Overhead Prediction**

Optimization method

Overhead prediction given a time scale.

Calculate the smoothing factor given the time scale.

$$n \leftarrow \frac{\tau}{d};\ \delta \leftarrow 2/(n+1)\ ;$$

For each monitor

Estimate the number flows and the mean size of 5-tuple flows during the last measurement period

Use these last observations to predict the mean size and the number of flows using the EWMA method

$$\overline{N}_k^{(t+1)} \leftarrow \alpha \hat{N}_k^{(t)} + (1-\alpha)\overline{N}_k^{(t)}$$

$$\overline{S}_k^{(t+1)} \leftarrow \alpha \hat{S}_k^{(t)} + (1-\alpha)\overline{S}_k^{(t)}$$

Derive a local prediction of the Overhead

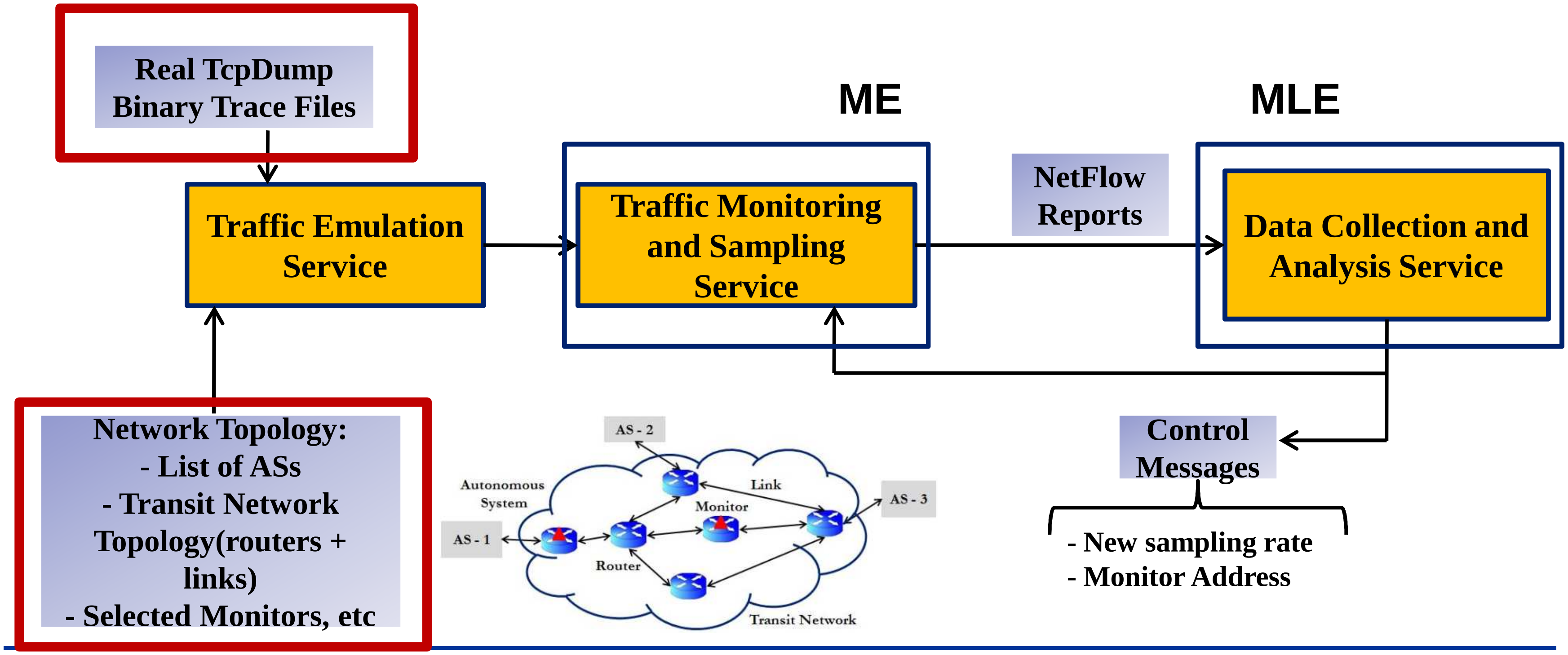
Derive the global prediction of the overhead

$$O_a \leftarrow \pi_a^f.\overline{N}_a + \pi_a^p.\overline{N}_a.\overline{\Sigma}_a\ ;$$

Experimental platform.

