Future Generation Computer Systems 111 (2020) 931–933

Contents lists available at ScienceDirect

Future Generation Computer Systems

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

Editorial

Economics of computing services: An overview of economic motivated solutions to Cloud Computing

1. Introduction

Advances in distributed systems technology have allowed for the provisioning of IT services with increasing flexibility at an unprecedented scale. As the global market for infrastructure, platforms and software services is continuously evolving supporting new application domains, the need to understand and deal with new implications and the multitude of new challenges is quickly growing. Some of these challenges are related to problems as diverse as the energy consumption involved in distributed systems operations or the definition of competition policies.

The combination of micro- and macro-economic principles with resource management strategies in computer science and engineering provides a fair basis for the analysis and the development of decision making models for large scale distributed systems. Economic resources involved in the production of services include a wide variety of smart devices, communication networks, and data centers [1]. Consequently, efficient resource management becomes a complex task. It has to consider the heterogeneity of resource-scarce devices in the edge, the network, and the architecture of computing resources, the need of orchestrating and solving software service platforms interoperability problems, and the design of cloud federation strategic alliances [2].

2. Articles on cloud economics during the past two years

Articles of past special issues and conference proceedings of the International Conference on the Economics of Grids, Clouds, Systems, and Services (GECON) addressed many of those topics. The purpose of the GECON conference series is to gather original research and build a strong multidisciplinary community in this increasingly important area of the future digital economy. The GECON conference series invites researchers and practitioners from academia and industry to present new research related to information technology and economics, in order to suggest solutions and to discuss lessons learned. The research comprises original work on extensions to existing technologies, successful deployments of technologies, economic analyses of technical systems, and on theoretical concepts in this interdisciplinary area between computer science and economics. Case studies, which demonstrated practical use of economic strategies, benefits and limitations, are also provided. The 2016 and 2017 GECON conferences addressed many of the challenges presented.

In particular, at the conference, business models for digital ecosystems have been discussed, requiring specific regulations and techniques for supporting them. For example, network service differentiation has been the focus of vivid debates. Although the need of Internet traffic has considerably increased in volume past years, most of this traffic is due to a very small number of content providers with strong quality requirements. In [3], Patrick Maillé and Bruno Tuffin in their paper entitled "Non-neutrality Pushed by Big Content Providers" highlight the challenge faced by Internet service providers (ISPs) in the light of large content providers publishing information about the quality of service that ISPs offer for accessing their services. In [4], the work "Insurance Pricing and Refund Sustainability for Cloud Outages", by Loretta Mastroeni and Maurizio Naldi, proposes to use insurance mechanisms for cloud providers to hedge the risks of violating SLAs with cloud users. The authors propose three performance metrics: number of outages, number of long outages and cumulative outage duration.

New business models are necessary for a fair distribution of revenue among members of alliances. Papers "Model for Incentivizing Cloud Service Federation" by Juan Pablo Romero et al. [5], and "Fairness in Revenue Sharing for Stable Cloud Federation" by Ram Govinda Aryal and Jörn Altmann [6], present respectively the variables and incentive schemes which determine the formation of *cloud federations*, where enterprises could reach a win-win situation by federating their cloud infrastructures; and design a profit maximization model, in which cloud service providers share their profits.

In this context, the proliferation of mobile applications and edge devices, which are closer to service clients, involves the user in these alliances. *Edge computing* has emerged as a solution to accommodate complex applications requirements by balancing data offloading and computation in data centers and edge devices. In [7], John Violos et al. present a knowledge extractor in their paper entitled "User Behavior and Application Modeling in Decentralized Edge Cloud Infrastructures". This work proposes a decision support framework developed in the BASMATI project that investigates user behavior (and in particular mobility) and the application model, as well as their combination, to the resource utilization patterns. Ioan Petri et al. in their paper "Incentivising Resource Sharing in Edge Computing Applications" [8], describe potential revenue models for micro-data centres to support service migration and service incoming request for edge based applications. The paper shows that the number of fog devices, network architecture and application specificities are factors that should be considered when developing business models. The contribution of Salman Taherizadeh et al. entitled "A Network Edge Monitoring Approach for Real-Time Data Streaming

https://doi.org/10.1016/j.future.2020.02.038
0167-739X/

Applications" [9] focuses on supporting network-edge monitoring to enable service adaption, primarily to meet user expectations on quality of service. The authors discuss how the use of such monitoring data can be used to alter data flow to/from the service.

