UNIVERSITEIT VAN AMSTERDAM

Master System and Network Engineering

Analysis of network measurement data

RP1 project report

Damir Musulin, Roy Duisters
{damir.musulin, roy.duisters}@os3.nl

February 7, 2011

Contents

Abstract

RIPE Atlas[1] is a new active network measurement system introduced by the RIPE(Réseaux IP Européens) NCC(Network Coordination Center)[2]. The RIPE Atlas system is currently in a prototype stage and actively in development. This study is done to analyse and gather results from the initial measurement data of RIPE Atlas.

The study made a comparison between the IPv4 and IPv6, based on performance and reliability. The study also detected events in the initial measurement data that may seemed interesting and correlated those to known network events.

To make the comparison between IPv4 and IPv6 as valid as possible, the initial measurement data has been synthesized. To make the measurement data comparable, comparison indexes have been calculated. These comparison indexes have been correlated to other comparison indexes, to generate the results and to ultimately compare IPv4 to IPv6.

The results show that IPv4 performs slightly better (approximately 1%) in terms of reliability (measured using packet loss) than IPv6. The results also show that the performance of IPv4 is higher (approximately 8%) than the performance of IPv6 (measured using network latency).

1 Introduction

The RIPE (Réseaux IP Européens) NCC (Network Coordination Center)[2] has been conducting active measurements from a network of approximately 100 Test Traffic Measurements (TTM) test boxes for many years now[3]. These test boxes perform measurements, including one-way delay and traceroute between the test boxes and major DNS servers. TTM aims to help ISPs with monitoring their networks and to plan transmission capacity.

Recently the focus for measurements is shifting form monitoring of the core of the network (like TTM, in the network of an ISP) to monitoring closer towards the edge of the network.

> *"I believe that there is a need for a more comprehensive approach covering the whole of the RIPE region. We need to answer questions based on topological location: "How is this DNS root server reachable from these Autonomous Systems (ASes)?" We also need to answer questions based on geographical location: "How are services in Germany reachable from Ukrainian cities?" This means we have to have multiple vantage points both in each AS in the RIPE region but also in each significant geographical area."*
> *- Daniel Karrenberg[3]*

RIPE Atlas is a prototype service for a new, large scale Internet measurement network[1]. RIPE Atlas consists of measurement nodes distributed around the Internet. It has the potential to organise thousands or even tens of thousands of nodes around the Internet[4].

The goal of RIPE Atlas is to produce an atlas of different kinds of high-resolution maps of the Internet. These maps will include geographical, topological, real-time and long-term information.

RIPE Atlas provides the possibility to measure connectivity from all measurement nodes. This information will be used as input data for the maps RIPE Atlas is to produce. Due the scale and design of RIPE Atlas, information can be obtained and compared, while this was previously near impossible.

This project focuses on the analysis of the initial data that has been gathered by RIPE Atlas. The RIPE Atlas measurement data will be used to create a comparison between IPv4 and IPv6.

2 RIPE Atlas

2.1 What is RIPE Atlas?

RIPE Atlas[1] is a new active network measurement system introduced by RIPE(Réseaux IP Européens) NCC(Network Coordination Center)[2]. The Atlas system consists of probes[5] that are distributed around the globe. These probes collect data from a number of predifined locations[1]. A Atlas probe is a small computer that collects data from a number of predefined locations. Figure 1 shows a picture of a prototype RIPE Atlas probe.

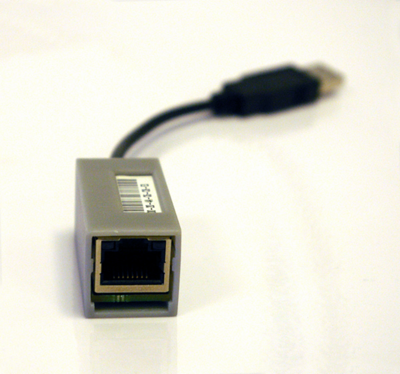

Figure 1: RIPE Atlas prototype probe [6]

A map of RIPE Atlas can be viewed in figure 2. This map a geographic chart of RIPE Atlas, showing how the probes are geographically spread. Figure 2 only shows the probes in the RIPE region. RIPE Atlas also collects data from a number of probes that are geographically outside of the RIPE service region.

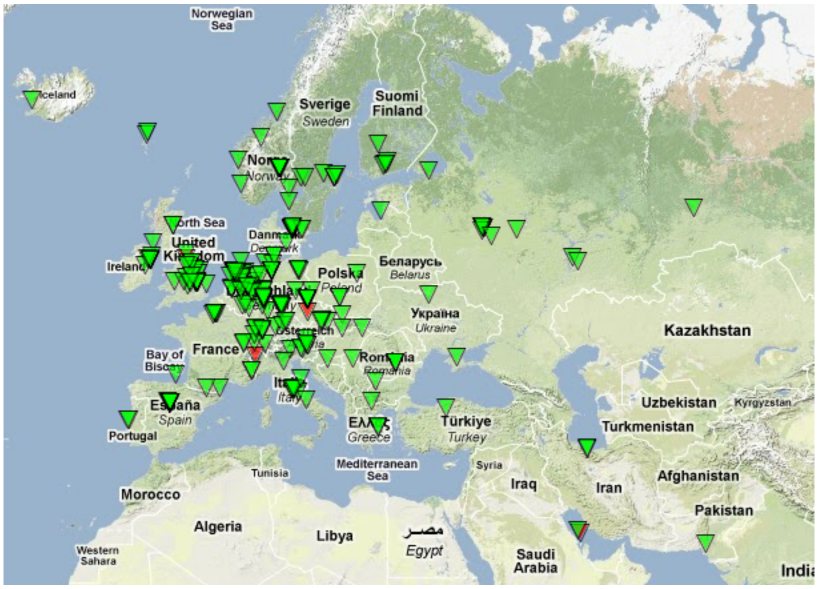

Figure 2: RIPE Atlas map [7]

The RIPE Atlas probes collect the following information:

- The network configuration of the probe itself;
- The uptime information of the probe itself;
- Round trip time measurements to the first and second hop of the probe;
- Round trip time measurements to a number of predetermined destinations.

2.2 The purpose of Atlas

RIPE Atlas is a new, prototype measurement system that can scale to potentially thousands of probes distributed around the world[1]. The purpose of RIPE Atlas is to create maps of the Internet based on the data collected from all the active probes[1].

3 Research methodology

This chapter describes the methodology that has been used to conduct the research. The chapter is organized in the following sub-chapters:

3.1 Data sample	The data sample subsection describes the sample of data that has been used to conduct the study. It shows the probe sample and the measurement sample that have been used as the sample for this study
3.2 Data-collection methods	The data-collection methods subsection describes how the data of the study has been gathered
3.3 Data analysis and synthesis	The data analysis and synthesis subsection shows how the data has been analysed and synthesized
3.4 Data interpretation and correlation	The data interpretation and correlation subsection describes the methods that have been used to make the data comparable and how this data has been correlated to show results
3.5 Caveats	The caveats subsection describes points of concern that might have a influence on the results
3.6 Research limitations	The research limitations subsection describes the limitations that influence the research
3.7 Chapter summary	The chapter summary shows a summary of the sections described in the methodology section

The main research question is defined as the following sentence:

- *How does IPv6 compare to IPv4 based on results that are being collected by RIPE Atlas?*

Rather than a research on how these protocols could perform in a modern day network, the research focuses on the current implementation of both protocols in the biggest network of all: the Internet.

The main research question is further divided into three subquestions:

- *How does the reliability of IPv6 compare to IPv4 based on results that are being collected by RIPE Atlas?*
- *How does the performance of IPv6 compare to IPv4 based on results that are being collected by RIPE Atlas?*
- *When events occur on IPv4 or IPv6, can these events be correlated to known network events or other sources of information?*

3.1 Data sample

This sections discuss the sample data that is used in the study. The subsections shows an example of a probe sample and an example of a measurement sample used in the research.

3.1.1 Probe sample

The research sample consists of measurement data of a total of 289 RIPE Atlas probes. The network measurements are obtained for the time period of 3 January up to 9 January 2011. Table 1 shows the number of probes and the total amount of measurements for the specified set.

Category	Number of probes	Total amount of measurements
All probes	289	˜ 1.5 million

Table 1: Probe sample

3.1.2 Measurement sample

The RIPE Atlas probes actively monitor a set of predefined destination addresses over both IPv4 and IPv6. Table 2 describes which destinations these probes monitor. At the moment of writing the probes are limited to measurements to these specific destination addresses.

In order to compare IPv4 and IPv6, the measurements from both protocols were investigated. As table 2 shows, the RIPE Atlas probes takes both IPv4 and IPv6 measurements from three destinations. These destinations are shown in bold in table 2. The results of these destinations will be analysed.

