CONCURRENCY AND COMPUTATION: PRACTICE AND EXPERIENCE
*Concurrency Computat.: Pract. Exper.* (2014)
Published online in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/cpe.3421

EDITORIAL

# A note on resource orchestration for cloud computing

Welcome to the special issue of Concurrency and Computation: Practice and Experience (CCPE) journal. This special issue compiles a number of excellent technical contributions that significantly advance the state-of-the-art in the areas of orchestrating cloud resources, composing new cloud services from existing ones, increasing energy efficiency via cloud resource orchestration, and developing cloud-based image processing solutions.

Over the past few years, cloud computing [1–4] has emerged as the latest and most dominant utility computing solution offering both hardware and software resources as virtualization-enabled services. Cloud computing providers such as Amazon Web Services and Microsoft Azure currently provide application owners the option of deploying their applications over a network of a virtually infinite resource pool with practically no up-front capital investment and with operating cost proportional to the actual use (i.e., implementing a pay-as-you-go model). An increasing number of cloud vendors offer information and communication technology (ICT) resources such as hardware (CPUs, GPUs, storage, and networks), software infrastructure (e.g., databases, webservers, stream-processing systems, and data-mining packages), and collaboration/communication applications (e.g., email, video on demand, and social networks) as infrastructure as a service (IAAS), platform as a service (PAAS), and software as a service (SAAS), respectively. This approach allows enterprises to easily, cost effectively, and reliably offer business services that are supported by computing and software resources that are provided and maintained by IAAS, PAAS, and SAAS providers. This makes cloud computing attractive to especially small and medium size enterprises (SMEs), as it allows them to focus more on their core business and less on ICT infrastructure.

One of the fundamental issues in exploiting cloud computing in this fashion is developing better Resource Orchestration (RO) [1–5] techniques and programming frameworks. More specifically, Resource Orchestration (RO) is ‘the set of operations that cloud providers (e.g., AWS) and application owners (e.g., Netflix) undertake (either manually or automatically via computer programs) for selecting, deploying, monitoring, and dynamically controlling configuration of hardware and software resources as a system of QoS assured components that can be seamlessly delivered to end-users’ [1]. Since RO operations span across all layers of cloud computing stack [1], an overall goal of RO is to ensure successful hosting and delivery of applications (SAAS) by managing the fulfillment of the QoS objectives of both the application owners (e.g., maximize availability, maximize throughput, minimize latency, and avoid overloading) and the Cloud resource providers (e.g., maximize utilization, maximize energy efficiency, and maximize profit).

One of the main complexities in Cloud resource management is that Cloud resources are typically identified by unique functional specifications, and then evaluated via their Quality of Service (QoS) properties. However, in practice, each resource may have multiple unique functional specifications that enable serving diverse user needs. For example, a song retrieval application is a Cloud resource that can be identified and enacted by the name of a song or via its lyrics, as users do may not know or remember the names of all song they want to find. The paper titled ‘A Service Evaluation Method for Cross-cloud Service Choreography’ [6] addresses this challenge by proposing a multifunctional specification solution for cross-cloud service choreography. This solution is based on a mixed integer programming model for multifunctional specification/identification that decomposes global resource and application constraints into local constraints. This allows constraint evaluations to be performed for cross-cloud service choreography. Experimental verification of this approach is also provided.

Provisioning cloud resources requires providers and consumers to reach an agreement on the service usage terms and conditions. Such agreements are captured as Service Level Agreements (SLAs). The paper titled ‘AutoSLAM – A Policy-based Framework for Automated SLA Establishment in Cloud

Environments' [8] proposes a policy-based framework for the automated establishment of SLAs in open, diverse, and dynamic cloud environments. The proposed Automated SLA Management framework allows providers and consumers to specify their requirements and capabilities in a flexible and expressive manner. It also supports multiple interaction models for SLA establishment, giving consumers and providers the flexibility to select those that are most appropriate in a given context, while simultaneously participating in multiple concurrent SLA interactions using different interaction models. In addition, this paper proposes a formal model for the underlying policies that is underpinned by (1) a corresponding physical model called WS-SLAM and (2) a reference system architecture that can use for its implementation. The paper validates the proposed framework by describing a Smart Cloud Purchaser prototype that can be used to automatically purchase computing resources from Amazon EC2 under different scenarios and contexts.

Once the resource management problems has been addressed, Cloud computing provides opportunities for developing sophisticated Cloud-based applications by combining simpler ones. However, the Cloud computing market is currently dominated by proprietary interfaces and concepts for orchestrating resources and application services. Therefore, solutions for dealing with cloud service heterogeneity and for increasing Cloud vendor-specific code portability are highly desirable. The paper titled 'Cloud Resource Orchestration within an Open-source Component-based Platform as a Service' [9] provides mechanisms for orchestrating cloud-enabled hardware and software resources via a recently developed open-source platform as a service.

Another issue in enabling Cloud-based service compositions is mitigating potential security threats and delivering sufficient performance. The paper titled 'Reputation-based Web Service Orchestration in Cloud Computing: A survey' [7] introduces a Web service selection and composition approach based on Web service reputation. This paper also presents a survey on some of the most relevant reputation-based Web service orchestration schemes for cloud computing. This includes a thorough analysis of the pros and cons of each of these existing approaches and a comprehensive comparison among all alterative solutions. Finally, the paper present current challenges and future research trends in the field of reputation-based cloud service orchestration.

