Soft Comput (2017) 21:4497–4499
DOI 10.1007/s00500-017-2739-8

EDITORIAL

AutoCompBD: Autonomic Computing and Big Data platforms

Florin Pop[1,2] · **Ciprian Dobre**[1,2] · **Alexandru Costan**[3]

Published online: 26 July 2017

Abstract The amount of data collected or generated by ICT systems is growing exponentially (today we reached a Petabyte Era and will soon enter the ExaScale one). Processing and storing ever-larger volumes of data introduces new challenges, and consequently, we need to constantly develop new technological means to face them. Massive parallel processing platforms are the answer and are already being developed over distributed systems (i.e., over cloud or fog computing). However, the problem is that such platforms need to support a wide variety of applications, coming with different processing requirements. Thus, self-* behavior is a must in this context, referring to self-managing characteristics of distributed computing resources, their capability to adapt to unpredictable changes while hiding intrinsic complexity to operators and users. This special issue is dedicated to dissemination and evaluation of advances in Autonomic Computing and Big Data platforms, supported by large-scale distributed systems (LSDS). Autonomic Computing is facilitated by self-management capabilities that modern LSDS introduce, such as self-configuration, self-healing, self-optimization, and self-protection properties. In LSDS, an important characteristic is dependability (defined in terms of reliability, availability, safety and security of the operating system). Increased dependability means the system has to be able to detect, recover, and tolerate every possible deviation from its normal operation, and a wide area of Autonomic Computing research is today dedicated to this subject. The models used in the development of systems with dependability capabilities combine monitoring, scheduling, data management, security, and fault tolerance. The challenge is that in Big Data platforms applications and users, and even the distributed resources themselves, introduce unpredictable dynamic behavior. Autonomic Computing is considered one great challenge today faced by the IT industry, in need of finding good answers to how to conquer the growing complexity of large-scale systems and how to adequately cope with the many issues faced by truly Big Data processing. All these topics challenge today researchers, due to the strong dynamic behavior of the user communities and of resource collections they use. The special issue is oriented on computer and information advances aiming to develop and optimize advanced system software, networking, and data management components to cope with Big Data processing and the introduction of Autonomic Computing capabilities for the supporting large-scale platforms. We consider that our special issue comes with new and novel added value in the domain of Autonomic Computing and Big Data platforms.

Keywords Scheduling algorithms · Resource management · Fault tolerance · Big Data · Heterogeneous distributed systems

✉ Florin Pop
florin.pop@cs.pub.ro; florin.pop@ici.ro

Ciprian Dobre
ciprian.dobre@cs.pub.ro; ciprian.dobre@ici.ro

Alexandru Costan
alexandru.costan@irisa.fr

[1] Computer Science Department, Faculty of Automatic Control and Computers, University Politehnica of Bucharest, Bucharest, Romania

[2] National Institute for Research and Development in Informatics, Bucharest, Romania

[3] IRISA/INSA Rennes, Rennes, France

1 Introduction

In the last decade, Big Data analytics has become an indispensable tool in transforming science, engineering, health

care, finance, and ultimately business itself, due to the unprecedented ability to extract new knowledge and automatically find correlations in massive datasets that naturally accumulate in our digital age. Several Big Data processing frameworks emerged facilitating the development of multi-step data pipelines using directly acyclic graph patterns. Intended for very large scales, they have to address a series of challenges, such as a scalable architecture, data location transparency, high throughput under concurrent access, and the storage of massive data with fine-grained access. Although these requirements are the prerequisites for any efficient data processing system, they also imply a high degree of complexity in the configuration and tuning of the system, with possible repercussions on the systems availability and reliability. Making the most out of these frameworks is challenging because efficient executions strongly rely on complex parameter configurations and on an in-depth understanding of the underlying architectural choices.

Such challenges can be overcome if the systems are outfitted with a set of self-management mechanisms that enable autonomic behavior, which can shift the burden of understanding and managing the system state from the human administrator to an automatic decision-making engine. A large-scale Big Data platform is a complex system that has to deal with changing rates of concurrent users, the management of huge data spread across hundreds of nodes or with malicious attempts to access or to damage stored data. Therefore, such systems could largely benefit from a self-adaptation component that enables autonomic behavior by adapting to unpredictable changes while hiding intrinsic complexity to operators and users.

However, enabling autonomic behavior for Big Data processing is not a trivial task. Self-adaptation is impossible without deep and specific knowledge of the state of both the system and the infrastructure where the system is running on. The challenge in this context is raised by the unpredictable dynamic behavior of users, applications, and resources. Moreover, such platforms need to support a wide variety of applications, coming with different processing requirements adding heterogeneity challenges on top of the scale ones.

2 This special issue

This special issue is dedicated to the dissemination and evaluation of these specific advances in Autonomic Computing for Big Data platforms. An important focus is on dependability (defined in terms of reliability, availability, safety and security of the operating system). Increased dependability means the system has to be able to detect, recover, and tolerate every possible deviation from its normal operation, and a wide area of Autonomic Computing research is today dedicated to this subject. The models used in the development of systems with such capabilities combine monitoring, scheduling, data management, security, and fault tolerance (Iordache et al. 2006). To this end, the special issue is oriented on computer and information advances aiming to develop and optimize system software, networking, and data management components to cope with Big Data processing and the introduction of Autonomic Computing capabilities for the supporting large-scale platforms.

The special issues follows the previous published work like AMSBA (advances in modeling and simulation for big data applications) (Florin et al. 2016), ARMCO (advanced topics in resource management for ubiquitous cloud computing: an adaptive approach) (Florin and Maria 2016), and MidHDC (advanced topics on middleware services for heterogeneous distributed computing) part 1 and part 2 (Florin et al. 2016, 2017).