Destination	IPv4 measurements	IPv6 measurements	Anycast address
k.root-servers.net	Yes	Yes	Yes
193.0.0.193	Yes	No	No
i.root-servers.net	Yes	No	Yes
f.root-servers.net	Yes	No	Yes
m.root-servers.net	Yes	Yes	Yes
l.root-servers.net	Yes	Yes	Yes
labs.ripe.net	Yes	No	No

Table 2: Measurement destinations per protocol

A measurement collected by the RIPE Atlas probes contains the following information:

- Probe identification number;
- Destination identification number;
- Unix time stamp;
- Minimum round trip time;
- Average round trip time;
- Maximum round trip time;
- Loss rate;
- Duplication rate.

3.1.3 Information needed for conducting the research

For the study the RIPE NCC needs to supply the data collected by the RIPE Atlas system. The study compares IPv4 and IPv6 based on performance and reliability, furthermore there will be a study into interesting events in the data from the probes. The research questions can be viewed in section 3 - Research methodology.

The results of the study are divided into:

- A general comparison study;
- A comparison per geographical location;
- A comparison per topological location;
- A search for ”interesting” events.

Data required for conducting the study:

- A dataset from the RIPE NCC with data from the probes, from 3 January up to 9 January 2011;
- A dataset from the RIPE NCC with the geographical locations of the probes;
- A dataset from the RIPE NCC with the IP numbers or Autonomous System numbers of the probes, to determine the topological location of the probes.

3.2 Data-collection methods

The study is dependent on the initial data that has been gathered by RIPE Atlas. Therefore, this initial measurement data needed to be obtained from the RIPE NCC. Only minimal additional information was collected (the AS numbers in which the probes reside). To gather the AS numbers of the probes, a whois client has been used to query the databases of regional Internet registries to determine the AS number (based on the source IP address of the probe).

3.3 Data analysis and synthesis

The dataset from RIPE contained the initial measurements of RIPE Atlas. The study focuses on a comparison between IPv4 and IPv6 (the research questions can be viewed in section 3 - Research methodology). The dataset contains measurements that have been gathered from 3 January up to 9 January 2011. This dataset needs to be analysed and synthesized to gather results from this information.

Since the research primarily focuses on a comparison of IPv4 and IPv6, only probes that have a connection on both protocols are analysed. To remove as many variables as possible for the comparison between IPv4 and IPv6, only ”native” connections are analysed. For example, probes connected to the Internet with a IPv6 tunnel are removed to make the comparison between IPv4 and IPv6 less prone to other variables.

Table 3 shows the amount of probes and measurements in the data set after the synthesis.

Category	Number of probes	Total amount of measurements
All probes	289	˜ 1.5 million
Dual-stack probes	132	˜ 0.68 million
Dual-stack probes without IPv6 tunnels	106	˜ 0.54 million

Table 3: Probe sample

3.3.1 Removal of probes without dual-stack connection

Since the study focuses on a comparison of IPv4 and IPv6, probes without the ability to take measurements on both protocols cannot be used for this comparison. The dataset from the RIPE NCC included data from all probes, including probes with connectivity on only IPv4 or only IPv6. To remove probes that are not dual-stack, a filter is applied to see if probes have values for both IPv4 and IPv6 connections. The probes without a dual-stack connection cannot be used for a comparison of both protocols and therefore removed from this dataset.

3.3.2 Removal of probes without a native IPv6 connection

Because probes with an IPv6 connection may use a tunnel to take measurements, the results would make the study less trustworthy, because the usage of a tunnel introduces another variable for the performance and reliability of IPv6. Therefore, there is a need to remove measurements that have been collected by probes that make use of a tunneled IPv6 connection. To detect IPv6 tunnels there was a need study which methods are available and if the methods are applicable to the dataset.

There are a number of methods that are available to detect IPv6 tunnels[8]. A method to detect tunnels is to use path MTU (Maximum Transmission Unit). Path MTU can be used to detect tunnels. Due to the extra tunnel overhead, less information can be sent in an IP packet and therefore the allowed MTU of connections that make use of a tunnel is generally lower than normal[8].

One other method that can be used for tunnel detection is endpoint spoofing to confirm that a connection is a tunnel. The endpoint spoofing allows of injection of packets into the tunnel to and other destination or yourself and thereby confirming the presence of a tunnel[8].

Another method is to check the IPv6 address of a probe and see if the IPv6 address belongs to a well-known IPv6 subnet that is used for IPv6 tunneling. Because the dataset from RIPE Atlas does not contain options for Path MTU or the detection of tunnels through endpoint spoofing, the only applicable option in this situation is tunnel detection based on the source IPv6 address. The problem with detection of tunnels based on IPv6 address is that only the tunnels will be found that is being actively looked for. Therefore, it cannot be determined how effective this technique is.

3.3.3 Removal of measurements without values

The data points of the probes that are dual-stack need to be checked to see if they contain a value for loss rate and latency. If the measurements do not contain any values for either IPv4 or IPv6, these measurements will be removed from the data set. The RIPE Atlas probes can be unplugged by the users or have lost their connection to the RIPE Atlas controller, making it impossible to store measurements. These events have a real possibility of occurring and have nothing to do with the reliability or performance of IPv4 or IPv6.

Therefore, measurements without data values for both IPv4 and IPv6 will be removed from the data set to make the comparison of IPv4 and IPv6 as valid as possible.

3.3.4 Detecting "interesting" events

The detection of "interesting" events is done by calculating the deviation of a measurement, from the median of a week worth of values of the minimal round trip time. If the measurement value deviates more than 50% from the median minimal round trip time, the value will be marked as "interesting" and researched.

The top 10 interesting events are taken based on the events that affected the most measurements. The median is used because the median has proven to be a more stable metric that is more resistant to spikes in latency than for example the mean. By setting these limits, it is possible to separate instability in the connection from the real "interesting" events, that can possibly be correlated to other probes within the same topological or geographical area.

3.4 Data interpretation and correlation

The "Data interpretation and correlation" section has been divided in two subsections: Comparison indexes and Data correlation. The first subsection describes the first step taken to transform the data into results, the calculation of a comparison index that can be used to compare IPv4 and IPv6. The second subsection describes the second step, whereas the comparison indexes are combined and correlated to each other.

3.4.1 Comparison indexes

To answer the research questions about the comparison of IPv4 and IPv6 there is a need to use a method that makes IPv4 and IPv6 comparable while also making it possible to make a valid comparison to other measurements.

Relative performance index To compare the performance of IPv4 with IPv6, per given measurement the average round trip time of IPv4 is divided by the average round trip time of IPv6. This results in a relative performance index that shows the difference between IPv4 and IPv6. The following formula has been used:

$$\text{Relative performance index} = \frac{averageroundtriptimeIPv6}{averageroundtriptimeIPv4}$$

Relative reliability index For the reliability comparison of IPv4 and IPv6, a method is used that is comparable to the method that has been used to calculate the relative performance. The difference is that the IPv6 loss rate is divided by the IPv4 loss rate. This is done for readability, since if IPv4 would be divided by IPv6, the performance index looks to contradict the the relative reliability index. By calculating the numbers by these formulas, one can easily identify whether IPv4 or IPv6 performs better in performance and/or reliability. If both comparison indexes show a comparison index that is higher than 1, IPv4's reliability and performance is better.The following formula has been used:

$$\text{Relative reliability index} = \frac{LossrateIPv4}{LossrateIPv6}$$

3.4.2 Data correlation

After the calculation of the comparison indexes, there is a need to combine these indexes to generate results. Three steps have been taken in order to correlate the data:

1. Calculation of the initial comparison index;
2. Calculation of a comparison index per destination;
3. Calculation of a comparison index for multiple probes.

Calculation of the initial comparison index The formulas (as listed in section 3.4.1 - Comparison indexes) are used to calculate a comparison index per measurement and per timestamp. Both table 4 and 5 show example calculations using fictional numbers. Table 4 shows a example performance index calculation. The first line shows probe 999, that has an average RTT of 30ms on IPv4 and 35ms on IPv6, from the K-root, on timestamp 1294012800. Using the formula for calculating the relative performance index (see section 3.4.1 - Comparison indexes), 35/30=1.17 is the relative performance index. The relative performance index (1.17) is over 1, indicating that IPv4 has a better performance in this example.

Probeid	Timestamp	Destination	Average RTT IPv6	Average RTT IPv4	Performance index
999	1294012800	K-root	30ms	35ms	1.17
999	1294012800	M-root	90ms	80ms	1.13
999	1294012800	L-root	40ms	33ms	1.21
...					

Table 4: Example performance index calculation

Table 5 shows a example reliability index calculation. Again, the first line shows probe 999, that has an loss rate of 1.0 on IPv4 and 0.9 on IPv6, from the K-root, on timestamp 1294012800. Using the formula for calculating the relative reliability index (see section 3.4.1 - Comparison indexes), 0.9/1=0.9 is the relative reliability index. The relative reliability index (0.90) is below 1, indicating that IPv6 has a better reliability in this example.

Probeid	Timestamp	Destination	Loss rate IPv6	Loss rate IPv4	Reliability index
999	1294012800	K-root	1.0	0.9	0.90
999	1294012800	M-root	0.9	1.0	1.11
999	1294012800	L-root	0.95	0.8	0.84
...					

