Computer Science and Information Systems 22(2):i–iv https://doi.org/10.2298/CSIS250200iR

Special Issue on Computer Systems and Resource Awareness ⋆ Editorial

Gordana Rakić[1], Maja H. Kirkeby[2], Padma Iyenghar[3,4], and Zoran Budimac[1]

[1] University of Novi Sad, Faculty of Sciences, Serbia
gordana.rakic@dmi.uns.ac.rs, zjb@dmi.uns.ac.rs
[2] Roskilde University, Computer Science, Denmark
kirkebym@acm.org
[3] University of Osnabrueck, Germany
[4] Innotec GmbH-TUV Austria Group, Germany
piyengha@uos.de

1. Introduction

Over the past decades, computer systems have evolved to provide high-performance and energy-efficient processing across a broad range of application domains, including mobile, embedded, data-center, and high-performance computing supporting smart, (self-)adaptive, and autonomous systems. However, realizing this potential requires system designers to be resource-aware and to manage the allocation of resources across system components under conflicting constraints. This presents a significant challenge: to understand and control trade-offs in the use of system resources-such as time, memory, energy, and data, both statically and dynamically.

Resource utilization must be considered throughout the systems' life cycle, including development, maintenance, and operation. Making informed trade-offs during specification, design, implementation, and execution requires a deep understanding of how resource usage decisions affect both individual components and the system as a whole. This is particularly important when multiple application threads execute concurrently, potentially competing for shared resources. Such awareness is crucial across sectors, including education, research, innovation, industry, and policy-making, and across all regions, for achieving sustainability in software and system engineering and usage.

The special issue "Computer Systems and Resource Awareness" has been edited and published under the scope of COST Action CA19135[5] CERCIRAS: Connecting Education and Research Communities for an Innovative Resource-Aware Society[6], funded by the European Cooperation in Science and Technology (COST) Association[7]. COST is a

⋆ This special issue publishes extended versions of selected papers presented at the 3 Workshops: (1) CERCIRAS 2021, (2) RAW 2022: Workshop on Resource Awareness of Systems and Society, and (3) RAW 2023: 2nd Workshop on Resource Awareness of Systems and Society.

[5] CA19135 https://www.cost.eu/actions/CA19135/

[6] CERCIRAS https://www.cerciras.org

[7] COST https://www.cost.eu/

funding agency for research and innovation networks. COST Actions help connect research initiatives across Europe and enable scientists to grow their ideas by sharing them with their peers. This boosts their research, career, and innovation.

CERCIRAS COST Action brings together over 400 participants representing more than 35 countries. CERCIRAS' goal is to build capacity and coordinate research to enhance resource awareness in computer and system engineering and usage. Achieving these goals requires action on two levels: (1) connecting fragmented research efforts to develop more holistic views on both problems and solutions; and (2) leveraging appropriate educational and technology assets to improve the understanding and management of resources by the academia, industry, and users of under-performing economies, ultimately contributing to economical, societal, and environmental benefits.

This special issue publishes extended versions of selected papers presented at the three editions of workshops organized under the scope of the COST Action CA19135, CERCIRAS: *Connecting Education and Research Communities for an Innovative Resource Aware Society*. The current RAW workshop series originated from the inaugural CERCIRAS Workshop, which was collocated with the CERCIRAS Open-the first Annual CERCIRAS Action meeting. The second edition of the CERCIRAS Workshop was renamed to RAW: *Workshop on Resource Awareness of Systems and Society*, and was held as part of ICT4S 2022: *ICT for Sustainability*[8]. The third edition, i.e., the 2nd workshop on Resource Awareness of Systems and Society (RAW 2023), was co-located with Euro-Par 2023: the 29th International European Conference on Parallel and Distributed Computing[9]. CERCIRAS supports the RAW workshops throughout the Action with a clear objective: to establish a discussion forum grounded in formal paper submission, peer review, and publication. The forum is intended to foster discussions on early scientific findings and their application in the area of resource-aware computing, and to extend its impact beyond the Action itself -bridging communities, disciplines, domains, and sectors while building and strengthening lasting connections.

The first article in this collection, titled "Hybrid Deployment Strategy for Software Updates to the Manufacturing Execution System Layer" by Petar Rajković et al., extends the authors' previous work on identifying common challenges in software update processes. That earlier work focused on the most effective update strategies running at the lowest (Internet of Things – IoT) and highest (Enterprise Resource Planning – ERP) levels with a special focus on energy efficiency. The further shift to the Manufacturing Execution System (MES) layer following the Industry 4.0 paradigm brought additional challenges and the contribution presented in this article. It describes a further development of the hybrid software deployment system when applied to the multi-connected levels (e.g. MES) that leads to positive results in network load distribution and significant reduction of efforts when a rollback is needed.

Mikhail Tatur et al. in "Comprehensive Approach to the Design of Information Systems and Optimization of Technical Solutions According to Many Criteria" deal with finding the optimal variant of a multiprocessor system as a multi-criteria optimization problem. The article describes several approaches to the comparative assessment of multiprocessor systems. The authors propose a ranking method to evaluate technical solutions. First, they rank individual alternatives for each optimization criterion separately,

[8] ICT4S 2022 `https://conf.researchr.org/home/ict4s-2022`
[9] Euro-Par 2023 `https://2023.euro-par.org`

and afterwards, they aggregate ordered ranked lists. The advantage of using the ranking methods is obtaining a complete technical solutions ranking based on their effectiveness according to several criteria. The article may potentially influence education, training, and practice for the domain of interest.

