

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.

Samenvatting

De meeste mensen denken bij het woord 'Internet' aan het world-wide web, aan e-mail, tegenwoordig vaak ook aan instant messaging, en misschien aan peer-to-peer applicaties of 'muziek downloaden'. Daarover gaat dit proefschrift niet. Het gaat over de onderliggende infrastructuur, die ervoor zorgt dat de genoemde toepassingen kunnen functioneren. Alle Internettoepassingen zijn namelijk gebaseerd op het Internet Protocol (IP). Dit protocol regelt het verzenden van brokjes informatie (*packets*) tussen computers die op het Internet zijn aangesloten. Een van de redenen dat het gebruik van het Internet zo'n grote vlucht heeft genomen is dat iedereen met een op het Internet aangesloten computer willekeurige toepassingen kan gebruiken, zolang die maar gebaseerd zijn op het Internet Protocol. Een nieuwe toepassing vereist geen aanpassingen aan het netwerk; alleen de computers waarmee je van die toepassing gebruik wilt maken, hoeven die toepassing te kennen. Dit betekent dat nieuwe toepassingen snel kunnen worden gebruikt, zonder dat gewacht hoeft te worden op toestemming of medewerking van de netwerkbeheerders. Bij het Internet heten die beheerders *Internet Service Providers* (ISP's). Iedere ISP is eigenaar van een deel van het Internet. Sommige ISP's zijn heel klein, en verlenen met het stukje van het Internet dat zij beheren toegang tot de rest van het Internet aan een kleine groep gebruikers, bijvoorbeeld mensen in één stad. Andere ISP's zijn veel groter en beheren een wereldwijd netwerk waarmee zij de netwerken van kleinere ISP's met elkaar verbinden. Wereldwijd zijn er honderdduizenden ISP's, die elk hun eigen stukje van de Internetinfrastructuur beheren. De ISP's zijn onderling van elkaar afhankelijk: een Internetgebruiker die vanuit Europa de website van het Witte Huis bekijkt, maakt daarvoor gebruik van de netwerken van een aantal ISP's die packets (brokjes informatie) via het IP-protocol aan elkaar doorgeven, van de gebruiker naar de computer waar de website op staat en weer terug.

Voor een ISP is het relevant om te weten of en in hoeverre de netwerken van andere ISP's goed functioneren. Bijvoorbeeld om te bepalen via de netwerken van welke andere ISP's de packets van eigen klanten het best kunnen worden doorgestuurd, zodat ze zo snel mogelijk op hun bestemming aankomen. Ook voor overheden is dit relevante informatie: de economie is in steeds sterkere mate afhankelijk van het functioneren van het Internet. Niet alleen wordt er door veel consumenten gebruik gemaakt van webwinkels om via Internet inkopen te doen, maar ook bedrijven onderling handelen bestellingen via het Internet af, en binnen bedrijven worden de informatiesystemen van verschillende vestigingen vaak via Internet aan elkaar gekoppeld. Het bepalen van de mate van functioneren van het netwerk van diverse ISP's is ech-

ter geen eenvoudige zaak: er zijn erg veel ISP's, en bovendien is het functioneren van het netwerk van een ISP bedrijfsgevoelige informatie die ISP's niet zomaar met hun concurrenten willen delen. Ook privacyoverwegingen spelen een belangrijke rol. Om toch informatie te kunnen krijgen over het functioneren van de netwerken die samen het Internet vormen, is het daarom nodig om metingen te verrichten aan de Internetinfrastructuur en de resultaten van die metingen te analyseren.

Dit proefschrift beschrijft een methode (of eigenlijk een aantal methoden) die gebruikt kunnen worden om een bepaald soort Internetmetingen te analyseren. Hiervoor is samengewerkt met RIPE NCC, een onafhankelijke organisatie die tot doel heeft de Internetinfrastructuur te ondersteunen door een forum te bieden voor technische coördinatie in het Europese deel van het Internet. RIPE NCC beheert in het 'Test Traffic Measurements'-project een aantal 'test boxes', die verspreid over het Internet metingen aan de infrastructuur uitvoeren. De belangrijkste gegevens die die metingen opleveren, zijn de routes tussen alle test boxes en de tijd die het een IP-packet kost om door het Internet te reizen van de ene naar de andere test box, voor alle combinaties van test boxes als oorsprong en bestemming. Omdat de test boxes verspreid zijn over het Internet, in het bijzonder bij grote in Europa actieve ISP's, bestrijken de meetgegevens een groot deel van de kern van het Europese Internet. Met behulp van de in dit proefschrift beschreven methoden kan meer gedetailleerde informatie over de mate van functioneren van de netwerken die tezamen het Internet vormen uit de ruwe meetgegevens worden gedestilleerd.

