

Unravelling Internet Infrastructure

detecting problem areas in the Internet infrastructure using
end-to-end measurements and delay tomography

PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus prof.dr.ir. J.T. Fokkema,
voorzitter van het College voor Promoties,
in het openbaar te verdedigen op maandag 4 april 2005 om 13.00 uur

door

Jan-Pascal VAN BEST

doctorandus in de natuurkunde
doctorandus in de informatica
geboren te Utrecht.

Dit proefschrift is goedgekeurd door de promotor:

Prof.dr. W.G. Vree

Samenstelling promotiecommissie:

Rector Magnificus,	voorzitter
Prof.dr. W.G. Vree,	Technische Universiteit Delft, promotor
Prof.dr. R.J. Meijer,	Universiteit van Amsterdam
Prof.dr.ir. H.J. Sips,	Technische Universiteit Delft
Prof.dr.ir. N.H.G. Baken,	Technische Universiteit Delft
Prof.dr.ir. W.A.H. Thissen,	Technische Universiteit Delft
Dr. L. Torenvliet,	Universiteit van Amsterdam
Dr. H.A.J.R. Uijterwaal,	RIPE NCC

This research was financially supported by the Directorate-General of Public Works and Water Management of the Dutch Ministry of Transport, Public Works and Water Management.

Published and distributed by:

Eburon Publishers

P.O. Box 2867

2601 CW DELFT

The Netherlands

e-mail: info@eburon.nl

www.eburon.nl

Author contact: janpascal@vanbest.org

Cover design: Ellen Jagtman

ISBN 90 5972 066 0

© 2005 Jan-Pascal van Best.

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, without the prior permission in writing from the proprietor.

Summary

The word 'Internet' is for most people indistinguishable from the world-wide web. Some people may include e-mail, currently often instant messaging, or maybe peer-to-peer applications or 'downloading music'. All these applications are not what this book is about. This book is about the underlying infrastructure, which enables all named applications. All Internet applications are based on the Internet Protocol (IP), the protocol that describes how pieces of information (*packets*) get transported between Internet-connected computers. One of the reasons the use of the Internet has seen such a large growth is that everyone with a computer that is connected to the Internet can make use of any application as long as it is based on IP. A new Internet application does not require changes to the network: just the computers that are directly involved in the application need to know about it. This means that new applications can be used rapidly, without the need to wait for permission or cooperation from the network administrators. For the Internet, those administrators are Internet Service Providers (ISPs). Each ISP owns a small or larger part of the Internet. Some ISPs are small, and they use the part of the Internet they administer to provide access to the rest of the Internet to a small group of customers, e.g., people in a single city. Other ISPs are much larger. Those ISPs have a world-wide network, with which they connect the networks of smaller ISPs. There are hundreds of thousands of ISPs around the world, and each of them administers its own piece of the Internet infrastructure. ISPs are dependent of each other: a European Internet user browsing the web site of the White House makes use of the networks of a number of ISPs that forward the user's packets to each other, using the IP protocol, from the user to the computer that hosts the web site and back again.

For an ISP, it is relevant to know whether and to what degree the networks of other ISPs are functioning well, e.g., to determine of which other ISPs the networks should be used to forward the packets of its own customers, to get those packets to their destination as soon as possible. Also for governments this is relevant information: the economy increasingly depends on the Internet. Not only do many consumers use web-based shops to do their shopping using the Internet, also companies handle orders via the Internet, and within companies the networks of local offices are often connected via the Internet. Determining the performance of the networks of all ISPs is a tricky business: there are many ISPs, and the performance of an ISP's network is sensitive business information. Privacy concerns also play an important part. To obtain information about the performance of the networks that together form the

Internet, it is necessary to perform measurements to the Internet infrastructure, and analyse the results of those measurements.