Another article that may have educational influence even across domains and sectors is entitled "Research on Problem Formulations in Resource-aware Problems Across Scientific Domains and Applications", published by Paweł Czarnul and Mariusz Matuszek. The authors focus on resource-related problems using one of the general problem and solution formulations, such as integer linear programming, greedy algorithms, dynamic programming, evolutionary algorithms, or machine learning oriented ones. The authors identify open research tasks across different domains based on identified formulations. They analyze each of these problems regarding resources to be optimized, specific optimization algorithms and domains. Finally, based on a broad review of literature, they consider the utilization of resources (time, cost, energy, human, computer, natural resources, and data) in problem formulations across domains and disciplines. Hence, the article may serve as an overview of algorithms applied on a specific domain or as an exemplification of the usage of alternative algorithms in related domains.

Petar Rajković et al. in the article entitled "Resource-Aware Design of an IoT Node for Use in Remote Industrial and Hazardous Areas" address the challenge of designing low-power IoT nodes that integrates hardware-software co-design. The design is focused on the replacement of standard hardware components with energy-efficient ones, where the charging of batteries is also optimized. It is based on standardized components for deployment in environments where energy efficiency and autonomy are critical. Such environments are usually remote, off-grid, industrial, and hazardous. On top of such hardware configuration, software components offer over-the-air updates and reconfiguration. The node is integrated into a cloud-based digital twin with centralized control over the complete operation cycle. Results show that the proposed node architecture reduces energy to 50%, while in some cases, consumption is reduced by up to one-tenth compared to starting designs. Finally, the article delivers a set of design recommendations for adapting the standard components for harsh environments.

"Demystifying Power and Performance Variations in GPU Systems through Microarchitectural Analysis", written by Burak Topcu et al. describes a novel memory performance and power monitoring tool for GPU programs. The tool is called GPPRMon and performs a systematic metric collection visualizes them and guides power and performance analysis for target executions based on simulation. The tool gathers metrics that reflect system and memory-related microarchitectural characteristics by monitoring individual instructions and reports collected performance and power consumption information, all dynamically at runtime. The mentioned visualization provides spatial and temporal views of the execution. Authors use the described tool to demonstrate performance and power analysis on GPU benchmark suites, focusing on memory-bound graph applications and resource-critical embedded programs. The results show the potential for application of the tool in memory-bound kernel identification, performance bottleneck analysis of a memory-intensive workload, performance power evaluation of an embedded application, and the impact of input size on the memory structures of an embedded system.

"Comparison and Analysis of Software and Hardware Energy Measurement Methods for a CPU+GPU System and Selected Parallel Applications" by Grzegorz Koszczał et al.

is an extension of the previous work on power capping in optimization of performance-energy metrics of deep neural networks training workloads. They use a professional power meter Yokogawa WT-310E, Intel RAPL and Nvidia NVML interfaces, to observe power consumption of multi-GPU and multi-CPU configurations, such as selected kernels from NAS Parallel Benchmarks for CPUs and GPUs and Horovod-Python Xception deep neural network training using several GPUs. The authors collect, compare and discuss results collected by both power measurement methods performed using: (1) 2 Intel Xeon CPUs and 8 Nvidia Quadro RTX 6000 GPUs, and (2) 2 Intel Xeon CPUs and 4 Nvidia Quadro RTX 5000 GPUs. Finally, the authors compare power consumption between hardware and software interfaces for CPU, GPU and mixed CPU+GPU workload configurations, using 1-40 threads for the CPUs and 1-8 GPUs.

Axel Wiedemann et al. in ”Manatee: A Multicore Interference Analysis Tool for Embedded SoC Evaluation” present the improvement of the RISC-V hardware development framework Chipyard by a tool which gives it awareness of alternation for shared resources during the design and development process of System on Chips. The tool moves the agile development focus of the framework to the effects of changes that are related to the shared resource alternation. The focus is moved by adding the capabilities for quick feedback upon a change regarding its effects on the shared resource. The prototype of the tool is tested and evaluated on a use case.

In the article titled ”Resource-Aware Object Detection and Recognition Using Edge AI Across the Edge-Fog Computing Continuum”, Dragan Stojanović et al. address the challenges of deploying deep neural networks for edge intelligence and traffic object detection and recognition on a video captured by edge device cameras. The authors mainly aim to analyse resource consumption and the achievement of resource awareness. They optimize computational resources across diverse edge devices within the edge-fog computing continuum while maintaining high object detection and recognition accuracy. The authors propose a method to exploit the edge-to-fog paradigm to distribute the inference workload across multiple tiers of the distributed system architecture. Implementation of the proposed method in the context of edge-fog-related solutions is evaluated based on several use cases. Results show improved accuracy of recognition and resource utilization, as well as adaptability of the system to dynamic traffic scenarios, ensuring real-time recognition performance even in challenging environments.

Finally, the article on ”Energy-efficiency of Software and Hardware Algorithms” by Maja H. Kirkeby et al. concludes this Special Issue. The authors compare the energy efficiency of algorithm implementations in hardware and software, with the aim of establishing a fair basis for comparison between the two. Specifically, the study investigates conditions under which time and energy efficiency diverge. To this end, they analyze implementations of the Heapsort sorting algorithm and Dijkstra’s path-finding algorithm written in C for the Raspberry Pi, and in Chisel for deployment on an FPGA. The results contribute to a deeper understanding of the trade-offs between performance and energy efficiency across different computational platforms.

cost EUROPEAN COOPERATION IN SCIENCE & TECHNOLOGY
Funded by the European Union
CERCIRAS