EURO²

CoE European Excellence in HPC Applications

EURO 4SEE.

SUCCESS STORIES 2025

EuroCC 2/EuroCC4SEE and CoE Success Stories

1st Edition, 2025

Publisher:

HLRS on behalf of the EuroCC 2/EuroCC4SEE and CASTIEL 2 Consortium

(Nobelstraße 19, 70569 Stuttgart)

Published: December 2025

Text and images: Members of EuroCC 2, EuroCC4SEE and CASTIEL 2 and CoE Consortia

Editor: Members of EuroCC 2, EuroCC4SEE, CASTIEL 2 Consortium

EuroCC 2 and EuroCC4SEE have received funding from the European High-Performance Computing Joint Undertaking (JU) under grant agreement No 101101903 and No 101191697. The JU receives support from the Eurpoan Union's Digital Europe Programme and Germany, Bulgaria, Austria, Croatia, Cyprus, Czech Republic, Denmark, Estonia, Finland, Greece, Hungary, Ireland, Italy, Lithuania, Latvia, Poland, Portugal, Romania, Slovenia, Spain, Sweden, France, Netherlands, Belgium, Luxembourg, Slovakia, Norway, Türkiye, Republic of North Macedonia, Iceland, Montenegro, Serbia, Bosnia and Herzegovina.
CASTIEL 2 has received funding from the European High-Performance Computing Joint Undertaking (JU) under grant agreement No 101102047. The JU receives support from the European Union's Digital Europe Programme and Germany, Italy, Spain, France, Belgium, Austria, Estonia.

Index

www.hpc-portal.eu

HPC IN EUROPE PORTAL

CHAPTER 1

ABOUT THE PROJECTS

Summary of Projects

About the EuroCC 2 Project[1]

Since 2023, EuroCC 2 and (since 2024) EuroCC4SEE[2] along with their preceding project EuroCC set up, support, and manage the network of 33 National Competence Centres (NCCs) in HPC (High-Performance Computing) across EU member states and associated countries. Their collective goal is to make HPC, High-Performance Data Analytics (HPDA) – collectively known as HPC+ – and Artificial Intelligence (AI) broadly available to different users from science, industry, public administration, and society.

EuroCC activities – coordination by the University of Stuttgart – mainly support NCCs in setting up their individual operational framework by facilitating access to currently available experience and expertise in HPC+ and AI at national and European levels. Ultimately, the main goal is to increase HPC related competences across Europe. EuroCC 2 and EuroCC4SEE work closely with the project CASTIEL 2, which not only supports NCCs but also the European HPC Centres of Excellence (CoEs).

EuroCC 2 and EuroCC4SEE @eurocc.bsky.social

About the HPC Centres of Excellence (CoEs)

The European HPC CoEs (currently 14 active CoEs) are pan-European projects organized by core expertise in specific application domains, motivated by grand scientific and societal challenges, and focus on preparing leading HPC simulation codes for the exascale era through showcasing important pilot use cases. The areas covered by the current set of CoEs range from astrophysics to plasma simulation, climate, geophysics and engineering down to medicine. Individual CoEs featured in this booklet are described in detail on page 146 and 147.

HPC CoE @hpccoe.bsky.social

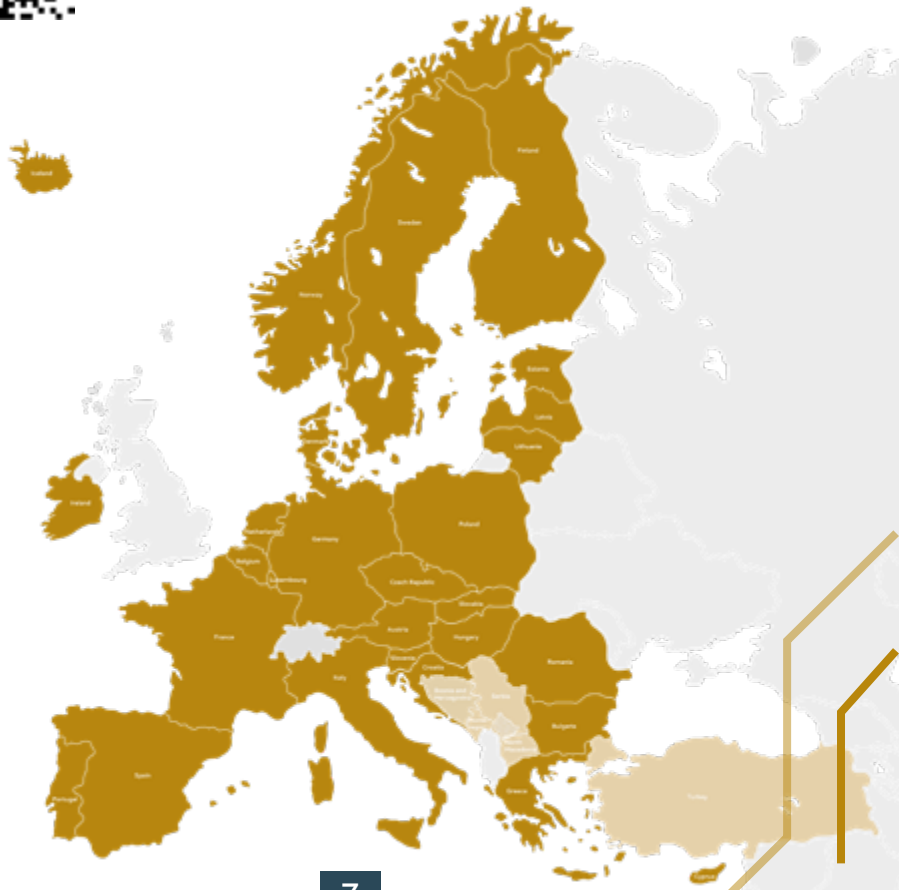
[1] *Funding acknowledgements for all projects are compiled on page 168 and 169.*

[2] *For simplicity, this booklet sometimes refers to "EuroCC activity" to encompass both EuroCC 2 and EuroCC4SEE projects as well as the preceding project EuroCC, if applicable.*

CASTIEL 2 – Coordination and Support Action

CASTIEL 2 coordinates and supports the NCC and CoE activities at the European level and fosters collaborations both between and among NCCs and CoEs to achieve maximum impact on their user communities and the European HPC landscape. CASTIEL 2 works closely with EuroCC 2/EuroCC4SEE to help elevate overall competence levels and promote best practices, with both NCC and CoE collaborations to highlight and promote their work. The aggregation of HPC+ competencies in this pan-European network is then leveraged by NCCs and CoEs to boost the global competitiveness of European industry, research, and public institutions, as showcased in this collection of success stories from a broad range of application areas and partner profiles.

Learn more about the projects here!
www.hpc-portal.eu



Introduction to the NCC Success Stories

One of the focal areas of the NCCs' work is using HPC to support measurement, detection, or prediction of **atmospheric and environmental** quantities, often coupled with their effect on ecosystems or agriculture. Examples investigating the **biosphere** include predicting effects of ground ozone on plant growth (p.34), predicting forest microclimate with high resolution to assess effects on the ecosystems (p.30), and a method for monitoring different species of butterflies (p.28). Moving on to **weather**, applications of HPC include weather forecasts using machine learning (p.32, p.36); improved surface wind predictions based on satellite data and AI (p.40); and forecasting areas with large waves to be avoided by shipping routes (p.26). An interesting idea is to mitigate hurricane damage through innovative "bubble curtains" designed to lower sea surface temperature (p.38). Also related to weather effects is using HPC simulations to model urban air pollution in Stockholm (p.42) – an effort that was later followed up by the HiDALGO2 CoE (p.158) using a more sophisticated workflow. Activities linked to **climate and natural hazard impacts** include an extraction of impact-related information from public sources (p.102), and an effort to digitize old weather and climate data from historical handwriting (p.104).

In the area of **agriculture and food**, we see applications like the analysis of a vast amount of satellite data to detect field boundaries (p.52), forecasting of fish location and quality employing either AI based on weather and oceanic data (p.48), or using acoustic data (p.54). Satellite data estimating soil nutrient levels can be used to support regenerative agriculture (p.50), and mass throughput analysis of genomic data present in honey permits to both detect counterfeit honey and monitor honey quality and bee wellbeing (p.46).

In the **energy sector**, an existing energy production planning system has been ported to HPC (p.58). Regarding energy production, efficiency of mixers for biogas digesters was optimized (p.60).

There are also numerous success stories in the field of **manufacturing and engineering**. For instance, in the footwear sector the transition from the time-consuming manual trial and error approach of optimising the design of shoe soles to a computational modelling approach was targeted (p.76). The

process for manufacturing polymer-based containers rigid plastic packaging was optimised (p.78), and the previously time-consuming and resource-intensive trial-and-error methods for the design of extrusion dies for plastic profiles with complex geometries was automated (p.80). Regarding modelling of materials, a bottleneck in dealing with crystal plasticity has been overcome (p.84), and a simulation of liquid metal jet anodes for X-ray machines has been developed (p.86). Further on, HPC simulations are used for optimising the processing of wood (p.64), improving energy efficiency or health conditions of buildings (p.66), or supporting marine and engineering industries on their path to decarbonisation (p.68). Scalability was improved for the simulation solution for an on-board battery charger for electric vehicles (p.70) and space engineering software, the latter demonstrated by computing the impact of space debris on satellites (p.72). A highly accurate wave-based acoustic simulation software was ported to an HPC system (p.74). Finally, an automated workflow for multi-GPU large-scale rendering of bakery equipment has been set up (p.82).

Text- and speech-based services represent a domain of growing importance, typically using Large Language Models (LLMs) as key technology: a reliable way was found to anonymise sensitive data in candidates' CVs while preserving the key qualifications (p.90) and end-to-end speech recognition models for the Scandinavian languages were trained (p.92). Manual managing and categorizing IT service tickets was replaced by an automated AI-based service trained on millions of Hungarian-language documents using HPC resources (p.96). Different foundation LLMs and tackling the challenges of fine-tuning generative models to produce embeddings for specific domains with great efficiency have been explored (p.98). Other applications of LLMs include minimisation of impact of errors or inaccuracies in the way company names and addresses are entered (p.100), processing and analyzing large datasets of legal documents (p.108), and improvements to models for multilingual Natural Language Processing (NLP) through modularization (p.106), and to build an AI-based assistant for administrative work in health care (p.94). LLM-related services are also mentioned in the paragraphs on environment and climate (p.102 and p.104), and in health care (p.116).

Another important area of NCC collaboration with industry and academia focuses on various aspects of the **health sector**, ranging from medical devices and Magnetic Resonance Imaging (MRI) image processing to drug development. Medical images are a focal point of activities, starting with HPC-based approaches for efficient handling of large amounts of medical images (p.114). The time for reconstruction of images from quantitative MRI was drastically reduced by optimising the reconstruction code for HPC (p.120). Advanced AI models are used for lesion detection, measurement, and segmentation from Computed Tomography (CT) scans (p.118) and for optimising the matching from MR images and recommending which part of the donor cartilage would best serve

which recipient (p.124). Moving on to the "wet lab" side of things, HPC is used to generate thousands of native and mutated protein structures of clinically relevant genes within a short time frame (6–12 months) (p.122), to identify molecules capable of inhibiting the interaction between the SARS-CoV-2 Spike protein and human ACE2 receptors (p.126), and to develop a realistic simulation of cell membranes on a molecular level with millions of atoms (p.112). Regarding commercial drug production, the extensive literature review necessary for compliance was automated (p.116).

Last but not least, a few of the success stories do not really fit the industry sector clusters mentioned before: Scaling the fine-tuning of LLMs for developing a math-tutor for students (p.140), improving resource allocation in IoT networks through an AI-driven scheduler (p.142), and implementing a financial risk assessment (p.138). Analysing flow of people inside of buildings improves occupancy of office space (p.132), detection of human poses has applications in urban security and industry (p.134), and algorithms for upcoming quantum computers can be tested already on today's supercomputers (p.136). Finally, large archeogenomics analyses permit a view back into our ancestry (p.130).

Introduction to the CoE Success Stories

The work of the CoEs (Centres of Excellence) is much more focused on the development and performance-porting of applications and workflows than on industrial transfer and take-up studies. Several stories cover workflow enhancements by adding adaptive mesh refinement capabilities as in EXCELLERAT P2 CoE (p.152, p.154) or by involving improved meshing capabilities in HiDALGO2 CoE (p.158). One particularly complex toolchain for creating a digital twin of infrared surveillance of fusion reactors is described by EXCELLERAT P2 CoE (p.156).

For climate codes, there are two stories from ESiWACE3: application code acceleration through dynamics load balancing between CPU and GPU execution of tasks (p.150), and the large speedups (12x – 20x) realized by porting a code to use multiple GPUs (p.148).

Modelling materials at atomic scale is a particularly demanding task. From this domain, we have in this booklet a story from MaX CoE about predicting the mutations of the COVID-19 virus (p.160) and a use case of theirs related to porting and adapting to a novel type of hardware, which also served as a measure for the hardware's performance in a co-design effort (p.162).

Optimized toolchains and workflows can be challenging to keep up-to date. CoE MultiXscale offers an automated solution with their EESSI tool, demonstrated for the Square Kilometre Array project (p.164).

› The world of HPC is full of surprising applications – dive into our selection, enjoy discovering all these amazing scientific, technological, and industrial innovations and improvements, and get inspired!
For further information or to get in contact with the concerned parties, you will find at the end of each story the contact details of the providers of the story.

CHAPTER 2

THE NATIONAL COMPETENCE CENTRES

THE NCCs

– NCC Austria

EuroCC Austria supports businesses, research institutions, and the public sector in leveraging HPC, HPDA, and AI. The NCC facilitates access to the EuroHPC infrastructure and national HPC systems operated by Austrian Scientific Computing (ASC).

– NCC Belgium

EuroCC Belgium is Belgium's NCC in the area of High-Performance Computing (HPC) and High-Performance Data Analytics (HPDA). It coordinates activities in all HPC-related fields and serves as a reference contact point on HPC/HPDA/AI at a national level.
EuroCC Belgium consortium and partners: Cenaero, CECI, and Vlaams Supercomputer Centrum (VSC).

– NCC Cyprus

The Cyprus NCC is led and coordinated by the Computation-based Science and Technology Research Center (CaSToRC), at the Cyprus Institute. CaSToRC pioneers the development of computational and large-scale data methodologies to advance scientific and technological disciplines. In parallel, it supports user communities in academia, industry and government in Cyprus to use HPC, advanced mathematical modelling, simulation, data science and HPDA for computational research and innovation.

– NCC Denmark

EuroCC Denmark, anchored at the Danish e-Infrastructure Consortium (DeiC) is a collaboration with the Universities of Aalborg, Aarhus, Copenhagen, and DTU. The NCC focuses on developing competencies in HPC, AI, and High-Performance Data Analytics (HPDA) for businesses, public administration, and the research community.
The goal is to provide access to supercomputing resources, offer training, and support the deployment of HPC applications. This initiative helps bridge the gap between industry and academia, fostering innovation and technological advancement in Denmark and Europe.

– NCC Germany

»Supercomputing in Deutschland (SIDE)«, the German NCC, strengthens High-Performance Computing (HPC), Artificial Intelligence (AI), and big data competence in Germany. Formed of the High-Performance Computing Center Stuttgart (HLRS), Jülich Supercomputing Centre (JSC), Leibnitz Supercomputing Centre (LRZ), Gauss Centre for Supercomputing e.V. (GCS), and Sicos BW, it supports industry, academia, and public institutions in utilizing these technologies, promotes knowledge transfer, and enhances European digital competitiveness.

– NCC Estonia

NCC Estonia coordinates HPC expertise at the national level. Its mission is to analyse, implement and coordinate all necessary activities and offer services to end users to cover their needs, from access to resources and technological consultancy to providing training courses for academia, public administrations and industry.

– NCC France

The French NCC: CC-FR, dedicated to HPC, HPDA, AI technologies and Quantum computing, brings together the community of technology providers and users. CC-FR federates the HPC, HPDA and AI ecosystem and supports SMEs on the use of intensive computing, High-Performance Data Analysis and Artificial Intelligence.

– NCC Hungary

The HPC Competence Centre (HPC CC) was established in 2020 and is operated by the Digital Government Development and Project Management Company (Digitális Kormányzati Fejlesztés és Projektmenedzsemnt Kft., DKF). It promotes the High Performance Computing (HPC) and introduces the HPC infrastructure to potential academic and industrial users. Its experts assess the incoming project applications and manage the appropriate allocation of national HPC capacities. Additionally, the Competence Centre maintains contact with users and international professional partners and represents the Hungarian HPC infrastructure and community in international professional forums.

– NCC Iceland

The EuroCC-supported NCC Icelandic High-Performance Computing (IHPC) serves as the hub for Iceland's HPC community, which includes experts in Artificial Intelligence (AI). IHPC's initiatives unite Iceland's national specialists and link with a European network of NCCs focused on HPC and AI. This collaboration offers a diverse range of services tailored to meet the unique requirements of industry, academia, and public sectors in Iceland.

– NCC Montenegro

EuroCC Montenegro, established at the University of Donja Gorica through the EU Horizon 2020 EuroCC project via EuroHPC JU, serves as the NCC for High-Performance Computing (HPC). Further supported by EuroCC 2 and EuroCC4SEE (DIGITAL EUROPE) and in partnership with the University of Montenegro, the NCC coordinates national HPC/HPDA/AI activities, acting as a central hub for industry, academia, and the public sector, fostering HPC competencies, expertise, access, and collaboration.

– NCC Netherlands

The NCC Netherlands aims at facilitating the knowledge sharing in the topics of HPC, HPDA and AI among all interested national parties (research institutes, SMEs, big industries, public administration and society in general), as well as to offer access to these competencies for other stakeholders at international level and be a driver for innovation. The NCC Netherlands acts as a starting point to get access to HPC, HPDA and AI resources in the Netherlands.

– NCC Norway

The Norwegian NCC is led by Sigma2, the national provider of large e-infrastructure services for research, with the independent research institutes NORCE and SINTEF as partners. The partnership within the Norwegian NCC is unique in how it combines the HPC competence – key for a competence centre – with domain competence provided by the research institutes, allowing the NCC to take on a larger portfolio of projects within most domains.

– NCC Portugal

The Portuguese NCC for high-performance computing (HPC) was established under the EuroCC project to advance and strengthen the computational skills of the Portuguese scientific community. Coordinated by the Portuguese National Funding Agency for Science, Research and Technology (FCT), the NCC brings together various organizations that contribute with their expertise and resources. Its primary aim is to make full use of the European High-Performance Computing Joint Undertaking (EuroHPC JU) resources, enhancing Portugal's role in cutting-edge computational research and innovation.

– NCC Slovenia

The Slovenian NCC operates within the Slovenian National Supercomputer Network – SLING which promotes the use of high-performance computers capabilities for research in science, industry, academia and the provision of public services. The most important task is raising the level of knowledge of users and increasing general awareness about the benefits of using high-performance computers.

– NCC Spain

The Spanish NCC provides high-performance computing (HPC), high-performance data analytics (HPDA), and Artificial Intelligence services tailored to the needs of industry (especially SMEs), Academia, and public administration in Spain. EuroCC activities in Spain are currently carried out through the collaborative efforts of five partners: BIFI, BSC, CESGA, SCAYLE, and the University of Cantabria. For more information, visit https://eurocc-spain.res.es/.

– NCC Sweden

ENCCS is the EuroCC NCC Sweden. It is hosted by RISE Research Institutes of Sweden and Linköping University. The centre is funded by the EuroHPC JU, the Swedish Research Council (Vetenskapsrådet) and the Swedish Innovation Agency (Vinnova).

– NCC Türkiye

Turkish National e-Science e-Infrastructure (TRUBA), operating under Turkish Academic Network and Information Center (TUBITAK ULAKBIM) is the coordinator of NCC Türkiye. Middle East Technical University (METU), Sabancı University (SU), and Istanbul Technical University National Center for High-Performance Computing (UHeM) are the third parties of the NCC. While METU is a public university based in Ankara, SU is a privately-funded university in Istanbul. ITU UHeM, also based in Istanbul, provides supercomputing and data storage services to academic and industrial users. The competencies include High-Performance Computing (HPC), High-Performance Data Analytics (HPDA), Artificial Intelligence (AI), CUDA, Materials Science, Computational Fluid Dynamics (CFD), and several other fields.

– NCC Finland

EuroCC Finland is part of the EuroCC 2 project, which establishes national HPC competence centres in different European countries. EuroCC Finland's mission is to support and improve the capabilities of Finnish business users to utilize the opportunities of High-Performance Computing, Data Analytics and Artificial Intelligence. Through EuroCC Finland, companies have access to the computing capacity of the EuroHPC LUMI Supercomputer. EuroCC Finland is operated by the CSC – IT Center for Science.

