The RIPE NCC Network Measurement Data Repository

Daniel Karrenberg

RIPE NCC

Outline

- Who

- What

- Where

- When

- Why

- How

Who

- RIPE NCC

- Neutral "Club" of ISPs

- More than just "Numbers are us"
  - We do Measurements
    - TTM
    - RIS
    - …
  - We publish analyses and raw data
  - …

- **Stable long-term funding base**

What

- Repository for Internet Related Measurement Data

- We will use it for our own raw data sets

- Open for other data sets that are

    - in danger to get lost,

    - in need a long-term stable home from the start,

    - supported by another good excuse,

    - particularly those that are long time series,

    - particularly those keyed on internet numbers.

- Start with O(100TB)

What

- Registration of data users
  - Lightweight for straightforward data sets
  - Custom for data with privacy/commercial concerns
  - Keep loose track of users and their work/results

- Initial Data Sets
  - RIS raw, TTM raw, dnsmon raw
  - Other RIPE NCC raw data sets
    - K-root related packet traces, v4/v6, DNS transfer size, …
  - Aukland / Waikato Trace Sets
  - NLANR ?

Where
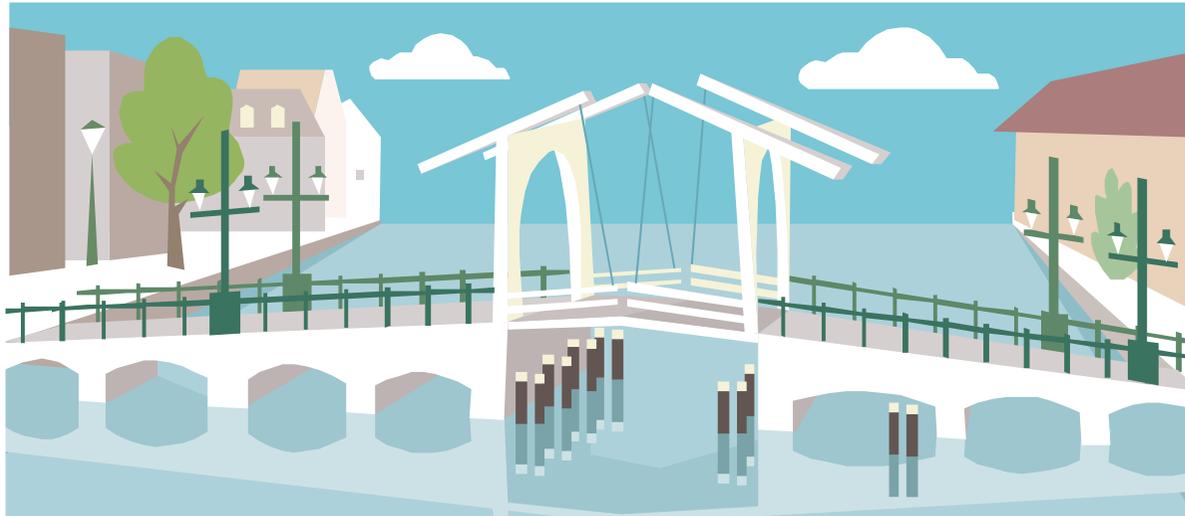
- Amsterdam

- AMS-IX

- May also be useful to replicate heavily accessed data sets from other continents

When

- Just starting up now
- Paper @ PAM

The RIPE NCC Internet Measurement Data Repository

Tony McGregor[1,2], Shane Alcock[1], and Daniel Karrenberg[2]

[1] University of Waikato, Hamilton, New Zealand
[2] RIPE NCC, Amsterdam, The Netherlands
(tonym, salcock)@cs.waikato.ac.nz
daniel.karrenberg@ripe.net

Abstract. This paper describes datasets that will shortly be available to the research community through an Internet measurement data repository operated by the RIPE NCC. The datasets include measurements collected by RIPE NCC projects, packet trace sets recovered from the defunct NLANR website and datasets collected and currently hosted by other research institutions. This work aims to raise awareness of these datasets amongst researchers and to promote discussion about possible changes to the data collection processes to ensure that the measurements are relevant and useful to the community.

1 Introduction

A core requirement of any Internet measurement project is to acquire appropriate measurement data. However, privacy and security concerns often prevent researchers from being able to collect the data themselves. It is very important, therefore, that organisations that collect useful measurement data are able to share it with the research community. Time that would otherwise be spent conducting measurements can instead be dedicated to the analysis of existing data. Shared access to measurement resources also promotes collaboration between researchers and allows studies that validate existing research results to be performed.

Why

- Needed to streamline own raw data publication

- Got request/idea from visiting researcher

- Recognised general utility of the facility

- Having interesting data locally available is nice

- Intensify / establish contacts with academia

- General relevance and appearance of RIPE NCC

Why

- **May very well lead to use/verification by ISPs**
  - Get data closer to operational ground truth

- **(Plug for labs.ripe.net)**

- **May lead to collecting data from ISPs later**
  - Grease publication of interesting data by ISPs
  - Possibly Aggregate & Anonymise

How

- Contact us if you know a data set that needs a home

- Same if you are interested in data we store

- Just an offer, no world domination intended 😉

- Faces available here:
  - Emile Aben
  - Robert Kisteleki
  - Daniel Karrenberg