MASTER THESIS

Estimating Inter-domain Network Delays by Passive Monitoring

prepared for the
Salzburg University of Applied Sciences
Degree Program
Information Technology & Systems Management

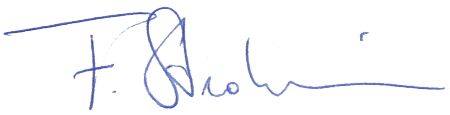
submitted by:
DI(FH) Felix Strohmeier

Head of Faculty:	FH-Prof. DI Dr. Gerhard Jöchtl
Supervisor:	Prof. Dr. (habil) Ulrich Hofmann

Salzburg, September 2010

Affidavit

Herewith I, Dipl.-Ing.(FH) Felix Strohmeier, declare that I have written the present master thesis fully on my own and that I have not used any other sources apart from those given.



First Name Surname	0910581012 Registration Number

Acknowledgement

"You fool some people sometimes
But you can't fool all the people all the time"
Bob Marley (1945-1981)
and Peter Tosh (1944-1987)

My acknowledgement is threefold: First, my thanks go to the team of Salzburg Research and my colleagues from the PRISM project being distributed among European reseach labs, universities and companies. The common work on the topic of Internet monitoring and related privacy issues enabled me to get the knowledge required for writing this thesis. Second, I would like to thank my advisor, Ulrich Hofmann, for providing constructive input and his support in creating the work around the topic of inter-domain delay estimation. Finally and most important, I want to gratefully thank my family. My wife Birgit, for giving the full support during the year of the master studies, and my two sons, Benedikt and Gregor for accepting my absence on several weekends required for taking courses and during summer time for writing this thesis.

Details

First Name, Surname:	DI(FH) Felix Strohmeier
University:	Salzburg University of Applied Sciences
Degree Program:	Information Technology & Systems Management
Title of Thesis:	Estimating Inter-domain Network Delays by Passive Monitoring
Academic Supervisor:	Prof. Dr. (habil) Ulrich Hofmann

Keywords

1^{st}	Keyword:	Internet Monitoring
2^{nd}	Keyword:	Internet User Privacy
3^{rd}	Keyword:	Anonymisation
4^{th}	Keyword:	Network Delay Estimation
5^{th}	Keyword:	Internet Topology

Abstract

This master thesis presents an approach to apply purely passive, privacy-preserving network monitoring to a performance monitoring problem. Internet service providers and network operators need to know, which services their customers are consuming on which quality level. In addition, the customer must be enabled to relate the experienced quality to the agreed service level parameters. The application of passive network monitoring for such tasks is useful: First, it can evaluate the real traffic of the customers, without injecting additional traffic to the network and second, measurement results are exactly based on the services the customers are accessing. On the other hand, passive network monitoring always includes the risk of compromising the customers' right for privacy and therefore often raises privacy concerns from the public. Nevertheless, recent privacy-preserving monitoring frameworks allow overcoming the "privacy-versus-utility" dilemma by decoupling data capturing, filtering and data processing in separate steps, always applying the required level of privacy. This work evaluates delay measurements based on passive monitoring and presents different possibilities to estimate the connection quality to remote Internet locations. The mechanism exploits passive RTT measurement results gathered from inter-domain traffic to remote Internet hosts and aggregates them. They system only reports reduced but relevant information to the operator, while discarding unnecessary information for privacy reasons. The remote locations represent aggregations on different administrative or geographic precision levels, like BGP prefixes, Autonomous Systems or countries.

Contents

List of Figures

List of Tables

1

Introduction

The Internet has become the most important communication infrastructure in the last two decades and is the medium with the quickest growth rate ever. The highly dynamic changes and extensions in the contents provided, has also a major impact on the underlying network infrastructure. To manage this, network operators develop and deploy intelligent and informative network management tools, mainly within their own administrative domains only. The configuration in the global environment is a complex task because of its distributed fashion, and failures like routing misconfigurations can happen. Due to the meshed network structure, such failures will usually not result in a complete loss of reachability of some parts of the Internet, but new bottlenecks can occur, and delays and packet losses can increase without immediate notice. Even in correctly configured environments, if new services become popular within a short range of time, low performance can result in user dissatisfaction due to bad service quality.

For these and similar reasons monitoring and measurement of traffic and performance in the Internet is as old as the Internet itself. Advances in the monitoring technologies, but mainly because of the raise of commercial and governmental interests turned the Internet from being an anonymous media ("On the Internet nobody knows you're a dog", Peter Steiner, 1993) to a place where users can be easily profiled and therefore need to be careful with their private information ("How the hell does Facebook know I'm a dog?", Rob Cottingham, 2010).

Network monitoring of Internet traffic is important to provide the network operators with information on the performance and load status of their networks, as well as to

detect intrusions or to classify traffic. In this context, the intention of monitoring is not to intrude the user privacy, but for example to gain knowledge about bottlenecks and potential performance improvement options.

Knowledge about the quality of experience (QoE) of the Internet customers provides a useful input for network operators to perform operative network management, but also the plan their networks for future requirements. One important parameter for the customers' QoE is the application response time, defined as the absolute value of the time difference between the user's action and the application reaction. Obviously, this parameter has many inputs like the performance of the application itself, the operating system delays, computing performance of the server, etc. Given that future software applications become more and more network dependent, the speed of the data transportation has a big impact on QoE. The main network level parameter of the application response time is the end-to-end delay of the packets on the network. Depending on the application, either the one-way delay (OWD) or the round-trip time (RTT) is of greater interest. This thesis concentrates on exploitation and evaluation of RTT measurements, which is important for many current and future Internet applications like cloud computing, distributed applications with centralised servers or largely distributed service mash-ups.

Measuring the RTT can be performed by different means. Generating useful information out of passive network monitoring has several advantages compared to measuring networks by active probing. First, with active probing additional traffic is sent to the network. This additional traffic may be treated differently by the network elements compared to the real user traffic. Second, this additional traffic generates additional load on the network, which carries no information but delivering only the results for the measurements. Avoiding such useless traffic and therefore reducing the overall load of networks reduces required resources for the overall Internet infrastructure. Third, the additional load introduced by active network probing biases the results of the measurements. Finally, and this is considered as the most important point, evaluation of artificially introduced traffic can never produce the same results compared to thorough evaluation of real user traffic.

Therefore, the possibilities of generating useful information by measurement of RTTs

to remote network locations by passive network monitoring are presented in this thesis. As soon as it comes to passive monitoring of Internet network traffic privacy concerns are raised, as Internet user data must be treated as sensitive personal data, and may be therefore only limitedly used by service providers (e.g. for accounting and billing reasons). Due to the recently developed privacy-preserving monitoring framework PRISM[1], the exploitation of information gathered by passive network monitoring enables many new possibilities of lawful, privacy-preserving applications.

Finally, knowledge about RTT is required by many applications (CDNs, peer-to-peer networks, Internet mapping, Identification of high-delay networks, etc.). In addition, network management applications or measurement platforms like iPlane [35] that rely on RTT data can exploit the generated information. Providing this information on different abstraction levels from the high granularity of single end-systems up to the more abstract levels like Autonomous Systems or even countries or continents can provide new views and allows e.g. to share this data among network operators without privacy concerns. This further enables to run applications from a multiple measurement points of view, rather than just having the single, provider-centric view.

The tools and methods developed and presented in this thesis will enable the network operators to answer the following questions:

- What are the important services demanded by the customers to what extend?
- Which Autonomous Systems provide the services used by the customers?
- What response time do the peering and transit networks provide to those services?
- What is the relation between performance and usage?
- Is there a need for improvements in the inter-domain connections of the network?

[1] `http://www.fp7-prism.eu`

2

Background Information and Deployment Scenarios

This chapter provides the required background information about AS-level Internet topology in Section 2.1 and introduces the parameter "round-trip time" in Section 2.2. Furthermore, Section 2.3 presents four deployment options of the proposed system.

2.1 AS-Level Internet Topology

The Internet is a decentralised infrastructure consisting of multiple separate but interconnected networks. While network administrators can use their own interior routing protocols, the routing between the administratively divided networks needs to follow a common standard. For this task the exterior routing protocol BGP (Border Gateway Protocol) has been standardised [49]. The separated administrative domains are so-called *Autonomous Systems* (AS).

The ASes were originally addressed with a 16bit AS number (ASN), i.e. from AS0 to AS65535, but the continuous growth of the Internet required the introduction of a 32-bit numbering scheme for AS numbers in December 2006 using a dotted notation (AS0.0-AS65535.65535). According to the ASN number report [25], currently more than 61.000 AS numbers have been allocated, but not all of them are announced in the Internet through BGP. The number of the announced Autonomous Systems is currently at about 35.000 ASes and is still growing. Figure 2.1 depicts the almost linear growth

of announced ASes since 1997. Compared to this the size of the BGP table, that is required for routing within the Internet currently contains almost 350.000 entries [26] for almost 700 billion Internet hosts worldwide [59].

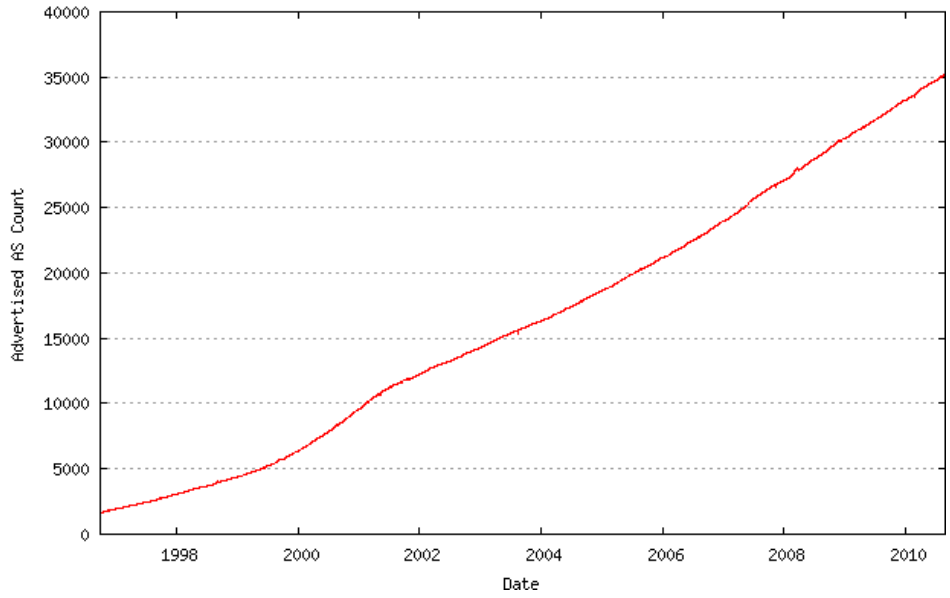


Figure 2.1: Number of Autonomous Systems announced in the Internet since 1997 [25].

Having nodes and links, the Internet on the IP-level can be modelled as a graph, where routers are the nodes and the physical connections between them are the links. The same model can be applied on the AS-level, where only the links between the border routers of the ASes are modelled and the nodes are the ASes themselves. An example of such model is illustrated in Figure 2.2. The ASes exchange traffic, which means that they act as source, sink and/or transit nodes. Section 5.1 provides some more details on the modelling of the AS-level Internet topology and related literature.

2.2 Round-Trip Time as Parameter for Quality of Experience and Quality of Service

This section provides details about the terms quality of service (QoS) and quality of experience (QoE) and explains how round-trip time (RTT) relates to those parameters. Roughly, QoE defines the *experience* perceived by the end-user, which also depends on the *service quality* of the network. This *QoS* is defined by a set of measurable network

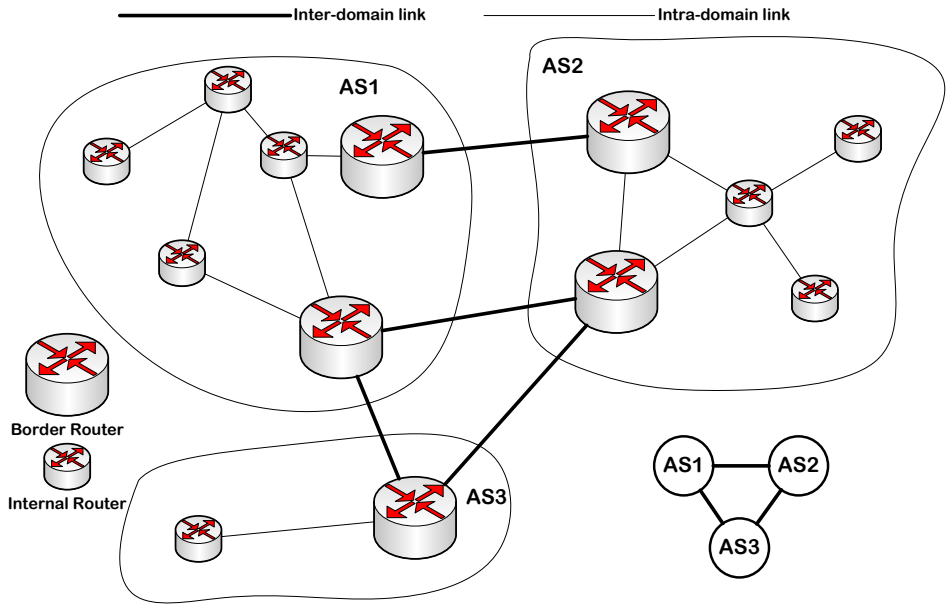


Figure 2.2: Example of a router-level and corresponding AS-level topology.

parameters, where RTT is one of them.

2.2.1 Quality of Experience

While QoS is already a very commonly used term for directly measurable quality parameters, "Quality of Experience" (QoE) has been introduced recently to describe not only the service quality, but also include user experience parameters like price and reputation and the degree of security of a service [16]. QoE can be generally defined as the user satisfaction during the use of a specific network based communication application or service. Network providers need to manage the QoE by performing two steps. First the QoE of the users must be measured, second the system must be able to be improved to provide a better satisfaction to the users, if necessary.

Different methods and tools exist to measure the user experience. Originally, real users were in charge of performing the tests, where different user groups tested the services (like audio or video applications) manually. During those tests, the users had to individually rate the quality and answer related questionnaires. Although such tests are still required and performed for several applications, those tests are expensive and time-consuming procedures. To avoid them, mathematical models have replaced those tests where possible. An example for such model is the E-Model for determining audio

quality specified by the ITU-T [28]. Based on a reference connection, this model takes several input parameters like impairment factors for delay and equipment to calculate the transmission rating factor ("R-Factor"). It is calculated as

$$R = Ro - Is - Id - Ie\_eff + A \tag{2.1}$$

where Ro is the basic signal-to-noise ratio and A is the advantage factor, which depends on the acceptance level of the user regarding the underlying communication system. The users for example accept in general a worse quality in case he or she uses a moving mobile equipment compared to a conventional fixed-line phone. Is, Id and Ie_eff are impairment factors that occur with simultaneously with the voice transmission Is, the delay of the voice signal Id and in the equipment and codec Ie_eff. The delay on the link between the communication partners is one important input factor of Id. The resulting R-Factor lies in the range of zero (bad quality) to 100 (excellent quality) and can be mapped to the user experience levels of the Mean-Opinion-Score (MOS). The MOS defines five different levels of user satisfaction: excellent (5), good (4), fair (3), poor (2) and bad (1).

It is important to mention that the QoE of the user in terms of delays or round-trip times does not only include the network delays but also delays introduced by the end-systems. This can include the time required to decode a live video-stream, encrypt or decrypt messages or to finish the rendering of a web page.

2.2.2 Quality of Service

Already in the late 1990s, the term "Quality of Service" (QoS) was introduced into the area of data networking, adopted by the ITU-T Recommendation X.641 [27] and the IETF mainly in RFC2205 on resource reservation [5] and RFC2474 about the differentiation of network services [41], both of them update several times by later RFCs. Several network-level QoS parameters have been standardised by the different standardisation bodies. The four major network QoS parameters are *packet delay* (One-way or Round-trip), *packet loss*, *delay variation* (sometimes called jitter) and *throughput*. Further derived parameters are packet reordering or specific packet loss patterns (e.g. loss

bursts or independent losses). Each of the listed parameters differently influences the performance of the network application. While some applications are resistant against large delays they can suffer from low throughput (e.g. file downloads). On the other hand, telephony applications have the opposite requirements, namely low delay and jitter, while they do not require high throughput. Other examples where high one-way or round-trip delays are annoying for the users are: Interactive TV, interactive gaming and basically all modern web applications with a high ratio of user interaction using asynchronous technologies like AJAX[1] which frequently performs web server interaction in the background, even if the user perceivably stays on a single web-site.

RTT is therefore only one out of a number of network performance parameters, but an important one. It can be very useful to discover poorly connected Internet destinations or to raise alarms in case of routing misconfiguration or similar errors. The correlation between RTT and other QoS parameters to derive additional information would be a logical next step but is out of scope of this thesis.

2.3 Deployment Scenarios

This section shows four example scenarios with different requirements that can apply the presented measurement and estimation methods. First, the end-host scenario allows Internet users to monitor their own connections. Second, the LAN scenario already includes measurements of multiple users, like the employees of a company. Furthermore, two provider scenarios are investigated, divided depending on the role of the provider. In the stub provider scenario only a single direction of RTT measurements will be of interest, while in the transit provider scenario RTTs between the measurement point and both of the communication partners can be exploited for delivering results. The terms *stub* and *transit* are taken from the classification of ASes as defined by Oliveira at SIGCOMM 2007, as follows: "A stub AS *only* appears as the last AS in an AS path, while a transit AS will appear in the middle of some AS paths." [42], page 10. The segregation into the four scenarios has been applied due to their specific requirements in terms of privacy and accuracy. The stub provider scenario will be the major reference

[1] Asynchronous JavaScript and XML

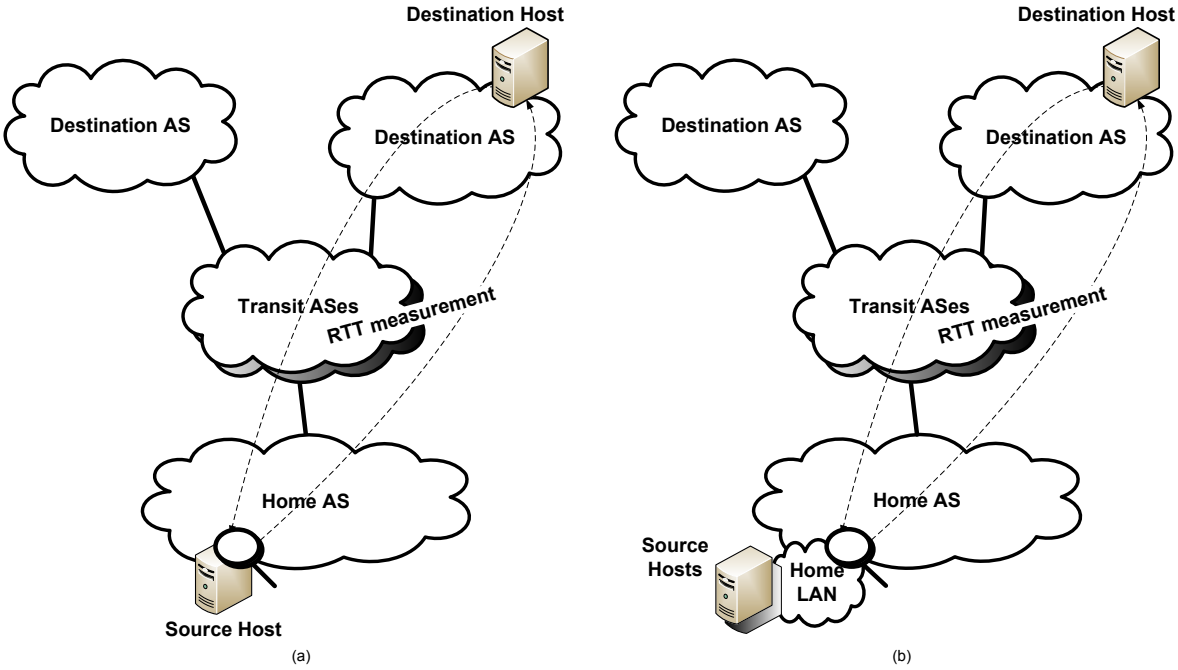


Figure 2.3: End-host (a) and LAN (b) RTT measurement scenario.

scenario in this thesis.

2.3.1 Scenario 1: Monitoring on an End-Host

The end-host scenario allows a single user to measure RTTs of the own traffic into the different ASes. It provides information to what remote hosts and networks the computer connects and how the RTT performs on those connections. However due to the limited amount of connections generated by a single user, the results delivered will suffer from quality. As privacy is not an issue in this scenario, it can be deployed easily.

An overview on the scenario is shown in Figure 2.3(a). The same scenario can be applied on the server side. It enables the service operator to know from which remote ASes customers are connecting to offered services and what RTT they experience.

2.3.2 Scenario 2: Monitoring on an Uplink from a Campus or Company Network

Scenario 2 is similar as Scenario 1 and is depicted in Figure 2.3(b). The difference to the previous scenario is that now the observation point is dislocated from the end-host(s) that generate traffic. This separation can already raise privacy concerns from the end-users, because the administrator of the campus or company LAN (denoted as "Home LAN" in Figure 2.3), who has access to the observation point is able to monitor the full communication data of the connected users. For the purpose of monitoring the RTT to the remote networks, the observation point can already eliminate the mapping from local IP addresses to the specific user.