Overview
NCC Success Stories

Section 3
Energy – 56

Section 4
Manufacturing and Engineering – 62

Section 5
Text- and Speech-based Services – 88

Section 6
Health Care – 110

Section 7
Other Fields – 128

Photo by **Malik Abdur rehaman** on **Unsplash**

Section 1
Environment, Climate and Weather

Improved Maritime Weather Forecast to Reduce Greenhouse Gas Emissions

NCC Austria

Industrial organisations involved:

The work was carried out following a strategic partnership between the two deep-tech startups AlongRoute Data I.K.E. and Neuralio A.I.
AlongRoute focusses on ultra-precise, AI-based marine forecasts for weather routing solutions. https://alongroute.com.
Neuralio A.I. is known for its expertise in creating advanced, custom AI solutions in the fields of Weather/Climate, Renewable Energy and Fintech. https://neuralio.ai

Technical/Scientific challenge:

Almost 3% of greenhouse gas emissions worldwide are attributable to maritime transportation, which is expected to quadruple in less than 30 years. An efficient way to cut emissions is to improve routing via early detection of adverse weather events. Traditional forecasting models cannot capture the complex oceanographic patterns, often resulting in inaccuracies, and lack the capacity to process extremely large datasets. The challenge was to improve the accuracy of oceanographic forecasts, specifically predicting Significant Wave Height (SWH) for optimized maritime operations.

Solution:

AlongRoute and Neuralio developed a hybrid Graph Neural Network-Gated Recurrent Unit (GNN-GRU) model, using Austria's supercomputer VSC-5. The AI model was trained on over 500 million data points and was able to generate reliable six-hour forecasts for SWH. It achieved a Mean Absolute Error (MAE) of 0.0071 and an R-squared (R^2) value of 0.98 against test data, demonstrating high accuracy and efficiency. Additionally, the Root Mean Squared Error (RMSE) of 0.015 further confirms the accuracy of the model's predictions. Two other models were trained, capable of predicting 12 respectively 24 hours ahead, but the 6-hour model was the most reliable. The model was trained on two NVIDIA A100 40GB GPUs attached to a host with 512GB of RAM, completing the process in approximately 13 hours and 50 minutes.

Business impact:

Ocean waves with a high SWH may submerge ships and destroy ocean or coastal infrastructure. It endangers human life, agriculture output, and the viability of aquaculture goods. As a result, precise forecasting of SWH is critical since it can assist in avoiding social and economic losses. Furthermore, SWH prediction can improve ship routes and might help to avoid rough seas, decreasing sailing time and fuel costs. As climate change is expected to influence oceanographic conditions, leading to more frequent and severe storms, the ability to accurately predict SWH will become increasingly important. With high-accuracy SWH predictions, maritime operations benefit from improved decision-making for safer, cost-effective routes. This predictive power is crucial for shipping companies aiming to reduce fuel consumption, avoid adverse weather, and enhance operational safety.

Benefits:

- Increased weather forecast reliability
- Reduced fuel consumption and greenhouse gas emissions
- Safer maritime navigation
- Reduced operational costs

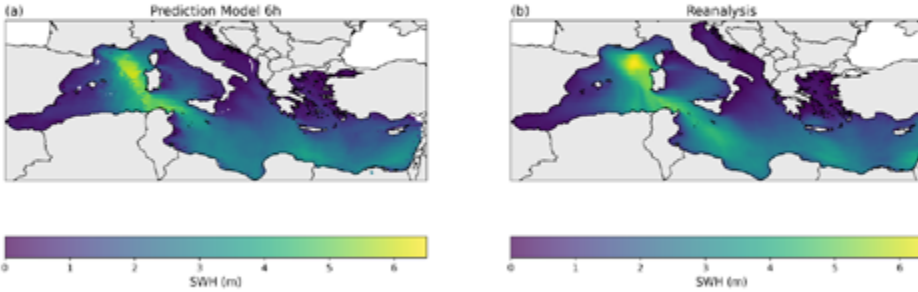


Figure 1: Model prediction on 14/01/2014
Figure 2: Reanalysis data of SWH for the same time

Keywords: Marine Weather Forecasting, Logistics, Significant Wave Height (SWH), Graph Neural Networks (GNNs), Gated Recurrent Units (GRUs), VSC-5
Industry Sector: Environment/climate/weather, Maritime
Technology: HPC, Deep Learning, Big Data

Contact:
Andreas Lindner,
andreas.lindner@advanced-computing.at

AI Model Achieves 97% Accuracy in Butterfly Species Detection

NCC Austria

Partners involved:

Viel-Falter Monitoring – Austria-wide butterfly monitoring programme managed by the Department of Ecology at the University of Innsbruck and financially supported by the Federal Ministry of Agriculture, Forestry, Environment and Water Management (BMLUK) as part of the Austrian Biodiversity Fund. In cooperation with the foundation Blühendes Österreich, an image-based butterfly recognition model was developed and evaluated.

Technical/Scientific challenge:

The project aimed to evaluate the accuracy of deep learning models in predicting butterfly species from images. The dataset, sourced from the Austrian Butterfly app (https://schmetterlingsapp.at) consists of over 500,000 expert-labeled images covering 180 species. Managing this vast dataset required substantial computational power, especially GPUs. The LEO5 HPC system at the University of Innsbruck was used, but initial experiments revealed lengthy training times when using only one GPU.

Solution:

EuroCC Austria recommended and implemented data parallelism, enabling multiple GPUs to train the model concurrently. On that base a Convolutional Neural Network was trained. Initial results on the LEO5 system were remarkable: training times per epoch dropped from two hours to under 12 minutes with four maximally utilized GPUs.

Further, after a row of tests on the LEO systems, EuroCC Austria established access to the LEONARDO supercomputer in Italy in order to leverage the parallelism full scope and perform extensive model and hyperparameter scans. A validation accuracy of 97 % for butterfly recognition was achieved.

Scientific impact:

The rapid growth of biodiversity data poses challenges for traditional species identification, which relies on time-consuming expert assessments. This project demonstrated that machine learning models can automate species recognition efficiently, reducing costs and time while improving data accessibility. Automated models allow species records to be analyzed promptly, offering feedback to app users, enhancing their learning, and boosting motivation.

Butterflies are important biodiversity indicators and improving the processing of related datasets supports their monitoring. In addition, it is possible to provide information on which species can be determined well and which are difficult to detect and might still need human expertise for determination. Different model architectures and data preparation steps were tested and the resulting insights might also be of interest for other species groups that can be determined from images.

The added value of the project lies primarily in testing various models, enabling the development of recommendations for practical applications. Additionally, the goal is to assess the overall effectiveness of the method, its performance for individual species, and its potential to reduce the workload of experts.

Benefits:

- Training time per epoch reduced from two hours to 12 minutes using four GPUs
- Accuracy of 97 % on the validation dataset
- Insights applicable to other image-based species recognition projects

Figure 1:
© Susanne Barkmann
The Small Copper (Lycaena phlaeas) has adapted its lifestyle to various environments, making it a true multi-talent when it comes to thriving in different conditions. As a result, it is one of the butterflies most frequently photographed by citizen scientists.

Figure 2:
© Susanne Barkmann
The Holly Blue (Celastrina argiolus) is a synanthrope, making it relatively easy to photograph as it is often seen in gardens.

Keywords: Biodiversity Monitoring, Machine Learning, Convolutional Neural Networks, Image detection, Citizen Science
Industry Sector: Ecology, Environmental Monitoring
Technology: HPC, HPDA, AI

Contact:
Andreas Lindner, andreas.lindner@advanced-computing.at
Friederike Barkmann, friederike.barkmann@uibk.ac.at

Mapping thermal buffering capacity of European forests

NCC Belgium

Scientific partners involved:

KU Leuven is a Belgian university founded in 1425, which conducts both fundamental and applied research. It is listed at #45 in THE World University Ranking and aims to actively participate in public and cultural debate while advancing a knowledge-based society. The university engages in public and cultural debate, promoting a knowledge-based society through its expertise, with operations spanning campuses, research parks, and hospitals.

The research lab sGlobe of KU Leuven aims to improve our understanding of the effects of global change on biodiversity and the functioning of terrestrial ecosystems. The lab, led by Prof. Koenraad Van Meerbeek, focuses on climate change and invasive species, two of the most important drivers of biodiversity loss.

Technical/Scientific challenge:

Scientists studying the environment frequently rely on macroclimatic data derived from standardised measurements in weather stations. However, these weather stations are designed to avoid the influence of external variables such as vegetation cover or topography on these measurements. They miss out on what is really important for many species living close to the ground, also known as the microclimate. A good example of this discrepancy is the temperature difference within and outside a forest during a scorching summer heatwave. This temperature difference is also called thermal buffering. sGlobe has developed new data that accounts for the thermal differences inside and outside forest ecosystems.

To study microclimate ecology and global change effects, sGlobe combines big data with state-of-the-art techniques like machine learning and species distribution modelling (SDM) to extract patterns from the data and answer ecological questions on large spatial scales (ecoinformatics).

According to Dr. Stef Haesen, postdoc at sGlobe, KU Leuven: *"Through modelling, we aimed to predict the thermal buffering capacity of European forests at a high spatial resolution of 25 x 25 m². This means that our study area consisted of more than 3 billion pixels, of which the monthly microclimate temperature had to be predicted individually. Reaching this objective without the computational capacity available at the Vlaams Supercomputer Centrum would have been impossible."*

Solution:

sGlobe used a combination of tiling and parallelisation within the cluster. Dr. Stef Haesen (postdoc at sGlobe, KU Leuven) highlighted: *"Europe is so big that trying to predict everything at once just isn't practical. To manage this, we divided the area into approximately 1,400 smaller, manageable tiles. We then used the nodes in the Genius cluster—specifically, its 36 cores allowed us to process 36 tiles concurrently. As each tile was completed, we immediately commenced processing on the next, maintaining continuous parallel computation. This approach significantly accelerated our analysis."*

Scientific impact:

Dr. Haesens's research generates ecologically relevant microclimate data used by ecologists in Belgium and Europe to gain better insights into the distribution of forest plant species. Ultimately, this approach allows for a more accurate assessment of the impact of climate change on these species.

Benefits:

- Ecologically relevant microclimate data
- Available at high spatial resolution
- Data openly available for the scientific community

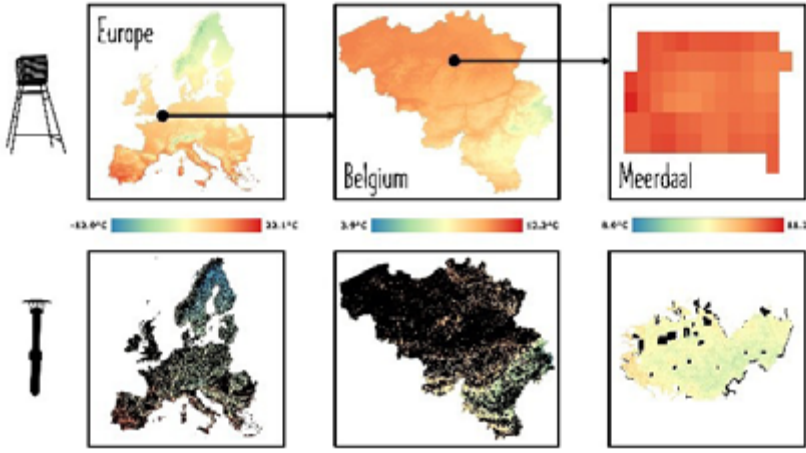


Figure 1: Image showing the difference between traditional climate data (depicted on the top line) and the newly modelled microclimate data (bottom line) by sGlobe, which is more detailed.
Web version: https://www.enccb.be/thermalbuffering

Keywords: Climate change, Forests, Thermal buffering, Microclimate, Machine Learning

Industry Sector: Environmental sciences/climate/ecology

Technology: HPC, Machine Learning

Contact:

contact@enccb.be

SAPIEN: Skillful Atmospheric Predictions with IntelligEnt Networks

NCC Denmark

Industrial organisations involved:

The Danish Meteorological Institute (DMI) is Denmark's official weather and climate service, providing reliable meteorological information and forecasts. Established in 1990, the DMI is responsible for weather warnings, climate monitoring, and oceanographic data. It supports public safety, environmental protection, and research by providing accurate weather predictions, climate analyses, and atmospheric observations for Denmark, Greenland, and the Faroe Islands.

Danish Meteorological Institute

Technical/Scientific challenge:

Traditional numerical weather prediction (NWP) models require extensive computational resources, often running on supercomputers for several hours. This leads to outdated forecasts with limited spatial and temporal resolution. Furthermore, the high energy consumption of these models raises sustainability concerns.

Nowcasting is fundamental for short-term weather predictions, particularly in rapidly changing conditions such as severe storms, heavy rainfall, and sudden temperature shifts. High-frequency, high-resolution nowcasts enable critical decision-making in industries like aviation, renewable energy, and emergency management, where even a few minutes of advanced warning can make a significant difference. The challenge was to develop an AI-based nowcasting solution that delivers accurate, high-resolution forecasts in near real-time while reducing computational costs.

Solution:

DMI gained access to the LUMI GPU-powered supercomputer to train two machine learning models — LDCast and SHADECast (Leinonen et al. (2023) and Carpentieri et al. (2025)) — for rainfall and solar radiation nowcasting. These models use deep learning architectures trained on weather observations (radar rainfall imagery and satellite cloud imagery), capturing complex atmospheric patterns based on observations.

By using AI-driven techniques, these models can generate predictions in minutes, significantly improving forecast timeliness and spatial precision. Unlike conventional approaches, which rely on physics-based simulations with high computational costs, these machine learning models learn from historical and real-time data to make rapid, data-driven forecasts. This allows for dynamic updates as new observations become available, ensuring that predictions remain as accurate and relevant as possible.

Business impact:

National weather services traditionally rely on high-performance computing (HPC) resources optimized for numerical weather prediction (NWP), using in house large-scale supercomputers dedicated to physics-based simulations. However, these institutions typically lack access to the specialized GPU infrastructure required for emerging AI-based forecasting techniques. This gap presents a major challenge, as AI-driven forecasting and nowcasting methods require high-speed parallel processing for training, which standard CPU-based HPC architectures are not designed to support.

Through EuroCC2, access to GPU-powered supercomputing resources, such as LUMI, has been fundamental in allowing the DMI weather models team to explore and develop machine learning approaches for weather nowcasting. Without this kind of access, the transition toward AI-enhanced meteorology would be significantly slower, limiting the ability compete in this rapidly evolving field. The availability of GPU resources is enabling the DMI team to experiment with deep learning architectures, train high-resolution models efficiently, and validate their performance against traditional nowcasting and forecasting techniques.

By integrating AI-based nowcasting into operational workflows, national weather services and private industry partners can achieve more rapid and precise weather predictions, benefiting sectors such as renewable energy, agriculture, and disaster management. The project also strengthens Europe's leadership in AI-driven meteorology, fostering new collaborations, funding opportunities, and industry adoption of next-generation forecasting technologies.

Benefits:

- Near real-time weather forecasting, reducing latency from hours to minutes
- Lower computational costs and improved energy efficiency compared to traditional NWP models
- Enhanced accuracy and spatial resolution

Figure 1: Snapshot of rainfall field from DMI radar imagery, the dataset which will be used to train one of the AI-based now-casting models for rainfall. www.dmi.dk
© DeiC

Keywords: Weather forecasting, Nowcasting, Meteorology, AI, Radar, Satellite

Industry Sector: Environment/climate/weather

Technology: Data-driven weather models (HPC, AI)

Contact:

marta.schulze@deic.dk

dennis.wollbrink@deic.dk

ncc-dk@deic.dk

webDO3SE: An Unprecedented Tool for Measuring the Global Impact of Tropospheric Ozone on Plant Life

NCC Germany

Scientific partners involved:

The Tropospheric Ozone Assessment Report (TOAR) is an international activity under the International Global Atmospheric Chemistry project, which aims to assess the global distribution and trends of tropospheric ozone and to provide data that are useful for the analysis of ozone impacts on health, vegetation, and climate. The TOAR data centre provides access to the TOAR database, which compiles air quality monitoring data from thousands of sites around the world.
https://toar-data.org/

Technical/Scientific challenge:

Ozone, a protective compound in earth's stratosphere, also occurs in the lowest layer of our atmosphere, the troposphere. Tropospheric ozone, though, is harmful to the climate, human health, and vegetation by reducing biomass, harvests, and biodiversity. Quantifying its global uptake by plants is crucial but challenging because it requires numerical modelling that correlates fine-resolution ozone measurements, plant specifics, and meteorological parameters, which are only available for individual sites and limited time periods, making quantification impractical and meaninglessly fragmented.

Solution:

The TOAR data infrastructure solves the problem of fragmented data. It couples the TOAR database, which is hosted at the Jülich Supercomputing Centre (JSC) clouds and contains one of the world's largest collections of ground-based ozone measurements, with the meteocloud, a collection of meteorological data also at JSC, as inputs to the ozone deposition model DO3SE. The workflow of webDO3SE is shown in *Figure 1*. Via a web interface, users can easily access the data and perform impact estimates themselves in their browser. They select a site and a species by REST query to webDO3SE. Then, all the parameters and inputs needed to run the model are automatically gathered on-line. The model output is then provided to the user directly in the browser for further analysis.

Scientific impact:

As climate change progresses, the greenhouse gases and compounds that beget tropospheric ozone will also likely increase. Historically, studying the far-ranging impacts of tropospheric ozone on food security, carbon sequestration, timber production, and protection against soil erosion, avalanches, and flooding has been hindered by fragmented access to data. Nevertheless, HPC infrastructure combined with TOAR data mitigates that fragmentation and enables the scaling of scientific studies to support public entities.

The TOAR tool, webDO3SE, is currently being used for a global deposition model intercomparison study and will in the future provide an unprecedently global assessment of ozone deposition of vegetation – the extent to which plants absorb and mitigate tropospheric ozone. Once webDO3SE is well established, it can also be used by environmental agencies to estimate tropospheric ozone-related crop damage and other negative impacts on vegetation. Over time the impacts of tropospheric ozone as well as any mitigation policies can be measured, tracked, and more accurately modeled to maximize future environmental and societal benefit.

Benefits:

- TOAR data combined with HPC resources and the meteocloud facilitates access to an unprecedented scope of tropospheric ozone data
- webDO3SE provides a unique interface for researching tropospheric ozone impact on vegetation
- Studying tropospheric ozone at a global level empowers environmental agencies

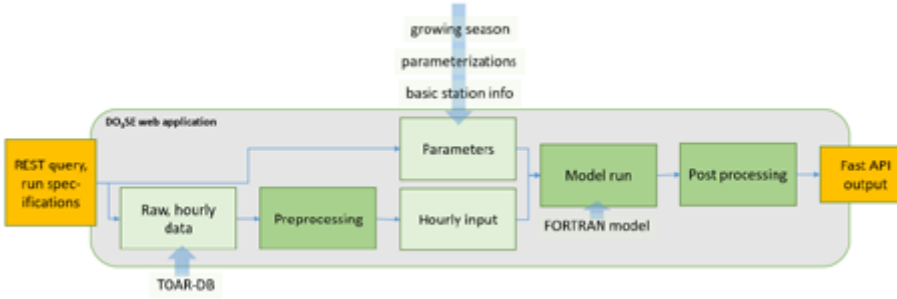


Figure 1: Scheme describing the productive workflow of webDO3SE, which is triggered by a user REST request. The model pipeline then runs automatically on JSC's cloud infrastructure and the results are displayed in the user's browser.

Keywords: Big data, HPC, HPDA, Climate, Sustainability
Industry Sector: Agriculture, Environment/climate/weather, Public services/Civil protection
Technology: HPC, HPDA

Contact:
contact@supercomputing-in.de

This project has received funding from the German Federal Ministry of Research, Technology and Space under grant agreement No 16HPC096K.
The sole responsibility for the content of this publication lies with the authors.

Forecasting the Future: How Compute Power Fuels Resilience and Innovation

NCC Netherlands

Industrial organisations involved:

The people at Beyond Weather are on a mission to revolutionize weather forecasting and to bring the future of climate insights to society. By combining cutting-edge AI with a deep understanding of meteorology, they are tackling some of the biggest challenges in predicting the weather.

Technical/Scientific challenge:

Imagine trying to predict the unpredictable — weather patterns beyond 14 days into the future. It's one of the toughest scientific challenges, requiring a blend of immense computing power, vast datasets, and innovative algorithms.

With the support of NCC Netherlands, the team of Beyond Weather started their journey into AI-based weather forecasting, taking crucial first steps toward unlocking new possibilities. The rise of Foundational Weather Models—a groundbreaking class of AI systems—is rewriting the rules of what's possible, creating opportunities to merge diverse data sources like satellite imagery, simulations, and real-time observations. But with such ambition comes a need for high-end GPUs to handle the sheer scale of computation.

Solution:

Rather than building a foundational weather model from scratch — a costly endeavor in a rapidly evolving field — Beyond Weather turned to open-source foundational models. These models, developed through collective innovation, are transforming industries, and they saw an opportunity to be able to always build on top of the state-of-the-art AI models.

With the compute credits and support of NCC Netherlands, they developed a unique approach: an efficient pipeline to fine-tune any model for predicting weather targets of interest. And they didn't stop there—the method goes beyond conventional fine-tuning on small datasets. Instead, Beyond Weather is integrating more and more datasets to create robust, accurate predictions that meet real-world needs.

This milestone wouldn't have been possible without access to cutting-edge GPU resources, enabling them to explore the frontiers of AI and weather forecasting. It's an exciting first step in a journey to reshape how we understand and respond to weather's complexities.