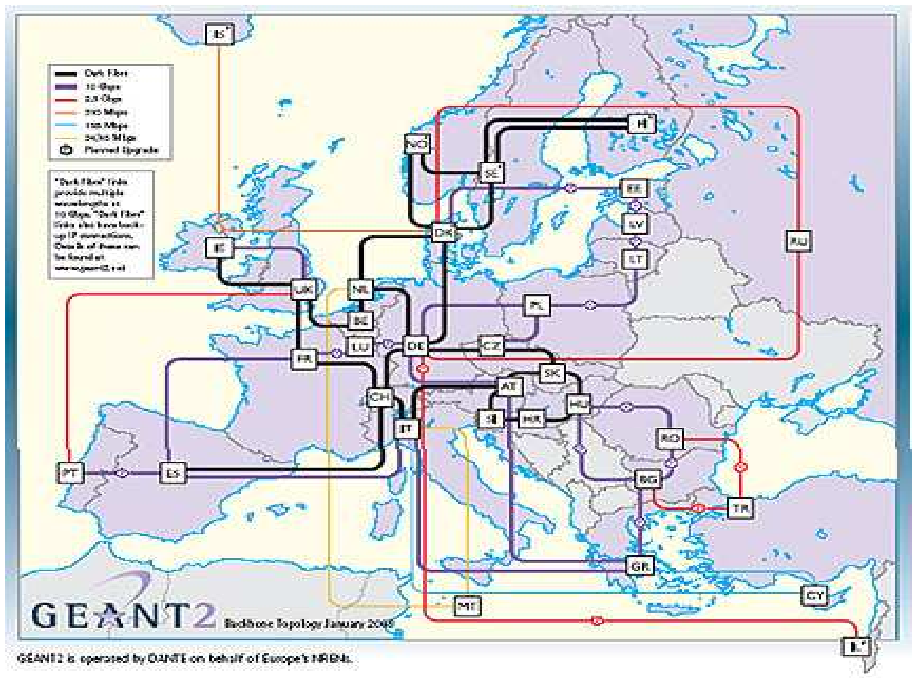
MonLab: http://planete.inria.fr/MonLab/

- Starting from a set of collected real traces (either from a single point during different periods of time or from different points) and a given network topology, is it possible to play real internet traffic (not synthetic) within the given topology while providing remote-controllable traffic monitoring and sampling capabilities for each router of the described topology.



Scenario description

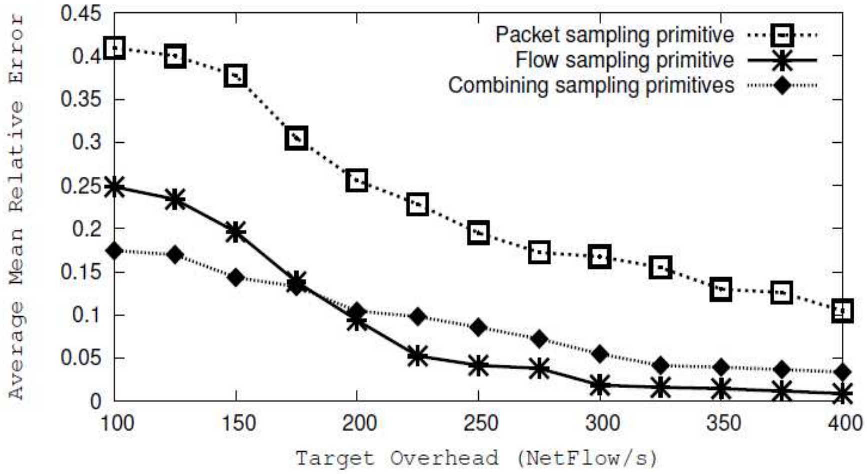
- We study the performance of our system on Geant topology.