With respect to *resource management*, Víctor Medel et al. in their paper "Client-side Scheduling on Kubernetes Based on Application Characterization" [10], propose to use an application characterization provided by users to allow a container management system, such as Kubernetes, to avoid resource contention. The paper "The Design and Evaluation of a Heaviness Metric for Cloud Fairness and Correct Virtual Machine Configurations" by Patrick Poullie, Burkhard Stiller [11] also assumes indirectly that an optimal utility function is given by the user. This work present a runtime prioritization mechanism for a fair assignment of resources to virtual machines according to their respective utility function and greediness. The paper "Cost Minimization of Virtual Machine Allocation in Public Clouds Considering Multiple Applications" [12], by Joaquín Entrialgo et al. presents a virtual machine allocation strategy to optimize the cost of multiple application deployment in public clouds. It is formulated as an optimization problem that takes the level of performance to be reached by a set of applications as inputs. It combines a long-term resource management approach (1 year) with a short-term one (next time slot, one hour).

Finally, *energy consumption* is consolidating as a relevant research area in distributed systems. The relevance of the subject is justified by the forecasts of electric consumption of communication technology, which in the worst case may contribute up to the 23% of greenhouse gas emissions in 2030 [13]. Muhammad Zakarya and Lee Gillamand in their work "An Energy Aware Cost Recovery Approach for Virtual Machine Migration" [14] study how energy efficiency can be optimized in datacenters with regards to a number of migration decisions that can be made. A number of migration heuristics are proposed to minimize energy use imposed by migration of virtual machines. The paper "Energy Efficiency Support Through Intra-layer Cloud Stack Adaptation" by Karim Djemame et al. reports on a cloud architecture to support energy efficiency at service construction, deployment and operation. The authors present an energy-aware cloud architecture alongside an intra-layer self-adaptation methodology tailored for SaaS, PaaS, and Iaa. To more related with energy efficiency papers have been selected in this special issue.

3. Articles of this special issue

Based on these topics of the GECON conferences, this special issue solicited contributions of interdisciplinary nature, combining business and economic aspects with engineering and computer-science related themes. Additionally, works with the largest potential impact presented in GECON 2016 and 2017 and the highest assessment in the reviewing process with three to four reviews per article were invited to be extended and to be significantly enhanced for this special issue. After an additional thorough reviewing process that involved many peer reviews and revisions, we have selected 6 papers. In this special issue, we have three papers related with resource management [15–17], two articles [18,19] related with energy efficiency, and one paper [20] that analyses an existing business ecosystem.

Daniele DAgostino et al. [15] present in their article titled "SoC-based Computing Infrastructures for Scientific Applications and Commercial Services: Performance and Economic Evaluations" an analysis of performance and economic aspects related to the adoption of low-power System-on-Chip (SoC) in IT infrastructures. In particular, the authors investigate the performance of the most representative SoCs for the computational intensive N-body benchmark and a real-life application taken from the field of molecular biology. The execution time and power consumption are compared for several high-performance and low-power processors using scientific computing and molecular biology applications. The analysis focuses on the trade-off between time-to-solution and energy-to-solution together with economical aspects.

Another cross-disciplinary work is "On Using Collaborative Economy for Test Cost Reduction in High Fragmented Environments" [20] by Kenyo Abadio Crosara Faria et al. The paper presents an analysis of an existing business ecosystem for testing applications and the proposal to overcome major limitations of the existing solution. It directly relates to the persisting topic of the lack of sufficient testing capabilities as a major threat to the future of all software based services and applications. The work generally revolves around the proposal for a new platform for a distributed mobile test environment based on a collaborative economy named Distributed Bug Buster (DBB).

In the work on "Reducing the Price of Resource Provisioning using EC2 Spot Instances with Prediction Models" [16], Javier Fabra et al. explored the opportunities of using Amazon Spot Instances for reducing the cost of a deployment. They proposed a model that can predict spot instance prices at different availability zones and a framework that, based on the aforementioned model, generates provisioning plans and deploys them on the Amazon Cloud.

