FLAMINGO

European Seventh Framework Network of Excellence

http://www.fp7-flamingo.eu/

WP5 — Network and Service Monitoring

Deliverable D5.1 — Initial deliverable on network and service monitoring

University of Twente, The Netherlands (UT)
Institut National de Recherche en Informatique et Automatique, France (INRIA)
University of Zurich, Switzerland (UZH)
Jacobs University Bremen, Germany (JUB)
Universität der Bundeswehr München, Germany (UniBwM)
University Politecnicà de Catalunia, Spain (UPC)
iMinds, Belgium (iMinds)
University College London, United Kingdom (UCL)

Project funded by the European Union under the
Information and Communication Technologies FP7 Cooperation Programme
Grant Agreement number ICT-FP7 318488

Document Control

Title:	D5.1 — Initial deliverable on network and service monitoring
Type:	Public
Editor(s):	Anna Sperotto
E-mail:	a.sperotto@utwente.nl
Doc ID:	D5.1
Delivery Date:	31.10.2013
Author(s):	Anna Sperotto, Rick Hofstede, Ricardo Schmidt Anthea Mayzaud, Anuj Sehgal, Björn Stelte Christos Tsiaras, Daniel Dönni, Daphne Tuncer, Guilherme Sperb Machado, Jair Santanna, Marinos Charalambides Mario Flores, Mario Golling, Maxim Claeys, Niels Bouten, Nikolay Melnikov, Radhika Garg, Rashid Mijumbi, Sebastian Seeber, Steven Latré

For more information, please contact:

Dr. Aiko Pras
Design and Analysis of Communication Systems
University of Twente
P.O. BOX 217
7500 AE Enschede
The Netherlands
Phone: +31-53-4893778
Fax: +31-53-4894524
E-mail: <a.pras@utwente.nl>

Legal Notices

Executive Summary

Monitoring is at the basis of any informed network and service management decision. However, nowadays networks are becoming progressively more complex, have higher traffic rates and are more frequently exposed to attacks. For these reasons, targeted monitoring solutions should be investigated to fulfil the requirements of the management of the Future Internet. This mindset is core to the research conducted in WP5.

The S.M.A.R.T. (Specific, Measurable, Achievable, Relevant, Timely) objectives (Section B.1.1.5 of the Description of Work) that are key to this WP are 1) integration of PhD students, and 2) producing scientific publications. We discuss the S.M.A.R.T. objectives in Section 2.1. The WP has fulfilled all S.M.A.R.T. objectives set for the first year. WP5 has at least two fully integrated PhD students (i.e., jointly supervised and financially paid by FLAMINGO). In addition, many collaborations among the PhD students have started, involving also PhD students that may be not financially paid by FLAMINGO, but are actively contributing to the work package. In addition, the WP has produced an outstanding scientific output, and the amount of published and submitted papers fulfills and exceeds the defined objective of 20 papers.

WP5 has also delivered excellent results with respect to WP5-specific objectives, which are reported in Section 2.2. Topics of research fulfilling the objectives have been identified and PhD students are currently actively performing research. Section 2.2 shows which objectives have been targeted in the first year and which ones will be targeted at a later stage in the project.

This deliverable presents the WP5 contributions in the context of the following Tasks: *Architectures and technologies for monitoring* (T5.1), *Collection and processing of monitoring data* (T5.2), and *Application domains for monitoring* (T5.3). We map the WP5 objectives to the related tasks in Section 2.3, where we also provide information on the topics presented in this deliverable.

The key achievements for the first year are:

- Many collaborations have started between PhD students (Section 3), which form the basis of the scientific research. Most of these collaborations are carried out in conjunction with other research WPs. Since, in the FLAMINGO view, “Network and Service Monitoring” and “Automated Configuration and Repair” are tightly coupled research areas, this WP works in close collaboration with WP6. The joint PhD collaborations are identified in D5.1 as well as in D6.1 by highlighting the contributions to each work package and its objectives. In addition, prototypes and open-source tools are a crucial component for some of the research carried out in this WP. For this reason, WP5 is also closely collaborating with WP1.
- WP5 has an excellent track record of publications. In the first year, WP5 and WP6 have already published 37 papers. In addition, seven papers are currently under review at major conferences and journals. Five publications are relative to research that has been conducted in collaboration with other EU projects, and four papers are jointly co-authored between FLAMINGO partners. Considering that several PhD collaborations have started in this first year, we expect the joint scientific output to further increase in coming years. An overview of the scientific output for WP5 is given in Section 2.1. The complete list of FLAMINGO publications can be found in D8.1.
- WP5 delivers scientific research with high impact within the network management community. This deliverable presents three examples of this. The first example is the research on intrusion detection and fingerprinting, which resulted in the open-source tool *SSHCure* (Section 5). The second example is the work performed on the *Network Flow Query Language*

(NFQL) (Section 7). Also in this case, the outcome is an open-source tool. The tools are developed in collaboration with WP1. Finally, WP5 researchers are involved in the FLAMINGO activity that we name the *Internet Traffic Statistics* project, a platform for the collection and publication of periodical traffic reports from major network operators worldwide, aiming at becoming the point of reference for the management community for conclusions about long-term trends in Internet traffic (Section 6).

Contents

1 Introduction

Network and service monitoring is at the basis of any informed management decision. As such, monitoring is one of the cornerstones of the management of the Future Internet, and one of the core research areas of FLAMINGO. The goal of this deliverable is to describe FLAMINGO's achievements in this research domain.

The deliverable is structured such as to give relevance to the objectives set for this WP. Section 2 reports which are the S.M.A.R.T. (Specific, Measurable, Achievable, Relevant, Timely) objectives and how the WP has successfully achieved them. It then summarizes the active research that it is taking place on the topics identified by the WP5-specific objectives.

The first S.M.A.R.T. objective is the *integration of PhD students*. In adherence to the Description of Work (DoW), at least two fully integrated PhD students are active in WP5. In addition, many PhD collaborations within the consortium have started during this first year. Information regarding these topics can be found in Section 3.

The second S.M.A.R.T. objective concerns the *scientific output* of the project. In the first year, WP5 and WP6 has published 37 papers, both at major conferences and in journals. In addition, seven other papers are currently under review. An overview of the status of the S.M.A.R.T. objective is given in Section 2.1. For a detailed list of the FLAMINGO published and submitted papers we refer the reader to D8.1.

Sections 4–7 highlight current research activities that are related to the WP5-specific objectives. In particular, Section 4 presents a monitoring architecture for security; Section 5 introduces SSHCure, an open-source tool for intrusion detection and fingerprinting that has been developed in collaboration with WP1; Section 6 reports WP5's efforts in data collection; and finally, Section 7 introduces a flow query language in the form of NFQL, an open-source tool also developed in collaboration with WP1.

Section 8 reports on the activities regarding the integration of European research that have been relevant for WP5. Finally, we conclude in Section 9.

2 Objectives and Tasks

This section presents an overview of the S.M.A.R.T. (Specific, Measurable, Achievable, Relevant, Timely) objectives for WP5 and the WP5-specific objectives. For the S.M.A.R.T. objectives, we indicate how these have been achieved for the first year. For the WP5-specific objectives, we summarise the activities that have taken place among the consortium partners, and we indicate the plans to address some of those objectives (2 and 7) in subsequent work.

2.1 S.M.A.R.T. Objectives

To comply to the S.M.A.R.T. objectives, WP5 has been active on the following topics:

- **Integration of Ph.D. students** – The Description of Work (Section B.1.1.5) states that "after 9 months each research WP will have identified at least two fully integrated Ph.D. students, which means that these students will be jointly supervised and financially paid by FLAMINGO". Since the beginning of the project, seven PhD students have joined FLAMINGO as *fully integrated PhD students*. These students, their affiliation and the co-supervising institution are listed in D8.1. For the FLAMINGO project, collaboration is at the basis of research. Therefore, PhD students are not working in isolation, but they are encouraged to collaborate with other institutions. For this reason, there is not a one-to-one match between a PhD student and a single WP. In addition, it is also important to highlight that PhD collaborations are taking place not only among fully integrated PhD students, but also with students that are not financially paid by FLAMINGO but that are actively contributing to the WP work. More information on the integration of PhD students, the PhD students active in the context of WP5 and their collaborations within the consortium can be found in Section 3.
- **Research** – The Description of Work (Section B.1.1.5) states that "after 18 month at least 20 scientific papers will be submitted / published". In the first year, the project has fulfilled and exceeded the expected number of publications. In the first year, WP5 and WP6 has published 37 papers, both at major conferences and in journals. In addition, seven other papers are currently under review. Given the strong collaboration between these WPs that is highlighted by the PhD collaborations described in Section 3, several publications presents both WP5 and WP6 aspects. For this reasons, we do not report the complete list of publications in this deliverable, but we refer the reader to D8.1. The scientific output reported in Table 1 is the outcome of collaborations with other European projects and European institutions. Table 2, on the other hand, reports the scientific output, in terms of published and submitted papers, that have been co-authored by more than one member of the FLAMINGO consortium.

[1] http://www.ict-mplane.eu/
[2] http://www.univerself-project.eu/
[3] http://www.smartenit.eu/
[4] http://www.eitictlabs.eu/
[5] http://cordis.europa.eu/fp7/ict/future-networks/documents/call8-projects/fusion-factsheet.pdf

Table 1: FLAMINGO publications in collaboration with other EU projects and institutions.

Authors	Title	Venue	EU project/ institution
I. Drago, E. Bocchi, M. Mellia, H. Slatman, A. Pras	Benchmarking Personal Cloud Storage [1]	ACM/SIGCOMM IMC 2013	mPlane[1]
M. Barrére, R. Badonnel, O. Festor	Vulnerability Assessment in Autonomic Networks and Services: A Survey [2]	IEEE Surveys & Tutorials	UniverSelf[2]
A. Lareida, T. Bocek, S. Golaszewski, C. Lüthold, M. Weber.	Box2Box – A P2P-based File-Sharing and Synchronization Application [3]	P2P 2013	SmartenIT[3]
G. Sperb Machado, T. Bocek, M. Ammann, B. Stiller	A Cloud Storage Overlay to Aggregate Heterogeneous Cloud Services [4]	LCN 2013	SmartenIT[3]
P. Poullie, B. Stiller	Fair Allocation of Multiple Resources Using a Non-monetary Allocation Mechanism [5]	AIMS 2013	SmartenIT[3]
C. Schmitt, B. Stiller, T. Kothmayr, W. Hu.	DTLS-based Security with two-way Authentication for IoT [6]	IETF	SmartenIT[3]
O. Festor, A. Lahmadi, R. Hofstede, A. Pras	Information Elements for IPFIX Metering Process Location (Internet Draft) [7]	IETF	EIT ICT Labs[4]
D. Tuncer, M. Charalambides, R. Landa, G. Pavlou	More Control Over Network Resources: an ISP Caching Perspective [23]	CNSM 2013	Fusion[5]

Table 2: Publications authored by multiple FLAMINGO partners.

Authors	Title	Venue	FLAMINGO partners
S. Seeber, A. Sehgal, B. Stelte, G. Dreo Rodosek, J. Schönwälder	Trust Computing Architecture for RPL in Cyber Physical Systems [8]	CNSM 2013	UniBwM, JUB
R. de O. Schmidt, N. Melnikov, R. Sadre, J. Schönwälder, A. Pras	Linking Network Usage Patterns to Traffic Gaussianity Fit [9]	PAM 2014 (under review)	UT, JUB
R. Mijumbi, J. Serrat, J.L. Gorricho, M. Claeys, F. de Turck, S. Latré	Design and Evaluation of Learning Algorithms for Dynamic Resource Management in Virtual Networks [10]	NOMS 2014 (under review)	UPC, iMinds
A. Mayzaud, A. Sehgal, R. Badonnel, I. Chrisment, J. Schönwälder	Mitigating DODAG Inconsistency Attacks in RPL Networks [11]	IPSN 2014 (under review)	INRIA, JUB

2.2 Workpackage Objectives

This section provides an high-level summary of the WP5-specific objectives. The objectives have been grouped in two categories. Section 2.2.1 describes the status of the objectives in which WP5 researchers are currently active in, both in term of research topics as well as academic activities in general. We name these as *ongoing objectives*. Section 2.2.2, on the other hand, includes the objectives for which activities have been planned to start at a later stage of the project (*open objectives*).

2.2.1 Ongoing Objectives

Objective 1: To integrate European research in the area of (flow-based) network and service monitoring – In collaboration with WP2, WP3 and WP4, WP5 has taken part in several activities in the context of network and service monitoring at European level. WP5 has collaborated with the *IRTF Network Management Research Group* (NMRG) by organizing and taking part to the **5th Workshop on the Usage of Netflow/IPFIX in Network Management**[6] (30th NMRG meeting) and the **1st Workshop on Large Scale Network Measurements**[7] (31st NMRG meeting). WP5 members have also been actively involved in the organization of the Dagstuhl Seminar **Global Measurement Framework**[8]; in the tutorials **Large-scale Measurement Platforms** (AIMS 2013) and **Management of the Internet of Things** (IM 2013); and finally in the organization of the 7th International Conference on Autonomous Infrastructure, Management and Security (AIMS 2013). Section 8 provides more information regarding the aforementioned activities.

Objective 3: To develop a generic distributed flow monitoring architecture – In the past, most Intrusion Detection Systems (IDSs) have been signature-based, inspecting (the payload of) every packet on a link, looking for predefined, malicious patterns. They have also mostly been developed as single observation points, both performing the collection and the analysis of the relevant data. Although this approach is, in some cases, still feasible from a technical perspective, it requires expensive hardware for dealing with high-speed links in backbone networks, for example. Our point of view with respect to this is that the monitoring and analysis of complex and highly dynamical systems, as the Internet, can benefit from the combined input of multiple, distributed and dedicated observation points, orchestrated by an intelligent manager. The research in this field is lead by the PhD collaboration between UT and UniBwM. More information about the proposed architecture can be found at Section 4.

Objective 4: To develop a flow query language for expressing temporal relationships of complex flow patterns – Cisco's NetFlow protocol and IETFs IPFIX open standard are the widely deployed techniques for collecting network flow statistics. Understanding certain patterns in these network statistics requires sophisticated flow analysis tools that can efficiently mine flow records. Existing solutions, however, lack of the flexibility to allow complex pattern-matching in large flow-based datasets. For this reason, JUB has developed a flow query language, the *Network Flow Query Language* (NFQL), specifically targeted to efficient and flexible flow mining. NFQL has been recently proposed in [12] and also presented at the *5th Workshop on the Usage of Netflow/IPFIX in Network Management*. More information on NFQL is available in Section 7 and, since it is available as an open-source tool, in D1.1.

