
Scalable Performance Signalling and Congestion Avoidance

SCALABLE PERFORMANCE SIGNALLING AND CONGESTION AVOIDANCE

MICHAEL WELZL
University of Innsbruck



Springer Science+Business Media, LLC

Library of Congress Cataloging-in-Publication Data

Scalable Performance Signalling and Congestion Avoidance

Michael Welzl

ISBN 978-1-4613-5117-7 ISBN 978-1-4615-0519-8 (eBook)

DOI 10.1007/978-1-4615-0519-8

Dissertation der Technischen Universität Darmstadt

Hochschulkennziffer D17

Copyright © 2003 by Springer Science+Business Media New York

Originally published by Kluwer Academic Publishers in 2003

Softcover reprint of the hardcover 1st edition 2003

All rights reserved. No part of this work may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, microfilming, recording, or otherwise, without prior written permission from the Publisher, with the exception of any material supplied specifically for the purpose of being entered and executed on a computer system, for exclusive use by the purchaser of the work.

Permission for books published in Europe: permissions@wkap.nl

Permissions for books published in the United States of America: permissions@wkap.com

Printed on acid-free paper.

Contents

| | |
|--|-----|
| Foreword by Jon Crowcroft | vii |
| Foreword by Max Mühlhäuser | ix |
| Acknowledgments | xi |
| 1. INTRODUCTION | 1 |
| 1.1 Reader's Guide | 2 |
| 1.2 Work done | 3 |
| 2. RELATED WORK | 7 |
| 2.1 Congestion Control in Packet Networks | 7 |
| 2.2 Problems and Issues | 11 |
| 2.3 A Trend towards Explicit Signalling | 20 |
| 2.4 Congestion Control in ATM: ABR | 24 |
| 2.5 Quality of Service | 26 |
| 2.6 QoS and Congestion Control | 32 |
| 2.7 TCP-friendly QoS: A Comparative Evaluation | 35 |
| 2.8 Explicit Traffic Signalling in Packet Networks | 38 |
| 3. PTP: A FRAMEWORK FOR EXPLICIT PERFORMANCE SIGNALLING | 45 |
| 3.1 End-to-end Arguments revisited | 45 |
| 3.2 Architecture | 46 |
| 3.3 Performance Parameters | 49 |
| 3.4 Policy Considerations | 52 |
| 3.5 The Performance Transparency Protocol | 57 |
| 3.6 Content Type Definitions | 68 |
| 3.7 Usage Scenarios | 75 |

| | | |
|-----|---|-----|
| 3.8 | Realisation | 84 |
| 4. | CONGESTION AVOIDANCE WITH DISTRIBUTED PROPORTIONAL CONTROL | 91 |
| 4.1 | Architecture | 91 |
| 4.2 | Design Phase 1: Fluid Model | 93 |
| 4.3 | Analysis | 98 |
| 4.4 | Design Phase 2: Discrete Model | 103 |
| 4.5 | Evaluation 1: Dynamic Behaviour | 110 |
| 4.6 | Evaluation 2: Long-term Performance | 118 |
| 4.7 | Final Comments | 137 |
| 5. | CONCLUSION | 141 |
| 5.1 | Future Work | 142 |
| | Appendices | 147 |
| A | The ns Network Simulator | 147 |
| A.1 | Overview | 147 |
| A.2 | Using <i>ns</i> | 148 |
| | References | 151 |
| | About the Author | 163 |
| | Index | 165 |

Foreword by Jon Crowcroft

This monograph is about research in the evolution of two communications trends, that of signaling, and that of end-system adaption to online feedback of network circumstances. The context is the Internet and its protocols and systems. Chapter two outlines the background to the Internet approaches to congestion control versus quality of service, and the range of options present in other network architectures that might be adopted here.

Signaling is generally concerned with everything to do with the control plane in networks; in other words, everything except the business of actually moving the users' data to and fro (the data plane). The Internet Protocol does not inherently possess a control plane. One could argue that the Internet Protocol consists of a per-packet piece of signaling apparatus (the packet header with a variety of fields that do control routing and scope), but that would be a rather extreme view. Instead, Internet signaling to date has been based in implicit behaviour or "side effects", such as routing exchanges, ICMP error messages or end-to-end protocol timeout events or other mechanisms to infer network behaviour.

Some years ago, a fully traditional signaling protocol called RSVP was devised by the IETF to try to offer a way to request explicit network performance in much the same way that Q.931 is used to signal requests for ISDN channels, or Q.2931 was defined for B-ISDN ATM networks. Dr Welzl has chosen to research a middle ground, by devising a new system that is in band (as RSVP was) but is also explicit, namely the Performance Transparency Protocol, PTP. In chapter three, he details the design of this system and evaluates its use by qualitative arguments from usage scenarios.