2.3.3 Scenario 3: Monitoring on the Inter-domain Link(s) from a Stub Network

The next step of aggregating multiple users is to place the observation point on the uplink of an Internet Service Provider (ISP) as shown in Figure 2.4(a). This scenario will deliver more accurate data due to the usually larger number of users. Consequently it has more available measurement values compared to Scenario 2. The main property of a stub provider is that it does not route traffic between networks, but only appears as a source or sink of traffic from the inter-domain point of view. Still a stub provider can be multi-homed, i.e. it has connections to multiple peering partners for load-balancing, backup or economical reasons. In case of a multi-homed network, it must be ensured that either the forward and backward routes are passing the same link, or monitoring data from both observation points must be combined in order to enable accurate passive RTT analysis that requires both directions of the traffic. Measuring the RTT from the observation point towards the end-user would mean to measure the performance of the "Home AS", which might be of interest in some use cases but usually ISPs would rely on other management tools to monitor the status of their own network.

As in Scenario 2, ISPs can only deploy such scenario when they have a privacy-preserving framework to ensure the privacy of the customers. Such framework has been

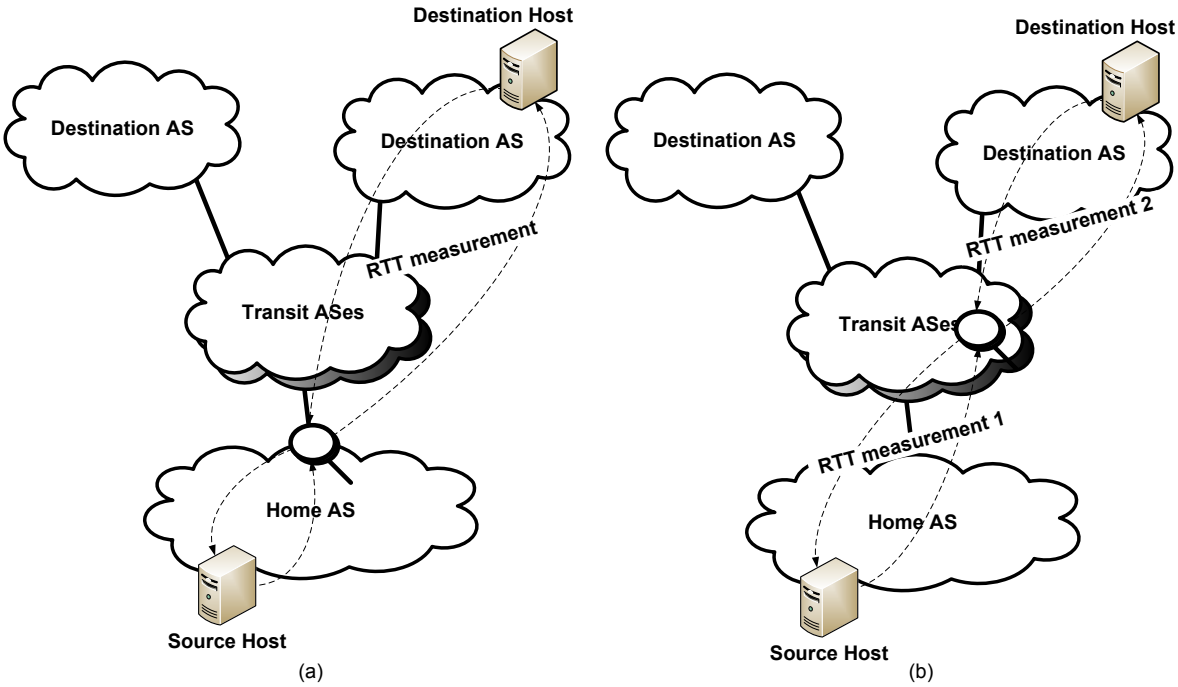


Figure 2.4: Stub (a) and transit (b) provider RTT measurement scenario.

implemented and evaluated in the PRISM project [48], further described in Chapter 6. The main objective of the PRISM project was to design and implement a privacy-preserving monitoring system. Based on this system, different use cases have been demonstrated, one of them for measuring delays to remote Autonomous Systems. The PRISM architecture strictly requires the definition of specific monitoring purposes and the association to user roles. This scenario will be used as the main reference in this thesis. Also the data analysis presented in Chapter 5 is based on such scenario.

2.3.4 Scenario 4: Monitoring on a Core Link of a Transit Network

Finally, the transit provider scenario can be applied to networks that route traffic from other networks, as shown in Figure 2.4(b). Monitoring in such scenario requires more advanced techniques than in the previously described ones. While the above scenarios can be easily managed from a single measurement point that can observe all traffic in both directions, core links often transport traffic on asynchronous routes. Sometimes there is even only one direction visible to the network owner, while the backwards

direction is routed on a different AS path. Monitoring RTT in such scenario would therefore require either to collect packet traces from many locations together and send them to one central point for calculations, or the application of less accurate estimations of RTT based only on the forward way of the TCP traffic. Literature on such methods is also available and will be discussed later in Section 3.4.

3

Estimation Methods for Round-Trip Times to Remote Network Locations by Passive Network Monitoring

The objective of this chapter is to present the details about how round-trip time can be measured and what estimation algorithms can be applied. Also a definition of this metric will be provided and how it can be reported from the observation point to some remote measurement management system. Some quantitative properties on RTTs that have already been detected in earlier works are described in Section 3.3. After that the thesis differentiates between RTT measurement and RTT estimation. RTT measurement means the exact measurement of the time a packet needs from a measurement point to a remote Internet end-system and return. Such measurement between two points in the Internet can be performed actively and passively. RTT estimation on the other hand defines the algorithms for the calculation of RTTs that cannot be directly measured.

3.1 RTT Measurement Methods

In this section the two basic methods on how RTT can be measured are differentiated are presented: *active* and *passive*. Such distinction is very popular in the whole research community working in the area of Internet monitoring and measurements.

3.1.1 Active RTT Measurement

For sake of completeness the active measurement of RTT is shortly described in this section. As with any active measurement, it must introduce additional traffic to the network and therefore needs to be configured carefully in order to not bias the user generated network traffic. The classical method of measuring RTT is the use of `ping`, which sends ICMP echo packets and measures the time until the answer of the remote host is received. This method has two major drawbacks. First, firewalls are often configured to block ICMP packets, which means that this method can lead to no results. Second, ICMP packets may be treated differently by routers compared to real data traffic and therefore experience more delay, leading to incorrect measurement results.

Another often used active measurement method is the use of `traceroute` in various implementations. The original implementation sends UDP packets with an increasing value set in the TTL field. Intermediate hosts between source and destination are expected to reduce the TTL field by one and reply with "ICMP Time Exceeded" messages when the TTL is expired. The replies can be used to evaluate the RTT between the sender and each of the intermediate host. This method experiences the same drawbacks as the usage of `ping`.

As a third method of active, but indirect delay measurements, the tool `king` can be mentioned. It allows to estimate the delays between various Internet end-hosts [22] by exploiting the domain name service (DNS). It is based on the assumption, that a large number of Internet hosts are close to their authoritative DNS servers. The measurements utilise recursive DNS queries, where name servers are forwarding the request to other name servers. Measuring this latency, and subtracting the RTT to the

first name server allows the calculation of the RTT between two name servers. Although the method may deliver largely inaccurate results in case the end-hosts are using remote DNS servers, it can be of interest for complementary measurements between locations that are not directly accessible.

Active RTT and delay measurements have large deployments by the Internet community, e.g. to build alert systems [24], for network tomography [40] or to discover the Internet topology on router level [53].

3.1.2 Passive RTT Measurement

Passive RTT measurements are based on real traffic in the network, without the injection of measurement packets. Measuring RTT in a passive way is in general more complex than by its active counterpart. Rather than using self-generated traffic probes, it requires the monitoring of traffic that includes at least the arrivals of two packets X and Y, where Y has been triggered by X, i.e. usually a packet that contains the acknowledgement of data sent in packet X. Therefore not all protocols enable passive RTT measurement. The two major transport protocols in the Internet are TCP and UDP, where TCP can be exploited for passive RTT measurement. In TCP each packet is marked with a sequence number (SEQ) to identify the offset of the byte that is currently sent. The receiver acknowledges the reception of the data by sending a packet with an acknowledgement number (ACK) identifying the next expected byte in the data stream. This enables the correlation of two packets that belong to one round-trip (SEQ-ACK). The time difference between those packets is then calculated as the current RTT. Due to the dynamic behaviour of TCP with delayed ACKs, sending and congestion windows and different TCP implementations such measurements cannot be done straight-forward, but usually are calculated estimates. Possible estimation algorithms to derive the RTT from passive monitoring are described in Section 3.4.

3.2 Definition and Reporting of RTT

Round-trip time is roughly defined as the time required for a packet to traverse along the network from a sender to a receiver and back. How they are defined in detail by the IPPM working group of the IETF and which global quantitative properties can be assumed is reported in this section. Finally, a method for reporting RTT using the IPFIX protocol is described.

3.2.1 The Round-Trip Delay Metric for IPPM

The metric for the round-trip time has been standardised by the IP performance metrics group (IPPM) of the Internet Engineering Task Force (IETF) in 1999. In the proposed standard, "A Round-trip Delay Metric for IPPM" (RFC2681 [3]), the motivation for measuring RTT[1] is emphasised by the following reasons:

- Some applications require low RTTs to perform well.
- High delay variations destroy the QoE of most interactive real-time applications
- RTT impacts the bandwidth provided by higher-level transport protocols
- The minimum RTT can be assumed as the (constant) sum of propagation and transmission delays, measured on lightly loaded networks.
- Values above the minimum indicate path congestion.

Also a few weaknesses are mentioned that this metric has compared to measurements using the one way delay metric, which is also specified by the IETF in RFC2679 [2]. The major drawback is that Internet paths are only seldom symmetric. Estimating the one-way delay from RTT is therefore impossible, e.g. by dividing the RTT by two into forward and return part. Even if packets in both directions are taking the same path, asymmetry is introduced due to different loads on the network or asymmetric links along the path, or due to different QoS provisions.

[1] In this thesis RTT is used as a synonym for both 'Round-trip time' and 'Round-trip delay'

As long as those weaknesses are taken into account, RTT measurements can still be utilised for the following reasons: RTT measurement tools are usually very easy to deploy, as there is no need for exact time synchronisation at the measurement hosts. This ease of deployment as well as the ease of interpretation keeps running RTT measurements interesting for providers. RTT also includes processing time in the destination, which may also impact the applications QoE.

In the RFC, the singleton part of the metric is defined as "Type-P-Round-trip-Delay" with three parameters: *Src* and *Dst*, specified by the IP addresses of the hosts and T, a time. In the definition of the metric, three cases are distinguished:

1. the "Type-P-Round-trip-Delay" from *Src* to *Dst* at T is dT
2. the "Type-P-Round-trip-Delay" from *Src* to *Dst* at T is undefined (informally, infinite)
3. the "Type-P-Round-trip-Delay" between *Src* and *Dst* at T

While in the first case, the source and destination hosts are explicitly identified together with the delay that was measured, the second case indicates a lost measurement, and the third case a measurement that has been made either from *Dst* to *Src* or from *Src* to *Dst*, i.e. without specification of the direction.

The definition of the Type-P packet is important, as the RTT heavily depends on the packet size. Long and short packets experience the same queuing and propagation delay, but significantly differ in their transmission delays.

Note that RFC2681 is intended to be implemented by means of active measurements, this thesis will concentrate on measuring RTT by passive means, as described in Chapter 3. Still, most parts of RFC2681 are also relevant for RTT measurements by passive monitoring.

3.2.2 Reporting RTT with the IPFIX Protocol

The *Internet Protocol Flow Information Export* (IPFIX) protocol [10] is a push-protocol designed for network monitoring. The architecture defines a *metering process* with an

exporter located at the *observation point* and a *collector*, which can be located remotely. The amount of information exported by the protocol has a high flexibility in order to cover current and future applications.

For the information transported by the IPFIX protocol *templates* are required to be defined for each structured data *set* that should be exported. *Templates* can be defined within the protocol header, i.e. they are not standardised but can be defined during runtime. A data *set* consists of multiple *information elements* (IE), each of those elements are defined by a name and a data type.

Exporting RTT information is not yet foreseen in the IPFIX standards. However, due to its flexibility it can be enhanced by self-defined fields. Many information elements and data types are already defined, which can be used as a basis. For the RTT case, just a single information element (RTT) had to be newly defined, while all the other information were selected from the existing information elements defined in RFC5102 [47]. Information elements that are used to monitor RTT on an IP address basis are listed in Table 3.1. The unit of RTT is microseconds. The Private Enterprise Number (PEN) 12325 was used to define Information Elements in the PRISM project, that are not yet listed with their identifiers (ID) in the IANA registry[2] (PEN=0).

PEN/ID	**Name**	**Data Type**	**Length**
0/323	observationTimeMilliseconds	unsigned64	8 octets
0/12	destinationIPv4Address	ipv4Address	4 octets
12325/199	roundTripTime	unsigned32	4 octets

Table 3.1: Template with IPFIX RTT Information Elements using IP addresses.

For a privacy-preserving data export without any IP address, another template shown in Table 3.2 was defined. It supports the export of AS numbers instead of IP addresses.

PEN/ID	**Name**	**Data Type**	**Length**
0/323	observationTimeMilliseconds	unsigned64	8 octets
0/17	bgpDestinationAsNumber	unsigned32	4 octets
12325/199	roundTripTime	unsigned32	4 octets

Table 3.2: Template with IPFIX RTT Information Elements does not export IP Addresses, but uses AS numbers.

[2] `http://www.iana.org/assignments/ipfix`

3.3 Quantitative Properties of RTT

Some general quantitative properties of the RTT by analysis of the results of a large number of ICMP and UDP traceroute probes are available in the report of the ISMA Workshop of October 2002[3]. They were sent from backbone monitors to a representative, globally distributed set of IP addresses. The talk was given by Andre Broido and presented additional findings, giving useful input for the AS level analysis made in this thesis. The general finding is that, RTTs "are to large extent independent of year, monitor location, sample size, time of the day and traceroute type". These properties, especially the independence of the time of the day, which is also the case in the result analysis given in Chapter 5.

In a more recent work in 2009, Maier, Feldmann, Paxson and Allman [36] performed a large scale study on residential DSL-based broadband Internet traffic. Among other performance parameters, they also studied the behaviour of the RTT based on the TCP SYN/ACK analysis. They distinct between local and remote RTTs, and discovered that local RTTs are substantially larger than the remote ones, which is partly caused by congested access links, but also due to the use of old (11Mbps) wireless access technology in users homes in dense population areas with numerous overlapping wireless networks. The data used within this thesis do not deliver such result, maybe because the network under investigation was from a cable network provider, not from a DSL provider. Other facts could be that the investigated data was not collected in such dense population areas, the the users did not heavily use wireless access, or the investigated provider had a less powerful uplink towards the major Internet backbones.

3.4 RTT Estimation Methods

Derived from the passive RTT measurements described in Section 3.1.2 different RTT estimations can be performed:

- Estimation of the end-to-end RTT between TCP sender and receiver

[3] `http://www.caida.org/workshops/isma/0210/ISMAagenda.xml`

- Estimation of the average RTT between observation point and remote networks or Autonomous Systems
- Estimation of the average RTT to different countries or continents

Measuring RTT by passive link monitoring can be done by several approaches. In the literature often the estimation of the end-to-end RTT is studied [29, 30, 58], in order to evaluate the performance of the TCP connection. The reasons for such approach are for example to estimate the retransmission timeout (RTO) of a connection or to estimate the available bandwidth of the path.

Jiang and Dovrolis [30] presented the handshake (SYN-ACK) estimation and the TCP slow-start estimation, requiring four segments at the beginning of the TCP connection with maximum segment size (MSS). Those algorithms generally provide one measurement per TCP connection.

Jaiswal et. al. [29] proposed an estimation technique for end-to-end RTT throughout the lifetime of a TCP connection, i.e. they are calculating a running RTT estimate. For their estimation method they need to know exactly which data packet is triggered by which ACK. This requires an estimation of the TCP congestion window, which is a complex task as it depends on the type of congestion control used in the TCP implementation and also on the packet loss. As there are situations where this task cannot be performed accurately, RTT estimations are being stopped and restarted.

Another important step has been investigated and evaluated by Shakkottai et. al. [52] and Veal, Li and Lowenthal [58]. Their algorithms allow to calculate end-to-end RTT estimates from packet captures containing only unidirectional TCP traffic. It relies on traffic patterns caused by the self-clocking mechanism from TCP. Such algorithms allow the measurement of the RTT throughout the lifetime of a TCP bulk data transfer session, while it cannot be applied to interactive sessions with only little data transfer.

The work presented in this thesis targets a different goal. Rather than estimating the end-to-end TCP RTT between TCP sender and receiver, the RTT between the measurement point and the network that hosts TCP multiple end-points is the parameter under investigation. The idea is based on the network operators view, where the RTT estimation between single end-systems (ESes) is less important for two main reasons:

First, the amount of generated data is too huge and varying in order to report all ES-level details to the operator and second, mid- to long-term network planning needs to be done on the level of AS peering agreements. The core contribution from this thesis is therefore not to provide another end-to-end RTT estimation method, but to derive an estimate of the average to a remote AS or country from a number of RTT measurements that are taken per end-system. The goal is to get a useful metric for evaluation of whole networks (and compositions of networks) rather than to single hosts, as such metrics are more important for ISPs for managing their peering partners and connections. Also, privacy-issues are less problematic, if individual end-systems are not under investigation. Therefore such AS-level estimation can be seen as another layer above the RTT measurements and estimations per end-system, using the values from the measurements below for the estimation per AS. Depending on the deployment scenarios presented in Section 2.3, different RTT measurement and estimation methods are more or less useful.

3.5 Data Aggregation

Each IP address seen in the Internet is connected to an end-host. This end-host or end-system has an administrative and geographic location. Figure 3.1 shows the hierarchy of the administrative structure of the Internet and the geographic relation. Administrative means, there is some company, private user or ISP that "owns" the IP address either permanently (static IPs), or only temporal for a given time interval (dynamic IPs). The administrative location can usually be found very easily using the `whois`-services from regional Internet registries like RIPE. Finding the exact geographic location is more difficult, due to mobile Internet and VPNs, it is rather impossible on a global scale. However, several services exist that provide mapping information between the IP address and the geographic location. Some selected services as well as additional meta-information that is available are described in Section 5.5.

Based on the hierarchy, data aggregation of the measurements per end-system (i.e. per IP address) can be performed. Different aggregation algorithms that can be selected to estimate the RTT are further described in this section. The single RTT observations

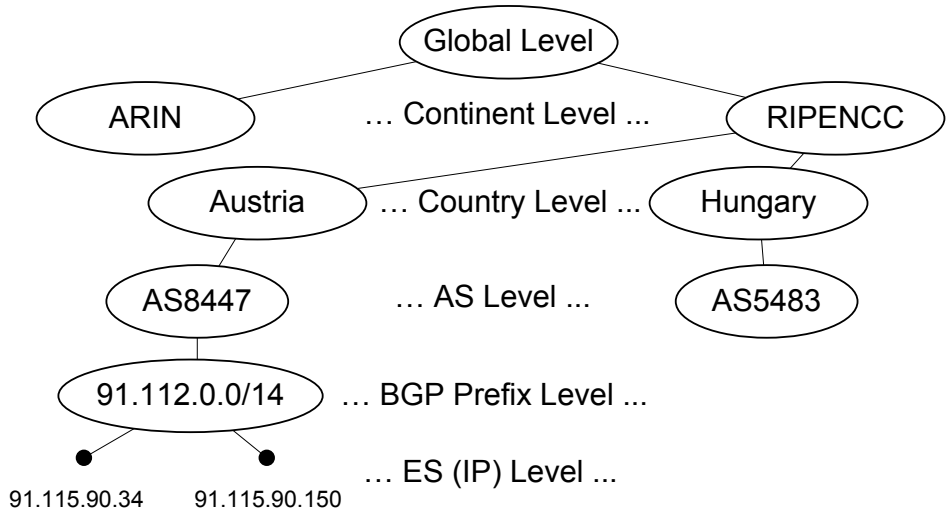


Figure 3.1: Simplified hierarchical assignment structure of the Internet. Due to the existence of transnational ISPs, the structure may have cross-connections. I.e. an end-system can be located in country X, while the associated AS can be located in country Y.

for inter-domain network delay estimation are produced in irregular time intervals. A well designed system requires to provide a *useful* moving estimation value at each time during the observation period, rather than a final *exact* estimation at the end of the measurement campaign. Generally spoken, RTTs cannot be considered as a bounded data set, but must be handled like a real-time, online data stream.

3.5.1 Extreme Values

Minimum measurement value and maximum measurement value are the extreme values from a measurement series. The maximum is not very expressive when measuring delays in the Internet as the maximum can become infinite per definition when packet losses occur. On the other hand, the minimum value can give important information about the status of the network. Measured minimum values reflect the length of the path, containing the *propagation delays* and the *link delays*. Propagation delays are independent on the packet length and capacity of the links, but depend on the properties of the transportation medium (fibre, copper, air). The link delays can be calculated by $\frac{L}{C}$, L being the packet length and C the link capacity. This means that with the knowledge about the packet length (which is available in the IP header), the end-to-end path capacity with empty queues all along the path can be calculated using the

measured minimum delay. The time to calculate the route (*routing delays*) and the time to wait on the router output link until it serves the packet (*queuing delays*) are the variable fractions, that do not appear in the minimum delay. Additional parameters to the minimum delay are required to reflect those delays in the measurement results.

3.5.2 Moving Averages

Multiple statistics are available to build averages on observations. Depending on the distribution of the data, some of them are more and others less meaningful. In this section, different averages are evaluated for case of RTT measurements.

3.5.2.1 Simple Moving Mean

The simple moving mean is the arithmetic mean value, moving along a sliding window. On the k-th position the mean of the round-trip time is calculated as:

$$\overline{RTT}_k = \frac{1}{n} \sum_{i=1+k-n}^{k} RTT_i \quad \forall\ k \geq n \tag{3.1}$$

n is the window size, that can be configured. Big window sizes provide smoothed values over a long history, while small window sizes only take recent values into account. A small window size will provide more jitter in the estimate than a big one, while a big one requires more values available in order to do the calculation. The storage requirement for such algorithm is constant and proportional to n. Outliers in RTT caused e.g. by packet loss must be removed before such algorithm, otherwise they would divert the simple moving mean from its real value easily. Due to the irregular appearance of measurement events, a time-based window should be preferred against an index-based window.