Ilia Pietri and Rizos Sakellariou address [17] in their article "A Pareto-based Approach for CPU Provisioning of Scientific Workflows on Clouds" the issue of workflow scheduling in heterogeneous computing environments like clouds. The particular considerations are on different frequency selection for the compute resources. The authors propose an algorithm called "Pareto-efficient Stepwise Frequency Selection" (PSFS) to provision CPU frequencies for a given number of CPU resources with the objective to find cost-efficient, Pareto-optimal solutions for different execution time vs. cost trade-offs. Simulations have been conducted to show cases, under which condition this approach is beneficial.

The article "Energy-aware Cost Prediction and Pricing of Virtual Machines in Cloud Computing Environments" [18] by Mohammad Aldossary et al. discusses an end-to-end approach to tackle energy efficiency in cloud computing. It suggests a particular architectural design to support energy-awareness and total cost prediction, an energy-aware model for VMs, a framework to predict usage and cost with the objective to minimize energy consumption in VMs. This paper presents also the cloud architecture to incorporate power consumption to the cost and facilitate an energy aware operation of resources. The main contributions are related to the presentation of a layered architecture, an energy-aware model to attribute energy consumption to heterogeneous running VMs, and the proposal and analysis of adoption of pricing energy-based pricing schemes.

Finally, in the paper "Modelling and prediction of resources and services state evolvement for efficient runtime adaptations" [19], Dimosthenis Kyriazis presents an approach to runtime adaptability for maintaining services' QoS based on adaptable monitoring. This, in turn, enables modelling and prediction on different parameters. The paper presents a deep analysis of the state and the progress beyond the current art in this field. The proposed and validated framework is based on widely extended tools and validation results are promising in terms of mechanisms to avoid over-provision, therefore presenting direct implications in costs for both cloud /fog/edge users and providers. The proposed mechanism is agnostic to the actual metric (or sets of) being monitored expanding its applicability at all levels of the stack from application to infrastructure resource.

