Guest Editor's Introduction: Special Section on Virtualization and Services for Cloud-Based Application Systems

Yiming Zhang, Rong Chang, and Paul Townend

Abstract—Cloud-based application systems are rapidly deployed worldwide in production use via virtualization and services computing technologies. The scaling demands for these application capabilities to the cloud providers, compound with differentiated requirements on the quality of services, have brought severe technical challenges. This special section focuses on the techniques of virtualization and services for cloud-based application systems, mainly including multi-scale resource management and sharing, elastic scheduling and allocation of computing and network resources, monitoring and diagnosis for cloud-based services, and cloud-based mobile systems. The articles of this special section illustrate recent advances in virtualization and services provisioning for cloud-based application systems. We expect that this special section will provide an integrated view of the state-of-the-art techniques, identify new challenges as well as opportunities, and promote collaboration among researchers in this field. We received 34 submissions and we finally accepted 4 articles. The acceptance rate is as low as 11.8 percent.

Index Terms—Virtualization, services computing, cloud computing, resource management, elastic scheduling, monitoring and diagnosis, mobile systems

1 Introduction

Cloud based application systems are deployed worldwide in production use via virtualization and services computing technologies. These systems utilize a variety of computing and network resources deployed in one or more Internet-based data centers and provide cost-efficient, dynamically-changed and real-time quality-assured application capabilities. However, the scale and utility characteristics [1] of rapidly-growing Internet applications have brought severe challenges to both virtualization and services computing technologies.

The scale of Internet applications, such as the user scale, data scale and task scale, have made virtualization become increasingly important to cloud based application systems. For example, it is common for Internet services (such as Facebook [2], Alibaba [3] and Twitter [4]) to have billions of users, and the amount of data of many online data-intensive (OLDI) applications [5] have exceeded the peta scale. Virtualization is the only way to isolate user requests, share resources, and satisfy differentiated performance requirements of various end users.

Clouds provide users with convenient IaaS (Infrastructure as a service), PaaS (Platform as a service) and SaaS (Software as a Service). In an Internet-based virtual computing environment (IVCE) [6], where application capabilities are delivered by dynamically composing remote third-party services whose providers could be anonymous or unknown to the end users [7], the quality-of-service (QoS) guarantees of the application systems are not only important but also difficult to achieve [8], making service-oriented solutions become less attractive for a large variety of cloud-based application systems.

To address such new challenges, cloud-based application systems need technical innovations in a wide range of virtualization and services computing technologies, including service-oriented design and development, resource-aware elastic scalability of cloud services, modeling of intra- and inter-cloud resources at different scales in a consistent manner, high reliability, availability, durability and consistency of IVCE application systems, trustworthiness of anonymous IVCE services, and so on.

This special section focuses on the virtualization and services computing technologies of cloud-based application systems, mainly including multi-scale resource management and sharing, elastic scheduling and allocation of computing and network resources, monitoring and diagnosis for cloud-based services, and cloud-based mobile systems. The articles introduced in this special section illustrate recent advances in virtualization and services computing of cloud-based application systems. We expect that this special section will provide an integrated view of the state-of-the-art techniques, identify new challenges as well as opportunities, and promote collaboration among researchers with common interests in this field.

We received 34 submissions and finally 4 of them are accepted. The acceptance rate is as low as 11.8 percent. The review process included two rounds of reviews. In the

- *Y. Zhang is with the School of Computer, National University of Defense Technology, Changsha 410073, China. E-mail: ymzhang@nudt.edu.cn.*
- *R. Chang is with IBM Research, San Jose, CA 95120-6099*
- *P. Townend is with the School of Computing, University of Leeds, Leeds LS2 9JT, United Kingdom*

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.2018.2878591

first round, each article submitted to this section received at least three independent reviews. We advanced to the second round 8 articles according to the reviews of the first round. In the second round, we discussed at the online editorial meeting and chose 4 candidates based both on their qualities and on the relevance to this special section. The 4 articles are eventually accepted after they addressed all the concerns of both reviewers and guest editors.

2 Special Issue Papers

The selected articles highlight the theme of the special section with significant research breakthroughs. They respectively address four specific technical challenges for virtualization and services in cloud-based application systems: performance diagnosis and analysis, traffic management and scheduling, resource consolidation and allocation, and mobile system virtualization.