3.5.2.2 Simple Moving Median

Like the simple moving mean, also the simple moving median needs to be configured with a window size n. Before it can be calculated, the RTTs must be sorted. The

median is the number on the "middle" position of the sorted set in case of an odd n. For example a set of $\{1, 3, 5, 6, 6\}$ has the same median as $\{1, 5, 5, 6, 99\}$, which is 5, while the mean values are 4.2 and 23.2 respectively. This example already shows that the median is more robust against single outliers. In case n is even, the mean value of the two middle values must be calculated to find the median value. Due to the sorting requirement, this algorithm has a more complexity compared to the moving mean, while the storage requirements are also constant and proportional to n.

3.5.2.3 Cumulative Moving Mean

The cumulative moving mean is the arithmetic mean value among the full history of observations. It can be calculated with a constant storage requirement as follows:

$$\overline{RTT}_k = \frac{1}{n} \left[RTT_k + (k-1)\, RTT_{k-1} \right] \quad \forall\ k > 1 \tag{3.2}$$

Like the above mentioned simple version from Formula 3.1, the cumulative moving mean quickly follows outliers.

3.5.2.4 Cumulative Moving Median

This median is similar to the simple moving median, but takes the full history into account. The problem of calculating this average is that the storage requirement increases with the number of observed values. Therefore it cannot be applied in practice.

3.5.2.5 Exponential Weighted Moving Average

The Exponential Weighted Moving Average (EWMA) is a commonly used statistic to calculate a moving average. Using this algorithm, the "weight" of the new value can be configured by the weighting factor α. It is, for example, also used to smooth the measured round-trip time for the calculation of the retransmission timeout of TCP [45]. The RTT after k observations is therefore calculated as follows:

$$EWMA_k = \alpha RTT_k + (1 - \alpha)\, EWMA_{k-1} \quad \forall\ k > 1 \tag{3.3}$$

The *EWMA* must be initialised, for example with the value of the first observation, or the mean value of the first 5 observations. Setting the α-value correctly depends on the number of measurements seen within a given time interval. The higher α, the higher new values will be weighted. In case of many observations in a short time frame, α can be set between 0.1 and 0.2. But if the breaks between two measurements increase, α must be increased as well, otherwise earlier measurements (e.g. from the day before) are weighted much higher than current observations. Also the elimination of outliers is required to usefully apply the EWMA algorithm. Otherwise single outliers increase the EWMA, and it takes a large number of new measurements to return the estimate to the true value.

4

Privacy-Issues Raised by Internet Monitoring and Regulative Framework

Internet traffic contains private communication between the Internet users. It therefore must be treated like any other private communication. This chapter will introduce the regulative framework that has been created by the European Parliament and Council to protect the individual personal data. As all European Directives, they are not directly enforced but allow the member states to create their own legislation with some minor adjustments to local laws.

The basic regulation framework document concerning privacy is European Directive 95/46/EC on the protection of individuals with regard to the processing of personal data and on the free movement of such data [17]. In this document, personal data and the processing of personal data are described. In Austria this directive is enforced by the "Datenschutzgesetz 2000 (DSG 2000)" [43].

In 2002, the data protection law has been refined for electronic communication with the directive 2002/58/EC. This directive harmonises the data protection regulations from the member states. It does not only apply for natural persons but also for legal persons. In Austria this directive is enforced by the "Telekommunikationsgesetz 2003 (TKG 2003)" [44].

Not directly relevant for the thesis but mentioned for completeness is that data protection laws are explicitly not applicable for activities that ensure the public security. Actions taken by the Member States "for the protection of public security, defence, State security and the enforcement of criminal law" [18] are definitely excluded from the application of the European privacy laws. This exception has been further extended in the European Directive 2006/24/EC on the retention of data generated or processed in connection with the provision of publicly available electronic communications services or of public communications networks. This enforces network operators to store connection data of their customers in order to track their communication behaviour. These data may not be commercially exploited by the operators.

In the following sections relevant terms and definitions from the European Directives on data protection in general and specifics about electronic communication, and how they are applied to Internet Monitoring.

4.1 Personal Data

In the European Directive 95/46/EC on the protection of individuals with regard to the processing of personal data and on the free movement of such data, *personal data* is defined as follows:

> *'personal data' shall mean any information relating to an identified or identifiable natural person ('data subject'); an identifiable person is one who can be identified, directly or indirectly, in particular by reference to an identification number or to one or more factors specific to his physical, physiological, mental, economic, cultural or social identity;* [17] Chapter I, Article 2(a).

Common examples for personal data are religion, health status, or personal preferences. It can be noted here that from continuously monitoring and analysing Internet traffic generated from a specific person, such information can be easily derived technically, e.g. from e-mail conversations or tracking of search terms. As soon as this information can be related to an individual person through combination of the data with identifiable data like a name, function, profession or address the data protection laws must be respected. As recommended by the data protection expert group ("Working Party"),

which has been established based on Article 29[1], also IP addresses must be treated as identifiable data, as persons can be indirectly identified based on his/her address. Such technologies are already widely in use for identification of criminal people and networks.

4.2 Data Processing

Again in Directive 95/46/EC, *processing of personal data* is defined as follows:

> *'processing of personal data' ('processing') shall mean any operation or set of operations which is performed upon personal data, whether or not by automatic means, such as collection, recording, organisation, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, blocking, erasure or destruction;* [17] Chapter I, Article 2(b).

Nowadays computer supported data processing is a common practice in almost all application areas, like in marketing, economics and research. As long as data processing is done statistically on anonymised data, i.e. there is no possibility to link back to an individual, there is no conflict with the data protection laws. Legal problems occur, if the direct relation to individuals can be done. One recent example of such legal case is the monitoring of private WiFi networks without the notification of the owners, as it has been performed by Google during capturing of photographs for their mapping and navigating project "StreetView". After claims from German data protection officers, Google stopped capturing WLAN data [60].

4.3 Privacy in Electronic Communication

In 2002 the EU directive about privacy regulation in electronic communication has been issued [18]. Main goal of this directive was to harmonise the partially already existing national laws on an European level. One of the main extensions was that in this directive not only the privacy from natural persons, but also from legal entities is

[1] Art.29 Data Protection Working Party:
`http://ec.europa.eu/justice_home/fsj/privacy/workinggroup/index_en.htm`

now covered. The national implementation of this directive in Austria has been made with in the "Telekommunikationsgesetz 2003 (TKG 2003)" [44].

4.4 Data Retention

The EU directive on data retention [19] defines the requirement of the European Union to store personal telecommunication data for a duration of six months up to two years. The collection of data needs to be done by the telecommunication operators, and includes all connection data (without content) including geographic location from all users independently any suspicious fact.

The data retention directive should have been implemented in national laws since March 2009. Due to many concerns, in Austria and other countries the corresponding law is still under discussion and has not been finalised yet. The European Commission has announced an evaluation of the data retention directive by mid of September 2010.

4.5 Application of the Regulative Framework on Internet Monitoring

The above issues motivated some of the work conducted in the PRISM project [48]. "Adequate" processing of private information is allowed, e.g. if they are required for accounting or billing purposes. The PRISM project built on this specification of dedicated monitoring purposes. Together with a reasonable access control infrastructure, Internet monitoring data can be used also for other operational tasks, like intrusion detection, traffic classification or performance monitoring.

The use of packet traces from operational networks for research and engineering tasks is often a problem. Too much application of anonymisation tools to the traffic does often remove important information required for new algorithms, like the detection of recent attacks. This problem is often referred to as the "privacy-versus-utility" dilemma.

Privacy-aware network monitoring will therefore be also an important topic in future. Today, search engine providers are facing problems with their data retention policies,

mainly about storing IP addresses. The concern is about to possibility to perform a profiling based on search terms, which is a valuable information for them to improve the services. In an open letter special concerns against search engine operators like Microsoft, Yahoo! and Google are raised by the Article 29 working party [32]. In case Internet service providers and network operators are collecting the same information based on passive network monitoring they will face the same problem.

5

Analysis and Aggregation of Measurement Results

This chapter starts with a state-of-the-art overview on available inter-domain topology models and the analyses that have been conducted in other works. From Section 5.2 to Section 5.4, a data set collected from an Austrian cable network provider is investigated on different levels. The analysis work for this thesis was performed using the 'R' environment[1]. The R-code used for calculation and generation of the graphs is available on the attached CD, directly embedded in the LaTeX-source of this document. The chapter concludes with Section 5.5 about the acquisition of meta information to IP addresses which is required to analyse the measurements on higher aggregation levels.

5.1 Inter-Domain Topology Analysis and Modelling

Before coming to a detailed analysis of RTT measurement data, it is required to understand how the Internet evolved during the recent years and which inter-domain topology models do exist. Starting with the theoretical models on the router-level, papers from Li, Alderson, Willinger and Doyle [1, 33] provide an approach on topology modelling based on statistics and graph theory. They are stating a first-principle theory on how networks are being planned based on practical constraints and present

[1] `http://www.r-project.org`

the topologies of two example backbone networks (CENIC and Abilene). This approach is complementary to many empirical studies based on Internet measurements, some of them described further below. One level up, the Internet topology can be also modelled on the AS-level. This abstraction hides the single links between the routers, and only shows the relations between the Autonomous Systems. The routing on such Inter-domain links is managed by the Border Gateway Protocol (BGP) [49]. Being a path-vector protocol BGP is aware of the full path to the final destination.

In 2005 and 2006 Chang presented his work on the establishment of traffic matrices on inter-AS level [6, 7] based on data from multiple repositories like RouteViews [57] and IRR [39], and applied generic graph growth models to the Internet AS graph. Traffic matrices analyse traffic flows regarding their source and destination, and were originally developed on the intra-AS level.

In 2007, Oliveira, Zhang and Zhang observed the evolution of the Internet AS topology presenting the results in [42]. According to this paper, one of the major challenges is to identify the *real* topology changes within the numerous *observed* topology changes that can happen not only at topology changes but also during visible route changes. It uses a birth-death process model considering three types of AS-level links: *Visible*, *Invisible* and *Hidden*. Beneath birth and death, a link can also be *revealed*, when it changes from hidden to visible. Links are invisible, if they are not announced by the peers due to routing policies. The detection of such invisible ("missing") links has been addressed by Cohen in [11].

A long term evolution study over 10 years based on RouteViews [57] and RIPE [51] data collected since 1997 respectively 1999 was presented by Dhamdhere and Dovrolis [12]. The papers report a linear growth in terms of ASes and inter-AS links since 2001, mainly at the periphery of the network. It also proposes an AS classification scheme according to their business type. The classes are: *Enterprise Customers*, *Small Transit Providers*, *Large Transit Providers*, *Access/Hosting Providers* and *Content Providers*. They built a classifier based on the customer degree and the peer degree of the ASes. Another classification method was proposed by Dimitropoulos, Krioukov, Riley and Claffy in [13]. Six different classes have been specified there: *Large ISPs*, *Small ISPs*,

Customer ASes, *Universities*, *Internet Exchange Points (IXPs)* and *Network Information Centers (NICs)*. Input parameters for their classification were several attributes of the ASes collected from the IRR [39] and the Oregon RouteViews Project [57]. The classification of AS is very useful as additional information for the evaluation of their RTT, in order to make a correct interpretation.

Zeitoun, Chuah, Bhattacharyya and Diot [61] studied the delay characteristics of inter-AS links, which are generally known as the performance bottlenecks in the core of the Internet. The three interesting findings of the paper were i) that there is no significant contribution to the end-to-end delays for most of the investigated links, ii) the few discovered exceptions are long-distance links, where the propagation delay is already significant and iii) there is no major day/night difference in the delay on those links. The most interesting result for this work is that the major part of the end-to-end delays is usually caused by a single AS along the path.

5.2 General Properties of the Investigated Data Set

In this section, the RTT measurements taken at a stub provider are analysed. The input data for the studies were taken during a measurement campaign in 2009 for a period of 10 days. It includes 9 million single RTT measurements, based on the handshake and termination methods to more than 32000 different end-systems located in more than 3200 different ASes. After filtering the RTT measurements towards the local AS and one AS that generated 2.5 measurement results per second to a single end-system (maybe because of an attack), the number of investigated measurements was finally 2.2 million.

For a better understanding of the measurement results, and how they can be used to estimate network delays on the AS-level, these measurement results were studied in detail per end-system. After some general properties of the data set provided in Section 5.2, the data was investigated on different aggregation levels. Starting on a per end-system basis ("perES"), the data has been aggregated in different aggregation levels. Major goal of this work to rate the connection quality to the target ASes and therefore the most important aggregation level is the AS-level ("perAS"). Still,

the aggregation per BGP-prefix is of interest ("perPrefix"), especially in large ASes. Therefore also this intermediate level has been taken into account for some studies. Finally, in order to find out whether higher aggregation levels make sense and also for the sake of visualisation, the data was also investigated by the country code ("perCC") and also per Regional Internet Registry ("perRIR"). The RIRs are currently world-wide five organisations which are responsible to delegate the Internet Number Resources and manage the AS numbers in their region: AfriNIC for Africa, APNIC for Asia and the Pacific region, ARIN for North America, LACNIC for Latin America and RIPE NCC for Europe, Middle East and Central Asia. As there are five RIRs, each of them mainly responsible for one continent, the last level can also be seen as like a continent level[2].

5.2.1 Timely Distribution of the Measurement Values

As a first analysis, the occurrence of the measurements was investigated. The observation point provided 2.2 million measurements to remote ASes in a total time of 10 days, which is a mean value of 2.5 measurements per second, i.e. a mean distance of 400ms between two measurement.

Figure 5.1 shows the total number of measurements per hour. The dotted vertical lines indicate midnights of days. Measurements started during Monday night. Both, the daily curves, as well as the weekly curve is visible. Also visible is the "long weekend" in the data. The Monday (after the dashed vertical line) during the measurements was a public holiday in Austria.

Next analysis is about the measurement values itself. 99,8% of the measurements are in the range of 2ms to 1s. The relative frequency of those RTTs is depicted in Figure 5.2. The peaks appearing at several levels could already indicate different duration of the RTT because of different countries or ASes.

To complete this picture with the values above one second, Figure 5.3 shows the appearance of large RTTs during the measurement campaign. As it can be seen from the graph, they appeared in bursts during the first two days and less frequent on the other

[2]Asia and Pacific counted as one continent then, while America is divided into two

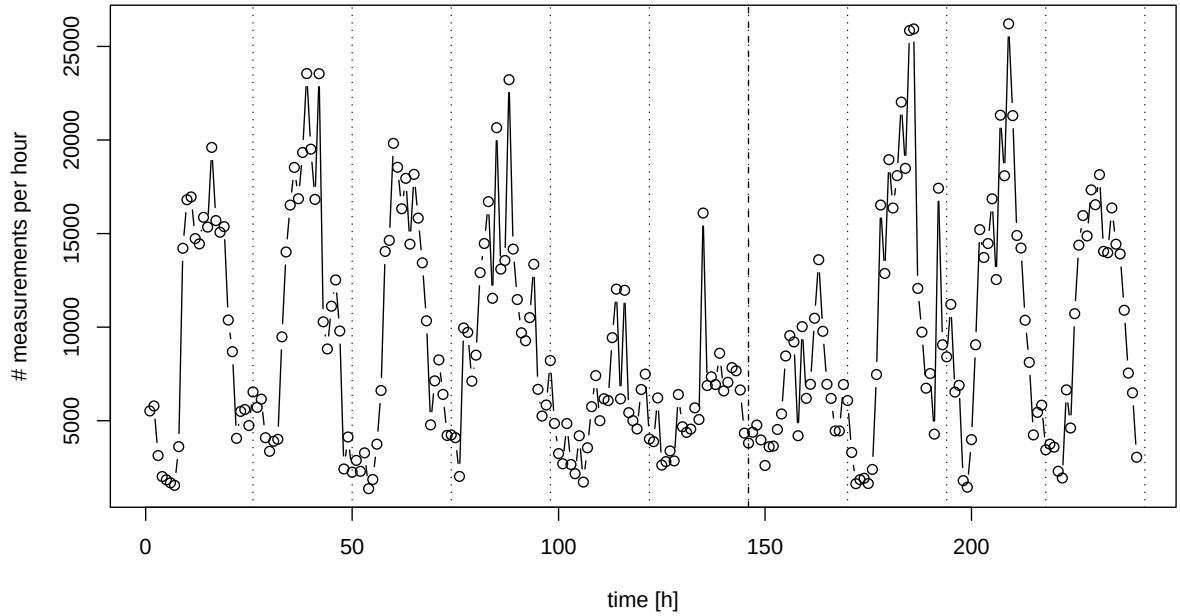


Figure 5.1: Total number of measurements generated per hour. It ranges between 1.300 and 26.000 measurements.

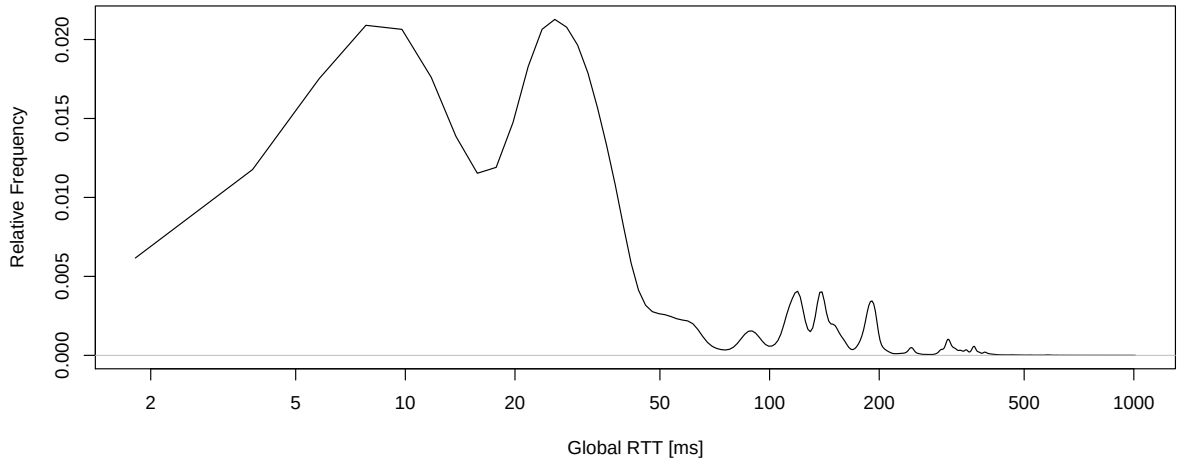


Figure 5.2: Relative frequencies of all RTT measurements below 1 second (99,8% of all measurements).

days. Therefore not all high delay values are single outliers, but may keep interesting information on the network and traffic situation that can be further studied.

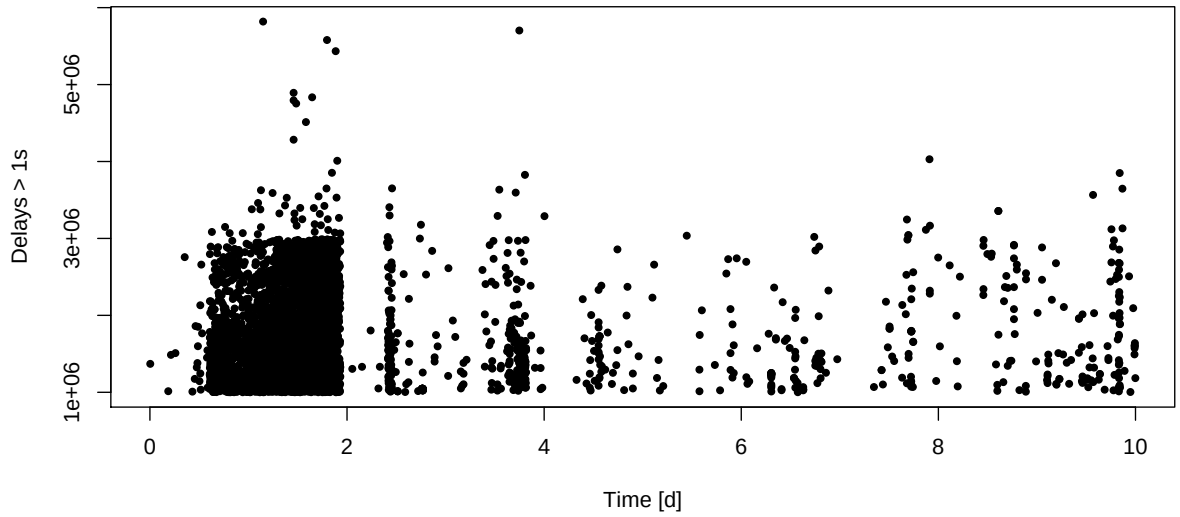


Figure 5.3: Appearance of large RTT measurements above one second (0,2% of all measurements).

Figure 5.4 shows the appearance of new ESes, as well as ASes. More than one third from the 3281 different ASes have been accessed within the first day. On the tenth day only 132 new ASes have been "detected".

5.3 RTT Analysis per End-System

Due to the nature of passive monitoring, measurement results only exist when users are generating traffic. This means, that measurement values can be very rare. In order to be still able to evaluate the delays to remote ASes, algorithms are necessary to make an estimation even if there was no measurement for e.g. several hours and "fade out" the estimate, if measurement values are too old. A detailed study about the significance of delay measurements was conducted by Choi et.al. [8]. They propose to report the 95% or 99% delay quantiles based on an estimation interval of 10-30 minutes for meaningful results of active measurement campaigns. If an estimation based on passive monitoring can be produced in a similar interval with a sufficiently high confidence level the active measurements can become unnecessary.

