

Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:

Ilager, S;Stankovski, V;Pallickarar, S;Buyya, R

Title:

Special issue: Elastic computing from edge to the cloud environments

Date:

2021-09-01

Citation:

Ilager, S., Stankovski, V., Pallickarar, S. & Buyya, R. (2021). Special issue: Elastic computing from edge to the cloud environments. *Software Practice and Experience*, 51 (9), pp.1849-1851. <https://doi.org/10.1002/spe.3012>.

Persistent Link:

<https://hdl.handle.net/11343/298640>

Special Issue: Elastic Computing from Edge to the Cloud Environments

Introduction

We are pleased to present a special issue that focuses on state-of-the-art research on **Elastic Computing from Edge to the Cloud Environments**. Today, a huge amount of data is being generated by the Internet of Things (IoT) devices such as smartphones, sensors, cameras, cars and robots [1]. In order to process the generated data, there exist Big Data platforms (such as Hadoop and Spark). Conventionally, they are deployed in centralised Data Centres, which, however, fall short of addressing time-critical requirements of the applications due to high latency between the Edge, where the data are generated and the Data Centres where they are processed [2]. The emerging Edge and Fog computing paradigms promise to solve this problem by seamlessly integrating hardware and software resources across multiple computing tiers, from the Edge to the Data Centre/Cloud. Since computing resources at the Edge may be power and capacity constrained, it is necessary to invent new lightweight platforms and techniques that seamlessly interact, sense, execute and produce results with very low latency, while at the same time address other high-level requirements of applications, such as security and privacy. Regarding these problems, there are many challenges that must be addressed with the invention of new architectures, methods, algorithms and solutions.

This special issue addresses various problems, challenges, new approaches and technologies to shape future directions for research, foster the exchange of ideas. In this special issue, six research articles are carefully selected after multiple rounds of peer-review to address some new and upcoming avenues of next-generation clouds. They include:

- Integrate and process data from underlying IoT platforms and services
- Improve the energy-efficient management of resources and tasks processing
- Address the QoS and time-critical aspects of smart applications
- Facilitate intelligent integration of information arising from various sources
- Address the requirements of very dynamic Big Data pipelines (e.g. moving smartphones, sensors, cars, robots with dynamically changing requirements for processing)
- Provide orchestration methods and scheduling policies that address dependability, reliability, availability and other high-level application requirements
- Adequately address the inherent variability of resources from the Edge to the Data Centres
- Provide new architectures which use the powerful computing resources of Data Centres, while at the same time providing optimal QoS to applications
- Implement distributed Artificial Intelligence methods from the Edge to the Data Centre/Cloud

This is the author manuscript accepted for publication and has undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: [10.1002/spe.3012](https://doi.org/10.1002/spe.3012)

This article is protected by copyright. All rights reserved.

A Summary of the Contributions

This special issue features six papers covering a range of topics including IoT application deployment frameworks in edge-cloud environments, cost optimisation models, and lightweight virtualisation model.

The first paper in this special issue titled “Edge-adaptable serverless acceleration for machine learning Internet of Things applications” [3] presents STOIC (serverless teleoperable hybrid cloud), an IoT application deployment and offloading system that extends the serverless model. The authors have developed a dynamic feedback control mechanism to precisely predict latency and dispatch workloads uniformly across edge and cloud systems using a distributed serverless framework. STOIC leverages hardware acceleration (e.g., GPU resources) for serverless function execution when available from the underlying cloud system. Finally, it configures the system to overcome deployment variability associated with public clouds. The system is evaluated using real-world machine learning applications and multi-tier IoT deployments (edge and cloud) and shown that it reduces overall execution time and achieves placement accuracy in the range of 92% to 97%.

The second paper in this special issue titled “Server Configuration Optimisation in Mobile Edge Computing: A Cost-Performance Tradeoff Perspective” [4] studies the problem of server configuration optimisation in Mobile edge computing (MEC) environments. The authors use M/M/m queuing models and establish the performance and cost models for the system. The article considers cost-constrained performance optimisation, and performance constrained cost optimisation based on multiple numerical algorithms. The numerical simulation-based experiments have shown this approach is able to balance the trade-off between investment cost and service quality.

