Felix F. Ramos
Victor Larios Rosillo
Herwig Unger (Eds.)

LNCS 3563

Advanced Distributed Systems

5th International School and Symposium, ISSADS 2005
Guadalajara, Mexico, January 2005
Revised Selected Papers

Springer

Felix F. Ramos   Victor Larios Rosillo
Herwig Unger (Eds.)

Advanced Distributed Systems

5th International School and Symposium, ISSADS 2005
Guadalajara, Mexico, January 24-28, 2005
Revised Selected Papers

Springer

Volume Editors

Felix F. Ramos
CINVESTAV Div. Guadalajara
Prol. Lopez Mateos Sur 590, Guadalajara, Jal. 45000, Mexico
E-mail: framos@gdl.cinvestav.mx

Victor Larios Rosillo
Universidad de Guadalajara, CUCEA
Department "Sistemas de Informacion"
799, Periferico Norte, Ed. L308, Zapopan, Jal. 45100, Mexico
E-mail: vlarios@acm.org

Herwig Unger
Universität Rostock, Institut für Informatik
Albert-Einstein-Str. 23, 18051 Rostock, Germany
E-mail: hunger@informatik.uni-rostock.de

Library of Congress Control Number: 2005932536

CR Subject Classification (1998): C.2.4, I.2.11, D.2.12, D.1.3, D.4, H.3, H.4

ISSN 0302-9743
ISBN-10 3-540-28063-4 Springer Berlin Heidelberg New York
ISBN-13 978-3-540-28063-7 Springer Berlin Heidelberg New York

Springer is a part of Springer Science+Business Media

springeronline.com

Printed in Germany

Typesetting: Camera-ready by author, data conversion by Olgun Computergrafik
Printed on acid-free paper SPIN: 11533962 06/3142 5 4 3 2 1 0

Lecture Notes in Computer Science 3563

Preface

It is our pleasure to present the papers accepted and presented at the 5th International School and Symposium on Advanced Distributed Systems (ISSADS) in this LNCS volume. The symposium was held in the city of Guadalajara, Mexico from January 24 to 28, 2005. The organization team was composed of members of CINVESTAV Guadalajara, Rostock University in Germany, the CUCEI and CUCEA campuses of Guadalajara University, and Instituto Tecnológico y de Estudios Superiores de Occidente, ITESO. The symposium is already a well-established annual meeting, at which scientists and people from the industrial field meet and discuss the progress of applications and the theory of distributed systems in a forum during the last week of January. This year, more than 250 people from 3 continents attended the conference. Most of them are scientists, teachers, students and engineers from the local industry.

The papers presented in the sessions of the symposium cover not only the subjects of distributed systems from the system level and applications, but also contributions from the area of theory and artificial intelligence concepts. These papers were selected out of more than 100 submissions. There was a selection filter in which each paper was evaluated by at least three members of the international Program Committee, who came from research institutions of good reputation all over the world.

We were pleased to listen to an invited talk given by Ali Hurson, a distinguished lecturer of ACM, on distributed data bases, and another one presented by Carlos Coello who become famous by its works in evolutive computation.

Besides the scientific sessions, ISSADS 2005 followed a tradition by offering six high-level tutorials covering the important topics of security, emerging technologies, knowledge management, multi-agent systems, design of complex systems, and modern Internet technologies. Furthermore, the conference was also a forum for presentations by the main industrial companies in the area of distributed systems – this year INTEL, HP, Texas Instruments and IBM took part in our event. All events enabled the participants to extend their knowledge in certain areas, to have a lot of worthwhile discussions, and to establish new, interesting and useful contacts.

We would like to extend our gratitude to all the members of the Program and Organizing Committees as well as their teams. We would like to note our special appreciation of Dr. Francisco Medina, Director of the Council of Science and Research of Jalisco, Mexico, for his financial support. We also thank J. Luis Leyva, Director of CINVESTAV, for all his help, and we thank Dionisio de Niz from ITESO, and Patricia Mendoza, Luis Gutierrez and Carlos Franco from Guadalajara University for their logistic support during all the events. Last but not least, we are grateful to the Chambers of Electronic and Information Technologies of Jalisco, Mexico. Gabriela Tagliaprieta put a lot of effort into supporting all participants. Finally, a special recognition is given to all the re-

searchers who submitted their papers to ISSADS 2005. They led to the success of ISSADS. We would be happy to see all of you and hopefully some new participants at ISSADS 2006, which will be held again in our nice city, Guadalajara, Mexico.