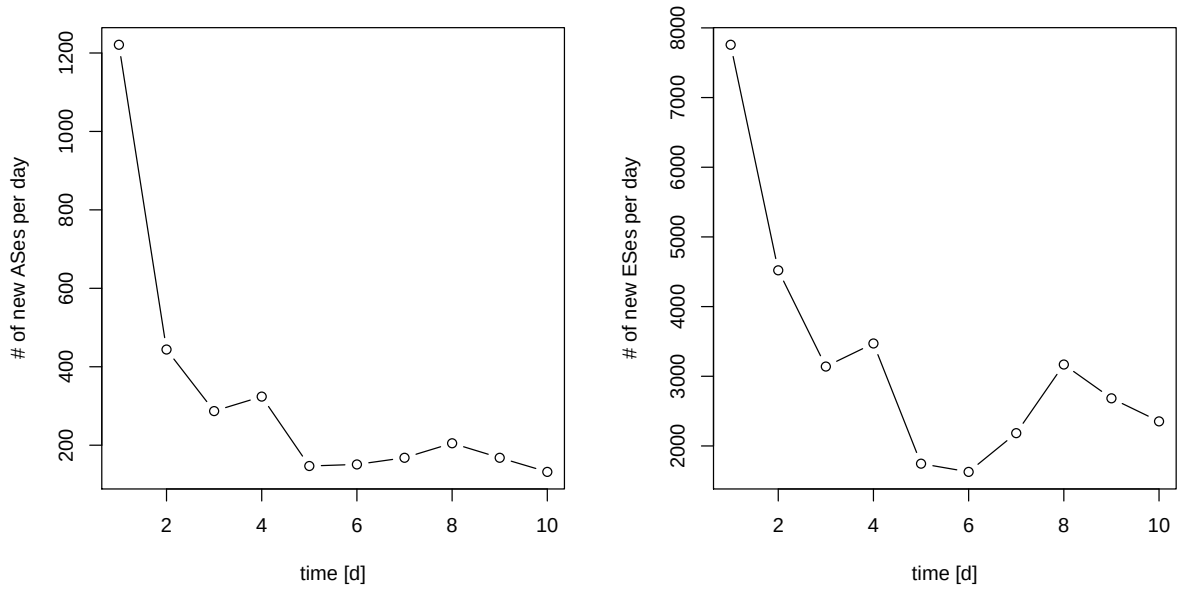


Figure 5.4: Number of newly seen Autonomous Systems (graph on the left) and end-systems (graph on the right) per day.

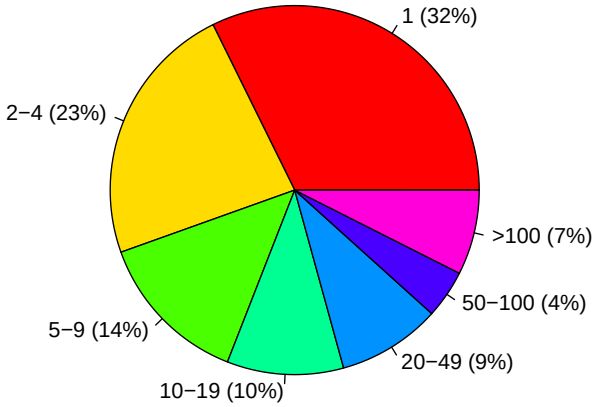


Figure 5.5: Number of measurements reported per End-System. The percentage value indicates the ratio of the total number of more than 32685 ESes.

5.3.1 Number of Measurements per End-System

Not all end-systems deliver a useful number of measurements to perform a detailed "perES"-study. It shoes that 7% of the end-systems produced more than 100 measurement results, while 32% only a single one. The distribution of the intensities collected over the whole 10 days of measurement is depicted in Figure 5.5. Having such rare events, switching to higher aggregation levels (BGP prefix or AS number) is useful to have a higher number of measurements per system under evaluation.

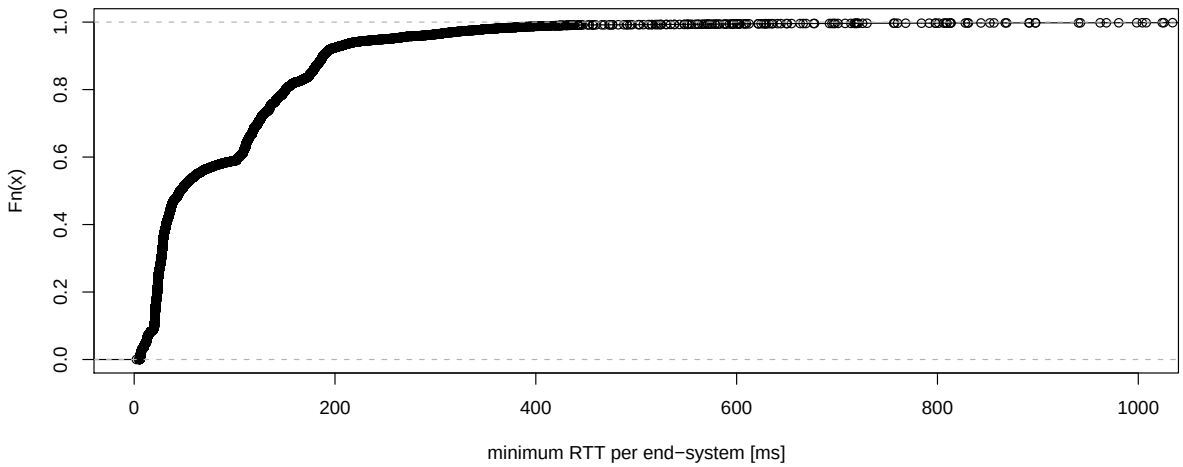


Figure 5.6: Empirical cumulative distribution function of the minimum RTT per end-system.

5.3.2 Minimum Measurement Value

As already mentioned in Section 3.5.1, the minimum measurement value plays an important role in the delay estimation. Given a reasonable amount of samples, the minimum can be interpreted as a constant delay value that includes the propagation delay and the link delay (packet length divided by the link capacity). The minimum delay is therefore directly proportional to the length of the packet. With knowledge about the packet length of the packets, the minimum delay can therefore be used to extrapolate RTT values also for different packet sizes.

Figure 5.6 shows the empirical cumulative distribution function Fn(x) (ECDF) from the minimum of the measured RTTs per end-system. There are some values above one second which have been removed for presentation reasons. An interesting "step" in the distribution function is at approximately 100ms: 60% of the end-systems have their minimum below. A possible reason for this step may be the separation between continental and inter-continental traffic.

5.3.3 Appearance of Measurement Events

The appearance of the measurement events per end-system is important in order to know how regular they are. While this can be controlled in case of active measurements, passive monitoring depends on the external trigger of traffic to generate such events. In this section, the time interval between two measurements generated by the same ES

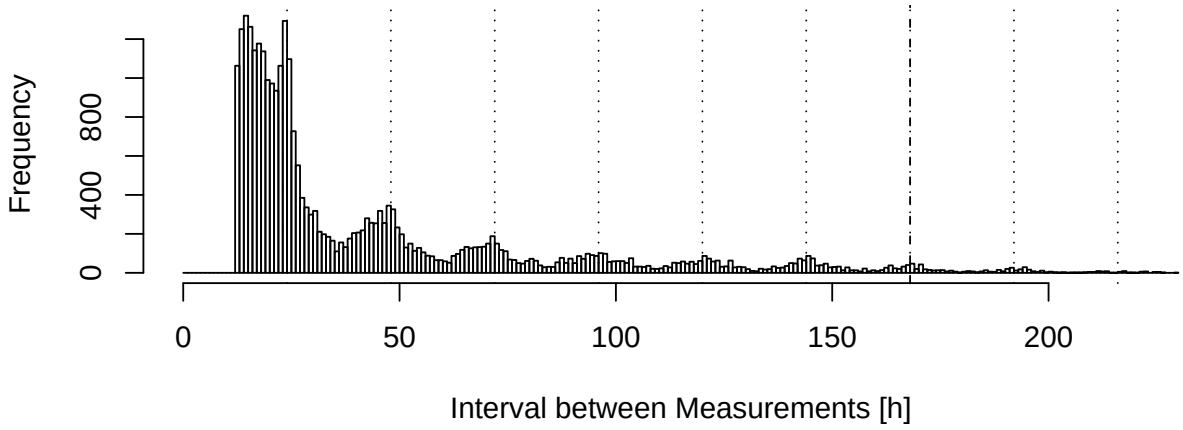


Figure 5.7: Histogram of the Timestamp Difference per ES, for values larger than 12 hours. Peaks at 24/48/72h etc. are visible.

is studied.

First interesting parameter is that already 95% of the measurement values occur within less than one hour to the previous measurement event of the same ES.

The distribution of the measurement interval between two measurements to the same ES is depicted in Figure 5.7. Vertical dotted lines are drawn at 24h/48h/72h/etc. and the vertical dashed line represents the interval of one week. Peaks in the daily intervals are visible. Note that this histogram only displays values above a time interval of 12h, which means that this graph only represents 1.3% of all values. Figure 5.13 in the next section shows the same on an AS level where daily peaks are still visible, but are less dominant compared to the analysis per ES.

The dominant group of measurements are very close to their respective predecessing measurement. The number of measurements with differences betweeen one and ten minutes to their predecessing measurement are depicted in Figure 5.8. Minor peaks of different height are at one, two, three and five minutes. Those and also the daily peaks are most likely related to regularly scheduled processes at the end-systems or their communication partners. This effect is smoothed on the AS-level as depicted in Figure 5.14 later on.

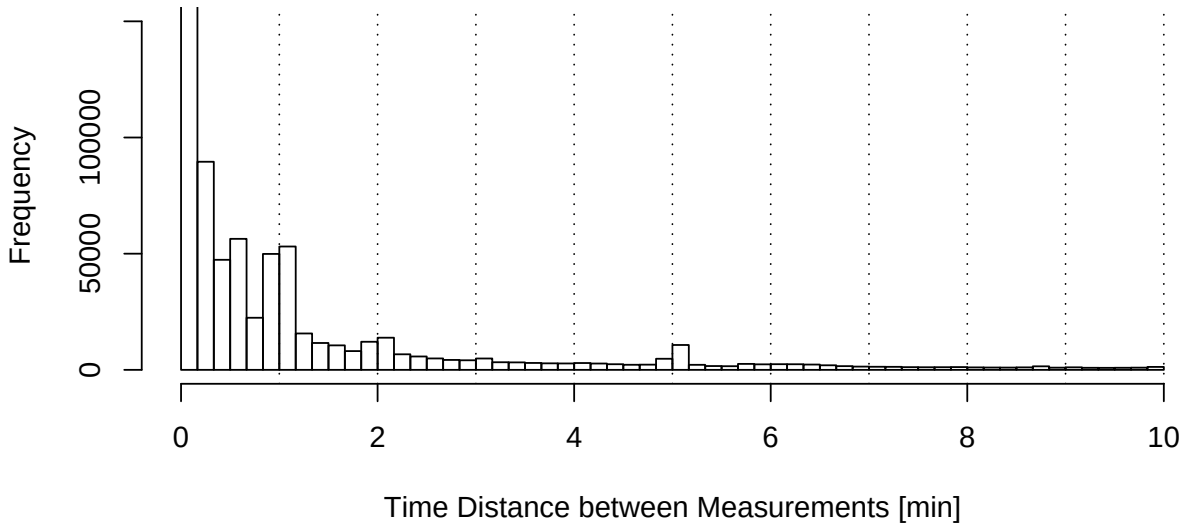


Figure 5.8: Histogram of the Timestamp Difference per End System, for values smaller than 10 minutes. Note that the large amount values below 10s (leftmost bin, 69% of all measurements) is cut to see the other effects.

5.3.4 Mean Values and Standard Deviation

Figure 5.9 shows the empirical standard deviations for each end-system with more than 30 values in relation to their mean values. As being expected, the higher mean values, the higher the variance of the values, in general. What is also visible in the figure is that there are a few ESes with low variations having high mean values, but several with a low mean value and very high variance. This effect is mainly visible because of single outliers and can be reduced by the application of the smoothing algorithm as described in Section 5.3.6 below.

In order to express the mean value as an estimate including a confidence interval, the samples must approximate a normal distribution. Testing for such distribution is described in the next section.

The $1 - \alpha$ confidence interval for the true mean in case of a normal distribution is calculated as:

$$\left[\overline{RTT} - \frac{\sigma}{\sqrt{n}} u_{1-\alpha/2} \; ; \; \overline{RTT} + \frac{\sigma}{\sqrt{n}} u_{1-\alpha/2}\right] \tag{5.1}$$

where $u_{1-\alpha/2}$ is the $(1 - \alpha/2)$-quantile of the normal distribution. For the common

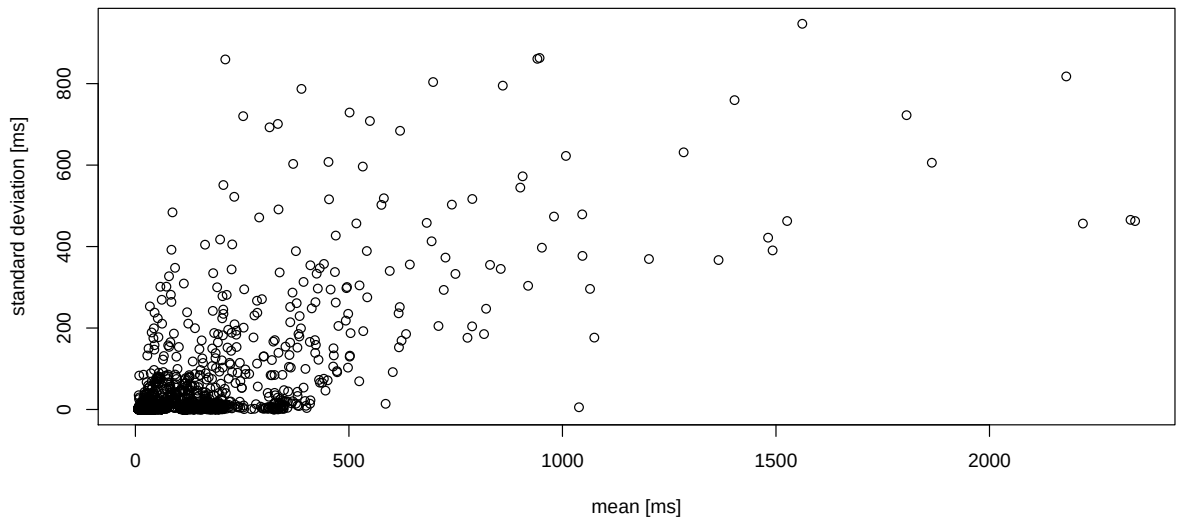


Figure 5.9: Plot of Mean Values (μ) and Standard Deviations (σ) per ES having more than 30 measurements.

significance level of α at 5% the value of $(1 - \alpha/2) = 0.975$. Calculating the 0.975-quantile of the normal distribution results in $u_{1-\alpha/2} = 1.96$.

Applied to the RTT measurements per end-system, the true RTT value within a 95%-confidence interval under the assumption of having an approximate normal distribution can be given as as:

$$RTT = \mu_{RTT} \pm 1.96\frac{\sigma_{RTT}}{\sqrt{n}} \tag{5.2}$$

where μ_{RTT} is the calculated mean value of the samples, σ_{RTT} is the calculated standard deviation and n is the number of measurements.

In Figure 5.10, two example histograms from two "busy" end-systems are shown. They are dedicated as busy, because all the depicted measurements weres collected within the same hour (5325 measurements in the left histogram and 12397 measurements in the right histogram). For the approximately normal distributed example (right histogram, AS33070), the true mean within the 95% confidence interval is

$$RTT = 150.156332 \pm 0.011097 = [150.145235; 150.167430] \tag{5.3}$$

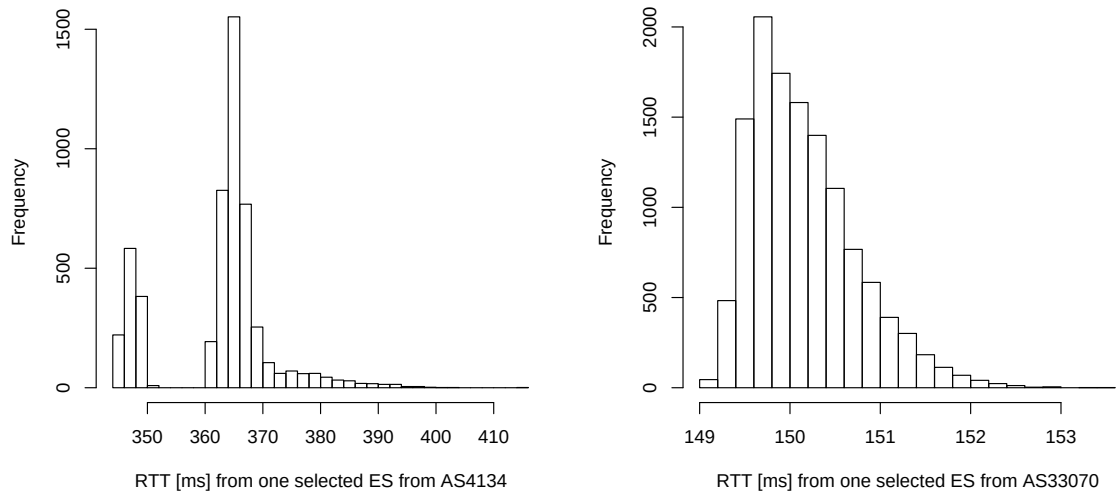


Figure 5.10: Example RTT Histograms of two heavily used End-Systems.

5.3.5 Distribution Testing

The relative frequency of all measurement values shown in Figure 5.2 displays a multimodal distribution with several peaks. Investigating the RTT distribution per ES provides a slightly different picture, but still the distribution of RTT per ES is not normal distributed and many cases. In addition to the previously used two example histograms from Figure 5.10, the complete set of normalised and logarithmised histograms of the RTT distributions per ES are provided on the CD attached to this thesis. To test against the normal distribution, the Kolmogorov-Smirnov goodness of fit (KS-GOF) test for unknown μ and σ can been applied, which is also qualified for a small number of samples[3], and is applicable also for unknown mean and variance values [23]. In the case that RTT is approximately normal distributed a confidence interval for the mean value can be indicated. The KS-GOF test basically calculates the maximum distance between the empirical and the theoretical distribution. If the maximum distance is larger than a given value (which depends on the significance level and the number of samples), the hypothesis of the distribution must be discarded.

When using the significance level of $\alpha = 5\%$, only in less than 20% of the cases in the

[3]Using a sample size of 30 measurements per end-system, about 5000 end-systems can be used for testing due to the number of measurements. Reducing the number of samples to 5, almost 15000 end-systems can be used.

given sample, the hypothesis of having a normal distributed sample must be discarded according to the KS-GOF test when using a sample size of 5. It can be expected that the RTT varies during the measurement period of 10 days. Therefore the KS-GOF test has not been applied per ES for the full duration of the measurement period, but instead using only RTT measurements taken with the same hour timeslot. From the 32000 end-systems times 240 hours from the given measurement sample, we have 60000 of such slots containing more than 5 measurements. Using those slots instead of taking samples from the whole 10-days period, in only about 15% of the cases, the hypothesis is discarded with $\alpha = 5\%$. Referring back to the example distributions depicted in Figure 5.10, the ES from AS33070 would return a normal-distributed sample with a much higher probability than the one from AS4134.

In case the hypothesis of being a normal distribution can be kept, the true mean values of the ESes can be supplied as an estimation parameter together with a confidence interval e.g. on the level of 95% as described above in Section 5.3.4.

5.3.6 Smoothing of Measurement Values

Extremely high values (outliers) are possible in the RTT measurements, mainly due to packet loss. Packet losses, especially during the connection establishment (TCP handshake) result are extremely painful for the quality experienced by the user. This is because of the required default value for the initial retransmission timeouts (RTO) of three seconds used by the standard-conform TCP implementations. However, those single events are not representative for the estimation of the AS-level delay. Therefore they should be taken out of the samples. Different smoothing strategies on the measurement values can be applied, where one of the effective ones is a running median. As the outliers appear as very single events, using three or five for the median window already perfectly degrades those outliers to their neighbouring values. Another strategy that has been applied in the PRISM project was to reduce new measurement values to the 95% quantile of the 100 previous values.

Discussions were introduced in the IETF to reduce the initial TCP RTO, and there are also some implementations using a smaller timeout in order to reduce the user

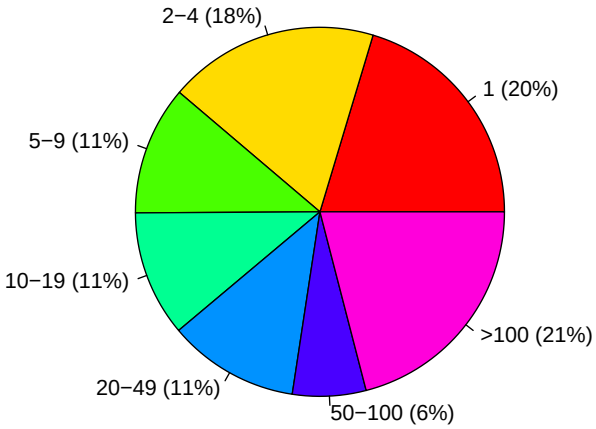


Figure 5.11: Number of measurements reported per Autonomous System. The percentage value indicates the ratio of the total number of 3281 ASes.

experienced delay after a packet loss during the three-way handshake [9].

5.4 RTT Analysis per Autonomous System

After going into the details of analysing single ESes, similar analysis has been performed on AS-level. First, the pie chart in Figure 5.11 depicts the number of measurements available per AS. Compared to the ES-level the number of systems with more than 5 measurements could be increased from 45% to 62%.