Table 5: Example reliability index calculation

Calculation of a comparison index per destination The second step in the process is to calculate a comparison index per destination. In the previous step a comparison index was calculated for every timestamp. The second step is to calculate a comparison index per destination, to compare how IPv4 compares to IPv6 per probe. These performance indexes can be combined into a single number, by using a number of different techniques. To show the results as accurate as possible, both the mean and median were calculated. This is done by using the previously calculated comparison indexes and calculating a mean and/or median of these values.

For example, consider the performance indexes of the previous example (as displayed in table 4). If the mean of these three values is calculated, the number 1.17 is the result. When the median is calculated, the same value turns out the be the result. However, over a larger data set the mean and median can show significant differences, as can be seen in the results of this paper.

Calculation of a comparison index for multiple probes The third (and final) step in the process is to calculate a comparison index for a set of probes. In the previous step a comparison index was calculated per probe, per destination. The third step combines the information of the different destinations and/or probes, to calculate performance indexes for a generic comparison, geographic or topological regions.

The comparison indexes that were calculated during the previous step will be used as input for this step. During this step both the mean and median were calculated of the previous of comparison indexes. This is done by using the previously calculated comparison indexes and calculating a mean and/or median of these values.

3.5 Caveats

After the synthesis of the data, there are a number of caveats that may influence the results. In this section these points of concern are brought up.

3.5.1 Caveats of the IPv4 and IPv6 comparison

To detect tunnels there was only one method applicable, detect tunnels based on IPv6 address. This is the only method applicable because the other mentioned methods need information from the probes that the probes do not have to determine if a tunnel is present. The problem that arises from this method is that you will only find the tunnels you are actively looking for, meaning that you will try to match IP address to known IP subnets of tunnels. The weakness in this method is that the filtered data still can harbor probes with IPv6 tunnels but that there is no way to detect them because the IPv6 address of the probe is not know as a well-known IPv6 address of a tunnel. Also, the effectiveness of this tunnel detection technique is unknown. The consequence of tunnels in the dataset means that IPv6 is subject to more variables and therefore the possibility exists that the performance and reliability of IPv6 are negatively impacted.

3.5.2 Caveats when detecting ”interesting” events

Detecting events is done with setting a upper and a lower limit based of the median minimum round trip time of the data set. The median is used because the median is more resistant to sudden increases in latency than for example the mean. With these limits it is possible to separate instability in the connection from events that are potentially interesting to research and correlate to other probes. To select the interesting events, the top 10 measurement points are taken based on measurement point with most deviation to the different destinations. The potential problem with this method is that there is a trade-off between filtering connection instability and detecting interesting events. It could be the case that there are interesting events that would be interesting to research, but do not surpass the defined threshold. It might also be the case that connection instability surpasses the defined threshold and will be marked as an interesting event.

3.6 Research limitations

The Atlas system is in the prototype stage[1]. Therefore, not all desired functionality has been implemented. Limitations that have influenced the study are:

- Intermediate paths;
- Anycast;
- IPv6 addresses of the dual-stack probes.

One of the limitations of the study is the inability to obtain the intermediate path a package takes from a probe to a root server. With this inability there is no way to determine how traffic is routed across the Internet. Therefore it is not possible to determine the possible cause, in case the performance or reliability of IPv4 or IPv6 differs. To solve this limitation traceroute could be used to determine the path a packet takes through the network.

An other limitation is the use of only IPv6 destinations that make use of anycast[9][10]. With the current dataset from the RIPE NCC there is no way to determine to witch instance of a root server a ICMP (Internet Control Message Protocol) packets travel. This study only looks at the IPv4 and IPv6 performance to the three anycasted DNS root servers, therefore it is unclear in how far these results can be extrapolated to the Internet as a whole.

3.7 Chapter summary

This chapter discussed the methodology that was used to conduct the study. The methodology chosen for this study is based on the research questions and the initial RIPE Atlas measurement data. To get to the desired data from the raw data, there was a need to specify what data was needed from the raw data. Because the study focused on a comparison of performance and reliability of native IPv4 and IPv6 connections, certain measurements have been removed from the data set, in order to make the comparison as valid as possible.

4 Reliability

This section describes the results of the reliability comparison between IPv4 and IPv6. It consists of the following sub-chapters:

4.1 Generic reliability	The generic reliability subsection shows an overall reliability comparison
4.2 Geographical region	The geographical region subsection shows a reliability comparison per geographical region
4.3 Topological region	The topological region subsection shows a reliability comparison per topological region (autonomous system)
4.4 Chapter summary	The chapter summary shows a summary of the sections described in the reliability section

To show the difference in reliability between IPv4 and IPv6, a reliability index has been calculated. The way these reliability indexes have been calculated is shown in section 3.4.1 - Comparison indexes. After the calculation of these reliability indexes, these indexes have been correlated and combined. The way these reliability indexes have been correlated and combined is shown in section 3.4.2 - Data correlation.

These reliability indexes shows whether the packet loss of IPv6 is higher or lower than the packet loss of IPv4. If the reliability index shows a value that is higher than 1, this means that IPv4 has a lower packet loss in comparison with IPv6. In case the performance index shows a value that is under 1, IPv6 has a lower packet loss than IPv4.

4.1 Generic reliability

This section describes the generic reliability comparison between IPv4 and IPv6. As described in section 4, the reliability index has been used as a number to describe the difference in reliability between IPv4 and IPv6.

The generic reliability of IPv4 and IPv6 is shown in table 6. As can be seen in the table, the mean performance indexes to the different destinations show a significant difference. For example, consider the mean reliability index of the measurements obtained from the K-root (1.01829) in comparison with the reliability index of the measurements obtained from the M-root (1.00876). Both reliability indexes show that IPv6 shows experiences a higher amount of packet loss than IPv4 and it shows that the destination of the measurements also plays a role in the results.

Destination	Amount of measurements	Mean reliability index
K-root	105 probes - ~107000 measurements	1.01829
M-root	105 probes - ~107000 measurements	1.00876
L-root	104 probes - ~106000 measurements	1.01297
Overall mean	105 probes - ~421000 measurements	1.01334

Table 6: Generic reliability comparison between IPv4 and IPv6

Reliability spread This paragraph describes the difference in reliability between IPv4 and IPv6 to the K-root in more detail. Table 7 shows these results.

Figure 3 shows the performance index values displayed in a graph. Essentially, figure 3 is a graphical presentation of the information of table 7. The horizontal axis shows the mean reliability index, rounded to three decimals. These numbers have been counted and their occurrence is presented on the vertical axis.

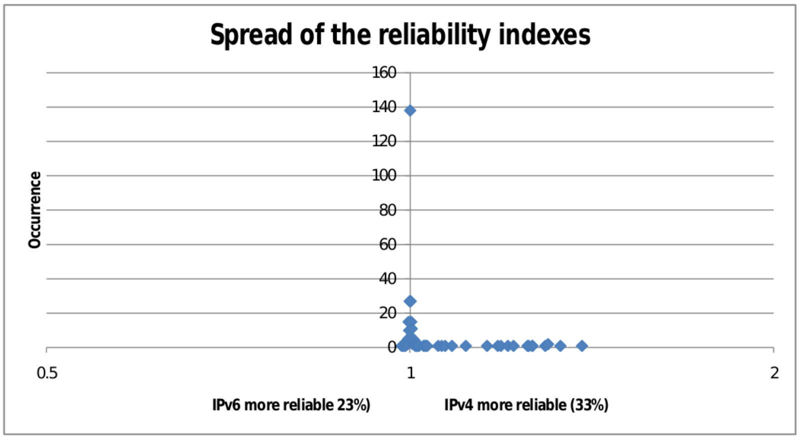

Figure 3: Spread of the mean reliability index

Reliability	Mean reliability percentage of probes
Higher IPv6 reliability	23.57%
Similar[1]reliability on both protocols	43.63%
Higher IPv4 reliability	32.80%
Mean	100%

Table 7: Generic reliability comparison between IPv4 and IPv6, based on the mean reliability index

Table 8 shows the reliability of IPv4 and IPv6, relative to the overall mean reliability index. Table 6 shows a overall mean reliability index of 1.01334 for the measurements that have been collected from all the probes, to all destinations. Table 8 shows how the probes are divided in the data set, relative to the overall mean reliability index of 1.01334.

The information in table 8 shows that about 90% of the probes is categorized in the ”Higher IPv6 reliability” category and about 9% is categorized in the ”Higher IPv4 reliability” category. This information tells that about 90% of the probes has a reliability index that shows better IPv6 reliability results than the overall mean reliability index suggests. Also note that 0% of the probes can be categorized in the ”Similar reliability on both protocols” category. This information shows that a relatively small set of probes (9%) has a significantly lower reliability on IPv6. This small set of probes has a relatively high influence on the mean reliability index.