Energy efficient orchestration of cloud resources is another emerging issue in Cloud computing, as energy consumption is one of the main costs of operating a Cloud. The paper titled 'Minimizing Total Busy Time in Offline Parallel Scheduling with Application to Energy Efficiency in Cloud Computing' [10] addresses this problem by modeling and formulating an energy efficient resource scheduling algorithm. The paper considers the situation where there are a fixed number of jobs to be scheduled on multiple identical computing machines, which have bounded capacities. Each job is associated with a start-time, an end-time, a process time, and demand for machine capacity. The goal is to schedule all of the jobs non-pre-emptively in their start-time and end-time windows, subject to cloud resource capacity constraints such that the total busy time of the machines is minimized. The scheduling problem aiming to minimize busy times is an NP-hard problem, and one of the best-known solutions for this problem is a five-approximation algorithm using a first-fit-decreasing algorithm. This paper proposes and proves a three-approximation algorithm, which includes a modified first-fit-decreasing-earliest algorithm and describes how to apply these results in a Cloud computing environment in order to improve the energy efficiency.

One of premier applications for cloud computing is image processing. This special issue also investigates the cloud-based image processing space in science, and it does this in the context of the Science Gateway for Massive Remote Sensing Image Processing (SGMRSIP) that is used to processes/analyze satellite images. More specifically, the paper titled 'A Web 2.0-based Science Gateway for Massive Remote Sensing Image Processing' [11] introduces a Web 2.0-based browser/server model that is designed for SGMRSIP. This incorporates a cloud-based image processing portal that is built on the top of a high performance computing cluster, and a client/server interaction model. The paper describes the design and implementation of the user interaction model and includes a verification study involving image processing workflows for satellite data analysis. These experimental results show improvements in the software scalability and interaction resulting to a better user experience when compared with the existing SGMRSIP. The final paper in this special issue is titled 'A General Metric and Parallel Framework for Adaptive Image Fusion in Clusters' [12] and is dedicated to the development of techniques (and related theoretical results) for automatic image fusion in the Cloud. Optimal



DOI: 10.1002/cpe

parameters for image fusion are needed since different parameters causing different fusion effects. The paper addresses the issues of the parameter setting for fusing images and quality assessment of the result.

We hope that the readers will find the articles of this special issue to be informative and useful.

## REFERENCES


1. Ranjan R, Benatallah B, Dustdar S, Papazoglou M. Cloud resource orchestration programming: overview, issues and directions. *IEEE Internet Computing*, IEEE Computer Society. (Accepted September 2, 2014)
2. Armbrust M, *et al*. A view of Cloud Computing, *Communications of the ACM Magazine* 2010; **53**(4):50–58, ACM Press, New York, USA.
3. Buyya R, Broberg J, Goscinski A (eds), Cloud Computing: Principles and Paradigms, 644 pages, ISBN-13: 978-0470887998, Wiley Press, USA, February 2011.
4. Wang L, Ranjan R, Chen J, Benatallah B. Cloud Computing: Methodology, Systems, and Applications. CRC Press, Taylor and Francis Group, October 3, 2011; 575 Pages, In Print (anticipated hardcopy publication date January 15, 2012). ISBN: 9781439856413
5. Wieder A, Bhatotia P, Post A, Rodrigues R.Orchestrating the deployment of computations in the Cloud with conductor. *In Proceedings of the 9th USENIX Conference on Networked Systems Design and Implementation* (*NSDI'12*), USENIX Association, Berkeley, CA, USA, 2012; 27–27.
6. Zhang S, Dou W, Chen J. A service evaluation method for cross-cloud service choreography. Concurrency Computation: Practice and Experience, 2013, doi: 10.1002/cpe.3091.
7. Mármol FG, Kuhnen MQ. Reputation-based web service orchestration in Cloud Computing: a survey. Concurrency Computation: Practice and Experience, 2013, doi: 10.1002/cpe.3177.
8. Chhetri MB, Bao Vo Q, Kowalczyk R. AutoSLAM – a policy-based framework for automated SLA establishment in Cloud environments, Concurrency Computation: Practice and Experience, 2013 doi: 10.1002/cpe.3171.
9. Petcu D, Panica S, Crăciun C, Neagul M, Şandru C. Cloud resource orchestration within an open-source component-based platform as a service, Concurrency Computation: Practice and Experience, 2013 doi: 10.1002/cpe.3175.
10. Tian W, Yeo CS. Minimizing total busy time in offline parallel scheduling with application to energy efficiency in cloud Computing, Concurrency Computation: Practice and Experience, 2013, doi: 10.1002/cpe.3176.
11. Miao Y, Wang L, Liu D, Ma Y, Zhang W, Chen L. A Web 2.0-based science gateway for massive remote sensing image processing, Concurrency Computation: Practice and Experience, 2013, doi: 10.1002/cpe.3049.
12. Wei J, Liu D, Wang L. A general metric and parallel framework for adaptive image fusion in clusters. *Concurrency Computation: Practice and Experience* 2014; **26**:1375–1387. doi:10.1002/cpe.3037.



RAJIV RANJAN

*CSIRO Digital Productivity and Services Flagship, North Road, Acton, ACT 2601, Australia*

E-mail: rranjans@gmail.com

RAJKUMAR BUYYA

*Cloud Computing and Distributed Systems (CLOUDS) Laboratory, Department of Computer Science and Software Engineering, The University of Melbourne, Melbourne, Australia*

SURYA NEPAL

*CSIRO Digital Productivity and Services Flagship, Cnr Viemiera and Pembroke Roads Marsfield, NSW 2122, Australia*

DIMITRIOS GEORGAKOPULOS

*School of Computer Science and Information Technology, Royal Melbourne Institute of Technology, Melbourne, Australia*