Fairness in Computer Networks

Inauguraldissertation
zur Erlangung des akademischen Grades
eines Doktors der Naturwissenschaften
der Universität Mannheim

vorgelegt von
Diplom Wirtschaftsinformatiker Robert R. Denda
aus Koblenz

Mannheim, 2003

Dekan:	Professor Dr. Jürgen Potthoff, Universität Mannheim
Referent:	Professor Dr. Wolfgang Effelsberg, Universität Mannheim
Korreferent:	Professor Dr. Stefan Fischer, Technische Universität Braunschweig

Tag der mündlichen Prüfung: 15. Januar 2004

Berichte aus der Telematik

Robert Denda

Fairness in Computer Networks

D 180 (Diss. Universität Mannheim)

Shaker Verlag
Aachen 2004

Bibliographic information published by Die Deutsche Bibliothek
Die Deutsche Bibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data is available in the internet at http://dnb.ddb.de.

Zugl.: Mannheim, Univ., Diss., 2004

Printed in Germany.

ISBN 3-8322-2581-1
ISSN 0948-700X

Shaker Verlag GmbH • P.O. BOX 101818 • D-52018 Aachen
Phone: 0049/2407/9596-0 • Telefax: 0049/2407/9596-9
Internet: www.shaker.de • eMail: info@shaker.de

Kurzfassung

Dienstgüte (Quality-of-Service) war im Rechnernetzwerkbereich eines der erfolgreichsten Forschungsgebiete der letzten zehn Jahre. Basierend auf den gewonnenen Forschungsergebnissen bedarf es nun hauptsächlich eines kontinuierlichen Standardisierungsprozesses und eines weitergehenden kommerziellen Einsatzes, um der sich entwickelnden Nachfrage nach Dienstgüte moderner Netzwerkanwendungen nachzukommen.

Das zur Dienstgüte komplementäre Gebiet *Fairness* wurde bisher jedoch nicht ausreichend tief betrachtet, wozu die vorliegende Dissertation einen wichtigen Forschungsbeitrag liefert.

Wir stellen die Verbindung zwischen den theoretischen Ergebnissen aus den zugehörigen Bereichen der Volkswirtschaftstheorie und den entsprechenden Anforderungen von Rechnernetzen her. Fairnessdefinitionen werden aufgezeigt und ihr Nutzen für den Netzwerkbereich evaluiert. Desweiteren beinhaltet die vorliegende Dissertation eine erste komplette Klassifizierung bestehender Arbeiten über Fairness in Rechnernetzen. Dafür haben wir Fairnessaspekte in die beiden Bereiche *Makro-Fairness* und *Mikro-Fairness* aufgeteilt.

Makro-Fairness bezieht sich auf eine allgemein faire Verteilung der Netzwerkressourcen, von der alle Nutzer und Anwendungen profitieren, wohingegen *Mikro-Fairness* sich auf Fairness im engeren Sinne für spezifische Anwendungen in Szenarien mit ausgeprägtem Wettbewerb, wie z.B. Echtzeithandelssysteme und Online-Auktionen, bezieht.

Im Bereich der Makro-Fairness liefert die vorliegende Dissertation zwei wichtige Forschungsbeiträge: einen neuartigen Architekturvorschlag, *iSSD*, der nutzerbasierte Fairness mit ähnlich niedriger Komplexität wie DiffServ erreicht, und ein neues TCP-freundliches Congestion Control Protokoll, *MPLBcc*. Die Fairnesseigenschaften der iSSD Architektur werden im Rahmen dieser Dissertation sowohl durch eine formale Analyse als auch durch Netzwerksimulationen bestätigt. Der zweite Forschungsbeitrag für den Bereich Makro-Fairness, MPLBcc, liefert für den speziellen Fall mehrerer gleichzeitiger Übertragungspfade bessere Ergebnisse als bestehende TCP-freundliche Algorithmen. Sowohl Simulationsergebnisse als auch Ergebnisse des Feldeinsatzes einer implementierten Variante von MPLBcc, die im Rahmen eines größeren Projektes für die Übertragung von digitalem Video über den Kabelnetzrückkanal eingesetzt wurde, werden aufgezeigt.

Bezüglich Mikro-Fairness, werden von uns entwickelte und getestete neue Algorithmen vorgestellt, die eine faire Übertragung von Netzwerkströmen des Typs *group-delay constrained traffic* ermöglichen, was besonders für wettbewerbsorientierte Anwendungen, wie z.B. Aktienhandelsanwendungen und Netzwerkspiele, hilfreich ist. Unter anderem wurde dazu ein neuer Algorithmus basierend auf *Circulating Envelopes* zur fälschungssicheren Bestimmung der Round-Trip-Time entwickelt.