The third paper in this special issue titled “EFFORT: Energy-efficient framework for offload Communication in mobile cloud computing” [5] proposes a task offloading mechanism from mobile to the remote cloud. The author's solution aims to solve the energy consumption of communication-intensive applications from mobile devices such as smartphones. The experimental evaluation is done by implementing a demonstration application in Android mobile OS. The results have shown that the proposed solution reduces the energy consumption of smartphones while executing applications and simultaneously reducing the communication cost.

The fourth paper in this special issue titled “Human Microservices: A framework for turning humans into service providers” [6] presents a framework facilitating the deployment of APIs on companion devices (smartphones and IoT devices). The author's framework proposes a new approach aiming to integrate humans in the IoT loop and facilitate computation units deployment in the devices that are closer to users, instead of remote clouds. The solution leverages existing standards for the rapid development of applications and improves software quality. Through a detailed case study on monitoring people's activity, the authors developed a software application following framework specification and demonstrated the feasibility of their proposed solution.

The fifth paper in this special issue titled “Function delivery network: Extending serverless computing for heterogeneous platforms” [7] introduces the extension to Function-as-a-Service (FaaS) model. The authors consider heterogeneous clusters and support heterogeneous

functions through a network of distributed heterogeneous target platforms called Function Delivery Network (FDN). The article proposes Function-Delivery-as-a-Service (FDaaS), a service model to deliver the function to the required to be targeted platform. It also ensures that Service Level Objective (SLO) energy efficiency requirements are met when scheduling functions. The new benchmarking system FDNInspecto is implemented which benchmarks the different distributed target platforms. The results have found that functions scheduling on edge reduce energy consumption by 17× without violating the SLO requirements when compared to a high-end target platform.

Finally, the last paper of this special issue titled “A Lightweight Virtualisation Model to Enable Edge Computing in Deeply Embedded Systems” [8] builds a virtualisation model for resource-constrained embedded devices. As the existing containers cannot be used in many deeply embedded systems (DES) due to an underlying operating system's (OSs) resource requirements including storage, memory, and processing power. To mitigate this issue, the authors present the Hellfire hypervisor, a lightweight virtualisation model that enables separation and improves the security of IoT applications on DES. The Hellfire simplifies essential components of virtualisation including compute, storage, and I/O among others. The experimental results conducted with the coremark benchmark show Hellfire has a small footprint of 23 KB while keeping a low average virtualisation overhead of 0.62% for multiple virtual machines execution.

To summarise, the articles in this special issue provide new definitions, architectural principles, systems and approaches to solve the challenges posed by emerging distributed systems such as edge and cloud computing, mainly driven by IoT workloads. They address important problems including lightweight platforms, energy efficiency, reliability and easy to integrate frameworks.

We hope that you will find this special issue truly useful and enjoyable.

Acknowledgements

We sincerely thank the editor-in-chief for helping us to organise this special issue. We thank the editorial office staffs for their continuous support. We are also thankful to all the authors for submitting their research works. More importantly, thanks to the reviewers for their valuable contributions in providing thoughtful comments and improving the quality of articles.

Shashikant Ilager
CLOUDS Lab, School of Computing and Information Systems
The University of Melbourne
Email - shashi.ilager@unimelb.edu.au
URL: shashikantilager.com

Vlado Stankovski
Faculty of Computer and Information Science
University of Ljubljana, Slovenia
Email: vlado.stankovski@fri.uni-lj.si
URL: <https://www.fri.uni-lj.si/sl/o-fakulteti/osebje/vlado-stankovski>

Shrideep Pallickarar
Department of Computer Science
Colorado State University, USA
Email - shrideep@cs.colostate.edu
URL: <http://www.cs.colostate.edu/~shrideep/>

Rajkumar Buyya
CLOUDS Lab, School of Computing and Information Systems
The University of Melbourne
Email: rbuyya@unimelb.edu.au
URL: <http://www.buyya.com>