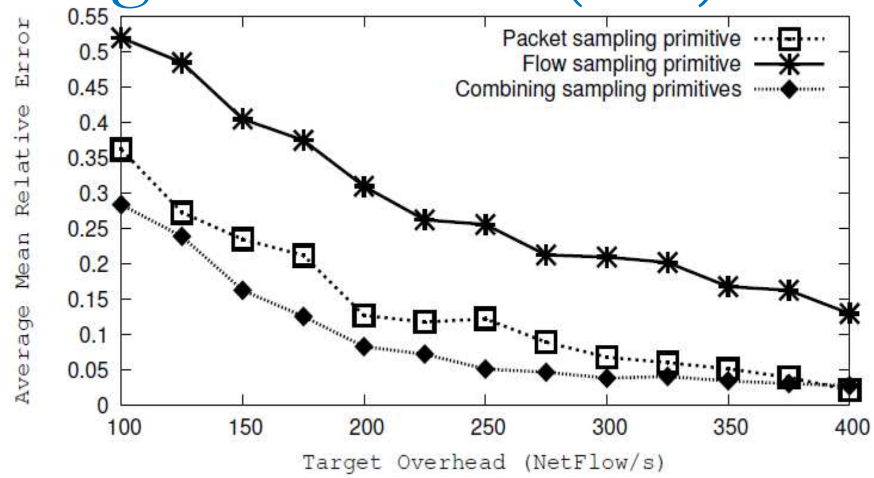
- Validation applications:
  - **Flow size estimation.**
  - **Flow counting**
  - **Heavy hitter detection**

Validation Results

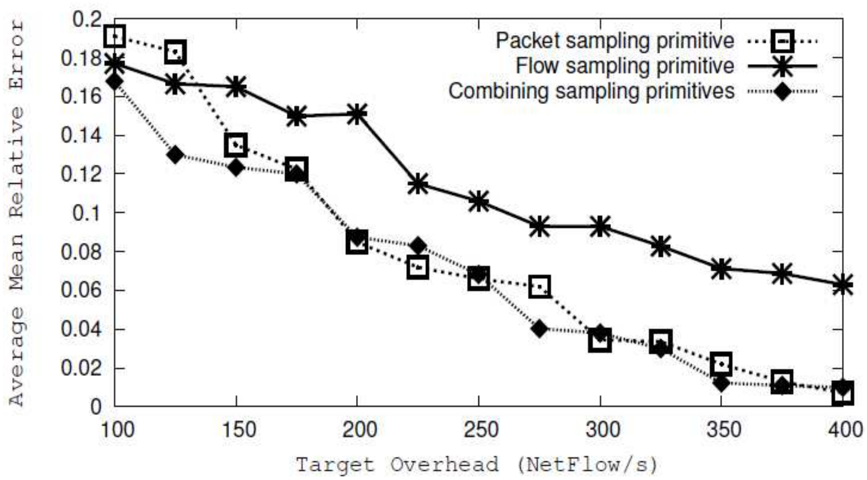
Mean relative error vs. Target overhead (TO).