Business impact:

Beyond Weather envisions a future where accurate long-range weather forecasts mitigate energy poverty, stabilize grids, and safeguard communities. As the energy sector increasingly relies on weather-dependent renewables, balancing supply and demand becomes a growing challenge, especially during extreme weather events. The AI-driven forecasts of Beyond Weather enables utilities to plan ahead, ensuring energy reserves meet demand and preventing blackouts that could leave millions in the cold.

The Beyond Weather technology supports energy traders and utilities to make proactive decisions, reducing financial losses while maintaining affordable energy prices. Forecasting weeks in advance helps stabilizing energy markets, mitigating the impact of price volatility on households. This fosters economic resilience and shields vulnerable communities from the cascading effects of energy crises.

Beyond the energy sector, these forecasts pave the way for smarter resource management, from optimizing water usage during droughts to protecting food supplies in agriculture. With every forecast, Beyond Weather aims to take another step towards building a more resilient and sustainable future for all.

Benefits:

- Accelerated AI development, saving training time and thereby reducing costs
- Reduced entry barriers, allowing rapid prototyping and product-market alignment.
- Processed 5TB of data & counting

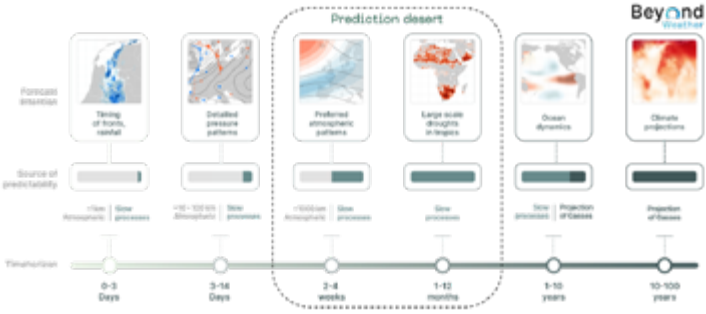


Figure 1: The challenge, predicting weather beyond 2 weeks is notoriously difficult. © 2024 Beyond Weather

Figure 2: Beyond Weather team. © 2024 Beyond Weather

- Keywords: Long range weather forecast, foundational weather models, finetuning, multimodality
- Industry Sector: Earth science, Energy, Environment/climate/weather, Agriculture, services, Software providers
- Technology: HPC, AI, High-end GPU

Contact:

Sem.vijverberg@beyond-weather.com

HPC and innovating technologies to mitigate hurricane impacts

NCC Norway

Industrial organisations involved:

OceanTherm develops technology that aims to mitigate the impacts of hurricanes. Bubble curtains have been a proven technology to keep ice out of Norwegian fjords by bringing warmer water with higher salinity up to the water surface through buoyancy introduced by pipes submerged in the fjord. The pipes provide a curtain of bubbles when air is pumped into the pipe and released through small holes.
OceanTherm is further developing this technology in tropical regions, to bring colder water to the surface, lowering the surface temperature to remove or reduce the energy available to the hurricanes passing over the surface.

Technical/Scientific challenge:

High surface water temperatures in tropical areas are the key source from which tropical storms absorb energy and grow into hurricanes of increasing magnitude, causing tremendous damage and taking many lives each year.
Deeper down in the ocean the water temperature is cooler. OceanTherm's technology can lift cold water to the sea surface, cooling it and thus reducing the energy available to a hurricane, preventing it from building its strength. Surface temperatures can hence be lowered by deploying bubble curtains, which bring colder water to the surface and reduce the hurricane's energy. Predicting a hurricane's path and where to deploy a bubble curtain, requires significant computational power, utilized in time to execute counter measurements. The large-scale nature of the problem makes it helpful to deploy workflows that utilise HPC to answer questions related to system design, performance and, ultimately, operation and preparedness.

Solution:

Building an automated workflow that downloads all the current datasets, execute each of the steps in sequence on HPC to ensure correct input data and provenance can be achieved and reduce time-to-solution and time to execute.

Business impact:

- Enabling execution of the business idea

Benefits:

- Speeding up time-to-solution
- Ensuring provenance and traceability
- Enabling the business idea
- Portability

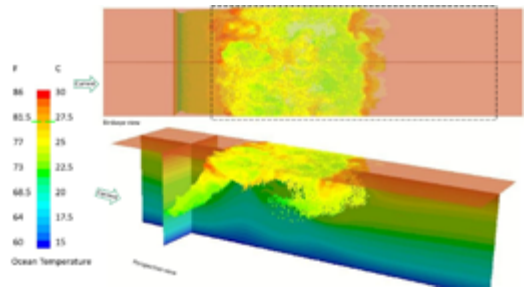


Figure 1: Simulating the possibility of changing the surface temperature
© Oceantherm

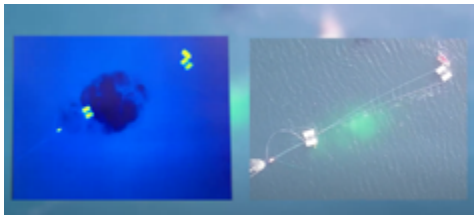


Figure 2: Sea trials to validate simulations

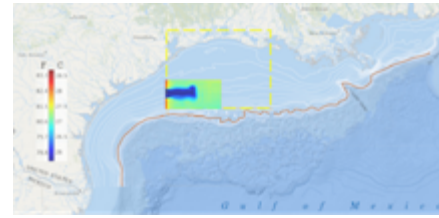


Figure 3: Simulating with real world data

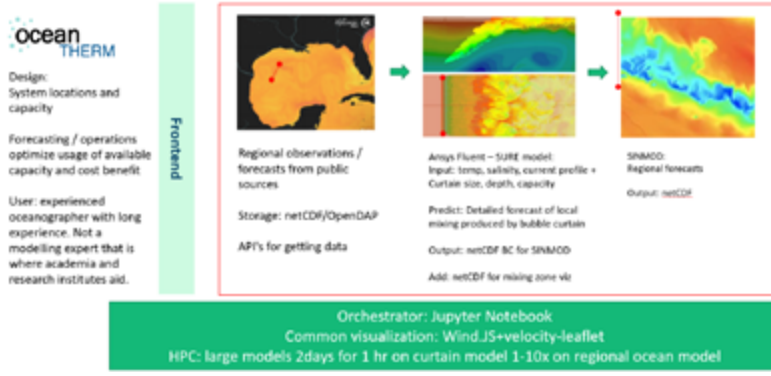


Figure 4:
The workflow

Keywords: HPC, Bubble curtain, CFD, automation, maritime, weather, multiphase flow

Industry Sector: Earth science, Environment/climate/weather, IT/HPC systems, services & software providers, Maritime, Public services/Civil protection

Technology: HPC, CFD, Workflow tools

Contact:
Paal Skjetne,
Paal.Skjetne@sintef.no

This project is co-funded by Norwegian research council through project ES723150.

Use of AI for satellite-based surface wind correction (NOW SYSTEMS)

NCC Spain

Industrial organisations involved:

Nologin Oceanic Weather Systems (NOW Systems) is an EC Copernicus Marine Service provider for the European Atlantic façade. Since 2018 they co-lead the IBI-MFC (Iberia-Biscay-Ireland Monitoring & Forecasting Centre) with Mercator Ocean International, delivering regional forecasts in collaboration with Meteo-France and CESGA. NOW Systems also became part (since 2015) of the Copernicus Marine INS-TAC (Thematic Assembly Center for in-situ data), with responsibilities (co-shared with Puertos del Estado) in the operation, validation and service evolution of in-situ observational products for this IBI area.

Technical/Scientific challenge:

This collaboration arises from the need to know the feasibility of using Artificial Intelligence techniques to improve the resolution and correction of winds generated by regional atmospheric models through the use of SAR satellite remote observation data, using the European Copernicus service.

Solution:

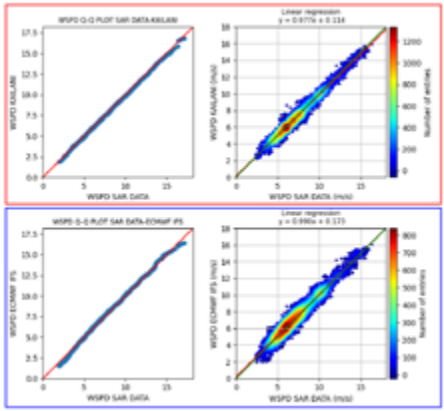
To this end, the study of various neural networks used in image enhancement for use in surface wind fields was carried out. These networks were modified by including both physical constraints and layers within the neural network itself as well as physical equations in the loss functions during training.
As a result, a neural network capable of generating a wind field at a higher resolution than the original model has been developed and trained, including the patterns observed in the SAR, and which can be generalized to any area of the planet.

Business impact:

- The HPC power optimizes and speeds up the preprocessing of the input data. Significant time reduction by an order of magnitude.
- The capacity of the GPU nodes allows training of large batches, minimizing errors during AI training.
- The processing speed of the GPU considerably reduces the training time of the different tests until the optimal architecture is found.

Benefits:

- High availability of nodes: allows to optimize the time spent on neural network development, ensuring that several experiments can be queued.
- Faster training: thanks to GPU-enabled nodes, which allows optimizing the architecture and obtaining more accurate results that improve coastal circulation and wave models.
- The use of HPC in all phases of the process reduces execution times and therefore optimizes R&D development costs.



Galicia (NE Atlantic) [GAL-D]

Validation period:
Sep. 2021 – Jan 2022

Name	SI [%]	RMSD [m/s]	Corr [-]
ECMWF	34.1	0.57	0.97
KAILANI	33.0	0.36	0.99

Figure 1: Comparison of metrics between the artificial intelligence model and the dynamic model (ECMWF) in the area of Galicia.

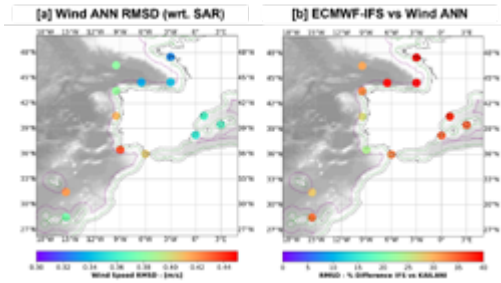


Figure 2: Performance of the AI wind model in the NE Atlantic and NW Mediterranean.
Left: RMSD between artificial intelligence and SAR data.
Right: Differences between AI and wind model.

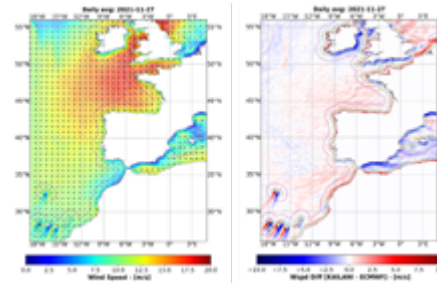


Figure 3: Performance of artificial intelligence in extreme events.
Left: Daily mean wind speed in the storm of 27 November 2021, preceded with Artificial Intelligence.
Right: Differences between Artificial Intelligence prediction and ECMWF dynamic model.

Keywords: Wind modelling, neural network, simulation, spatial resolution, satellite data

Industry Sector: Earth science

Technology: AI, HPC

Contact:

eurocc@res.es

Analysing Air Pollution Flow

NCC Sweden

Industrial organisations involved:

Stockholms luft- och bulleranalys (SLB-analys) is a unit at the Environment and Health Administration of the City of Stockholm. The unit is responsible for monitoring outdoor air quality in the city. SLB-analys also runs the regional system of air quality monitoring on behalf of the East Sweden Air Quality Management Association (Östra Sveriges Luftvårdsförbund) and assists its municipalities with services such as measurements and model calculations.

Technical/Scientific challenge:

Air pollutants of interest in urban environments are typically aerosol particles, measured as the mass of particles smaller than 2.5 (PM2.5) or 10 micrometer (PM10), and nitrogen oxides (NO, NO2). The dispersion of these (and other) pollutants highly depends on meteorological parameters, in particular air flows. Computational Fluid Dynamics (CFD) program packages, such as OpenFOAM and MISKAM, are used to simulate wind fields for wind comfort simulations and dispersion studies of air pollutants in complex urban environments. SLB-analys intends to run scaling tests on EuroHPC resources to push the boundaries of CFD simulations using OpenFOAM to larger spatial domains and higher complexity. Larger domains and higher-level turbulence models will be more computationally demanding which motivates this application for HPC resources.

Solution:

In this project, SLB-analys intends to perform sensitivity and scaling tests using EuroHPC resources. First, they will perform benchmark tests with much larger spatial domains of urban areas. Current computing resources either do not allow them to perform such simulations with large spatial domains, or it takes prohibitively long time to perform such simulations. Thus, they aim at performing these tasks on EuroHPC resources, which will improve the time-to-solution for such simulations. Depending on the performance, if necessary, they will make certain optimizations to ensure that all the simulations are carried out efficiently on the HPC cluster. Finally, SLB-analys will perform sensitivity tests with different turbulence models using the optimized version of OpenFOAM in order to improve the quality of the results on the dispersion of air pollutants.

Business impact:

The current challenge for SLB-analys is to perform simulations with large spatial domains and with complex turbulence models. The access to HPC resources, will firstly enable the ability to perform simulations with large spatial domains and with complex turbulence models. Second, by using more complex turbulence models, the accuracy of the results will be improved so that it will be possible to investigate whether new developments will meet air quality limits and propose measures to improve air quality in sensitive areas. Thirdly, using these simulation methods will allow them to investigate the role of complex urban structures (tunnels, bridges, etc.) in air pollution dispersion, ultimately aiming at mitigating health and environmental impacts.

Benefits:

- Simulating larger urban area becomes possible
- Time-to-solution is greatly reduced
- Better results used for air quality assessment

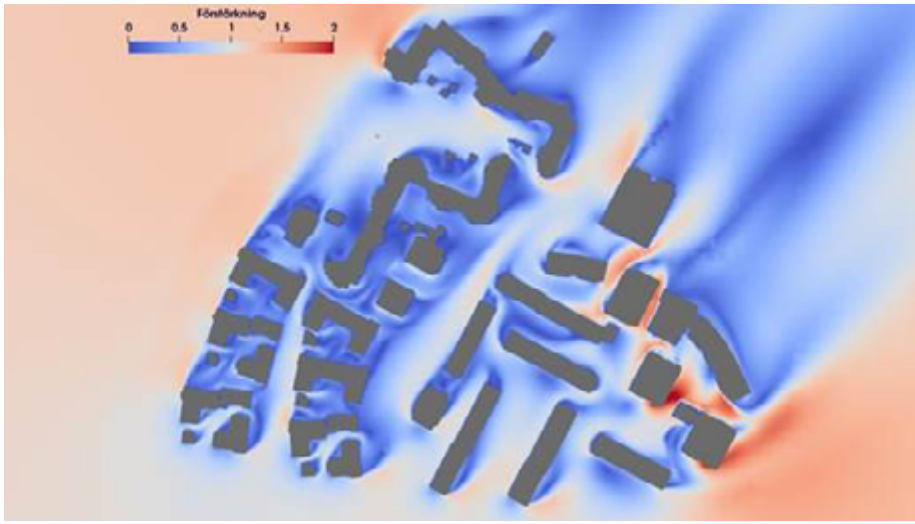


Figure 1: Example of a simulated air flow in an urban environment using OpenFOAM (method: RANS/k-epsilon). The amplification factor (color-coded) shows the wind velocity at two meters above ground level relative to a situation with wind over the same terrain without buildings, for wind from south-west.

Keywords: Air quality, Computational fluid dynamics, Environment, OpenFOAM, MISKAM

Industry Sector: Environment/climate/weather

Technology: HPC

Contact:

Apostolos Vasileiadis,

apostolos.vasileiadis@ri.se

Photo by **no one cares** on **Unsplash**

Section 2
Agriculture and Food

A DNA test to detect counterfeit honey

NCC Estonia

Industrial organisations involved:

Celvia is an Estonian company and Research Institution owned by the University of Tartu, the Tartu University Hospital and the Nova Vita Clinic, among others. Celvia's mission is to develop and deliver innovative science-based products and services for patients, clinicians and businesses. The company combines new knowledge with state-of-the-art technology and unique skills, resulting in competitive genetic tests that meet high standards. They collaborate with researchers and clinicians from Estonia and abroad to ensure the high quality of basic research.

Technical/Scientific challenge:

Honey contains approximately 0.01% DNA, which can be extracted and sequenced to produce up to 20 million DNA sequencing reads. These sequences can be computationally analysed using a taxonomic classifier to get insights into the biological composition of honey, aid in monitoring honeybee pathogens, and assess authenticity. However, this computational analysis is resource-intensive as the taxonomic classifier alone relies on a database of approximately 400 GB in size, and the analysis involves multiple steps. The company needs to process a high volume of samples on a weekly basis.

Solution:

Using HPC, the company processes 32 samples in parallel through a singularity containerized Nextflow workflow comprising seven subtasks. The most resource-intensive subtask demands 400 GB of RAM and 8 CPUs per sample. To streamline operations, they have automated the detection of new data, downloading, and initiation of analysis using crontab (*crontab is an operating system configuration that defines a schedule for running automated tasks at specified times*). Additionally, HPC resources are utilized to securely store raw data and analysis results, ensuring data integrity and confidentiality with backup systems in place.

Business impact:

Honey is one of the most commonly counterfeited foods, creating a significant global challenge. Counterfeit honey is produced at a fraction of the cost of authentic honey, driving down market prices and threatening the livelihoods of honest beekeepers. To address this issue, HPC solutions offer the computational power necessary to analyse large volumes of honey samples on a weekly basis. These analyses help combat honey fraud by verifying authenticity and providing valuable taxonomic insights, enabling beekeepers to better understand and protect the quality of their honey and monitor honeybee health.

Benefits:

- HPC computational resources enable rapid processing of large volumes of weekly samples for honey analysis service, reducing results reporting time
- Automation of the analysis process eliminates the need for specialists, allowing office staff to simply download results when notified
- With petabytes of secure storage, HPC ensures safe data storage, including backups and access restrictions

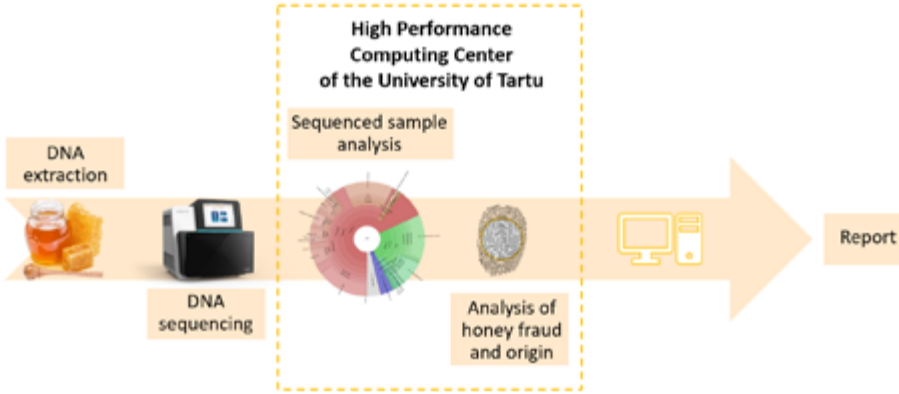


Figure 1: Honey analysis as a service: steps from honey sample to report.

Keywords: Poultry management, HPC, predictive analytics, resource optimization, data-driven forecasting, market alignment

Industry Sector: Agriculture, Food and Drink

Technology: HPC, HPDA

Contact:
Ülar Allas,
ylar.allas@ut.ee

Using AI and HPC to forecast fish location and quality, maximizing value and reducing waste

NCC Iceland

Industrial organisations involved:

GreenFish is a software company using artificial intelligence, big data analysis and HPC to predict the location, quality and quantity of pelagic and groundfish species with an eight-day forecast period. The forecasts reduce vessel time at sea, reducing costs and carbon emissions, while maximizing catch value and minimizing bycatch. With the support of IHPC, the Icelandic SME GreenFish received access to the EuroHPC JU system DEEP at JSC, Germany. GreenFish has been awarded the 2025 North Atlantic Seafood Innovation Awards.
https://greenfish.is/

Technical/Scientific challenge:

The fishing industry traditionally relies on conventional, manual methods to locate and choose fishing grounds. This approach is time-consuming, fuel-intensive and, due to its arbitrary nature, is unlikely to yield better results for years to come. In addition, understanding ocean and weather conditions plays a critical role in this process, and the human mind is not capable of taking them all into account when making decisions. GreenFish uses AI and HPC to model and predict fishing locations, integrating all relevant factors.

Solution:

GreenFish provides fishermen with reliable forecasts which predict the location of hard-to-find pelagic species alongside with the quality of demersal species (e. size, fillet yield, number of parasites etc.). GreenFish also provides fishermen with the most accurate weather and oceanographic data available on the market. This allows fishermen to optimize their trips and maximize value extracted from the ocean.

Business impact:

GreenFish currently holds a vast amount of fishing and oceanographic data from various sources. The mission is to develop models for each species to be able to create reliable and useful fishing forecasts for the industry.

This would be impossible without the use of HPC, as the model development requires a vast amount of training with large data sets. The IHPC support to access EuroHPC JU systems allows GreenFish to train their models quickly and effectively, ensuring that fisheries can expect forecasts models that reduce vessel time at sea, lowering costs and carbon emissions, while maximizing catch value and minimizing bycatch soon.