[6] `https://trac.tools.ietf.org/group/irtf/trac/wiki/NetworkManagementResearchGroupAg&LgBerlin`
[7] `https://trac.tools.ietf.org/group/irtf/trac/wiki/NetworkManagementResearchGroupAg&LgZurich32`
[8] `http://www.dagstuhl.de/13472/`

Objective 5: To collect (anonymized) monitoring data – Several data collection activities are taking place among the FLAMINGO partners. Many of these happen in the context of the Joint Security Lab (see D1.1). A different initiative, which has the potential of achieving high visibility in the network management community, is the *Internet Traffic Statistic* project carried out at UT. The *Internet Traffic Statistics* aims at filling the gap left in the community by the, now discontinued, *Internet2 NetFlow: Weekly Reports*, which were one of the few sources able to provide long term trends on Internet traffic. The *Internet Traffic Statistics* will provide aggregated traffic statistics periodically collected at ISP and large backbone operators. The project has been especially welcomed by Research and Education Networks worldwide, and currently has the support from SURFnet (The Netherlands), CESNET (Czech Republic), DeiC (Denmark), RNP (Brasil) and GÉANT (the pan-European Research and Education Network).

Objective 6: To create annotated traces to assess the quality of different Intrusion Detection Systems – UT is currently involved in researching how traffic characteristics are affecting the sensitivity of Intrusion Detection Systems. An example is a work carried out by R. Hofstede and L. Hendriks (M.Sc. student at UT), which focuses on fingerprinting and annotating SSH attacks, i.e., scans and brute-force attacks generated by various attack tools. This is done by means of honeypots and traffic traces from production systems with active SSH daemons. The knowledge acquired from this master thesis will be valuable for the development and the performance assessment of SSHCure, a behavior-based Intrusion Detection System (IDS) for the detection of SSH dictionary attacks. (Section 5, and Obj 8). SSHCure is used for several PhD collaborations (Section 3.1), and an integral part of our network monitoring architecture (Section 5).

Objective 8: To propose novel solutions for intrusion detection and fingerprinting – UT has developed a behavior-based Intrusion Detection System (IDS) for the detection of SSH dictionary attacks, named SSHCure. This IDS is novel in the sense that it is flow-based and the only IDS of its kind that is able to identify whether an attack has been successful or not (i.e., whether an SSH login by an attacker has been successful). Moreover, the IDS has been released as open-source software. More details on SSHCure are provided in Section 5.

Objective 9: To propose and study monitoring frameworks for IaaS, PaaS and SaaS Clouds (i.e., to allow elastic management of cloud infrastructures) – UT, in collaboration with the mPlane project[9], is active in the characterisation of personal cloud storage services. Personal cloud storage services are data-intensive applications already producing a significant share of Internet traffic. Several solutions offered by different companies attract more and more people. However, little is known about the capabilities of each service, its architecture, and above all, the performance implication of the design choices. The outcome of this research was a paper published at ACM/SIGCOMM Internet Measurement Conference 2013 (IMC, acceptance rate 23%) [1]. The paper presents a methodology to study cloud storage services and it compares 5 popular offers, revealing different system architectures and capabilities. The implications on performance of different designs are assessed executing a series of benchmarks. The results show no clear winner, with all services suffering from some limitations or having potential for improvement. In some scenarios, the upload of the same file set can take seven times more, wasting twice as much capacity [1].

[9] http://www.ict-mplane.eu/

2.2.2 Open Objectives

Objective 2: To create and maintain articles within Wikipedia and other online systems in this area – WP5 plans to identify a list of possible topics which are either currently not covered in Wikipedia pages or that can be improved considering the expertise of the project members. This objective will be addressed starting from the second year, and it is linked to the FLAMINGO dissemination activities through popular online media. The decision of which topic to address, however, will be taken based on several factors, among which one of the most important is the knowledge generated in the project. Since knowledge generation has started in the first year, this objective has been postponed to the second year.

Objective 7: To investigate the applicability of different AI and machine learning techniques for flow analysis – WP5 has in plan of conducting research on this topic starting from early 2014. An important role in this objective will be taken by INRIA, which will conduct research on the topics of SCADA (Supervisory Control and Data Acquisition) anomaly detection and anomaly detection applied to network flows and DNS. The research on SCADA networks aims at developing a methodology to automatically discover a pattern of behaviour of a running SCADA system through the analysis of traffic messages traveling in its control loop network. The extracted patterns of behaviour will be used to automatically generate anomaly detection specifications to identify deviations in a running SCADA system. The work on flow-based anomaly detection will investigate graph-based approaches for describing host relationships, and it will for example use advanced machine-learning methods for fingerprinting. Also in the context of this research area, the FLAMINGO partners will investigate possibilities for collaborations.

2.3 Tasks and Objectives Mapping

Table 3 summarises the status of the S.M.A.R.T. objectives relative to WP5 (Section 2.1) and the WP5-specific objectives (Section 2.2). For each of the considered objectives, Table 3 indicates if the objective has been achieved (S.M.A.R.T objectives), or if there are WP activities that are contributing to the objective (WP5-specific objectives). For the WP5-specific objectives, Table 3 shows to which of the Tasks in the DoW the objective is contributing to. Finally, the table acts as a guide for the reader to locate the sections of this deliverable that provide additional information on a specific objective.

Table 3: Mapping of objectives and tasks.

Objective	Task 5.1	Task 5.2	Task 5.3	Status	Additional Material
S.M.A.R.T. Objective 1				Achieved	Section 3
S.M.A.R.T. Objective 2				Achieved	D8.1
WP Objective 1				Ongoing	Section 8
WP Objective 2				Open	
WP Objective 3	X			Ongoing	Section 4
WP Objective 4			X	Ongoing	Section 7
WP Objective 5		X		Ongoing	Section 6
WP Objective 6		X		Ongoing	
WP Objective 7		X	X	Open	
WP Objective 8			X	Ongoing	Section 5
WP Objective 9			X	Ongoing	

3 Integration of PhD Students

The integration of PhD students is a key SMART objective for this WP. We initially outline the PhD collaborations within the consortium in Section 3.1 and provide a more detailed description of the collaborations in Section 3.2.

For an overview of the *fully integrated* PhD students, including their affiliations and the co-supervisors, we refer the reader to D8.1.

3.1 PhD Student Collaborations

In the course of the first year of the project, the FLAMINGO PhD students and their supervisors have been invited to develop collaborations with other partners in the consortium. The collaborations have been identified during the FLAMINGO meeting that took place in February 2013 (see D8.1), and they are graphically described in Figure 1. Tables 6 – 9 in Appendix A provide an overview of the collaborations with respect to the WP5, WP6 and WP7 objectives that they contribute to.

Table 4 gives an overview of the PhD students active in WP5 and it lists the collaborations that have been started among members of the consortium. Throughout this deliverable, but also in D6.1 and D7.1, we refer to the collaboration by means of acronyms. For completeness, Table 4 also reports the acronym of the collaborations in which each PhD student is involved. In Section 3.2, we provide a description of the collaborations.

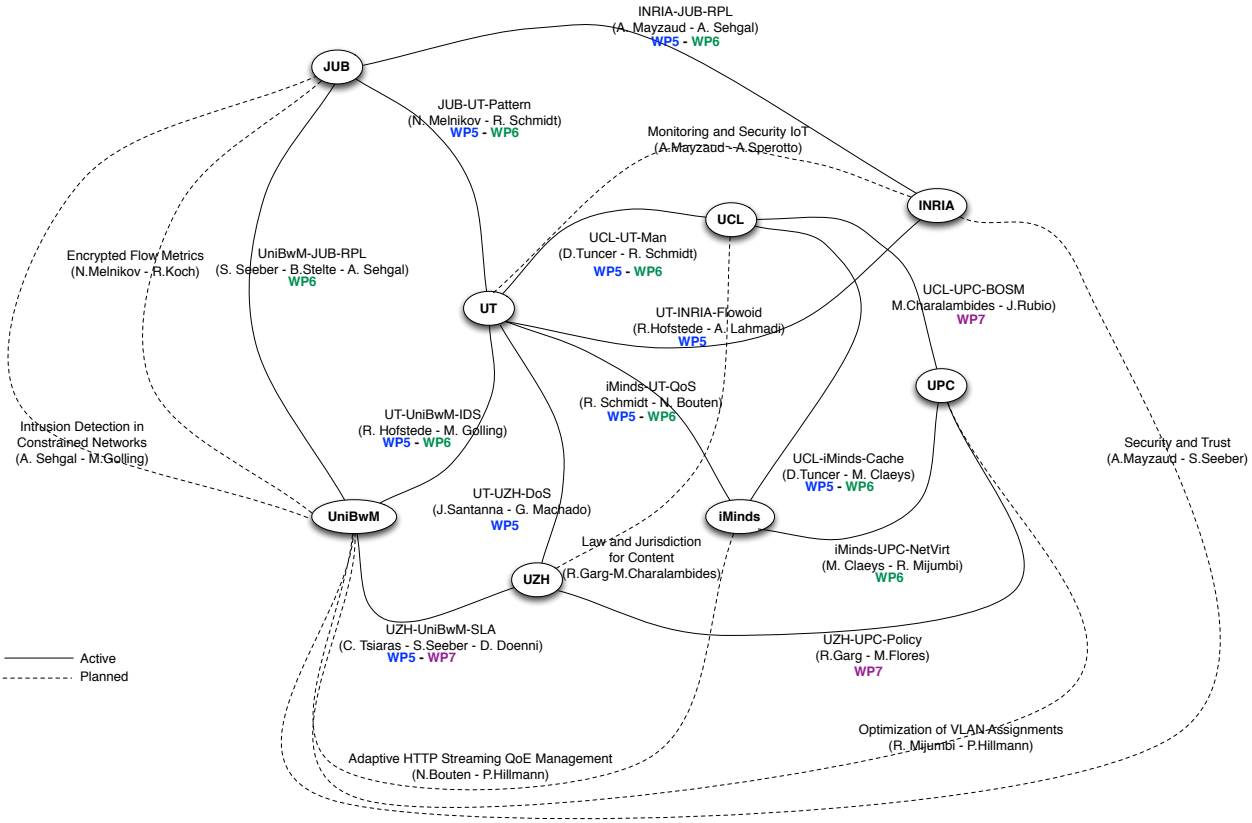


Figure 1: Overview of PhD collaborations.

Table 4: PhD students in WP5, and collaborations with consortium partners.

Name	Affiliation	Collaborations	Collaboration Acronyms
Anthéa Mayzaud	INRIA	JUB	INRIA-JUB-RPL
Nikolay Melnikov	JUB	UT	JUB-UT-Pattern
Rick Hofstede	UT	UniBwM, INRIA	UT-INRIA-Flowoid UT-UniBwM-IDS
José Jair C. de Santanna	UT	UZH	UT-UZH-DoS
Anuj Sehgal	JUB	UniBwM, INRIA	INRIA-JUB-RPL UniBwM-JUB-RPL
Mario Golling	UniBwM	UT	UT-UniBwM-IDS
Christos Tsiaras	UZH	UniBwM	UZH-UniBwM-SLA
Guilherme Sperb Machado	UZH	UT	UT-UZH-DoS
Daphne Tuncer	UCL	iMinds, UT	UCL-iMinds-Cache UCL-UT-Man
Maxim Claeys	iMinds	UCL, UPC	iMinds-UPC-NetVirt UCL-iMinds-Cache
Niels Bouten	iMinds	UT	iMinds-UT-QoS
Sebastian Seeber	UniBwM	JUB, UZH	UZH-UniBwM-SLA UniBwM-JUB-RPL
Daniel Dönni	UZH	UniBwM	UZH-UniBwM-SLA
Ricardo Schmidt	UT	JUB, UCL, iMinds	iMinds-UT-QoS UCL-UT-Man JUB-UT-Pattern
Björn Stelte	UniBwM	JUB	UniBwM-JUB-RPL
Rashid Mijumbi	UPC	iMinds	iMinds-UPC-NetVirt

3.2 Description of the Collaborations

This section describes the PhD collaborations among FLAMINGO partners that have been defined during the first year of the project. Since the collaborations involve elements of both WP5 (*Network and Service Monitoring*) and WP6 (*Automatic Configuration and Repair*), this section will appear in both this deliverable and deliverable D6.1.

3.2.1 Linking Network Usage Patterns to Traffic Gaussianity Fit (JUB-UT-Pattern)

The Gaussianity of traffic aggregates is a desirable characteristic in the domain of network traffic modelling due to the wide adoption of Gaussian models. Past works have been extensively researched the Gaussian property of traffic aggregates, its advantages for proposing traffic models and how this can be disturbed by traffic bursts. To the best of our knowledge, however, never a work has tried to connect traffic Gaussianity, or lack thereof, to network usage patterns. This knowledge would be valuable in understanding the limitations of current traffic models in presence of certain traffic. Therefore, the aim of this collaboration is to find out the potential connections between traffic bursts and poor Gaussianity , and also to point out what sort of host and/or application traffic pattern creates such disruptions. Preliminary results show that Gaussianity fit can be directly connected to presence or absence of extreme traffic bursts. The results also show that even in a more homogeneous network (i.e., hosts with similar access rates) we can identify extreme traffic bursts that might ultimately compromise Gaussianity fit. A paper on the topic of this collaboration is currently under review for PAM 2014 [9].

This work contributes to WP5 since it is based on the collection of packet-level traffic traces from many different locations around the globe.

The collaboration also contributes to WP6, since it aims at establishing a “rule-of-thumb” to estimate whether traffic is Gaussian or not, based on conclusions from data analysis such as hosts behavior and applications usage.

3.2.2 Energy-aware Traffic Management (UCL-UT-Man)

UCL has developed an adaptive resource management approach, which can reduce the energy consumption of core IP networks [13]. The approach is based on the reconfiguration of traffic splitting ratios at the edges of the network, so that traffic is distributed over a subset of router line cards, while unused ones enter sleep mode. UT has been working on methods to estimate the required bandwidth of network links based on flow-level traffic measurements [14].

The goal of the collaboration between UCL and UT is to propose a decentralised system for traffic management supported by a link dimensioning approach for better reallocation of link resources. The system takes decisions on traffic splitting within a backbone network focusing on energy efficiency by reducing the number of required line cards. Currently, the system assumes that traffic averages represent the required capacity to be allocated per flow. Link dimensioning would allow the management system to have better estimations of required capacity, e.g., considering quality of service metrics, which would ultimately be used on the reconfiguration of traffic splitting.

