

Commenced Publication in 1973

Founding and Former Series Editors:

Gerhard Goos, Juris Hartmanis, and Jan van Leeuwen

Editorial Board

David Hutchison

Lancaster University, UK

Takeo Kanade

Carnegie Mellon University, Pittsburgh, PA, USA

Josef Kittler

University of Surrey, Guildford, UK

Jon M. Kleinberg

Cornell University, Ithaca, NY, USA

Friedemann Mattern

ETH Zurich, Switzerland

John C. Mitchell

Stanford University, CA, USA

Moni Naor

Weizmann Institute of Science, Rehovot, Israel

Oscar Nierstrasz

University of Bern, Switzerland

C. Pandu Rangan

Indian Institute of Technology, Madras, India

Bernhard Steffen

University of Dortmund, Germany

Madhu Sudan

Massachusetts Institute of Technology, MA, USA

Demetri Terzopoulos

New York University, NY, USA

Doug Tygar

University of California, Berkeley, CA, USA

Moshe Y. Vardi

Rice University, Houston, TX, USA

Gerhard Weikum

Max-Planck Institute of Computer Science, Saarbruecken, Germany

Can Türker Maristella Agosti
Hans-Jörg Schek (Eds.)

Peer-to-Peer, Grid, and Service-Oriented in Digital Library Architectures

6th Thematic Workshop
of the EU Network of Excellence DELOS
Cagliari, Italy, June 24-25, 2004
Revised Selected Papers



Springer

Volume Editors

Can Türker

Functional Genomics Center Zurich, UNI / ETH Zurich
Winterthurerstrasse 190, Irchel, Y32 H06, 8057 Zurich, Switzerland
E-mail: tuerker@fgcz.ethz.ch

Maristella Agosti

University of Padua, Department of Information Engineering
Via Gradenigo 6/a, 35131 Padua, Italy
E-mail: maristella.agosti@unipd.it

Hans-Jörg Schek

University for Health Sciences, Medical Informatics and Technology
Institute for Information Systems
Eduard Wallnöfer-Zentrum 1, 6060 Hall i. Tirol, Austria
E-mail: schek@umit.at

Library of Congress Control Number: 2005931330

CR Subject Classification (1998): H.2, H.4, H.3, H.2.4, H.5

ISSN 0302-9743
ISBN-10 3-540-28711-6 Springer Berlin Heidelberg New York
ISBN-13 978-3-540-28711-7 Springer Berlin Heidelberg New York

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, re-use of illustrations, recitation, broadcasting, reproduction on microfilms or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer. Violations are liable to prosecution under the German Copyright Law.

Springer is a part of Springer Science+Business Media

springeronline.com

© Springer-Verlag Berlin Heidelberg 2005
Printed in Germany

Typesetting: Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, India
Printed on acid-free paper SPIN: 11549819 06/3142 5 4 3 2 1 0

Preface

Newer computing areas like peer-to-peer, grid, and service-oriented computing provide a number of opportunities and challenges for architectures of future digital libraries. Peer-to-peer data management allows for loosely coupled integration of information services and sharing of information such as recommendations and annotations. Grid computing middleware is needed because certain services within digital libraries are complex and computationally intensive, e.g., extraction of features in multimedia documents to support content-based similarity search or for information mining in biomedical data. The service-orientation provides mechanisms to describe the semantics and usage of information services and to combine services into workflow processes for sophisticated search and maintenance of dependencies. Elements of all three directions should be combined in a synthesis for future digital libraries architectures.

This volume contains selected and revised papers from the Sixth Thematic Workshop of the EU Network of Excellence DELOS on Digital Library Architectures, which was held in S. Margherita di Pula (Cagliari), Italy, 24–25 June 2004. This workshop was co-located with the 12th Italian Symposium on Advanced Database Systems (SEBD 2004) and organized jointly by the DELOS Network of Excellence and the Department of Information Engineering of the University of Padua, Italy. DELOS (<http://www.delos.info>) is an interdisciplinary EU FP6 Network of Excellence with a broad vision: future digital libraries should enable any citizen to access human knowledge any time and anywhere, in a friendly, multi-modal, efficient and effective way. The main objective of DELOS is to define and conduct a joint program of activities in order to integrate and coordinate the ongoing research activities of the research teams in the field of digital libraries for the purpose of developing next generation digital library technologies, and the papers in this proceedings volume address a broad range of issues in this field.

This volume presents and discusses relevant architectural aspects that must be taken into account when designing Digital Library Management Systems (DLMS) able to face the challenges the final users impose on future DLMS. The authors of the papers have made an effort to concentrate on and deal with the architectural areas where present Digital Library systems present specific problems to be solved.

The impact on open problems in this area of digital libraries made by the papers in this volume is relevant, as the reader will discover. Here are some of the topics covered:

- The relation of the EGEE (Enabling Grids for E-science) project to digital libraries is explained and examples of how the project infrastructure can be used in the field are highlighted.
- The fundamental tenets of Service-Oriented Architecture (SOA) and their relevance to Internet-scale computing (or Grid computing) are addressed and clarified together with the presentation of the application of SOA principles to building Internet-scale applications using Web Services technologies and how to avoid software pitfalls.