Benefits:

- Saves GreenFish thousands of EUR per month in computing power rental
- Enables GreenFish to accelerate product development, ensuring continued industry support from the start with an early proof of concept
- Reduces startup costs
- Allowed GreenFish to be daring in their development, since there was no shortage of training time
- The use of HPC drastically improves the accuracy of our forecasts

Figure 1: Presentation of how GreenFish prediction assists the fishing industry

Keywords: Forecasts Fishing

Industry Sector: Environment/climate/weather, IT/HPC systems, Services & software providers, Maritime

Technology: AI, HPC

Contact:

Reza Hassanian, seh@hi.is

Hemanadhan Myneni, myneni@hi.is

Making regenerative agriculture accessible worldwide through AI-driven technology

NCC Netherlands

Industrial organisations involved:

Spatialise is taking a big step in making the agricultural sector more sustainable through the use of advanced AI technology. NCC Netherlands is delighted to support a young company that offers both cost-saving solutions for farmers, and makes a significant contribution to reducing the environmental footprint of agriculture. By building a digital soil analysis technology that makes it possible to accurately measure the organic carbon (SOC) and other essential nutrients such as nitrogen, phosphorus, and potassium (NPK) in the soil. This allows farmers to optimise their use of fertilisers, resulting in fewer emissions and cost savings.

SPATIALISE

Technical/Scientific challenge:

Spatialise is developing a satellite-powered monitoring tool that tracks soil nutrients in agricultural lands using machine learning. By training artificial neural networks on global satellite data, the tool predicts soil nutrient levels worldwide. To improve accuracy in specific regions, the global model is then fine-tuned with regional data. Training the global model requires massive amounts of data and large neural networks which makes the training computationally very expensive. Furthermore, fine tuning this model to perform well in all areas of interest includes many rounds of training, which also requires large computational ressources.

Solution:

Spatialise addresses this challenge by utilizing Snellius, the Dutch national supercomputer, to train and fine-tune artificial neural networks for soil nutrient prediction. The global model serves as a foundation and is refined through transfer learning with region-specific datasets. This process involves extensive hyperparameter tuning (i.e. *a process that fine-tunes the internal settings of neural networks to achieve better accuracy*), requiring multiple iterations to optimize performance across diverse environments. By leveraging Snellius' HPC capabilities and SLURM-based compute-job scheduling (i.e. a system tool used in supercomputers to efficiently allocate computing resources), Spatialise efficiently distributes workloads, accelerates training, and ensures scalability in handling large datasets and complex neural architectures.

Business impact:

Spatialise previously relied on Databricks on Azure, which provided a comprehensive framework with integrated pipelines but proved cost-prohibitive for large-scale computational workloads. By leveraging Snellius, Spatialise has significantly reduced computing costs while gaining access to high-performance resources for training and fine-tuning neural networks. The use of Snellius has improved hyperparameter tuning, leading to better model performance. Additionally, it has helped identify key areas for research focus, such as optimizing satellite-based features, improving training methods, and refining transfer learning strategies. This has accelerated development and provided valuable insights to enhance research efficiency.

Benefits:

- Cost savings: Reduced computational expenses compared to cloud-based solutions like Databricks on Azure
- Improved efficiency: By leveraging Snellius' HPC resources, Spatialise can rapidly optimize model performance through hyperparameter tuning which is a process that fine-tunes the internal settings of neural networks to achieve better accuracy. This accelerates training times and enhances predictive capabilities.
- Optimized research: Identified key areas for improving training, satellite features, and transfer learning
- Enhanced scalability: Enabled large-scale computations for more accurate and region-specific soil predictions

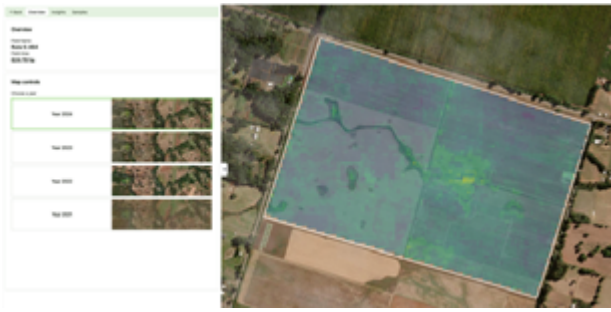


Figure 1: Soil Monitoring Platform of Spatialise
© 2024 Spatialise

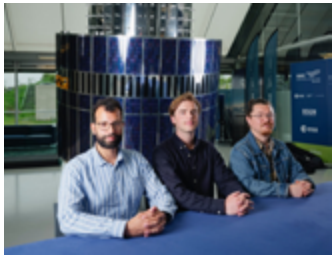


Figure 2: Spatialise team members
© 2024 Spatialise

Keywords: Satellites, artificial neural networks, transfer learning, hyperparameter tuning

Industry Sector: Agriculture, Space, Earth Observation Satellites

Technology: HPC, AI

Contact:
Frank Sloot,
frank@spatiali.se,
+31 6 52 42 51 35

Digifarm Revolutionises Agriculture with Artificial Intelligence

NCC Norway

Industrial organisations involved:

Digifarm is one of Norway's leading ag-tech start-up and has spent the last years developing the technology for automatically detecting the highest accuracy field boundaries, seeded acres and deforestation using deep neural network models and super-resolved EO data.
Through their research and development, DigiFarm has developed an AI model that can automatically detect field boundaries between fields. They train a deep neural network to identify boundaries and different elements in a field, such as grain, water, trees and deforestation. With 4 million hectares of training data from 57 countries, their model has become very large and requires significant computational power, or computational capacity, for training.

Technical/Scientific challenge:

Training neural network to detect and classify boundary layers requires a large and diverse dataset to produce an accurate model. Training a model with a large dataset to produce high-accuracy neural network is extremely resource intensive and not possible on a GPU (Graphical processing unit) typically used on private computers, and not even on a stand-alone data centre grade GPU. An AI workflow typically has multiple steps, from fetching the training data, preprocessing the data, training the model, postprocessing the results and uploading results. All these discreet steps have different requirements for computing resources, which is fine to run on an owned workstation. Workflows designed for workstations are not suitable to run on cloud or HPC (High-performance computing) resources, as shared resources should be released when not in use to avoid allocating expensive resources when not needed by the workflow. Splitting a serial workflow into discrete tasks with dependencies, allows a larger set of shared resources to process the workflows cost and more efficiently through more parallelization, spanning multiple physical components in a larger supercomputer.

Figure 1: Detecting field boundaries through AI
© Spatialise

Solution:

Migrating Digifarm´s training workload to HPC, and optimizing the workflow to run on parallel infrastructure, allowed for larger datasets used during training, and additional classifications to be added to the model. Furthermore, migrating the workload from national-level infrastructure to European-level infrastructure on LUMI, Europe's most powerful supercomputer, increases the ability to achieve astonishing accuracy when training models. LUMI is one of the world's leading AI platforms, equipped with some of the highest-performing data processors (GPUs) available on the market. Supercomputers like LUMI use highly advanced and expensive GPUs, which are far more powerful than those in home PCs. These GPUs provide dramatically faster calculations, essential for Machine Learning. The training process is significantly accelerated through access to large amounts of interconnected GPUs, and bigger models can be trained while saving time. Running the training process on LUMI further improves accuracy in the model, significantly surpassing previously achieved accuracy.

Business impact:

The results DigiFarm has achieved on LUMI are extraordinary. It will shorten the path from testing to a market-ready product by about 6 months, while increasing quality of their products.

"*LUMI represents a turning point for us and will undoubtedly greatly impact our ability to deliver innovative solutions for the agricultural sector. Access to LUMI allows us to accelerate development and significantly improve our model's precision, [...] We have improved the model's accuracy by 4.2% in just a few months. This is a major achievement in deep learning, as it becomes more difficult to achieve higher accuracy the longer a model is trained,*" says Nils Helset, founder and CEO of DigiFarm. Further improvements of accuracy has been achieved since then.

With these results, Digifarm was nominated for three awards in 2024, and was awarded the price for HPC Innovation Excellence by Hyperion research and HPC user group, and the HPCwire editor's choice award for Best HPC collaboration (Academia/Government/Industry). Digifarm also presented their results at the EuroHPC user days.

Benefits:

- Shorter time (about 6 months) to market through more compute capacity available
- Higher accuracy by 4.2% through abilities to train larger models
- Reduced cost trough optimized use of shared resources

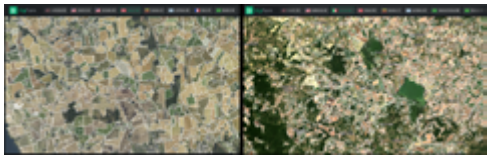


Figure 2: Fields do not look the same in Norway and Italy

Figure 3: Increased accuracy with larger training dataset, while reducing time due to larger parallel resources

Keywords: Accuracy, performance, cost, time-to-market, competitiveness

Industry Sector: Agriculture, Automotive, Raw materials

Technology: HPC, AI

Contact:

Roger Kvam,

roger.kvam@sigma2.no

This project is co-funded by Norwegian research council through project ES723150.

Time Development of Biomass Estimation using Acoustic Measurements (Sustainovate)

NCC Norway

Scientific partners involved:

Sustainovate is a European consultancy firm that promotes sustainable marine business development and innovation by bridging the gap between science, industry, and government. It advocates data-driven technologies for industrial fisheries, offering tools for autonomous data collection, processing, analysis, and collaboration using cloud computing, IoT, and AI. The firm specializes in managing complex data types like hydroacoustic and visual imagery data for this purpose.

Technical/Scientific challenge:

The main challenge is to demonstrate the added value of a fish-mapping-model in the process of estimating fish biomass based on acoustic recordings during commercial fisheries on North Sea herring. A continuously collected amount of 5+TB of acoustic data needs to be analysed in real time by complex tools. We investigate the feasibility to realize the complex analysis chains using a workflow engine, where possibilities for human interaction facilitating collaborations with third parties can be realized.

Solution:

Machine learning methods were employed to estimate fish biomass within a specific geographical area and time frame: fishing seasons in July-August for 7 years in the northern North Sea. Two spatio-temporal datasets were compared: acoustic measurements from trawlers and scientific surveys. The scientific survey data provided more homogeneous spatial coverage, while the industry data was biased towards fishing grounds.

Scienfitic impact:

The study develops a first glance on the idea of giving biomass estimates using statistical methods. Thus, we provide a comparison in terms of total estimated biomass for each year between scientific and commercial vessel observations. The timespan corresponds to fishing seasons (July–August) for 7 years in the northern North Sea. A python-based framework has been put in place to easily integrate more upcoming data.

The methodology developed in the project will be used to improve Sustainovate's scientific toolbox and expertise in the field of provision of annual fish biomass indices. It will increase the company's competitiveness in that sector and will contribute significantly to sustainable fisheries.

Benefits:

- Framework comparing Biomass information from Fishing fleet and Scientific survey
- Reusability
- Improved data analysis
- Gain of expertise in Biomass estimation in the context of Marine science

Keywords: HPC, Clustering methods, Fishing, Biomass estimation, North Sea herring

Industry Sector: Environment/climate/weather, Food and drink, Life sciences, Maritime

Technology: HPC, Clustering methods, Bootstrapping

Contact:

Klaus Johannsen,

kljo@norceresearch.no

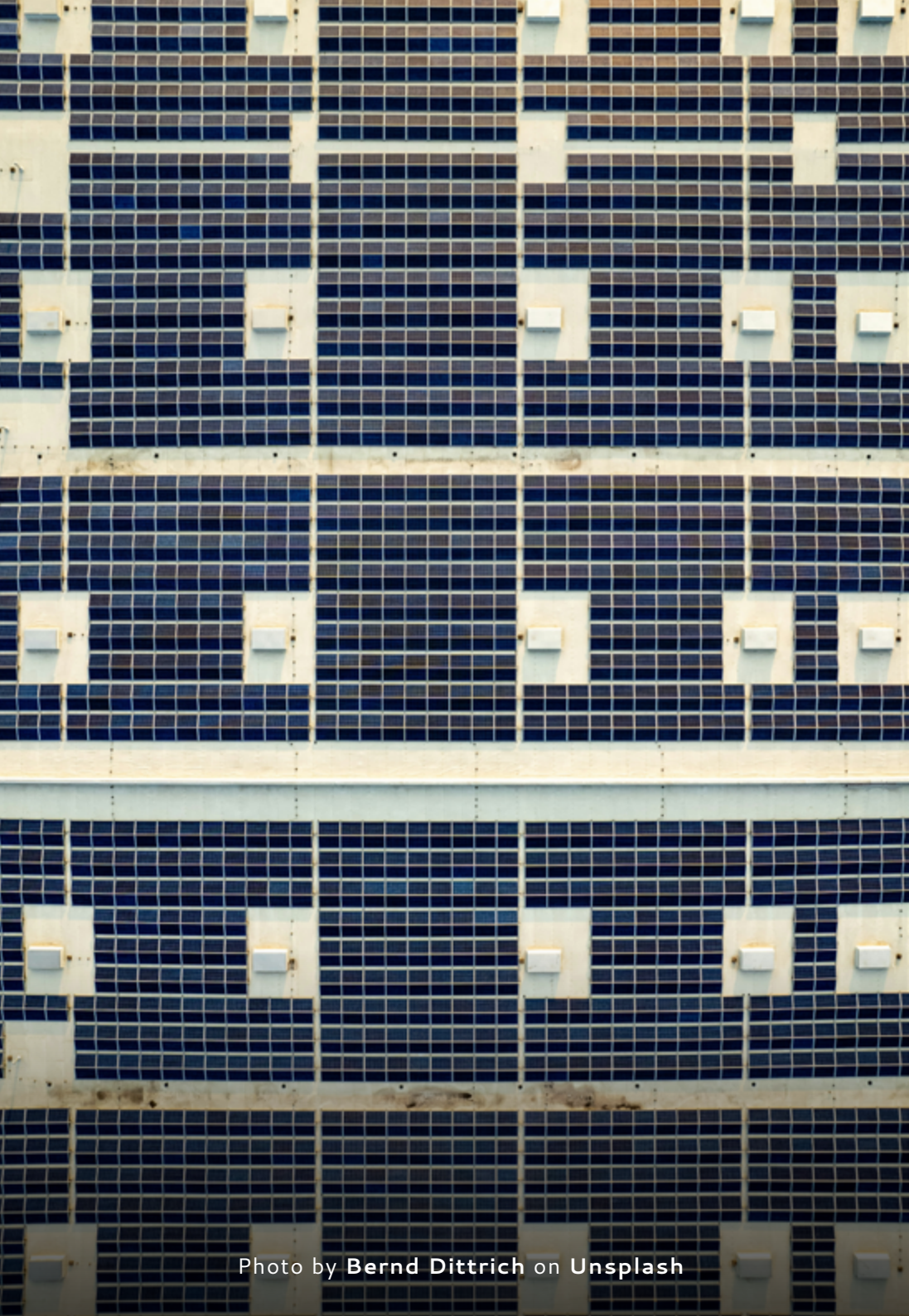


Photo by **Bernd Dittrich** on **Unsplash**

Section 3
Energy

Transition of power production simulation system to HPC

NCC Norway

Industrial organisations involved:

SINTEF Energy Research is an applied research institute dedicated to creating innovative energy solutions. The Institute specializes in research-based knowledge and infrastructure both in Norway and internationally, with the aim of providing its clients with solutions and services that increase value and strengthen their competitiveness. SINTEF Energy Research is a part of the SINTEF group, one of Europe's largest independent research institutes. SINTEF is a non-profit research foundation where no owners can withdraw dividends. Financial profits are invested in scientific equipment and expertise. Development of simulation models and tools for power production and planning have been a core business for SINTEF Energy Reserarch for more than 40 years. The tools offered are used by the majority of power marked actors in the nordics and baltics for production planning and engaging in the Nord Pool day-ahead and intraday trading system.

Technical/Scientific challenge:

The long legacy of simulation systems for the electrical power marked has resulted in a simulation system focused on single machine systems with transitions from VAX to Windows, and lately to Linux. The simulation system involves an advanced stocastic-dynamic model, encompassing a large linear program to be iterated for a multitude of production trajectories and price scenarios. Due to the nature of the Nord Pool power marked, the actors have a two-hour window to enact the simulations and prepare bids. This limits the complexity of the simulations that can be run in order to complete in time for being used for decision support. Although there has been an emphasis on using Message Passing Interface in later years to facilitate multi-node simulations, the full scale HPC context has not been explored. Platform- and compiler independence, build system and library structure was not prepared for being used in the typical HPC setting with module system (Lmod) and queue management (Slurm).

Solution:

A proof-of-concept was established by refactoring parts of the simulation system using CMake as a build automation tool. Adaptions were made to use the module system (LMOD) and job queue system (Slurm). This allowed testing and benchmarking for day-ahead production planning simulations for typical marked scenarios on the Betzy supercomputer at Sigma2. It also allowed investigation on how to scale up for more complex simulations with significantly higher granularity – a complexity that is prohibitive with the current business operations due to time and computation resource constraints.

Business impact:

Preliminary results show that simulations requiring typically two hours of simulation time can easily be reduced to less than two minutes on a HPC system. The CPU-efficiency when scaling to multi-node simulations for higher granularity and model complexity can reach a CPU-utilization of 60-80%. This scaling is dependent on the model complexity; more complex models benefit from higher number of nodes. Facilitating the transformation of power system simulation tools to use HPC can allow the industry to streamline the software tooling for production planning. It also can facilitate simulation models that yield more accurate results with higher granularity within the available time constraints of the Nord Pool's power marked bidding processes.

Benefits:

- Standardization of simulation code and build tools to use typical HPC context and queuing system
- HPC allows running more complex simulations with higher granularity
- Transition to HPC allows the power marked industry to off-load the computation to existing compute offerings

Figure 1: Day ahead trading market with time critical window of operation between 10:00 and 12:00. © SINTEF ENERGY

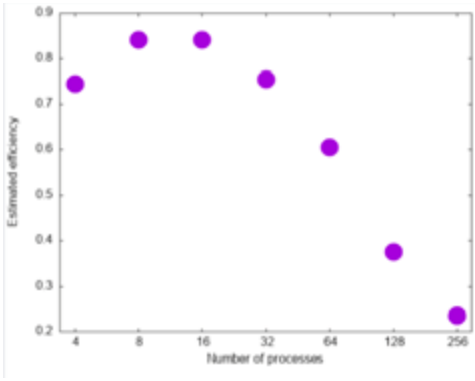


Figure 2: Scaling dependent on simulation size

Keywords: Energy, Power marked, Production planning, Nord Pool
Industry Sector: Power marked planning, Electricity production
Technology: HPC, AI

Contact:
Olaf Trygve Berglihn,
oberg@sigma2.no

This project is cofounded by Norwegian research council through project ES723150.

Efficiency of mechanical mixers in biogas digesters

NCC Türkiye

Industrial organisations involved:

EYS Makina has been producing machinery and equipment for the management and recycling of organic waste since 2001 and offers innovative solutions to its customers by following technological developments. It operates in the fields of animal origin fertilizer management, compost production facilities, biogas production facilities, recycling, solid waste disposal facilities, waste water treatment, and industrial separation applications. EYS is one of the the leading companies (https://e-y-s.com) in waste management, especially in developed countries such as the USA, Canada, Germany, Japan, England, Belgium, France, Switzerland, and China.

Technical/Scientific challenge:

In this project, EYS Makina is challenged to respond to a customer request in a very short amount of time. The mixing characteristics of large-scale renewable biogas energy plant reactors need to be analyzed before the construction and deployment of mixers. EYS R&D team rose to this challenge using open source CFD software and HPC resources, analyzed multiple scenarios within a short period of time, and continues to work on multiple different scenarios for EYS Metal's future needs.

Solution:

Analyses are carried out in different configurations using dive and tangential angle parameters. The simulations have to be carried out in three dimensions and unsteady using an accurate turbulence model. In this study, the large eddy simulation (LES) approach is followed to study three-dimensional turbulent flows. The most optimal results were obtained when the number of cells per processor was around 10,000. It is also noted that the number of cores per node is important, as it is appropriate to work with a full node from the cluster to achieve an optimal communication load. The flow is modeled as periodic because it contains periodic repetitions.

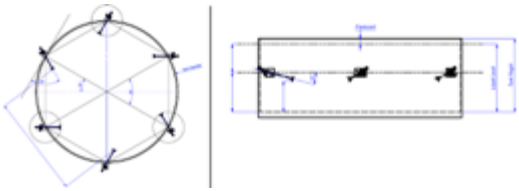


Figure 1: Mixer positioning

Business impact:

EYS Makina benefits from HPC resources in both design and application cases. In this project, EYS used HPC resources to investigate and optimize the mixing processes inside biogas digesters. Parallel, scalable, open-source CFD tools are utilized to achieve this goal.

While tackling this practical engineering problem, EYS Makina is also aiming to improve their HPC skills and reach a higher and more efficient level of HPC resource usage. In order to do so, EYS Makina:

- closely monitored scalability and parallel performance metrics.
- tried to increase utilization of automated operations, such as on-the-fly data analysis, file transfer processes, or using certain software applications remotely on HPC servers, etc.
- tried to streamline steps in their flow chart, aiming to minimize human intervention in the process (e.g., geometry generation, mesh generation, solver setup, etc.).