This work is being carried out in the scope of WP5 and WP6, addressing the following aspects:

- The collaboration contribute to the collection of anonymised monitoring data. This data comprises packet-level traffic captures used to validate the link dimensioning procedure and also to produce synthetic traffic to be used in simulations of the management system. The need for synthetic traffic is due to difficulties in acquiring traffic captures of an entire backbone network. This aspect contributes to WP5.
- This activity extend previous work on adaptive resource management and enhance the energy-awareness control loop, by providing better link load estimates. This can allow the control loop to make more energy-friendly reconfiguration decisions. The performance of the approach will be evaluated in terms of the number of line cards that can enter sleep mode, as well as link utilisation. This aspect contributes to WP6.

3.2.3 Intrusion Detection Systems (UT-UniBwM-IDS)

This joint research activity is a collaboration between UT and UniBwM. Intrusion detection is nowadays commonly performed in an automated fashion by IDS [15]. Several classifications for IDSs are common. One of these classifications focuses on the kind of data that is used for performing intrusion detection. The first class of IDSs mainly uses packet headers (flows) for intrusion detection. While these flow-based IDSs have a high-performance and are usually little privacy-intrusive, they are typically affected by a high number of undetected attacks (false negatives; see Figure 2). In contrast to flow-based IDSs, payload-based IDSs are capable of performing extensive layer-7-detection (and, therefore, have a lower false negative rate), but at the expense of a much higher system requirements as well as a violation of privacy [16].

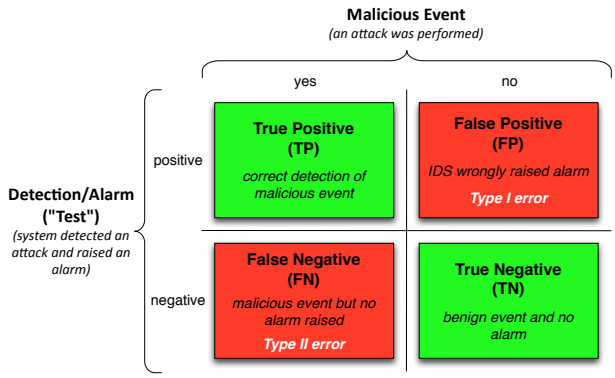


Figure 2: Categories of Alarms ("Confusion Matrix").

Given these observations, performing intrusion detection in high-speed networks is a challenging task. While many payload-based IDSs are working well at the backend of service provider networks, the backbone is often characterized by communication links with high-speed connections and thus requires well equipped IDS in order to be capable of handling 100 Gbps or more, for example [17]. Within this collaboration, it is planned to create a framework for distributed intrusion detection in high-speed networks by combining especially flow-based and payload-based intrusion detection. As already stated, in addition to monetary aspects, legal issues in general and privacy issues in particular are also important reasons, why payload-based IDS are rarely deployed in high-speed networks today [16].

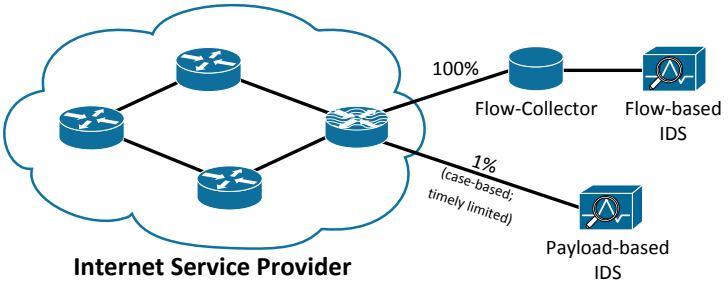


Figure 3: Simplified Scenario.

In order to overcome these disadvantages, this collaboration tries to make use of both approaches (both flow-based and payload-based intrusion detection) in a multi-layered approach. As depicted in Figure 3, the approach is centered around the ideas that (i) the first layer comprises flow-based intrusion detection, which performs detection based on the entire packet stream (100%) and that (ii) depending on the result of the flow-based detection, the payload-based IDS is activated for a certain period of time to investigate the anomaly of the flow-based IDS in more detail (1%) - in order to verify or falsify the result of the flow-based IDS. As network attacks can last shortly and a switch has to be made from flow-based to packet-based detection, detection has to be performed in real-time.

This work is being carried out in the scope of WP5 and WP6, addressing the following aspects:

- The collaborative work contributes mainly to WP5: Network and Service Monitoring by performing multi-layered IDS. Especially objective 8 - novel solutions for IDS is addressed with this collaboration. As the long-term goal of this collaboration is also to link different managers with each other, this addresses objective 3 "to develop a generic distributed flow monitoring

architecture". Regarding this objective, in [18] an Evaluation of State of the Art IDS-Message Exchange Protocols was already performed.

- For WP6, [18] also contributed to objective 3 "to develop an inventory of approaches for automated configuration and repair". By using cloud-based solutions (as described in [19, 20]), requirements for cloud-based Services have been investigated and architectural approaches specific for this application domain have already been partially developed (also addressing objective 9 "to propose and study automated configuration and repair in the context of the management of clouds (especially Interclouds)") As it is planned to develop an "automated" architecture that can be used in different administrative boundaries, objective 5 "to develop new architectures for automated configuration and repair approaches across administrative boundaries" is addressed as well.

3.2.4 Towards A Trust Comupting Architecture for RPL in Cyber Physical Systems (UniBwM-JUB-RPL)

Cyber Physical Systems (CPSs) are widely expected to be formed of networked resource constrained devices. In order to suit the constraints of such networks, the Internet Engineering Task Force (IETF) developed the Routing Protocol for Low power and Lossy Networks (RPL) and Low-power and Lossy Networks (LLNs). Security in CPSs is important for maintaining the integrity and privacy of data, while also improving network resiliency to attacks. Even though RPL provides support for integrity and confidentiality of messages, details regarding key management and signatures are not covered. Since complexity and size is a core concern in LLNs, off-loading the security features to a Trusted Platform Module (TPM) can make it possible to include sophisticated security provisions in an RPL implementation.

This collaboration develops mechanisms to use the security mechanisms of a TPM in order to secure the communication in an RPL network. The secure exchange of control messages in an RPL network is supported in the standard defined mechanism, but there is no instruction which describes the key management and signatures. Our approach secures the exchange of control messages by using our defined mechanism in conjunction with the facilities of a TPM. The design of a trust establishment and key exchange mechanism around the implied trust of a TPM to provide keys for secure RPL nodes, is a main task of this research. With this approach, the usage of a TPM on Resource Constrained Devices reduces the processing load on the main processor. The goal of this examination is the prevention of the dissemination of misleading routing information, which can affect the availability of the whole network. As a next step, the previous developed idea will be deployed on real hardware devices to evaluate the solution in comparison to other approaches. This is necessary to prove the existing simulation results.

The collaboration fits within WP6, since it develops a mechanism, specifically, targeted to RPL networks, to secure the communication. This approach is applicable in RPL networks which are used in wireless sensor networks.

3.2.5 Flowoid: a NetFlow/IPFIX Probe for Android-based Devices (UT-INRIA-Flowoid)

Analysis of the network behaviour of applications running on a smartphone device requires the collection of information about data leaving the device and where it is sent. Ciscos NetFlow and the more recent IPFIX are flow export technologies that have seen a rapid adoption and widespread integration in many campus, enterprise and backbone networks. To be able to export and analyse mobile device characteristics (such as its location at the moment of certain network activity), the

NetFlow and IPFIX protocols have to be extended. The flow exporter, flow collector and analysis application need to be aware of these extensions as well. The work in this collaboration has been divided between INRIA and UT. On the one hand, INRIA is responsible for developing a flow exporter tailored to Android devices. On the other hand, UT is responsible for the flow collector and analysis application. The major achievements of the collaboration are:

- Development of a NetFlow and IPFIX metering process for Android devices;
- Extension of nfdump/Nfsen and SURFmap with location support;
- IETF draft describing a set of information elements for IPFIX metering process location [7].

The following aspects of this collaboration fall within WP5. In this work, INRIA has developed Flowoid, a NetFlow and IPFIX metering process tailored to Android devices. The probe associates geolocation data with each observed network flow, consisting of the GPS coordinates of the mobile device, among others. This information is exported together with the traditional fields defined in the NetFlow and IFPIX: IP version, source and destination addresses, the number of exchanged bytes, the type of protocol, the number of exchanged packets, the source and destination ports, and the duration of a flow. In addition, it contains 7 additional fields that denote the identifier of the device: the identifier of the localization method, a timestamp, the integer part of the latitude, the decimal part of the latitude, the integer part of the longitude, and the decimal part of the longitude. UT has extended the state-of-the-art flow collector nfdump/NfSen with location support, allowing us to analyse the flows exported by the Flowoid probe. In a previous work UT has developed a network monitoring tool based on the Google Maps API, named SURFmap [21], which adds a geographical dimension to flow data and displays the data on a map. Since SURFmap [21] already supports network traffic geolocation (i.e. adding physical locations of hosts to network data), this tool has been extended to visualize the locations of devices on a map. This will allow us to visualize network traffic of mobile device with respect to devices locations.

3.2.6 Flow-based Traffic Measurements for In-Network Video Quality Adaptation (iMinds-UT-QoS)

HTTP Adaptive Streaming (HAS) services allow the quality of streaming video to be automatically adapted by the client application in face of network and device dynamics. A major obstacle for deploying HAS in managed networks, is the purely client-driven design of current HAS approaches, which leads to excessive quality oscillations, globally suboptimal behavior, and the inability to enforce management policies. These challenges can be tackled by steering the quality selection from within the network. iMinds already deployed a distributed in-network management heuristic, which is able to reduce the number of switches with a factor 5 and increase the quality up to 30%. One of the shortcomings of the current deployment, however, is the assumption of a static bandwidth for each link and the absence of cross-traffic. To overcome this issue, this collaboration will use flow-based measurements to measure and predict the per-link throughput, which is available for HAS traffic. Using these predictions the in-network video quality adaptation can divide the resources among the different HAS clients based on a provider's policy.

This work contributes to WP5 because real world traffic traces will be collected and used to validate the link dimensioning approach.

This work contributes to WP6 in the following. The goal of this collaboration is to develop a distributed algorithm/heuristic, which is able to divide the resource among the various HAS clients subject to a providers policy. Using the measurements provided by WP5, each agent is able to

perform a local optimization based on the throughput predictions. The different distributed agents share this network information and their local decisions with each other to be able to automatically react to changes in the network environment. Using the measurements and predictions on the current and future cross-traffic, the agents can make estimations on the residual bandwidth that can be shared among the different HAS sessions. These local estimations and the shared network information serve as input to the algorithm which limits the quality of the HAS sessions crossing the managed resources. This leads to a more stable quality selection at the client, since oscillations due to changed network environments are avoided.

3.2.7 Study of DODAG Inconsistency Attacks in RPL Networks (INRIA-JUB-RPL)

The growing interest for the Internet of Things has resulted in the large-scale deployment of Low power and Lossy Networks, such as wireless sensor networks and home automation systems. A new routing protocol called RPL for IPv6 over Low power Wireless Personal Area Networks (6LoWPAN) has been specifically designed by the IETF Routing Over Low Power Lossy Networks (ROLL) working group to deal with these requirements.

RPL forms a tree like topology for routing packets, called a Destination Oriented Directed Acyclic Graph (DODAG). In order to detect possible loops, also referred to as DODAG inconsistencies, RPL uses IPv6 header options to track the direction of the packet and any rank errors. Specifically, the O-Bit option is used to track direction of the packet, i.e. upwards or downwards in a tree. If an upwards packet is received from a node with a rank lower than the current node, an inconsistency is detected. As such, the node will set the R-bit option, used to track rank errors, and forward the packet. If the next receiving node also detects an inconsistency in the direction of the packet, and the R-bit option is also set, this node will drop the packet and reset the trickle timers used by RPL.

Such a reset of trickle timers leads to an increase in the number and frequency of control packets being sent and received in the DODAG, thereby also impacting the already low available energy. A malicious node can create artificial DODAG inconsistencies by manipulating these IPv6 header options, thereby leading to increased overhead, denial of service and even black-hole attacks that are hard to detect.

The objective of this joint scenario is (1) to establish a state-of-the-art overview about security attacks against RPL networks, (2) to identify the key parameters that are required to detect these attacks, (3) develop a mitigation strategy to reduce the effect of such attacks, (4) develop an approach for children nodes to detect when a parent node might be malicious and (5) to experiment and evaluate the developed solutions.

A review of the state-of-the-art in RPL security has been conducted and this has led to the identification of DODAG inconsistency attacks that can lead to increased overhead and energy consumption, denial of service and even black-hole attacks.

Based on this, scenarios were constructed to study the performance of the RPL network when such attacks are carried out. Via an implementation in Contiki, it was identified that the mitigation strategy proposed by RPL, which involves ignoring packets with the appropriate IPv6 header after a fixed threshold is reached, uses an arbitrary value for the threshold. A new function that dynamically scales this threshold was developed to improve performance of the network while under attack. A comparitive study between the (1) no threshold, (2) fixed threshold and (3) dynamic threshold scenarios has been performed.

An approach to counter the black-hole attack has also been developed. This approach allows nodes to periodically enter a promiscuous mode in order to verify that the parent is not modifying packets in malicious ways, i.e. setting incorrect IPv6 header options. When such manipulation is

detected, it is countered by blacklisting the parent, informing the neighborhood and triggering a rebuild of the DODAG. An implementation of this algorithm in Contiki is underway, following which an evaluation will be performed. A paper on the topic of this collaboration is currently under review for IPSN 2014 [11].

This research contributes to WP5 and WP6:

- The collaboration is developing methods to identify malicious nodes that might attempt to infiltrate and negatively impact an RPL based network. Since there are many possible attack vectors, there are different approaches being designed to identify and subsequently mitigate the effect of these attacks. Upon successful identification, we are also working on an approach to blacklist such malicious nodes from being able to rejoin the network. These aspects of the collaboration fits within WP5.
- The collaboration approaches for RPL are being developed and tested on the IEEE 802.15.4 + 6LoWPAN platform, which is expected to form the basis of the Internet of Things. Similarly, all the work is being currently tested on the TelosB network node, which is a resource constrained device used commonly in the IoT and WSN areas. In the future, the collaboration is also expected to investigate resource constrained devices. Since resource constrained devices need targeted solutions for reconfiguration, this collaboration also contribute to WP6.

3.2.8 SLA Fulfillment Mechanism (UZH-UniBwM-SLA)

One research goal of this research activity is the definition of a mechanism that detects if an Service Level Agreements (SLA) is violated in the context of a voice service over mobile networks. The research is motivated by the need for an SLA violation detection mechanism to identify that a QoS-guarantee is fulfilled or not. The need for this mechanism derives from the Auction-based Charging User-centric System (AbaCUS) [22].