Currently, End Systems ("Hosts") in the Internet make use of the Transmission Control Protocol for the vast majority of their applications. This runs on top of the Internet Protocol, and is generally therefore only able to use those services and signaling functions which are exposed by the API for IP. With the extension of the IP service to include explicit signaling, more explicit adapta-

tion is possible in end systems. Thus in chapter four, Congestion Avoidance with Distributed Proportional Control is investigated, which allows far more rapid and more accurate adaption to network behaviour by end users, through the use of explicit quantitative feedback from within the network. The approach is thoroughly evaluated through a variety of modeling and simulation, and measurement experiments. The monograph finishes with some discussion of the future for these two topics.

This is an important area for network engineers and researchers to understand. The Internet is still evolving, and the range and heterogeneity of links and nodes is ever expanding, with the introduction of optical core networks, and wireless access in the Local and Wide Area, and efficient utilisation, combined with good response is something all users and providers require as more enterprises in the domestic and commercial world come to depend on this world wide communications infrastructure.

This dissertation makes a significant contribution to that evolution.

Jon Crowcroft
University of Cambridge
Cambridge, UK, 11 June 2003

Foreword by Max Mühlhäuser

The Internet has mastered the growth from an initial four to almost one billion interconnected computers — an unparalleled success and a victory over any other network architecture. Its protagonists believe in the key success factor being the simplicity and rather state-less nature of the core Internet Protocol IP. IP is connection-less and lacks rather common features of network layer protocols such as congestion control and signaling — or at least, these features exist only in a very rudimentary form. As a consequence, the “users” of IP i.e. transport layer protocols such as TCP or even the application programs are supposed to take care of such additional functionality. This concept finds its limits when desired behavior of the network depends on collaborative observation of certain rules and procedures by virtually all users, and when routers — which usually do not use transport layer protocols — must be included in such procedures.

Particularly the trend towards using the Internet for synchronous audio and video streaming has exposed painful limits of the above-mentioned concepts. Decent live streaming support in the Internet requires a number of Quality-of-Service guarantees to be fulfilled and is a complex issue, but in an oversimplified view it boils down to avoiding delay of data packets in congested queues on the way from source to sink. Years of research have concentrated on reserving resources for prioritized streaming data in the net. This research was challenged by a battle of Internet protagonists seeking to preserve their success factor, the simplicity and (rather) stateless nature of the Internet. Per-Flow states, the ideal key concept for streaming support, are today considered intolerable in the general Internet.

Michael Welzl has investigated the area described and took a different approach. He researched ways of moving Internet congestion control further in the spectrum from congestion removal towards congestion avoidance. Any Internet user today can experience the fact that congestion control in the Internet often happens at a time when congestions have already occurred. Michael

Welzl's dream was and his vision is an Internet where applications are told early enough to reduce traffic, such that congestions do not occur and the net remains close to its optimal throughput.

Given the above outline, three obvious questions remain: i) what is the minimal extension to the Internet network layer and most "stateless-like" approach for realizing the vision? ii) how can end systems (transport layer protocols or applications) benefit from this extension and which rules do they have to observe? iii) how can the vision — now a concrete approach — be deployed in a network with a billion nodes? The present book gives answers to these questions: the first one is addressed with the Performance Transparency Protocol (PTP), a proposed network layer extension, cf. chapter three; the second one is answered in chapter four, around the introduction of Congestion Avoidance with Distributed Proportional Control (CADPC); as to the last question, sections 3.4 and 3.7 provide valuable insight. One last point is worth mentioning in this preface: the present book backs its claims and findings with an extremely thorough and extensive practical effort. Michael Welzl has implemented a PTP testbed and an elaborate simulation system; he has conducted extensive simulation experiments and proven the value of his approach in a very expressive way. All in all, the present book deserves to become a historic step towards readying the Internet for its novel applications — applications which it has already started to host, but which it ought to host much better in an attempt to continue living up to its fame.

Max Mühlhäuser
Technical University of Darmstadt
Darmstadt, Germany, 13 June 2003

Acknowledgments

I could not have written this book without the help of the following people:

Ph. D. thesis Supervisor: Max Mühlhäuser

Second Ph. D. thesis supervisor: Jon Crowcroft

Help with PTP: Alfred Cihal

Help with CADPC: Andrew Barnhart (who sent me a hard copy of his ATM forum contribution), Ralf Hauber (who inspired me to write a Chiu/Jain diagram simulator), Thomas Mayr, Bernhard Quatember and Martin Zwicknagl

PTP Linux implementation: Markus Lindorfer (router),
Bernhard Prischl (end node)

AVCS implementation: Gerhard Stummer

English enhancement: Simon Bailey

Encouragement and input: Arjan Durresi, Raj Jain, Byung Kim, Timothy Mangan, Jörg Widmer and many others who replied to my private mails and to my queries in the end2end interest mailing list

Inspiration: The guy who cut my floor tiles beside me while I designed the CADPC mechanism and all New Orleans musicians (and, in particular, drummers) who played on albums that I listened to while working on this thesis

I am especially grateful to all my friends and colleagues (most notably Petra Ratzenböck) who had to endure listening to an immense amount of boring details about PTP, CADPC, QoS and Congestion Control in general. Last but not least, I would like to thank Alex Greene for his support during the final stage of publication.