SPARCITY

NEWSLETTER

Issue 3 | June 2023

KOÇ UNIVERSITY · Sabancı Üniversitesi · inesc id lisboa · [simula . research laboratory] · LMU LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN · GRAPHCORE

In this issue:

EDITORIAL

Dear readers,

Launched in April 2021, this 3-year project is funded by the European High Performance Computing Joint Undertaking (EuroHPC JU) under the 2019 call of Extreme Scale Computing and Data Driven Technologies for research and innovation actions.

The SparCity project aims at creating a supercomputing framework that is providing efficient algorithms and coeherent tools specifically designed for maximizing the performance and energy efficiency of sparse computations on emerging HPC systems, while also opening up new usage areas for sparse computations in data analytics and deep learning.

SparCity delivers a coherent collection of innovative algorithms and tools for enabling both high efficiency of sparse computations on emerging hardware platforms.

The SparCity Coordination,

Didem Unat

In-person Meeting | September 1-3, 2022 (Istanbul, Turkey)

PAST EVENTS

☒ SIAG Supercomputing Spotlights Webinar|February 15, 2023

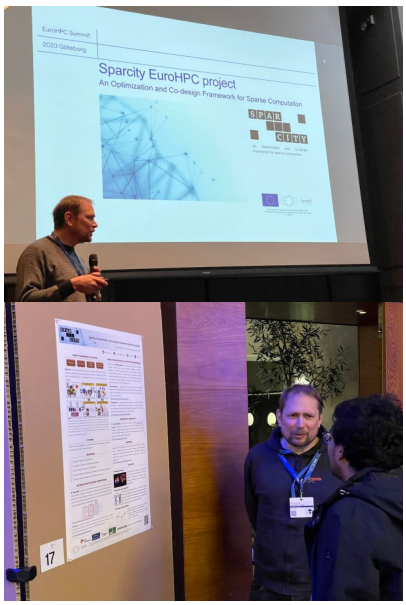
Prof. Didem Unat gave a webinar entitled "Supercomputing Software for Moore and Beyond" at Supercomputing Spotlights, a new webinar series. This is an outreach initiative of SIAG/Supercomputing.

Video

☒ SIAM Meeting | February 27, 2023

Dr Johannes Langguth from Simula Research Laboratory had a lecture: "ML Accelerator Hardware: A New Model For Parallel Sparse Computations?" at the SIAM Meeting, Amsterdam, The Netherlands. You can find the abstract here.

☒ EuroHPC Summit | March 20-23, 2023



The SparCity project was represent at the EuroHPC Summit in Gothenburg, Sweden. In this Summit members from all areas of the European HPC community, from researchers to industry to decision makers, congregated to share in discussions about the state of play of European HPC, its achievements and European priorities going forwards. Dr Johannes Langguth from Simula Research Laboratory presented the project in an oral communication and a poster session.

☒ SAB Member visit at Koç University | May, 2023

Prof. Paul Kelly from Imperial College London and one of the SparCity Scientific Advisory Board members visited Koç University for a couple of days. Prof. Didem Unat's team had the chance to present our project, among others, and discuss several ideas and ongoing steps.

☒ HPC student Club at Istanbul Technical University | May 4, 2023

Prof. Didem Unat gave a Seminar about SparCity and other projects at the HPC student Club at Istanbul Technical University. This initiative included 20 students.

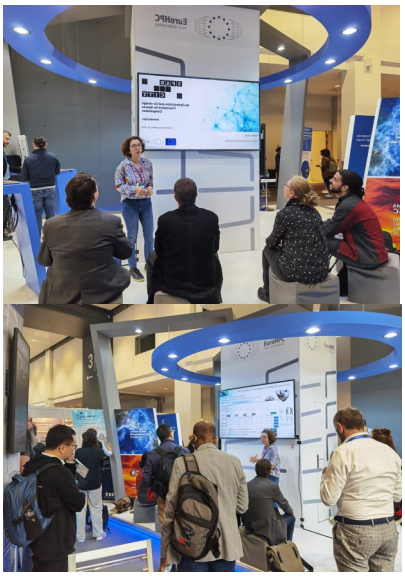
☒ Student club at Middle east technical University | May 5, 2023

Prof. Didem Unat gave a similar seminar about SparCity and other projects at the Student club at Middle east technical University. This initiative included 30 students.

☒ ISC 2023 | May 21-25, 2023

Prof. Didem Unat, Dr Are Magnus Bruaset and Dr Sara Tanqueiro were present at the ISC High Performance 2023. This international conference and exhibition fosters the growth of a global HPC community of technology providers and users.

Prof. Didem presented the SparCity project at the EuroHPC Booth (see more in the "Communication, Dissemination & Outreach" section.