Furthermore, this joint research activity aims to determine suited actions in respect to charging when a violation is detected. Facilitating the first goal on the level of traditional circuit-switched mobile phone calls would demand insight in a Mobile Network Operator's (MNO) infrastructure. Since this is currently not possible, the decision to focus on Voice-over-IP (VoIP) services over mobile networks has been taken. To our knowledge, nowadays there is not a QoS related metric for VoIP services over mobile networks. Thus, this work of SLA Fulfillment Mechanism aims to provide the respective metric as well as a prototype of the evaluation mechanism.

In respect to the WP5, the joint research activity aims to monitor the environment of the participants of an VoIP call which establishes the connections via a mobile network. Additionally the statistics of a VoIP call are monitored if accessible. This data will be used to identify if an SLA violation has taken place during the last VoIP call. In respect to the used end device the mobile signal-strength and the battery level will be monitored. These monitored values will help to identify possible SLA violations.

Furthermore, for the needs of the SLA Fulfillment Mechanism joined research project and the Value of Service (VoS) PhD project (Daniel Dönni) a measurement application will be designed and implemented. The application will be developed for the Android platform and allow the end-user to determine the connectivity, one-way delay and loss, round-trip delay and loss, delay variation, loss patterns, packet reordering, bulk transport capacity, link bandwidth capacity, and packet duplication metrics. The measurement application will use UDP packets to determine the above metrics, except for the bulk transport capacity which relies on TCP. Furthermore, novel metrics will be

developed to capture price and price-performance aspects of IP networks. Details are subject to current research.

The measurement application will be equipped with a graphical user interface which allows the end-user to specify measurement parameters. In particular, it will be possible for the end-user to define a measurement schedule which contains a list of metrics that shall be measured. For each metric it will be possible to specify the following:

- whether a single measurement or a series of measurements shall be conducted
- what destination address shall be used
- what payload size shall be used

Depending on the metric, additional configuration options might be provided.

For the measurement application to function correctly, two additional components will be implemented. The first one is the measurement server which acts as a counterpart to the measurement application. Its main responsibility is to accept and reflect packets submitted by the measurement application. The second one is the measurement database which is responsible for storing the measurement data. The database will also contain a suitable schema for storing prices offered by network operators to capture pricing aspects as well.

All three components of the measurement platform will be implemented adhering to best-practice design and coding standards. This includes but is not limited to the definition of suitable application components, the implementation of unit tests using JUnit, as well as a proper documentation using JavaDoc. Furthermore, the measurement application will be made available in Google play, such that it can be easily downloaded and used by the research community.

The collaboration between the members of the SLA Fulfillment Mechanism and the Value-of-Service (VoS) member also contributes to WP7 by proposing business and regulation actions in case of SLA violation identification and price-performance metrics for IP networks. Additional details can be found at D7.1, Section 4.7 and Section 4.9.

3.2.9 Cache Management (UCL-iMinds-Cache)

Current content delivery services operated by large Content Delivery Networks (CDN) providers can exert enormous strain on Internet Service Provider (ISP) networks. This is mainly attributed to the fact that CDN providers control both the placement of content in surrogate servers spanning different geographic locations, as well as the decision on where to serve client requests from (i.e. server selection). These decisions are usually taken by using only limited information about the carrier networks, and this can adversely affect network usage.

UCL has developed an approach by which ISPs can have more control over their resources [23]. This involves operating a limited capacity CDN service within ISP networks by deploying caching points in the network. Empowering ISPs with caching capabilities can allow them to implement their own content placement and server selection strategies which will result in better utilisation of network resources. The work has investigated content placement strategies that can be used by the ISPs to manage the placement of content items in the various network caching locations according to user demand characteristics.

In this joint research activity, UCL and iMinds are extending the scenario previously considered by focusing on the case where a large-scale ISP leases caching capacity to multiple content providers

(CP). The objective of this work is to develop a new cache management strategy that can be used by the ISP to decide on the cache capacity allocation and content placement configuration that can minimise its network resource usage given the capacity requirements of each CP and the characteristics of the user demand.

Initial work has mainly focused on formally modelling the problem as an Integer Linear Program (ILP) problem, forming a common basis for the development of more simple and lightweight heuristic algorithms. The performance of the different heuristics will be evaluated based on realistic traffic traces provided by iMinds and will be compared according to several parameters.

This work is being carried out mostly in the scope of WP6. In the context of the proposed scenario, the cache management approach to be developed will implement a control loop that automates the configuration of in-network caching points. The performance of the approach will be evaluated over a number of parameters using real traffic traces. The collection of such traces contributes to WP5.

3.2.10 Management of Virtualized Networks (iMinds-UPC-NetVirt)

This joint research activity is a collaboration between iMinds and UPC. The work focusses on virtual network embedding. In virtual network embedding, a Virtual Network Provider (VNP) acts as a mediator between service providers (SPs) and infrastructure providers (InPs). Virtual network requests are launched by the service providers and requests contain requirements on node and link capacities. Service providers target to receive a virtual network, fulfilling the request, that minimizes the embedding costs. The Virtual Network Provider reserves substrate resources from the infrastructure provider to be able to embed the virtual networks. Embedding solutions are based on the resource capacities requested by the service providers. Actual load will however vary over time, leading to situations where a lot of substrate resources remain unused.

The first goal of this collaboration therefore is to develop a dynamic embedding algorithm, able to dynamically adapt the embedding solution to the actual demands, perceived by network monitoring. This will optimize substrate resource usage and increase the acceptance rate of virtual network requests. For this purpose, the researchers will apply a multi-agent reinforcement learning approach. This research has resulted in a paper, currently under review for NOMS 2014 [10].

For the next steps of this research, the collaborations will take a broader look at the network virtualization problem by considering a service provider point of view. Instead of embedding a virtual network topology onto a substrate network, the researchers will consider an additional step where the virtual network topology is constructed based on the service requirements. As a first step, a catalogue of service enablers and requirements that can be considered for embedding is being compiled. Based on this catalogue, the researchers will narrow down the specific problem definition.

The goal and approach of this collaboration lie within WP6. This is clearly visible when considering the dynamic embedding algorithms for automatically identifying the optimal network configuration that the collaboration is developing.

3.2.11 TraceMan-based Monitoring of DoS Attacks (UT-UZH-DoS)

Distributed Denial of Service attack (DDoS) is one of the major threats to the Internet. In general, this kind of attack aims to deny the ability of a host (target) to respond to legitimate network traffic. Especially, when a DDoS is based on exhausting network resources, it is very difficult, and often

ineffective, to mitigate with an on-premise solution (in which a target itself stops the attack), since the effect of the attack will also affect all the services and users in the same network.

In order to attempt to mitigate this kind of attacks, solutions that detect and stop the malicious traffic before reaching the target are considered. DDoS Network Protections (DNP) are examples of these solutions that redirect and filter the malicious traffic during attacks. An example of DNP is the online service CloudFlare[10].

DNP solutions have been adopted by a wide range of companies, and it seems also by companies that exchange a considerable amount of private data with customers, such as banks. However, few knowledge is is available on the practical use of DNP. Questions that raise are, for example, if these DNP solutions redirect malicious traffic to third-party domains, and, if so, where the traffic is redirected to, and what is – if any – the impact of employing DNP. Therefore our goal is to understand how DNP solutions work in practice, if they are reliable and what it means for the end users when their are handled by a DNP.

The planned approach for this research is to monitor targets, which are protected by a DNP solution, especially during DDoS attacks. This can be achieved by long term monitoring of services known to be using DNP, or in a lab environment, where DDoS can be generated on-demand. The monitoring should encapsulate several dimensions (e.g., round-trip-time, volume of traffic generated) , but being focused on where the malicious traffic is going through, for example by means of one of the open source tools developed in this project, TraceMan (see D1.1).

At this stage, the collaboration and its goals have been identified, but the research work is planned to start later in the project. Given its strong measurement focus, this collaboration fits in WP5.

3.3 Collaboration Activities

In this section we list the activities that have taken place in relation to the aforementioned collaborations.

- Face-to-face FLAMINGO meetings have taken place during the IM 2013 conference in Gent, in particular:
  - **UT-UZH**: the goal of the meeting was to investigate possible collaboration between the UZH PhD student Guilherme Sperb Machado and the researchers from UT. The outcome of this meeting has later concretised in the collaboration **UT-UZH-DoS**
  - **UT-INRIA**: the goal of the meeting was to investigate possible collaboration on the topic of network security and modeling between the PhD student Anthéa Mayzaud (INRIA) and the UT.
  - **UCL-iMinds**: the goal of the meeting was to define the work plan for the collaboration **UCL-iMinds-Cache**. During the meeting the research topics under investigation were discussed in more detail including cache management algorithms and mechanisms to dynamically scale the cache capacity.
  - **UT-iMinds**: R. Schmidt (UT) and N. Bouten (iMinds) discussed the different approaches for traffic measurements (sFlow, packet-based) and how they can be integrated within the ns-3 based HTTP Adaptive Streaming simulation framework (**iMinds- UT-QoS**). This meeting was at the basis for the collaboration **iMinds-UT-QoS** which progressed in the next months.

[10] https://www.cloudflare.com/

- Nikolay Melnikov (JUB), has visited the UT for one week in June 2013 in the scope of the PhD collaboration with Ricardo Schmidt. The topic of the collaboration **JUB-UT-Pattern** is establishing the contribution of individual hosts to the Gaussianity properties of the traffic aggregate. This research is carried out in collaboration with Dr. R. Sadre of the University of Aalborg. Regular phone calls are taking place.
- Jair Santanna (UT) has visited the Jacobs University Bremen for a face-to-face meeting with the co-supervisor J. Schönwälder. The goal of the meeting was to discuss the Jair Santanna's PhD topic and his current research.
- In the context of the collaboration on **UZH-UniBwM-SLA**, bi-weekly telcos are taking place. In the mid of November one week in Zurich is planned to evaluate mobile measurement environments and to get possible inputs for conference papers. Also a visit from Christos Tsiaras and Daniel Dönni is planned in Munich in the early 2014.
- Rashid Mijumbi (UCL) has visited iMinds for two weeks in the scope of the collaboration **iMinds-UPC-NetVirt**, with the goal of defining research directions for a collaboration between the two institutions, on the topic of virtual network embedding. The outcomes are detailed discussions to clear definitions of the collaboration timeline including public targets. Additionally a paper to NOMS has been submitted [10].
- UT and iMinds are currently collaborating **iMinds-UT-QoS** on combining flow-based traffic measurement with In-Network Video Quality Adaptation. Researchers from the two institutions have regular phone calls to discuss the progress of the collaboration.
- Rick Hofstede (UT) has visited UniBwM two days in February 2013 in the scope of Intrusion Detection Systems in the collaboration **UT-UniBwM-IDS** with Mario Golling. The outcomes are further collaboration plans and ideas for joint papers. In the future more shorter visits in Enschede and Munich, depending on paper deadlines, are planned.
- Radhika Garg (UZH) is planning a visit at UPC in the scope of the collaboration **UZH-UPC-Legal** in the context of legal regulations.
- Daphne Tuncer (UCL) is planning to visit at iMinds to prepare a joint paper in the scope of the collaboration **UCL-iMinds-Cache**.
- Anthea Mayzaud (INRIA) visited JUB for one week in June 2013 in the scope of the collaboration **INRIA-JUB-RPL**. During this visit they defined the exact collaboration aspects. They discussed a topology for simulation study with multiple scenarios that take into account different attack and data-delivery rates. Additionally they modified the Contiki RPL implementation to enable the evaluation of possible DAG inconsistency attacks and the structure for a full paper was defined.

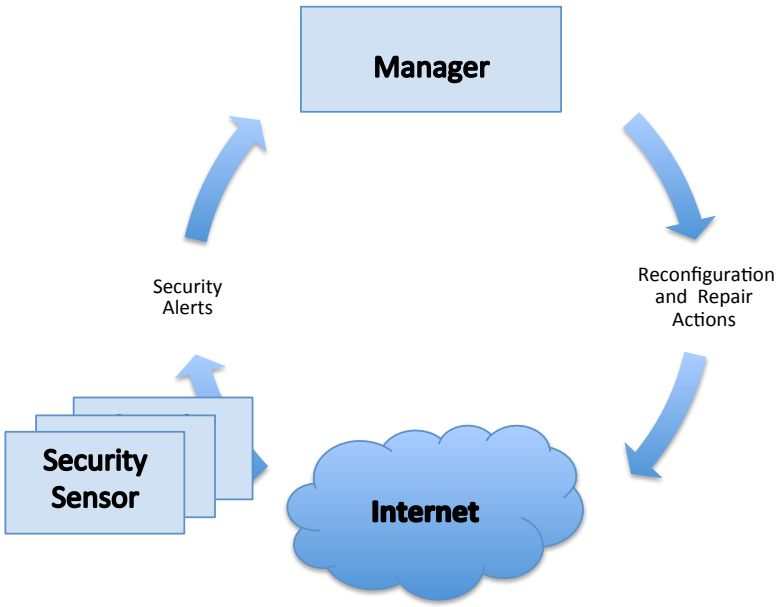


Figure 4: High-level monitoring architecture for security.

4 Monitoring Architecture for Security

In this and the following Sections 5–8, the reader will find more detailed information on the status of several research activities carried out during the first year of the project (see Table 3). The topic of the current section is the design of a monitoring architecture for security.

In the past, most Intrusion Detection Systems (IDSs) have been signature-based, inspecting (the payload of) every packet on a link, looking for predefined, malicious patterns. They have also mostly been developed as single observation points, both performing the collection and the analysis of the relevant data. Although this approach is, in some cases, still feasible from a technical perspective, it requires expensive hardware for dealing with high-speed links in backbone networks, for example. Our point of view with respect to this is that the monitoring and analysis of complex and highly dynamical systems, as the Internet, benefit from the combined input of multiple, distributed and dedicated observation points. We depict this at high-level in Figure 4. These observation points, or sensors, might be Intrusion Detection Systems, spam filters, firewall logs and network traffic information such as network flows and DNS traffic. Each sensor would be specialized on the detection of a specific attack, or on the analysis of a specific type of traffic. As indicated in Figure 4, since we are now dealing with heterogeneous sensors, there is the need of processing the alerts in an intelligent manner. For this reason, we have introduced a *Manager* component. The *Manager* will be in charge of collecting, processing and acting upon the input alerts it receives from the sensors. The *Manager* will also be able to reconfigure, activate and deactivate the sensors accordingly to the current security situation in the network. Compared to centralized solutions, the one we propose is a less expensive alternative that is extensible to several type of sensors.