This thesis describes a method, or rather a set of methods, that can be used to analyse a certain class of Internet measurements. For these measurements there has been cooperation with RIPE NCC, an independent, not-for-profit membership organisation that supports the infrastructure of the Internet through technical coordination in (mainly) the European part of the Internet. RIPE NCC, in its ‘Test Traffic Measurements’ project, maintains a number of ‘test boxes’, which are distributed around the Internet and perform measurements on the Internet infrastructure. The most important data delivered by these measurements are the paths through the Internet between all pairs of test boxes, and the time an IP packet take to travel from each of the test boxes to each of the others. Because the test boxes are distributed around the Internet, located mainly at major European ISPs, the measurement data reaches a large part of the core of the European Internet. Using the methods described in this thesis more detailed information about the performance of the networks that together form the Internet can be distilled from the raw measurement data.

Chapter 1 starts with an introduction to the subject matter of this thesis, and states the goals of the research described in this thesis:

Goal 1. Contribute to a better understanding of the Internet infrastructure for the benefit of both ISPs and policy makers, using Internet measurements.

Goal 2. Assist the Internet community in detecting slow links in the global infrastructure.

To reach these research goals, this thesis attempts to answer the following research question:

Research Question. What Internet measurements and analysis methods can be used to obtain a better understanding of the Internet infrastructure, to the benefit of both policy makers and ISPs? Specifically, how can slow links in the Internet be detected in a practical way?

The following chapters of this thesis each treat intermediate research questions, the answers to which together form the answer to the main research question.

Chapter 2 provides some basic insight to the Internet infrastructure, topology, protocols, and performance. Special attention is given to the way it is determined for each packet how it can best reach its destination. This chapter also describes the Internet measurements performed by the Test Traffic Measurements project, including its main limitations and ways to overcome those limitations as far as possible.

In Chapter 3 a number of mathematical methods for the analysis of Internet measurements are presented. First a way to represent Internet measurement mathematically is introduced; next a method to automatically draw network graphs is presented. But the major part of this chapter is devoted to a number of methods for *network*

delay tomography: reconstructing the performance of network links inside the Internet infrastructure using only end-to-end measurements, such as the measurements of the Test Traffic Measurements project. There is no measurement equipment attached to each network node of each ISP, but a limited number of test boxes that perform measurements from one to the other like a series of X-rays. In this thesis, the focus is on link delays as performance parameters of links in the Internet, i.e., the time Internet packets take to traverse each link on their path. The problem of determining the delay of each individual network link from the measurement data can be mathematically represented as a system of linear equations. For a typical measurement data set, this linear system consists of thousands of equations in thousands of unknowns, with mathematical properties that make the system unsolvable. This thesis proposes a number of methods to obtain information from this large number of equations. The first is to use a least squares approach; as a variant, a non-negative least squares approach can be used, in which only results larger than or equal to zero are allowed; the third method attempts to find combinations of links for which it is possible to determine the delay of the combination; the fourth and last method tries to determine upper and lower bounds for the delays of individual network links.

In Chapter 4 a software program, written mainly in Java, is presented in which these mathematical methods are implemented. Using this program raw Internet measurement data can be imported and analysed using the mathematical methods.

Chapter 5 shows to what degree the computer program and the methods contained in it are able to determine the delays of individual links in the Internet infrastructure using measurement data from the Test Traffic Measurements project. As had been expected on mathematical grounds, none of the methods is able to deliver exact results for the link delays. The least squares method does deliver results correct within a factor of two. For the detection of slow links this is more than adequate. Finding combinations of individual links for which it is possible to exactly determine the performance proved possible, but the process takes a lot of computing time and delivers only a limited number of combinations. The method that tries to determine upper and lower bounds to the delays of individual links does not give much useful results, since most of the lower bounds found are zero, and most of the upper bounds are so high they do not provide much information.

Chapter 6 discusses the results of this thesis. The most important results are that it is possible, to a certain extent, to determine information about the links delays from only end-to-end measurements, which can be used to detect slow links in the Internet infrastructure. As another result, insight into the cohesion between the ISPs that together form the Internet can be given. Both results are useful for ISPs in their decision making about with which other ISPs to connect their networks, and for governments in adequately assessing the need and the possible effects of intervention in the, both economically and socially, very relevant Internet.