[1] The sample is considered ”similar” when it deviates less than 0.05% from 1 (the index whereas the reliability of IPv4 and IPv6 is equal)

Reliability (relative to the overall mean reliability index)	Mean reliability percentage of probes
Higher IPv6 reliability	90.45%
Similar[2]reliability on both protocols	0.32%
Higher IPv4 reliability	9.24%
Mean	100%

Table 8: Generic reliability comparison between IPv4 and IPv6, relative to the overall mean reliability index of the week

4.2 Geographical region

This section compares the reliability of IPv4 and IPv6 per country. As described in section 4 Reliability, the reliability index has been used as a number to compare and describe the differences in reliability between IPv4 and IPv6.

Table 9 shows a generic performance comparion between IPv4 and IPv6, grouped per country. As can be seen in the table, the reliability between IPv4 and IPv6 shows clear differences per geographical region. For example, Sweden shows interesting results, indicating that IPv6 experiences slightly less packet loss in comparison with IPv4. The Netherlands also shows interesting results, indicating that the packet loss the IPv6 service experiences is very close to the packet loss the IPv4 service experiences.

Country	Amount of measurements	Mean reliability index
Germany	19 probes - ~58000 measurements	1.02251
The Netherlands	15 probes - ~46000 measurements	1.00007
Spain	7 probes - ~21000 measurements	1.10196
Czech Republic	6 probes - ~18000 measurements	1.00215
Sweden	5 probes - ~15000 measurements	0.99924

Table 9: Generic reliability comparison between IPv4 and IPv6 per country

Reliability to K-root Table 10 shows a reliability comparison between IPv4 and IPv6 of measurements obtained from the K-root server. It is interesting to note that in most countries the reliability of IPv4 and IPv6 is very similar, with the exception of Spain. In Spain the reliability index (1.12340) shows a clear advantage for IPv4, to the K-root.

Country	Amount of measurements	Mean reliability index
Germany	19 probes - ~19000 measurements	1.02643
The Netherlands	15 probes - ~15000 measurements	1.00041
Spain	7 probes - ~7000 measurements	1.12350
Czech Republic	6 probes - ~6000 measurements	1.00233
Sweden	5 probes - ~5000 measurements	1.00014

Table 10: Generic reliability comparison between IPv4 and IPv6 per country to the K-root server

[2] The sample is considered ”similar” when it deviates less than 0.05% from the mean reliability index of the week

Reliability to M-root Table 11 shows a reliability comparison between IPv4 and IPv6 of measurements obtained from the M-root server. This table shows that there are two countries where the reliability of IPv6 shows a slight advantage for IPv6, Sweden and The Netherlands. The rest of the counties show a slight advantage for IPv4, to the M-root destination.

Country	Amount of measurements	Mean reliability index
Germany	19 probes - ~19000 measurements	1.02208
The Netherlands	15 probes - ~15000 measurements	0.99944
Spain	7 probes - ~7000 measurements	1.07976
Czech Republic	6 probes - ~6000 measurements	1.00129
Sweden	5 probes - ~5000 measurements	0.99587

Table 11: Generic reliability comparison between IPv4 and IPv6 per country to the M-root server

Reliability to L-root Table 12 shows a reliability comparison between IPv4 and IPv6 of measurements obtained from the L-root server. The results of the measurements obtained from the L-root server show that there are no countries where the reliability of IPv6 shows an advantage in comparison with IPv4, to the L-root. However, the differences in reliability between IPv4 and IPv6 are very small in these countries, with the exception of Spain, that shows a larger advantage for IPv4 to the L-root.

Country	Amount of measurements	Mean reliability index
Germany	19 probes - ~19000 measurements	1.01894
The Netherlands	15 probes - ~15000 measurements	1.00126
Spain	7 probes - ~7000 measurements	1.10263
Czech Republic	6 probes - ~6000 measurements	1.00283
Sweden	5 probes - ~5000 measurements	1.00172

Table 12: Generic reliability comparison between IPv4 and IPv6 per country to the L-root server

4.3 Topological region

This section compares the reliability of IPv4 and IPv6 per topological region. As described in section 4 Reliability, the reliability index has been used as a number to compare and describe the differences in reliability between IPv4 and IPv6.

Table 13 shows a comparison between IPv4 and IPv6, grouped per topological region (Autonomous System). As the table shows, there are differences in the reliability of IPv4 and IPv6 when these are compared per autonomous systems. In most autonomous systems the differences between IPv4 and IPv6 are quite small (the reliability indexes show a ~1% difference), with the exception of AS-H, where the reliability index (1.14181) shows a clear advantage for IPv4.

AS	Country	Amount of measurements	Mean reliability. index
AS-A	The Netherlands	6 probes - ˜18000 measurements	0.99993
AS-B	Germany	4 probes- ˜12000 measurements	1.00295
AS-C	The Netherlands	3 probes- ˜9000 measurements	1.00075
AS-D	France	3 probes- ˜9000 measurements	1.00003
AS-E	Slovenia	2 probes- ˜6000 measurements	0.99708
AS-F	Iran	2 probes- ˜6000 measurements	1.00035
AS-G	Denmark / Sweden	2 probes- ˜6000 measurements	0.99998
AS-H	Germany	2 probes- ˜6000 measurements	1.14181
AS-I	Czech Republic	2 probes- ˜6000 measurements	1.00022
AS-J	Slovakia	2 probes- ˜6000 measurements	1.00009
AS-K	Portugal	2 probes- ˜6000 measurements	1.00085
AS-L	Spain	2 probes- ˜6000 measurements	1.00232
AS-M	Canada	2 probes- ˜6000 measurements	0.99967
AS-N	Germany	2 probes- ˜6000 measurements	0.99822

Table 13: Generic reliability comparison between IPv4 and IPv6 per autonomous system

Reliability to K-root Table 14 shows the results of measurements to the K-root server, combined per autonomous system that contains two or more RIPE Atlas probes. This table shows that the reliability index of IPv4 and IPv6 to the K-root hardly deviates more than 1%. An exception is AS-H where the reliability index shows a difference of approximately 10% in favor of IPv4.

AS	Country	Amount of measurements	Mean reliability. index
AS-A	The Netherlands	6 probes - ˜6000 measurements	1.00006
AS-B	Germany	4 probes- ˜4000 measurements	1.00696
AS-C	The Netherlands	3 probes- ˜3000 measurements	0.99997
AS-D	France	3 probes- ˜3000 measurements	1.00017
AS-E	Slovenia	2 probes- ˜2000 measurements	0.99844
AS-F	Iran	2 probes- ˜2000 measurements	1.00578
AS-G	Denmark / Sweden	2 probes- ˜2000 measurements	1.00107
AS-H	Germany	2 probes- ˜2000 measurements	1.10465
AS-I	Czech Republic	2 probes- ˜2000 measurements	1.00009
AS-J	Slovakia	2 probes- ˜2000 measurements	1.00000
AS-K	Portugal	2 probes- ˜2000 measurements	1.00093
AS-L	Spain	2 probes- ˜2000 measurements	1.00446
AS-M	Canada	2 probes- ˜2000 measurements	0.99990
AS-N	Germany	2 probes- ˜2000 measurements	1.00000

Table 14: Reliability comparison between IPv4 and IPv6 per autonomous system to the K-root server

Reliability to M-root Table 15 shows the results of measurements to the K-root server, combined per autonomous system that contains two or more RIPE Atlas probes. Again, like the reliability measurements of IPv4 and IPv6 per autonomous system to the K-root server show, the measurements to the M-root server hardly deviate more than 1% from each other. Again, AS-H is an exception, because these measurements show that the reliability index deviates approximately 17% in favor of IPv4, to the M-root server.

AS	Country	Amount of measurements	Mean reliability. index
AS-A	The Netherlands	6 probes - ~6000 measurements	0.99897
AS-B	Germany	4 probes- ~4000 measurements	1.00146
AS-C	The Netherlands	3 probes- ~3000 measurements	1.00087
AS-D	France	3 probes- ~3000 measurements	0.99991
AS-E	Slovenia	2 probes- ~2000 measurements	0.99609
AS-F	Iran	2 probes- ~2000 measurements	0.99278
AS-G	Denmark / Sweden	2 probes- ~2000 measurements	0.99819
AS-H	Germany	2 probes- ~2000 measurements	1.17050
AS-I	Czech Republic	2 probes- ~2000 measurements	0.99996
AS-J	Slovakia	2 probes- ~2000 measurements	1.00005
AS-K	Portugal	2 probes- ~2000 measurements	0.99934
AS-L	Spain	2 probes- ~2000 measurements	1.00022
AS-M	Canada	2 probes- ~2000 measurements	1.00061
AS-N	Germany	2 probes- ~2000 measurements	0.99585

Table 15: Reliability comparison between IPv4 and IPv6 per autonomous system to the K-root server

Reliability to L-root Table 16 shows the results of measurements to the K-root server, combined per autonomous system that contains two or more RIPE Atlas probes. Again, the measurements show that the reliability of IPv4 and IPv6 shows small differences. The reliability measurements hardly deviate more than 1% from each other. Again, the exception is AS-H where the reliability index deviates approximately 15% in favor of IPv4, to the L-root.

