

Report on the 10th Conference on Database Systems for Business, Technology, and the Web (BTW 2003)

Gerhard Weikum
University of the Saarland
Saarbrücken, Germany
weikum@cs.uni-sb.de

Harald Schöning
Software AG
Darmstadt, Germany
Harald.Schoening@softwareag.com

Erhard Rahm
University of Leipzig
Leipzig, Germany
rahm@informatik.uni-leipzig.de

Introduction

The 10th Conference on Database Systems for Business, Technology, and the Web (BTW 2003 for short) took place from February 26-28, 2003 in Leipzig, a city in East Germany of great cultural heritage: it is the city of the late work of Johann Sebastian Bach, the location of Auerbachs Keller made famous by Goethe's Faust, the place where Napoleon was defeated in the Battle of Nations in 1813, and the origin of the German reunion starting with the Monday Demonstrations in 1989. The BTW 2003 conference was hosted by the University of Leipzig.

The bi-annual BTW conference is the major forum of the database communities in Germany, Austria, and Switzerland, and brought together more than 270 scientists and practitioners from this central European region and beyond. The program included 3 keynotes, 14 long and 17 short papers in the scientific track that were selected by the program committee from about 80 submissions, 11 papers in the industrial program, 2 contributions by the recipients of the BTW dissertation award, a special program for undergraduate students, 4 tutorials, and 2 co-located workshops. The conference was co-located with the 13th German Conference on Communications in Distributed Systems, bringing two research communities together.

The proceedings are published in the Lecture Notes in Informatics series of the German Computer Society. All papers, most of which are in English, and most presentations are

available online at <http://www.btw2003.de/proceedings/proceedings.en.shtml> and via DBLP at <http://www.informatik.uni-trier.de/~ley/db/conf/btw/btw2003.html>. The conference site <http://www.btw2003.de/> contains detailed material including photos of participants.

Keynotes

The first keynote, given by Frank Leymann from IBM Böblingen and entitled "Web Services: Distributed Applications without Limits", drew a fascinating picture of the brave new world of Web Services. Frank started by giving an overview of the virtual component model that underlies Web Services and the standardized protocol stack including SOAP, WSDL, BPEL4WS, and UDDI. He expressed his firm opinion that Web Services will be the very first standard for distributed applications that is universally accepted and will truly unify the IT industry, thus bringing dramatic benefits for interoperability, application development productivity, and ease of long-term maintenance. In particular, Web Services will boost our abilities to compose higher-level applications from existing components, both stateless as well as flow-oriented and long-lived. Also, Web Services will support a seamless integration with other component technologies such as J2EE or .Net via automatic wrapping and bridging.

In the second part of his talk Frank went beyond the current state of the art and sketched a roadmap for advanced broker services that will automatically match a pro-

grammer's or programmed component's request for a specific service with the available services and their quality-of-service properties (aka. policies). This matchmaking should be driven by mapping expected behavior based on a user's policy specification onto advertised behavior of existing services. The long-term vision is that the Web or the emerging Grid will provide a global service bus that can revolutionize our way of building distributed business-to-business applications and workflows. In addition to presenting this very inspiring vision, Frank also pointed out a number of technical issues for research opportunities.

The second keynote, given by Alon Halevy from the University of Washington and entitled "Data Integration: a Status Report", reviewed recent progress on this long-standing problem and pointed out major challenges that our research community should intensively tackle. He emphasized that we have gained a much better understanding of mediator architectures and query processing over multiple heterogeneous sources, with XML more and more serving as a common denominator. Alon also discussed the industrial impact of these recent advances, and included some nice anecdotes from his other life as a startup entrepreneur.

In the second part of his talk, Alon reconsidered some of the hardest challenges that remain unsolved in the area of data integration and pointed out some highly intriguing and promising research avenues. As one can easily guess, semantic schema reconciliation is one of these challenges. One approach that Alon outlined was to do schema integration in a peer-to-peer manner rather than in the traditional centralized mindset. So schema mappings exist only between pairs of interacting peers, and queries have to be forwarded and dynamically translated along their way in a P2P network. The rationale for this approach is that many small problems are easier to solve than one very big prob-

lem. A second, equally fascinating, direction that Alon pointed out is to apply machine learning techniques for schema reconciliation. Here Alon raised the key question what we could learn about schema matchings if we had available a huge corpus of existing database schemas (each one possibly with some characteristic instances). Obviously, he couldn't give a final answer to this question, but he certainly convinced the audience that this is a very important and fruitful research avenue, especially for doctoral students.

The third keynote, by Peter Lockemann from the University of Karlsruhe, addressed a rather non-fashionable but extremely important long-term problem: "Information System Architectures: from Art to Science". He first pointed out that there are very few good examples for clearly structured large-scale data management systems that are designed from clean principles. In view of the emerging problems of poor manageability and maintainability of the large software systems that our economy and society critically depend on, Peter urged both academic and industry people to consider system architecture as an absolutely top-priority item on our long-term research agenda. He then offered a framework for design guidelines and a high-level methodology for thinking about system components, their functionality and interactions, and also their quality of service guarantees.

Peter argued that the data model and the requirements of the target application classes, on one hand, and the available hardware technology, on the other hand, determine the system architecture to a large extent. As an example, the prevalent architecture of relational database engines, with a query optimizer and processor, a storage engine with sophisticated caching, and a transaction manager, has evolved the way it is because of performance having been the primary design goal, the declarative nature of queries, and the access gap between disk and

memory technologies. If we now were to build a new kind of XML engine with mostly navigational access and, perhaps, security as the primary concern of the applications, we would arrive at a fairly different architecture. To further underline the viability of his hypotheses, Peter also sketched architectures for hippocratic (i.e., privacy-centered) and incorruptible (i.e., consistency-centered) database engines that he systematically derived from his design guidelines. Peter's thought-provoking presentation, which was on the last day of the conference, triggered lively discussion in the audience, and gave all conference participants plenty of "thought food" for their trips home.

Scientific Program, Industrial Program, and Dissertation Awards

The papers and presentations in the scientific program covered a wide spectrum of topics including the usual ones like XML and Web Services, but also some that are rarely covered in database conferences such as data management in life sciences or combining information retrieval and database technologies and even a few talks on neglected but equally important problems off the beaten paths such as memory tuning for nested query execution. The industrial program featured the major vendors of the database system industry, but also several startups and database application perspectives from two big players of the insurance and the marketing service industries.

The three-day conference was concluded with the presentation of the BTW dissertation awards. This is a prize for the past two years' best German doctoral theses in the field of database and information systems; it is awarded by the German Computer Society (GI) and the winners are selected by the steering committee of the GI chapter on information systems. The prize comes with a cheque of 2500 Euros, donated by IBM

Germany. The winners for the 2001/2002 period were Reinhard Braumandl from the University of Passau for his thesis on "Quality of Service and Optimization in Data Integration Systems" and Jens Lechtenbörger from the University of Münster for his thesis on "Data Warehouse Schema Design". Both of these research projects have also been published in top-rate international conferences and journals. Before receiving their awards and cheques, Reinhard and Jens gave very nice short presentations on their excellent dissertations and thus provided a wonderful role model for the next generation of doctoral students who are now most motivated to strive for similar success.

Outlook

With almost 300 participants the BTW 2003 conference was one of the best attended in the 18 years of the BTW series, and an impressive success especially in the current recession of the IT industry. The next BTW conference, the 11th one in spring 2005, will celebrate the 20th anniversary of the BTW series, and will take place in Karlsruhe in the Southwest of Germany. The call for papers will appear in summer 2004, and the deadline for submissions will be in fall 2004.