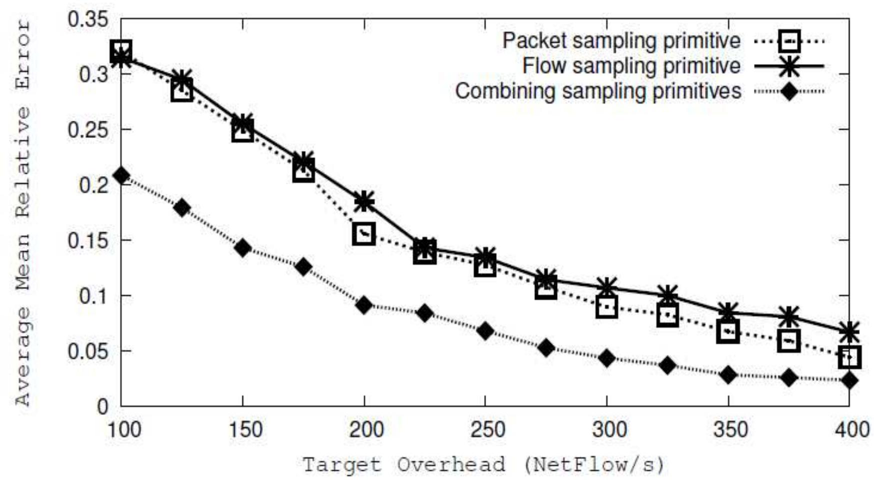
(a) ***Flow counting***



(b) ***Flow size estimation***



(c) ***Heavy hitter detection***



(d) ***Global accuracy***

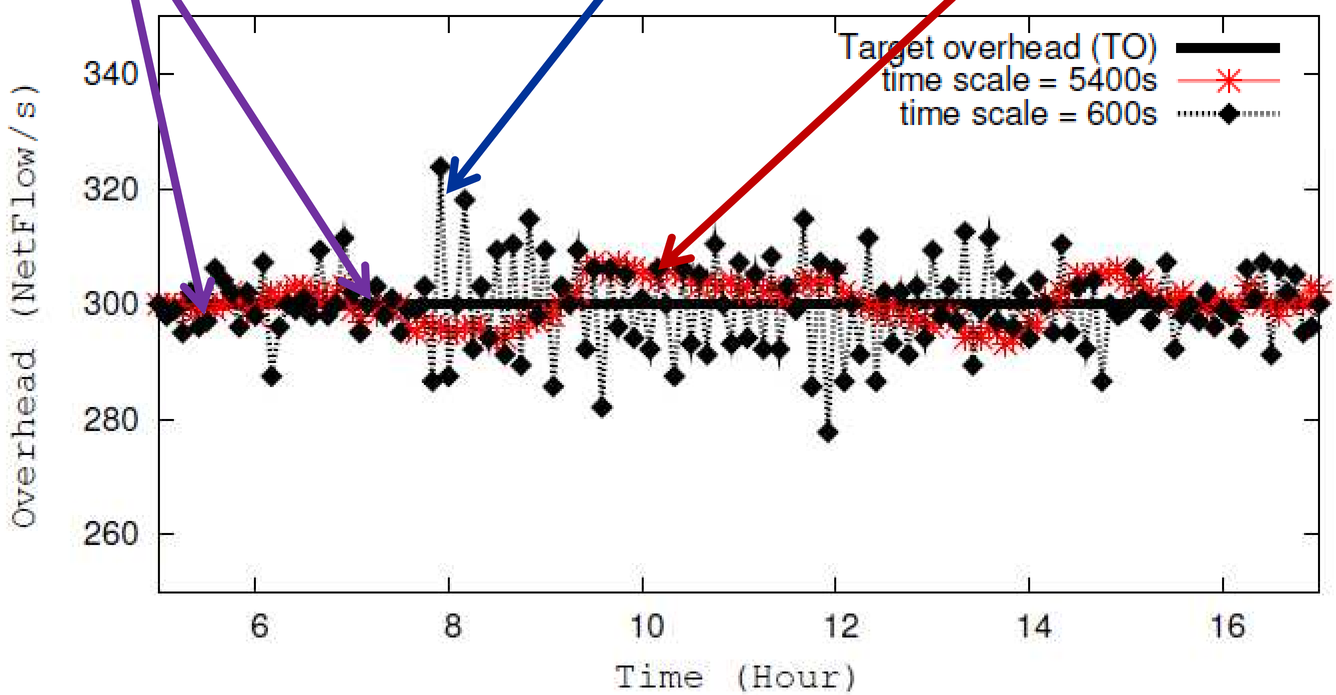
Validation Results

Resulting overhead vs. time using two time scales.

The system maintains the overhead around the TO.

Stable behavior

Oscillating behavior



Validation Results
Sensitivity analysis study

- To characterize qualitatively and quantitatively the impact of an input parameter on the system output and haw it compares with the impact of the other parameters .
- The main idea of FAST (Fourier Amplitude Sensitivity Test) is to assign to each parameter a distinct integer frequency (characteristic frequency). Then for a specific parameter, the variance contribution can be singled out of the model output with the help of the Fourier transformation

Parameter	symbol	range	impact
Target Overhead	$\mathcal{TO}$	$[20, 500]$	0.431
Time scale	τ	$[60s, 7200s]$	0.1147
Computation period	d	$[60s, 300s]$	0.0234
Min sampling rate	SR_{min}	$[0, 0.01]$	0.0876
Max sampling rate	SR_{max}	$[0.01, 1]$	0.0935

Parameters of the experiment.

Conclusion

Summary

Our adaptive monitoring system:

- Coordinates responsibilities between the different monitors and sampling primitives.
- Keeps the overhead around the target value
- Improves accuracy according to tasks requirements and importance.
- Tracks changes in the traffic.

Perspectives:

Our adaptive monitoring system:

- Extension to more applications
  - Flow size distribution
  - Anomaly detection
- Distribute the control

A multi-task adaptive monitoring system combining different sampling primitives.

Thank you !

Imed LASSOUED (Imed.Lassoued@.inria.fr)
Chadi BARAKAT (Chadi.Barakat@.inria.fr)