Next, Figure 5.12 shows the empirical cumulative distribution functions from the minimum, mean and maximum of the measured RTTs per Autonomous System. Compared to the minimum curve per ES in Figure 5.6, the minimum converges quicker against one, already at 500ms. This means that single outliers are less probably to appear on the higher aggregation level. The same sharp bend at 100ms is visible. Also visible is that the mean value is closer to the minimum than to the maximum, due to the assymetric distribution of the measurement values.

Figures 5.13 and 5.14 present the apperance of measurement events for all ASes, i.e. the time interval between two measurements to the same target AS. On ES-level both, daily peaks and per-minute peaks were visible in Figures 5.7 and Figure 5.8. On AS-level the situation is different and those peaks are not that significant anymore, as shown in Figure 5.13 for the daily peaks and Figure 5.14 for the per-minute peaks.

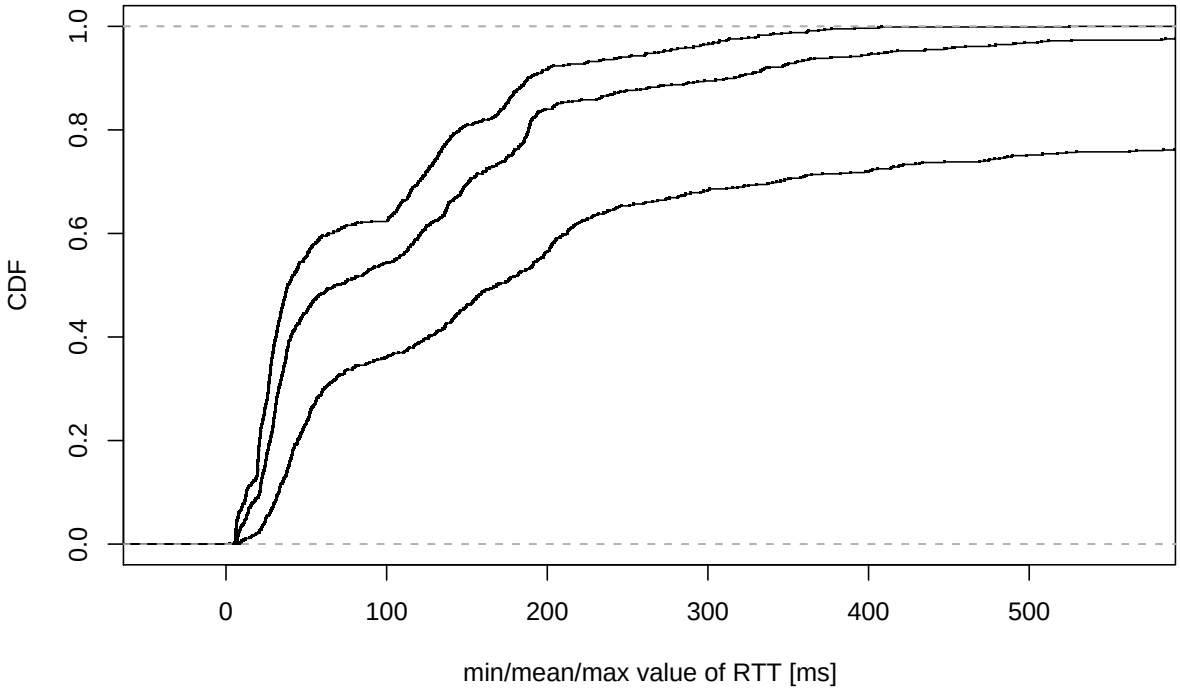


Figure 5.12: Plot of `min` (upper), `mean` (middle) and `max` (lower) empirical cumulative distribution function of the RTT per Autonomous System.

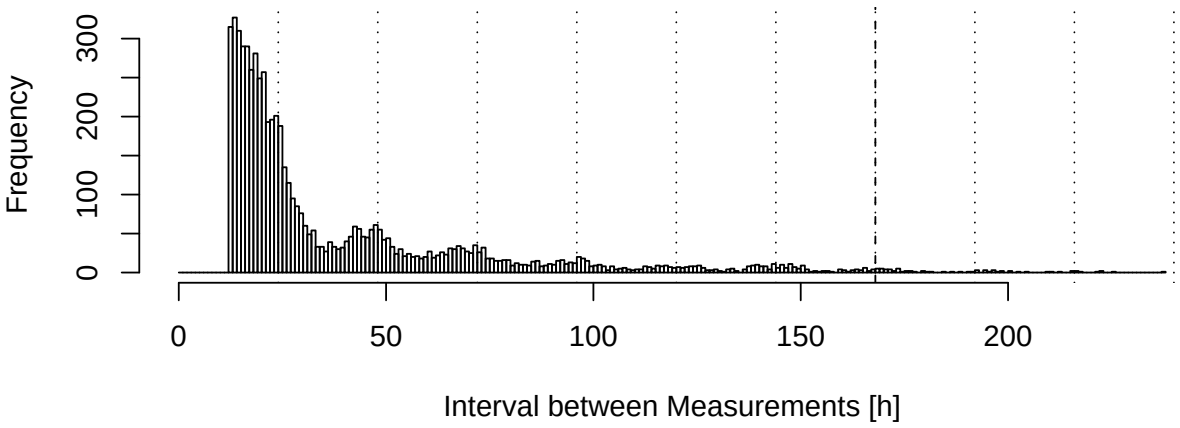


Figure 5.13: Histogram of the Timestamp Difference per Autonomous System, for values larger than 12 hours.

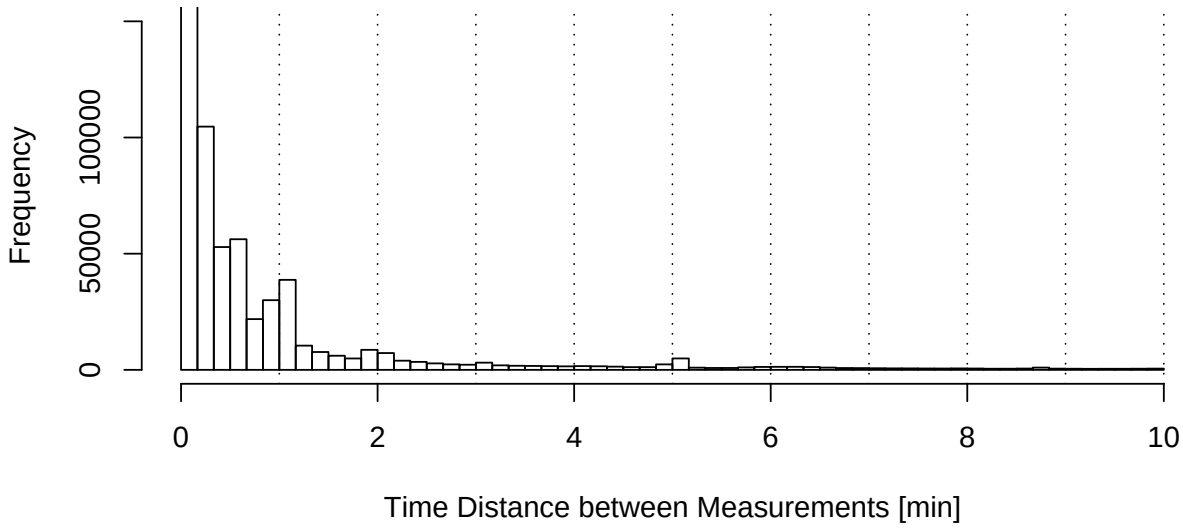


Figure 5.14: Histogram of the Timestamp Difference per Autonomous System, for smaller than 10 minutes, peaks like at the ES level are not significant here. Note that the large amount values below 10s (leftmost bin, 79% of all measurements) is cut.

In Table 5.1, six ASes are listed, where interesting behaviours became visible during the investigation of the data set. They are also a representative set of differently behaving ASes. It must be noted here that the mentioned behaviours are no globally valid expressions for those target ASes, but are only a limited interpretation from a single observation point of view.

5.4.1 Mean Values and Standard Deviation

This section provides some information on statistical parameters on the AS-level. In Table 5.2, the parameters for the selected AS from Table 5.1 are listed.

To provide the full picture, Figure 5.15 has been generated. Again, systems with less than 30 measurements have been discarded from this plot. Compared the Figure 5.9, which provides the same plot on the ES-level, less extremes are visible, i.e. there are less low mean values with high variance. Also high mean values with low variances do not appear anymore.

ASN	AS Name	Behaviour
8075	Microsoft	Among the networks with most measurements available in total. Between 3000 and 11000 measurements per day. 745 ES in 26 networks, on 6 different but constant levels between 40ms and 120ms. Such AS consists of many networks distributed around the globe.
8437	UTA	Very low RTT in general (6-8ms) distributed among 64 end-systems in 12 different networks. One ES generates high values up to 400ms within a three minutes time interval. Such case can have two causes: Either the complete ES is considered as outlier, or the connection was overloaded/impaired during the time interval.
32934	Facebook	Two levels of RTTs, both of them high in relation to the number of measurements and the popularity of the service. The levels exactly correspond to two separate networks: `69.63.184.0/21` with 13 end-systems has an RTT around 120ms and `69.63.176.0/21` with 67 end-systems around 190-200ms.
15169	Google	Most measurement events of all ASes (7000-14000 per day), largest number of different end-systems
5403	APA	Commonly used service in Austria (includes the online platform from the Austrian National Television orf.at), low RTTs, 82 end-systems, all of them in the same prefix.
10157	Yahoo! Korea	Geographically remote, high RTTs, regular measurement events over full measurement period.

Table 5.1: Description on the behaviours of six selected Autonomous Systems.

AS Name	**min**	**max**	μ	σ/μ	**RTTs**	**NWs**	**ESes**
Yahoo! Korea	331	529	350	9%	2.768	1	5
APA	6	161	7	24%	32.470	1	86
Facebook	116	351	176	18%	52.456	2	87
UTA	7	398	10	130%	35.609	14	88
Microsoft	35	3119	98	59%	71.996	27	745
Google	20	1121	33	92%	101.616	23	1153

Table 5.2: Statistics of selected Autonomous Systems. Unit of min, max and μ is $[ms]$. The right part of the table contains the counts of measurements (RTTs), different network prefixes (NWs) and end-systems (ESes).

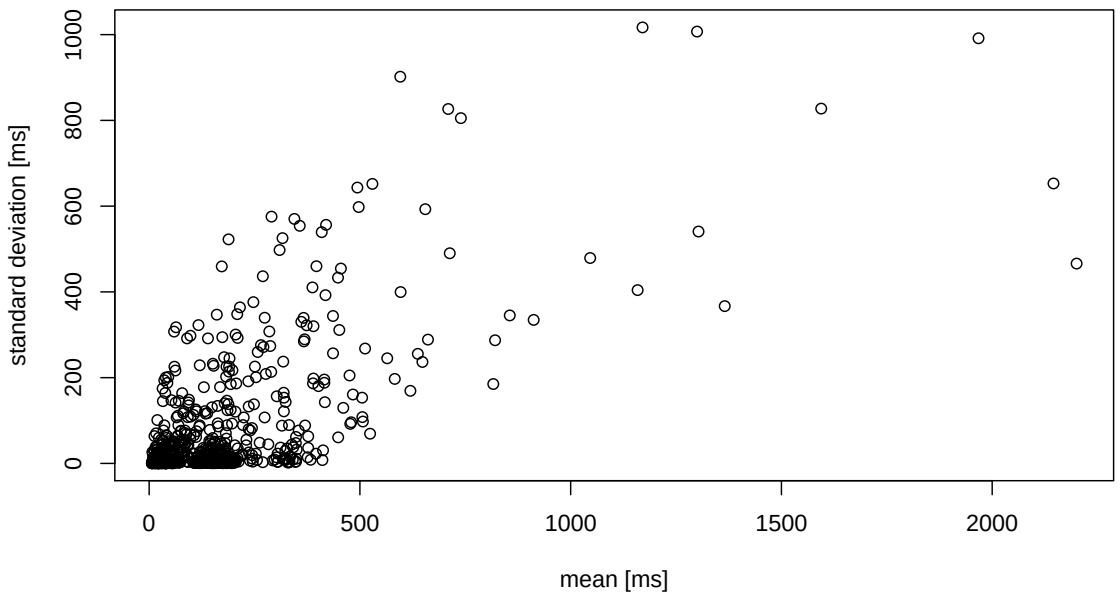


Figure 5.15: Plot of Mean Values and Standard Deviations per Autonomous System (with more than 30 measurements).

5.4.2 BGP Prefixes per Autonomous System

Table 5.2 provides the number of BGP prefixes interpreted as separate networks for the selected ASes in column “NWs”. The BGP-level study was introduced, because several ASes have shown in their RTTs that different RTT levels exist. They are either be based on the (connection) performance of different ESes or different, dislocated networks (i.e. BGP prefixes). In less than 40% of the used Autonomous Systems more than one BGP prefix has been used. Those ASes have been analysed on a “perPrefix”-basis and have been compared to the “perES”-basis. Visualising the RTT measurements with different colours based either on the prefix or the full IP address, some ASes (AS13008, AS20940) show a clear classification by BGP prefix, while others (AS15169) still have different levels within one BGP prefix. Such visualisations are available in colour on the attached CD for several ASes.

5.5 Acquisition of IP Address Meta-Information

An IP address implies additional information, which often has only temporal validity. As already mentioned in Chapter 4, IP addresses can even be treated as personal identifiable information (PII). Although many anonymisation algorithms have been proposed and implemented to make IP addresses unrecoverable for unauthorised people, there were at least the same number of attempts to revert the IP addresses from anonymised traces to their original numbers. For aggregation of measurement data on higher abstraction levels, the IP addresses need to be mapped to a more general information, which cannot be reverted by a 1:1 mapping.

The following list contains meta-information about IP addresses that are provided by public or commercial services:

- **AS number**: The number of the Autonomous System to which the IP address belongs. The mapping can change if networks are administratively moved from one AS to another.
- **AS name**: The identifying name of the Autonomous System, being a 1:1 relation to the AS number.
- **BGP prefix**: Identifier of the IP networks the IP address belongs to.
- **Allocation date**: The date, when the IP address has been initially provided by the RIR. Since the introduction of Classless Inter-domain Routing, unused IP addresses from large networks can be returned and reallocated.
- **Regional Internet Registry**: One of the five regional registries responsible for reallocating IP addresses they get from the IANA pool.
- **Country Code**: A country code, usually for the administrative location of the IP address. See Section 3.5 for more details.
- **Region Code**: A subcountry code, like a region or state. E.g. Austria is divided into 9 regions (one for each state).
- **City**: A city name, in the US coupled with the Metro Code.
- **Lat/Lon**: Approximates of the GPS coordinates. The usually maximum accuracy are the coordinates of the guessed city.

- **ISP**: The local provider of the IP address.
- **Organisation (whois)**: The administrative owner of the IP address.
- **netspeed**: The guessed connection speed of the IP address.
- **Domain name (nslookup)**: The corresponding DNS name of the IP address.

Additional information can be derived from the information above, like timezone data or postal codes. As it can be seen from the list, several fields can only be provided as approximates or on low accuracy levels.

Depending on the life-time of services, the at least the two following levels need to be considered:

- **Fluctuating IPs**: IP addresses that are distributed by stub providers to their customers are often withdrawn and reassigned after some hours, so that they can serve more customers with a smaller pool of IP address. Those customers cannot run a permanent service on the same IP address.
- **Quasi-fix IPs**: Long-term running web services deployed on servers with static IPs. Those addresses tend to change only on a monthly to yearly basis, or even not at all.

In the following some example services are listed that provide meta-information to IP addresses. This list is not intended to be complete, but provide an idea what public and commercial services with what granularity of information are freely available.

5.5.1 RIPE Database

The *RIPE database* provided by the European Regional Internet Registry, contains registration information about IP addresses and AS numbers, allocated by RIPE NCC. RIPE NCC is an independent, not-for-profit membership organisation that supports the infrastructure of the Internet through technical coordination in its service region, covering Europe, Middle East and parts of Asia[4]. The RIPE database [50] is the operational `whois` database and can be searched by providing IP address or AS numbers to its

[4] `http://www.ripe.net`

web-interface. For illustration, sample queries and results are shown in Appendix A.2. A related project to query Internet number resources is the recently launched Internet Number and Resource Database (INRDB) [31]. It provides information not only from the operational RIPE database, but also from IANA and the RIPE RIS project [51].

5.5.2 Team Cymru

Team Cymru[5] is an US-based research firm specialised on Internet security. Among many other mainly security related services, they offer an IP-to-ASN mapping service, providing the AS number plus additional information to a given IP address. The mapping service is based on information that is collected by BGP feeds gathered by more than 50 BGP peers. The update interval of the database is four hours. The service can be requested with an IP address or an AS number. When providing an IP address, the following information is returned by the mapping service.

- **BGP Origin ASN**: The Autonomous System Number (ASN) of the provided IP address.
- **BGP Peer ASN**: Possible mapping of the peer ASNs, that are one hop away from the BGP origin ASN's prefix. This feature may be useful for further analysis to investigate the IP's upstreams, but the website mentions that the method is far from being perfect.
- **BGP Prefix**: The BGP prefix of the provided IP address in the a.b.c.d/n notation.
- **Prefix Country Code**: The assigned country code to the AS, in a two-digit format, as obtained by the RIRs using the ISO 3166-1-alpha-2 code[6].
- **Prefix Registry**: The assigned Regional Internet Registry (RIR), in lower-case letters, i.e. one of "`afrinic`", "`apnic`", "`arin`", "`lacnic`" or "`ripencc`".
- **Prefix Allocation date**: The date when the prefix or AS was allocated to the corresponding AS number.

[5] `http://www.team-cymru.org`
[6] `http://www.iso.org/iso/country_codes/iso_3166_code_lists/`

Austria's city accuracy in the MaxMind databases	GeoLite	GeoIP
Correctly Resolved Within 40km of True Location	74%	79%
Incorrectly Resolved More Than 40km from True Location	23%	20%
Not Covered on a City Level	3%	1%

Table 5.3: Accuracy levels of MaxMind Databases: GeoLite City and GeoIP City for Austria. Values for all available countries are specified in [38].

- **ASN Description**: The descriptive name for the Autonomous System as provided by the daily regenerated CIDR-Report [26].

The service is made accessible through different interfaces. For quick queries of single IPs, http and https interfaces are provided, to make bulk queries, a `whois` server is maintained, which can be queried using the netcat tool [21], or directly by using TCP or UDP. Using this method, the service claims that bulk queries with around 100.000 IP addresses can return the results in less than a minute [55].

5.5.3 MaxMind GeoIP

MaxMind[7], a privately held company describes itself as an industry-leading provider of geographic location detection tools. Main customers of their GeoIP product are Internet service providers that want to pinpoint the location of their customers and visitors to the granularity of their city in real-time. Beneath their commercial products, they also provide free open source databases GeoLite Country and GeoLite City. GeoLite County can is available for download in both binary and CSV file format. GeoLite City is provided as CSV file with 130MB. The accuracy mentioned on the product web-site is 99,5% on the country level, nominal a much better fit than it can be achieved by resolving the country by the location of the AS allocation. The city level is less accurate and specified differently per country. Figures for Austria are given in Table 5.3. The accuracy of city levels is not satisfying, but is only of limited interest for RTT measurements, except for huge countries like USA or Russia.

[7] `http://www.maxmind.com`

6

The Implementation of the Autonomous System Delay Finder and Integration in the PRISM System Architecture

This chapter contains the description of the Autonomous System Delay Finder (ASDF), which has been implemented in the framework of the PRISM project. Due to the requirements of the PRISM privacy framework, the ASDF is separated into two distinct tasks. The lower layer, which implements the calculation of the RTT based on packet capturing information is implemented in the PRISM front-end. From there the information is forwarded via IPFIX to the core of the ASDF, which does the mapping from the IP addresses to higher-level information (like ASNs and countries, etc).

Before going into the details of the ASDF, an overview on the PRISM project, the architecture and other example monitoring applications is provided in Section 6.1. Afterwards, Section 6.2 describes different use cases of ASDF with the specification of monitoring purposes. Finally, Section 6.3 about the integration of ADSF into the PRISM system concludes this chapter.

6.1 An Overview on the PRISM Project

PRISM (**Pri**vacy-aware **S**ecure **M**onitoring) targeted on breaking the dilemma of privacy-versus-utility of Internet monitoring. PRISM was an international research project in the seventh framework programme (FP7) of the European Union. In this framework, PRISM contributed to the theme "Secure, Dependable and Trusted Infrastructures". The project has been carried out by eight project participants from six European countries between March 2008 and June 2010. Many findings presented in this thesis are based on this cooperative work.

To show the technical possibility to design a privacy-preserving monitoring system without loosing the utility of today's existing monitoring applications, an integrated architecture has been presented, prototypically implemented and validated by running multiple applications, partially with specific adaptations to the PRISM system. To proof the overall concept, the project furthermore assessed the provided system in the dimensions of regulation, performance and security.

The remainder of this section shortly describes the PRISM architecture as well as three example applications to demonstrate the applicability of the system.

6.1.1 The PRISM Architecture

To achieve the project goals, a two-tier monitoring architecture has been designed. It includes three major components. The entity on the lowest level is called *front-end* and is bound to the link to be monitored. The PRISM front-end is designed to cope network speeds up to ten Gigabit per second, and provides already the initial step of data reduction. Only data that is strictly required for further processing for one of the predefined monitoring purposes, is encrypted and forwarded to the *back-end*. The back-end's main responsibilities are to safely store the retrieved data, manage the access to this data and also further process the monitored data for specific requirements. The required information for these tasks is provided by the third main component, the *privacy-preserving controller* (PPC), representing the "Source of Authority". The PPC is located outside of the data flow, but hosts the semantic information about possible

monitoring purposes, the users and their roles and manages authorisation in the whole PRISM system. Attached to the PRISM components are the network link itself on the front-end side and the external monitoring application on the back-end side. The external monitoring application can either be an existing legacy application, or a new application specifically implemented for the PRISM system. Figure 6.1 shows the flow of the data-plane communication as well as the front-end and the back-end tasks. As protocol between the front-end and the back-end IETF's IP Flow Information Export (IPFIX) protocol [10] has been selected. IPFIX is a standard protocol for measurement and monitoring data export and provides great flexibility and extensibility.