- A series of experiments confirm that SOA and a service-oriented component architecture is indeed applicable to building flexible, effective and efficient digital library systems, by evaluating issues of simplicity and understandability, reusability, extensibility and performance.
- Digital libraries in healthcare represent an important type of future implementation. This type of system hosts an inherently large and continually growing collection of digital information. Especially in medical digital libraries, this information needs to be analyzed and processed in a timely manner. Sensor data streams, for instance, providing continuous information on patients, have to be processed online in order to detect critical situations. A novel information management infrastructure based on a hyperdatabase system that combines the process-based composition of services and operators needed for sensor data stream processing with advanced grid features is presented to solve this type of challenge.
- The problem of collaborative search across a large number of digital libraries and query routing strategies in a peer-to-peer (P2P) environment is faced and experiments with the MINERVA prototype testbed study the benefits and costs of P2P search for keyword queries.
- Similarity search in metric spaces represents an important paradigm for content-based retrieval of many applications. Scalable and distributed new types of indexes are proposed and experimented by exploiting parallelism in a dynamic network of computers.
- Similarity search can benefit from the support of an infrastructure that combines various information technologies like databases, service-oriented architectures, peer-to-peer and grid computing. Query distribution and load balancing based on domain-specific knowledge can be exploited by such an infrastructure and it is shown that it is possible to reduce query response times.
- Since digital libraries are dispersed over several peers of a steadily increasing network, dedicated peers may provide specialized services. Examples of this sort of system would be a system that performs specialized image similarity searches or a system that manages annotations in an automatic way in order to support users and their annotative practices.

Finally, we would like to thank all those who contributed directly or indirectly to this volume. Our thanks goes to all the authors for submitting their papers to these post-proceedings as well as to all the members of the program committee for reviewing and helping in selecting the papers.

June 2005

Can Türker
Maristella Agosti
Hans-Jörg Schek

Organization

Program Committee

Maristella Agosti	University of Padua, Italy
Elisa Bertino	University of Milan, Italy
Donatelli Castelli	CNR-ISTI, Italy
Wilhelm Hasselbring	OFFIS Oldenburg, Germany
Yannis Ioannidis	University of Athens, Greece
Martin Kersten	CWI Amsterdam, The Netherlands
Erich Neuhold	FhG Darmstadt, Germany
Hans-Jörg Schek	UMIT Innsbruck, Austria (chair)
Heiko Schuldt	UMIT Innsbruck, Austria
Can Türker	FGCZ Zurich, Switzerland (co-chair)
Gerhard Weikum	MPI Saarbrücken, Germany
Pavel Zezula	Masaryk University Brno, Czech Republic

External Referees

S. Balko
G. Brettlecker

Table of Contents

An Overview of the EGEE Project <i>Bob Jones</i>	1
Grid Computing Using Web Services Technologies <i>Savas Parastatidis, Paul Watson, Jim Webber</i>	9
Similarity Grid for Searching in Metric Spaces <i>Michal Batko, Claudio Gennaro, Pavel Zezula</i>	25
Organisation-Oriented Super-Peer Networks for Digital Libraries <i>Ludger Bischofs, Ulrike Steffens</i>	45
A Combined Hyperdatabase and Grid Infrastructure for Data Stream Management and Digital Library Processes <i>Manfred Wurz, Gert Brettlecker, Heiko Schuldt</i>	63
The MINERVA Project: Towards Collaborative Search in Digital Libraries Using Peer-to-Peer Technology <i>Matthias Bender, Sebastian Michel, Christian Zimmer, Gerhard Weikum</i>	80
Distribution Alternatives for Superimposed Information Services in Digital Libraries <i>Sudarshan Murthy, David Maier, Lois Delcambre</i>	96
Towards a Service-Oriented Architecture for the Adaptive Delivery of Heterogeneous Cultural Resources <i>Jérôme Godard, Frédéric Andrès, Elham Andaroodi, Katsumi Maruyama</i>	112
Analysis and Evaluation of Service Oriented Architectures for Digital Libraries <i>Hussein Suleman</i>	130
A System Architecture as a Support to a Flexible Annotation Service <i>Maristella Agosti, Nicola Ferro</i>	147
A Service-Oriented Grid Infrastructure for Multimedia Management and Search <i>Michael Mlivoncic, Christoph Schuler, Can Türker, Sören Balko</i>	167

StreamOnTheFly: A Network for Radio Content Dissemination <i>László Kovács, András Micsik, Máté Pataki, Robert Stachel</i>	188
Supporting Information Access in Next Generation Digital Library Architectures <i>Predrag Knežević, Bhaskar Mehta, Claudia Niederée, Thomas Risse, Ulrich Thiel, Ingo Frommholz</i>	207
Query Trading in Digital Libraries <i>Fragkiskos Pentaris, Yannis Ioannidis</i>	223
Moving Digital Library Service Systems to the Grid <i>Leonardo Candela, Donatella Castelli, Pasquale Pagano, Manuele Simi</i>	236
Author Index	261