Guest Editorial: Cloud Services Meet Big Data

Keqing He, *Senior Member, IEEE*, Liang-Jie Zhang, *Fellow, IEEE*, Schahram Dustdar, *Fellow, IEEE*, and Yutao Ma, *Member, IEEE*

THE concept of Cloud Service represents a prime facility and feature of services in a cloud computing environment that can be made available to users on demand. Due to the flexibility of cloud computing in scaling IT resources up and down, cloud services gradually become valuable to attract the gaze of researchers and engineers from both academia and industry when they are faced with dynamically changing business requirements. Different stakeholders, such as consumers, providers, and operators, are generating a vast amount of data on such services per minute on the Internet, which increasingly comes to show the "4V" characteristics of big data. Therefore, new methodologies and techniques are urgently required for designing, validating, developing, testing, and deploying cloud services on demand in this specific scenario based on big data, as well as for efficiently being adaptive to business dynamics and users' explicit and implicit requirements.

The objective of this Special Issue (SI) on Cloud Services Meet Big Data is to solicit innovative and promising methods and techniques related closely to cloud services in the era of Big Data. We expect that the issue can promote the visibility and relevance of this noteworthy direction in the interdisciplinary field between Services Computing and other emerging disciplines. As a result, the call of this SI received forty-nine submissions from more than ten countries and regions. Every submission underwent a thorough screening process that contains at least two rounds of peer review by at least three reviewers. Finally, we selected seventeen high-quality manuscripts and organized them into two separate parts.

The first part of this SI focuses on three hot research topics, namely service selection and composition, service security and privacy, and service operation and monitoring, including ten regular papers that address the most challenging issues in these fields from the perspectives of the theoretical research and applications. A brief introduction to the ten articles is presented as follows.

The first four articles fall within the scope of service selection and composition. The authors of the four papers attempted to investigate this topic, one of the fundamental issues in the field of Services Computing, to make cloud services efficient in processing big data.

The first article, entitled "Policy-Aware Service Composition: Predicting Parallel Execution Performance of Composite Services," by Mai Xuan Trang, Yohei Murakami, and Toru Ishida (an IEEE Fellow and an IEICE Fellow), proposes a model embedding service policies into formulae to calculate the performance of composite services under parallel execution. According to the output of the model, the optimal degree of parallelism (DOP) for a given composite service can be predicted adequately. The experimental results on real-world translation services indicate that the proposed model performs well in identifying the optimal DOPs for composite services.

In the second article entitled "Automated Testing of WS-BPEL Service Compositions: A Scenario-Oriented Approach," Chang-ai Sun, Yan Zhao, Lin Pan, Huai Liu, and Tsong Yueh Chen present a scenario-oriented testing approach that can generate test cases for WS-BPEL (the Web Services Business Process Execution Language) based service compositions automatically. Also, the authors develop a prototype tool implementing the proposed approach and demonstrate its applicability and effectiveness using an empirical study.

Chen Lv, Wei Jiang, Songlin Hu, Jiye Wang, Guoliang Lu, and Zhiyong Liu (a fellow of the CCF (China Computer Federation)) investigate the dynamic evolution problem of service compositions in their article entitled "Efficient Dynamic Evolution of Service Composition." The authors introduce the continuous query mechanism into QoS (Quality of Service)-aware automatic service composition and then propose an event-driven continuous query algorithm to cope with different types of dynamic Web services efficiently. The evaluation using both real-world QoS data and synthetic Web service data shows the advantage of the proposed algorithm over a state-of-the-art solution.

In "Effective BigData-Space Service Selection over Trust and Heterogeneous QoS Preferences," the final article on service selection and composition, Hongbing Wang, Chao Yu, Lei Wang, and Qi Yu attempt to tackle the problem of service selection in a heterogeneous preference- and trust-based big data space. They develop a novel multi-objective optimization method to make a trade-off decision between service trust and QoS preference, according to the degree of matching with user requirements. Besides, the authors experiment with the QWS dataset and synthetic

- *K. He and Y. Ma are with the School of Computer Science, Wuhan University, Wuhan 430072, China. E-mail: {hekeqing, ytma}@whu.edu.cn.*
- *L.-J. Zhang is with the Kingdee International Software Group Company Limited, Shenzhen 518057, China. E-mail: zhanglj@kingdee.com.*
- *S. Dustdar is with the Distributed Systems Group, TU Wien, Wien A-1040, Austria. E-mail: dustdar@dsg.tuwien.ac.at.*

For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org, and reference the Digital Object Identifier below.
Digital Object Identifier no. 10.1109/TSC.2018.2837699

data to demonstrate the effectiveness of the proposed approach.