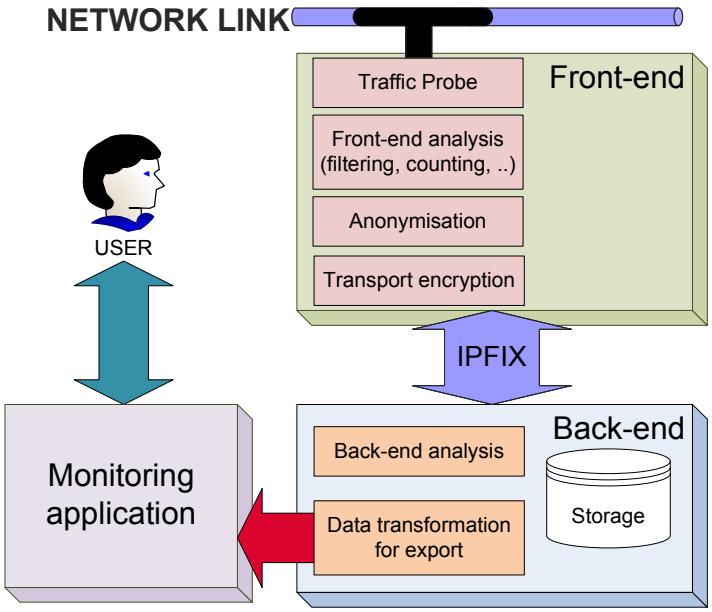


Figure 6.1: The PRISM Architecture components and their interaction [14].

6.1.2 Example Monitoring Applications in PRISM

Three example applications were carefully selected among the major areas of passive monitoring applications and have been implemented in the PRISM context as a reference for demonstration. The three areas are "intrusion detection", "traffic classification" and "performance monitoring". The applications are shortly described as follows. Details are provided in the PRISM project deliverable by Dorfinger et. al. in [15]. The first use case was to implement a privacy-preserving intrusion detection

software based on Snort[1]. The software was split into separate parts, some of them running in the front-end and others in the back-end. As second application a Skype detection engine has been connected to PRISM to demonstrate the traffic classification scenario. The main goal here was to keep the Skype detection engine untouched and completely external to the PRISM architecture by running it as legacy application. Only the input traffic has been modified to reduce the information as much as possible. The goal was to already pre-filter packets, which do definitely not contain Skype traffic, and only forward packets to the "real" detection engine those flows and packets that contain Skype traffic with a reasonable probability. The third scenario was the Autonomous System Delay Finder (ASDF) to demonstrate a privacy-preserving performance monitoring scenario. This application was implemented from scratch and was therefore directly adjusted to the requirements of the PRISM architecture. Components running on the front-end, the back-end and as external application were exactly tailored to the PRISM needs. It is described in detail in in the following section.

6.2 ASDF Monitoring Scenarios and Purposes in PRISM

Two different use case scenarios are defined in this section, which are afterwards broken down into specific monitoring purposes as required by the PRISM architecture. The monitoring purposes were associated to predefined user roles. These scenarios have also been presented in [54].

6.2.1 Scenarios for Privacy-Preserving Network Delay Evaluation

To observe and investigate the behaviour of the network delay, two different usage scenarios are outlined. The first one is used to get an overview on how remote ASes are connected to the monitoring location, the latter one is designed to do a step-by-step evaluation of poorly reachable ASes. Due to the architectural model of PRISM,

[1] `http://www.snort.org`

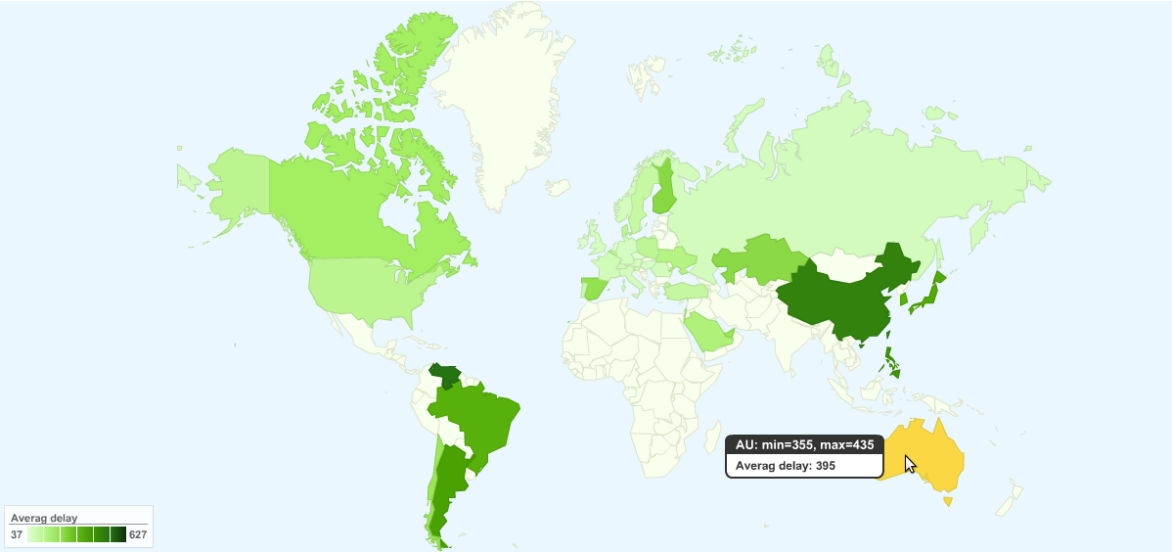


Figure 6.2: Example Map to visualise the AS Delay per Country using Google Chart Tools. The map shows the average delay to all ASes for each country. The measurement point used to generate this map was active for more than one hour and was located in Italy.

each use case needs to be built from one or more specific monitoring purposes with associated user roles. Depending on the user role, a monitoring purpose is allowed to be executed or not. Those monitoring purposes are further described in Section 6.2.2. The two scenarios are an Internet reachability map and the detection and investigation of problematic Autonomous Systems.

6.2.1.1 An Autonomous System Reachability Map

The first use case to look at is the visualisation of an intensity map that provides locations of ASes in different colours, where a user can get a live overview on the reachability of different countries. Countries are based on the administrative location of the AS, which do not necessarily correspond to the physical one. Despite of this limitation, such map provides a good impression to which regions in the world the network customers communicate and how those remote regions are connected by the ISP. An example AS Reachability Map is shown in Figure 6.2. An additional table (not shown in the figure) allows the user to investigate the details about each of the ASes, and can be sorted per country, RTT or the name/number of the ASes.

6.2.1.2 Detection and Investigation of Problematic AS

The second possible use case demonstrated in PRISM is a troubleshooting scenario. When given thresholds of delays are passed, a company customer care staff member receives notifications, and in further steps the problems behind can be investigated. Customer care staff is usually triggered by a customer experiencing problems with his connection. Knowledge about problematic target ASes help to prevent from searching for solutions of an end-user specific problems, while an AS-wide problem exists. In this case, a more privileged staff member from the network management section will be able to get a detailed look on the IP addresses involved in communication with the problematic AS, and can perform a detailed analysis on this. To see if the problem was persistent in the past or not, additional historical data can be requested for the AS in question.

6.2.2 Definition of Monitoring Purposes

Before implementing the two described usage scenarios in the PRISM context, the definition of monitoring purposes is required. Stored in the PRISM ontology [34], they are connected to the access control model. There it is strictly defined, which user is allowed to access what purposes during what time. As an initial step for ASDF four monitoring purposes have been defined: One for the visualisation of a reachability map for remote ASes (and their administrative country), one for alarm generation after threshold crossing, one for the long-term collection of measurement data, and finally one for detailed analysis, where all measurements including the IP addresses are reported. Three roles from the PRISM ontology were selected to allow these purposes, namely staff members from “company customer care” (lowest privileges), “technical customer support” and the “network management section” (highest privileges). In this section, those monitoring purposes and their mapping to the access control are described. The defined access model is related to the real staff member roles applied by an operational provider. The definition of monitoring purposes is required to control and log access to prevent misuse of the monitoring system.

6.2.2.1 AS Reachability Map

The reachability map reports an estimation of the delay to remote ASes based on the recently measured RTTs. If enabled, the PRISM system periodically reports the status of each AS accessed in the past. The report includes the name and the number of the AS, the country code and RIR for the AS, the timestamp of the latest measurement to this AS, the RTT in milliseconds and the number of measurements, which were used to calculate the RTT value. ASes are reported only if a certain number of measurements (e.g. 20) have been received. This purpose may be run by users in the technical customer support role or higher privileged personnel.

6.2.2.2 AS Reachability Map With IP Addresses

This is the most verbose purpose and is therefore restricted to selected, trustworthy staff members. It reports all RTT measurements to all remote ASes with their name, number, country and RIR. Every RTT report includes the remote IP addresses, the timestamp and the RTT in microseconds. Due to privacy-sensitive information of IP addresses, only members from the network management section are allowed to perform this monitoring task. As additional security mechanism for starting such monitoring tasks the existence of an alarm can be required. Due to access logging all operations stay traceable to uncover illegal activities.

6.2.2.3 AS Alarm Generation

This purpose is limited to report remote ASes, where the delay evaluation passed a threshold, defined either globally or separately for single ASes. The basic idea of this task is to limit the exported information for two reasons. First, the monitoring system should only report ASes which require performance improvements, and secondly to allow users with limited privileges, like the company customer care staff to be made aware of problems to specific remote locations.

6.2.2.4 AS Historical Data

The final purpose is designed to allow a long-term investigation of the delay per AS. It is required to investigate the progress of the RTT to remote ASes and to determine, whether the RTT value changed only short-term, periodically or if there is already a longer-term problem. Once such task is started, the monitoring system collects RTT data per AS and stores it timestamped into a local, encrypted database. This is the only purpose, where data is stored within the monitoring system, while the above purposes report their results directly to the user in real-time. Technically, this purpose is divided into two sub-purposes. One for storing the data, and one for requesting the data from the storage. Both are only allowed to be used by the network management section.

6.3 Integration of ASDF in the PRISM System

This section describes how the functionalities of the ASDF has been split among the PRISM components. Components running on the front-end, the back-end and in the external application have different security and trust levels. Another advantage of the strict cut between collection and processing allows also to include alternative sources of RTT measurements, like active `traceroute`s or `ping`s.

6.3.1 Estimation of RTT in the PRISM Front-End

The PRISM front-end is the most sensitive part of the architecture because it monitors the traffic on the operational network links. Beneath being highly secure, the front-end must also be high-performance and able to cope with Gigabits of traffic per second in real-time. Therefore the front-end algorithms must be highly optimised. As mentioned in Section 3.4, several algorithms exist to get RTT measurements from a passive network monitor. The methods need to be selected depending on the amount of data and the link location (stub or transit network). The PRISM field-trial was conducted on an uplink from a stub provider in Italy, which provides bidirectional traffic from their

customers to the Internet. In this environment the SYN-ACK estimation algorithm (as described in [30]), and a similar method based on the connection termination procedure of TCP have been considered important. Both have been implemented in the `rtte` application, called *HANDSHAKE* and *TERMINATION*. Some implementation details are provided in the following.

The PRISM front-end has a *capturing unit* and a *processing unit*, which are the same also for different applications running in the PRISM environment. In the ASDF case, the capturing unit captures all TCP flows and already drops the payload of the packets. The processing unit analysis provides the results of RTT measurements (in microseconds) per IP address, based on TCP flows received from the capturing unit. Its design consists of a robust subject/observer pattern. This allows one object[2] to be an observer of another object, which is the subject under observation. The observer will get notified in case of any registered events triggered from the subject. In the `rtte` context, the subject and observers have been assigned with the following tasks: The *subject* is the entity in charge of capturing and classifying TCP segments on a per-flow basis, while the *observers* implement algorithms in charge of making different RTT estimations. Such a design is extensible, in that it is very easy to either add additional algorithms (as observers) or to convert the application from single thread to multithread in order to take advantage of multiple CPU cores or multiple CPUs.

Internet traffic can be provided to `rtte` by means of any traffic source: The application makes use of pcap library to collect packets and therefore both, pcap files and network devices like *eth0* are valid inputs. The exported RTT estimations are triples of: i) timestamp, ii) the remote IP address and iii) the RTT between observation point and remote IP address. These are sent to an IPFIX collector using the template specified in Section 3.2.2 or to an XML file, depending on the options provided at command line. The command line options allow the architectural framework to call the monitoring sub-system in the front-end with different configuration parameters. This can be used to adapt the configuration for different monitoring purposes.

[2]The implementation was done in C++ and *object* refers to C++ objects here.

6.3.2 Aggregation of Measurements in the PRISM Back-End

As IP addresses must be treated as personal information, data aggregation is performed in an embedded processing component of the PRISM back-end before exporting the data to the external application. From the different levels of aggregation granularity that can be considered based on the IP address information the levels of AS numbers and countries are evaluated in the ASDF scenarios.

Like the exported information from the front-end, the back-end continuously receives triples of timestamp, IP address and RTT, which are processed in batches of 30 seconds. The back-end has two major tasks. The IP-address-to-AS mapping and the calculation of an RTT estimate per AS from the number of measurements received from the front-end.

To map IP addresses that are gathered from the measurements into AS numbers, information from an external mapping service is required. The chosen service for ASDF is provided by *Team Cymru* as already described in Section 5.5.2. For ASDF the `whois`-interface was chosen as it allows bulk queries for multiple IP addresses. Depending on the number of requested IP addresses, each query includes additional addresses that do not appear on the monitored network, for privacy reasons. In order to reduce the load on the external service, ASDF keeps a local cache of mappings between BGP prefixes and AS information. Entries have also stored a timestamp to allow renewing the cached information after expiration. The local database providing this cache contains two tables connected by a many-to-many relation, one for the AS information (including the threshold used to generate alarms), and one for the BGP prefixes. The simple physical database model for the local ASDF cache is depicted in Figure 6.3.

The second task is the estimation of the RTT for an AS, which is based on all measurements that have been mapped to this AS. For the calculations, different statistical functions are provided, like moving means, moving medians, or exponentially moving averages with or without outlier elimination as described in Section 3.4. For the PRISM trials the EWMA has been used, where outliers have been downgraded in the calculation based on the 95%-quantile. New values have been weighted with α of 0.2.

Figure 6.3: Physical database model for the local ASDF cache. The left table stores AS-related information while the right one stores information about address prefixes. The middle table represents the relation between those two.

One problem in the IP to AS mapping that is reflected in the ASDF is that multihomed networks and hosts may return multiple AS numbers. From the available information it is not possible anymore to dedicate, through which of the ASes the packets were routed. Such measurements are identified and can be either dropped or related to one or all associated ASes.

As described in Section 6.2.2, ASDF produces different results based on the chosen monitoring purpose during the aggregation step. As an example, the investigation of routing problems as discussed in Section 6.3.3. It is based on generated alarms, once given thresholds of the RTT are passed. Thresholds can be defined globally or separately for each of the ASes.

6.3.3 Identification of Routing Problems in an External Application

Once an alarm is generated due to a poorly reachable AS, additional investigation can be performed to locate the problem. Poor connection quality to remote ASes can have various reasons. Problems either exist only short-term due to traffic or routing conditions or long-term due to the network topology or even physical constraints. Short-term means on the protocol level (milliseconds to minutes) or network management level (hours to days), while long-term in this context means on network planning level (weeks to month). The obvious physical constraint responsible for long delays to remote ASes is the geographic distance to the measurement point. More interesting

reasons are temporary or permanently overloaded routers or links along the path, but also overloaded end-systems. The most interesting reason for long and varying delays is the fact that routes are unstable. Such instabilities were investigated e.g. in [46].

First the user needs to distinguish, whether the alarm was generated because of a single end-system or because of multiple end-systems in the reported AS, in order to evaluate the seriousness of the alarm. Therefore the detailed monitoring purpose which reports also the remote IP addresses must be activated. This will show the different IPs of the AS in trouble, and how the RTT is distributed amongst them. In case the problem is on a whole target AS, further investigation can be performed by starting additional active measurements. The drawback of such approach is that measurements can be done only from the subjective viewpoint. When querying further external data sources, an objective view can be gathered. One comfortable way to get information of distributed data sources is the *MOMENT mediator*, which enables a semantic uniform access to distributed data sources of Internet measurement data [20]. The MOMENT mediator can answer semantic SPAQRL queries towards concepts available in the MOMENT ontology. Two example queries are given below: "Return all measurements where the source IP address was in a specific range" is expressed in SPARQL like:

```
PREFIX xsd: <http://www.w3.org/2001/XMLSchema>
PREFIX MD: <http://www.fp7-moment.eu/MomentDataV2.owl#>
SELECT DISTINCT ?a WHERE {?a a MD:SourceIP;
  MD:SourceIPValue ?x.
  FILTER (?x >"3339139328"^^xsd:int )
  FILTER (?x <"3339139349"^^xsd:int ) }
LIMIT 100
```

A query about the AS path like "Give me all AS-path which terminate to in AS 9551" is expressed in SPARQL as follows:

```
PREFIX MD: <http://www.fp7-moment.eu/MomentDataV2.owl#>
SELECT ?aspath { ?path MD:AsPathValue ?aspath.
  FILTER (regex(?aspath,".+ 9551","i")) }
ORDER BY DESC(?aspath)
LIMIT 100
```

The quality of the MOMENT mediator results depends on the quality of the data in the underlying data sources. Therefore the information about the actual data source

still needs to be verified by the user. However the big advantage of using the mediator with the underlying MOMENT ontology is to overcome the heterogenity of data sources. Data representation (units and data types, etc.), information model (database schemas, file system structures, etc.) as well as query language (CSV files, SQL, web services, etc.) are largely different in the common data sources. An ontology turns the distributed information into structured knowledge. Different measurement units can be transformed automatically, like dotted IP addresses into integer values. All mediated data sources can be accessed by raw or pre-defined SPARQL queries.

7

Evaluation of the Architecture

This section evaluates some problems that can arise due errors during the RTT measurement during the mapping of IP addresses to higher aggregated identifiers like AS numbers.

7.1 Measurement Errors

The measurement errors investigated may occur due to performance or accuracy limitations either of the capturing hardware or the processing software. Also software implementation errors may deliver some errornous results as described in Section 7.1.4. Those errors are only discovered, if TCP does not behave as expected, but produces some sequence of packets that never occurred during the testing period. Therefore it may be difficult to detect those errors under normal testing conditions.

7.1.1 Timestamping

Packets can be timestamped on several places during their traversal through the host system. In the PRISM use case, the timestamping happened already during the capturing process at the very lower end of the network stack. However, running `rtte` on a usual Linux system using the pcap library, timestamping can be a problem, as it depends on the operating system, the driver, and the hardware. However, the timestamping issue must always seen together with the time to be measured. As long as the

relation between measured time and timestamp accuracy is less than 1:10, it should be save for correct interpretation of the measurements for most of the applications. For very time-critical applications, the timestamp accuracy must be accordingly more accurate.

7.1.2 Packets Lost during the Capturing Process

Standard network interfaces produce buffer overwrites in the kernel, in case the data cannot be delivered to the capturing process. In order to capture 100% of the packets on the link, a dedicated monitoring hardware must be used. There is high-speed network capturing hardware available on the market, like the DAG®cards[1] from Endace, or the Network Processor from Intel®[2].

7.1.3 End-System Delays

As already discussed in Section 1, also the end-systems add their delays as one component to the results of the RTT measurements. Their inclusion has both a drawback and an advantage. The advantage is that the end-system delay is important to be measured, as it also influences the QoE of the user. The drawback is that end-system delays cannot be separated from the network delays, which means that a single low performing end-system can drop the overall performance of the measured AS. Therefore high RTTs from one end-system must be weighted less than high RTTs from many end-systems in the same AS.

In order to evaluate how the load (CPU, I/O, memory and hard disk) of the server can influence the measurement results, a test under laboratory conditions was made. Therefore a web sever (Apache) has been set up, which has been requested every second by using `jmeter` [56]. The test has been conducted with a zero and full loaded CPU in the HTTP server. The results are depicted in Figure 7.1. The thin solid line shows the empirical frequency of RTTs measured every second to the web server process running on an unloaded CPU. The thin dashed line (most right one) show the results produced

[1] `http://www.endace.com/endace-dag-high-speed-packet-capture-cards.html`
[2] `http://www.intel.com/design/network/products/npfamily/ixp425.htm`

with a fully loaded CPU/IO/memory and hard disk[3]. Interesting are the results, where the web server was requested many times (100 times per second, thick lines), also in loaded and unloaded CPU states. Probably due to some performance optimisation mechanism of the system, the results are even better than under unloaded conditions. The frequency curves show that the RTT measured by `rtte` is slightly higher the loaded condition. Under the bottom line, differences in the mean value are in the range of 30 microseconds and therefore it can be assumed that the load situation of the end-host does not significantly influence the measured RTT when using the TCP SYN-ACK algorithm.