February, 2005

Félix F. Ramos C.
Victor Larios Rosillo
Herwig Unger

Program Committee

Chair Félix Francisco Ramos Corchado, Cinvestav Guadalajara
Co-chair Victor Manuel Larios Rosillo, CUCEA, Universidad de Guadalajara
Editor Herwig Unger, Rostock University, Germany

Scientific Committee

Anastase Adonis
Alain Bui
Anbulagan
Arbad Farhad
Gilbert Babin
Barradas Hector Ruiz
Jean Paul Barthès
Nadia Bennani
Thomas Böhme
Pierre Boulanger
Marc Bui
Humberto Cervantes
Hugo Coyote
Diaz Michel
Didier Donsez
Drira Khalil
Estivill Castro Vladimir
Eduardo Fernandez
Alexander Gelbukh
Emmanuelle Grislin
Guzmán Arenas Adolfo
Guilherme Bittencourt
Hervé Luga
Hubert Kihl
Kim Kane
Jean-Luc Koning
Peter Kropf
Sylvain Lecomte
Linpeng Huang
Guerrero Miguel López
Aurelio López
M.L. Ernesto López
Maralin Jimenez Ricardo
M. Nakano Miyatake
René Mandiau
Andrei N. Tchernykh
Elizabeth Pérez Cortéz
Meana Hector Pérez
Pires Ramos Milton
Ren Fuji
Román Graciela
Scalabrin Edson Emilio
Sybertin Blanc
Tazi Said
Thierry Delot
Torres Deni
Helena Unger
Joaquin Vila
Ricardo Vilalta
Thierry Villemur
Young-Hwan Park
Yves Duthen
Andreas Zekl

Organization

Public Relations and Logistics-Chair Ana Gabriela Tagliapietra, CINVESTAV Guadalajara
Logistics and Webmasters Joel Espinosa and Jorge Hernández, CINVESTAV Guadalajara

Lecture Notes in Computer Science

For information about Vols. 1–3593

please contact your bookseller or Springer

Vol. 3728: V. Paliouras, J. Vounckx, D. Verkest (Eds.), Integrated Circuit and System Design. XV, 753 pages. 2005.

Vol. 3710: M. Barni, I. Cox, T. Kalker, H.J. Kim (Eds.), Digital Watermarking. XII, 485 pages. 2005.

Vol. 3703: F. Fages, S. Soliman (Eds.), Principles and Practice of Semantic Web Reasoning. VIII, 163 pages. 2005.

Vol. 3702: B. Beckert (Ed.), Automated Reasoning with Analytic Tableaux and Related Methods. XIII, 343 pages. 2005. (Subseries LNAI).

Vol. 3698: U. Furbach (Ed.), KI 2005: Advances in Artificial Intelligence. XIII, 409 pages. 2005. (Subseries LNAI).

Vol. 3697: W. Duch, J. Kacprzyk, E. Oja, S. Zadrożny (Eds.), Artificial Neural Networks: Formal Models and Their Applications - ICANN 2005, Part II. XXXII, 1045 pages. 2005.

Vol. 3696: W. Duch, J. Kacprzyk, E. Oja, S. Zadrożny (Eds.), Artificial Neural Networks: Biological Inspirations - ICANN 2005, Part I. XXXI, 703 pages. 2005.

Vol. 3687: S. Singh, M. Singh, C. Apte, P. Perner (Eds.), Pattern Recognition and Image Analysis, Part II. XXV, 809 pages. 2005.

Vol. 3686: S. Singh, M. Singh, C. Apte, P. Perner (Eds.), Pattern Recognition and Data Mining, Part I. XXVI, 689 pages. 2005.

Vol. 3684: R. Khosla, R.J. Howlett, L.C. Jain (Eds.), Knowledge-Based Intelligent Information and Engineering Systems, Part IV. LXXIX, 933 pages. 2005. (Subseries LNAI).

Vol. 3683: R. Khosla, R.J. Howlett, L.C. Jain (Eds.), Knowledge-Based Intelligent Information and Engineering Systems, Part III. LXXIX, 1397 pages. 2005. (Subseries LNAI).

