

Web Content Caching and Distribution

Web Content Caching and Distribution

Proceedings of the 8th International Workshop

Edited by

Fred Douglass

*IBM T.J. Watson Research Center,
Hawthorne, NY, U.S.A.*

and

Brian D. Davison

*Department of Computer Science and Engineering,
Lehigh University, Bethlehem, PA, U.S.A.*



KLUWER ACADEMIC PUBLISHERS

NEW YORK, BOSTON, DORDRECHT, LONDON, MOSCOW

eBook ISBN: 1-4020-2258-1
Print ISBN: 1-4020-2257-3

©2004 Springer Science + Business Media, Inc.

Print ©2004 Kluwer Academic Publishers
Dordrecht

All rights reserved

No part of this eBook may be reproduced or transmitted in any form or by any means, electronic, mechanical, recording, or otherwise, without written consent from the Publisher

Created in the United States of America

Visit Springer's eBookstore at:
and the Springer Global Website Online at:

<http://www.ebooks.kluweronline.com>
<http://www.springeronline.com>

Contents

A Message from the Workshop Chairs	ix
Credits	xi
Contributing Authors	xiii
 Part 1 – Mobility	
Mobility-aware server selection for mobile streaming multimedia content distribution networks <i>Muhammad Mukarram Bin Tariq, Ravi Jain, Toshiro Kawahara</i>	1
Performance of PEPs in cellular wireless networks <i>Pablo Rodriguez, Vitali Fridman</i>	19
 Part 2 – Applications	
Edge caching for directory based Web applications: Algorithms and performance <i>Apurva Kumar, Rajeev Gupta</i>	39
Computing on the edge: A platform for replicating Internet applications <i>Michael Rabinovich, Zhen Xiao, Amit Aggarwal</i>	57
Scalable consistency maintenance for edge query caches: Exploiting templates in Web applications <i>Khalil Amiri, Sara Sprenkle, Renu Tewari, Sriram Padmanabhan</i>	79
 Part 3 – Architectures	
Proxy+: Simple proxy augmentation for dynamic content processing <i>Chun Yuan, Zhigang Hua, Zheng Zhang</i>	91
Synopsis: Multicast cloud with integrated multicast and unicast content distribution routing <i>Dan Li, Arun Desai, Zheng Yang, Kenneth Mueller, Stephen Morris, Dmitry Stavisky</i>	109
Synopsis: A large enterprise content distribution network: Design, implementation and operation <i>Jacobus Van der Merwe, Paul Gausman, Chuck Cranor, Rustam Akhmarov</i>	119

Synopsis: Architectural choices for video-on-demand systems <i>Anwar Al Hamra, Ernst W. Biersack, Guillaume Urvoy-Keller</i>	129
---	-----

Part 4 – Multimedia

Dynamic cache reconfiguration strategies for a cluster-based streaming proxy <i>Yang Guo, Zihui Ge, Bhuvan Urgaonkar, Prashant Shenoy, Don Towsley</i>	139
---	-----

Stream engine: A new kernel interface for high-performance Internet streaming servers <i>Jonathan Lemon, Zhe Wang, Zheng Yang, Pei Cao</i>	159
---	-----

Streaming flow analyses for prefetching in segment-based proxy caching to improve delivery quality <i>Songqing Chen, Bo Shen, Susie Wee, Xiaodong Zhang</i>	171
--	-----

Part 5 – Customization

Subscription-enhanced content delivery <i>Mao Chen, Jaswinder Pal Singh, Andrea LaPaugh</i>	187
--	-----

Cooperative architectures and algorithms for discovery and transcoding of multi-version content <i>Claudia Canali, Valeria Cardellini, Michele Colajanni, Riccardo Lancellotti, Philip S. Yu</i>	205
---	-----

Synopsis: User specific request redirection in a content delivery network <i>Sampath Rangarajan, Sarit Mukherjee, Pablo Rodriguez</i>	223
--	-----

Part 6 – Peer-to-Peer

Friendships that last: Peer lifespan and its role in P2P protocols <i>Fabian E. Bustamante, Yi Qiao</i>	233
--	-----

Synopsis: A fine-grained peer sharing technique for delivering large media files over the Internet <i>Mengkun Yang, Zongming Fei</i>	247
---	-----

Part 7 – Performance and Measurement

Proxy-cache aware object bundling for Web access acceleration <i>Chi Hung Chi, HongGuang Wang, William Ku</i>	257
--	-----

Synopsis: A case for dynamic selection of replication and caching strategies <i>Swaminathan Sivasubramanian, Guillaume Pierre, Maarten van Steen</i>	275
---	-----

Synopsis: Link prefetching in Mozilla: A server-driven approach <i>Darin Fisher, Gagan Saksena</i>	283
---	-----