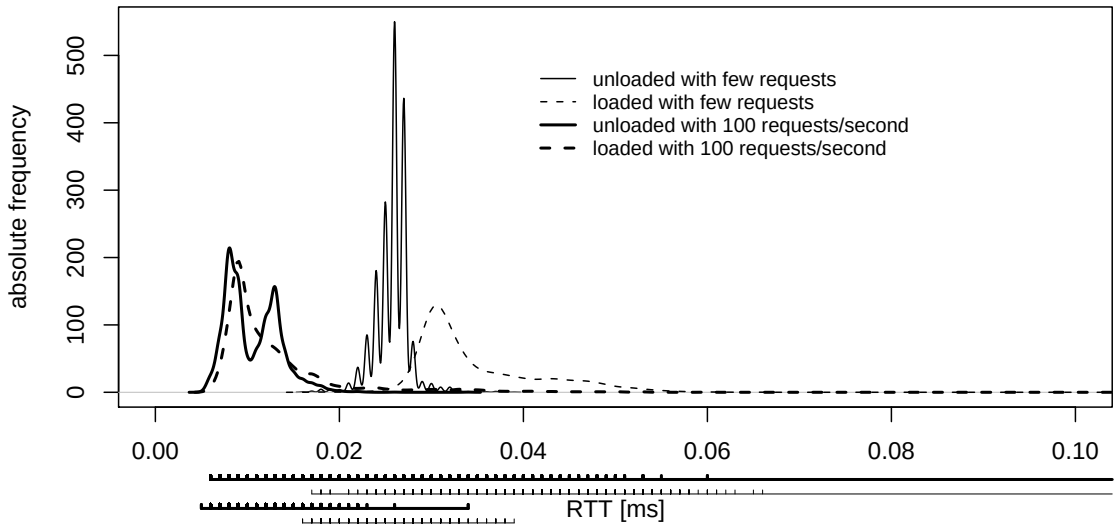


Figure 7.1: Histogram of the measurement results for end-system internal delay evaluation.

7.1.4 Double RTT Measurements for Single Initial SYN

SYN-ACKs may appear twice (or even more often) to acknowledge the same SYN. This can happen for example when the ACKs from the client is being lost after the observation point. In that case, the `rtte` implementation produced two measurement results. Such measurement values need to be removed. The example timestamp at

[3]The `stress` command has been used with the following parameters:
`stress --cpu 8 --io 4 --vm 2 --hdd 1 --timeout 30s`

0.260 seconds as shown below has even three identical measurement values. Accuracy has been lost during ASDF, where timestamps are stored only in milliseconds, although many capturing platforms can accurately report the timestamps in microseconds. A look into the raw data turned out, that with microseconds accuracy, only two values are really identical (i.e. even same source and destination port), but one with a much higher RTT. It turn out that a second ACK was received after some seconds, and another entry for the initial SYN has been made. It can be expected, that the `rtte` implementation can cope with lost packets and usually does not do duplicate reporting. However, for the analysis in this thesis, 34 erroneous values were kept by the software and have been removed manually. Based on the total number of measurements, the error rate produced by this bug is at 0.0015%. The bug in `rtte` has been reported to the implementers. An example of such conversation is given below:

```
0.260 IP B.3503 > A.443: S 0:0(0) win 5840 <mss 1460,sackOK,timestamp 0 0,nop,wscale 0>
0.446 IP A.443 > B.3503: S 0:0(0) ack 1 win 5792 <mss 1460,sackOK,timestamp 0 0,nop,wscale 7>
0.447 IP B.3503 > A.443: . ack 1 win 5840 <nop,nop,timestamp 18 0>
0.450 IP B.3503 > A.443: P 1:110(109) ack 1 win 5840 <nop,nop,timestamp 19 0>
0.989 IP B.3503 > A.443: P 1:110(109) ack 1 win 5840 <nop,nop,timestamp 73 0>
2.069 IP B.3503 > A.443: P 1:110(109) ack 1 win 5840 <nop,nop,timestamp 181 0>
4.230 IP B.3503 > A.443: P 1:110(109) ack 1 win 5840 <nop,nop,timestamp 397 0>
4.553 IP A.443 > B.3503: S 0:0(0) ack 1 win 5792 <mss 1460,sackOK,timestamp 1025 397,nop,wscale 7>
4.554 IP B.3503 > A.443: . ack 1 win 5840 <nop,nop,timestamp 429 1025,nop,nop,sack 1 {0:1}>
8.550 IP B.3503 > A.443: P 1:110(109) ack 1 win 5840 <nop,nop,timestamp 829 1025>
10.542 IP A.443 > B.3503: S 0:0(0) ack 1 win 5792 <mss 1460,sackOK,timestamp 2525 829,nop,wscale 7>
10.543 IP B.3503 > A.443: . ack 1 win 5840 <nop,nop,timestamp 1028 2525,nop,nop,sack 1 {0:1}>
17.191 IP B.3503 > A.443: P 1:110(109) ack 1 win 5840 <nop,nop,timestamp 1693 2525>
23.551 IP A.443 > B.3503: S 0:0(0) ack 1 win 5792 <mss 1460,sackOK,timestamp 5525 1693,nop,wscale 7>
23.552 IP B.3503 > A.443: . ack 1 win 5840 <nop,nop,timestamp 2229 5525,nop,nop,sack 1 {0:1}>
35.474 IP B.3503 > A.443: P 1:110(109) ack 1 win 5840 <nop,nop,timestamp 3421 5525>
47.755 IP A.443 > B.3503: S 0:0(0) ack 1 win 5792 <mss 1460,sackOK,timestamp 11576 3421,nop,wscale 7>
47.757 IP B.3503 > A.443: . ack 1 win 5840 <nop,nop,timestamp 4649 11576,nop,nop,sack 1 {0:1}>
70.038 IP B.3503 > A.443: P 1:110(109) ack 1 win 5840 <nop,nop,timestamp 6877 11576>
95.958 IP A.443 > B.3503: S 0:0(0) ack 1 win 5792 <mss 1460,sackOK,timestamp 23626 6877,nop,wscale 7>
95.960 IP B.3503 > A.443: . ack 1 win 5840 <nop,nop,timestamp 9468 23626,nop,nop,sack 1 {0:1}>
139.167 IP B.3503 > A.443: P 1:110(109) ack 1 win 5840 <nop,nop,timestamp 13789 23626>
139.353 IP A.443 > B.3503: . ack 110 win 46 <nop,nop,timestamp 34476 13789>
211.409 IP A.443 > B.3503: . 1:1449(1448) ack 110 win 46 <nop,nop,timestamp 52487 13789>
211.409 IP A.443 > B.3503: P 1449:2336(887) ack 110 win 46 <nop,nop,timestamp 52487 13789>
211.411 IP B.3503 > A.443: . ack 1449 win 8688 <nop,nop,timestamp 21012 52487>
211.412 IP B.3503 > A.443: . ack 2336 win 11584 <nop,nop,timestamp 21012 52487>
211.682 IP B.3503 > A.443: . 110:1558(1448) ack 2336 win 11584 <nop,nop,timestamp 21039 52487>
211.683 IP B.3503 > A.443: P 1558:2283(725) ack 2336 win 11584 <nop,nop,timestamp 21039 52487>
211.873 IP A.443 > B.3503: . ack 1558 win 68 <nop,nop,timestamp 52604 21039>
```

Also the TERMINATION algorithm, which reports the RTTs measured during the connection closure using the FIN and FIN-ACK packets, made problems during the evaluation of the measurement values. The destination address for the measured RTT is sometimes exchanged with the source address, depending on which of the two communicating hosts initiates the connection closure. As they represent just a little portion

compared to the measurements taken by the HANDSHAKE algorithm ($\approx 1.4\%$), those measurements were also removed before the analyses taken in Chapter 5.

7.1.5 Initial RTO

As already mentioned in Section 5.3.6, outliers are being produced mainly by packet losses. Standard TCP implementations set the initial retransmission timeout to three seconds, i.e. when no SYN-ACK has been received within the initial RTO, a second SYN is being sent. The evaluations of RTT measurements have shown a significant decrease of the amount of measurements above three seconds most probably caused by this initial timeout (also visible in Figure 5.3). This means that the packet capturing process ignores the second SYN packet from the same connection for the RTT calculation, but takes the time difference between the finally arriving SYN-ACK and the initial SYN packet. This behaviour might lead to wrong interpretation of the measurement results, as not the RTT is that high, but a packet loss occurred. Like passive RTT estimation there are also works on passive packet loss estimation [4].

7.2 Mapping Errors

Address prefixes and routes are announced and withdrawn quite frequently from the ISPs, due to mergers or other reorganisations in the global Internet business. Weekly changes of prefix announcements clustered by ASes are available in the CIDR Report [26]. Because of this fluctuation the local cache of the mapping function can be out of date quickly, and therefore needs to be regularly updated to avoid mapping failures. An improvement against constantly updating the local cache within specific time intervals would be to perform a regular check on changing prefixes, e.g. by requesting the CIDR Report. The updates in the local cache can then be limited to those networks that have withdrawn or announced new routes recently.

Also the mapping to countries can contain errors, as many ISPs operate on an international level. Therefore the mapping to countries on IP level compared to AS level

can largely differ. In the data set used in Chapter 5 more than 32000 unique IP addresses were present. In less than 15% of the IP addresses (and also less then 15% of the measurements) the country mapping between IP-level and AS-level differs. It therefore makes sense to consider both mappings seperately, depending on what conclusion needs to be drawn from the country code. Using a commercial service like GeoIP Country from MaxMind can again increase the accuracy of the mapping from 99,5% to 99,8% [38], except for AOL IP addresses that have a global mapping to the US country code. However, the real accuracy cannot be finally proven, especially due to the existence of mobile Internet and VPNs, IP addresses can actually be in totally different locations than the mapping information suggests.

Of course mapping errors become even larger in cases where the packet capturing process and the mapping process are timely decoupled.

8

Summary and Conclusion

The thesis presented the work on passive estimation of delays to remote Autonomous Systems. The application of passive monitoring technologies in operational networks can be problematic due to privacy reasons. The decoupling of the packet capturing process with pre-processing, the data evaluation and the data presentation process was proposed and prototypically implemented by the PRISM project. One of the implemented scenarios was the Autonomous System Delay Finder (ASDF), which required an algorithm to estimate the delay to remote Autonomous Systems based on the passive measurement of the round-trip time to remote hosts, identified by their IP addresses. To design such algorithm the delays have been first investigated per end-system, to see whether there is a relation between the delays of end-systems belonging to the same AS. A relation between RTT and AS could be proven, but the results showed that especially for large and heavily used ASes, a distinction on finer levels of granularity e.g. on BGP-prefix level, or even on a per end-system level is required.

8.1 Innovation Potential

The Internet has become the most important communication infrastructure, and moves forward into application areas, where best-effort services without performance monitoring are not acceptable anymore. As an example, emerging standards like the IEC/FDIS

80001-1 from ISO[1], about the application of risk management for IT-networks, incorporating medical devices requires specific activities on the safety and effectiveness of the underlying IT-network. Even if provided services already deliver good performance, it will become more important in the future to monitor SLAs for specific services and to declare some level of trust or risk for the network. Network monitoring, including performance monitoring, will therefore require even more attention in future networks, also those connected globally via the Internet. On the other hand, for many applications and especially from the field like the given example, user privacy is also extremely important. Network monitoring without some privacy-preserving framework will not be acceptable for the users, and therefore Internet service providers and network operators need to take corresponding actions.

The presented work allows generating innovative products in the area of Internet measurements. It supports Internet service providers in the privacy-preserving evaluation of the quality experienced by their customers. Countermeasures in case of low quality connections can be taken. Although such measurements can (and are) already performed in current deployments, the use of passive monitoring technologies can conflict with EU privacy-laws. As described in Chapter 4 of the thesis, storing and processing of IP address are problematic because of the possibility of deriving personal identifiable information (PII). Observing the customer requests on a higher aggregation level without the possibility to revert this information back to identify single users enables network monitoring with legal conformity.

Therefore, providers currently can deploy infrastructures for passive network monitoring only for internal use and for a limited set of applications and without the evaluation of IP addresses. The PRISM project prototypically developed a monitoring infrastructure that performs data collection and processing in a secure environment at the monitoring point and only exports the amount of information required by the monitoring application of the user. This allows the providers to make specific data available to a larger amount of users and even share this data with other providers and network operators. Sharing measurement data raises another concern. Beneath the protection of the privacy of the users, Internet providers are very sensitive about publishing or

[1] `http://www.iso.org/iso/catalogue_detail.htm?csnumber=44863`

sharing performance data. With a PRISM-like monitoring system, it is possible to share data based on non-disclosure agreements (NDAs), where all participating parties can grant restricted access to their data. Sharing data with the peering partners that are operating the neighbour networks helps to pinpoint performance bottlenecks in the inter-domain traffic.

The specific performance-related measurement data as produced by ASDF has another potential innovation: Current works either work on a per end-system basis, which is a too high granularity of data. The aggregation per Autonomous System, or even higher degrees of aggregation based on meaningful algorithms directly provides the data interesting for the network operator to answer the questions already stated at the beginning of the work:

- What are the important services demanded by the customers to what extend?
- Which Autonomous Systems provide the services used by the customers?
- What response time do the peering and transit networks provide to those services?
- What is the relation between performance and usage?
- Is there a need for improvements in the inter-domain connections of the network?

Finally, future work based on the provided implementation is useful for various further applications in different areas, for examples the detection of anomalies. Undesirable traffic crossing the network like denial of service attacks or port scans typically produce many connections to and/or from the same Internet host. The analysis performed in the thesis has even uncovered one of such events coincidentally.

8.2 Conclusion and Future Work

Several findings presented in this thesis would be interesting to be studied in more detail, but also enhancements in new directions, like the introduction of more metrics for AS classification would be of interest.

A first step of AS classification can be already done based on the data produced for this thesis. ASes can fore example not only be classified by there RTT performance

but also by other parameters, like the relative number of measurements (to identify the “degree of usage”) or the number end-systems or BGP prefixes visible.

Examples for new metrics, that can be investigated are

- Improved passive RTT measurements algorithms that are not only based on the three-way handshake, but consider the complete TCP flow
- New estimation methods for jitter
- New estimation methods for packet loss

The inclusion of several metrics would help to provide a more meaningful classification of Autonomous Systems. However, the overall objective needs to be kept in mind, which is to enable the ISPs to deploy such a system easily and to be able to draw meaningful high-level conclusions from the provided data.

Bibliography

[1] Alderson, D., Li, L., Willinger, W., and Doyle, J. C.: *Understanding internet topology: principles, models, and validation.* IEEE/ACM Trans. Netw., 13(6):1205–1218, 2005, ISSN 1063-6692. http://dx.doi.org/10.1109/TNET.2005.861250.

[2] Almes, G., Kalidindi, S., and Zekauskas, M.: *A One-way Delay Metric for IPPM.* RFC 2679 (Proposed Standard), September 1999. http://www.ietf.org/rfc/rfc2679.txt.

[3] Almes, G., Kalidindi, S., and Zekauskas, M.: *A Round-trip Delay Metric for IPPM.* RFC 2681 (Proposed Standard), September 1999. http://www.ietf.org/rfc/rfc2681.txt.

[4] Benko, P. and Veres, A.: *A Passive Method for Estimating End-to-End TCP Packet Loss.* In *Global Telecommunications Conference, 2002. GLOBECOM '02. IEEE*, pages 2609–2613 vol.3, November 2002. http://dx.doi.org/10.1109/GLOCOM.2002.1189102.

[5] Braden, R., Zhang, L., Berson, S., Herzog, S., and Jamin, S.: *Resource ReSerVation Protocol (RSVP) – Version 1 Functional Specification.* RFC 2205 (Proposed Standard), September 1997. http://www.ietf.org/rfc/rfc2205.txt, Updated by RFCs 2750, 3936, 4495.

[6] Chang, H.: *An economic-based empirical approach to modeling the internet's inter-domain topology and traffic matrix.* PhD thesis, University of Michigan, Ann Arbor, MI, USA, 2006, ISBN 0-542-56869-1. Adviser-Jamin, Sugih.

[7] Chang, H., Jamin, S., Morley, Z., and Willinger, W.: *An empirical approach to modeling inter-as traffic matrices.* In *IMC '05: Proceedings of the 5th ACM SIGCOMM conference on Internet Measurement*, pages 12–12, Berkeley, CA, USA, 2005. USENIX Association.

[8] Choi, B. Y., Moon, S., Cruz, R., Zhang, Z. L., and Diot, C.: *Quantile sampling for practical delay monitoring in Internet backbone networks.* Comput. Netw., 51(10):2701–2716, 2007, ISSN 1389-1286. http://dx.doi.org/10.1016/j.comnet.2006.11.023.

[9] Chu, H.K. J.: *Tuning TCP Parameters for the 21st Century.* Presentation held on the 75th IETF meeting in Stockholm, Sweden, July 2009. http://tools.ietf.org/agenda/75/slides/tcpm-1.pdf.

[10] Claise, B.: *Specification of the IP Flow Information Export (IPFIX) Protocol for the Exchange of IP Traffic Flow Information.* RFC 5101 (Proposed Standard), January 2008. http://www.ietf.org/rfc/rfc5101.txt.

[11] Cohen, R. and Raz, D.: *The Internet dark matter: on the missing links in the AS connectivity map.* In *Proceedings of the 25th Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM 2006)*, April 2006, ISBN 1-4244-0221-2. http://dx.doi.org/10.1109/INFOCOM.2006.234.

[12] Dhamdhere, A. and Dovrolis, C.: *Ten years in the evolution of the internet ecosystem.* In *IMC '08: Proceedings of the 8th ACM SIGCOMM conference on Internet measurement*, pages 183–196, New York, NY, USA, 2008. ACM, ISBN 978-1-60558-334-1. http://dx.doi.org/10.1145/1452520.1452543.

[13] Dimitropoulos, X., Krioukov, D., Riley, G., , and Claffy, K.: *Revealing the Autonomous System Taxonomy: The Machine Learning Approach.* In *In Proceedings of Passive and Active Measurement Conference (PAM), 2006.*, 2006. http://pamconf.net/2006/papers/s5-dimitropoulos.pdf.

[14] Dorfinger, P., Schmoll, C., and Strohmeier, F.: *Privacy-Aware Network Monitoring.* ERCIM News, April 2009, ISSN 0926-4981.

[15] Dorfinger, P., Strohmeier, F., Moosbrugger, A., Gojmerac, I., Trammell, B., Boschi, E., Bianchi, G., Procissi, G., Teofili, S., Nobile, E., Lioudakis, G., Gogoulos, F., Antonakopoulou, A., Kaklamani, D., Venieris, I., Matachowski, M., and Khavtasi, S.: *Final monitoring applications specification and analysis.* PRISM Project Deliverable D3.2.3, 2009.

[16] EU FP7 ICT Project Consortium: *FP7 ICT project "Perimeter" FP7 - 224024.* http://www.ict-perimeter.eu/ (accessed 15.07.2010).

[17] European Parliament and Council: *Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data.* EC Directive, 1995. http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:31995L0046:EN:HTML.

[18] European Parliament and Council: *Directive 2002/58/EC of 12 July 2002, concerning the processing of personal data and the protection of privacy in the electronic communications sector (Directive on privacy and electronic communication).* EC Directive, 2002. http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:32002L0058:EN:HTML.

[19] European Parliament and Council: *Directive 2006/24/EC of the European Parliament and of the Council of 15 March 2006 on the retention of data generated or processed in connection with the provision of publicly available electronic communications services or of public communications networks and amending Directive 2002/58/EC.* EC Directive, 2006. http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2006:105:0054:01:EN:HTML.

[20] Ferreiro, Á., Fichtel, T., Vergara, J. López de, Mátray, P., Strohmeier, F., Tropea, G., and Weinsberg, U.: *Semantic Unified Access to Traffic Measurement Systems for Internet Monitoring Service*. In *ICT-MobileSummit 2009 Conference Proceedings*, June 2009, ISBN 978-1-905824-12-0.

[21] Giacobbi, G.: *The GNU Netcat project*. `http://netcat.sourceforge.net/` (accessed 23.07.2010), 2004–2006.

[22] Gummadi, K. P., Saroiu, S., and Gribble, S. D.: *King: estimating latency between arbitrary internet end hosts*. SIGCOMM Comput. Commun. Rev., 32(3):11–11, 2002, ISSN 0146-4833. `http://dx.doi.org/10.1145/571697.571700`.

[23] Hartung, J., Elpelt, B., and Klösener, K. H.: *Statistik - Lehr- und Handbuch der angewandten Statistik*. R. Oldenbourg Verlag München Wien, 13th edition, 2002.

[24] Holleczek, T., Venus, V., and Naegele-Jackson, S.: *Statistical Analysis of IP Delay Measurements as a Basis for Network Alert Systems*. In *Proceedings of the 2009 IEEE International Conference on Communications (ICC 2009)*. IEEE, June 2009. `http://dx.doi.org/10.1109/ICC.2009.5199487`.

[25] Huston, G.: *The 32-bit AS Number Report*. `http://www.potaroo.net/tools/asn32/` (accessed 23.07.2010), 2010.

[26] Huston, G., Bates, T., and Smith, P.: *The CIDR Report*. `http://www.cidr-report.org/` (accessed 23.07.2010), 2010.

[27] ITU: *Information technology Quality of Service: Framework (ITU-T Recommendation X.641)*. International Telecommunications Union, December 1997. `http://www.itu.int/rec/T-REC-X.641-199712-I/en/`.

[28] ITU: *The E-model: a computational model for use in transmission planning (ITU-T Recommendation G.107)*. International Telecommunications Union, April 2009. `http://www.itu.int/rec/T-REC-G.107`.

[29] Jaiswal, S., Iannaccone, G., Diot, C., Kurose, J., and Towsley, D.: *Inferring TCP Connection Characteristics through Passive Measurements*. In *Proceedings of the 23rd Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM 2004)*, volume 3, pages 1582–1592. IEEE, March 2004, ISBN 0-7803-8355-9. `http://dx.doi.org/10.1109/INFCOM.2004.1354571`.