Vol. 3682: R. Khosla, R.J. Howlett, L.C. Jain (Eds.), Knowledge-Based Intelligent Information and Engineering Systems, Part II. LXXIX, 1371 pages. 2005. (Subseries LNAI).

Vol. 3681: R. Khosla, R.J. Howlett, L.C. Jain (Eds.), Knowledge-Based Intelligent Information and Engineering Systems, Part I. LXXX, 1319 pages. 2005. (Subseries LNAI).

Vol. 3678: A. McLysaght, D.H. Huson (Eds.), Comparative Genomics. VIII, 167 pages. 2005. (Subseries LNBI).

Vol. 3677: J. Dittmann, S. Katzenbeisser, A. Uhl (Eds.), Communications and Multimedia Security. XIII, 360 pages. 2005.

Vol. 3675: Y. Luo (Ed.), Cooperative Design, Visualization, and Engineering. XI, 264 pages. 2005.

Vol. 3674: W. Jonker, M. Petković (Eds.), Secure Data Management. X, 241 pages. 2005.

Vol. 3672: C. Hankin, I. Siveroni (Eds.), Static Analysis. X, 369 pages. 2005.

Vol. 3671: S. Bressan, S. Ceri, E. Hunt, Z.G. Ives, Z. Bellahsène, M. Rys, R. Unland (Eds.), Database and XML Technologies. X, 239 pages. 2005.

Vol. 3670: M. Bravetti, L. Kloul, G. Zavattaro (Eds.), Formal Techniques for Computer Systems and Business Processes. XIII, 349 pages. 2005.

Vol. 3665: K. S. Candan, A. Celentano (Eds.), Advances in Multimedia Information Systems. X, 221 pages. 2005.

Vol. 3664: C. Türker, M. Agosti, H.-J. Schek (Eds.), Peer-to-Peer, Grid, and Service-Orientation in Digital Library Architectures. X, 261 pages. 2005.

Vol. 3663: W.G. Kropatsch, R. Sablatnig, A. Hanbury (Eds.), Pattern Recognition. XIV, 512 pages. 2005.

Vol. 3662: C. Baral, G. Greco, N. Leone, G. Terracina (Eds.), Logic Programming and Nonmonotonic Reasoning. XIII, 454 pages. 2005. (Subseries LNAI).

Vol. 3661: T. Panayiotopoulos, J. Gratch, R. Aylett, D. Ballin, P. Olivier, T. Rist (Eds.), Intelligent Virtual Agents. XIII, 506 pages. 2005. (Subseries LNAI).

Vol. 3660: M. Beigl, S. Intille, J. Rekimoto, H. Tokuda (Eds.), UbiComp 2005: Ubiquitous Computing. XVII, 394 pages. 2005.

Vol. 3659: J.R. Rao, B. Sunar (Eds.), Cryptographic Hardware and Embedded Systems – CHES 2005. XIV, 458 pages. 2005.

Vol. 3658: V. Matoušek, P. Mautner, T. Pavelka (Eds.), Text, Speech and Dialogue. XV, 460 pages. 2005. (Subseries LNAI).

Vol. 3655: A. Aldini, R. Gorrieri, F. Martinelli (Eds.), Foundations of Security Analysis and Design III. VII, 273 pages. 2005.

Vol. 3654: S. Jajodia, D. Wijesekera (Eds.), Data and Applications Security XIX. X, 353 pages. 2005.

Vol. 3653: M. Abadi, L.d. Alfaro (Eds.), CONCUR 2005 – Concurrency Theory. XIV, 578 pages. 2005.

Vol. 3652: A. Rauber, S. Christodoulakis, A M. Tjoa (Eds.), Research and Advanced Technology for Digital Libraries. XVIII, 545 pages. 2005.

Vol. 3649: W.M.P. van der Aalst, B. Benatallah, F. Casati, F. Curbera (Eds.), Business Process Management. XII, 472 pages. 2005.

Vol. 3648: J.C. Cunha, P.D. Medeiros (Eds.), Euro-Par 2005 Parallel Processing. XXXVI, 1299 pages. 2005.

Vol. 3646: A. F. Famili, J.N. Kok, J.M. Peña, A. Siebes, A. Feelders (Eds.), Advances in Intelligent Data Analysis VI. XIV, 522 pages. 2005.