Benefits:

- Reduced response time to our customers
- Increase in customer satisfaction
- Strengthening market competition
- New customer demands on digester analysis

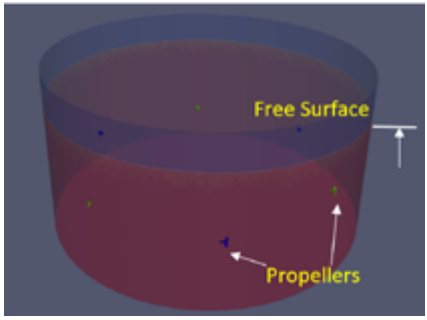


Figure 2: The full model of the digester

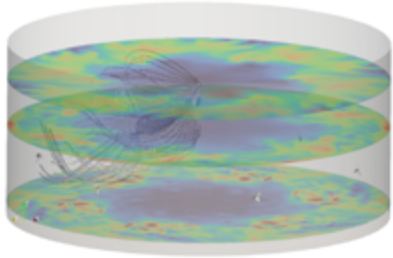


Figure 3: Snapshot of the flow field at different cross sections

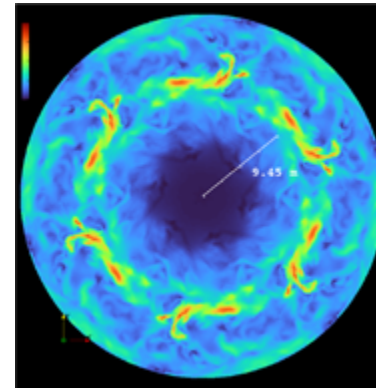


Figure 4: Instantaneous flow field represented with velocity magnitude at constant z-section

Keywords: Computational Fluid Dynamics, Large eddy simulation, Submersible axial mixer, Biogas Digester, Mixing Efficiency

Industry Sector: Renewable Energy, Manufacturing & Engineering

Technology: HPC

Contact:

Engin Orçun Kozaka, EYS,

okozaka@e-y-s.com

Ayse Gul Gungor, ITU UHeM,

ayse.gungor@itu.edu.tr

Merve Demirtaş, TUBITAK ULAKBIM,

merve.demirtas@tubitak.gov.tr

Photo by **Pavel Neznanov** on **Unsplash**

Section 4
Manufacturing and Engineering

HPC supporting the future of engineered wood products

NCC Finland

Industrial organisations involved:

Raute Oyj is a global group that develops and supplies technology and services to the wood industry. Raute is the market leader in plywood and veneer beam manufacturing technology and the only company in the world able to offer customers a complete factory-level solution. Raute combines technology, analytics and services to ensure more efficient, predictable and profitable production for its customers.

Technical/Scientific challenge:

The company encountered a R&D issue that needed to be solved quickly. In manufacturing there is a lead time of hours between production stages, providing the critical window in which data must be analyzed to prevent contamination of the physical wood products. The time-critical nature of the analysis requires efficient computation.

Solution:

The company decided to try supercomputing and the EuroHPC LUMI supercomputer was considered worth a chance. Raute subsequently contacted NCC Finland for computing resources. One calculation round on the LUMI Supercomputer was done ten times faster than with the traditional Raute workstation based methods.

Business impact:

Raute's mission, as a manufacturer and supplier of technology solutions for veneer and plywood, is to help its customers optimize and develop their entire production line. This means that the company will continue to process ever larger amounts of data, requiring AI based tools and technology for quality control and improvement.

According to Raute: "*The aim is also to enable customer factories around the world to use raw materials as efficiently as possible, but with as little impact on the environment as possible, i.e. through energy savings. In the future, HPC supported machine learning and the possibilities they offer will add value to the development of customer facing services.*"

Benefits:

- 10 times faster calculation
- Scalable processing using HPC
- Improved quality of end product by filtering out defective materials
- Ecological values and the fact that LUMI system uses certified, 100% renewable energy

Keywords: Wood industry, Plywood, Manufacturing, Analytics, Machine learning, LUMI

Industry Sector: Manufacturing & engineering, minerals and forest-based products

Technology: HPC

Contact:

Dan Still,

Dan.Still@csc.fi

The acquisition and operation of the EuroHPC supercomputer is funded jointly by the EuroHPC joint Undertaking, through the European Union's Connecting Europe Facility and the Horizon 2020 research and innovation programme, as well as the Participating States FI, BE, CH, CZ, DK, EE, ES, NL, NO, PL, SE.

Supercomputing for Built Environment Simulations

NCC Finland

Industrial organisations involved:

Sweco AB is a leading European consultancy company specialising in engineering and architecture for the built environment and industry. The company uses computational fluid dynamics (CFD) for advanced airflow simulations and industrial processes with the help of high-performance computing.

Technical/Scientific challenge:

Computational fluid dynamics can be used to simulate the flow phenomena of gases or liquids. By simulating flow rates, pressures and other turbulence quantities, it is possible to improve the energy efficiency of data centres, safety of hydrogen power plants, indoor environmental conditions in residential and office buildings, or wind comfort in the built environment. With CFD, a virtual prototype corresponding to the physical reality can be produced in the early stages of design. This reduces risks in decision-making and guides the design in the right direction.

Realistic and accurate models and simulations require substantial computing capacity. Cloud computing services have their limitations: the available runtimes may be too short for the modelling needs or using them may become expensive.

Solution:

"*Workflows specifically optimized for HPC minimize the bottlenecks of computing. In his way, the resources required for CFD – in practice, several hundred processor cores – can be used for a large number of simultaneous computation tasks*," Sweco explains.

They also add "*For our clients, this translates into significant added value, in terms of design quality, schedules and costs*".

Business impact:

Obtaining the results of computational fluid dynamics simulations typically takes several days or even weeks. In practice, the supercomputer enables up to ten times higher efficiency and the use of more accurate models.

"*It is indeed hard to imagine CFD without a high-performance computing environment,*" says Sweco, an experienced user of HPC technology.

Benefits:

- Benefits of modern technologies
- Best possible tools for solving our clients' challenges that require very specific expertise
- Added value in design quality and schedules
- Cost-efficiency

Keywords: CFD, LUMI, Simulations, Engineering

Industry Sector: Manufacturing & engineering

Technology: HPC

Contact:

Dan Still,

Dan.Still@csc.fi

The acquisition and operation of the EuroHPC supercomputer is funded jointly by the EuroHPC joint Undertaking, through the European Union's Connecting Europe Facility and the Horizon 2020 research and innovation programme, as well as the Participating States FI, BE, CH, CZ, DK, EE, ES, NL, NO, PL, SE.

Leverage from the EU 2014-2020

Wärtsilä Finland used supercomputing for engine property simulations: superior capacity, cost efficiency, and sustainability

NCC Finland

Industrial organisations involved:

Wärtsilä Corporation is a global leader in innovative technologies and lifecycle solutions for the marine and energy markets. The company emphasises innovation in sustainable technology and services helping its customers to continuously improve their environmental and economic performance.

Technical/Scientific challenge:

"*We deal with very complex and computationally demanding problems which often include mesh motion, chemical reactions, moving particles, and multi-phase flows. We are aiming to improve our cost efficiency and effectiveness in demanding simulations. The scale and complexity of our simulations have to fit within the required timeframes, too*," says Wärtsilä Finland.

Wärtsilä wanted to solve some challenges and enhance their computational capabilities and to explore what expertise the LUMI and CSC have to offer.

"*In our collaboration with CSC, we aimed at evaluating how LUMI's resources could help us to perform high-fidelity simulations more efficiently and cost-effectively*," explains Wärtsilä Finland.

The test runs also serve Wärtsilä's ambitions of becoming carbon neutral by 2030.

Solution:

The Wärtsilä's Thermofluids & Simulations team uses OpenFOAM, an open-source software for CFD simulations. OpenFOAM can be used to simulate internal combustion (IC) engines including piston and valve motion – a functionality that Wärtsilä Finland has developed in collaboration with CFD-Direct. "*Our partnership with CFD Direct has been key to unlocking the full potential of OpenFOAM. Combined with LUMI's computational power, we can effectively tackle large-scale simulations.*"

The company was really positively surprised by the results. The improved efficiency was clearly visible in their test runs. „*We were able to run our simulations much faster, meeting tight project deadlines and accelerating our R&D processes*", adds Wärtsilä Finland. They concluded by saying, "*We also assessed the cost-efficiency of our in-house infrastructure compared to the LUMI and CSC's services. We compared the costs, scalability, and performance benefits of each approach, which provided us with a clear understanding of the potential advantages of using HPC. CSC's responsive service desk ensured that any technical challenges were addressed promptly, minimizing downtime and enhancing overall productivity.*"

Business impact:

"*By leveraging advanced CFD techniques, we analyse complex fluid dynamics, combustion processes, and thermal interactions to optimize performance, efficiency, and sustainability. Our work plays a crucial role in driving Wärtsilä's mission to decarbonize marine and energy industries, ensuring innovative and environmentally friendly solutions for a cleaner future.*

Our focus is towards ramping up our industrial R&D computations in LUMI with the applications already identified. Obviously, our goal is to maximize the use of simulations and minimize carbon intensive energy consumption and costs. If we are successful, we should be able to leverage the use of any capacity offered to us," Wärtsilä Finland says.

Benefits:

- Improved efficiency in simulations
- Accelerated R&D processes
- Reduced energy consumption, sustainability
- Reduced costs

Keywords: RDI, Simulation, Sustainability, Moving particles, Cost efficiency

Industry Sector: Manufacturing & engineering, Energy and Maritime

Technology: HPC, CFD (Computational Fluid Dynamics)

Contact:

Dan Still,

Dan.Still@csc.fi

The acquisition and operation of the EuroHPC supercomputer is funded jointly by the EuroHPC joint Undertaking, through the European Union's Connecting Europe Facility and the Horizon 2020 research and innovation programme, as well as the Participating States FI, BE, CH, CZ, DK, EE, ES, NL, NO, PL, SE.

Leverage from the EU 2014–2020

European Union European Regional Development Fund

REGIONAL COUNCIL OF KAINUU

Development of an on-board battery charger

NCC France

Industrial organisations involved:

The NCC CC-FR is managed by TERATEC the European Pole of Competence in high performance digital simulation, in partnership with CERFACS the European Center for Advanced Research and Training in Scientific Computing, with Inria Academy, a continuing education program dedicated to open source software, with CRIANN, the Regional Computing Center of Normandy, and ROMEO the Regional Computing Center of the University of Reims Champagne-Ardenne.

INFINERGIES is an engineering and design company specialized in power electronics. Its services cover all the stages of a new product development project: analysis, definition of requirements, design and simulation, prototype testing and validation.

Technical/Scientific challenge:

Involved in the development of an on-board battery charger for electric vehicles, Infinergies was faced with the problem of the cumulative duration of the simulations to be carried out. The product must be able to charge the battery, but also supply electricity to the grid, or power sockets in the vehicle, regardless of the battery's state of charge. This requires a large number of simulations to test these numerous operating cases.

Solution:

Over a one-month period, launch of 140 calculations (single-core) with an average duration of 12 hours, thanks to:

· the available memory (60 GB of memory required for each calculation, i.e. half a Myria supercomputer calculation node),

· the ability to run several calculations simultaneously,

· continuous support from CC-FR for setting up calculations on the supercomputer, and for linking calculations together.

Business impact:

Infinergies had a very tight schedule. The help they received from CC-FR and CRIANN in setting up their simulations on the supercomputer was extremely effective, and then the access to the computing power enabled them to solve their industrial problem within the deadline.

Benefits:

- Simultaneous simulation of numerous operating points
- Compliance with the end customer's specifications in a short timeframe

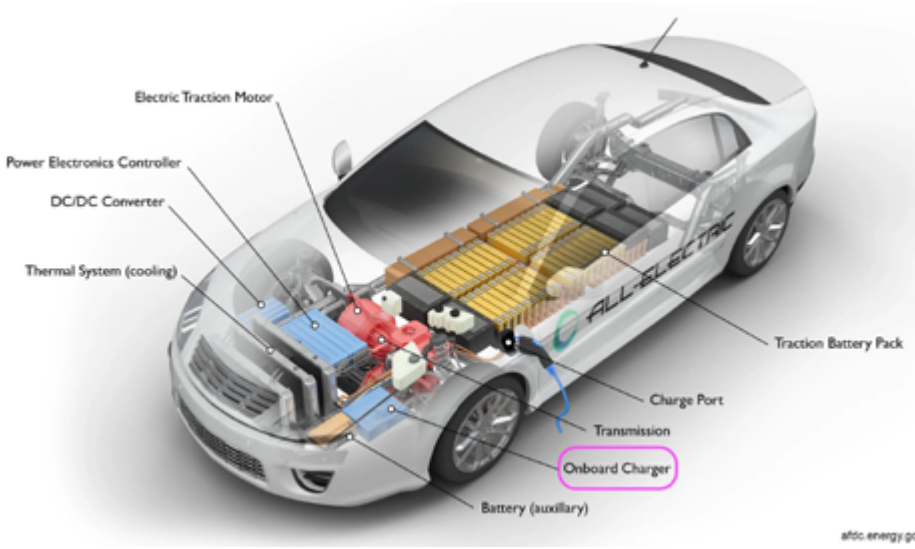


Figure 1: Key Components of an All-Electric Car
Image source: https://afdc.energy.gov/vehicles/how-do-all-electric-cars-work
Illustration identifying various components involved in the project.

More information: https://cc-fr.eu/accompagnement/

Keywords: HPC, Simulation, Battery Charger, Electrical Vehicle
Industry Sector: Industry/Transportation
Technology: HPC, Simulation

Contact:
Dr Karim Azoum
Karim.azoum@teratec.fr
+33 762 740 360

Development of explicit non-linear reference software for space applications

NCC France

Industrial organisations involved:

The NCC CC-FR is managed by TERATEC the European Pole of Competence in high performance digital simulation, in partnership with CERFACS the European Center for Advanced Research and Training in Scientific Computing, with Inria Academy, a continuing education program dedicated to open source software, with CRIANN, the Regional Computing Center of Normandy, and ROMEO the Regional Computing Center of the University of Reims Champagne-Ardenne.

ABSTRAO is a SME offering software and consultancy services for numerical simulation for the physics of extreme stress.

Technical/Scientific challenge:

ABSTRAO, a SME located in Toulouse, is specialized in integrating numerical modelling into the analysis and optimization of systems subjected to extreme loading. The company develops an explicit non-linear reference software dedicated to space applications, they needed a first evaluation of the scalability of a multi-GPU HPC version of the code.

Solution:

At the end of 2023, the company consulted its Regional Computing Center, CALMIP, which is one of the 21 Mesocenters involved in CC-FR. This partnership has enabled the company to take advantage of the Mesocenter's infrastructure and technical expertise.

A study of strong and weak scalability has been carried out on MESONET ARM supercomputer "Turpan" which is hosted at CALMIP. The calculations also made it possible to demonstrate the applicability of the software to simulate hypervelocity impacts of space debris on satellite protective structures.

Business impact:

The collaboration between CALMIP (one of the 21 Mesocenters involved in CC-FR) and ABSTRAO has been decisive in the first evaluation of the ABSTRAO HPC Solver performances. This partnership has enabled a comprehensive and rigorous assessment of the solver's multi-GPU HPC capabilities.

Benefits:

- Very good weak and strong scalability (90%) on 30xA100 on Turpan - MESONET ARM supercomputer
- Significant progress in code maturity

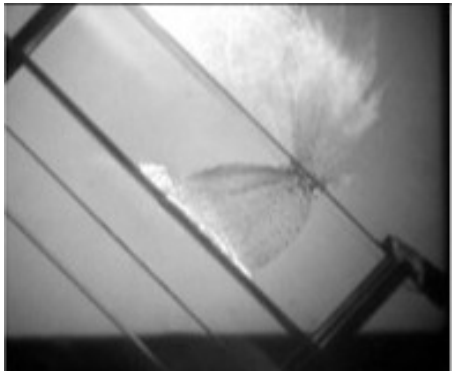


Figure 1: Hypervelocity impact test result at 7.2km/s

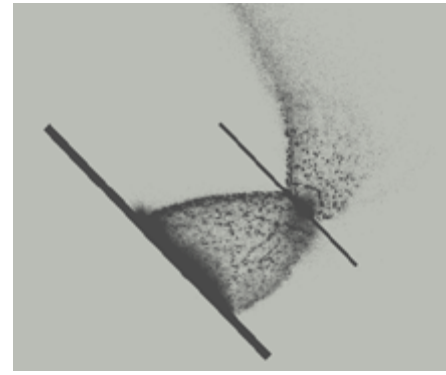


Figure 2: ABSTRAO Solver result

More information: https://cc-fr.eu/accompagnement/

Keywords: HPC, Simulation, Software

Industry Sector: Space Engineering

Technology: HPC, Simulation

Contact:

Dr Karim Azoum

Karim.azoum@teratec.fr

+33 762 740 360

Next generation of sound simulation for digital twin, accelerated by HPC

NCC Iceland

Industrial organisations involved:

Treble Technologies leverages the world's fastest wave-based sound simulation technology, enabling unprecedented efficiency in designing and optimizing sound and acoustics. With the support of IHPC, the Icelandic SME Treble received access to the HPC system DEEP at JSC, Germany – a system that received funding from the EuroHPC Joint Undertaking.
https://www.treble.tech/

Technical/Scientific challenge:

Wave-based acoustic simulation was traditionally too slow and impractical for anything beyond very small spaces, while geometric acoustic platforms struggled with accuracy at low frequencies and required days to compute complex, high-quality impulse responses for auralization, which limited their applications. Training AI models in sound applications was also difficult to measure and had large data.

Solution:

The Treble cloud-based tool accelerates audio model development by generating rich, realistic datasets without traditional data collection for AI training and testing. Using precise simulation technology, it models diverse audio scenes for robust real-world performance in tasks like speech recognition and noise suppression. Treble introduces an innovative hybrid sound engine that merges wave-based methods with geometrical acoustics, delivering unparalleled precision and scalability for creating acoustic simulations. This cutting-edge platform enables seamless virtual prototyping, empowering audio companies to identify design flaws early in the development process and significantly enhance the performance of their equipment, such as building design and the automotive industry.

Business impact:

Access to HPC through IHPC and the JSC system DEEP has enabled Treble Technologies to deliver substantial advances in sound simulation technology. Traditional acoustic modelling faced limitations in speed and accuracy, particularly for complex or large-scale environments. By leveraging HPC, Treble's solutions, such as Treble's acoustic simulation platform and the Treble SDK, achieve simulations considerably faster than competing tools, while delivering unparalleled accuracy. This has drastically reduced the need for physical prototypes and enabled efficient integration into design workflows.

Treble's technology, accelerated by HPC, enables innovation across industries such as construction, automotive, consumer electronics, and pro audio. The ability to generate synthetic audio data efficiently has also reduced time-to-market for AI-powered sound technologies, including speech recognition and adaptive sound systems. HPC integration empowers stakeholders with intuitive, scalable tools that improve decision-making, design quality, and overall product performance while supporting more sustainable and efficient development practices.

Benefits:

- Time savings: Sound simulations are now completed faster than traditional finite element analysis solvers
- Cost efficiency: Significant reduction in the need for physical prototyping and manual labor in model preparation
- Product optimization: Enhanced precision allows for earlier detection of design flaws, improving final product quality
- Waste reduction: Streamlined workflows reduce rework and associated material waste in industries such as construction and manufacturing

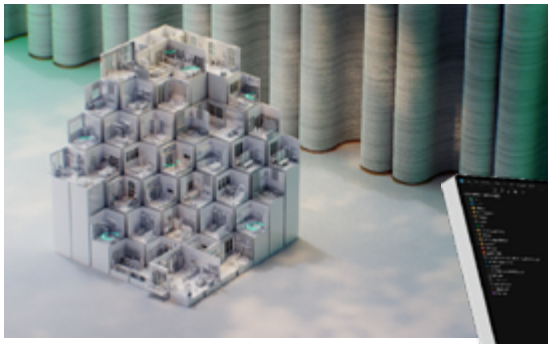


Figure 1: Rapid creation of diverse, highly realistic audio datasets and simulations

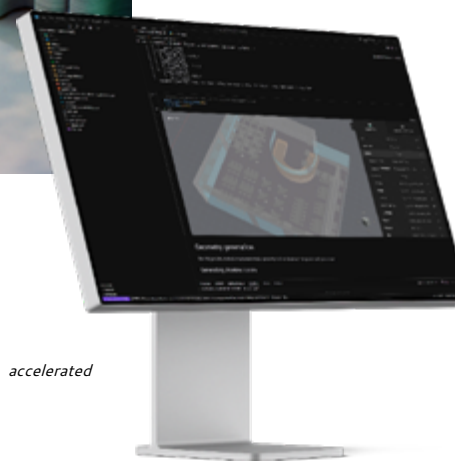


Figure 2: HPC empowers Treble for accelerated audio simulation and higher accuracy

Keywords: HPC, wave-based simulation, sound rendering, synthetic audio data, acoustic design

Industry Sector: Automotive, Construction, IT/HPC systems, Smart Cities, Consumer Electronics, Pro audio

Technology: HPC, AI, SaaS, sound simulation, synthetic data generation

Contact:

Reza Hassanian, seh@hi.is

Hemanadhan Myneni, myneni@hi.is

Grateful acknowledgment is given for the support and contribution of the European Digital Innovation Hub Iceland (EDIH-IS) under Grant Agreement No. 101083762.
This project has received funding from the European High-Performance Computing Joint Undertaking (JU) under grant agreement No 951732. The JU receives support from the European Union's Horizon 2020 research andinnovation programme and Germany, Bulgaria, Austria, Croatia, Cyprus, the Czech Republic, Denmark, Estonia, Finland, Greece, Hungary, Ireland, Italy, Lithuania, Latvia, Poland, Portugal, Romania, Slovenia, Spain, Sweden, the United Kingdom, France, the Netherlands, Belgium, Luxembourg, Slovakia, Norway, Switzerland, Turkey, Republic of North Macedonia, Iceland, Montenegro.