[30] Jiang, H. and Dovrolis, C.: *Passive estimation of TCP round-trip times*. SIGCOMM Comput. Commun. Rev., 32(3):75–88, 2002, ISSN 0146-4833. `http://dx.doi.org/10.1145/571697.571725`.

[31] Kisteleki, R., Karrenberg, D., Wilhelm, R., and Refice, T.: *The Internet Number Resource Database (INRDB)*. `http://labs.ripe.net/Members/kistel/content-intro-inrdb-internet-number-resource-database` (accessed 23.07.2010), 2009–2010.

[32] Kohnstamm, J.: *Letter from the Article 29 Working Party addressed to search engine operators (Google, Microsoft, Yahoo!)*, May 2010. http://ec.europa.eu/justice_home/fsj/privacy/docs/wpdocs/others/2010_05_26_letter_wp_google.pdf.

[33] Li, L., Alderson, D., Willinger, W., and Doyle, J.: *A first-principles approach to understanding the internet's router-level topology.* In *Proceedings of the 2004 conference on Applications, technologies, architectures, and protocols for computer communications, SIGCOMM '04*, pages 3–14, New York, NY, USA, 2004. ACM, ISBN 1-58113-862-8. http://dx.doi.org/10.1145/1015467.1015470.

[34] Lioudakis, G. V., Gogoulos, F., Antonakopoulou, A., Kaklamani, D. I., and Venieris, I. S.: *Privacy Protection in Passive Network Monitoring: An Access Control Approach.* In *International Conference on Advanced Information Networking and Applications Workshops*, pages 109–116, Los Alamitos, CA, USA, 2009. IEEE Computer Society, ISBN 978-0-7695-3639-2. http://doi.ieeecomputersociety.org/10.1109/WAINA.2009.158.

[35] Madhyastha, H. V., Isdal, T., Piatek, M., Dixon, C., Anderson, T., Krishnamurthy, A., and Venkataramani, A.: *iPlane: an information plane for distributed services.* In *OSDI '06: Proceedings of the 7th symposium on Operating systems design and implementation*, pages 367–380, Berkeley, CA, USA, 2006. USENIX Association, ISBN 1-931971-47-1. http://portal.acm.org/citation.cfm?id=1298455.1298490.

[36] Maier, G., Feldmann, A., Paxson, V., and Allman, M.: *On dominant characteristics of residential broadband internet traffic.* In *IMC '09: Proceedings of the 9th ACM SIGCOMM conference on Internet measurement conference*, pages 90–102, New York, NY, USA, 2009. ACM, ISBN 978-1-60558-771-4. http://dx.doi.org/10.1145/1644893.1644904.

[37] MaxMind Inc.: *GeoIP City/ISP/Organization Demo.* http://www.maxmind.com/app/lookup_city (accessed 23.07.2010), 2002–2010.

[38] MaxMind Inc.: *GeoLite City Accuracy for Selected Countries.* http://www.maxmind.com/app/geolite_city_accuracy (accessed 23.07.2010), 2002–2010.

[39] Merit Network, Inc.: *Internet Routing Registries.* http://www.irr.net (accessed 13.07.2010), 2000–2010.

[40] Morató, D., E. Magaña, Izal, M., Aracil, J., Naranjo, F., Astiz, F., Alonso, U., Csabai, I., Haga, P., Simon, G., Steger, J., and Vattay, G.: *The European Traffic Observatory Measurement Infrastructure (ETOMIC): A Testbed for Universal Active and Passive Measurements.* In *Proceedings of Testbeds and Research Infrastructures for the DEvelopment of NeTworks and COMmunities (TRIDENTCOM 2005)*, pages 283–289, Los Alamitos, CA, USA, 2005. IEEE Computer Society, ISBN 0-7695-2219-X. http://dx.doi.org/10.1109/TRIDNT.2005.34.

[41] Nichols, K., Blake, S., Baker, F., and Black, D.: *Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers.* RFC 2474 (Proposed

Standard), December 1998. http://www.ietf.org/rfc/rfc2474.txt, Updated by RFCs 3168, 3260.

[42] Oliveira, R. V., Zhang, B., and Zhang, L.: *Observing the evolution of internet as topology.* SIGCOMM Comput. Commun. Rev., 37(4):313–324, 2007, ISSN 0146-4833. http://dx.doi.org/10.1145/1282427.1282416.

[43] Österreichisches Parlament: *Datenschutzgesetz 2000 (DSG 2000), BGBl. I Nr. 165/1999.* Österreichisches Bundesgesetzblatt, 2000.

[44] Österreichisches Parlament: *Telekommunikationsgesetz 2003 (TKG 2003), BGBl. I Nr. 70/2003 idF. BGBl. I Nr. 133/2005.* Österreichisches Bundesgesetzblatt, 2003.

[45] Paxson, V. and Allman, M.: *Computing TCP's Retransmission Timer.* RFC 2988 (Proposed Standard), November 2000. http://www.ietf.org/rfc/rfc2988.txt.

[46] Pucha, H., Zhang, Y., Mao, Z. M., and Hu, Y. C.: *Understanding network delay changes caused by routing events.* SIGMETRICS Perform. Eval. Rev., 35(1):73–84, 2007, ISSN 0163-5999. http://dx.doi.org/10.1145/1269899.1254891.

[47] Quittek, J., Bryant, S., Claise, B., Aitken, P., and Meyer, J.: *Information Model for IP Flow Information Export.* RFC 5102 (Proposed Standard), January 2008. http://www.ietf.org/rfc/rfc5102.txt.

[48] Rao, S., Khavtasi, S., Bianchi, G., Teofili, S., Procissi, G., Di Pietro, A., Gojmerac, I., Hyytia, E., Boschi, E., Trammel, B., Lioudakis, G., Strohmeier, F., and Schmoll, C.: *Privacy-Preserving Network Monitoring Architecture for the Future Internet.* In *ICT-MobileSummit 2009 Conference Proceedings*, June 2009, ISBN 978-1-905824-12-0.

[49] Rekhter, Y., Li, T., and Hares, S.: *A Border Gateway Protocol 4 (BGP-4).* RFC 4271 (Draft Standard), January 2006. http://www.ietf.org/rfc/rfc4271.txt.

[50] Réseaux IP Européens Network Coordination Centre (RIPE NCC): *RIPE Database.* http://www.ripe.net/db/index.html (accessed 23.07.2010), 1992–2010.

[51] Réseaux IP Européens Network Coordination Centre (RIPE NCC): *Routing Information Service (RIS).* http://www.ripe.net/ris/ (accessed 23.07.2010), 1999–2010.

[52] Shakkottai, S., Srikant, R., Brownlee, N., and Claffy, A. Broido K. C.: *The RTT Distribution of TCP Flows in the Internet and its Impact on TCPbased Flow Control.* Technical report, The Cooperative Association for Internet Data Analysis (CAIDA), 2004. "http://www.caida.org/publications/papers/2004/tr-2004-02/tr-2004-02.pdf".

[53] Shavitt, Y. and Shir, E.: *DIMES: let the internet measure itself.* SIGCOMM Comput. Commun. Rev., 35(5):71–74, 2005, ISSN 0146-4833. http://dx.doi.org/10.1145/1096536.1096546.

[54] Strohmeier, F., Bonelli, N., and Salvador-Gonzales, A.: *AS-Level Network Delay Evaluation with Privacy-preserving Passive Network Monitoring.* Presented at the 21st Tyrrhenian Workshop on Digital Communications: Trustworthy Internet, Ponza, Italy, September 2010.

[55] Team Cymru Research NFP: *IP to ASN Mapping.* `http://www.team-cymru.org/Services/ip-to-asn.html` (accessed 23.07.2010), 2010.

[56] The Apache Software Foundation: *Apache JMeter.* `http://jakarta.apache.org/jmeter/` (accessed 15.07.2010), 1999–2010.

[57] University of Oregon: *Route Views Project.* `http://www.routeviews.org/` (accessed 13.07.2010), 2005–2010.

[58] Veal, B., Li, K., and Lowenthal, D.: *New Methods for Passive Estimation of TCP Round-Trip Times.* In *In Proceedings of the Passive and Active Measurement Workshop (PAM), 2005.*, 2005. `http://pamconf.net/2005/PDF/34310124.pdf`.

[59] Washington, DC: Central Intelligence Agency: *The World Factbook 2009.* `https://www.cia.gov/library/publications/the-world-factbook/index.html` (accessed 14.08.2010), 1981–2010.

[60] Wilkens, A.: *Google stoppt Sammlung von WLAN-Daten (in German).* Heise online News, May 2010. `http://www.heise.de/newsticker/meldung/Google-stoppt-Sammlung-von-WLAN-Daten-1000683.html`.

[61] Zeitoun, A., Chuah, C., Bhattacharyya, S., and Diot, C.: *An AS-level study of Internet path delay characteristics.* In *In Proceedings of IEEE Globecom*, pages 1480–1484, 2004.

List of Abbreviations

AJAX Asynchronous JavaScript and XML

ASN Autonomous System Number

AS Autonomous System

BGP Border Gateway Protocol

CDN Content Delivery Network

CIDR Classless Inter-Domain Routing

CPU Central Processing Unit

DNS Domain Name Service

ECDF Empirical Cumulative Distribution Function

ES End-System

EU European Union

IANA Internet Assigned Numbers Authority

ICMP Internet Control Message Protocol

IETF Internet Engineering Task Force

INRDB Internet Number Resource Database

IP Internet Protocol

IPFIX IP Flow Information Export

IPPM IP Performance Metrics

ISP Internet Service Provider

ITU-T Telecommunication Standardisation Sector of the International Telecommunication Union

IXP Internet Exchange Point

KS-GOF Kolmogorov-Smirnov Goodness of Fit

MOS Mean-Opinion-Score

MSS Maximum Segment Size

NIC Network Information Center

OWD One-Way Delay

PII Personal Identifiable Information

QoE Quality of Experience

QoS Quality of Service

RFC Request for Comments (IETF document)

RIPE NCC RIPE Network Coordination Centre

RIPE Réseaux IP Européens

RIS Routing Information Service

RTO Retransmission Timeout

RTT Round-Trip Time

TCP Transmission Control Protocol

TTL Time To Live

UDP User Datagram Protocol

Appendix

A

Example Results of Services Providing Meta-Information to IP Addresses

In this appendix example results of the mapping services are shown. Queries have been made to the following IP addresses:

- **91.115.90.150**: IP address in the Telekom Austria Customer Network
- **209.85.229.147**: IP address hosting a Google Search Engine Web Server
- **194.232.104.139**: IP address from the Austrian Television On-line Services orf.at

A.1 MaxMind GeoIP Demo Query and Result

Table A.1 shows the information provided by the free on-line demo service of the GeoIP database. The accuracy of the returned data depends on the region where the IP addresses are located.

A.2 RIPE Database Query and Result

In this section, the results of the queries on the RIPE database are presented. Only IP addresses allocated by RIPE NCC produce informative results in the first step. IP address from other ranges require a subsequent query to a different service.

Hostname/IP	91.115.90.150	209.85.229.147	194.232.104.139
Country Code **Country**	AT Austria	US United States	AT Austria
Region Code **Region**	05 Salzburg	CA California	09 Wien
City **Postal Code**	Salzburg -	Mountain View 94043	Vienna -
Latitude **Longitude**	47.8000 13.0333	37.4192 -122.0574	48.2000 16.3667
ISP	Telekom Austria TA AG	Google	APA - Austria Presse Agentur reg. Gen.m.b.H.
Organisation	Highway Customers	Google	Local Area Network of APA
Metro Code **Area Code**	- -	807 650	- -

Table A.1: Reply from MaxMind GeoIP Demo. Results are returned as HTML table [37].

A.2.1 Example Query on a Telekom Austria Highway Customer IP address

The query to the Telekom Austria Customer network ("Highway 194") returns beneath the AS number also the following details.

```
% This is the RIPE Database query service.
% The objects are in RPSL format.
%
% The RIPE Database is subject to Terms and Conditions.
% See http://www.ripe.net/db/support/db-terms-conditions.pdf
% Note: This output has been filtered.
%       To receive output for a database update, use the "-B" flag.
% Information related to '91.115.0.0 - 91.115.255.255'
inetnum:         91.115.0.0 - 91.115.255.255
netname:         TA-HIGHWAY-SPEED
descr:           Highway Customers
descr:           Telekom Austria TA AG
country:         AT
admin-c:         HMH25-RIPE
tech-c:          AAH12-RIPE
tech-c:          DAH12-RIPE
tech-c:          HMH25-RIPE
status:          ASSIGNED PA
remarks:         please contact abuse@aon.at for criminal use, portscan, SPAM, etc.
mnt-by:          AS8447-MNT
mnt-lower:       AS8447-MNT
role:            Host Master Highway
address:         Telekom Austria TA AG
address:         Arsenal Objekt 24
```

```
address:          1030 Vienna
address:          Austria
phone:            + 43 (0)59059 10
fax-no:           + 43 1 7962565
abuse-mailbox:    abuse@aon.at
remarks:          for database maintenance please contact
remarks:          <  hostmaster @ aon.at  >
admin-c:          VM404-RIPE
tech-c:           MA3804-RIPE
tech-c:           AJ2061-RIPE
tech-c:           HH1035-RIPE
tech-c:           RH186-RIPE
nic-hdl:          HMH25-RIPE
mnt-by:           AS8447-MNT
role:             Domain Admin Highway
address:          Telekom Austria TA AG
address:          Arsenal Objekt 24
address:          1030 Wien
address:          Austria
phone:            +43(0)59059 169340
fax-no:           +43(0)59059 169347
abuse-mailbox:    abuse@aon.at
admin-c:          WC82-RIPE
tech-c:           CW6434-RIPE
tech-c:           WC82-RIPE
nic-hdl:          DAH12-RIPE
mnt-by:           AS8447-MNT
role:             Abuse Admin Highway
address:          Telekom Austria TA AG
address:          Postfach 1001
address:          1011 Wien
address:          Austria
phone:            +43 (0)59059 159130
fax-no:           +43 (0)59059 169347
abuse-mailbox:    abuse@aon.at
admin-c:          WC82-RIPE
tech-c:           WC82-RIPE
nic-hdl:          AAH12-RIPE
remarks:          ************************************************
remarks:          * CONTACT FOR CRIMINAL USE, PORTSCAN, SPAM, ETC. *
remarks:          ************************************************
mnt-by:           AS8447-MNT
% Information related to '91.112.0.0/14AS8447'
route:            91.112.0.0/14
descr:            HIGHWAY194
origin:           AS8447
remarks:          ========================================
remarks:          please report abuse incidents (eg network
remarks:          scanning, spam originating, etc.) to
remarks:          abuse@aon.at
remarks:          ========================================
mnt-by:           AS8447-MNT
```

A.2.2 Example Query on a Google IP Address

In this case, the allocation of the IP address has not been done by RIPE NCC. Therefore a generic answer is returned, without any specific information.

```
% This is the RIPE Database query service.
% The objects are in RPSL format.
%
% The RIPE Database is subject to Terms and Conditions.
% See http://www.ripe.net/db/support/db-terms-conditions.pdf

% Note: This output has been filtered.
%       To receive output for a database update, use the "-B" flag.

% Information related to '0.0.0.0 - 255.255.255.255'

inetnum:         0.0.0.0 - 255.255.255.255
netname:         IANA-BLK
descr:           The whole IPv4 address space
country:         EU # Country is really world wide
org:             ORG-IANA1-RIPE
admin-c:         IANA1-RIPE
tech-c:          IANA1-RIPE
status:          ALLOCATED UNSPECIFIED
remarks:         The country is really worldwide.
remarks:         This address space is assigned at various other places in
remarks:         the world and might therefore not be in the RIPE database.
mnt-by:          RIPE-NCC-HM-MNT
mnt-lower:       RIPE-NCC-HM-MNT
mnt-routes:      RIPE-NCC-RPSL-MNT
source:          RIPE # Filtered

organisation:    ORG-IANA1-RIPE
org-name:        Internet Assigned Numbers Authority
org-type:        IANA
address:         see http://www.iana.org
remarks:         The IANA allocates IP addresses and AS number blocks to RIRs
remarks:         see http://www.iana.org/ipaddress/ip-addresses.htm
remarks:         and http://www.iana.org/assignments/as-numbers
e-mail:          bitbucket@ripe.net
admin-c:         IANA1-RIPE
tech-c:          IANA1-RIPE
mnt-ref:         RIPE-NCC-HM-MNT
mnt-by:          RIPE-NCC-HM-MNT
source:          RIPE # Filtered

role:            Internet Assigned Numbers Authority
address:         see http://www.iana.org.
e-mail:          bitbucket@ripe.net
admin-c:         IANA1-RIPE
tech-c:          IANA1-RIPE
nic-hdl:         IANA1-RIPE
remarks:         For more information on IANA services
remarks:         go to IANA web site at http://www.iana.org.
```

```
mnt-by:          RIPE-NCC-MNT
```

A.2.3 Example Query on an ORF.AT IP Address

The web servers of the Austrian national TVs' on-line services are located in an Autonomous System registered by the Austrian press agency (APA).

```
% This is the RIPE Database query service.
% The objects are in RPSL format.
%
% The RIPE Database is subject to Terms and Conditions.
% See http://www.ripe.net/db/support/db-terms-conditions.pdf

% Note: This output has been filtered.
%       To receive output for a database update, use the "-B" flag.

% Information related to '194.232.104.0 - 194.232.104.255'

inetnum:         194.232.104.0 - 194.232.104.255
netname:         APA-LAN
remarks:         INFRA-AW
descr:           Local Area Network of APA
country:         AT
admin-c:         AN6666-RIPE
tech-c:          AN6666-RIPE
status:          ASSIGNED PA
mnt-by:          AS5403-MNT
source:          RIPE # Filtered

role:            APA Network Admin
address:         APA - IT Informations Technologie GmbH
address:         Laimgrubengasse 10
address:         A-1060 Vienna
phone:           +43 (1) 36060 6666
e-mail:          noc@apa.at
remarks:         trouble:      Information: http://www.apa-it.at
remarks:         trouble:      Questions and bug reports ...
                                        mailto:hotline@apa.at
admin-c:         EF1420-RIPE
admin-c:         HT13-RIPE
admin-c:         ME2435-RIPE
tech-c:          EF1420-RIPE
tech-c:          HT13-RIPE
tech-c:          ME2435-RIPE
nic-hdl:         AN6666-RIPE
mnt-by:          AS5403-MNT
source:          RIPE # Filtered

% Information related to '194.232.0.0/16AS5403'
```

```
route:          194.232.0.0/16
descr:          AT-APA-960125
origin:         AS5403
mnt-by:         AS5403-MNT
org:            ORG-AAPA1-RIPE
source:         RIPE # Filtered

organisation:   ORG-AAPA1-RIPE
org-name:       APA - Austria Presse Agentur reg. Gen.m.b.H.
org-type:       LIR
address:        APA - Austria Presse Agentur reg.GmbH
                Laimgrubengasse 10
                A-1060 Vienna
                Austria
phone:          +43 1 36060 6666
fax-no:         +43 1 36060 6699
e-mail:         noc@apa.at
admin-c:        EF1420-RIPE
admin-c:        HPB2-RIPE
admin-c:        HT13-RIPE
admin-c:        JS2437-RIPE
admin-c:        ME2435-RIPE
mnt-ref:        AS5403-MNT
mnt-ref:        RIPE-NCC-HM-MNT
mnt-by:         RIPE-NCC-HM-MNT
source:         RIPE # Filtered
```

A.3 IP/ASN Mapping Team Cymru Query and Result

The result for the example IP addresses from querying the database from Team Cymru is shown below. The information is precise and limited to some fields of major interest.

```
[Querying v4.whois.cymru.com]
[v4.whois.cymru.com]
AS   |IP             |BGP Prefix     |CC|Registry|Allocated |AS Name
8447 |91.115.90.150  |91.112.0.0/14  |AT|ripencc |2006-09-04|TELEKOM-AT Teleko[...]
15169|209.85.229.147 |209.85.228.0/23|US|arin    |2006-01-13|GOOGLE - Google Inc.
5403 |194.232.104.139|194.232.0.0/16 |AT|ripencc |1996-01-25|AS5403 APA-Media-[...]
```

Multiple query interfaces are available, for example a graphical web interface but also machine accessible interfaces for DNS queries or direct `whois`-queries, which allow bulk requests. Results are provided in easily plain text and can be easily parsed. They are based on other data sources (from BGP peers and Regional Internet Registries). This service was best suited for the implementation of ASDF.

Curriculum Vitae

DI(FH) Felix Strohmeier received his master degree from the Department of Telecommunications Engineering at the Salzburg University of Applied Sciences in 1998. His diploma thesis contributed to the development of the interconnection between the Internet and the Siemens PSTN switch. Within 1998 he joined the Advanced Networking Center (ANC) at Salzburg Research as Researcher. In this position the focus of his work is on Internet quality testing and measurements, contributing to several Austrian national and European projects as well as to the development of the software platform MINER (http://miner.salzburgresearch.at). He coordinated the FP6 IST project MOME, a co-ordination action on Internet measurement tools and data. Currently, Felix Strohmeier is technically managing MOMENT (http://www.fp7-moment.eu) and contributing to PRISM (http://www.fp7-prism.eu); both of them projects in the FP7 ICT programme endorsing the Bled declaration "Towards a European approach to the Future Internet". These projects are carried out in co-operation with major industry partners from the telecommunications area, like Telefnica, Hitachi Europe or Ericsson. In this context Felix Strohmeier co-authored several research papers on international conferences, two of them recently published at the ICT Mobile Summit 2009. In parallel to his work at Salzburg Research, he lectured on network protocols and services at the Salzburg University of Applied Sciences from 2005-2008. Since 2008, he is member of the management committee of the COST Action IC0703, an academic forum about "Data Traffic Monitoring and Analysis (TMA): theory, techniques, tools and applications for the future networks".