Future Generation Computer Systems 87 (2018) 416–419

Contents lists available at ScienceDirect

Future Generation Computer Systems

journal homepage: www.elsevier.com/locate/fgcs

Editorial

Emerging trends, issues and challenges in Internet of Things, Big Data and cloud computing

Anna Kobusińska [a,*], Carson Leung [b], Ching-Hsien Hsu [c], Raghavendra S. [d], Victor Chang [e]

[a] Institute of Computing Science, Poznań University of Technology, Poznań, Poland
[b] University of Manitoba, Winnipeg, MB, Canada
[c] Chung Hua University, Hsinchu, Taiwan
[d] University Visvesvaraya College of Engineering, Bangalore, Karnataka, India
[e] International Business School Suzhou and Research Institute of Big Data Analytics, Xi'an Jiaotong-Liverpool University, Suzhou, China

ARTICLE INFO

Keywords:
Big Data
Cloud computing
Internet of Things
Data analysis
Cloud computing platforms
Applications and management

ABSTRACT

Although Big Data, IoT and cloud computing are three distinct approaches that have evolved independently, they are becoming more and more interconnected over time. The convergence of IoT, Big Data and clouds provides new opportunities and results in development of new applications in many fields, including business, healthcare, sciences and engineering. At the same time, various challenges are faced during processing and management of massive amounts of data, as well as during their storage in cloud environments. This special issue presents novel research approaches related to Big Data, IOT and cloud computing. It also discusses the encountered problems and open issues.

1. Introduction

Cloud computing has emerged as an important computing paradigm, enabling ubiquitous convenient on-demand access through Internet to a shared pool of configurable computing resources [1,2]. In this paradigm, software (applications, databases, or other data), infrastructure and computing platforms are widely used as services for data storage, management and processing. They provide a number of benefits, including reduced IT costs, flexibility, as well as space and time complexity. To benefit, however, from numerous promises that cloud computing offers, many issues have to be resolved, including architectural solutions, performance optimization, resource virtualization, providing reliability and security, ensuring privacy, etc [3–5].

Another significant technology trend that nowadays is gaining increasing attention is Internet of Things (IoT) [6,7]. In IoT, intelligent and self conguring embedded devices and sensors are interconnected in a dynamic and global network infrastructure, enabling scalability, flexibility, agility and ubiquity in fields of massive scale multimedia data processing, storage, access and communications. IoT is driving new interest in Big Data [8–10], by generation of enormous amount of new types of data being generated by sensors and other input devices, which have to be stored, processed and accessed. The need to monitor, analyze and act upon these data brings many issues like data confidentiality, data verification, authorization, data mining, secure communication and computation.

The future development of cloud computing systems is more and more influenced by Big Data and IoT [11,12]. There are research and industrial works showing applications, services, experiments and simulations in Clouds that support the cases related to IoT and Big Data [13–15]. Provision of above issues presents a new set of emerging problems and challenges that are expected to be identified and addressed. Therefore, the aim of this special issue is to present and discuss novel ideas and research outcomes on all aspects of Big Data, Internet of Things and cloud computing, as well as to identify new research topics. In particular, this special issue aims to examine the prospects and challenges that arise during the conjunction of the modern cloud applications with the field of Internet of Things and Big Data. Promoting the submission of the ongoing work with the existing important theoretical and practical results, along with position papers and case studies of already present verification projects, this special issue will highlight the art in this domain. As one of the goals, this special issue intends also to convene researchers and practitioners to review the diverse range of features of security, privacy, trust and reliability in IoT and Cloud. It also examines significant theories, scrutinies technology enablers, formulates significant application and devise new methods to overcome the major problems that this research area poses.

2. Brief review of special issue content

The special issue consists of invited top conference papers from SC2-2016 and IoTDBS 2017 conferences, as well as papers from the

* Corresponding editor.
E-mail address: Anna.Kobusinska@cs.put.poznan.pl (A. Kobusińska).

https://doi.org/10.1016/j.future.2018.05.021
0167-739X/

open call. In the response to the call for papers, 128 high-quality manuscripts submitted by various authors, and encompassing the most varied topics within the scope of the special issue were received. All manuscripts underwent a rigorous peer review process to ensure they meet the standards and quality of the FGCS journal. A throughout analysis of the research contributions of submitted papers, and the compliance with the scope and relevance to the special issue was conducted. Also, the innovations of manuscripts have been considered carefully to finalize the selection process. Eventually, we were pleased to select 15 high-quality manuscripts, which focus on advancing the latest research on the topic of Internet of Things, Big Data and cloud computing. The acceptance rate of the papers was 11.7%.

In the first paper, the real-time monitoring and profiling of big data applications and frameworks was considered. Many big data applications and frameworks involve heavy use of system resources. To improve performance of these applications and frameworks, Enes et al. in [16] presented a new framework, called BDWatchdog, to monitor system resources and profile big data applications in real time. BDWatchdog provides a process-based accurate analysis to visually characterize the performance of both big data workloads and frameworks by combining time series for resource monitoring and Java Virtual Machine (JVM) flame graphs for source code profiling. The proposed framework is architected in such a way that it allows scalable analysis of big data-scale applications across clusters, spotting of resource and code bottlenecks, as well as characterization of big data applications.