In Hoofdstuk 1 wordt een inleiding gegeven op het proefschrift, en worden de doelen opgesteld van het in dit proefschrift beschreven onderzoek:

Doel 1. *Bijdragen aan een beter begrip van de Internetinfrastructuur ten behoeve van zowel ISP's als beleidsmakers met gebruik van Internetmetingen.*

Doel 2. *Ondersteunen van de Internetgemeenschap in het opsporen van langzame verbindingen in de wereldwijde infrastructuur.*

Om deze onderzoeksdoelen te kunnen bereiken wordt in dit proefschrift een antwoord gezocht op de volgende onderzoeksvraag:

Onderzoeksvraag. *Welke Internetmetingen en -analysemethoden kunnen gebruikt worden om een beter begrip van de Internetinfrastructuur te verkrijgen, ten behoeve van zowel beleidsmakers als ISP's? In het bijzonder, hoe kunnen langzame verbindingen in het Internet in een uitvoerbare wijze opgespoord worden?*

De volgende hoofdstukken van dit proefschrift behandelen telkens een deelvraag van deze onderzoeksvraag.

In Hoofdstuk 2 worden relevante begrippen uit de Internetinfrastructuur, -topologie, -protocollen en -performance geïntroduceerd en uitgelegd. In het bijzonder wordt ingegaan op de manier waarop in het Internet wordt bepaald volgens welke route IP-packets hun doel het beste kunnen bereiken. Verder worden de Internetmetingen van het Test Traffic Measurements-project verder toegelicht, onder andere met

betrekking tot de beperkingen die aan verschillende vormen van Internetmetingen kleven.

In Hoofdstuk 3 worden een aantal wiskundige methoden gepresenteerd voor de analyse van Internetmeetgegevens. Ten eerste wordt een manier om Internetmeetgegevens wiskundig weer te geven geïntroduceerd; vervolgens wordt een methode om netwerkplaatjes automatisch te tekenen gepresenteerd. Het grootste deel van dit hoofdstuk is echter gewijd aan een aantal methoden voor *network delay tomography*: het achterhalen van informatie over de snelheid van verbindingen binnen in het Internet door alleen gebruik te maken van zogenaamde *end-to-end*-metingen, zoals de metingen van het Test Traffic Measurements-project: er wordt geen meetapparatuur aan elk netwerkknooppunt van elke ISP gehangen, maar een beperkt aantal test boxes licht met onderlinge metingen als een serie röntgenfoto's het Internet door. In dit proefschrift wordt als indicator voor de snelheid van elke verbinding de *link delay* gebruikt, de tijd die het IP-packets kost om een bepaalde verbinding te passeren. Het probleem om uit de meetresultaten te achterhalen wat de link delay van elke individuele verbinding is, valt wiskundig gezien te omschrijven als een stelsel lineaire vergelijkingen. Voor een typische set meetresultaten gaat het dan om duizenden vergelijkingen met duizenden onbekenden, die bovendien niet zomaar oplosbaar zijn. In dit proefschrift worden een aantal methoden voorgesteld om toch informatie uit dit grote aantal vergelijkingen te halen. De eerste is om een kleinste-kwadratenbenadering te doen; als variant daarop kan een kleinste-kwadratenbenadering worden gedaan waarbij alleen resultaten groter dan of gelijk aan nul worden toegestaan; de derde methode zoekt naar combinaties van individuele verbindingen waarbij de totale link delay van de combinatie wél exact te achterhalen is; en de vierde en laatste methode probeert boven- en ondergrenzen voor de link delays van de individuele verbindingen te achterhalen.

In Hoofdstuk 4 wordt een voornamelijk in Java geschreven computerprogramma gepresenteerd waarin deze wiskundige methoden zijn geïmplementeerd. Met behulp van dit programma kunnen ruwe Internetmeetgegevens worden ingelezen en aan de hand van de wiskundige methoden worden geanalyseerd.

Hoofdstuk 5 laat zien in hoeverre het computerprogramma en de daarin verwerkte methoden in staat stellen om de link delays van individuele netwerkverbindingen te bepalen aan de hand van ruwe meetgegevens uit het Test Traffic Measurements-project. Zoals op wiskundige gronden was verwacht, is geen van de methoden in staat om exacte resultaten te leveren. Wel levert het gebruik van de kleinste-kwadratenmethode resultaten op die binnen een factor twee correct zijn. Voor het opsporen van langzame verbindingen is dit meer dan voldoende. Het vinden van combinaties van individuele verbindingen waarvoor de totale link delay wel exact vast te stellen is, werkt, maar kost zeer veel rekentijd en levert bovendien maar een beperkt aantal van die combinaties op. De methode die boven- en ondergrenzen probeert te achterhalen levert weinig nuttige resultaten op, doordat de gevonden ondergrenzen vaak nul zijn en de bovengrenzen zo hoog dat ze weinig informatie opleveren.

Hoofdstuk 6 bespreekt de resultaten van dit proefschrift. De belangrijkste zijn dat het tot op zekere hoogte mogelijk is om uit enkel end-to-end metingen informa-

tie over de link delays van individuele Internetverbindingen te destilleren, waardoor langzame verbindingen kunnen worden opgespoord. Bovendien kan er inzicht gegeven worden in de samenhang tussen de ISP's die samen het Internet vormen. Beide vormen van informatie zijn van nut voor ISP's bij het selecteren van partners om hun netwerken aan te koppelen, en voor overheden bij het goed inschatten van de noodzaak en de mogelijke effecten van ingrijpen in het economisch en maatschappelijk uiterst relevante Internet.