Two of the most common topics of questions regarding cloud services are security and privacy. The second three articles address specifically the issues related to service security and privacy from the viewpoints of both providers and users.

Lirim Osmani, Salman Toor, Miika Komu, et al. present a case study of the DII-HEP secure cloud infrastructure in "Secure Cloud Connectivity for Scientific Applications." The other authors of this paper include Matti J. Kortelainen, Tomas Linden, John White, Rasib Khan, Paula Eerola (a member of the Royal Swedish Academy of Sciences and a member of the Finnish Society of Sciences and Letters), and Sasu Tarkoma. To facilitate hybrid cloud scenarios for scientific applications, the authors propose an approach to scale out a private cloud deployment to public clouds securely. Also, they conduct a series of experiments based on the DII-HEP production cloud to test the performance and stability of their infrastructure.

In the second article on service security and privacy, "Data-Driven and Feedback-Enhanced Trust Computing Pattern for Large-Scale Multi-Cloud Collaborative Services," Xiaoyong Li, Huadong Ma (a fellow of the CCF), Wenbin Yao, and Xiaolin Gui investigate the trust computing problem of multi-cloud collaborative services. They propose and develop a Data-driven and Feedback-Enhanced Trust (DFET) computing pattern across multiple data centers. Besides the theoretical analysis that the DFET pattern is highly dependable against garnished and bad-mouthing attacks, the authors build a prototype system to demonstrate its feasibility.

The verification of the integrity and retrievability of remote data in cloud storage systems remains challenging for different cloud service providers. The third article, entitled "Dynamic Proofs of Retrievability for Coded Cloud Storage Systems," by Zhengwei Ren, Lina Wang, Qian Wang, and Mingdi Xu, proposes a dynamic proof of retrievability scheme that supports public auditability and communication-efficient recovery from data corruptions. The formal security analysis and extensive experimental evaluations indicate that the proposed scheme is practical for use in cloud storage systems.

The last three articles belong to one of the most popular topics in the field of Services Computing, service operation and monitoring, and they attempt to tackle the problems of reliability monitoring, workflow scheduling, and data replica placement related closely to cloud services.

In "A Data Set for User Request Trace-Oriented Monitoring and its Applications," Jingwen Zhou, Zhenbang Chen, Ji Wang, Zibin Zheng, and Michael R. Lyu (an IEEE Fellow and an ACM Fellow) release a publicly available dataset called TraceBench and the corresponding collection system called MTracer. TraceBench is a fine-grained dataset of user request traces collected from a real-world cloud storage service. Besides, the authors validate the usability and authenticity of TraceBench in several specific trace-oriented monitoring scenarios, such as anomaly detection, performance problem diagnosis, and temporal invariant mining.

Zhongjin Li, Jidong Ge, Haiyang Hu, Wei Song, Hao Hu, and Bin Luo attempt to tackle the problem of energy consumption in clouds in their article entitled "Cost and Energy Aware Scheduling Algorithm for Scientific Workflows with Deadline Constraint in Clouds." The authors propose a Cost- and Energy-Aware Scheduling (CEAS) algorithm to reduce the execution cost and energy consumption of workflows as much as possible while satisfying the deadline constraint. Also, they evaluate the CEAS algorithm using CloudSim and four real-world scientific workflow applications, suggesting that it outperforms the baseline approaches.

The last article of the first part of this SI, entitled "A Genetic Algorithm Based Data Replica Placement Strategy for Scientific Applications in Clouds," formulates the problem of data replica placement using a tripartite graph-based model. Lizhen Cui, Junhua Zhang, Lingxi Yue, Yuliang Shi, Hui Li, and Dong Yuan propose a genetic algorithm-based data replica placement strategy for scientific applications in clouds. Besides, the experimental results indicate that the proposed strategy performs better than the random placement strategy used in the Hadoop Distributed Files System (HDFS).