☒ RISC-V HPC23 Workshop| May 25, 2023

Alexandre Rodrigues from INESC-ID participated in the first international workshop on RISC-V for HPC. The goal of this workshop is to continue building the community of RISC-V in HPC, sharing the benefits of this technology with domain scientists, tool developers, and supercomputer operators.

COMMUNICATION, DISSEMINATION & OUTREACH

♦SparCity Video

April, 2023

We released the SparCity promotional video, which presents the our project in a nutshell. Watch now to learn how we're revolutionizing HPC in Europe!

Video

♦ #MeetTheTeam and #CodeVideos Series

We published on YouTube, project website and social media new #MeetTheTeam videos with Eren Yenigül and Tuğba Torun from Sabanci University.

We also released new #CodeVideos about SparseBase, Cardiac Simulations, Epistasis and Yloc.

New videos will be released every month where you can learn more about a different team member and the work developed within SparCity.

#MeetTheTeam videos

#CodeVideos

♦ ISC 2023

May 22-24, 2023

Our project was present at the EuroHPC-JU Booth at ISC 2023 in Hamburg. Prof. Didem Unat and Dr Sara Tanqueiro had the opportunity to share our project with the scientific academia and industry. Thanks to EuroHPC-JU for organising the participation of all projects in this conference.

PUBLICATIONS

♦Andreas Thune, Sven-Arne Reinemo, Tor Skeie and Xing Cai (2023). Detailed Modeling of Heterogeneous and Contention-Constrained Point-to-Point MPI Communication. IEEE Transactions on Parallel and Distributed Systems.
DOI: 10.1109/TPDS.2023.3253881

♦Luk Burchard, Kristian Gregorius Hustad, Johannes Langguth and Xing Cai (2023). Enabling Unstructured-Mesh Computation on Massively Tiled AI-Processors: An Example of Accelerating In-Silico Cardiac Simulation. Front. Phys. Sec. Statistical and Computational Physics.
DOI: 10.3389/fphy.2023.979699

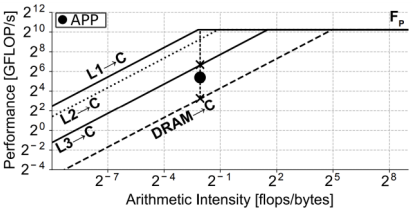
♦Muhammad Aditya Sasongko, Milind Chabbi, Paul H. J. Kelly and Didem Unat (2023). Precise event sampling-based data locality tools for AMD multicore architectures. Concurrency Computat Pract Exper.
DOI: 10.1002/cpe.7707

♦Sergej Breiter, Josef Weidendorfer, Minh Thanh Chung and Karl Fürlinger (2023). A Profiling-Based Approach to Cache Partitioning of Program Data. Parallel and Distributed Computing, Applications and Technologies. PDCAT 2022. Lecture Notes in Computer Science, vol 13798. Springer, Cham.
DOI: 10.1007/978-3-031-29927-8_35

RESOURCES

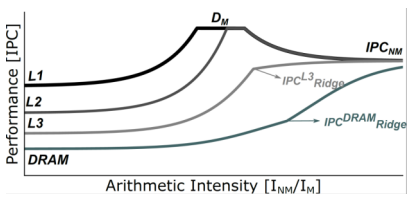
Roofline Models

Roofline models are tools for assessing the maximum efficiency of general-purpose and domain-specific architectures, such as multi-core, many -core, and accelerator processor designs. Such models combine information about the maximum compute performance and peak memory bandwidth of the architecture with the arithmetic intensity represented as the ratio of the number of floating-point operations performed relative to the number of bytes of transferred across different levels of the memory hierarchy. By depicting these relationships in an intuitive plot, one can easily judge whether a specific kernel is bounded by the available memory bandwidth or compute performance, and which optimisations to apply in order to fully exploit the architecture capabilities. To aid in this process, the Cache-aware Roofline Model[1] includes several compute and memory ceilings, typically representing the maximum perfor-mance achievable with different parallelisation and SIMD vectorization strategies, as well as strategies for efficient memory transfers and data reuse for different levels of the memory hierarchy.



While roofline models have been around for more than a decade, they have until now had an unleashed potential for accurate estimation of the achievable performance of sparse kernels. Such sparse kernels are algorithmic com-ponents that explicitly take advantage of the sparse nature of the underlying data and thereby optimize the kernel's memory usage and computational burden by clever implementations.

In SparCity, specially crafted micro-benchmarks and advanced tools for system monitoring are tightly coupled in order to extract the characteristics of a targeted computer system running sparse computations. This allows the construction of realistic upper bounds for the memory bandwidth and compute performance for such workloads. These bounds can differ substantially from the theoretic upper bounds found in the system manuals and factsheets, which are usually derived under idealised assumptions. This has led to the introduction of the Sparse Cache-Aware Roofline Model, which explicitly adapts the upper bounds to sparse kernels and takes into account kernel-specific reuse of data across different levels of the memory hierarchy. SparCity has also proposed the Mansard Roofline Model[2], which uncovers a minimum set of architectural features that must be considered to provide insightful, but yet accurate and realistic, modeling of performance upper bounds for modern processors.