¥641.92元

Table of Contents

Database Systems

Distributed and Parallel Algorithms

Real-Time Distributed Systems

Cooperative Information Systems

Fault Tolerance

Information Retrieval

Modeling and Simulation

Wireless Networks and Mobile Computing

Artificial Life and Multi-agent Systems

Database System Architecture – A Walk Through Time: From Centralized Platform to Mobile Computing Keynote Address

Ali R. Hurson and Yu Jiao

Department of Computer Science and Engineering
The Pennsylvania State University, University Park, PA 16802, USA
{hurson,yjiao}@cse.psu.edu

Abstract. Classical distributed database systems monolithically offer distribution transparency and higher performance. This is made possible by making data available and closer to the application domain(s) that uses it by means such as the data distribution, duplication, and fragmentation. However, with the advances in technologies this monolithic and top down approach becomes insufficient. In the new networked computational environment, the data distribution issue has evolved into data integration from several heterogeneous and autonomous data sources. Heterogeneous distributed databases are designed to deal with issue of data integration and interoperability. They are developed to allow timely and reliable access to heterogeneous and autonomous data sources in an environment that is characterized as "sometime, somewhere."

The concept of mobility, where users access information through a remote connection with portable devices, has introduced additional complexities and restrictions to the heterogeneous distributed database system. This keynote address first introduces a three dimensional space to classify and identify the evolution of different classes of database systems. It also extensively discusses Heterogeneous Distributed Database Systems (HDDBS) and Mobile Data Access Systems (MDAS). Finally, it will address several research issues and their potential solutions.

1 Information Everywhere, Computers Everywhere

Advances in computation and communication technologies have enabled users to access information at anytime, from anywhere. However, this flexibility comes at the expenses of new challenges. These challenges stem from the diversity in the range and the exponential growth of information that is available to a user at any given time. The spread of computer networks, the wide breadth of access devices with different physical characteristics, and the extensive need for information sharing have created a demand for cooperation among pre-existing, distributed, heterogeneous, and autonomous information sources in an infrastructure that is characterized by:

- Low bandwidth,
- Frequent disconnection,
- High error rates,
- Limited processing resources, and
- Limited power sources.

Binary data was the main data format in 1970s and it has evolved into text, images, multi media, and sensor data during the past decades. In mid 1980s, it was estimated

F.F. Ramos et al. (Eds.): ISSADS 2005, LNCS 3563, pp. 1–9, 2005.

that the U.S. Patent Office and Trademark has a database of size 25 terabytes (1 terabytes = 10^{12} bytes) subject to search and update. In 1990s, it was estimated that the NASA's Earth Observing Project would generate more than 11,000 terabytes of data during the 15-year time-period. Recently, it was estimated that the amount of new information generated in 2002 was about 5 exabytes (1 exabyte = 10^{18} bytes).

The diversity in representation, the growth in size, and the increased availability of information also introduced new challenges in areas such as security and resource management (e.g., power and network bandwidth). The cost associated with viruses, unsolicited emails, and other attacks has grown exponentially. For example, it is estimated that 7000 new computer viruses were discovered in 2003 and the FBI approximated that computer viruses cost businesses $27 million during that time. As another example, in 2003, unsolicited email cost businesses $20 billion worldwide due to lost productivity, system overhead, user support, and anti-spam software.

The availability of heterogeneous, autonomous, and partially unreliable information in various forms and shapes brings out the following challenges:

- How to locate information intelligently, efficiently, and transparently?
- How to extract, process, and integrate relevant information efficiently?
- How to interpret information intelligently?
- How to provide uniform global access methods?
- How to support user and data source mobility?

2 Database Systems Taxonomy

Different parameters can be used to classify the architecture of data base systems. We classify data base systems along the following three parameters:

- **Physical infrastructure:** This dimension refers to the underlying platform composed of homogeneous/heterogeneous processing devices interconnected through different communication medium. The processing devices ranging from powerful parallel machines to portable units communicating with each other via a wide variety of communication medium ranging from land-based connection to wireless medium.
- **Services:** Along this dimension one can distinguish two classes in which either there is no distinction between services, or there is a distinction between user processes and data processes.
- **Distribution:** This dimension distinguishes distribution of processing, data, and control (also known as autonomy). Note that data distribution also includes data fragmentation and data replication.