Synopsis: A generalized model for characterizing content modification dynamics of Web objects	293
<i>Chi Hung Chi, HongGuang Wang</i>	

Part 8 – Delta Encoding

Server-friendly delta compression for efficient Web access	303
<i>Anubhav Savant, Torsten Suel</i>	

Evaluation of ESI and class-based delta encoding	323
<i>Mor Naaman, Hector Garcia-Molina, Andreas Paepcke</i>	

Author Index	345
--------------	-----

A Message from the Workshop Chairs

Dear Participant:

Welcome to the 8th International Web Caching and Content Delivery Workshop. Since our first meeting in 1996, this workshop has served as the premiere forum for researchers and industry technologists to exchange research results and perspectives on future directions in Internet content caching and content delivery. This year we received 46 submissions, of which 15 have been selected as full-length papers and 8 as synopses. We extend our thanks to the authors of the selected papers, all of which are included in these proceedings. In addition to technical presentations, we are pleased to have Bill Weihl of Akamai to present the keynote address, and a panel discussion on uncachable content organized by Zhen Xiao of AT&T Labs – Research.

While originally scheduled to be held in Beijing, China, the workshop moved to the US this year as a result of the concerns over the SARS virus. We are indebted to our industrial sponsor, IBM, for providing the facilities in which to hold the workshop. The T.J. Watson Research Center that serves as our venue spans three sites across two states, and is the headquarters for the eight IBM research labs worldwide. We are also grateful to the members of the program committee for helping to select a strong program, and to the members of the steering committee who continue to provide advice and guidance, even as plans are made for next year's workshop.

In past years, we have found great topics and fruitful discussion as people from industry and academia interact. We are confident that you will experience the same at this year's workshop.

Brian D. Davison	Fred Douglass
<i>General Chair</i>	<i>Program Chair</i>

Credits

General Chair

Brian D. Davison, *Lehigh University*

Program Chair

Fred Douglass, *IBM T.J. Watson Research Center*

Program Committee

Martin Arlitt, *University of Calgary*

Remzi Arpaci-Dusseau, *University of Wisconsin*

Chi-Hung Chi, *National University of Singapore*

Mike Dahlin, *University of Texas at Austin*

Fred Douglass, *IBM T.J. Watson Research Center*

Zongming Fei, *University of Kentucky*

Leana Golubchik, *University of Southern California*

Jaeyeon Jung, *MIT LCS*

Dan Li, *Cisco Systems, Inc.*

Guillaume Pierre, *Vrije Universiteit, Amsterdam*

Weisong Shi, *Wayne State University*

Oliver Spatscheck, *AT&T Labs – Research*

Renu Tewari, *IBM Almaden Research Center*

Amin Vahdat, *Duke University*

Geoff Voelker, *University of California, San Diego*

Zhen Xiao, *AT&T Labs – Research*

Steering Committee

Azer Bestavros, *Boston University*

Pei Cao, *Cisco*

Jeff Chase, *Duke University*

Valentino Cavalli, *Terena*

Peter Danzig, *University of Southern California*

John Martin, *Network Appliance*

Michael Rabinovich, *AT&T Labs – Research*

Wojtek Sylwestrzak, *Warsaw University*

Duane Wessels, *The Measurement Factory*

Keynote Speaker

William Weihl, *Akamai Technologies, Inc.*

Panel Moderator

Zhen Xiao, *AT&T Labs – Research*

Panelists

Indranil Gupta, *University of Illinois, Urbana-Champaign*

Arun Iyengar, *IBM Research*

Michael Rabinovich, *AT&T Labs – Research*

Torsten Suel, *Polytechnic University*

William Weihl, *Akamai Technologies, Inc.*

Session Chairs

Chi-Hung Chi, *National University of Singapore*

Brian D. Davison, *Lehigh University*

Fred Douglass, *IBM T.J. Watson Research Center*

Zongming Fei, *University of Kentucky*

Michael Rabinovich, *AT&T Labs – Research*

Pablo Rodriguez, *Microsoft Research, Cambridge*

Oliver Spatscheck, *AT&T Labs – Research*

Torsten Suel, *Polytechnic University*

External Reviewers

Benjamin Atkin

Limin Wang

Yan Chen

Craig Wills

Subhabrata Sen

Kun-Lung Wu

Andrew Tridgell

Contributing Authors

Amit Aggarwal	<i>Microsoft</i>
Anwar Al Hamra	<i>Institut Eurecom</i>
Rustam Akhmarov	<i>AT&T Labs – Research</i>
Khalil Amiri	<i>Imperial College London</i>
Ernst W. Biersack	<i>Institut Eurecom</i>
Fabian E. Bustamante	<i>Department of Computer Science, Northwestern University</i>
Pei Cao	<i>Cisco Systems, Inc.</i>
Claudia Canali	<i>University of Parma</i>
Valeria Cardellini	<i>University of Roma “Tor Vergata”</i>
Mao Chen	<i>Department of Computer Science, Princeton University</i>
Songqing Chen	<i>College of William and Mary</i>
Chi Hung Chi	<i>National University of Singapore</i>
Michele Colajanni	<i>University of Modena and Reggio</i>
Chuck Cranor	<i>AT&T Labs – Research</i>
Arun Desai	<i>Cisco Systems, Inc.</i>
Zongming Fei	<i>Department of Computer Science, University of Kentucky</i>