In the following, we will describe how the high-level monitoring architecture for security can be adapted in the context of the research conducted in the FLAMINGO consortium. The architecture in Section 4.1 focuses on network measurements and intrusion detection. For the reconfiguration aspects that can also be associated with such architecture (e.g., network reconfiguration prompted by a security threat) , we refer the reader to D6.1.The FLAMINGO architecture itself is presented in Section 4.1. In such an architecture, intrusion detection based on network flows plays a key role. Export of network flows is nowadays supported on most high-end packet forwarding devices,

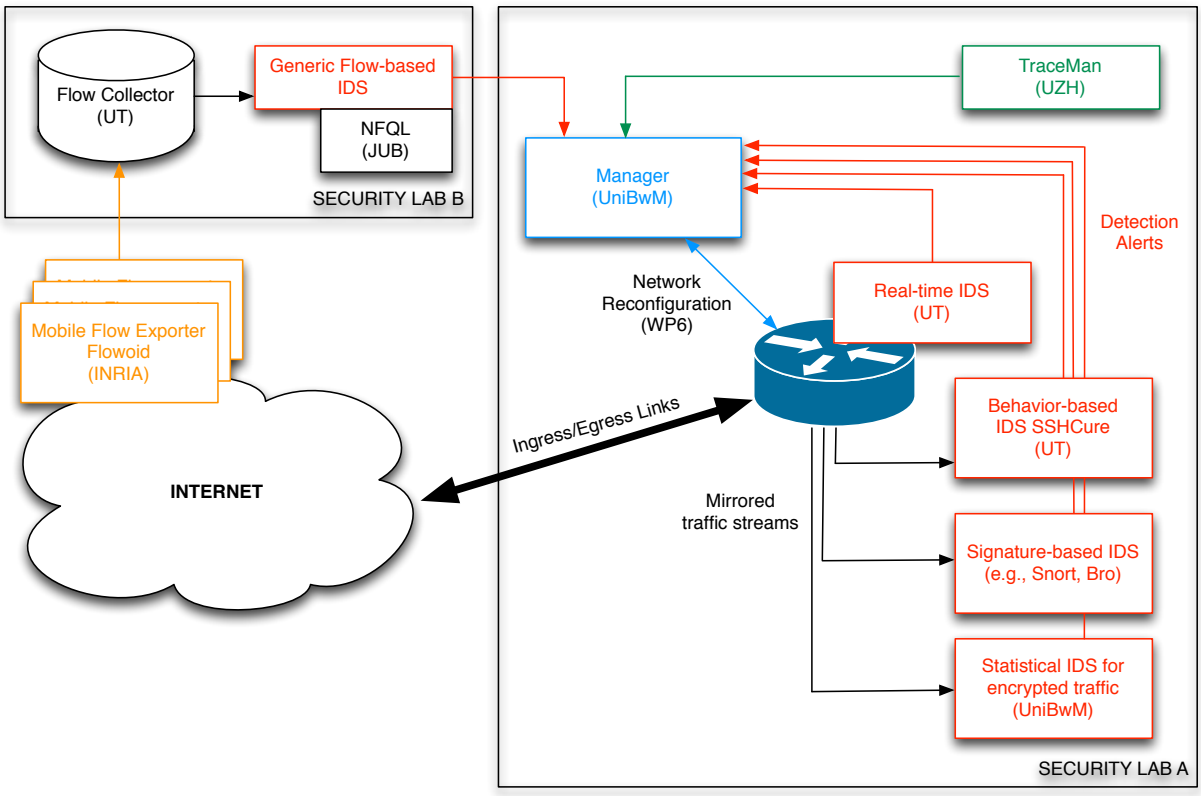


Figure 5: Flamingo flow-based monitoring architecture for security.

which allows for rapid and simple deployment. In Section 4.2, we discuss the various sensors that are and could be deployed to improve the likelihood of detecting attacks. It should be mentioned that the architecture is in an early state. Our approach is to bootstrap the current architecture with the ongoing research, as presented here, and extend it step by step with the research output of the FLAMINGO project, and the other participants in the NoE. This is an ongoing process that will last the whole lifetime of the project.

4.1 FLAMINGO Monitoring Architecture

The monitoring architecture presented in Figure 5 and discussed in this section combines several types of IDSs. First, a real-time, flow-based IDS is part of the router and scans *all* network traffic for intrusions and reports the results back to a manager. This IDS is able to perform a high-level scan on the full traffic stream to determine which traffic mixes require more attention, even in networks with throughputs of 10 Gbps and higher. Based on the detection results, a manager – a software process – decides which traffic mix is of interest for other IDSs, such as other (potentially more specific) flow-based IDSs or signature-based IDSs. By configuring the router that interconnects the infrastructure, the traffic streams to the other IDSs are pre-filtered. This has the following important advantage: there is no need anymore to invest in expensive sensors and monitoring systems that are able to cope with high-speed links and ever-increasing throughputs. Besides controlling the various traffic streams, the manager can also decide which IDS type is most efficient for analyzing the incident reported by the real-time IDS. In the end, as soon as the manager can be sure that a certain traffic mix is malicious, this traffic can be dropped by the router (e.g., by means of Access Control Lists (ACLs)). As such, the manager is the core of the proposed monitoring architectures,

and is able to control all IDSs and the firewall autonomously.

Besides the manager being the "brain" of the system and controlling the traffic streams, our architecture consists of the following components, shown in Figure 5:

- **Real-Time IDS** – This real-time intrusion detection, developed by UT and presented in [24], is part of the router. It is flow-based, and able to make a preliminary scan on all the network traffic passing through the router. As such, it can notify the manager about incidents or signs of potential incidents.
- **Sensors** – As soon as the manager of the system receives information about signs of (potential) incidents, it will select which sensor(s) is/are most appropriate for the particular incident. For example, in case the *Real-Time IDS* identifies a brute-force attack against an SSH daemon within the observed network, the manager can decide to activate *SSHCure*, a flow-based IDS for SSH (Section 5), which will analyze the pre-filtered SSH data in detail. In other situations where a signature-based IDS is more appropriate, for example, or traceroute measurements can be used to enhance the detection of the monitoring system, the respective sensors can be enabled.

In the next subsection, we will discuss the various sensor types in more detail.

4.2 Distributed Sensors

The architecture presented before has been designed such that any type of sensor can be added to it, to improve the manager's knowledge about the security state of the network. Compared to the *Real-time IDS*, the sensors perform on-demand, targeted analysis. So far, we have considered the following sensors:

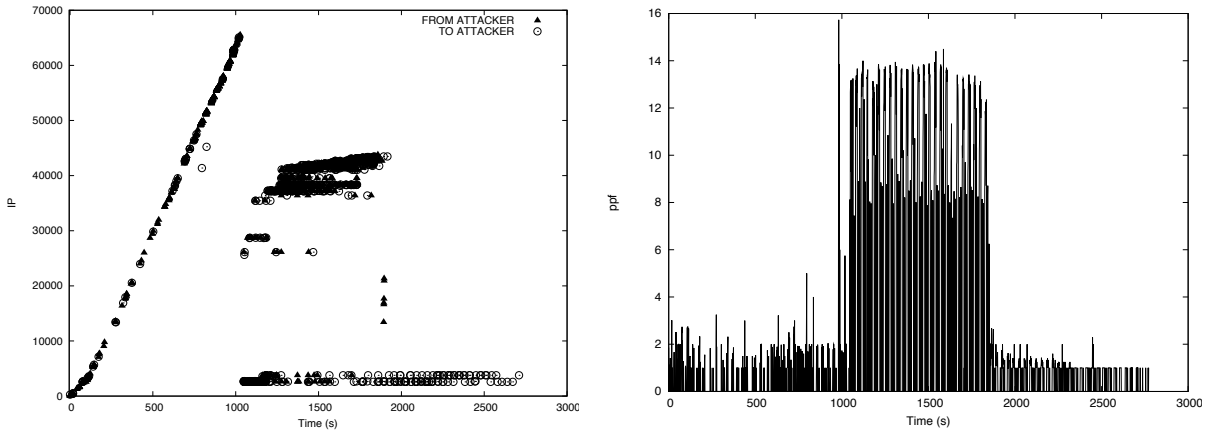
- **Behavior-based IDS SSHCure** (UT) – SSHCure is a flow-based IDS that has developed by UT and presented in [25]. It is able to detect SSH scans, brute-force attempts, and compromises in flow data, and fingerprint the tool used by the attackers. A detailed description on it is provided in Section 5.
- **Signature-based IDS** – Signature-based IDSs, or payload-based IDS, have been developed for years and open-source variants are considered in our architecture. Well-known examples are Snort[11] and Bro[12] [26]. Due to their intensive detection engines, they require expensive hardware to be operable in high-speed networks. By pre-filtering the data, as is the case in our architecture, regular machines can be run to still be able to use these IDSs.
- **Statistical IDS for encrypted traffic** (UniBwM) – UniBwM has an IDS under development that is capable of detecting intrusions in encrypted traffic [27, 28, 29]. Given the nature of encrypted traffic, this IDS is behavior-based, just like SSHCure.
- **TraceMan** (UZH) – TraceMan is an active measurement tool designed by UZH for collecting traceroute measurements. These measurements can be used to detect routing problems or network performance problems, for example. In the context of our monitoring architecture for security, it can be used to confirm the presence of large network attacks that affect the performance of the network such that it can be measured in terms of increased delays.

[11] http://www.snort.org/
[12] http://www.bro.org/

- **Mobile Flow Exporter Flowoid** (INRIA) – Flowoid is a flow exporter specifically designed by INRIA for operation on mobile devices. It can be executed on any Android device and provides a means for exporting network traffic without interference of a mobile network operator. Mobile devices, which are distributed by nature, export flow data to a flow collector designed by UT, on which a generic flow-based IDS can be deployed. Also NFQL, a flow query language designed by JUB, can be used by that IDS to query the data on the flow collector (Section 7).

Since the architecture is modular with respect of the deployed sensor types, it can be easily extended with other sensors. Sensors that we have in mind so far are those for specific types of (popular) attacks. They could be activated once the manager is informed by the *Real-Time IDS* about the presence of those attacks. In the future, we plan to extend the current architecture such as to include the knowledge generated by FLAMINGO partners with respect to monitoring, detection and fingerprinting of network attacks, and we will evaluate the possibility of including also domain-specific networks, such as, for example in the context of the IoT.

In the following section, we will detail one of the proposed sensors, namely *SSHCure*.



(a) Time-series for IP addresses (b) Time-series for packets-per-flow (PPF)

Figure 6: Traffic generated by SSH attack, from [30].

5 Intrusion Detection and Fingerprinting

The presence of network attacks on the Internet are a main area of concern for network managers and security teams. Some attacks are small and harmless, and can therefore be considered 'background noise' in the network. Others have a larger impact, as they result in blocked network connectivity or compromise of one or more systems. It is therefore crucial to have monitoring systems in place that warn in case of such an event or intrusion.

At small-scale, intrusion detection systems (IDSs) can be deployed on end hosts, where they have access to and monitor the hosts' resources. However, in large-scale deployments, such an approach is management-wise unfeasible; The average campus network, for example, has up to 65k hosts connected to the network, which makes it time-consuming and complex to maintain IDSs on all these hosts. A promising approach that overcomes this limitation is the use of network-based and network-wide IDSs. These IDSs observe the whole network from a particular *observation point*. More precisely, in order to cope with the ever-increasing link capacities, it is common to consider only summaries of the network traffic – commonly referred to as traffic 'flows'. The fact that only flow information is available, and therefore no payload, enforces the use of network behavior analysis (NBA, i.e., identify anomalous behavior from benign behavior), rather than signature-based detection, in which an IDS tries to find predefined patterns. In addition, the fact that the amount of encrypted data is increasing makes signature-based detection impossible.

A typical example of a protocol that uses encryption and therefore requires the use of behavior-based intrusion detection, is Secure SHell (SSH). By means of SSH, a hacker can gain access to and potentially full control over remote hosts. Once compromised, a hacker can sabotage not only the host itself, but also use it for attacking other systems. The detection of intrusions, especially in the case of SSH, is therefore crucial for preventing damage to hosts and networks. Previous work has demonstrated and validated the detection of SSH dictionary attack by means of flow data [30]. This work classifies SSH attacks into three phases:

Scan phase – An attacker scans an IP address block in order to find hosts running an SSH daemon (TCP port 22).

Brute-force phase – An attacker tries to login to a small subset of the scanned hosts, using a large number of username/password combinations.

Die-off phase – After a successful login, there is still traffic between the attacker and the compromised target. This residual traffic is due to commands being executed by the attacker on the target host.

The fact that the three phases identified in [30] exist in practice is shown in Figure 6a, which shows an SSH attack in a campus network with roughly 65k IPv4 addresses. The *scan* phase takes place from the beginning of the attack until $t \approx 1000s$, immediately followed by the *brute-force* phase, which terminates at $t \approx 1750s$. Finally, residual traffic, part of the *die-off* phase, is present on the network until $t \approx 2750s$ after the beginning of the attack. The fact that this pattern can be very well detected in a flow-based (or behavior-based) manner is demonstrated in Figure 6b: the three attack phases show a clearly distinguishable pattern in terms of packets per flow (in the case of SSH this refers to the number of exchanged packets between attacker and target within a single connection). This characteristic can be used to identify attacks in a flow dataset.

Based on the work presented in [30], we have implemented a flow-based IDS: SSHCure[13] [25]. SSHCure builds upon the foundations of the popular open-source flow collection framework NfSen[14]. It is able to detect SSH dictionary attacks in semi-real-time and provides a network-wide overview of intrusions in monitored networks. In addition, it is the only open-source IDS that is able to detect whether SSH attacks have been successful or not. This is a very important aspect for network managers and security teams, as they receive many alerts and warnings per day, often false ones. By notifying them about successful attacks, they can focus on those incidents that require particular attention, i.e., those that can damage the network and other hosts.

Many different SSH attack tools are available in the Internet. SSHCure is not only able to detect the attack, but it also identify the specific tool fingerprint. This clearly shows the advantage of behavior-based attacks: attacks by unknown attack tools can still be detected, potentially without being able to fingerprint the particular tool, i.e., provide the tool's name.

A screenshot of SSHCure's *Dashboard*, i.e., its main page on which network managers and security teams can find a summary of all SSH attacks in their network, is shown in Figure 7. More precisely, the screenshot shows the state of the UT campus network over one week, together with statistics on the most frequent attackers and targets. Please note that for privacy reasons the original IP addresses have been removed by the screenshot. Given the number of incidents, it is clear that an IDS is an important tool for monitoring the network. The UT campus network is in this respect not an exception; We have learned that SSHCure is being deployed in an increasing number of networks, both national and international, ranging from small-business networks to backbones.