Moving into the field of domain-specific processors designed bottom-up for AI workloads, the Graphcore IPU Roofline Model addresses all main IPU architecture characteristics and execution phases in order to accurately estimate the performance, power consumption and energy efficiency for intile computations.

ComDetective

ComDetective is a communication matrix generation tool that leverages hardware features in commodity CPUs to detect inter-thread data movement, while avoiding the drawbacks of prior work by being more accurate and introducing low time and space overheads. https://github.com/ParCoreLab/ParCoreTools

Reusetracker

ReuseTracker is a tool measuring the reuse distance – a widely used metric for data locality – in parallel shared-memory applications by taking into account the cache coherence. It can help programmers optimize their code by mapping the detected reuses to their locations in source code and improving the data locality of their applications.
https://github.com/ParCoreLab/ParCoreTools

References: [1] A. Ilic, F. Pratas, and L. Sousa. "Cache-aware roofline model: Upgrading the loft." *IEEE Computer Architecture Letters* 13.1 (2014): 21-24. [2] D. Marques, A. Ilic, and L. Sousa. "Mansard roofline model: Reinforcing the accuracy of the roofs." *ACM Transactions on Modeling and Performance Evaluation of Computing Systems* 6.2 (2021): 1-23.

UPCOMING EVENTS

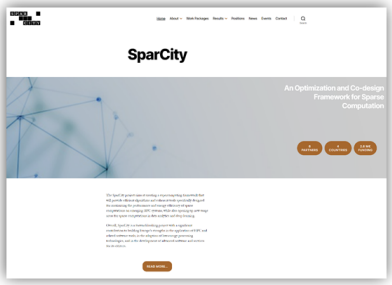
♦EuroHPC Collaboration Meeting, Turin
June 6-7, 2023

♦SparCity In-person Meeting, Lisbon, Portugal
June 15-16, 2023

♦ First Workshop on Tools for Data Locality, Power and Performance (TDLPP 2023)
EuroPar 23, Limassol, Cyprus
August 28 - 29, 2023
Know more here

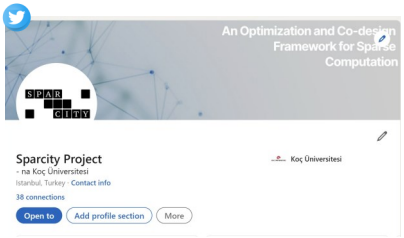
♦Future is Sparse: Methods and Tools for Sparse Computations
SC 2023, Denver, Colorado
November 12 - 17, 2023
Know more here

WEBSITE



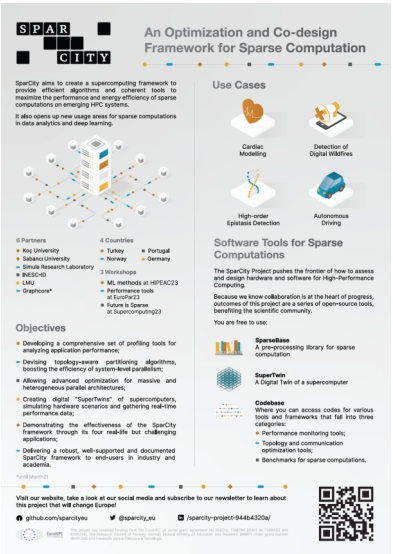
https://sparcity.eu

SOCIAL MEDIA

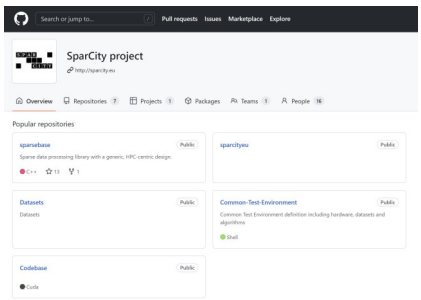


PROMOTIONAL MATERIALS

SparCity released new promotional material



REPOSITORY



https://github.com/sparcityeu

MORE INFORMATION

https://sparcity.eu
sparcity-project-group@ku.edu.tr

This project has received funding from the European High-Performance Computing Joint Undertaking under grant agreement No 956213.

Koç University and Sabancı University are supported by the Turkish Science and Technology Research Centre Grant No 120N003 and 220N254, respectively.

Simula and Graphcore are supported by the Research Council of Norway. Ludwig-Maximilians-Universität München (LMU) is supported by the German national Funding agency (BMBF). INESC-ID is supported by Fundação para a Ciência e a Tecnologia (FCT).