Introduction

Salvatore Orlando, Gabriel Antoniu, Amol Ghoting, and Maria S. Perez

Topic chairs

The manipulation and handling of an ever increasing volume of data by current data-intensive applications require novel techniques for efficient data management. Despite recent advances in every aspect of data management (storage, access, querying, analysis, mining), future applications are expected to scale to even higher degrees, not only in terms of volumes of data handled but also in terms of users and resources, often making use of multiple, pre-existing autonomous, distributed or heterogeneous resources. The notion of parallelism and concurrent execution at all levels remains a key element in achieving scalability and managing efficiently such data-intensive applications, but the changing nature of the underlying environments requires new solutions to cope with such changes. In this context, this topic sought papers in all aspects of data management (including databases and data-intensive applications) that focus on some form of parallelism and concurrency. Each paper was reviewed by four reviewers and, after discussion, we were able to select four regular papers.

The accepted papers address relevant issues on various topics such as effective data compression, GPU-based data indexing, distributed collaborative data filtering and parallel query processing.

The paper entitled "Compressing the Incompressible with ISABELA: In-situ Reduction of Spatio-Temporal Data" by S. Lakshminarasimhan et al. proposes an effective method for In-situ Sort-And-B-spline Error-bounded Lossy Abatement (ISABELA) of scientific data that is widely regarded as effectively incompressible. ISABELA achieves an accurate fitting model that guarantees a ¿= 0.99 correlation with the original data and leverages temporal patterns in scientific data to compress data by 85%, while introducing only a negligible overhead on simulations in terms of runtime. The authors demonstrate that the proposed method outperforms existing lossy compression methods, such as Wavelet compression.

The second paper, entitled "kNN Query Processing in Metric Spaces using GPUs" by R. Barrientos addresses the idea of using GPUs to accelerate brute-force searching algorithms for metric-space databases. It shows how to improve existing GPU implementations and explores the viability of using GPUs in this context. The paper discusses the performance of both brute-force and indexing-based algorithms that take into account the intrinsic dimensionality of the elements of the database.

The third paper, entitled "An Evaluation of Fault-Tolerant Query Processing for Web Search Engines" by M. Marin et al. addresses strategies to perform parallel query processing in large scale Web search engines. The paper studies

E. Jeannot, R. Namyst, and J. Roman (Eds.): Euro-Par 2011, LNCS 6852, Part I, pp. 351–352, 2011.
© Springer-Verlag Berlin Heidelberg 2011

the suitability of such strategies for the case where processor replication is used to improve query throughput and to support fault-tolerance.

The forth paper, entitled "Performance Optimizations for Distributed Collaborative Filtering" by A. Narang et al. focuses on the usage of collaborative-filtering-based recommender systems by Internet-oriented companies for automatic predictions about user interests: the idea is to infer data from information about like-minded users. The paper presents a distributed algorithm that uses collaborative filtering for soft real-time distributed co-clustering. The proposed algorithm is optimized for multi-core cluster architectures.

We take this opportunity to thank the authors who submitted a contribution, the Euro-Par Organizing Committee, as well as the referees whose relevant comments and efforts substantially contributed to the effectiveness of the evaluation process and to the quality of the resulted program for this topic.