The first article, entitled "Straggler Root-Cause and Impact Analysis for Massive-scale Virtualized Cloud Datacenters" by Garraghan et al., focuses on the long tail problem and studies the frequency and impact that task stragglers impose on cloud datacenters. The article filters and characterizes straggler behaviors, and proposes detailed analysis of straggler's root-cause to ascertain an intrinsic understanding of stragglers within real systems. Based on these findings, this article proposes a method for straggler detection in large-scale distributed systems which successfully mitigates the effects of extreme stragglers. An empirical analysis of long tail identification in two large-scale production cloud datacenters comprising more than 12,000 servers shows the impact of stragglers caused by data skew. The authors get some interesting results such as the percentage of affected jobs by stragglers and by high resource utilization. The authors combine both offline analytics and online agent based monitoring/analytics in order to improve the timeliness of straggler identification. The approach is capable of identifying task stragglers caused by data skew rapidly and accurately at runtime, and can be integrated into state-of-the-art straggler mitigation techniques in order to enhance the temporal properties and timeliness of job execution. This article proposes an important problem, and presents reasonable solution and convincing results validated by large-scale trace-driven simulations.

The second article, entitled "Flow Classification for Software-Defined Data Centers Using Stream Mining" by Chao et al., studies how to identify long (elephant) flows which will be rerouted to improve network utilization in large-scale cloud systems. Traffic management is important to effectively utilize the high bandwidth provided by datacenters. Traditionally, identifying elephant flows requires either significant monitoring overhead or endhost modifications. This article proposes a stream-mining-based elephant flow detection and scheduling system called FlowSeer. The key idea is to use the flows' first few packets to train the streaming classification models, in order to predict the rate and duration of the flows. FlowSeer enables the controller and switches to perform cooperative prediction. The proposed method significantly reduces the monitoring overhead and hardware/end-host modifications. FlowSeer uses the controller and switches to perform cooperative prediction and most decisions can be made by individual switches locally, reducing both detection latency and signaling overhead. Analysis shows that FlowSeer requires less than 100 flow table entries at each switch to enable cooperative prediction, and hence can be implemented on off-the-shelf switches. The evaluation via both experiments in realistic virtual networks and trace-driven simulation shows that FlowSeer improves the throughput by multiple times over previous studies.

The third article, entitled "Cost-Efficient Consolidating Service for Aliyun's Cloud-Scale Computing" by Yan et al., describes the server consolidation service (C4) of the Aliyun cloud, which is critical for improving energy efficiency and resource utilization of its Internet-scale cloud computing. At the scale of Aliyun, server consolidation has to address several challenges. First, the migration time and cost of local-storage-based services is very high. Second, the resource utilization of migrating services changes over time which may bring risks of migration oscillation. Third, practical constraints (such as the strict maintenance time window) in Aliyun's production environment make server consolidation more complicated. These challenges are realistic, and the motivation is tightly coupled with the challenges. This article investigates a large-scale production Aliyun cloud, analyzes a large amounts of statistical data collected from the system, and draws some interesting conclusions, e.g., the maintenance window, and the migrating instance size distribution. This article further proposes the migration cost model to address these challenges by collecting statistics and analyzing the user patterns, utilization, and migration cost. The proposed model not only considers the traditional migration costs, but also accounts for the cost that is introduced by the utilization of local storage. The authors has implemented the C4 consolidating service, which has deployed and run at the Aliyun cloud for over a year.

The fourth article, entitled "HyperCo: Optimizing Network Performance in ARM-Based Mobile Virtualization" by Yao, et al., finds that in the ARM based mobile virtualized environment the context switch between the host and the guest triggered by hypercall brings significant CPU overload and thus degrades I/O performance under intensive workloads. The problem of optimizing both network throughput and latency for network-intensive applications is very important for ARM based mobile virtualized environment. To address the problem, this article proposes the HyperCo approach, a software-only method that optimizes network I/O performance through controlling the number of hypercalls and achieves a desirable tradeoff between throughput and latency. HyperCo reduces the number of hypercalls by using adaptive hypercall coalescing. HyperCo remarkably improves the network throughput under intensive workloads while keeping a low penalty by adapting coalescing interval dynamically at the low frequency of network I/O requests. The authors implement HyperCo by modifying the front-end driver of virtio-net on KVM/ARM platform.

Acknowledgments

This special section cannot succeed without the collective effort and support of many individuals and organizations. We would like to thank all the authors for their submissions, and all the reviewers for their great efforts to provide insightful and valuable reviews. We also want to thank the editorial board of *IEEE Transactions on Services*

Computing, the Editor-in-Chief, Prof. Ling Liu, and Miss Christine Kurzawa. In particular, we would like to thank Prof. Yuanyuan Yang, Dr. Huiba Li and Prof. Kai Chen for their effort and contribution to this special section.