This special issue contains the following papers, described based on the main ideas presented in their abstracts.

- Autonomic deployment decision making for big data analytics applications in the cloud (Lu et al. 2015). The paper proposes a deployment decision-making solution for BDA applications in the cloud as follows: a novel language, named DepPolicy, to specify runtime deployment information as policies; a model for the deployment decision-making problem as a constraint programming problem using MiniZinc; a decision-making algorithm that can make different deployment decisions for different jobs in a way that maximizes overall utility while satisfying all given constraints (e.g., cost limit); the design and implementation of a decision-making middleware, named DepWare, for BDA application deployment in the cloud.
- Sensing service architecture for smart cities using social network platforms (Chifor et al. 2016). The paper presents a security architecture where social networks enable an adaptive sensing-as-a-service system. Smart city objects, which need additional information beside the one obtained from the sensors, collect it from the most trusted nodes on the social network.
- An evaluation of cloud-based mobile services with limited capacity: a linear approach (Skourletopoulos et al. 2016). The paper evaluates different cloud-supported mobile services subject to limited capacity, as the selection of a service may introduce additional costs, such as those that derive from the additional amount of memory required for processing.
- Analysis of power consumption in heterogeneous virtual machine environments (Negru et al. 2016). This paper studies the efficiency of data-intensive applications, and the penalties, in terms of power consumption, that are

introduced by different degrees of heterogeneity of the virtual machines characteristics in a cluster.

- A simple model to exploit reliable algorithms in cloud federations (Rubio-Montero et al. 2016). The paper presets a simple model together with a methodology to couple scheduling software with GWpilot, a framework that allows the personalized characterization of cloud resources that those algorithms require, overcoming their lack of trustworthiness in the information provided by the cloud services.
- Cloud services composition through cloud patterns: a semantic-based approach (Di Martino et al. 2016). This paper presents a methodology for the discovery and composition of cloud services, guided by cloud patterns.
- ATAC4Cloud: a framework for modeling and simulating autonomic cloud (Chainbi et al. 2016). The paper presents a cloud simulator supporting autonomic behaviors and integrating a workload generator that builds benchmarks to test the cloud infrastructure. The underpinning of this work is the synergy existing between agent technology and Autonomic Computing to develop self-adaptive cloud systems. ATAC4Cloud is developed as an extension of CloudSim.

Acknowledgements The guest editors would like to thank all authors for their valuable contributions and efficient cooperation in the preparation of this special issue. We appreciate the efforts all referees for to provide timely, constructive, thoughtful, and knowledgeable reviews and for their substantial contributions to the final revised versions appearing in this special issue. We express our gratitude and thank to them for their hard work. We address our personal warm regards to Prof. Antonio Di Nola, the Editor-in-Chief of Soft Computing Journal and to Aniello Castiglione (managing editor) for their support and advises, as well as to the editorial and managerial journal team from Springer for their assistance and excellent cooperative collaboration to produce this valuable scientific work. The authors would like to acknowledge the contribution of the following projects: DataWay (PN-II-RU-TE-2014-4-2731) and MobiWay (PN-II-PT-PCCA-2013-4-0321).

Compliance with ethical standards

Conflict of interest All authors declare that they have no conflict of interest.

References

Chainbi W, Chihi H, Azaiez M (2016) ATAC4CLOUD: a framework for modeling and simulating autonomic cloud. Soft Comput 1–12. doi:10.1007/s00500-016-2451-0

Chifor B-C, Bica I, Patriciu V-V (2016) Sensing service architecture for smart cities using social network platforms. Soft Comput 1–10. doi:10.1007/s00500-016-2053-x

Di Martino B, Cretella G, Esposito A (2016) Cloud services composition through cloud patterns: a semantic-based approach. Soft Comput 1–14. doi:10.1007/s00500-016-2264-1

Iordache GV, Boboila MS, Pop F, Stratan C, Cristea V (2006) A decentralized strategy for genetic scheduling in heterogeneous environments. In: OTM confederated international conferences "On the move to meaningful internet systems". Springer, pp 1234–1251

Lu Q, Li Z, Zhang W, Yang LT (2015) Autonomic deployment decision making for big data analytics applications in the cloud. Soft Comput 1–12. doi:10.1007/s00500-015-1945-5

Negru C, Mocanu M, Cristea V, Sotiriadis S, Bessis N (2016) Analysis of power consumption in heterogeneous virtual machine environments. Soft Comput 1–12. doi:10.1007/s00500-016-2129-7

Pop F, Iacono M, Gribaudo M, Kołodziej J (2016) Advances in modelling and simulation for big-data applications (AMSBA). Concurr Comput Pract Exp 28(2):291–293

Pop F, Potop-Butucaru M (2016) ARMCO:advanced topics in resource management for ubiquitous cloud computing: an adaptive approach. Future Gener Comput Syst 54:79–81

Pop F, Zhu X, Yang LT (2016) MIDHDC: advanced topics on middleware services for heterogeneous distributed computing. Part 1. Future Gener Comput Syst 56:734–735

Pop F, Zhu X, Yang LT (2017) MIDHDC: advanced topics on middleware services for heterogeneous distributed computing. Part 2. Future Gener Comput Syst 74:86–89

Rubio-Montero AJ, Rodríguez-Pascual MA, Mayo-García R (2016) A simple model to exploit reliable algorithms in cloud federations. Soft Comput 1–13. doi:10.1007/s00500-016-2143-9

Skourletopoulos G, Mavromoustakis CX, Mastorakis G, Batalla JM, Sahalos JN (2016) An evaluation of cloud-based mobile services with limited capacity: a linear approach. Soft Comput 1–8. doi:10.1007/s00500-016-2083-4