2.1 Centralized Databases

A centralized database system is the one that runs on a single computer platform and does not interact with other computer systems. Based on our taxonomy, a centralized database is characterized by: its single processing unit, without distinction between its services, and without any notion of distribution. The centralized database systems can be further classified as, single-user configuration and multi-user configuration. Naturally, database systems designed for single-user configuration do not provide many facilities needed for a multi-user system e.g., concurrency control, and security.

2.2 Parallel Databases

Let us inject processing distribution into the scope of the centralized database system. This brings out the so called parallel database systems, in which several processing resources in cooperation with each other are intended to resolve users' requests. Note that in this environment there exist no notion of data and control distribution. In addition, there is no distinction between services that are provided by the database management system. Parallel configurations are aimed at improving the performance and throughput by distribution of a task at different granularity (fine, medium, or coarse) granularity among several processing units. The literature has introduced four classes of parallel systems:

- Shared Memory (tightly coupled) – All processors share a common global memory.
- Shared Disk (loosely coupled) – All processors share a common set of disks.
- Share Nothing – The processors share neither a common memory nor common disks.
- Hierarchical – A hybrid of the other models.

As noted in the literature; the shared memory configuration is not scalable and the communication network is the system bottleneck. Raid technology can be used to improve performance and reliability of the disk subsystem. The shared nothing configuration offers scalability at the expenses of high inter-processor communication cost. Finally, in the hierarchical topology, at the higher level, system acts as a shared nothing organization and at the lower level, each node could be a shared memory and/or a shared disk system.

2.3 Client/Server Topology

In sections 2.1 and 2.2 we looked at two configurations that did not make any distinction between the services provided by the database management system. Along the *services*, we distinguish two classes of functions:

- The data functions (the back end processes) – query processing, query optimization, concurrency control, and recovery.
- The user functions (the front end processes) – report writer and Graphical User Interface facilities.

This brings out the so called client/server topology that has functionality split between a server and multiple clients. The client/server topology can be grouped into multiple client/single server and multiple client/multiple server configurations. Functionality and processing capability of the client processors and communication speed between the client and server also distinguishes two classes of the client/server topology, namely, transaction server (thin client) and data server (fat client). The client/server topology is one step towards distributed processing. It offers a user-friendly environment, simplicity of implementation, and high degree of hardware utilization at the server side.

2.4 Peer-to-Peer Topology

This topology is a direct evolution of the client/server topology. Note that in the client/server topology functionality is split into user processes and data processes, in which user processes handle interaction with the user and data processes handle interaction with data. In a Peer-to-Peer topology, one should expect to find both classes of processes placed on every machine. From a data logical perspective, client/server topology and Peer-to-Peer topology provide the same view of data - data distribution transparency. The distinction lies in the architectural paradigm that is used to realize this level of transparency.

2.5 Distributed Databases

Distributed databases are based on data distribution. It brings the advantages of distributed computing to the database management domain. A distributed system is a collection of processors, not necessarily homogeneous, interconnected by a computer network. Data distribution is an effort to improve performance by reducing communication costs and to improve quality of service in case of network failure. Based on our taxonomy, a distributed database system has the following characteristics: data is distributed (possibly replicated and/or fragmented) stored in locations close to the application domain(s) that uses it (e.g. increased availability), processors do not share resources (i.e., disks and memory) and processes are more distinct, and the underlying platform is possibly parallel. In comparison to parallel systems in which processors are tightly coupled and constitute a single database system, a distributed data base system is a collection of loosely coupled systems that share no physical components.

How to distribute data in order to improve performance, reliability, and accessibility, and how to provide transparency are the key issues in the design of distributed databases. Table 1 enumerates some important issues that one needs to consider in the design of distributed database systems.

Table 1. Issues in distributed database systems

ISSUES	REMARKS
Data Distribution	How data should be distributed/replicated/fragmented in order to improve performance/reliability/accessibility
Distribution Transparency	Distribution transparency includes network (location), replication, and fragmentation transparency
Keeping Track of Data	Keeping track of the data distribution, fragmentation, and Replication
Replicated Data Management	Which replica to access, how to maintain consistency, and how to control number of replica
Database Recovery	Recover from individual site crashes and system failures
Query Processing and Query Optimization	Query resolution, generation of sub-queries, and aggregation of partial results
Transaction Management	Concurrency control protocols
Security	Authenticate users, enforce authorization and access control, and auditing

2.6 Multidatabases

Adding control distribution to the definition of distributed databases as discussed in section 2.5 results in an environment with the following characteristics: