

EDITORIAL

Scalable optimization in grid, cloud, and intelligent network computing – foreword

Global optimization in large-scale distributed systems requires massive amounts of computations for complex objective functions. Conventional global optimization based on stochastic algorithms cannot guarantee an actual global optimum with a finite searching iteration. Therefore, scalability is a desirable feature for the optimization techniques in highly distributed dynamic environments, where the storage and computing capabilities can be spread over a wide geographical area. They must dynamically adapt to organizational relationships and real-world uncertainties.

Intelligent Networks, such as grids, peer-to-peer, ad hoc networks, constellations, and clouds enable the flexible routing and charging, advanced user interactions and the aggregation and sharing of geographically distributed resources. Collectively owned and managed by distinct organizational bodies, such complex large-scale distributed systems typically encompass computational resources from different institutions, enterprises, and individuals and are governed by heterogeneous administrative policies and regulations. System management techniques must therefore be able to group, predict, and classify different sets of rules, configuration directives, and environmental conditions to impose dissimilar usage policies on various users and resources. They must effectively deal with various optimization criteria, users' requirements, massive data processing, and, finally, uncertainties in system information that may be incomplete, imprecise, and fragmentary. Next information technology architectures, such as green cloud-to-cloud systems and green mobile clouds, provide elastic and in fact unlimited resources, including storage, as various services to cloud users with possible minimal energy utilization. However, both cloud users and cloud service providers are almost certain to be from different trust domains. Therefore, a secure user-enforced data access control mechanism must be provided before cloud users have the liberty to outsource sensitive data to the cloud for storage and further processing.

With the advent of intelligent networks, where efficient interdomain operation and high scalability of the whole system are the most important features, it is arguably required to investigate novel methods and techniques to enable secure access to data and resources, flexible communication, efficient scheduling, self-adaptation, decentralization, and self-organization.

This special issue herewith presents six research papers with novel concepts in the analysis, implementation, and evaluation of the next generation of intelligent scalable techniques for data-intensive processing and global optimization problems in large-scale distributed systems.

The first three papers discuss novel scalable solutions of data-intensive global optimization problems in well-known large-scale network environments. The presented techniques and their implementations are based on formal mathematical and logical models with the new optimization criteria (energy conservation), semantic rules and ontology, and modern synchronization modules of parallel computational processes. Li *et al.* in [1] introduced a methodology for improvement of the performance of the dynamic core of Global/Regional Assimilation and Prediction System (GRAPES) – the Numerical Weather Prediction system used by Chinese Meteorology Administration. The system performance is formally modeled as a sequence of large, sparse linear systems formulated by the discretization of global 3D Helmholtz equation. The authors developed a solver that enables an effective synchronization of the numerical processes at the global units of the system. The results of simple empirical analysis show good scalability of the proposed methodology achieved by using up to 6144 active cores in GRAPES. In [2], the authors present a framework for the energy-aware system management in backbone networks. The energy optimization problem is formulated as a

general mathematical programming problem with various constraints and control parameters. Dynamic voltage and frequency scaling method is implemented for minimizing the energy utilization at global and local levels of the management system along with a wide range of the resolution methodologies. All possible energy saving decisions of the system units are directly specified, together with decisions concerning traffic assignment to particular links. The results of the experiments show the best performance of the system in the case of concentration of the network traffic on a minimal subset of network components. The problem of massive processing of huge volumes of data in the Internet is discussed in [3]. Dong and Hussein propose an ontology-based Web crawler and Web page classifier with an embedded semisupervised learning module. This module enables the continuous enrichment of the definitions of ontological concepts in crawling and Web page classification process. The semantic relevance of crawling topics and Web pages is specified by semantic similarity and probabilistic models.

The remaining three papers address the big-data paradigm from various perspective. Bilal *et al.* [4] benchmark some well-known data center network architectures and categorically state their pros and cons. With this knowledge, the authors propose future advancements pertaining to the network architecture of data centers. In [5], a generic data-structure oriented programming template is discussed for supporting massive remote sensing data. The authors have built the case that the templates provide distributed abstractions for large remote sensing image data with complex data structures. The performance of their technique is improved by developing efficient parallel input/output (I/O) directly to and from the distributed data structures. Zhang *et al.* [6] have discussed an advanced data center architecture that harness the power of multiple data centers. The key technology that they advocate to manage such a large-scale distributed computing system is to build on both groups of distributed data centers/clusters that are equipped with data center or cluster resource manager. Additional security and access control procedures are put in place to provide a seamless interaction between various domains. In addition to a structure, the domain data centers are organized as collaborative modules, which enables processing of workflow workloads.

We believe that all of the papers presented in this Special Issue ought to serve as a reference for students, researchers, and industry practitioners interested or currently working in the evolving and interdisciplinary area of scalable computing and intelligent networking. We hope that the readers will find new inspiration for their research.

We are grateful to all the contributors of this issue. We thank the authors for their time and efforts in the presentation of their recent research results. We also would like to express our sincere thanks to the reviewers, who have helped us to ensure the quality of this publication. Our special thanks go to Prof Geoffrey C. Fox (Editor-in-Chief) and all of the editorial and management team of *Concurrency and Computation: Practice and Experience* Wiley journal for their great support throughout the entire publication process.

REFERENCES

1. Li L, Xue W, Ranjan R, Jin Z. A scalable Helmholtz solver in GRAPES over large scale multi-core cluster. DOI: 10.1002/cpe.2979.
2. Niewiadomska-Szynkiewicz E, Sikora A, Arabas P, Kołodziej J. Control system for reducing energy consumption in backbone computer network. DOI: 10.1002/cpe.2964.
3. Dong H, Hussein FK. SOF: a semi-supervised ontology-learning-based focused crawler. DOI: 10.1002/cpe.2980.
4. Bilal K, Khan SU, Zhang L, Li H, Hayat K, Madani SA, Min-Allah N, Wang L, Chen D, Iqbal M, Xu C-Z, Zomaya AY. Quantitative comparisons of the state of the art data center architectures. DOI: 10.1002/cpe.2963.
5. Ma Y, Wang L, Liu D, Yuan T, Liu P, Zhang W. Distributed data structure templates for data-intensive remote sensing applications. DOI: 10.1002/cpe.2965.
6. Zhang W, Wang L, Song W, Ma Y, Liu D, Liu P, Chen D. Towards building a multi-datacenter infrastructure for massive remote sensing image processing. DOI: 10.1002/cpe.2966.

JOANNA KOŁODZIEJ

Institute of Computer Science, Cracow University of Technology
ul. Warszawska 24, 31-155 Cracow, Poland
 E-mail: jokolodziej@pk.edu.pl

EDITORIAL

SAMEE ULLAH KHAN

*Department of Electrical & Computer Engineering, North Dakota State University
Fargo, USA*

EROL GELENBE

*Dept. Electrical & Electronic Engineering, Imperial College
London SW7 2BT, UK*

EI-GHAZALI TALBI

*LIFL - UMR LILLE 1/CNRS 8022, Cité Scientifique - Bâtiment M3
59655 Villeneuve d'Ascq Cédex - FRANCE*