The research summarised in this section can be found in the publications [25, 30]. SSHCure is one of the open-source tools developed in collaboration with WP1.

[13] `http://sshcure.sf.net/`
[14] `http://nfsen.sf.net/`

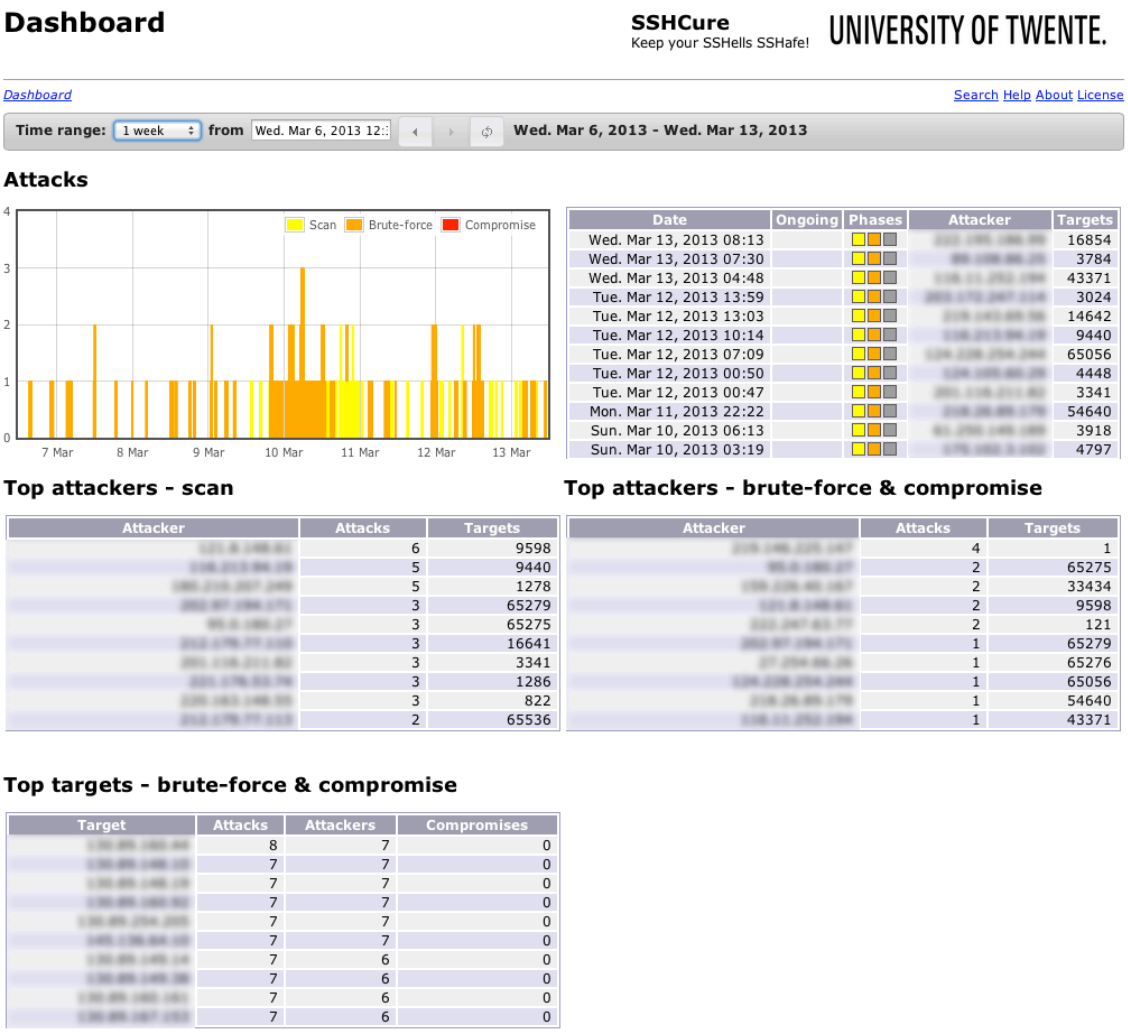


Figure 7: Screenshot of SSHCure, where IP addresses have been removed for privacy reasons.

6 Data Collection

Among the FLAMINGO partners, several data collection activities are taking place. In this section, we provide an overview of the type of traces that are being collected in the context of the Joint Security Lab (Section 6.1). We then focus on the *Internet traffic statistics*, a platform for the collection and publication of periodical traffic reports from major network operators worldwide (Section 6.2).

6.1 Data Collection and Joint Security Lab

Several data collection activities are taking place in the context of the Joint Security Lab. In line with the goals of the Joint Security Lab, such measurements are not made public but they can be shared with the consortium partners upon request and typically when a researcher is visiting an institution part of the Joint Security Lab. We report here the main measurements that are in place within the Joint Security Lab (Table 5) and we refer the reader to Deliverable 1.1 for a detailed description of the infrastructure.

Table 5: Traffic monitoring activities within the Joint Security Lab.

Traces	UT	INRIA	UniBwM
NetFlow/IPFIX	X	X	X
sFlow	X		X
SNMP	X		
Honeypots data	X	X	X
DNS	X		
IDS alerts		X	X
Blacklists		X	
Network Telescope			X
Server Logfiles			X

6.2 Internet Traffic Statistics

Internet traffic statistics can provide valuable information to network analysts and researchers about the way networks are used nowadays. In the past, such information was provided by Internet2 in a public website called *Internet2 NetFlow: Weekly Reports*. The website reported traffic statistics from the Abilene network on a weekly basis, for the period starting in the early 2000s to April 2010. The network connected 230 research institutes with a 10Gb/s link. Although these reports were limited to the behavior of the Abilene traffic, the *Internet2 NetFlow: Weekly Reports* had become a reference point for researchers because they were one of the few institutions to provide long term measurement data relative to a large backbone. Unfortunately, Internet2 discontinued Weekly Reports in April 2010 and, to the best of our knowledge, there is no information of this kind available to the public nowadays. The FLAMINGO consortium, and the University of Twente in particular, has seen in this situation the occasion for proposing an advanced and improved version of the original weekly reports. Our goal is to have long term aggregated (flow-based) measurements from major ISPs in the world, in particular National Research and Education Networks (NRENs). We believe that having such long terms measurements from several backbone can be of interest for both researchers as well as operators and it is also a service for the network management community worldwide.

The *Internet Traffic Statistics* is an ongoing effort carried on at the University of Twente. The following steps have been taken:

- **System architecture and data acquisition procedure -** The UT has designed the system architecture and the data acquisition procedure that will form the basis of the *Internet Traffic Statistics* framework. The system architecture is composed by 4 main blocks as shown in Figure 8. The first block is the actual traffic measurement that is performed at the operators side. The system architecture relies on traffic measurements at the flow level, by means of technologies such as Cisco's NetFlow or IPFIX-based flow probes. These types of probes are typically already in place in large ISPs, that use them for monitoring their network. The flow measurements are the input for the next block of the architecture that calculates statistics from measurements using a simple and open source Python script. Statistics are then stored in a JSON file. The second block can be located either at the operators side or at the UT side. The advantage of having it at the UT server is that it gives UT full control over the procedure, which can, for example, be repeated if any inconsistency is found. However, this is not a scalable solution when considering that the *Internet Traffic Statistics* wants to include data from many different operators. Therefore, the script can also run at the operator server, process the flow data and generate statistics at the operator side; then UT can retrieve

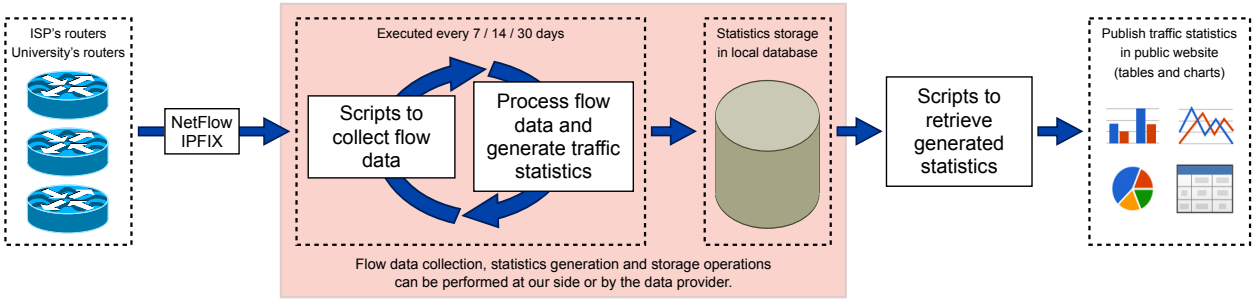


Figure 8: Internet Traffic Statistics: data collection procedure.

the JSON file only and collect the pre-processed data. In the third block, the JSON file is processed and all the information is stored in a MySQL database. Finally, the fourth block is responsible for presenting the data to the end user. The user can request statistics from specific periods via the website. A request is done to the database and data is presented in the format of tables and plots. In addition, the user has the option to download the requested statistics data in plain text.

- **User interface design -** The UT has designed and implemented a preliminary version of the user interface for the *Internet Traffic Statistics*. The statistics will be available by means of a web interface. A prototype version of the interface is visible at `http://stats.simpleweb.org`.

- **Contact with operators -** The success of the *Internet Traffic Statistics* relies on the support of network operators worldwide. UT has been in contact with several network operators, and in particular with NRENs. Several NRENs have welcomed the idea of the *Internet Traffic Statistics* and they offered helpful input in defining which traffic metrics can be at the same time (i) useful for operators and the scientific community and (ii) can be openly released without privacy concerns. Currently, the *Internet Traffic Statistics* has the support of the following NRENs, which will start providing data as soon as the data acquisition procedure is finalized:

  - SURFnet, the Dutch Research and Education Network[15]
  - CESNET, the Czech Education and Scientific Network[16]
  - GÉANT, the pan-European research and education network that interconnects Europe's National Research and Education Networks (NRENs)[17]
  - DeiC, Danish e-infrastructure cooperation[18]
  - RNP, the Rede Nacional de Ensino e Pesquisa (Brasilian Education and Scientific Network)[19]

During the second year of the project, UT will consolidate the data acquisition and processing workflow, and it will automatise the data collection from the interested partners. Currently, RNP has made available a test dataset, while other operators, DeiC and CESNET in particular, are testing the data pre-processing script in their networks. We expect the post-processed data, i.e., the statistics, to amount to 90-100 MB per week per partners.

[15] http://www.surfnet.nl
[16] http://www.cesnet.cz
[17] http://www.geant.net/
[18] http://www.deic.dk/
[19] https://www.rnp.br/

7 Flow query language

Cisco's NetFlow protocol and IETF's IPFIX open standard are the widely deployed techniques for collecting network flow statistics. Understanding certain patterns in these network statistics requires sophisticated flow analysis tools that can efficiently mine flow records. Network Flow Query Language (NFQL), recently proposed in [12] and also presented at the *5th Workshop on the Usage of Netflow/IPFIX in Network Management*, can process flow records, aggregate them into groups, apply absolute or relative filters, and invoke Allen interval algebra rules to merge group records. In [12], the authors introduce an efficient implementation of the query language and they evaluate its performance with respect to a suite of benchmarks against contemporary flow-processing tools.

7.1 Shortcomings of Current Flow-Processing Tools

NFQL aims at overcoming the shortcomings common to several existing flow-processing tools. Such shortcomings are:

- **Limited filtering rules**. Tools such as ntop[20], FlowScan[21], NfSen[22] and Stager[23] offer filtering capabilities. However, the power of filtering rules in the aforementioned tools is mostly limited to absolute comparisons of flow attributes. As a result, relative comparison amongst different flows or querying for timing relationship among them is not possible.
- **Limited grouping and merging operations**. Similarly to NFQL, the tool SiLK offers, beside filtering, also grouping and merging operations. However the grouping and merging operations can only be performed using an equality operator. This restriction allows the tool to perform optimization such as using hash tables to perform lookups. NFQL on the other hand, provides a much richer set of comparisons operations, such as *equal*, *not equal*, *greater than*, *less than*, *greater or equal*, *less or equal*.

7.2 The NFQL Processing Pipeline

The NFQL Processing Pipeline, depicted in Figure 9, is based on the following steps:

- The **splitter** is responsible to duplicate and forwarding the input stream to several branches of the processing pipeline.
- The **filter** blocks perform absolute filter on the input flow-records data, by comparing separate fields of a flow record against provided values
- The **grouper** blocks aggregate the flow records they receive from the filter blocks. The output of a grouper is a group-flow.
- The **group-filter** block perform filtering operations on group-flows. The group-flows that satisfy the filter are passed to the merger block.

[20] http://www.ntop.org/
[21] http://www.caida.org/tools/utilities/flowscan/
[22] http://nfsen.sourceforge.net/
[23] http://freecode.com/projects/stager

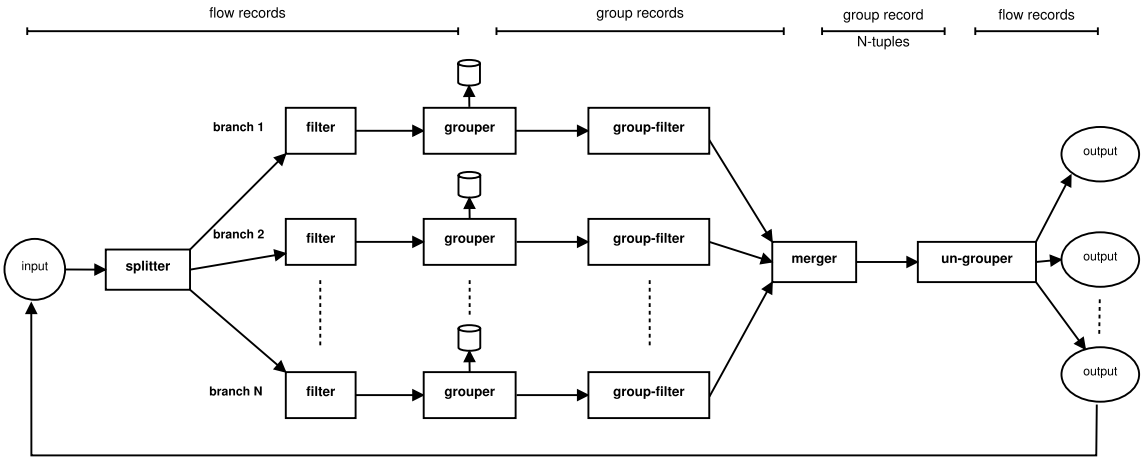


Figure 9: NFQL pipeline.