In short, we expect that the fascinating research contributions of the articles in this SI would offer potential insights into the further research of Services Computing and spark broad discussion in the emerging area of cloud services when meeting big data.

ACKNOWLEDGMENTS

We appreciate very much the valuable work of more than 120 reviewers from all over the world. Again, we are very grateful to all of them for submitting high-quality reviews in time. Besides, we wish to express our gratitude to the former Editor in Chief, Prof. Ling Liu, the Editor in Chief, Prof. James Joshi, the Associate Editor in Chief, Dr. Rong N. Chang, and the Administrator, Ms. Christine Kurzawa of *IEEE Transactions on Services Computing* for their constant and valuable help.

Keqing He received the PhD degree from Hokkaido University, Japan, in 1995. He is a professor with the School of Computer Science, Wuhan University, China. His research interests include services computing and software engineering. Until now, he has co-authored three research monographs (in Chinese) and more than 200 peer-reviewed papers (in English, Chinese, and Japanese) and published five ISO technical standards (ISO/IEC 19763-3,5,7,8,9) as well as eight Chinese national standards. Also, he received many domestic and international prizes and awards, such as the ISO/IEC Certificate of Appreciation for project editor and the second Prize of National Scientific and Technological Progress. He is a senior member of the IEEE and a distinguished member of the CCF, and he has been a founding vice chair of the CCF Technical Committee on Services Computing (TCSC) and an associate editor of *IEEE Transactions on Services Computing*.

Liang-Jie Zhang received the PhD degree on pattern recognition and intelligent control from Tsinghua University, in 1996. He is currently a senior vice president, CTO, and director of Research at Kingdee International Software Group Company Limited, where he leads innovations of services transformation and products' architectures. Prior to this position, he was a research staff member and program manager of application architectures and realization at IBM T.J. Watson Research Center and chief architect of Industry Standards at IBM Software Group, where he led the R&D of the custom solution engagement cloud and service-oriented architecture solutions. He has more than 50 granted patents in the areas of e-commerce, Web services, rich media, data management, cloud computing, artificial intelligence, and information appliances, and he has published more than 160 technical papers in journals, book chapters, and conference proceedings. He chaired the IEEE Computer Society's Technical Committee on Services Computing from 2005 to 2008. He has served as the editor-in-chief of the *International Journal of Web Services Research* and is the founding editor-in-chief of *IEEE Transactions on Services Computing*. He was elected as an IEEE fellow in 2011, and in the same year won the Technical Achievement Award "for pioneering contributions to Application Design Techniques in Services Computing" from the IEEE Computer Society.

Schahram Dustdar is full professor of Computer Science heading the Distributed Systems Group with the TU Wien, Austria. From 2004-2010 he was also honorary professor of information systems with the Department of Computing Science, University of Groningen (RuG), The Netherlands. From Dec 2016 until Jan 2017 he was a visiting professor with the University of Sevilla, Spain and from January until June 2017 he was a visiting professor with UC Berkeley. He is co-editor-in-chief of the new *ACM Transactions on the Internet of Things* as well as editor-in-chief of *Computing* (Springer). He is an associate editor of *IEEE Transactions on Services Computing*, *IEEE Transactions on Cloud Computing*, *ACM Transactions on the Web*, and *ACM Transactions on Internet Technology*, as well as on the editorial board of *IEEE Internet Computing* and *IEEE Computer*. Dustdar is the recipient of the ACM Distinguished Scientist award (2009), the IBM Faculty Award (2012), an elected member of the Academia Europaea: The Academy of Europe, where he is chairman of the Informatics Section, as well as an IEEE fellow (2016).

Yutao Ma received the PhD degree in computer science from Wuhan University, China, in 2007. He is now an associate professor with the School of Computer Science, Wuhan University. He was with the Institute of China Electronic System Engineering Corporation (Beijing) as a postdoctoral fellow and has been a visiting scholar in the Department of Electronic and Computer Engineering, Lehigh University, USA. His research focus is on the development and maintenance of large-scale software service systems. He has published more than fifty peer-reviewed papers and received two best paper awards at international conferences. In addition to a member of the ACM and the IEEE, he is now a senior member of the CCF and a member of the CCF TCSC.