POLITECNICO DI MILANO

DEPARTMENT OF ELECTRONICS, INFORMATICS AND BIOENGINEERING

DOCTORAL PROGRAMME IN INFORMATION TECHNOLOGY

POST-CLOUD COMPUTING: ADDRESSING RESOURCE MANAGEMENT IN THE RESOURCE CONTINUUM

Doctoral Dissertation of:

Michele Zanella

Supervisor:

Prof. **William Fornaciari**

Tutor:

Prof. **Cristina Silvano**

The Chair of the Doctoral Program:

Prof. **Barbara Pernici**

Cycle XXXIII

First version: October 15th, 2021
Revised version: January 8th, 2022
Ph.D. defense date: January 27th, 2022

Acknowledgments

FIRST of all, I would like to thank my advisor, *Prof. William Fornaciari*, who gave me the opportunity to have this research experience, involved me in international European projects, conferences, trips, summer schools and has been a mentor during these years at *HEAPLab*, providing me suggestions to improve my research and professional skills.

In this regard, I would like to mention my historical office colleagues and friends, *PhD Federico Reghenzani* and *PhD Giuseppe Massari*, for the works, missions, and leisure activities done together. And I would also like to mention *Domenico* and *Filippo* for sharing part of their research journey with me. Of course, I would also like to thank all HEAPLab members for their welcome, cooperation and friendship, and all the european parterns of the projects I was involved in.

I would also like to thank the reviewers of this thesis, *Prof. Fazio* and *Prof. Bertozzi*, for their valuable work and suggestions, and also the tutors and the board of professors of the Doctorate programme.

A special thanks to my parents, *Antonella* and *Lucio*, for their unconditional love and support during my studies and life. And also to my grandmother *Maria Luisa* and my aunt *Santina* for their helpfulness and closeness.

I also special mention to my friends *Daniele*, *Ilaria* and *Simone*, *Ilaria* and *Davide* (and daughters) for their patience and the good times we lived together in these almost

20 years, even online during this pandemic.

In addition, I would like to thank three people who patiently and constantly support me in all my professional and life activities: *don Enrico Bonacina*, *don Ivano Tagliabue* and *don Riccardo Miolo*.

Finally, I would like to thank all those who have not been directly mentioned in these acknowledgements but who are *still* in my life and in some way have helped me and shared with me some moments of this journey.

“Everything that can be counted does not necessarily count;
everything that counts cannot necessarily be counted.”

(Attributed to Albert Einstein, probable author William Bruce Cameron)

Abstract

DUE to the massive amount of data generated by very pervasive IoT and mobile devices combined with a high transfer rate and real-time requirements of emerging scenarios, Cloud computing is showing some limitations. In this sense, post-cloud computing solutions (e.g., Edge and Fog computing) move part (or all) of the computation closer to the data source, making them a very hot research topic. Even if there are some tentative frameworks and standardization proposals, there are no homogeneous architectural models to integrate the various paradigms or, in many cases, they are based on proprietary solutions. Moreover, current solutions implement in part (or not at all) fine-grained resource management techniques, which are necessary to deal with energy-constrained devices. For these reasons, this thesis proposes the BarMan framework as a cooperative approach to integrate different run-time managed levels of resources from the Cloud to Edge continuum.

However, dealing with the aforementioned distributed and multi-level systems means having different kinds of resource heterogeneity. In this sense, this research aims to propose a novel resource-aware and task-based programming model to overcome the current state-of-the-art limitations. Thus, since the correlation between heterogeneity and post-cloud scenarios, the programming model has also been extended for developing and integrate distributed applications. The programming model can integrate applications with the BarbequeRTRM, which monitors and manages devices resources.

Furthermore, in the Fog scenario, mobile devices become part of the computing system because they can be exploited by lower-level or nearby devices to perform part of the computation. On the other hand, mobile devices increase their computational power, still being affected by their energy budget limitation. In this regard, the rest of this research work aims to enable efficient integration of mobile devices at the Fog level through the run-time management of the application's execution, device's resources allocation, and energy consumption while considering the application's performance and requirements.

Finally, to meet the research community demand for real use-cases and hardware test-beds, we applied our approach to a developed video surveillance application and to a large-scale emergency scenario, evaluating them on self-built Fog test-beds and through simulation tools.

Abstract (in italian)

Il progressivo aumento dei dati, generati dai sempre più pervasivi dispositivi mobili e IoT unito ai requisiti real-time sempre più stringenti degli scenari emergenti, mostra il limite tecnologico dell'infrastruttura basta sul Cloud. Per superare queste limitazioni, le più recenti e interessanti soluzioni (Edge e Fog computing) spostano parte della computazione più vicino alle sorgenti di dati. Sebbene in letteratura vi siano proposti dei framework, questi molte volte non presentano modelli architetturali omogenei o sono basati su soluzioni proprietarie. Inoltre, questi ritrovati spesso implementano solo in parte una gestione fine delle risorse, necessaria per integrare dispositivi con energia limitata. Dunque questa tesi presenta BarMan come un framework per integrare senza soluzione di continuità differenti livelli di risorse dalle piattaforme Cloud ai dispositivi edge.

Tuttavia un sistema così distribuito e multilivello richiede di gestire risorse eterogenee. A tal proposito questo lavoro presenta un innovativo modello di programmazione, basato su applicazioni modulari e integrato con il sistema di gestione delle risorse BarbequeRTRM, e di distribuzione trasparente dei moduli.

Negli scenari di Fog computing, i dispositivi mobili possono essere utilizzati da altri dispositivi vicini per svolgere parte della computazione. Nonostante ciò, sebbene negli anni essi abbiano aumentato la loro potenza computazionale, continuano ad avere vincoli energetici. Pertanto la restante parte della ricerca si è concentrata sull'integra-

zione efficiente di questi dispositivi nell'infrastruttura Fog, attraverso la gestione a runtime dell'esecuzione delle applicazioni, dell'allocazione delle risorse e della potenza consumata, tenendo conto delle prestazioni e dei requisiti della singola applicazione.

Infine, data la necessità nella comunità scientifica di avere scenari di applicazione reali, abbiamo adattato il nostro approccio ad uno scenario di videosorveglianza e ad un sistema di gestione delle emergenze, valutandoli su dispositivi reali e attraverso software di simulazione.