Innovating Portugal's Footwear Industry

NCC Portugal

Industrial organisations involved:

Atlanta, a Portuguese company founded in 1995, excels in innovative sole production for the footwear industry. With advanced facilities, they craft 20,000 pairs daily, blending design, quality, and technical expertise.

Technical/Scientific challenge:

The Portuguese footwear industry faces a major technical challenge in optimizing the design of shoe soles. The traditional trial-and-error method used to design injection molds for soles is time-consuming, resource-intensive, and fails to consistently produce optimal results. The challenge lied in transitioning from empirical approaches to computational modeling to improve production efficiency and product quality.

Solution:

By using HPC and process modeling, the footwear industry was able to simulate the injection molding process before physical experimentation. This simulation helped optimize parameters such as mold geometry, air vent locations, and gate positioning. The approach allowed detailed analysis of flow front evolution and weld line positioning, reducing the number of iterations needed to manufacture optimal molds. The collaboration between the University of Minho and Atlanta – Componentes para Calçado, Lda, within the Greenshoes 4.0 project, aimed to introduce this modeling phase into the design of shoe soles, cutting down on time, materials, and costs.

Business impact:

The current empirical method used in the footwear industry is inefficient and expensive. By incorporating HPC and computational modeling, the sector can revolutionize its approach to designing shoe soles, making it faster, more efficient, and precise. Simulations enable manufacturers to gather data that would be impossible through physical testing alone, allowing for better decision-making in optimizing mold designs. This digital transformation, promoted by projects like Greenshoes 4.0, will increase production efficiency, reduce waste, and enhance product quality. Ultimately, using HPC in the footwear industry will strengthen its international competitiveness and facilitate the adoption of advanced manufacturing techniques.

Benefits:

- Optimized mold design for better precision
- Reduced time and material costs
- Enhanced product quality by avoiding weak regions
- Improved process efficiency and faster production cycles
- Informed decision-making with detailed simulation data
- Support for innovation in manufacturing techniques
- Contribution to the digital transformation of the footwear industry

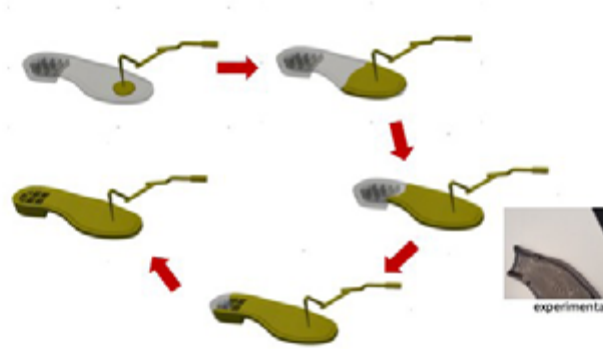


Figure 1: Simulation of the injection molding process using HPC

Figure 2: Optimized molds that resulted from the simulations

Keywords: Footwear industry, injection molding, high-performance computing, computational modeling, digital transformation

Industry Sector: HPC, HPDA

Technology: HPC, AI, SaaS, sound simulation, synthetic data generation

Contact:

Lígia Breda Melo / NCC Portugal,

ligiabredamelo@lip.pt / eurocc-portugal@lip.pt

Extrusion Blow Molding Simulator

NCC Portugal

Industrial organisations involved:

Founded in 1976, Logoplaste is a pioneer in the manufacturing of rigid plastic packaging, delivering innovative and eco-friendly solutions that reduce CO2 emissions and support a circular economy.

Technical/Scientific challenge:

In the extrusion blow molding (EBM) process used for manufacturing polymer-based containers, one of the main challenges lies in predicting the complex behavior of polymeric materials. Traditionally, design teams have relied on trial-and-error methods to define processing parameters, resulting in long setup times and significant resource consumption. Simulating material behavior and optimizing process parameters with computational accuracy has become increasingly important, especially for complex geometries and advanced material models.

Solution:

A research team from the University of Minho, in collaboration with the Logoplaste Innovation Lab, developed and implemented numerical simulation tools for the EBM process to improve the prediction of polymer behavior and enhance manufacturing efficiency. Initial versions of these tools yielded promising results for simpler geometries, but refining the constitutive models significantly increased computation time, beyond the capacity of desktop computing.

To address this, the team ported the simulation tools to high-performance computing (HPC) systems, boosting computational speed and enabling the integration of optimization algorithms. This approach made it possible to perform simulations within timeframes compatible with industrial needs.

Business impact:

Currently, the EBM process relies heavily on trial-and-error experimentation, which is time-consuming and inefficient. By leveraging HPC systems to simulate the entire process with high accuracy, the research team aims to drastically reduce the time required for process setup and optimization. The improved simulation tools will allow manufacturers to experiment with complex geometries and material behaviors, leading to more efficient design processes and better product quality. In the long run, the use of HPC in EBM simulation will enhance production efficiency, reduce waste, and improve time-to-market, giving industries a competitive edge.

Benefits:

- Optimized process parameters, reducing trial-and-error
- Shorter time-to-market for products
- Enhanced accuracy in design and material predictions
- Increased resource and cost efficiency
- Scalable simulations using HPC
- Continuous improvement through ongoing tool development
- Data-driven decisions via optimization tools

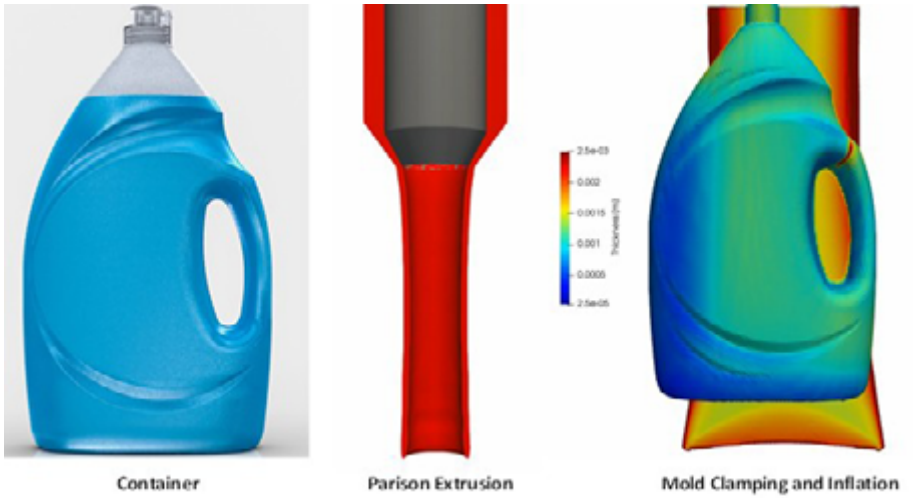


Figure 1

Keywords: Extrusion blow molding, computational rheology, polymer processing, process optimization, high-performance computing

Industry Sector: Manufacturing

Technology: HPC, HPDA

Contact:

Lígia Breda Melo / NCC Portugal,

ligiabredamelo@lip.pt / eurocc-portugal@lip.pt

HPC for Profile Extrusion

NCC Portugal

Industrial organisations involved:

Soprefa is a Portuguese SME specialized in the production and distribution of plastic profiles for a large variety of applications.

Wolf Dynamics is an Italian SME specialized in consulting services in computer-aided engineering, multi-physics simulations, numerical optimization, data analytics, and interactive data visualization.

Technical/Scientific challenge:

The design of extrusion dies for plastic profiles with complex geometries, essential for industries such as automotive, healthcare, and construction, traditionally relies on trial-and-error methods. This approach is time-consuming, resource-intensive, and often results in suboptimal products, making it difficult to meet industrial demands for cost efficiency and reduced time-to-market, especially for complex, customized profiles.

Solution:

The experiment implemented a simulation framework using HPC systems to optimize the design of extrusion dies. By leveraging open-source computational libraries like OpenFOAM and optimization tools like Dakota, Soprefa was able to run hundreds of design iterations within a single day, which would have been impossible with traditional methods. The simulation framework also integrated with CAD tools to streamline communication between design and HPC systems, allowing the automatic optimization of extrusion dies. The process was successfully tested on both existing and new designs, significantly improving efficiency.

Business impact:

Prior to using HPC, the design process for plastic profile extrusion dies was slow, costly, and limited in accuracy. With the adoption of HPC-based computational modeling, Soprefa has reduced product development time by 30-40% and lowered costs by 23%. The ability to run multiple simulations in parallel enables the optimization of complex die designs more efficiently, improving product quality and production speed. The success of this experiment highlights the potential for broader application of HPC in polymer processing and other industries. The collaboration between the University of Minho and Wolf Dynamics has not only solved key industrial challenges but also expanded their expertise in computational modeling, positioning them for future research and industrial projects.

Benefits:

- Improved profile extrusion die design using simulation, optimization, and HPC systems
- 30-40% faster product time to market (3 to 2 months)
- 40% reduction in raw materials (1 Ton to 600 kg)
- 23% cost reduction (€18,000 to €14,000 per tool)
- Greater independence in extrusion die design, protecting Soprefa's know-how

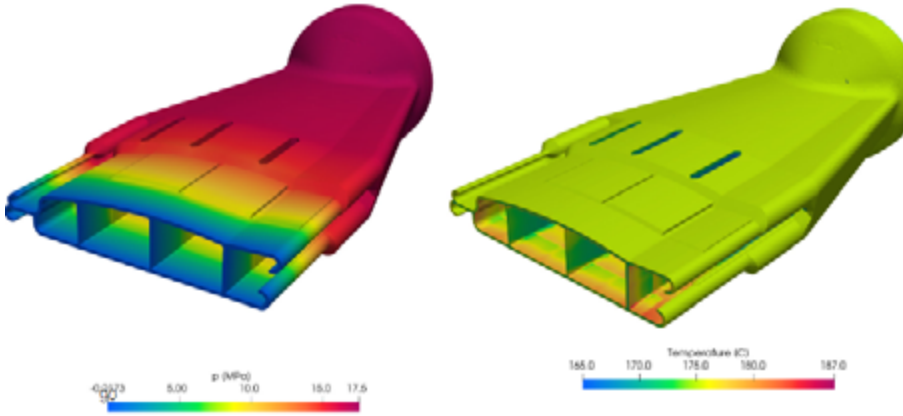


Figure 1: Profile extrusion is the technique employed to manufacture constant cross-section thermoplastic profiles, which has a vast range of major applications

Keywords: Extrusion dies, high-performance computing, plastic profiles, computational modeling, process optimization

Industry Sector: Manufacturing

Technology: HPC, CFD Simulation

Contact:

Lígia Breda Melo / NCC Portugal,

ligiabredamelo@lip.pt / eurocc-portugal@lip.pt

The work of the SME was performed within a business experiment funded by the FF4EuroHPC project. The project has received funding from the European High-Performance Computing Joint Undertaking (JU) under agreement No 951745. The JU receives support from the European Union's Horizon 2020 research and innovation program and Germany, Italy, Slovenia, France, Spain.

Rendering job preparation for multiple GPUs in an HPC cluster

NCC Slovenia

Industrial organisations involved:

GOSTOL – GOPAN is a global provider of integrated industrial solutions for medium-sized and large industrial bakeries, including equipment for dough mixing, dividing, molding, proofing, baking and cooling. Complete adaptability to the buyer's needs, more than 70 years of experience in the field of baking industry, technically and technologically sophisticated, energy-efficient equipment, innovativeness and reliability are our main qualities.

Technical/Scientific challenge:

Various visualizations, animations and high-fidelity visual designs of products the company produces are needed for marketing, exhibitions and presentation purposes. Rendering is done on local workstations which don't allow processing in a desired timeframe because of the lack of computer resources. The obvious solution is using a GPU cluster and understanding how much can we speed up this process. Blender has been chosen as a software solution since it is free and open-source.

To use multiple GPUs in Blender, it is required to prepare multiple jobs, one for each GPU involved in the computation. In case of a large number of frames to be processed on many GPUs, manual job preparation becomes very time-consuming but also leading to user errors and not optimal resource allocation. For this reason, the automation of job preparation is needed.

Solution:

In this case we wanted to achieve two goals. The first was to demonstrate to the customer that using a GPU cluster is a much faster and cheaper solution than using a powerful workstation. We achieved this by comparing the time needed to finish the same computations on one workstation and on multiple GPU nodes.

The second was to prepare a simple and intuitive application that allows the user to upload the blend file, define the frame range to be rendered and the number of GPU nodes available. The application prepares the job scripts for each node by splitting the desired frame range into multiple sets of frame ranges for each GPU on a single node. This way we can automate the process of job scripts preparation, optimize resources allocation and evenly distribute the workload among nodes and GPUs.

Business impact:

Based on the described PoC, we could demonstrate that using HPC services on pay-per-use billing model instead of using on-premise workstations would have significant advantages for the current customer needs. Various rendering jobs are needed only a few times in a span of one year and at the same time renders are needed fast which is crucial for meeting the deadlines. With more compute power we could also deliver better and more attractive visualizations.

Additionally, the shared nature of HPC resources means that users can scale up their usage as needed, paying only for the compute time they use rather than investing in expensive, high-end hardware that might sit idle during periods of low demand. This flexible usage model can lead to lower overall costs compared to maintaining a powerful local environment. Moreover, HPC environments are optimized for peak usage, ensuring that resources are available when they're needed most. Overall, using HPC not only improves productivity and results but also provides a more cost-effective and scalable solution for demanding creative workflows.

Having an intuitive GUI makes the job preparation easier, faster and resilient to user errors, allowing also non-experienced users to be able of conducting computational jobs on HPC systems.

Benefits:

- Drastically cuts simulation time
- Lowers operational costs
- Enhances visualizations
- Boosts response to peak demands

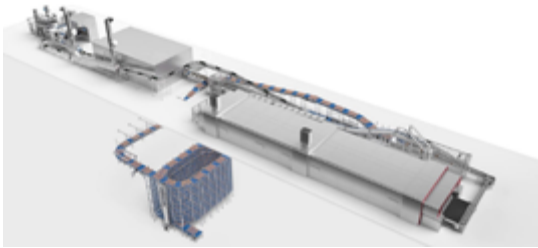


Figure 1: Industrial bakery line for free-baked long loaves

Figure 2: Simple application that allows automatic Blender jobs preparation

Keywords: Rendering, Blender, Visualizations, Automatation of jobs preparation, GPU computing

Industry Sector: IT/HPC systems, services & software providers, Manufacturing & engineering

Technology: HPC

Contact:

Tristan Pahor, NCC Slovenia,

Tristan.pahor@arctur.si

Alleima: Simulating material behavior

NCC Sweden

Industrial organisations involved:

Alleima AB, formerly known as Sandvik Materials Technology, is a leading manufacturer specializing in advanced stainless steels, powder-based alloys, and special alloys. The company provides high-performance materials and solutions for a number of industries. These include aerospace, automotive, chemical processing, oil and gas, and medical technology. With a focus on innovation and sustainability, Alleima aims to enhance performance and productivity for its customers through its materials technology expertise.

Technical/Scientific challenge:

Alleima AB accessed LUMI supercomputer for their material behavior simulations using DAMASK software. Alleima received important help from ENCCS material science expert Wei Li.

Achieving precise material properties involves extensive mechanical testing to estimate peak forces in specific working processes and to determine the mechanical properties of the final product. However, this process is hindered by lengthy lead times and high costs. Currently, predicting low-level deformation mechanisms and strain hardening requires laborious characterization work, contributing to extended timelines and increased expenses to attain the desired material properties.

The bottleneck in material characterization significantly impacts the precision of the materials design, relying heavily on the availability of cutting-edge materials characterization techniques. While nondestructive characterization is the preferred method, it remains a challenge in many cases, hindering progress in material property determination.

Solution:

To expedite the development of new materials, both the knowledge and methodology of material design using computer-aided tools are crucial. Leveraging recent advancements in crystal plasticity modeling provides a robust and efficient approach to material development. It offers a promising solution to overcome the challenges posed by traditional testing methods. This project employs Damask software on LUMI supercomputer for crystal plasticity modeling.

Figure 1: Processing of the alloys including hot and cold processes. Credits: Alleima AB.

Business impact:

The scientific case of the project revolves around utilizing DAMASK for crystal plasticity calculations, with input from stacking fault energy (SFE) calculations. The primary motivation is to overcome the challenges associated with estimating SFE in materials with values close to or below 0, where conventional methods yield substantial errors. Using DAMASK, the project aims to simulate tensile test curves to accurately predict forces in designated working processes. This integrated approach improves the reliability of material behavior predictions in practical applications.

Benefits:

- Reducing lengthy lead times and high costs associated with extensive mechanical testing and labor-intensive characterization work
- Utilizing advanced modeling techniques to improve the prediction of deformation mechanisms and strain hardening, enhancing the precision of materials design and ensuring desired material properties are achieved more efficiently
- Understanding the mechanical behavior of materials through advanced modeling enables the optimization of the production routes, resulting in decreased costs and reduced waste of energy, materials, and resources

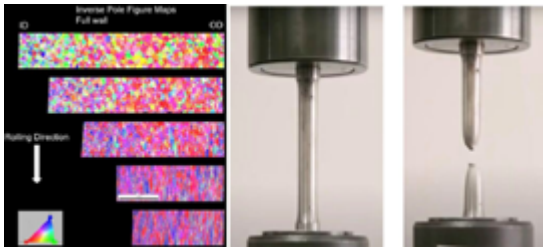


Figure 2: Example of mechanical tests and characterisation techniques (e.g. EBSD). Credits: Alleima AB.

Figure 4: Examples of Alleima's products. Credits: Alleima AB.

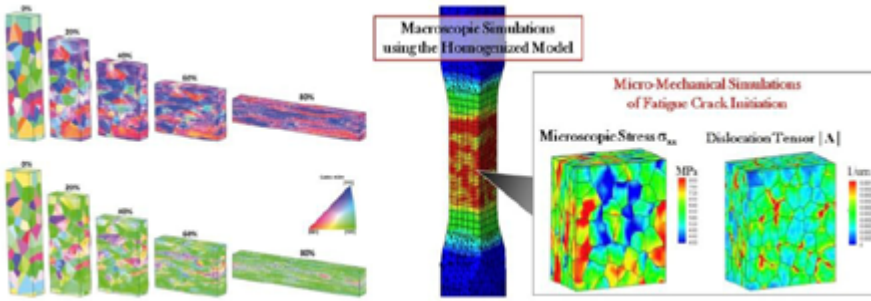


Figure 3: Example of crystal plasticity simulations (Damask software). Credits: Alleima AB.

Keywords: Material behavior, Damask, LUMI, simulations

Industry Sector: Material sciences, Manufacturing & engineering

Technology: HPC

Contact:

Apostolos Vasileiadis,

apostolos.vasileiadis@ri.se

Simulating liquid metal jets

NCC Sweden

Industrial organisations involved:

Excillum AB is a global leader in the field of advanced microfocus and nanofocus X-ray sources. The company produces cutting-edge high-end X-ray sources utilizing the patented metal jet anode technology.

Technical/Scientific challenge:

The metal jet anode is a flowing liquid metal that replaces the standard solid anode and enables much higher electron-beam power and can therefore generate higher X-ray flux. The challenge is to control the shape characteristics of the moving liquid in the design of the X-ray machines.

To aid the design, computational fluid dynamics (CFD) calculations are performed. The calculations include velocity, pressure, turbulence distribution, and liquid-vacuum interface design parameters. The CFD simulations is a fast and cost-effective approach for evaluating a large number of designs.

The grid size, together with the time-stepping size decides the resolution of a simulation. During the simulations, micro-scale spatial resolution and micro-second temporal resolution provide accurate fluid dynamics information regarding jet behaviors such as deformation, shifting, twisting in a high vacuum environment. In one of the studies, they will investigate how deviations from the ideal nozzle shape will influence the jet shape. A simulation to answer this question requires a significant amount of computing resources.

Solution:

The CFD application is developed using existing libraries provided by OpenFOAM v1812. The Excillium team will first launch a series of benchmark tests to assess the performance of OpenFOAM for the user case. Depending on the results, they may further carry out some related optimization work. The goal is to find the optimal setup for performance given a certain resolution of the simulation.

Business impact:

Accessing a supercomputer will allow Excillium to conduct several numerical experiments for parameter studies which saves both time and costs compared to conducting the same number of physical experiments in a lab. By this, they will shorten the lead time between the concept & feasibility phase and the development phase in the product life cycle, which is a key to achieving efficient development.

