Internet Delay Measurements using Test Traffic

Data Disclosure Policy

*Henk Uijterwaal**

RIPE-NCC

June 16, 1998

Document: RIPE-180.ps

Version 1.0

1 Introduction

This document describes the data-disclosure policy for the Test-Traffic Project [1]. This policy describes who can access the data from the project, what one can do with the data and which conditions must be fulfilled before the data can be published outside the RIPE-meetings and working groups.

This policy is based on the ideas described in [1] and has been extensively discussed in a BoF at RIPE-30. Drafts of this document have been circulated amongst the ISP's who participated in the project in May 1998. It is assumed that ISP's who will join the project after May 1998, agree with this policy before deciding to join the project.

Collecting data with the test-boxes means collecting data about ISP's and the performance of their networks. We realize that this is a delicate matter as no ISP wants to see an analysis that puts the performance of his networks in a bad daylight, in particular if the scientific merits of the analysis cannot be proven. On the other hand, the results of the test-traffic project can be a valuable tool for both day-to-day operations as well as long term planning, we certainly do not want to be too restrictive about what can be done with the data.

The basis of our data-disclosure policy is that, at the moment, the test-traffic project should be considered a scientific experiment. We do measurements and collect data that we believe is correct. The analysis focusses on describing the data and finding parameters that describe the overall network performance.

*Email: henk@ripe.net

However, we have not proven this (yet) and until such time, one should not use the data to judge the performance of an ISP.

2 Access to the data

2.1 Participating ISP's

As stated in an earlier document [1], each ISP hosting a test-box will have access to the data collected with the test-box at his site. There will be two ways in which the ISP can access the data:

1. **Using a telnet connection to the test-box.** This method gives the ISP access to the delay and routing information as it is being recorded by the test-box within seconds after the data has been taken. However, as we explained in [1] each test-box can only record *incoming* delay measurements and *outgoing* routing vectors. No information about outgoing delay measurements will be available.

2. **Access to the data-base.** This method will give access to the results of both *incoming* and *outgoing* delay measurements and routing vectors. However, this requires that the data is first collected and processed at a central point, so the data will not be available immediately.

The data that is available to participating ISP's will include IP-numbers of the test-boxes that the box communicated with, their position and other details like that.

Instructions for accessing the data are available from `http://www.ripe.net/test-traffic/Host_testbox/access.html`. As a security feature, data will only be made available to machines with IP-numbers that have been specified by the ISP in advance.

2.2 Others

All others will have access to an anonymous version of the data, that is data where IP-numbers, location of the test-boxes and all other information that can be used to trace where the boxes are located has been removed from the data.

Again, details on how to access the data will be available from `http://www.ripe.net/test-traffic/Host_testbox/access.html`.

2.2.1 At a later stage

It has been suggested that, at a later stage, old data should be available to everybody without restrictions. *If this suggestion is approved,* everybody can access data that has been taken at least N months ago, including information about the location of test-boxes.

The idea behind this proposal is that it gives everybody a chance to analyze the data and test ideas on how to improve the Internet using old data. If N is sufficiently high (at least several months, perhaps even a year), then there will have been so many changes in the networks that information about networks is probably outdated and certainly not confidential anymore.

Note that this idea is only a suggestion and will not be implemented in the near future without prior consultation with the sites hosting the test-boxes.

3 Analysis and publication of the data

The data can be used freely for any analysis that one considers interesting. One is free to discuss the analysis inside the organization that did the analysis or the relevant RIPE working groups.

Before an analysis is presented to the outside world, the analysis will have to be verified. This means that the organization (including the RIPE-NCC) that did the analysis, will have to provide a write-up of the analysis that includes enough detail for anybody to independently re-implement the analysis and verify its conclusions. This write-up will be circulated amongst (a subset of) the ISP's participating in the project for a peer review.

If there are objections to an analysis, it will be discussed with the authors what changes in the analysis will have to be made in order to make it acceptable. If an ISP still disagrees, they can ask that data related to their site is removed from the analysis. However, no single ISP can veto the publication of an analysis by another ISP or the RIPE-NCC.

If data is published, it should include as little references to names of other ISP's, IP-addresses of test-boxes and routers, and the like. Note that it will never be possible to make the data completely anonymously.

Most request for publishing data anonymously comes from the for-profit community. In the non-profit community, there appears to be far less resistance against publications where data can be traced back to specific sites.

It should be noted that the review process takes time. Anybody planning to present data at a conference should keep this in mind.

4 Changes in the policy

As the analysis of the data moves along, we expect to get a better understanding of the correct interpretation of the data. We also plan to cross-check the results and eliminate possible sources of experimental errors. We therefor expect that the data-disclosure policy will have to be changed as the project moves along.

Changes in the data-disclosure policy can be suggested by the participating ISP's or the RIPE-NCC. All changes will be discussed with the sites participating in the project at that point. All parties agree on the change, then this document will be revised. It is our goal that the new policy will be acceptable to everybody participating in the project.

The new data-disclosure policy will only apply to data taken after the change in the policy.

5 Conclusion.

This document described our data disclosure policy. When a site agrees to host a test-box, it is assumed that this site agrees with this policy. This document presumably contains all kinds of legal holes which can be exploited. The idea behind this document, however, is that one shall use the data as one would treat the output of a scientific experiment, not as a means to attack fellow ISP's.

References

[1] H. Uijterwaal, O. Kolkman, "Internet Delay Measurements using Test-Traffic, Design Note", RIPE-158.