The second paper [17], considers resource allocation and task scheduling in cloudlet systems. The paper is motivated by the observation that the ability to provide cloud service is critical for the emerging mobile cloud computing (MCC) systems. Although the cloudlet possesses adequate resources to simultaneously process multiple mobile requests, it is not as sufficient as a remote cloud data center. Zhang et al. introduced a load-aware resource allocation and task scheduling (LA-RATS) strategy for cloudlet-based MCC systems. The proposed solution enhances the quality of mobile cloud service by adaptively allocating resource in a MCC system for delay-tolerant and delay-sensitive mobile applications based on the dynamic behavior and load profile of cloudlets. Specifically, it applies heuristic and meta-heuristic technology for scheduling tasks and allocating resources in cloudlet under normal load. It migrates tasks of delay-tolerant applications under cloudlet overload. To raise the utilization of the cloudlet, Zhang et al. also proposed a tree generation-based task backfilling mechanism to enable full use of the idle resource via a backward shifting strategy and to avoid unnecessary queue growth for virtual machines.

The following two papers of this special issue are devoted to virtualization problem. Network virtualization enables to group multiple physical networks into one virtual network, or separate a single physical network into multiple logical networks. Among different virtual network approaches, the Hybrid Virtual Networks (HVN) gain a lot of attention. G. Sun et al. in [18], investigated how to provide Hybrid Virtual Networks support both multicast and unicast traffic. It is a meaningful topic since network virtualization solves the problem of current network structure and multicast coexist with unicast traffic in the network. The authors proposed both ILP and heuristic solutions, a relatively complete simulation results are given.

Also, the interconnected cloud approach attracts a lot attention. Interconnected clouds enable to host computing resources, and share the load in private clouds, providing a high-performance, and cost-efficiency at the same time. But, they also have to deal with the problem that different cloud providers have different kinds of Virtual Machine (VM) instance types, pricing models, and management interfaces. Chih-Tien Fan et al. proposed a job management system for federated clouds that exploits agent technology to interconnect different cloud environments [19]. In the paper [19], the proposed solution VM instances are selected based on the job's deadline constraint. The authors proposed to use the rough set theory to predict job length based on historical data, such that the scheduler has the knowledge to see whether the system can meet the deadline constraint of a submitted job. Since this paper does not assume the job execution time is known, the proposed approach is thus novel and practical for job management on federated clouds. The experimental result show that the proposed approach works well for various types of jobs.

Further, this special issue tackles the problem of security and reliability of cloud environments. The chosen papers introduce different approaches and methods to provide those features. First, in [20], the secure policy execution is examined. The policy-based management is an important issue in the collaborative clouds that allow multiple users from different domains to access and share files. To carry out or enforce the policy decisions, a Cloud Policy Decision Point (CPDP) service is currently often used. But, this process reveals information about the policies to the third-party clouds. M. Alam et al. in [20], proposed using reusable garbled circuits (RGC) to protect the policy information from being revealed. The proposed protocol, privacy aware cross tenant access control (PaCTAC), contains three phases, the policy generation phase, and an attribute generation phase, and then a policy evaluation phase. The protocol makes use of the ABE RGC scheme. By using PaCTAC, it is possible to prevent a CPDP from learning about the policy in transactions, which take place in the CPDP.

In turn, increasing the reliability of cloud environments is discussed in [21]. The paper focuses on the cloud resources and cloud management capabilities, which are exposed via RESTful services. The reliability of such solutions plays an important role. The paper describes a new idea for increasing reliability of RESTful web services in clouds environment. Since the simple restart of a failed processing (as a new instance) from the very beginning, is usually insufficient and often leads to unacceptable inconsistency of processing state, the authors of [21] proposed the system mechanism that copes with this problem. The paper contributes with three main issues: RESTful recovery consistency model, application of the model in ReServE service and formal and empirical evaluation of the proposed recovery protocol. The introduced formal definition of RESTful recovery consistent model may be a formal basis for any forthcoming research.

The succeeding paper [22], discusses the source location privacy issue. There are many security-critical wireless sensor network applications in the current era of big data. Hence, source location privacy (SLP) has become important. Gu et al., in the considered paper proposed a viable 2-step framework for SLP-aware routing protocol selection. It uses the existing library of performance profiles of various routing algorithms and decision theoretic heuristics to access trade-offs. With this framework, the protocols are first profiled and filtered by capturing their performance under various protocol configurations and metric sets. Then, decision theoretic heuristics are applied to characterize and select the most appropriate SLP routing protocols for different applications in wireless sensor networks by removing dominated protocols and formalizing the notion of relevance with suitable weights and utility functions of attributes.

Finally, among proposed solutions that increase security and reliability, the erasure coding is considered in [23]. Erasure coding is one of the techniques offering data protection of large scale distributed cloud storage systems. This technique is used to store and retrieve data in the cloud storage. Existing solutions, such as striping encoding and replicating encoding, either provide a poor read/write performance or generate high network and I/O overhead. The model proposed by F. Xu et al., called incremental encoding, is a solution that provides a good performance similar to the replicating encoding, while achieves a low overhead as in the