Correspondence
Shashikant Ilager
shashi.ilager@unimelb.edu.au
Orcid: <https://orcid.org/0000-0003-1178-6582>

References

1. Gubbi, Jayavardhana, Rajkumar Buyya, Slaven Marusic, and Marimuthu Palaniswami. "Internet of Things (IoT): A vision, architectural elements, and future directions." *Future generation computer systems* 29, no. 7 (2013): 1645-1660.
2. Satyanarayanan, Mahadev. "The emergence of edge computing." *Computer* 50, no. 1 (2017): 30-39.
3. Zhang, Michael, Chandra Krintz, and Rich Wolski. "Edge-adaptable serverless acceleration for machine learning Internet of Things applications." *Software: Practice and Experience* (2020). DOI: <https://doi.org/10.1002/spe.2944>
4. He, Zhenli, Kenli Li, Kegin Li, and Wei Zhou. "Server configuration optimization in mobile edge computing: A cost-performance tradeoff perspective." *Software: Practice and Experience* (2021). DOI: <https://doi.org/10.1002/spe.2951>
5. Malik, Saif UR, Hina Akram, Sukhpal Singh Gill, Haris Pervaiz, and Hassan Malik. "EFFORT: Energy efficient framework for offload communication in mobile cloud computing." *Software: Practice and Experience* (2020). DOI: <https://doi.org/10.1002/spe.2850>
6. Laso, Sergio, Javier Berrocal, José García-Alonso, Carlos Canal, and Juan Manuel Murillo. "Human microservices: A framework for turning humans into service providers." *Software: Practice and Experience* (2021). DOI: <https://doi.org/10.1002/spe.2976>
7. Jindal, Anshul, Michael Gerndt, Mohak Chadha, Vladimir Podolskiy, and Pengfei Chen. "Function delivery network: Extending serverless computing for heterogeneous platforms." *Software: Practice and Experience* (2021). DOI: <https://doi.org/10.1002/spe.2966>
8. Tiburski, Ramão T., Carlos R. Moratelli, Sérgio F. Johann, Everton de Matos, and Fabiano Hessel. "A lightweight virtualization model to enable edge computing in deeply embedded systems." *Software: Practice and Experience* (2021). DOI: <https://doi.org/10.1002/spe.2968>

Special Issue: Elastic Computing from Edge to the Cloud Environments

Introduction

We are pleased to present a special issue that focuses on state-of-the-art research on ***Elastic Computing from Edge to the Cloud Environments***. Today, a huge amount of data is being generated by the Internet of Things (IoT) devices such as smartphones, sensors, cameras, cars and robots [1]. In order to process the generated data, there exist Big Data platforms (such as Hadoop and Spark). Conventionally, they are deployed in centralised Data Centres, which, however, fall short of addressing time-critical requirements of the applications due to high latency between the Edge, where the data are generated and the Data Centres where they are processed [2]. The emerging Edge and Fog computing paradigms promise to solve this problem by seamlessly integrating hardware and software resources across multiple computing tiers, from the Edge to the Data Centre/Cloud. Since computing resources at the Edge may be power and capacity constrained, it is necessary to invent new lightweight platforms and techniques that seamlessly interact, sense, execute and produce results with very low latency, while at the same time address other high-level requirements of applications, such as security and privacy. Regarding these problems, there are many challenges that must be addressed with the invention of new architectures, methods, algorithms and solutions.