4. Conclusion

This special issue has required a lot of work from many people, who we would like to thank very much. Foremost, we would like to thank Peter Sloot, as the Editor-in-Chief of the journal Future Generation Computer Systems (FGCS), for his continuing support on the topic of economics of computing systems. Furthermore, we would like to thank Winky Huang of Elsevier for her support in the publishing process. Last but not least, we would like to thank very much all of our reviewers for their effort in giving feedback to the authors and, thereby, improving the quality of the papers of this special issue.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] Z. Sanaei, S. Abolfazli, A. Gani, R. Buyya, Heterogeneity in mobile cloud computing: taxonomy and open challenges, IEEE Commun. Surv. Tutor. 16 (1) (2013) 369–392.
[2] N. Haile, J. Altmann, Evaluating investments in portability and interoperability between software service platforms, Future Gener. Comput. Syst. 78 (2018) 224–241.
[3] P. Maillé, B. Tuffin, Non-neutrality pushed by big content providers, in: Economics of Grids, Clouds, Systems, and Services - 14th International Conference, GECON 2017, Biarritz, France, September 19–21, 2017, Proceedings, Springer, 2017, pp. 29–39.
[4] L. Mastroeni, M. Naldi, Insurance pricing and refund sustainability for cloud outages, in: Economics of Grids, Clouds, Systems, and Services — 14th International Conference, GECON 2017, in: Lecture Notes in Computer Science, vol. 10537, Springer, 2017, pp. 3–17.
[5] J.P.R. Coronado, J. Altmann, Model for incentivizing cloud service federation, in: Economics of Grids, Clouds, Systems, and Services - 14th International Conference, GECON 2017, in: Lecture Notes in Computer Science, vol. 10537, Springer, 2017, pp. 233–246.
[6] R.G. Aryal, J. Altmann, Fairness in revenue sharing for stable cloud federations, in: Economics of Grids, Clouds, Systems, and Services — 14th International Conference, GECON 2017, in: Lecture Notes in Computer Science, vol. 10537, Springer, 2017, pp. 219–232.
[7] J. Violos, V.M. de Lira, P. Dazzi, J. Altmann, B. Al-athwari, A. Schwichtenberg, Y. Jung, T.A. Varvarigou, K. Tserpes, User behavior and application modeling in decentralized edge cloud infrastructures, in: Economics of Grids, Clouds, Systems, and Services - 14th International Conference, GECON 2017, in: Lecture Notes in Computer Science, vol. 10537, Springer, 2017, pp. 193–203.
[8] I. Petri, O.F. Rana, J. Bignell, S. Nepal, N. Auluck, Incentivising resource sharing in edge computing applications, in: Economics of Grids, Clouds, Systems, and Services - 14th International Conference, GECON 2017, in: Lecture Notes in Computer Science, vol. 10537, Springer, 2017, pp. 204–215.
[9] S. Taherizadeh, I.J. Taylor, A. Jones, Z. Zhao, V. Stankovski, A network edge monitoring approach for real-time data streaming applications, in: Economics of Grids, Clouds, Systems, and Services - 13th International Conference, GECON 2016, in: Lecture Notes in Computer Science, vol. 10382, Springer, 2016, pp. 293–303.
[10] V.M. Gracia, C. Tolón, U. Arronategui, R. Tolosana-Calasanz, J.Á. Bañeares, O.F. Rana, Client-side scheduling based on application characterization on kubernetes, in: Economics of Grids, Clouds, Systems, and Services - 14th International Conference, GECON 2017, in: Lecture Notes in Computer Science, vol. 10537, Springer, 2017, pp. 162–176.
[11] P. Poullie, B. Stiller, The design and evaluation of a heaviness metric for cloud fairness and correct virtual machine configurations, in: Economics of Grids, Clouds, Systems, and Services - 13th International Conference, GECON 2016, in: Lecture Notes in Computer Science, vol. 10382, Springer, 2016, pp. 193–207.
[12] J. Entrialgo, J.L. Díaz, J. García, M. García, D.F. García, Cost minimization of virtual machine allocation in public clouds considering multiple applications, in: Economics of Grids, Clouds, Systems, and Services - 14th International Conference, GECON 2017, in: Lecture Notes in Computer Science, vol. 10537, Springer, 2017, pp. 147–161.
[13] A. Andrae, T. Edler, On global electricity usage of communication technology: trends to 2030, Challenges 6 (1) (2015) 117–157.
[14] M. Zakarya, L. Gillam, An energy aware cost recovery approach for virtual machine migration, in: Economics of Grids, Clouds, Systems, and Services - 13th International Conference, GECON 2016, in: Lecture Notes in Computer Science, vol. 10382, Springer, 2016, pp. 175–190.
[15] D. DAgostino, A. Quarati, A. Clematis, L. Morganti, E. Corni, V. Giansanti, D. Cesini, I. Merelli, SoC-based computing infrastructures for scientific applications and commercial services: Performance and economic evaluations, Future Gener. Comput. Syst. 96 (Harv. Bus. Rev. 2017) (2019) 11–22, http://dx.doi.org/10.1016/j.future.2019.01.024.
[16] J. Fabra, J. Ezpeleta, P. lvarez, Reducing the price of resource provisioning using EC2 spot instances with prediction models, Future Gener. Comput. Syst. (ISSN: 0167-739X) 96 (2019) 348–367, http://dx.doi.org/10.1016/j.future.2019.01.025.
[17] I. Pietri, R. Sakellariou, A pareto-based approach for CPU provisioning of scientific workflows on clouds, Future Gener. Comput. Syst. (ISSN: 0167-739X) 94 (2019) 479–487, http://dx.doi.org/10.1016/j.future.2018.12.004.
[18] M. Aldossary, K. Djemame, I. Alzamil, A. Kostopoulos, A. Dimakis, E. Agiatzidou, Energy-aware cost prediction and pricing of virtual machines in cloud computing environments, Future Gener. Comput. Syst. (ISSN: 0167-739X) 93 (2019) 442–459, http://dx.doi.org/10.1016/j.future.2018.10.027.
[19] D. Kyriazis, Modelling and prediction of resources and services state evolvement for efficient runtime adaptations, Future Gener. Comput. Syst. (ISSN: 0167-739X) 94 (2019) 1–10, http://dx.doi.org/10.1016/j.future.2018.09.035.
[20] K.A.C. Faria, R. de Aquino Gomes, E.N. de Andrade Freitas, A.M.R. Vincenzi, On using collaborative economy for test cost reduction in high fragmented environments, Future Gener. Comput. Syst. (ISSN: 0167-739X) 95 (2019) 502–510, http://dx.doi.org/10.1016/j.future.2019.01.023.

Jörn Altmann
Seoul National University, Seoul, Republic of Korea

José Ángel Bañares
University of Zaragoza, Zaragoza, Spain

Konstantinos Tserpes
Harokopio University, Athens, Greece