AS	Country	Amount of measurements	Mean reliability. index
AS-A	The Netherlands	6 probes - ~6000 measurements	1.00078
AS-B	Germany	4 probes- ~4000 measurements	1.00043
AS-C	The Netherlands	3 probes- ~3000 measurements	1.00142
AS-D	France	3 probes- ~3000 measurements	1.00000
AS-E	Slovenia	2 probes- ~2000 measurements	0.99672
AS-F	Iran	2 probes- ~2000 measurements	1.00249
AS-G	Denmark / Sweden	2 probes- ~2000 measurements	1.00066
AS-H	Germany	2 probes- ~2000 measurements	1.15028
AS-I	Czech Republic	2 probes- ~2000 measurements	1.00061
AS-J	Slovakia	2 probes- ~2000 measurements	1.00021
AS-K	Portugal	2 probes- ~2000 measurements	1.00228
AS-L	Spain	2 probes- ~2000 measurements	1.00228
AS-M	Canada	2 probes- ~2000 measurements	0.99851
AS-N	Germany	2 probes- ~2000 measurements	0.99882

Table 16: Reliability comparison between IPv4 and IPv6 per autonomous system to the K-root server

4.4 Chapter summary

The reliability chapter discussed the results of the study into the reliability comparison of IPv4 and IPv6. The results were gathered and combined using the methods as described in section 3 - Methodology. The reliability chapter is divided in three subchapters that each compare IPv4 and IPv6 in a different way.

The first subchapter (4.1 - Generic reliability) describes the generic reliability and shows an overall view of the reliability of IPv4 in comparison with IPv6. The overall mean reliability index shows that the reliability is approximately 1% higher on IPv4 than on IPv6.

The second subchapter (4.2 - Geographical region) shows a reliability comparison between IPv4 and IPv6 per geographical region. The results clearly show that the reliability of IPv4 and IPv6 differs per geographical region. Of the five geographical regions (in this case, countries) that contain enough probes to make a comparison per geographical region possible (a minimum of 5 probes), one country shows that the IPv6 reliability is higher than the reliability of IPv4. In the remainder of the countries IPv4 shows to have a higher reliability.

The third subchapter (4.3 - Topological region) shows a reliability comparison between IPv4 and IPv6 per topological region. The results clearly show that the reliability of IPv4 and IPv6 differs per topological region. Of the 14 topological regions (autonomous systems) that have enough probes to make a comparison per topological region possible, five autonomous systems show a (marginally) better reliability for IPv6. The remainder of the autonomous systems show a higher reliability for IPv4.

5 Performance

This section describes the results of the performance comparison between IPv4 and IPv6. It consists of the following sub-chapters:

5.1 Generic performance	The generic performance subsection shows an overall performance comparison
5.2 Geographical region	The geographical region subsection shows a performance comparison per geographical region
5.3 Topological region	The topological region subsection shows a performance comparison per topological region (autonomous system)
5.4 Chapter summary	The chapter summary shows a summary of the sections described in the performance section

To show the difference in performance between IPv4 and IPv6, a performance index has been calculated. The way these performance indexes have been calculated is shown in section 3.4.1 - Comparison indexes. After the calculation of these performance indexes, these indexes have been correlated and combined. The way these performance indexes have been correlated and combined is shown in section 3.4.2 - Data correlation.

These performance indexes shows whether the latency of IPv6 is higher or lower than the latency of IPv4. If the performance index shows a value that is higher than 1, this means that IPv4 performs better in comparison with IPv6. In case the performance index shows a value that is under 1, IPv6 performs better than IPv4.

5.1 Generic performance

This section describes the generic performance comparison between IPv4 and IPv6. As described in section 5, the performance index has been used as a number to describe the difference in performance between IPv4 and IPv6.

The generic performance of IPv4 and IPv6 is shown in table 17. As can be seen in the table, the mean and median performance indexes show a significant difference. For example, the mean and median performance indexes of the L-root destination show a difference of 1.64. Since certain outliners in measurement data influence these results, both the mean performance index and the median performance index have been given to provide an accurate view of the results.

Destination	Amount of measurements	Mean perf. index	Median perf. index
K-root	105 probes - ~107000 measurements	1.72	1.08
M-root	105 probes - ~107000 measurements	1.68	1.02
L-root	104 probes - ~106000 measurements	2.89	1.25
Overall mean/median	105 probes - ~421000 measurements	2.10	1.08

Table 17: Generic performance comparison between IPv4 and IPv6

Performance spread This paragraph describes the difference in performance between IPv4 and IPv6 in more detail. Table 18 shows the probes categorized in three categories: higher IPv6 performance, similar performance on both protocols and higher IPv4 performance. As seen in table 17, the results show a significant difference between different techniques for calculating the performance index.

Table 18 shows a performance comparison between IPv4 and IPv6, based on the mean and median performance index. These results show the amount of probes whereas IPv6 performs similar, higher or lower than IPv4 on both the mean and median performance index. Figure 4 shows the performance index values displayed in a graph. Essentially, figure 4 is a graphical presentation of the information that is summarized in table 18. The horizontal axis shows the mean performance index, rounded to one decimal. These numbers have been counted and their occurrence is presented on the vertical axis.

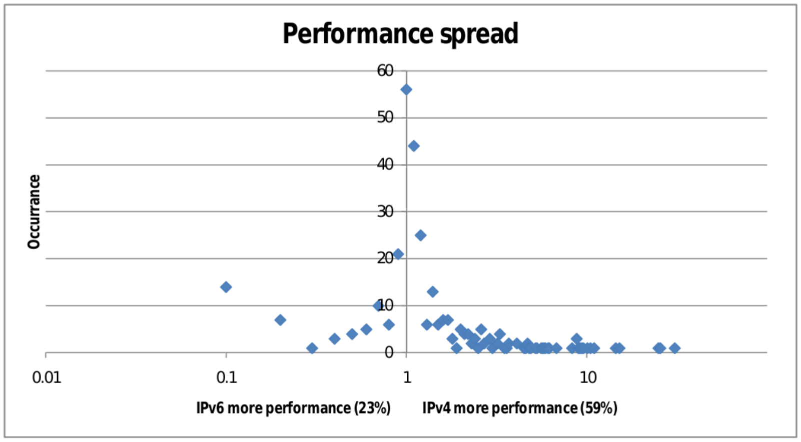

Figure 4: Spread of the mean performance index

When comparing the mean and median performance index in table 18, the results based on the median performance index show that the IPv6 results seem to be consistently higher in comparison with the results based on the mean performance index. Still, even with the median performance index, the majority of the probes show a higher performance over IPv4.

Performance	Mean perf. index percentage of probes	Median perf. index percentage of probes
Higher IPv6 performance	22.92%	25.48%
Similar[3]performance on both protocols	17.83%	19.43%
Higher IPv4 performance	59.24%	55.10%
Total	100%	100%

Table 18: Performance comparison between IPv4 and IPv6, based on the mean and median performance index

Table 19 shows the performance of IPv4 and IPv6, relative to the overall performance index. Table 17 shows a overall mean performance index of 2.10 and a overall median performance index of 1.08 for the measurements that have been collected from all probes, to all destinations. Table 19 shows how the probes are divided in the data set, relative to the overall mean and median performance indexes.

[3] The sample is considered ”similar” when it deviates less than 5% from 1 (the index whereas the performance of IPv4 and IPv6 is equal)

The information in table 19 shows that about 78% of the probes is categorized in the "Higher IPv6 reliability" category based on the mean performance index and about 20% is categorized in the "Higher IPv4 reliability" category. This information tells that about 78% of the probes has a performance index that shows better IPv6 performance results than the overall mean performance index suggests. Also note that 2% of the probes can be categorized in the "Similar reliability on both protocols" category. This information shows that a relatively small set of probes (20%) has a significantly lower performance on IPv6. This small set of probes has a relatively high influence on the mean reliability index.

When comparing these numbers to the results of the median performance index, a different pattern emerges. The probes are spread more evenly among the different categories.

Mean and median performance (relative to the overall mean/median performance index)	Mean perf. index percentage of probes	Median perf. index percentage of probes
Higher IPv6 performance	77.71%	40.76%
Similar[4]performance on both protocols	2.23%	14.97%
Higher IPv4 performance	20.06%	44.27%
Total	100%	100%

Table 19: K-root performance comparison between IPv4 and IPv6, relative to the overall mean/median performance index of the week

5.2 Geographical region

This section compares the performance of both IPv4 and IPv6 per geographical region. As described in section 5 - Performance, the performance index has been used as a metric to describe the difference in latency between IPv4 and IPv6. The probes have been divided into geographical areas based on the country in which they reside. In order to make the countries comparable and keep the amount of measurement data to a reasonable amount, countries that have five or more probes are compared.

Table 20 shows a generic performance comparison between IPv4 and IPv6 per country, to all destinations combined. Again, the mean and median performance index numbers show a significant difference. The performance between IPv4 and IPv6 shows large differences between the different geographical regions. The results of the Czech Republic are particularly interesting, since this is the only country where the median performance index of shows a slight advantage of IPv6 over IPv4.

The following paragraphs elaborate the performance measurements of IPv4 and IPv6 in more detail.