This special issue addresses various problems, challenges, new approaches and technologies to shape future directions for research, foster the exchange of ideas. In this special issue, six research articles are carefully selected after multiple rounds of peer-review to address some new and upcoming avenues of next-generation clouds. They include:

- Integrate and process data from underlying IoT platforms and services
- Improve the energy-efficient management of resources and tasks processing
- Address the QoS and time-critical aspects of smart applications
- Facilitate intelligent integration of information arising from various sources
- Address the requirements of very dynamic Big Data pipelines (e.g. moving smartphones, sensors, cars, robots with dynamically changing requirements for processing)
- Provide orchestration methods and scheduling policies that address dependability, reliability, availability and other high-level application requirements
- Adequately address the inherent variability of resources from the Edge to the Data Centres
- Provide new architectures which use the powerful computing resources of Data Centres, while at the same time providing optimal QoS to applications
- Implement distributed Artificial Intelligence methods from the Edge to the Data Centre/Cloud

A Summary of the Contributions

This special issue features six papers covering a range of topics including IoT application deployment frameworks in edge-cloud environments, cost optimisation models, and lightweight virtualisation model.

The first paper in this special issue titled “Edge-adaptable serverless acceleration for machine learning Internet of Things applications” [3] presents STOIC (serverless teleoperable hybrid cloud), an IoT application deployment and offloading system that extends the serverless model. The authors have developed a dynamic feedback control mechanism to precisely predict latency and dispatch workloads uniformly across edge and cloud systems using a distributed serverless framework. STOIC leverages hardware acceleration (e.g., GPU resources) for serverless function execution when available from the underlying cloud system. Finally, it configures the system to overcome deployment variability associated with public clouds. The system is evaluated using real-world machine learning applications and multi-tier IoT deployments (edge and cloud) and shown that it reduces overall execution time and achieves placement accuracy in the range of 92% to 97%.

The second paper in this special issue titled “Server Configuration Optimisation in Mobile Edge Computing: A Cost-Performance Tradeoff Perspective” [4] studies the problem of server configuration optimisation in Mobile edge computing (MEC) environments. The authors use M/M/m queuing models and establish the performance and cost models for the system. The article considers cost-constrained performance optimisation, and performance constrained cost optimisation based on multiple numerical algorithms. The numerical simulation-based experiments have shown this approach is able to balance the trade-off between investment cost and service quality.

The third paper in this special issue titled “EFFORT: Energy-efficient framework for offload Communication in mobile cloud computing” [5] proposes a task offloading mechanism from mobile to the remote cloud. The author’s solution aims to solve the energy consumption of communication-intensive applications from mobile devices such as smartphones. The experimental evaluation is done by implementing a demonstration application in Android mobile OS. The results have shown that the proposed solution reduces the energy consumption of smartphones while executing applications and simultaneously reducing the communication cost.

The fourth paper in this special issue titled “Human Microservices: A framework for turning humans into service providers” [6] presents a framework facilitating the deployment of APIs on companion devices (smartphones and IoT devices). The author’s framework proposes a new approach aiming to integrate humans in the IoT loop and facilitate computation units deployment in the devices that are closer to users, instead of remote clouds. The solution leverages existing standards for the rapid development of applications and improves software quality. Through a detailed case study on monitoring people’s activity, the authors developed a software application following framework specification and demonstrated the feasibility of their proposed solution.

The fifth paper in this special issue titled “Function delivery network: Extending serverless computing for heterogeneous platforms” [7] introduces the extension to Function-as-a-Service (FaaS) model. The authors consider heterogeneous clusters and

support heterogeneous functions through a network of distributed heterogeneous target platforms called Function Delivery Network (FDN). The article proposes Function-Delivery-as-a-Service (FDaaS), a service model to deliver the function to the required to be targeted platform. It also ensures that Service Level Objective (SLO) energy efficiency requirements are met when scheduling functions. The new benchmarking system FDNInspecto is implemented which benchmarks the different distributed target platforms. The results have found that functions scheduling on edge reduce energy consumption by 17× without violating the SLO requirements when compared to a high-end target platform.

Finally, the last paper of this special issue titled “A Lightweight Virtualisation Model to Enable Edge Computing in Deeply Embedded Systems” [8] builds a virtualisation model for resource-constrained embedded devices. As the existing containers cannot be used in many deeply embedded systems (DES) due to an underlying operating system's (OSs) resource requirements including storage, memory, and processing power. To mitigate this issue, the authors present the Hellfire hypervisor, a lightweight virtualisation model that enables separation and improves the security of IoT applications on DES. The Hellfire simplifies essential components of virtualisation including compute, storage, and I/O among others. The experimental results conducted with the coremark benchmark show Hellfire has a small footprint of 23 KB while keeping a low average virtualisation overhead of 0.62% for multiple virtual machines execution.