Desweiteren wurde im Rahmen dieser Dissertation eine neuartige Architekturkomponente für Mikro-Fairness konzipiert und implementiert, die zweierlei Ziele erfüllt: zum einen erlaubt sie, Fairness für bestehende Netzwerkanwendungen ohne deren Veränderung zu erreichen und, zum anderen, dient sie als eine Forschungsplattform, mit der gezeigt wird, dass in speziellen Fällen das Fairnessproblem durch eine Transformation in ein Problem des optimalen Handels auf einem virtuellen Markt gelöst werden kann. Diese Architekturkomponente, *Fairness Proxy*, ist unser Forschungsbeitrag für einen generelleren anwendungsorientierten Ansatz für Mikro-Fairness.

Abstract

During the last decade Quality-of-Service (QoS) issues have been investigated successfully within the research community, and it only requires a further continuous process of standardization and commercial deployment to fulfill the evolving demands of modern networking applications concerning QoS.

Nonetheless, we find that the complementary issue of fairness has to date been addressed only insufficiently. With this thesis we contribute to research in that area.

We provide the link between theoretical results from the field of political economics and the needs of computer networks. Fairness definitions are given and evaluated for their use in networking. In addition, we present a first comprehensive classification of existing work on fairness in computer networks. For that reason, we divide fairness into the two dimensions: *macro-fairness* and *micro-fairness*. By *macro-fairness* we mean the distribution of network resources in a generally fair manner from which all users and applications benefit, whereas *micro-fairness* aims at fine-grained fairness for specific applications in competitive scenarios, examples of which are real-time trading systems, online auctions, etc.

Our research contribution in the field of macro-fairness is two-fold: We introduce a novel architectural proposal called *iSSD*, which achieves user-based fairness of a complexity similar to that of DiffServ, and we present our new TCP-friendly congestion control mechanism called *MPLBcc*. We validate the iSSD architecture for providing macro-fairness by means of formal analysis and network simulations. Our second research contribution to macro-fairness, MPLBcc, outperforms other TCP-

friendly mechanisms in the special case where multiple simultaneous transmission paths are available. We present both simulation results as well as field experience from the implementation of a variant of MPLBcc that was integrated as part of a bigger project for digital video transmission over the cable network return channel.

With regard to micro-fairness, we designed and validated new algorithms that achieve fair group-delay constrained traffic, which is especially useful for competitive applications such as stock trading or network games. As part of that work, we developed an algorithm for estimating round-trip time that uses *circulating envelopes* to help prevent cheating. Furthermore, we designed and implemented a novel architectural component for micro-fairness that fulfills two goals: it allows to add fairness to existing networking application scenarios without modifications to the original client-server application, and, secondly, it serves as a research platform, demonstrating that in specific cases, the problem of fairness can be solved by mapping it to the problem of optimal trading on a virtual market. This component, which we call *fairness proxy*, is our research contribution towards a more general application-layer approach to micro-fairness.

Acknowledgements

I gratefully acknowledge the support and collaboration of all the people without whom this thesis would not have been possible.

First and foremost, I wish to thank Professor Wolfgang Effelsberg for his guidance, excellent advice and many helpful suggestions throughout the years of my Ph.D. I especially thank him for his open-minded way of remotely supporting me during the final years and his ability to share a contagious fascination for the generation of innovative ideas. I also thank Professor Stefan Fischer for his suggestions during the initial phase of my thesis.

I am grateful to my colleagues Dr. Christoph Kuhmünch, Dr. Jörg Widmer and Professor Dr. Martin Mauve for fruitful discussions and our collaboration on a variety of issues related to the field of my research. Many thanks also to Betty Weyerer for her very valuable help concerning linguistic matters during the final phase of the Ph.D.

I am indebted to Dr. Albert Banchs for our brainstorming sessions on quality-of-service and fairness, during one of which he invented the terms ”macro-fairness” and ”micro-fairness”. Albert also co-authored various papers on the initial SSD architecture. Thanks to my other former colleagues at NEC, especially to Dr. Heinrich Stüttgen, who contracted me for work related to the early phase of my Ph.D.

Furthermore, I am deeply grateful to my friend Professor Joaquín Luque for his ever warm and welcoming support, both on private and on professional matters; I also

thank all my collaborators at the University of Seville, especially for their support during the Muvitran project.