Country	Amount of measurements	Mean perf. index	Median perf. index
Germany	19 probes - ˜58000 measurements	2.59	1.33
The Netherlands	15 probes - ˜46000 measurements	2.93	1.05
Spain	7 probes - ˜21000 measurements	1.66	1.40
Czech Republic	6 probes - ˜18000 measurements	1.46	0.98
Sweden	5 probes - ˜15000 measurements	2.45	1.15

Table 20: Generic performance comparison between IPv4 and IPv6 per country

[4] The sample is considered "similar" when it deviates less than 5% from the mean/median performance index of the week

Performance to K-root Table 21 shows the results of the measurements to the K-root server, measured from probes in a specific country. Again, differences in performance between different geographical regions can be observed. The RIPE Atlas data shows that (according to the median performance index) the IPv6 performance to the K-root server is higher over IPv6 in the Czech Republic country, while the performance of IPv4 is higher in the other countries.

Country	Amount of measurements	Mean perf. index	Median perf. index
Germany	19 probes - ˜19000 measurements	3.05	1.68
The Netherlands	15 probes - ˜15000 measurements	1.77	1.13
Spain	7 probes - ˜7000 measurements	1.43	1.18
Czech Republic	6 probes - ˜6000 measurements	0.82	0.80
Sweden	5 probes - ˜5000 measurements	1.63	1.18

Table 21: Performance comparison between IPv4 and IPv6 per country to the K-root server

Performance to M-root The results from the measurements to the M-root server, divided per country are shown in table 22. Again, large differences in the IPv4 and IPv6 performance between geographical areas can be observed.

Also, while the median IPv6 performance in Sweden to the K-root server shows a slight advantage for IPv4 (a median performance index of 1.18 - table 21), the median performance index towards the M-root server shows that IPv6 has a higher performance. The RIPE Atlas data shows that the performance index from the Czech Republic to the M-root server shows similar results as that are observed to the K-root server (as can be seen in table 21. The median performance index of Sweden shows interesting results as well, since the median performance index of 0.71 shows that the median IPv6 performance is higher to the L-root server from Sweden.

Country	Amount of measurements	Mean perf. index	Median perf. index
Germany	19 probes - ˜19000 measurements	2.51	1.20
The Netherlands	15 probes - ˜15000 measurements	3.12	1.02
Spain	7 probes - ˜7000 measurements	1.89	1.60
Czech Republic	6 probes - ˜6000 measurements	2.08	0.88
Sweden	5 probes - ˜5000 measurements	0.74	0.71

Table 22: Performance comparison between IPv4 and IPv6 per country to the M-root server

Performance to L-root Table 23 shows the results of the measurements to the K-root server, measured from probes in a specific country. Whereas the Czech Republic has shown a higher median performance for IPv6 to the K-root and M-root servers (table 21 and 22), the performance of IPv4 is higher towards the L-root server. Also, it is interesting to note that the median performance index of The Netherlands shows that the performance of IPv6 to the L-root server is higher than the performance of IPv4 to the same root server. Also, it is interesting to note that the performance of IPv4 in Sweden to the L-root server shows a rather large advantage for IPv4, in comparison with IPv6.

Country	Amount of measurements	Mean perf. index	Median perf. index
Germany	19 probes - ˜19000 measurements	1.80	1.14
The Netherlands	15 probes - ˜15000 measurements	3.61	0.97
Spain	7 probes - ˜7000 measurements	1.67	1.55
Czech Republic	6 probes - ˜6000 measurements	1.48	1.51
Sweden	5 probes - ˜5000 measurements	5.00	6.20

Table 23: Performance comparison between IPv4 and IPv6 per country to the L-root server

5.3 Topological region

This section compares the performance of IPv4 and IPv6 per topological region. The probes have been divided into topological regions based on the autonomous system in which they reside. In order to make the autonomous systems comparable and keep the amount of measurement data to a reasonable amount, autonomous systems that have two or more probes are compared.

Table 24 shows the autonomous systems that contain two or more probes. This table shows some interesting results, since there seems to be a big difference in the performance between IPv4 and IPv6 per autonomous systems. For example, the median performance index of AS-N shows a value of 0.51, while the rest of the autonomous systems have median performance index values of between 0.98 and 2.41. AS-N is the only autonomous system in which IPv6 performs significantly better than IPv4.

AS	Country	Amount of measurements	Mean perf. index	Median perf. index
AS-A	The Netherlands	6 probes - ˜18000 measurements	1.04	1.01
AS-B	Germany	4 probes- ˜12000 measurements	2.81	2.41
AS-C	The Netherlands	3 probes- ˜9000 measurements	3.69	1.21
AS-D	France	3 probes- ˜9000 measurements	3.39	1.02
AS-E	Slovenia	2 probes- ˜6000 measurements	0.94	0.98
AS-F	Iran	2 probes- ˜6000 measurements	1.15	1.17
AS-G	Denmark / Sweden	2 probes- ˜6000 measurements	2.08	2.16
AS-H	Germany	2 probes- ˜6000 measurements	6.84	2.10
AS-I	Czech Republic	2 probes- ˜6000 measurements	1.05	0.98
AS-J	Slovakia	2 probes- ˜6000 measurements	1.04	1.01
AS-K	Portugal	2 probes- ˜6000 measurements	0.82	1.02
AS-L	Spain	2 probes- ˜6000 measurements	1.59	1.47
AS-M	Canada	2 probes- ˜6000 measurements	1.14	1.06
AS-N	Germany	2 probes- ˜6000 measurements	0.61	0.51

Table 24: Generic performance comparison between IPv4 and IPv6 per autonomous system

Performance to K-root Table 25 shows the results of measurements to the K-root server, combined per autonomous systems that contains two or more RIPE Atlas probes. Again, interesting results can be observed when comparing autonomous systems. For example, AS-K and AS-L are particularly interesting, since these both show a significantly higher performance over IPv6 when compared to IPv4.

AS	Country	Amount of measurements	Mean perf. index	Median perf. index
AS-A	The Netherlands	6 probes - ~6000 measurements	1.08	1.07
AS-B	Germany	4 probes- ~4000 measurements	2.53	2.25
AS-C	The Netherlands	3 probes- ~3000 measurements	1.36	1.22
AS-D	France	3 probes- ~3000 measurements	0.66	0.68
AS-E	Slovenia	2 probes- ~2000 measurements	1.50	1.62
AS-F	Iran	2 probes- ~2000 measurements	1.10	0.97
AS-G	Denmark / Sweden	2 probes- ~2000 measurements	3.28	3.18
AS-H	Germany	2 probes- ~2000 measurements	5.01	1.33
AS-I	Czech Republic	2 probes- ~2000 measurements	0.96	0.98
AS-J	Slovakia	2 probes- ~2000 measurements	1.02	0.99
AS-K	Portugal	2 probes- ~2000 measurements	0.37	0.36
AS-L	Spain	2 probes- ~2000 measurements	0.72	0.46
AS-M	Canada	2 probes- ~2000 measurements	1.05	1.04
AS-N	Germany	2 probes- ~2000 measurements	1.23	1.24

Table 25: Performance comparison between IPv4 and IPv6 per autonomous system to the K-root server

Performance to M-root Table 26 shows the results of measurements to the M-root server, combined per autonomous systems that contains two or more RIPE Atlas probes. It is interesting to note that there are significant differences between the mean and median performance indexes in some autonomous systems. Consider AS-C where the mean performance index shows a value of 8.85 and the median shows a value of 1.31. AS-N shows some interesting results, since both the mean and median performance index show values of 0.10, indicating a rather large performance advantage for IPv6 in this autonomous system to the M-root.

AS	Country	Amount of measurements	Mean perf. index	Median perf. index
AS-A	The Netherlands	6 probes - ~6000 measurements	1.06	1.01
AS-B	Germany	4 probes- ~4000 measurements	1.88	1.74
AS-C	The Netherlands	3 probes- ~3000 measurements	8.85	1.31
AS-D	France	3 probes- ~3000 measurements	1.02	1.02
AS-E	Slovenia	2 probes- ~2000 measurements	0.57	0.57
AS-F	Iran	2 probes- ~2000 measurements	1.16	1.17
AS-G	Denmark / Sweden	2 probes- ~2000 measurements	0.70	0.69
AS-H	Germany	2 probes- ~2000 measurements	12.08	6.38
AS-I	Czech Republic	2 probes- ~2000 measurements	0.90	0.88
AS-J	Slovakia	2 probes- ~2000 measurements	1.01	1.01
AS-K	Portugal	2 probes- ~2000 measurements	1.03	1.03
AS-L	Spain	2 probes- ~2000 measurements	1.86	1.57
AS-M	Canada	2 probes- ~2000 measurements	1.09	1.08
AS-N	Germany	2 probes- ~2000 measurements	0.10	0.10

Table 26: Performance comparison between IPv4 and IPv6 per autonomous system to the M-root server

Performance to L-root Table 27 shows the results of measurements to the M-root server, combined per autonomous systems that contains two or more RIPE Atlas probes. As the performance comparison to the M-root server showed, AS-N shows a value of 0.49 for both the mean and median performance index, indicating a rather large advantage for IPv6, to the L-root in this autonomous system.