To summarise, the articles in this special issue provide new definitions, architectural principles, systems and approaches to solve the challenges posed by emerging distributed systems such as edge and cloud computing, mainly driven by IoT workloads. They address important problems including lightweight platforms, energy efficiency, reliability and easy to integrate frameworks.

We hope that you will find this special issue truly useful and enjoyable.

Acknowledgements

We sincerely thank the editor-in-chief for helping us to organise this special issue. We thank the editorial office staffs for their continuous support. We are also thankful to all the authors for submitting their research works. More importantly, thanks to the reviewers for their valuable contributions in providing thoughtful comments and improving the quality of articles.

Shashikant Ilager
CLOUDS Lab, School of Computing and Information Systems
The University of Melbourne
Email - shashi.ilager@unimelb.edu.au
URL: shashikantilager.com

Vlado Stankovski
Faculty of Computer and Information Science
University of Ljubljana, Slovenia
Email: vlado.stankovski@fri.uni-lj.si
URL: <https://www.fri.uni-lj.si/sl/o-fakulteti/osebje/vlado-stankovski>

Shrideep Pallickarar
Department of Computer Science
Colorado State University, USA
Email - shrideep@cs.colostate.edu
URL: <http://www.cs.colostate.edu/~shrideep/>

Rajkumar Buyya
CLOUDS Lab, School of Computing and Information Systems
The University of Melbourne
Email: rbuyya@unimelb.edu.au
URL: <http://www.buyya.com>

Correspondence
Shashikant Ilager
shashi.ilager@unimelb.edu.au
Orcid: <https://orcid.org/0000-0003-1178-6582>

References

1. Gubbi, Jayavardhana, Rajkumar Buyya, Slaven Marusic, and Marimuthu Palaniswami. "Internet of Things (IoT): A vision, architectural elements, and future directions." *Future generation computer systems* 29, no. 7 (2013): 1645-1660.
2. Satyanarayanan, Mahadev. "The emergence of edge computing." *Computer* 50, no. 1 (2017): 30-39.
3. Zhang, Michael, Chandra Krintz, and Rich Wolski. "Edge-adaptable serverless acceleration for machine learning Internet of Things applications." *Software: Practice and Experience* (2020). DOI: <https://doi.org/10.1002/spe.2944>
4. He, Zhenli, Kenli Li, Keqin Li, and Wei Zhou. "Server configuration optimization in mobile edge computing: A cost-performance tradeoff perspective." *Software: Practice and Experience* (2021). DOI: <https://doi.org/10.1002/spe.2951>
5. Malik, Saif UR, Hina Akram, Sukhpal Singh Gill, Haris Pervaiz, and Hassan Malik. "EFFORT: Energy efficient framework for offload communication in mobile cloud computing." *Software: Practice and Experience* (2020). DOI: <https://doi.org/10.1002/spe.2850>
6. Laso, Sergio, Javier Berrocal, José García-Alonso, Carlos Canal, and Juan Manuel Murillo. "Human microservices: A framework for turning humans into service providers." *Software: Practice and Experience* (2021). DOI: <https://doi.org/10.1002/spe.2976>
7. Jindal, Anshul, Michael Gerndt, Mohak Chadha, Vladimir Podolskiy, and Pengfei Chen. "Function delivery network: Extending serverless computing for heterogeneous platforms." *Software: Practice and Experience* (2021). DOI: <https://doi.org/10.1002/spe.2966>
8. Tiburski, Ramão T., Carlos R. Moratelli, Sérgio F. Johann, Everton de Matos, and Fabiano Hessel. "A lightweight virtualization model to enable edge computing in deeply embedded systems." *Software: Practice and Experience* (2021). DOI: <https://doi.org/10.1002/spe.2968>