Benefits:

- Parameter studies to avoid lab tests aiming at design optimization, which saves working hours and cost on test materials
- Numerical experiments to revel new knowledge related to improving the configuration of key components in the product
- Ability to run simulation of different kind, e.g., CFD, optical simulation, particle simulation, in a single platform, which increase efficiency

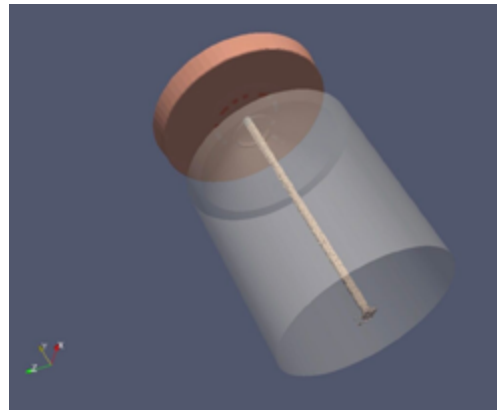


Figure 1: An overview of a liquid metal jet running in a vacuum chamber, revealed by detailed simulations

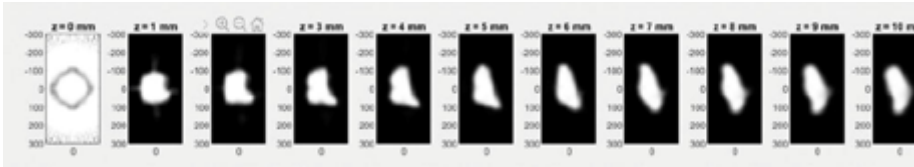


Figure 2: Cross-sections of a liquid metal jet subject to deformation due to unwanted defects in the nozzle, revealed by detailed simulations

Keywords: Simulation, OpenFOAM, CFD, Comptutational fluid dynamics

Industry Sector: Manufacturing & engineering

Technology: HPC

Contact:

Apostolos Vasileiadis,

apostolos.vasileiadis@ri.se

Photo by **Glenn Carstens-Peters** on **Unsplash**

Section 5
Text- and Speech-based Services

Automated Data Anonymisation Framework for Human Resources

NCC Cyprus

Industrial organisations involved:

Hello Radius is an AI-powered hiring platform that accelerates and secures recruitment by using skill-matching and anonymisation technology to reduce bias and protect candidate data. Designed for Human Resources (HR) teams, it streamlines candidate sourcing, screening, and shortlisting to make hiring faster, fairer, and more cost-effective.

Technical/Scientific challenge:

Hello Radius needed a reliable way to anonymise sensitive data in candidate CVs, such as names, addresses, and phone numbers, to ensure compliance with data privacy standards and reduce hiring bias. The challenge was to accurately anonymise this information while preserving the key qualifications, skills, and experience essential for fair candidate evaluation.

Solution:

NCC Cyprus offered consultation services to Hello Radius, which focused on investigating a range of tools for effective data anonymization. Microsoft Presidio and LangChain provided good candidate solutions, due to their strong capabilities in detecting and anonymising sensitive data. The integration of these tools could provide a solution that safeguards personal information while keeping valuable, non-identifying data intact.

Business impact:

The use of Microsoft Presidio and LangChain for data anonymisation has greatly improved Hello Radius's efficiency and compliance with data privacy laws. By protecting sensitive information from candidate CVs in an efficient manner, the company reduces the risk of data breaches and legal issues while promoting a fairer hiring process that focuses on skills and experience rather than personal details The suggested approach supports the company's mission to make AI-powered recruitment safer and fairer.

Benefits:

- Enhanced data privacy compliance
- Efficient and scalable anonymisation
- Improved data integrity for decision-making
- Reduced hiring bias

Keywords: Data anonymisation, secure recruitment, compliance with data privacy, AI, LLMs

Industry Sector: Services & software providers

Technology: Artificial Intelligence, LLMs

Contact:

eurocc-contact@cyi.ac.cy

Funded by the European Union. This work has received funding from the European High Performance Computing Joint Undertaking (JU) and Germany, Bulgaria, Austria, Croatia, Cyprus (co-funded by the EU within the framework of the Cohesion Policy Programme "THALIA 2021-2027"), Czech Republic, Denmark, Estonia, Finland, Greece, Hungary, Ireland, Italy, Lithuania, Latvia, Poland, Portugal, Romania, Slovenia, Spain, Sweden, France, Netherlands, Belgium, Luxembourg, Slovakia, Norway, Türkiye, Republic of North Macedonia, Iceland, Montenegro, Serbia under grant agreement No 101101903.

Unlocking the Power of ASR: From Audio Data to Text Transcription

NCC Denmark

Industrial organisations involved:

Dictus ApS is dedicated to delivering the best speech recognition technology for Nordic languages, serving clients across diverse sectors, including government, broadcasting, logistics, and healthcare. Their mission involves developing and refining speech recognition models, utilizing advanced technologies like Wav2Vec and Whisper to meet specific needs such as language variations, domain-specific jargon, and varying acoustics. Trusted by institutions like the Danish Parliament, Norwegian Storting, DR, TV2, Postnord, and medical professionals, Dictus consistently demonstrates expertise in speech-to-text solutions, whether for dictation or audio recordings.

Technical/Scientific challenge:

Dictus would like to train end-to-end speech recognition models, such as Wav2vec and Whisper, for the Scandinavian languages. However, on their own hardware, they can only reasonably hope to train models on, say, a thousand hours of speech. Dictus has much more data than that at the Dictus disposal – for Danish only, an order of magnitude more.

In addition, Dictus would like to improve the robustness of the models by adding augmented (i.e., synthetically manipulated) data. The attempts at data augmentation they have made on smaller data sets have been promising, but so far it has been impractical to scale this approach.

Solution:

Training on a cluster of GPUs would drastically reduce training time and make it feasible to exploit much more of the data they have at their disposal, as well as the data augmentation schemes they foresee. In addition, for the first time, they see a practical possibility of training a common model for several Scandinavian languages, exploiting transfer learning.

Business impact:

Dictus has identified a major business opportunity in offering Scandinavian speech recognition at a quality comparable to much larger languages in the world. They already have developed advanced ASR solutions in operation for several large customers in Scandinavia. However, they face tough competition from the tech giants, boasting far greater hardware resources available for model training. With access to the LUMI supercomputer, Dictus will get the opportunity to train and fine-tune ASR models with Scandinavian training data, as well as focus on the peculiarities of these languages.

Benefits:

- Reducing the error rate of the recognized text
- Reducing the resources needed for post-processing and finalizing of the text results
- Strengthening Tech Independence and Data Sovereignty
- Ensure GDPR-compliant solutions for public and private customers

- Keywords: Language and speech technology, Data Protection & GDPR-compliant solutions, Acoustic model training, AI end2end speech recognition, Scandinavian ASR models
- Industry Sector: Services & software providers, Life sciences, Health care
- Technology: HPC

Contact:
marta.schulze@deic.dk
dennis.wollbrink@deic.dk
ncc-dk@deic.dk

Improving the efficiency of healthcare client encounters with the LUMI supercomputer

NCC Finland

Industrial organisations involved:

Gosta Labs is a Finnish health technology company developing machine learning models to improve patient care and medical practice. The company's flagship product is Gosta Aide, an artificial intelligence assistant with language models tailored to the needs of the social and healthcare sector, automating the creation of client and patient records. This frees up the time of care staff for client encounters, thus supporting the delivery of human-centred, high-quality care.

Technical/Scientific challenge:

Health and social care professionals need to spend up to half of their daily working time on administrative tasks such as creating client and patient records. There can be differences in the recording practices of professionals in different organisations. The aim was to harmonise them and develop an AI assistant to help with administrative tasks in healthcare. Gosta Labs wanted to generate client and patient records in several European languages. While existing solutions are already in use (mainly US), many of them use closed, large language templates with insufficient data protection. There is a need for high-level data protection and a controlled language model training process.

Solution:

Gosta Labs has used the LUMI supercomputer to develop task-specific language templates. Training task-specific machine learning models requires substantial computing power. LUMI was chosen as the best platform for Gosta Labs to train its models. During the project, the company created several task-specific models that allowed replacement of the previous large, closed models. They improved their models' performance and resource efficiency, while also offering an environment-independent alternative to closed large language models from a data protection perspective. They also achieved the scalability needed to produce the necessary European data sets for training and to train their models for multiple European languages. The models they develop will be able to run in much lighter environments and will be more energy efficient too.

Business impact:

Gosta Labs´ industrial research project supported by Business Finland, helped them to launch their LUMI work in spring 2024. Later the company raised €1.2 million in pre-seed funding to accelerate its wider development and internationalisation. The company aims to become Europe's leading AI developer in healthcare. The first models have already been developed and deployed in healthcare in Finland and Switzerland.

"*The next step for Gosta Labs is to extend its support to more European languages. We want to continue to develop new models and further develop our models on LUMI. We will also improve our capabilities in the automated measurement of impact data*," says Gosta Labs. "*Our customers have been genuinely excited about our strong European expertise and our ability to develop machine learning models for healthcare, so that not all solutions need to rely on global providers*," says Gosta Labs.

"*In addition, our development efforts have been praised by data protection teams in healthcare organizations*," adds Gosta Labs.

Benefits:

- Created several task specific models allowing replacement of large, closed models
- Improvement of models´ performance and resource efficiency
- Organisational environment independency and improved data protection
- Improved scalability to train models for several European languages

Keywords: Health care, AI assistant, Technology development, Task-specific European language models, Machine learning models

Industry Sector: Health care, technology services & software providers

Technology: HPC, AI, machine learning

Contact:

Dan Still,

Dan.Still@csc.fi

Tiina Leiponen,

Tiina.Leiponen@csc.fi

The acquisition and operation of the EuroHPC supercomputer is funded jointly by the EuroHPC joint Undertaking, through the European Union's Connecting Europe Facility and the Horizon 2020 research and innovation programme, as well as the Participating States FI, BE, CH, CZ, DK, EE, IS, NL, NO, PL, SE.

Leverage from the EU 2014-2020

This project has received support from Business Finland.

AI-Powered IT Service Optimization with Language Modelling

NCC Hungary

Industrial organisations involved:

DAM Invisible Technology Plc. is a cutting-edge IT solutions provider specializing in the design, development, and management of complex IT infrastructures, with a strong focus on security, automation, and operational efficiency.

At the core of DAM Invisible Technology's innovation is IR:IS, an advanced AI-powered platform that transforms operations by automating workflows, enhancing security, and improving system resilience. Unlike traditional IT management tools, IR:IS is designed to integrate seamlessly with existing infrastructures, proactively identifying and resolving issues before they impact business continuity.

Technical/Scientific challenge:

Manually managing and categorizing IT service tickets is time-consuming and inefficient. The challenge was to develop an AI-based system capable of automatically classifying and prioritizing tickets while ensuring compliance with strict Service Level Agreements (SLAs). With a large but imbalanced dataset, traditional methods struggled with accuracy and scalability.

Key methodologies include the use of BERT-based classifiers and Masked Language Modeling (MLM) to enhance ticket processing accuracy. Initial experiments with pre-existing models (e.g., hubert-base-cc and NYTK\PULI-BERT-large) demonstrated limitations, prompting the creation of proprietary models IRIS-BERT-base and IRIS-BERT-large.

Solution:

By leveraging HPC resources on the Komondor supercomputer, DAM Invisible Technology Plc. trained advanced AI models, including BERT-based classifiers and Masked Language Models (MLM). The IRIS-BERT models, trained on millions of Hungarian-language documents, improved ticket classification accuracy and automated key IT service processes. Advanced training techniques, such as curriculum learning and loss function optimization, enhanced model performance while HPC enabled rapid, large-scale training.

Results show substantial improvements in classification performance, with metrics like accuracy and F1 Macro significantly surpassing earlier statistical models. Future developments include the completion of training for IRIS-BERT-large, expected to further enhance system capabilities.

Business impact:

The integration of AI into IT service management not only minimizes manual workload but also significantly accelerates ticket resolution, ensuring a more efficient and streamlined operational framework. By automating the classification and prioritization of tickets, organizations can strategically reallocate engineering resources toward high-value initiatives, enhancing overall productivity and innovation.

The ongoing advancements in IRIS-BERT-large are set to drive even greater efficiency, making IT operations more scalable, cost-effective, and resilient. Backed by cutting-edge machine learning models, this system continuously refines its decision-making capabilities, reducing human error and improving response accuracy.

By leveraging HPC capabilities for the development, through Komondor, DAM Invisible Technology ensures that its AI models operate with exceptional precision and speed, further amplifying the impact of automation in enterprise IT environments.

Benefits:

The integration of these innovations into the IR:IS platform underscores the potential for AI-powered efficiency gains in IT service management.

- Up to 50% reduction in service desk workload, enabling IT teams to focus on proactive system improvements rather than repetitive tasks
- 15% increase in engineering capacity, allowing for faster project delivery and innovation
- 20-100% reduction in SLA-related penalties, cutting down unnecessary operational costs while improving service reliability
- A 10% increase in efficiency directly contributes to doubling the profit generated by human-driven operations

Category	BERT Classifier	BERT MLM Classifier	HF Sequence Classification	
Base Model	hubert-base-cc	hubert-base-cc	IRIS-BERT-base	
Metrics	Accuracy	Accuracy	Accuracy	F1 Macro
OSL	63,41%	79,24%	**85,58%**	58,54%
OSLT	77,33%	81,21%	**88,97%**	72,19%
TYPE	88,85%	84,56%	**90,30%**	69,44%
PRIORITY	77,71%	79,60%	**91,35%**	77,86%
USER_GRADE	52,50%	57,46%	**69,63%**	66,74%
Average	71,96%	76,42%	**85,17%**	

Figure 1: IRIS-BERT results in relation to Human performance. © and source: DAM Invisible Technology data.

Figure 2: IRIS-BERT models achieve comparable results to previous Hungarian BERT models, and our large model performs best in the categorization tasks. © and source: DAM Invisible Technology data, Weights and Biases.

Keywords: DAM IT, IR:IS, AI, IT Service Management, Ticket Classification, BERT, NLP, Automation, Efficiency

Industry Sector: IT service provider

Technology: HPC, AI

Contact:

András Dorn (CEO), dorn.andras@damit.hu

Márk Antal (Competence Center Leader), antal.mark@damit.hu

Optimizing AI-based RAG Agents: Finetuning LLMs with High-Performance Computing

NCC Netherlands

Industrial organisations involved:

Zeta Alpha, based in the Amsterdam Science Park, is an advanced AI research and product lab with strong ties to academic and global research networks. They are developing a cutting-edge Enterprise Retrieval Augmented Generation (RAG) and AI Search platform, helping knowledge-intensive organizations to unlock insights and reuse internal and external knowledge by applying Generative AI in a secure internal platform. This empowers researchers, analysts, and decision-makers to easily navigate vast knowledge landscapes and make more informed decisions. With a focus on innovation, they aim at redefining AI's role in enhancing decision-making processes, building new product categories that revolutionize enterprise knowledge management.

Technical/Scientific challenge:

Zeta Alpha's Software-as-a-Service (SaaS) product "Neural Discovery Platform" links information from internal documents, scientific papers, source code, and practical use cases, leveraging state-of-the-art deep-learning-based natural language processing (NLP) models. By integrating smart AI-centric search, advanced analytics, recommendations, question answering, and summarization, it synthesizes insights from scattered content, reducing the time required to stay up-to-date and offering a comprehensive field overview.

To achieve these goals, Zeta Alpha conducts R&D projects in Neural Search, RAG, an approach that enhances AI-generated responses by retrieving relevant information, and AI Agents, intelligent autonomous systems designed to process and synthesize insights from data. A key component of these systems is embeddings, numerical representations of words, sentences, or documents that capture their semantic meaning and relationships, enabling AI models to understand and generate text more effectively.With support from NCC NL, Zeta Alpha started the process to explore different foundation LLMs and tackle the challenges of fine-tuning Generative models to produce embeddings for specific domains with great efficiency. By generating embeddings tailored to specific domains, such as medicine or finance, Zeta Alpha aims to enable its platform to provide highly accurate and relevant results for users searching for information in those areas. This process requires increased access to GPU resources to optimize performance and enhance model accuracy.

Solution:

The SaaS platform developed by Zeta Alpha uses customized vector embedding models to power semantic retrieval for Retrieval Augmented Generation (RAG). The quality of the embedding models on a specific domain and language determines the accuracy of the RAG system's answers, reduces language model hallucinations (a phenomenon in AI where language models generate incorrect or fabricated information), and hence increases user and organizational trust in their AI platform.

Instead of using dedicated resources on public clouds to experiment with new approaches, Zeta Alpha started a collaboration with NCC NL to leverage the national HPC ecosystem, including consultancy, training, and access to world-class GPU resources. Through this collaboration, Zeta Alpha gained insights into the solution's performance and efficiency by exploring the HPC architecture (memory and parallel I/O, scalability on advanced GPUs), which is more difficult to accomplish in public clouds.

The project was successful in quickly creating a model (Zeta-Alpha-E5-Mistral) that achieves a competitive score on the global Massive Text Embedding Benchmark (MTEB) Leaderboard. The model was open sourced, along with the instructions on how to reproduce the work. Zeta Alpha's project is ongoing, aiming to create even more advanced models, particularly those with strong multi- and cross-lingual performance across all European languages, and effective retrieval on key business domains (chemistry, legal search, high tech manufacturing, life sciences, R&D).

Business impact:

The business impact of using HPC for language models is transformative, driving innovation, reducing costs, improving customer experience, and unlocking new revenue streams. Organizations that integrate HPC into their AI strategies gain a significant competitive advantage by processing and analyzing larger datasets more efficiently. With the insights from the proof-of-concept results, Zeta Alpha can make informed, data-driven investment decisions, ensuring resources are allocated effectively to maximize growth and further enhance AI capabilities. Additionally, supporting European startups in the Neural Search and Embedding space contributes to the EU's goal of achieving technological autonomy.

Benefits:

- HPC accelerates the training and fine-tuning of language models
- Faster processing allows frequent updates to models
- HPC ensures that LLM achieve high levels of accuracy and efficiency, directly enhancing customer-facing applications like chatbots, virtual assistants, and recommendation systems

Figure 1: Zeta Alpha Team in their office in Amsterdam. @ 2024 Zeta Alpha

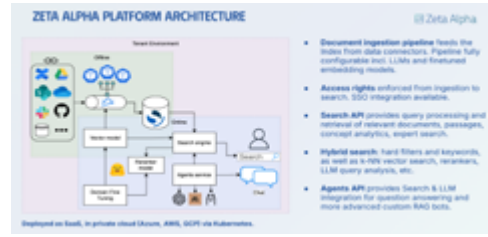


Figure 2: Zeta Alpha platform architecture @ 2024 Zeta Alpha

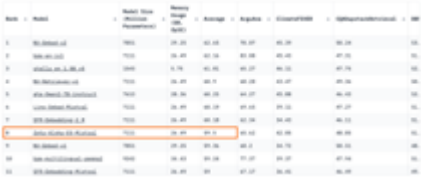


Figure 3: MTEB Leaderboard - Retrieval tasks @ 2024 Zeta Alpha

For more information: https://www.zeta-alpha.com/post/fine-tuning-an-llm-for-state-of-the-art-retrieval-zeta-alpha-s-top-10-submission-to-the-the-mteb-be

Keywords: Proof of concept, conceptual exploration, Exploration of performance, efficiency and accuracy, Reproducible Workflow

Industry Sector: IT/HPC systems, services & software providers

Technology: HPC, AI, High-end GPU

Contact:

Erik Kentie, erik.kentie@surf.nl

Jakub Zavrel, zavrel@zeta-alpha.com

Artificial Intelligence gives Norwegian Customs improved control of import and export of goods

NCC Norway

Industrial organisations involved:

Norwegian Customs work to prevent illegal imports and exports of goods and facilitate correct and efficient imports and exports of goods according to the regulatory framework of the Norwegian government.

Technical/Scientific challenge:

Verifying compliance with regulations regarding the import and export of goods is complex and involves 20 to 30 different public agencies at the border.

When declarations are filled out there are often errors or inaccuracies in the way company names and addresses are entered, resulting in various ways to write the same sender. This makes it challenging to obtain a comprehensive overview of goods imported to and exported from Norway by the same entity.

Maintaining an accurate and up-to-date archive of foreign entities and their associated declarations presents a significant challenge.

Solution:

Entity resolution is a branch of artificial intelligence that aims to identify data records belonging to the same real-world object. In this context it is used for identifying foreign senders and receivers. This is particularly valuable when merging data from diverse sources, aiding in customer management and in fraud detection.

By applying entity resolution to the Norwegian customs archive of declaration records, we were able to identify declarations coming from the same foreign parties, beyond what is possible with filtering techniques.

Norwegian Customs benefit from entity resolution by detecting similar foreign entities, thereby capturing a more comprehensive data representation of the flow of goods.

Business impact:

For Norwegian Customs, which annually declare large amounts of goods in and out of the country, it is of great importance to ensure that the goods declarations recorded are correct. The machine learning algorithms attempt to correct the inaccuracies that occur when goods declarations are filled out and read.