AS	Country	Amount of measurements	Mean perf. index	Median perf. index
AS-A	The Netherlands	6 probes - ~6000 measurements	0.97	0.97
AS-B	Germany	4 probes- ~4000 measurements	4.02	4.55
AS-C	The Netherlands	3 probes- ~3000 measurements	0.86	0.86
AS-D	France	3 probes- ~3000 measurements	8.48	5.84
AS-E	Slovenia	2 probes- ~2000 measurements	0.75	0.74
AS-F	Iran	2 probes- ~2000 measurements	1.20	1.25
AS-G	Denmark / Sweden	2 probes- ~2000 measurements	2.26	2.16
AS-H	Germany	2 probes- ~2000 measurements	3.44	1.85
AS-I	Czech Republic	2 probes- ~2000 measurements	1.29	1.30
AS-J	Slovakia	2 probes- ~2000 measurements	1.09	1.10
AS-K	Portugal	2 probes- ~2000 measurements	1.05	1.06
AS-L	Spain	2 probes- ~2000 measurements	2.20	1.99
AS-M	Canada	2 probes- ~2000 measurements	1.28	1.24
AS-N	Germany	2 probes- ~2000 measurements	0.49	0.49

Table 27: Performance comparison between IPv4 and IPv6 per autonomous system to the L-root server

5.4 Chapter summary

The performance chapter discussed the results of the study into the performance comparison of IPv4 and IPv6. The results were gathered and combined using the methods as described in section 3 - Methodology. The performance chapter is divided in three subchapters that each compare IPv4 and IPv6 in a different way.

The first subchapter (5.1 - Generic performance) describes the generic performance and shows an overall view of the performance of IPv4 in comparison with IPv6. The overall median performance index shows that the median IPv6 latencies are approximately 8% higher than the IPv4 latencies.

The second subchapter (5.2 - Geographical region) shows a performance comparison of IPv4 and IPv6 per geographical region. The results clearly show that the performance of IPv4 and IPv6 differs per geographical region. Of the five geographical regions (in this case, countries) that contain enough probes to make a comparison per geographical region possible (a minimum of 5 probes), one country shows that the IPv6 latency is under the latency of IPv4 (thus IPv6 has a higher performance). In the remainder of the countries IPv4 shows to have a higher performance.

The last subchapter (5.3 - Topological region) shows a performance comparison of IPv4 and IPv6 per topological region. The results show clear differences in the performance of IPv4 and IPv6 per topological region. Of the 14 topological regions (autonomous systems) that have enough probes to make a comparison between topological regions possible, three autonomous systems show a better performance for IPv6. The remainder of the autonomous systems show a higher performance for IPv4.

6 Interesting events

This section describes the findings of interesting events in the Atlas system. This chapter consists of the following sub-chapters:

6.1 General overview of the interesting events	A general overview of the events
6.2 Information about each event	An overview of the information about each event
6.3 Correlating to known network events	Correlating of interesting events with known network events
6.4 Chapter summary	The chapter summary shows a summary of the sections described in the "Interesting events" section.

To detect interesting events, the methodology described in section 3.3.4 is used. The methodology requires that a deviation value is chosen, to check the most interesting events from less interesting events. To detect these interesting events, a deviation of 50% has been selected. All measurement values that deviate more than 50% from the median minimum round trip time of the week are marked as possible interesting events.

For this study, the value of 50% for the deviation has proven to be a good trade-off between the possibility of missing possible interesting events and a large amount of samples caused by occasional connection instability of a single measurement.

6.1 General overview of the interesting events

This section describes the general overview of the 10 most interesting events that are found. To obtain the 10 most interesting events, the top 10 timestamps with the most probes reporting a deviation of more than 50% were selected. Because the study is focusing on a comparison between IPv4 and IPv6, the K, L and M root server are used as the destination server, since the probes collect measurements from these root servers on both IPv4 and IPv6.

Event	Duration	Number of affected measurements	Number of affected probes	Number of affected AS	Number of affected countries
A	30 minutes	76	60	41	16
B	75 minutes	57	35	24	9
C	15 minutes	40	26	18	10
D	15 minutes	36	28	21	10
E	15 minutes	36	28	20	10
F	15 minutes	35	26	20	11
G	15 minutes	35	27	17	10
H	15 minutes	34	26	20	10
I	15 minutes	33	26	19	9
J	15 minutes	33	23	17	10

Table 28: General overview of the interesting events.

6.2 Information about each event

In this section a overview is given per interesting event. In this overview the probes that reported a deviation of over 50% are divided per root server, showing to which destination the event had the highest impact.

Interesting event A Event A is the biggest event seen be the Atlas system. The data in table 29 shows that for event A the highest impact was observed by measurements from the K-root destination. Of the 60 probes that showed this event, 39 probes report a deviation of more than 50% for IPv4 and 20 report a deviation of more than 50% on IPv6. The event was mainly visible on the K-root destination. Event A shows that the event was visible on both IPv4 and IPv6, so the event influenced both protocols.

Destination	IP version	Number of affected probes	Number of affected AS	Number of affected countries
K-root	4	39	25	8
K-root	6	20	16	8
M-root	4	1	1	1
M-root	6	4	3	2
L-root	4	7	6	5
L-root	6	5	5	4

Table 29: Overview of interesting event A per root server.

Interesting event B Event B is an interesting event that was visible for about 75 minutes. The information of event B is shown in table 30. This event shows that the deviation limit to the M-root IPv4 was triggered for 18 probes, but for one probe was influenced by this event on IPv6 to the M-root destination. For the L-root server 16 probes reported a deviation of more than 50% for IPv4 and 3 probes on IPv6. The highest impacted destinations are root servers M and L.

Destination	IP version	Number of affected probes	Number of affected AS	Number of affected countries
K-root	4	6	6	5
K-root	6	1	1	1
M-root	4	18	8	2
M-root	6	1	1	1
L-root	4	16	8	5
L-root	6	3	3	2

Table 30: Overview of interesting event B per root server.

Interesting event C Table 31 shows event C. The numbers for event C show an even spread except for the IPv4 connection to the L-root server. There is a difference, since more probes probes have seen the event on measurements from the L-root destination. The probes that report a deviation to the L-root are spread in nine different autonomous systems.

Destination	IP version	Number of affected probes	Number of affected AS	Number of affected countries
K-root	4	7	6	5
K-root	6	4	4	3
M-root	4	6	6	6
M-root	6	5	4	3
L-root	4	13	9	7
L-root	6	5	5	5

Table 31: Overview of interesting event C per root server.

Interesting event D The information about event D can be viewed in table 32. Like in event C, the probes that have seen the event are evenly spread among the different destinations. It is interesting to note that like event C, the IPv4 connection to the L-root server show a higher number of probes that have noticed the event. Also the K-root server shows a higher number of probes that have noticed the event.

Destination	IP version	Number of affected probes	Number of affected AS	Number of affected countries
K-root	4	13	9	7
K-root	6	1	1	1
M-root	4	4	4	4
M-root	6	1	1	1
L-root	4	15	12	6
L-root	6	2	2	1

Table 32: Overview of interesting event D per root server.

Interesting event E The information about event E can be viewed in table 33. Event E is an event that had a duration of approximately 15 minutes. The event shows that there is a larger amount of IPv4 probes that have seen the event with the K-root and L-root destination, in comparison with the IPv4 measurements from the M-root.

What also is interesting to note is that no probes have noticed the event on IPv6 on the L-root destination, while 14 probes have noticed the event on the L-root destination on IPv4.

Destination	IP version	Number of affected probes	Number of affected AS	Number of affected countries
K-root	4	15	10	7
K-root	6	3	3	3
M-root	4	3	3	2
M-root	6	1	1	1
L-root	4	14	11	5
L-root	6	0	0	0

Table 33: Overview of interesting event E per root server.

Interesting event F The information about event F can be viewed in table 34. Event F shows a similar pattern as event E. The pattern shows an event spread over all destinations on both protocols, except for the K and L-root destination over IPv4.

Destination	IP version	Number of affected probes	Number of affected AS	Number of affected countries
K-root	4	11	8	7
K-root	6	2	2	2
M-root	4	5	5	5
M-root	6	3	3	3
L-root	4	13	10	5
L-root	6	1	1	1

Table 34: Overview of interesting event F per root server.

Interesting event G The information about event G can be viewed in table 35. Event G shows a pattern whereas the event is spread over all destinations, except for the IPv4 connection to the L-root destination. It is also interesting to note that the IPv4 connection to the L-root has more probes that have noticed the event.

Destination	IP version	Number of affected probes	Number of affected AS	Number of affected countries
K-root	4	6	3	3
K-root	6	3	3	2
M-root	4	3	3	3
M-root	6	4	3	3
L-root	4	15	11	9
L-root	6	4	4	4

Table 35: Overview of interesting event G per root server.