Darin Fisher	<i>IBM</i>
Vitali Fridman	<i>Microsoft Research, Cambridge</i>
Hector Garcia-Molina	<i>Department of Computer Science, Stanford University</i>
Paul Gausman	<i>AT&T Labs – Research</i>
Zihui Ge	<i>Department of Computer Science, University of Massachusetts at Amherst</i>
Yang Guo	<i>Department of Computer Science, University of Massachusetts at Amherst</i>
Rajeev Gupta	<i>IBM India Research Lab</i>
Zhigang Hua	<i>Institute of Automation, Chinese Academy of Sciences</i>
Ravi Jain	<i>DoCoMo Communications Laboratories USA</i>
Toshiro Kawahara	<i>DoCoMo Communications Laboratories USA</i>
William Ku	<i>National University of Singapore</i>
Apurva Kumar	<i>IBM India Research Lab</i>
Riccardo Lancellotti	<i>University of Roma “Tor Vergata”</i>
Andrea LaPaugh	<i>Department of Computer Science, Princeton University</i>
Jonathan Lemon	<i>Cisco Systems, Inc.</i>
Dan Li	<i>Cisco Systems, Inc.</i>
Stephen Morris	<i>Cisco Systems, Inc.</i>
Kenneth Mueller	<i>Cisco Systems, Inc.</i>
Sarit Mukherjee	<i>Microsoft Research, Cambridge</i>

Mor Naaman	<i>Department of Computer Science, Stanford University</i>
Sriram Padmanabhan	<i>IBM Santa Teresa Lab</i>
Andreas Paepcke	<i>Department of Computer Science, Stanford University</i>
Guillaume Pierre	<i>Department of Mathematics and Computer Science, Vrije Universiteit, Amsterdam</i>
Yi Qiao	<i>Department of Computer Science, Northwestern University</i>
Michael Rabinovich	<i>AT&T Labs – Research</i>
Sampath Rangarajan	<i>Lucent Technologies Bell Laboratories</i>
Pablo Rodriguez	<i>Microsoft Research, Cambridge</i>
Gagan Saksena	<i>AOL</i>
Anubhav Savant	<i>CIS Department, Polytechnic University</i>
Bo Shen	<i>Hewlett-Packard Laboratories</i>
Prashant Shenoy	<i>Department of Computer Science, University of Massachusetts at Amherst</i>
Jaswinder Pal Singh	<i>Department of Computer Science, Princeton University</i>
Swaminathan Sivasubramanian	<i>Department of Mathematics and Computer Science, Vrije Universiteit, Amsterdam</i>
Sara Sprenkle	<i>Duke University</i>
Dmitry Stavisky	<i>Cisco Systems, Inc.</i>
Torsten Suel	<i>CIS Department, Polytechnic University</i>
Muhammad Mukarran Bin Tariq	<i>DoCoMo Communications Laboratories USA</i>
Renu Tewari	<i>IBM Almaden Research Center</i>

Don Towsley	<i>Department of Computer Science, University of Massachusetts at Amherst</i>
Bhuvan Urgaonkar	<i>Department of Computer Science, University of Massachusetts at Amherst</i>
Guillaume Urvoy-Keller	<i>Institut Eurecom</i>
Jacobus Van der Merwe	<i>AT&T Labs – Research</i>
Maarten van Steen	<i>Department of Mathematics and Computer Science, Vrije Universiteit, Amsterdam</i>
HongGuang Wang	<i>National University of Singapore</i>
Zhe Wang	<i>Cisco Systems, Inc.</i>
Susie Wee	<i>Hewlett-Packard Laboratories</i>
Zhen Xiao	<i>AT&T Labs – Research</i>
Zheng Yang	<i>Cisco Systems, Inc.</i>
Mengkun Yang	<i>Department of Computer Science, University of Kentucky</i>
Philip S. Yu	<i>IBM T.J. Watson Research Center</i>
Chun Yuan	<i>Microsoft Research Asia</i>
Xiaodong Zhang	<i>College of William and Mary</i>
Zheng Zhang	<i>Microsoft Research Asia</i>