- The **merger** block performs relative filtering on the N-tuples of groups formed from the N streams passed on from the group-filters as input.
- Finally, the **ungrouper** expands group-flows to their respective individual flows for presenting them to the user.

7.3 Evaluation

Figure 10 shows an example of the performance of NFQL with respect to existing flow-processing tools. The test focused on the filter stage, since this is present in all the considered flow-processing tools. The tools have been tested using 20M flow records from the publicly available Trace 7 at the SimpleWeb [24] repository. The input trace was compressed at ZLIB_LEVEL 5 using the zlib suite. It was also converted to nfdump and SiLK compatible formats keeping the same compression level. The suite was run on a machine with 24 cores of 2.5 GHz clock speed and 18 GiB of memory.

It can be seen that the performance of the filter stage in NFQL is comparable to that of flow-tools and SiLK. SiLK takes less time on lower ratios, probably due to the fact that SilK, as well as nfdump, also use their own file format. nfdump appears to be significantly faster than the rest. This is because nfdump uses the lzo compression scheme [31]. It is therefore clear that adding lzo compression will likely improve NFQLs filter performance.

For further details on the research summarised in this section, we refer the reader to the FLAMINGO publication [12]. NFQL is one of the open-source tools developed in collaboration with WP1.

[24] http://simpleweb.org/

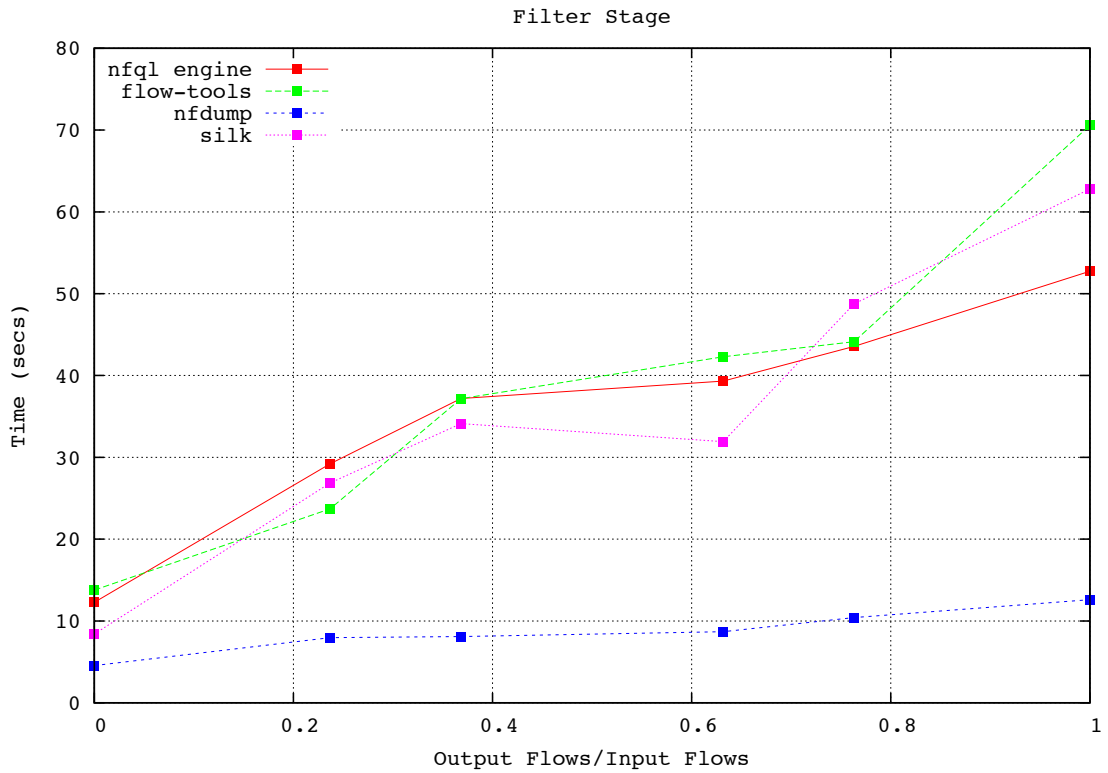


Figure 10: NFQL performance in comparison with other flow-processing tools.

8 Integration of European Research

This section reports on the International activities (Section 8.1) and the collaborations with other EU projects and institutions (Section 8.2) that have involved WP5.

8.1 International Activities

In collaboration with WP2, WP3 and WP4, the consortium has been actively involved in several international activities focussing on diverse aspects of network and service monitoring. We report here the ones that have been relevant, in topic and for the partners participation, to WP5, and we refer to the respective WP2, WP3 and WP4 deliverables for the general overview of these activities:

- The **5th Workshop on the Usage of NetFlow/IPFIX in Network Management** (30th NMRG meeting, July 30, 2013, Berlin), co-located with IETF 87. The workshop investigates how NetFlow/IPFIX is used in practice in various aspect of network monitoring and management. Also this year the workshop succeeded in bringing together researchers, operators and manufacturers to exchange their hands-on experience. The workshop was chaired by Ramin Sadre (Aalborg University) and Aiko Pras (University of Twente). The partners JUB, INRIA and UniBwM have actively participated to the workshop by presenting their research.
- The **1st Workshop on Large Scale Network Measurements** (31st NMRG meeting, October 14, 2013, Zürich), co-located with the 9th International Conference on Network and Service Management (CNSM 2013). The goal of the workshop is bringing together researchers and in particular PhD students working on topics related to large scale infrastructures for network

measurements. The workshop is organized by Jürgen Schönwälder in collaboration with the EU projects Leone[25] and mPlane[26]. The partners JUB, UT, and UZH participated to the workshop by presenting their research.

- The Dagstuhl seminar **Global Measurement Framework**[27] (Schloss Dagstuhl, November 17-20, 2013). The seminar will bring together researchers from industry, academia, and regulators across continents and across different backgrounds to discuss the state of the art in measurements and their exploitation, measurement and analysis techniques, privacy and anonymization. The seminar is organized, in collaboration with the EU projects Leone[25] and mPlane[26], by: Philip Eardley (BT Research, GB), Marco Mellia (Polytechnic University of Torino, IT), Jörg Ott (Aalto University, FI), Jürgen Schönwälder (Jacobs University Bremen, DE) and Henning Schulzrinne (Columbia University, US).

- V. Bajpai and N. Melnikov (JUB) have given a tutorial at the 7th International Conference on Autonomous Infrastructure, Management and Security (AIMS 2013) on the topic **Large-scale Measurement Platforms**.

- J. Schönwälder, A. Sehgal (JUB) have given a tutorial at the IFIP/IEEE International Symposium on Integrated Network Management (IM 2013), Ghent, Belgium, May 2013, on the topic **Management of the Internet of Things**.

- Martin Waldburger (UZH) and Anna Sperotto (UT) have taken part in the organization of the **7th International Conference on Autonomous Infrastructure, Management and Security** (AIMS 2013, Barcelona, Spain), as Conference Co-chair and PhD Student Workshop Co-chair, respectively. In addition, Joan Serrat (UPC) has been general chair of the conference and several members of the consortium have acted as TPC members.

8.2 Collaborations with Other EU Projects and Institutions

Given the focus of WP5 on network and service measurements, FLAMINGO has also collaborated with other EU project that are active on the topic of measurements. In particular, FLAMINGO collaborated with:

- The FP7 Project Leone[25] – From global measurements to local management (grant no. 317647).

- The FP7 Project mPlane[26] – Building an Intelligent Measurement Plane for the Internet (grant no. 318627).

- The FP7 Project UniverSelf[28] – Realizing autonomics for Future Networks (grant no. 257513).

- The FP7 Project SmartenIT[29] – Socially-aware Management of New Overlay Application Traffic combined with Energy Efficiency in the Internet (grant no. 317846).

FLAMINGO, Leone and mPlane contributed to the organization of the **1st Workshop on Large Scale Network Measurements** and of the Dagstuhl seminar **Global Measurement Framework**

[25] http://www.leone-project.eu/
[26] http://www.ict-mplane.eu/
[27] http://www.dagstuhl.de/13472
[28] http://www.univerself-project.eu/
[29] http://www.smartenit.eu/

(see Section 8.1). The collaborations with other EU projects also had as an outcome several published and submitted papers, as highlighted in Section 2.1.

Finally, FLAMINGO is collaborating with the **EIT ICT Labs**[30]. Rick Hofstede (UT) and Abdelkader Lahmadi (INRIA) have been researching on how to integrate geographical information into network flows. An Internet Draft with title “Information Elements for IPFIX Metering Process Location” has been submitted as result of this collaboration [7].

[30] http://www.eitictlabs.eu/

9 Conclusions

This deliverable describes WP5 achievements in the field of network and service monitoring. After the first year, WP5 has met the S.M.A.R.T. objectives (regarding the integration of PhD students and the scientific output) and research is actively ongoing in all the research areas related to the WP5-specific objectives. Two of the WP5-specific objectives have been planned for the second year of the project, and for those the project members have identified a clear plan on how to address them in their future work. The scientific outcome of the WP, both relatively to the number of publications as well as the quality of the venues, is outstanding.

With respect to the activities relative to this WP, FLAMINGO has demonstrated to take an active position in the network management community. This is proven by activities that have potentially high impact on the community, such as the described open-source tools and the *Internet Traffic Statistics*. WP5 researchers have been involved in several measurements and monitoring activities at European level, and WP5 has established collaborations with several EU projects. These activities are witness by a Dagsthul seminar, two NMRG meetings and workshops, as well as several tutorials and conferences organisation.

The PhD collaborations are without any doubt one of the major achievement for the first year. Although on paper it is clear that there are advantages in collaborating with other institutions, in practice this type of activities may often become secondary when dealing with everyday occupations. In the context of this project, this is not the case. We have been very positively surprised because the project has exceeded the our initial expectations in terms of the numer of collaborations. Moreover, the PhD students are very active in this respect. In year two, we expect the collaborations to reach a further degree of maturation, which will ensure also an excellent research output. It is therefore clear that WP5 will make sure that existing and possibly new PhD collaborations play a major role also in year two.

Abbreviations

6LoWPAN	IPv6 over Low power Wireless Personal Area Networks
ACL	Access Control List
API	Application Programming Interface
CDN	Content Delivery Network
CPS	Cyber Physical Systems
CP	Content Provider
DDoS	Distributed Denial of Service attack
DNP	DDoS Network Protections
DODAG	Destination Oriented Directed Acyclic Graph
EIT	European Institute of Technology
Gbps	Giga bits per second
HAS	HTTP Adaptive Streaming
HTTP	Hyper-text Transfer Protocol
ICT	Information and Communications Technology
IDS	Intrusion Detection System
ILP	Integer Linear Program
IPFIX	Internet Protocol Flow Information Export
ISP	Internet Service Provider
InP	infrastructure Provider
JSON	JavaScript Object Notation
LLN	Low-power and Lossy Networks
MNO	Mobile Network Operator
NFQL	Network Flow Query Language
NMRG	Network Management Research Group
NREN	National Research and Education Network
QoS	Quality-of-Service
ROLL	Routing Over Low Power Lossy networks
RPL	Routing Protocol for Low power and Lossy Networks
S.M.A.R.T	Specific Measurable Achievable Relevant Timely

SCADA Supervisory Control and Data Acquisition

SLA Service Level Agreement

SNMP Simple Network Management Protocol

SP Service Provider

SSH Secure SHell

TCP Transmission Control Protocol

TPM Trusted Platform Module

UDP User Datagram Protocol

VNP Virtual Network Provider

VoIP Voice-over-IP

VoS Value of Service

Acknowledgments

This deliverable is based on input from the WP5 Partners of the FLAMINGO consortium. A particolar acknowledgement goes to all the PhD students that have not only provided textual input, but that are working on a daily basis on the challenging research topics that we report.

A Overview of PhD Collaborations and Objectives

Tables 6 – 9 in this Appendix provide an overview of the PhD collaborations established during the first year of the FLAMINGO project. The tables establish an acronym for each collaboration and they summarize the WP-specific objectives that a collaboration contributes to, for WP5, WP6 and WP7.

Collaboration Name	Acronym	Involved PhDs	WP5 Objectives	WP6 Objectives	WP7 Objectives
Linking Network Usage Patterns to Traffic Gaussianity Fit	JUB-UT-Pattern	• Nikolay Melnikov (JUB) • Ricardo Schmidt (UT)	• To collect (anonymized) monitoring data	• To develop information models, algorithms, learning techniques and knowledge description approaches as enablers for automated configuration and repair	
Energy-aware Traffic Management	UCL-UT-Man	• Daphne Tuncer (UCL) • Ricardo Schmidt (UT)	• To collect (anonymized) monitoring data	• To evaluate automated configuration and repair approaches as being part of the autonomic control loops	
Intrusion Detection Systems	UT-UniBwM-IDS	• Rick Hofstede (UT) • Mario Golling (UniBwM)	• To develop a generic distributed flow monitoring architecture	• To develop information models, algorithms, learning techniques and knowledge description approaches as enablers for automated configuration and repair	

Table 6: WP5-WP6-WP7 PhD collaborations - Part 1

Collaboration Name	Acronym	Involved PhDs	WP5 Objectives	WP6 Objectives	WP7 Objectives
Towards A Trust Comupting Architecture for RPL in Cyber Physical Systems	UniBwM-JUB-RPL	• Sebastian Sebeer (UniBwM) • Björn Stelte (UniBwM) • Anuj Sehgal (JUB)		• To specify guidelines about the applicability of approaches for automated configuration and repair to specific application domains • To apply the developed approaches to several application domains such as of (i) content-aware networking, (ii) cloud-based services and (iii) wireless sensor networks	
Flowoid: a NetFlow/IPFIX probe for Android-based devices	UT-INRIA-Flowoid	• Rick Hofstede (UT) • Abdelkader Lahmadi (INRIA)	• To collect (anonymized) monitoring data		
Flow-based Traffic Measurements for In-Network Video Quality Adaptation	iMinds-UT-QoS	• Niels Bouten (iMinds) • Ricardo Schmidt (UT)		• To develop information models, algorithms, learning techniques and knowledge description approaches as enablers for automated configuration and repair	

Table 7: WP5-WP6-WP7 PhD collaborations - Part 2

Collaboration Name	Acronym	Involved PhDs	WP5 Objectives	WP6 Objectives	WP7 Objectives
Study of DAG Inconsistency Attacks in RPL Networks	INRIA-JUB-RPL	• Anuj Sehgal (JUB) • Anthéa Mayzaud (INRIA)	• To propose novel solutions for intrusion detection and fingerprinting.	• To apply the developed approaches to several application domains such as of (i) content-aware networking, (ii) cloud-based services and (iii) wireless sensor networks. • To develop an inventory of approaches for automated configuration and repair	
SLA Fulfillment Mechanism	UZH-UniBwM-SLA	• Christos Tsiaras (UZH) • Sebastian Sebeer (UniBwM) • Daniel Dönni (UZH)	• To collect (anonymized) monitoring data		• Propose business and regulation actions in case of SLA violation identification
Cache Management	UCL-iMinds-Cache	• Daphne Tuncer (UCL) • Maxim Claeys (iMinds)	• To collect (anonymized) monitoring data	• To evaluate automated configuration and repair approaches as being part of the autonomic control loops	

Table 8: WP5-WP6-WP7 PhD collaborations - Part 3

Collaboration Name	Acronym	Involved PhDs	WP5 Objectives	WP6 Objectives	WP7 Objectives
Management of Virtualized Networks	iMinds-UPC-NetVirt	• Maxim Claeys (iMinds) • Rashid Mijumbi (UPC)		• To develop new architectures for automated configuration and repair approaches across administrative boundaries • To develop information models, algorithms, learning techniques and knowledge description approaches as enablers for automated configuration and repair	
DDoS Network Protection	UT-UZH-DoS	• Jair Santanna (UT) • Guilherme Sperb Machado (UZH)	• To collect (anonymized) monitoring data • To create annotated traces to assess the quality of different Intrusion Detection Systems (IDS) • To investigate the applicability of different AI and machine learning techniques for flow analysis • To propose novel solutions for intrusion detection and fingerprinting		

Table 9: WP5-WP6-WP7 PhD collaborations - Part 4

B Internet Traffic Statistics

This appendix includes the poster, by M. Hoogesteger, R. Schmidt, A. Sperotto and A. Pras, that has been presented at the Terena Networking Conference 2013 [32].