Interesting event H For event H, table 36 shows that the probes that have noticed the event are located at the IPv4 connections to the K-root and L-root destination. It is also interesting to note that the M-root server does not show as many probes that have noticed the event over IPv4, in comparison with the probes that have noticed the event over IPv6.

Destination	IP version	Number of affected probes	Number of affected AS	Number of affected countries
K-root	4	11	8	5
K-root	6	2	2	2
M-root	4	5	5	5
M-root	6	1	1	1
L-root	4	14	11	6
L-root	6	1	1	1

Table 36: Overview of interesting event H per root server.

Interesting event I The information about event I can be viewed in table 37. Event I is an event that had a duration of approximately 15 minutes. Event I is showing that more probes that have noticed the event on IPv4 on the K-root and L-root destinations. The M-root destination on IPv4 shows in comparison with the K-root and the L-root on IPv4 no high number of probes that have noticed the event.

Destination	IP version	Number of affected probes	Number of affected AS	Number of affected countries
K-root	4	12	9	7
K-root	6	2	2	2
M-root	4	4	4	2
M-root	6	2	2	2
L-root	4	14	11	5
L-root	6	0	0	0

Table 37: Overview of interesting event I per root server.

Interesting event J The information about event J can be viewed in table 38. Like event I, event J shows a higher number of probes have noticed the event on IPv4 for the K-root and the L-root destination. This event is barely visible on IPv6 for the K-root and L-root destinations.

Destination	IP version	Number of affected probes	Number of affected AS	Number of affected countries
K-root	4	13	9	6
K-root	6	1	1	1
M-root	4	4	4	2
M-root	6	1	1	1
L-root	4	14	11	5
L-root	6	0	0	0

Table 38: Overview of interesting event J per root server.

Summary of the 10 interesting events What is interesting in the top 10 of interesting events, is the fact that most probes notice these events on the K-root and the L-root destinations, with the exception of event B, whereas the there were more probes that had noticed the event on the M-root destination. It is known that the M-root server has fewer anycast instances around the world[11] than the K-root server and the L-root server.

Because the study has limitations (as described in section 3.6 - Research limitations) there is no way to determine the real cause of these differences.

6.3 Correlation to known network events

To correlate the 10 interesting events to known network events, there is a need for external information. The study found two possible sources of information that could be used to correlate these events to known network events. These systems are:

- RIPE NCC DNS Monitoring Services;

- RIPE NCC Routing Information Service.

The RIPE NCC DNS Monitoring Services[12] is a system to monitor the DNS root servers. Another system that was found is the RIPE NCC RIS(Routing Information Service)[13].The RIS server is used to collect data about the BGP routing table of the Internet. With the data from the RIPE RIS server it is possible to see BGP updates. The problem that this study has are the limitations as described in section 3.6 it is very difficult to pinpoint known network events to the interesting events that where extracted from the dataset. Mainly the usage of only IPv6 anycast destinations make it hard to pinpoint events, since the it is unknown to which DNS instance a packet gets routed to.

If there was a possibility to see which physical DNS root server instance a probe gathers its measurement from, it would be possible to correlate these events to the BGP routing table using RIS. For now it is near impossible to study which network event is responsible for triggering an interesting event, since it is unknown from which physical DNS root server instance the probe gets obtains its measurements. When these limitations are resolved in the future, the correlation of interesting events to network events has a better chance of succeeding.

6.4 Chapter summary

The study into interesting events has produced an interesting pattern, generally a higher number of probes notice a event on IPv4 to the K-root and L-root destination than to the M-root destination. Because the study is limited in its ability to see which path the measurements of a probe takes, as described in section 3.6 - Research limitations, it is near impossible to find the root cause of these events in an efficient way.

The correlation of these events to network events is also subject to the same limitations. These limitations make correlating near impossible because there is no data on how the measurements of the probes travel through the Internet to the destinations, therefore it is unknown to which network events a measurement can subject to.

7 Conclusion and recommendations

7.1 Conclusion

The main research question is defined as the following sentence:

- *How does IPv6 compare to IPv4 based on results that are being collected by RIPE Atlas?*

The main research question is further divided into three subquestions. The combination of these three subquestions provide an anwer to the main research question.

- *How does the reliability of IPv6 compare to IPv4 based on results that are being collected by RIPE Atlas?*

The results show that IPv4 performs slightly better (approximately 1%) in terms of reliability (measured using packet loss) than IPv6. The study showed that the destinations the probes take measurements from also play a rather large role in the outcome of the measurements.

- *How does the performance of IPv6 compare to IPv4 based on results that are being collected by RIPE Atlas?*

The results show that the performance of IPv4 is higher (approximately 8%) than the performance of IPv6 (measured using network latency). The study showed that the destinations the probes take measurements from also play a rather large role in the outcome of the measurements.

- *When events occur on IPv4 or IPv6, can these events be correlated to known network events or other sources of information?*

The study into interesting events shows that it is near impossible to correlate events with known network events, because it is not known which anycast instance the probe takes its measurements from. Also, the intermediate paths a measurement of a probe takes to get to a destination are unknown, further complicating the study. Therefore it is not known to which known network events a measurement can be subject to.

7.2 Recommendations

Tunnel detection Detecting IPv6 tunnels is important to make the comparison between IPv4 and IPv6 as valid as possible. In this study tunnel detection was done on basis of filtering known IPv6 tunnel addresses, as described in section 3.3.2 - Removal of probes without a native IPv6 connection. The method used in this research was the only viable method to detect IPv6 tunnels, due to current limitations of RIPE Atlas. For future research in the comparison of IPv4 and IPv6, a good form of tunnel detection is needed to remove IPv6 tunnels. Once it is possible to detect tunnels with a high degree of accuracy the comparison of IPv4 and IPv6 will be more valid.

Unicast destination One limitation the research faced is the anycast limitation as described in section 3.6 - Research limitations. One way to circumvent this problem, is to add a unicast destination that the probes take measurements from over IPv6. Another way is to implement a technique in RIPE Atlas to detect which physical anycast root server answers the query.

Trace route functionality Traceroute functionality would allow future research to detect the intermediate path a measurement has taken. By implementing traceroute functionality, many of the limitations that were part of this study (as described in section 3.6 - Research limitations) will be resolved.

7.3 Future research

Events noticed by probes by the K-root and the L-root destinations In paragraph 6.2 - Information about each event there is an interesting finding. Most of the probes that notice interesting events, notice these on the K-root and the L-root destination over IPv4, while the M-root destination only noticed one event where there where more probes that saw the event on the M-root destination than on the K-root and L-root destination. Because of limitations described in section 3.6 - Research limitations, it is impossible for this study to find the root cause of this pattern. Once the limitations described in section 3.6 - Research limitations are resolved, it would be interesting and possible to find out the root cause of this pattern.

Tunnel detection Once there is a tunnel detection method that can remove IPv6 tunnels, a comparison between IPv4 and IPv6 would become more valid. For future research it would be interesting to see if there are methods that can be applied to the data from Atlas system or to the firmware of the Atlas probes to mark IPv6 tunnels and to research how effective these techniques are.

References

[1] RIPE NCC. RIPE Atlas - Home. http://atlas.ripe.net/, 2010. [Online; accessed 03-Februari-2011].

[2] RIPE Network Coordination Centre. Welcome to RIPE.NET. http://www.ripe.net/, 2011. [Online; accessed 05-January-2011].

[3] Daniel Karrenberg. Active Measurements Need More Vantage Points . http://labs.ripe.net/Members/dfk/active-measurements-need-more-vantage-points, 2010. [Online; accessed 05-January-2011].

[4] Robert Kisteleki. Analysis of network measurement data. http://staff.science.uva.nl/~delaat/sne-2010-2011/, 2010. [Online; accessed 03-January-2011].

[5] RIPE NCC. Atlas Probe. http://labs.ripe.net/Members/dfk/a-small-probe-for-active-measurements, 2010. [Online; accessed 25-January-2011].

[6] RIPE NCC.

[7] RIPE NCC. RIPE Atlas - FAQ. http://atlas.ripe.net/faq, 2010. [Online; accessed 05-January-2011].

[8] Lorenzo Colitti. IPv6 tunnel discovery. http://www.dia.uniroma3.it/~compunet/tunneldiscovery/TunnelDiscovery.pdf, 2003. [Online; accessed 26-January-2011].

[9] W. Milliken C. Partridge, T. Mendez. Host Anycasting Service. http://www.ietf.org/rfc/rfc1546.txt, 1993. [Online; accessed 07-January-2011].

[10] T. Hardie. Distributing Authoritative Name Servers via Shared Unicast Addresses. http://tools.ietf.org/html/rfc3258, 2002. [Online; accessed 07-January-2011].

[11] root servers.org. root-servers.org. http://root-servers.org, 2010. [Online; accessed 26-January-2011].

[12] RIPE NCC. RIPE NCC DNS Monitoring Services. http://dnsmon.ripe.net/dns-servmon/, 2010. [Online; accessed 29-January-2011].

[13] RIPE NCC. RIPE Atlas - Projects - RIS. http://www.ripe.net/ris/, 2010. [Online; accessed 05-January-2011].