My work benefited indirectly from various sojourns in Canada and my early days of struggling with concurrent programming, for which I am indebted to Professor Peter Buhr.

Many thanks also to my friends and colleagues of the former R & D group at Enditel where I worked throughout the last three years of my Ph.D. Enditel provided me with a job that made the pursuit of my Ph.D. a really interesting challenge; without that it would have been only half the experience.

Finally, but most importantly, I want to thank my parents Monika and Ronald as well as my aunt Maria for their continuous support in all imaginable ways. Likewise, I thank Guadalupe and her family, my brother and my sister, as well as all my German, Spanish and Canadian friends for their significant non-technical help and understanding during these years. It is all of you who finally made my Ph.D. possible.

Contents

Abbreviations

ABR	Available Bit Rate
ACK	Acknowledgement
AN	Access Node
APON	ATM Passive Optical Networks
ATM	Asynchronous Transfer Mode
CBQ	Class-Based Queueing
CBR	Constant Bit Rate
CSFQ	Core-Stateless Fair Queueing
CSMA/CA	Carrier Sense Multiple Access with Collision Avoidance
CTS	Clear-to-Send
CHOKe	CHOose and Keep for responsive flows, CHOose and kill for unresponsive flows
DFWMAC	Distributed Foundation Wireless Medium Access Control
DiffServ	Differentiated Services
DPS	Dynamic Packet State
DRR	Deficit Round Robin
DSLAM	Digital Subscriber Line Access Multiplexer
DWRR	Deficit Weighted Round Robin
ECN	Explicit Congestion Notification
FDM	Frequency Division Multiplexing
FEC	Foward Error Correction
FFQ	Frame-based Fair Queueing
FIFO	First-in-first-out
FLID-DL	Fair Layered Increase/Decrease with Dynamic Layering
FQ	Fair Queueing
FRED	Flow Random Early Drop
GPS	Generalized Processor Sharing
HDLC	High-level Data Link Control
IKE	Internet Key Exchange
IP	Internet Protocol
iSSD	Improved Scalable Share Differentiation
ISP	Internet Service Provider
LAN	Local Area Network
LDA+	Loss-Delay Based Adaption
LPM	Loss Path Multiplicity
LPR	Linear Proportional Response
LQD	Longest Queue Drop

LTS	Layered Transmission Scheme
MAC	Medium Access Control
MACAW	Media Access Protocol for wireless LANs
MLDA	Multicast Loss-Delay Based Adaption
MPEG	Motion Pictures Expert Group
MPLBcc	Multi-path Load-Balanced congestion control
MTCP	Multicast TCP
MUVITRAN	MUlti-path VIdeo TRANsmision
NAT	Network Address Translation
NCA	Nominee-Based Congestion Avoidance
NTP	Network Time Protocol
OBLI	Origin-Based Linear Interpolation
OLT	Optical Line Termination
ONU	Optical Network Unit
OSPF	Open Shortest Path First
PCC	Probabilistic Congestion Control
PGMCC	Pragmatic general multicast congestion control
PGPS	Packet-by-packet Generalized Processor Sharing
POTS	Plain Old Telephony Service
PQ	Priority Queueing
QoS	Quality of service
RAS	Remote Access Server
RED	Random Early Detection
RFQ	Rainbow Fair Queueing
RIO	RED with In-profile / Out-profile marking
RIP	Routing Information Protocol
RLA	Random Listening Algorithm
RLC	Receiver-Driven Layered Congestion Control
RM	Resource Management
RSVP	Resource Reservation Protocol
RTCP	Real Time Control Protocol
RTP	Real Time Protocol
RTS	Request-to-Send
RTT	Round Trip Time
SACK	Selective Acknowledgement
SBLI	Secant-Based Linear Control
SCFQ	Self-Clocked Fair Queueing
SCORE	Stateless Core
SDH	Synchronous Digital Hierarchy
SRED	Stabilized RED
TEAR	TCP Emulation at Receivers
TFRP	TCP-friendly Transport Protocol
TFRC	TCP-Friendly Rate Control Protocol
TCP	Transmission Control Protocol
TDM	Time Division Multiplexing
TUF	Tag-based Unified Fairness
UDP	User Datagram Protocol

USD	User-Share Differentiation
VC	Virtual Clock
WAN	Wide-Area Network
WDM	Wavelength Division Multiplexing
WFQ	Weighted Fair Queueing
WRR	Weighted Round Robin

List of Figures

List of Tables