This poster presents the project for recreating an extended version of the Internets traffic weekly reports, which we name the *Internet Traffic Statistics*. The main idea of this work is to collect data from different sources around the world, generate network traffic statistics and make them available in a public repository with a friendly user interface. The aim is to have continuous traffic reports that, after a long period of data collection, will allow us to understand the evolution of Internet traffic. Traffic data will be collected at, for example, Internet Service Providers, NRENs and universities, allowing us to have a broad picture of the Inernet usage. Traffic statistics will be generated from flow data (e.g., NetFlow). The poster illustrates the entire process from traffic data collection to statistics generation and presentation. With this project, we aim to once again provide such valuable information to network engineers, analysts and research community.

Reports on Internet Traffic Statistics

Martijn Hoogesteger, Ricardo de O. Schmidt, Anna Sperotto and Aiko Pras
Design and Analysis of Communication Systems, University of Twente, The Netherlands
m.hoogesteger@student.utwente.nl, r.schmidt@utwente.nl

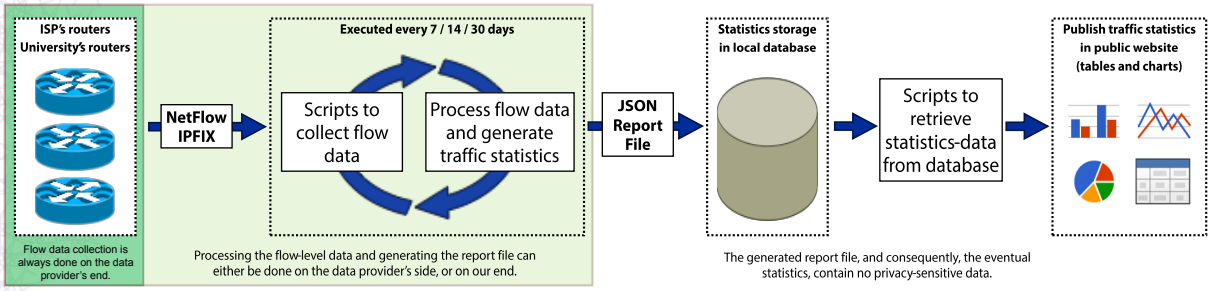
1. Motivation

- Internet2's Weekly Report served as a very important public source of information for network analysts and researches; however, it was discontinued in April 2010
- There is no other public repository that provides such valuable statistics about Internet traffic
- Our goal is to recreate these Internet reports using more data from networks around the globe, providing an even broader view of Internet usage

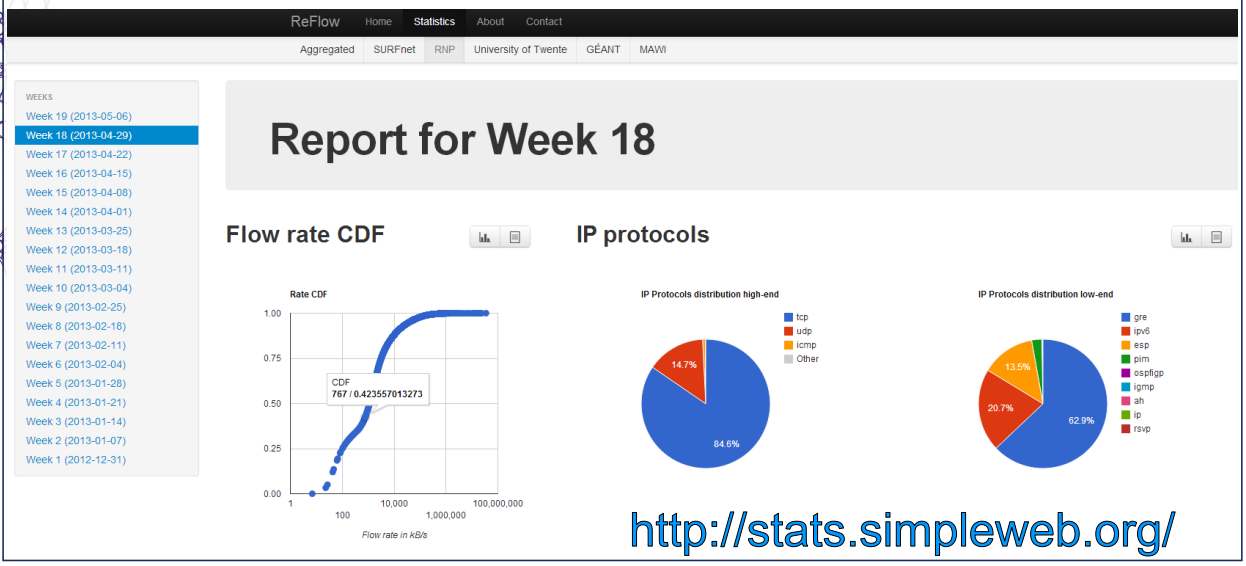
2. Internet Traffic Reports

- Traffic data is collected at observation points in different networks
- The data comprises **flow-level traffic measurements** in the format of **NetFlow / IPFIX**
- Flow data is collected and processed into **reports on statistics**.
- Reports are collected, stored and eventually **presented in a public website**, providing several visualizations of the statistics.

3. The Complete Procedure



4. Website and User Interface



http://stats.simpleweb.org/

Acknowledgements

SURFNET, RNP, UNIVERSELF, GÉANT, FLAMINGO, CESNET

CTIT

UNIVERSITY OF TWENTE.

References

[1] I. Drago, E. Bocchi, M. Mellia, H. Slatman, and A. Pras. Benchmarking Personal Cloud Storage. In *ACM/SIGCOMM Internet Measurements Conference 2013 (IMC 2013)*, 2013.

[2] M. Barrere, R. Badonnel, and O. Festor. Vulnerability Assessment in Autonomic Networks and Services: A Survey. *Communications Surveys & Tutorials, IEEE*, 2013.

[3] A. Lareida, T. Bocek, M. Waldburger, and B. Stiller. RB-tracker: A fully distributed, replicating, network-, and topology-aware P2P CDN. In *Integrated Network Management (IM 2013), 2013 IFIP/IEEE International Symposium on*, pages 1199–1202, 2013.

[4] G. Sperb Machado, T. Bocek, M. Ammann, and B. Stiller. A Cloud Storage Overlay to Aggregate Heterogeneous Cloud Services. In *Proc. of the 38th IEEE Conference on Local Computer Networks (LCN 2013)*, Oct. 2013.

[5] P. Poullie and B. Stiller. Fair Allocation of Multiple Resources Using a Non-monetary Allocation Mechanism. In *Proc. 7th International Conference on Autonomous Infrastructure, Management and Security (AIMS 2013)*. 2013.

[6] C. Schmitt, B. Stiller, T. Kothmayr, and W. Hu. DTLS-based Security with two-way Authentication for IoT (Internet Draft). http://tools.ietf.org/html/draft-schmitt-two-way-authentication-for-iot-01, October 2013.

[7] O. Festor, A. Lahmadi, R. Hofstede, and A. Pras. Information Elements for IPFIX Metering Process Location (Internet Draft). http://tools.ietf.org/html/draft-festor-ipfix-metering-process-location-01, July 2013.

[8] S. Seeber, A. Sehgal, B. Stelte, G. Dreo, and J. Schönwälder. A Trust Computing Architecture for RPL in Cyber Physical Systems. In *Proc. of the 9th International Conference on Network and Service Management (CNSM 2013)*, Oct 2013.

[9] R. de O. Schmidt, N. Melnikov, R. Sadre, J. Schönwälder, and A. Pras. Linking network usage patterns to traffic gaussianity fit. In *PAM 2014 (submitted to)*.

[10] R. Mijumbi, J. Serrat, J.L. Gorricho, M. Claeys, F. De Turck, and S. Latré. Design and Evaluation of Learning Algorithms for Dynamic Resource Management in Virtual Networks. In *14th IEEE/IFIP Network Operations and Management Symposium (NOMS 2014) (submitted to)*.

[11] A. Mayzaud, A. Sehgal, R. Badonnel, I. Chrisment, and J. Schönwälder. Mitigating DODAG Inconsistency Attacks in RPL Networks. In *ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN 2014) (submitted to)*.

[12] V. Bajpai, J. Schauer, and J. Schönwälder. NFQL: A Tool for Querying Network Flow Records. In *Proc. of the 13th IFIP/IEEE Symposium on Integrated Network and Service Management (IM 2013)*, May 2013.

[13] M. Charalambides, D. Tuncer, L. Mamatas, and G. Pavlou. Energy-aware adaptive network resource management. In *IFIP/IEEE International Symposium on Integrated Network Management (IM 2013)*, May 2013.

[14] R. de O. Schmidt, R. Sadre, A. Sperotto, and A. Pras. Lightweight Link Dimensioning using sFlow Sampling. In *Proc. of the 9th International Conference on Network and Service Management (CNSM 2013)*, Oct. 2013.

[15] K. Scarfone and P. Mell. Guide to Intrusion Detection and Prevention Systems (IDPS). *NIST Special Publication*, 800(2007):94, 2007.

[16] M. Golling and B. Stelte. Requirements for a Future EWS-Cyber Defence in the Internet of the Future. In *Cyber Conflict (ICCC), 2011 3rd International Conference on*, 2011.

[17] GÉANT. Breakthrough GÉANT Network Marks Ten Years of Success: High Bandwidth pan-European Research Network Continues Advances with 100 Gbps Plans . `TenYearsOfSuccess`, November 2010.

[18] R. Koch, M. Golling, and G. Dreo Rodosek. Evaluation of State of the Art IDS Message Exchange Protocols. In *International Conference on Communication and Network Security (CNS 2013)*, 2013.

[19] G. Dreo, M. Golling, W. Hommel, and F. Tietze. ICEMAN: An architecture for secure federated inter-cloud identity management. In *IFIP/IEEE International Symposium on Integrated Network Management (IM 2013)*, May 2013.

[20] G. Dreo, M. Golling, and W. Hommel. MuSIC: An IT security architecture for inter-community clouds. In *IFIP/IEEE International Symposium on Integrated Network Management (IM 2013)*, May 2013.

[21] R. J. Hofstede and T. Fioreze. SURFmap: A Network Monitoring Tool Based on the Google Maps API. In *IFIP/IEEE International Symposium on Integrated Network Management (IM 2009)*, June 2009.

[22] C. Tsiaras and B. Stiller. Challenging the Monopoly of Mobile Termination Charges with an Auction-based Charging and User-centric System (AbaCUS). *NetSys 2013 - Networked Systems*, 2013.

[23] D. Tuncer, M. Charalambides, R. Landa, and G. Pavlou. More Control Over Network Resources: an ISP Caching Perspective. In *Proc. of the 9th International Conference on Network and Service Management (CNSM 2013)*, Oct. 2013.

[24] R. Hofstede, Bartoš V, A. Sperotto, and A. Pras. Towards Real-Time Intrusion Detection for NetFlow/IPFIX. In *Proc. of the 9th International Conference on Network and Service Management (CNSM 2013)*, May 2013.

[25] L. Hellemons, L. Hendriks, R. Hofstede, A. Sperotto, R. Sadre, and A. Pras. SSHCure: A Flow-Based SSH Intrusion Detection System. In *Proc. of the 6th IFIP WG 6.6 International Conference on Autonomous Infrastructure, Management, and Security, AIMS 2012*, pages 86–97. June 2012.

[26] V. Paxson. Bro: a system for detecting network intruders in real-time. *Computer Networks*, 31(23–24):2435–2463, 1999.

[27] R. Koch and G. Dreo. User identification in encrypted network communications. In *Proceedings of the 6th International Conference on Network and Service Management (CNSM 2013)*, Oct. 2010.

[28] R. Koch and G. Dreo. Command Evaluation in Encrypted Remote Sessions. In *Fourth International Conference on Network and System Security, (NSS 2010)*, Sept. 2010.

[29] R. Koch and G. Dreo. Security System for Encrypted Environments (S2E2). In *Recent Advances in Intrusion Detection, 13th International Symposium (RAID 2010)*, Sept. 2010.

[30] A. Sperotto, R. Sadre, P.-T. de Boer, and A. Pras. Hidden Markov Model modeling of SSH brute-force attacks. In *Proc. of the 20th IFIP/IEEE Int. Workshop on Distributed Systems: Operations and Management (DSOM '09)*, 2009.

[31] P. Deutsch and J-L. Gailly. ZLIB Compressed Data Format Specification version 3.3. RFC 1950 (Informational), May 1996.

[32] M. Hoogesteger, R. de O. Schmidt, A. Sperotto, and A. Pras. Internets Traffic Statistics Reports. TERENA Networking Conference 2013 (TNC 2013), Poster, June 2013. https://tnc2013.terena.org/core/poster/20.