REFERENCES

[1] Y. Zhang, L. Chen, X. Lu, and D. Li, "Enabling routing control in a dht," *IEEE J. Select. Areas Commun.*, vol. 28, no. 1, pp. 28–38, Jan. 2010.

[2] D. Beaver, S. Kumar, H. C. Li, J. Sobel, P. Vajgel, et al., "Finding a needle in haystack: Facebook's photo storage," in *Proc. 9th USENIX Symp. Operating Syst. Des. Implementation*, 2010, vol. 10, pp. 1–8.

[3] Z. Zhang, C. Li, Y. Tao, R. Yang, H. Tang, and J. Xu, "Fuxi: A fault-tolerant resource management and job scheduling system at internet scale," *Proc. VLDB Endowment*, vol. 7, no. 13, pp. 1393–1404, 2014.

[4] A. Java, X. Song, T. Finin, and B. Tseng, "Why we twitter: Understanding microblogging usage and communities," in *Proc. 9th WebKDD/1st SNA-KDD 2007 Workshop Web Mining Social Netw. Anal.*, 2007, pp. 56–65.

[5] M. Chowdhury, S. Kandula, and I. Stoica, "Leveraging endpoint flexibility in data-intensive clusters," in *Proc. ACM SIGCOMM Comput. Commun. Rev.*, vol. 43, no. 4, 2013, pp. 231–242.

[6] X. Lu, H. Wang, J. Wang, J. Xu, and D. Li, "Internet-based virtual computing environment: Beyond the data center as a computer," *Future Generation Comput. Syst.*, vol. 29, no. 1, pp. 309–322, 2013.

[7] Y. Zhang and L. Liu, "Distributed line graphs: A universal technique for designing dhts based on arbitrary regular graphs," *IEEE Trans. Knowl. Data Eng.*, vol. 24, no. 9, pp. 1556–1569, Sep. 2012.

[8] B. Bamba, L. Liu, P. Pesti, and T. Wang, "Supporting anonymous location queries in mobile environments with privacygrid," in *Proc. 17th Int. Conf. World Wide Web*, 2008, pp. 237–246.

Yiming Zhang received the BSc and MSc degrees in mechanics engineering, in 2001 and 2003, and the PhD degree in computer science, in 2008, all from the Chinese National University of Defense Technology (NUDT), Changsha, Hunan, China. He is currently an associate professor with the School of Computer, NUDT. His current research interests include cloud computing and networking. He received the China Computer Federation (CCF) Distinguished Dissertation Award in 2010. He has published more than 20 books and more than 30 technical articles in journals and conference proceedings. He is the program chair for iVCE Workshop 2016 and was ICDCS11 PC member and P2P10 external PC member. He has participated in more than 10 industrial projects and helped to develop many commercial systems and software tools. He was a visiting professor with Microsoft Research Asia in 2011 and at the computer lab of University of Cambridge in 2012 and 2013.

Rong Chang received the BS degree from the National Chiao Tung University, Taiwan, and the PhD degree from the University of Michigan, USA. He is with IBM Research, researching on innovative services computing technologies for IoT Cloud services. Before joining IBM, he was with Bellcore researching on personal ubiquitous application services for B-ISDN clouds. He is a member of IBM Academy of Technology and his research encompasses the areas of distributed computing, enterprise clouds, and service management optimization. He is an editor-in-chief of the *International Journal of Cloud Computing* and an associate editor of the *IEEE Transactions on Services Computing* and the *International Journal of Services Computing*. He co-authored more than 30 patents and 40 refereed technical articles. He is chair of the IEEE Technical Committee on Services Computing, general chair of the 2015 IEEE World Congress on Services, and distinguished engineer of ACM.

Paul Townend received the PhD degree, in 2006. He has been an active computer scientist since the year 2000, publishing more than 50 internationally peer-reviewed articles and articles. His research spans a wide range of topics within the distributed Ssystems domain, with an emphasis on big data analytics, cloud computing, decision support, energy-efficient computing, and dependability. He has served as general chair of iVCE 2013, IEEE SOSE 2015, IEEE ISORC 2015, and IEEE ISORC 2016, and is currently a university academic fellow with the University of Leeds, United Kingdom. He has been team leader of the Distributed Systems and Services Group at Leeds since 2004, and has been lead researcher on numerous successful research projects. He jointly leads several UK-China research collaborations including the highly-successful COLAB project, and was a visiting researcher with Penn State University, in 2011.