It is crucial for the Customs Authorities to be able to identify all shipments received or sent by the same actor. This applies to both Norwegian and foreign actors. NORCE supported the Customs Authorities in developing new methods.

Entity Resolution is fundamental for integrating information from multiple sources. Linking objects across data sources is not a trivial task. An effective algorithm for Entity Resolution means that the Customs Authorities can better utilize all their data sources.

Benefits:

- 9 million customs declarations annually can be processed efficiently
- Achieve a comprehensive data representation of the flow of goods
- Improved control of illegal import and export of goods

Figure 1: Illustrative photo of containers inspected by Norwegian Customs

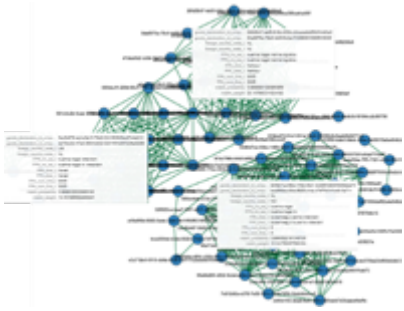


Figure 2: Machine learning algorithms cluster entities based on similarity

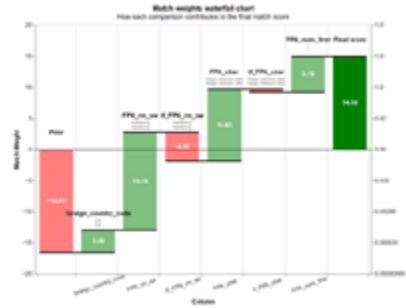


Figure 3: Weight matching for similar entities

Keywords: Entity resolution, goods declaration, Machine learning

Industry Sector: IT/HPC systems, Public services/Civil protection

Technology: AI, Machine Learning, HPC

Contact:

Jeremy Cook,

jeco@norceresearch.no

Harnessing AI to Mitigate Climate Impacts

NCC Sweden

Scientific partners involved:

RISE Research Institutes of Sweden is a state-owned research institute with over 3000 employees and offices located throughout Sweden. The organization is dedicated to fostering innovation and advancing knowledge across various scientific and technological fields, to foster research and innovation to serve Swedish industrial and societal sectors towards strengthening Sweden's global competitiveness, while aligning with sustainability objectives.

The Natural Language Processing (NLP) Group at RISE carries out research on fundamental topics such as resource-efficiency, multilingualism and language model-evaluation, and on applications such as political survey analysis, detection of adverse effects of medicines from medical data, and retrieval of impacts of extreme climates events by mining online texts.

Technical/Scientific challenge:

Extreme climate events cause adverse environmental and socio-economic impacts. Mitigating these impacts requires understanding how they link to climatic drivers. However, existing climate hazard impact databases have limitations in terms of completeness and accuracy. While a wealth of detailed climate impact information is available from online text sources such as news websites and reports, its extraction by human experts is time-consuming, and they are thus not typically used in impact databases. The NLP group seeks to overcome this gap by producing an impacts database of climate extremes through mining of online text sources.

Solution:

The Natural Language Processing Group at RISE was awarded 3500 GPU node hours at the EuroHPC JU system Leonardo. They will use LLMs to extract information on extreme climate events from online texts.

By using their language models (LLMs) on Leonardo, the NLP group at RISE will develop and evaluate an approach to extract impact data, such as place, time, number of people affected and monetary costs. The NLP group will use the open-source suite of Llama 2 models for this purpose and tune using an annotated set of impact data.

Scientific impact:

The NLP Group at RISE is advancing climate impact research by using state-of-the-art large language models (LLMs) on the EuroHPC JU system Leonardo. This enables systematic extraction of critical data—such as location, time, affected populations, and costs—from unstructured online text sources.

By addressing the limitations of existing climate hazard databases, this work contributes to creating a comprehensive, high-quality database essential for policymakers and disaster management professionals. It also highlights how HPC resources can transform AI research into actionable solutions, setting a benchmark for integrating LLMs in real-world applications.

Benefits:

- The supercomputing resources will enable the NLP Group to load full LLMs, rather than slimmed-down versions thereof
- The NLP Group will be able to run multiple experiments within a reasonable amount of time. Insights gained from these experiments will inform future experiments
- Faster turnaround of experiments will improve the research productivity, the capabilities of the models developed, and, in extension, the quality of the impact data extracted

Keywords: LLM, Climate mitigation, NLP, climate impact
Industry Sector: Computer science
Technology: AI

Contact:
Apostolos Vasileiadis,
apostolos.vasileiadis@ri.se

Digitizing old weather and climate data

NCC Sweden

Scientific partners involved:

The Swedish Meteorological and Hydrological Institute (SMHI) is an expert authority with a global perspective and a vital task of predicting changes in weather, water and climate. The Institute is a hub for applied research, providing forecasts, warnings and support for decision making – all based on science, advanced technology and large quantities of weather- and climate-data.

Technical/Scientific challenge:

Nearly all meteorological agencies in the world, including SMHI, possess troves of archival data of observations spanning decades in paper format. The ambition of the project is to optimize and train a sufficiently accurate machine learning model which can handle different forms of tabular data, convert handwritten-text and produce machine-readable files. This would aid and accelerate the digitization work from the paper archives into data, which is done manually as of now. As a result of the project, SMHI aims at digitizing numerous historical weather observations that will help a better understanding of climate, especially of the occurrence of extreme weather events.

Solution:

Dawsonia is a table-detection and handwritten text recognition (HTR) project aimed at data-rescue of old weather journals. It specializes in digitization of handwritten numeric data in the form of tables. Unlike simple optical character recognition (OCR) of printed prose, the project is also studying the structural layout of the text in tables and digitizing the handwritten text within them (*Figure 2*). SMHI aims to use a combination of image processing and machine learning to achieve this. The digitization pipeline is implemented in Python, using well-known open-source scientific libraries such as scikit-image and TensorFlow.

The allocation at LUMI enables the team to optimize a pipeline for the digitization. Here, optimal complexity for the different tasks (table-detection and HTR) within the pipeline will be required. Expanding the training dataset and discovering the optimal model parameters will involve several training and evaluation cycles for which GPUs are beneficial.

Business impact:

The code and the project, although young, have attracted interest both externally, from meteorological agencies abroad, and internally, from SMHI from groups dealing with paper documents in tabular format. If successful it can accelerate the process of digitization from many such archives.

Benefits:

An HPC allocation enables rapid testing and development of the product. For example, in the current version of the code, the handwritten text recognition (HTR) neural network has been tested on both a CPU cluster of 8 cores and a GPU in LUMI.

- On the CPU, training the neural network takes 11 hours.
- On the GPU the whole training takes only 1 hour.

Having a GPU at disposal allows for faster tuning hyperparameters of this model. Alternatively, using other openly available neural networks also requires some extra training (also known as transfer learning) where GPUs are handy.

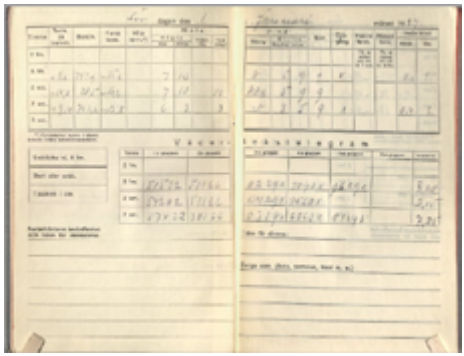


Figure 1: Example of two pages from a scanned weather journal dated 1st January 1927

Figure 2: Tables detected from the scan of Figure 1

Figure 3: Digitization pipeline

Keywords: Weather data, Climate data, Image processing, Digitisation

Industry Sector: Earth science, Environment/climate/weather

Technology: AI, HPC

Contact:

Ashwin Mohanan,

ashwin.mohanan@ri.se

Unlocking Multilingual Models

NCC Sweden

Scientific partners involved:

RISE Research Institutes of Sweden is a state-owned research institute with over 3000 employees and offices located throughout Sweden. The organization is dedicated to fostering innovation and advancing knowledge across various scientific and technological fields, to foster research and innovation to serve Swedish industrial and societal sectors towards strengthening Sweden's global competitiveness, while aligning with sustainability objectives.

Technical/Scientific challenge:

The Modular Multilingual Models (MMM) project explores modularization as a key to making models for multilingual Natural Language Processing (NLP) more effective, by delivering higher quality for more languages, and more efficient, by requiring less resources to do so. The focus is on techniques related to model architecture (static modularization) and training and tuning procedures (dynamic modularization). In addition, the interplay between these techniques and methods for improving efficiency, in particular model compression is studied. The specific aims of the project are to develop multilingual models that (a) perform on par with comparable monolingual models, (b) can be dynamically extended with new languages, and (c) can be compressed for efficient deployment.

Solution:

The project is organized into three main research themes, each of which will involve large-scale experimental studies carried out by training and evaluating large language models on the HPC cluster. The first research theme deals with modularization via model architecture, which is referred to as static modularization. The second research theme deals with modularization via training and tuning procedures, which we refer to as dynamic modularization. The third research theme deals with the interaction of modularization and different techniques for improving efficiency, in particular techniques for model compression.

Scientific impact:

Having access to supercomputers is necessary to pursue research in natural language processing today, given the methodological shift towards using large language models for essentially all natural language processing tasks.

Through this project, innovative modularization techniques are being developed, enabling models to achieve performance levels comparable to monolingual systems while supporting dynamic adaptation to new languages.

These advancements contribute to the broader NLP community by addressing significant challenges in multilingual language processing, such as scalability, efficiency, and inclusivity. By enabling efficient training, fine-tuning, and deployment, the project paves the way for more accessible and sustainable language models, particularly for underrepresented languages. The outcomes of this research will enhance multilingual communication technologies and serve as a foundation for future innovations in AI-driven global connectivity.

Benefits:

- Fast evaluation of large language models
- Dynamic modularisation
- Improving efficiency

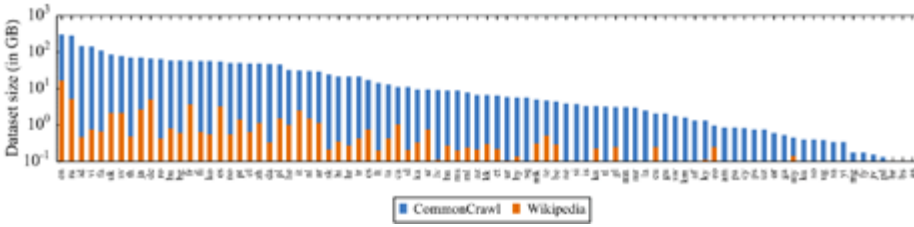


Figure 1: The image shows the amount of data available for different languages included in CommonCrawl and Wikipedia data sets for training massively multilingual language models. (From Conneau et al. [2020] Unsupervised Cross-Lingual Learning at Scale. Proceedings of ACL.)

Keywords: AI, LLM, LLM evaluation

Industry Sector: Computer science

Technology: AI

Contact:

Apostolos Vasileiadis,

apostolos.vasileiadis@ri.se

AI for law

NCC Sweden

Industrial organisations involved:

Dagg AI is a Swedish AI consulting firm that specializes in strategy development and the integration of artificial intelligence technologies to meet the business objectives of their clients. The company provides tailored solutions based on industry best practices and fosters collaborative partnerships to drive technological advancement and product innovation.

Technical/Scientific challenge:

Dagg AI is currently working with a client, Aloi.law, a forward-thinking company focused on revolutionizing the way law firms draft and manage contracts. A key challenge for Aloi.law is the need to efficiently process and analyze large datasets of legal documents. In an industry where accuracy and speed are crucial competitive factors, the ability to utilize advanced computational resources for processing vast amounts of data is a critical component of their solution.

Solution:

The project is organized into three main research themes, each of which will involve large-scale experimental studies carried out by training and evaluating large language models on the HPC cluster. The first research theme deals with modularization via model architecture, which is referred to as static modularization. The second research theme deals with modularization via training and tuning procedures, which we refer to as dynamic modularization. The third research theme deals with the interaction of modularization and different techniques for improving efficiency, in particular techniques for model compression.

Business impact:

The use of HPC resources is expected to significantly enhance Aloi.law's ability to deliver faster, more accurate contract solutions. By reducing the time required to draft and review contracts, Aloi.law will be able to offer its services at a competitive advantage, ultimately improving productivity and quality in the legal contract process. This innovation has the potential to reshape how legal information is processed and utilized, offering considerable benefits for law firms and their clients.

Benefits:

- Time Savings: By automating the drafting and review process for legal contracts, Aloi.law significantly reduces the time required to manage complex legal workflows
- Enhanced Accuracy: Advanced AI models improve the precision of contract analysis and drafting, reducing errors and minimizing legal risks
- Productivity Boost: Streamlined workflows and faster processing enable law firms to handle higher volumes of contracts, enhancing overall business productivity
- Innovative Edge: The partnership demonstrates a scalable and cutting-edge approach to integrating AI and HPC technologies, creating competitive advantage and positioning Aloi.law as a market leader

Keywords: AI, Law, Innovation, DaggAI
Industry Sector: IT/HPC systems, services & software providers
Technology: AI, HPC

Contact:
Erik Holmström,
erik.holmstrom@ri.se

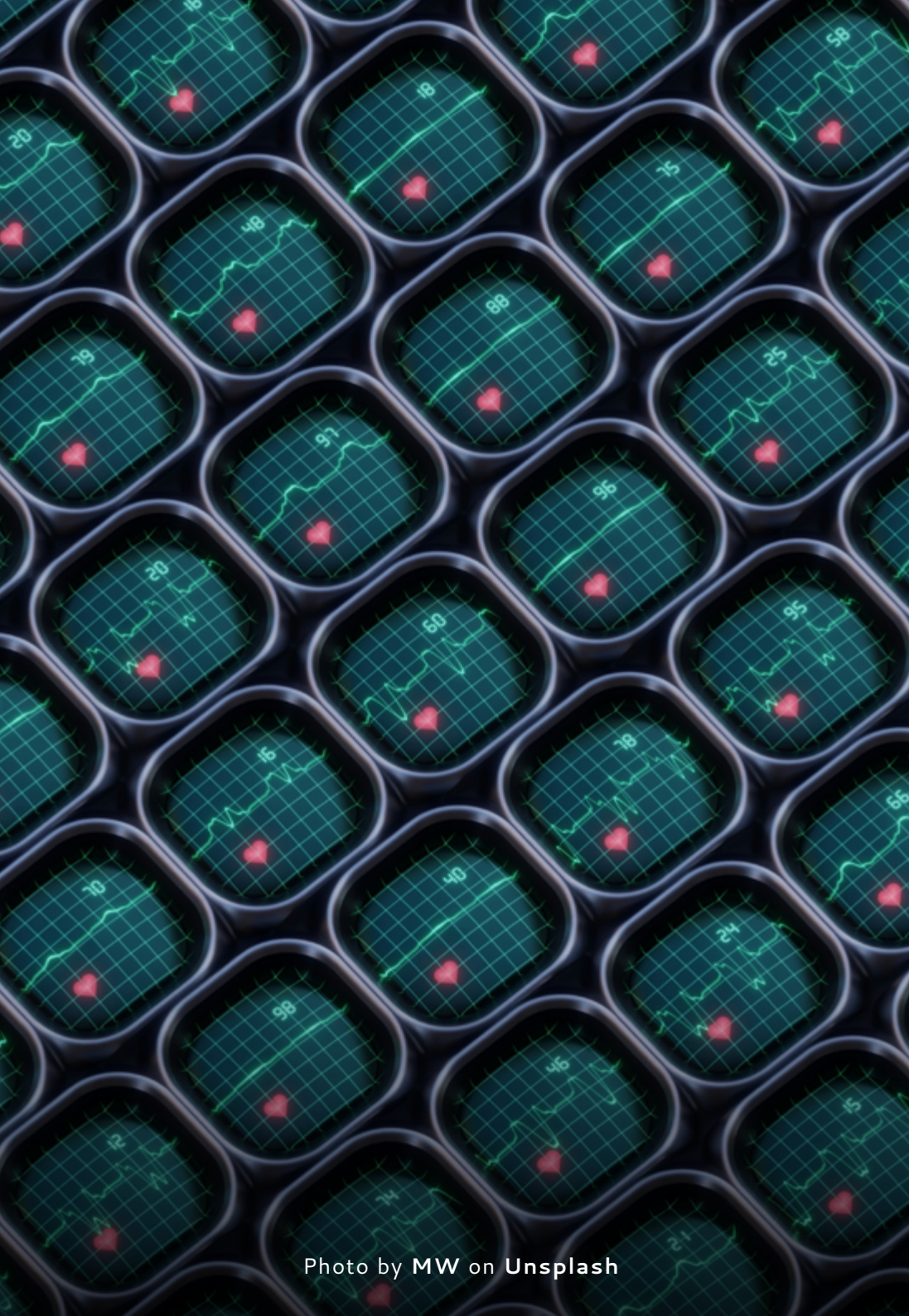


Photo by **MW** on **Unsplash**

Section 6
Health Care

Towards better disease detection using multi-million atom simulations

NCC Belgium

Partners involved:

The Laboratory of Theoretical Chemistry (LCT), directed by Prof. B. Champagne, developed an expertise in theoretical and quantum chemistry. Its research axes are centred on the development and application of methods to evaluate and interpret the properties responsible for optical, nonlinear optical, and electrical phenomena in molecules, supramolecular assemblies, polymers, and molecular crystals. Many of these investigations are integrated into multidisciplinary research projects aiming at designing new materials with outstanding properties. The laboratory is part of the Namur Institute of Structured Matter (NISM) at the university of Namur.

Technical/Scientific challenge:

Cell membranes are highly complex and heterogeneous structures containing many kinds of lipids and proteins, and their behaviour is not yet fully understood. Moreover, their composition can change due to some diseases such as Alzheimer's or Parkinson's. As atomistic details of the membranes are not reachable with experiments, computer simulations are needed to obtain an accurate model of the morphology of a membrane. However, a realistic representation of a cell membrane requires an atomistic model, considering all the atoms of molecules such as lipids, proteins, and water. Millions of atoms are needed to obtain a representative membrane composition as many different lipids are present in a membrane. Until recently, it was impossible to study such massive systems.

Solution:

Molecular dynamics (MD) simulations were performed using the NAMD & VMD codes on the pre-exascale EuroHPC supercomputer LUMI. 20 millions of CPU.hours were needed to compute about 300 nanoseconds of exploitable (production) data. The modelled membranes were composed of about 3 million atoms, including 50 kinds of lipids. Moreover, 20 chromophores were added to the membrane in order to reveal the membrane morphology and its evolution. Indeed, experimentally, as these chromophores interact strongly with light, and this interaction is affected by their surroundings, their close environment in the membrane can be probed using lasers. This relationship has been addressed by the simulations.

Scientific impact:

Different geometries of the membrane have been obtained from the MD simulations and it shows a curvature fluctuation of the membrane but no variation of the thickness. The results of the MD simulations have then been used to investigate the optical response of the chromophores using a quantum chemistry code. The spectroscopy that is used to determine the optical response of the chromophores is based on an optical process called second harmonic generation (SHG). This spectroscopy is highly sensitive to the molecular environment of the active probes and has the advantage of using low-energy photons, limiting the damage to the cell tissue. After these first two calculations, the MD and the optical response simulations, the amount of data collected was enormous. Appropriate machine learning tools were then used to determine the correlation between the lipid composition near the chromophore or the orientation of the chromophore, and its SHG response. A better understanding of how to determine the morphology of the membrane based on the SHG response has therefore been achieved. As the membrane model represented a healthy tissue, further research will be necessary to determine the morphology and the SHG response of membranes having compositions characteristic of a particular disease. These results have been published in a peer-reviewed journal.

Article: C. Bouquiaux, B. Champagne, and P. Beaujean. J. Chem. Inf. Model. 2024, 64, 518 (10.1021/acs.jcim.3c01568).

Benefits:

- Demonstration of the potential of multi-million atoms simulations for biomedical applications
- Better understanding of the atomic structure of a cell membrane
- Better understanding of the optical responses of chromophores embedded in a membrane depending on its composition

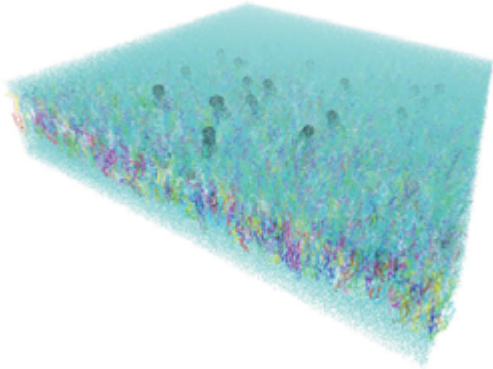


Figure 1: Overview of the membrane used in the simulation, containing 20 chromophores (in black), for a total up to 3 millions atoms.

Keywords: Chromophores, Lipids, cell membranes, nonlinear optical response, molecular dynamics, quantum chemistry

Industry Sector: Biotechnology/Bioinformatics

Technology: HPC, classical and quantum mechanics, machine learning

Contact:

contact@enccb.be

Analysis of MRI images using HPC

NCC Cyprus

Industrial organisations involved:

The German Medical Institute (or German Oncology Center) is a leading medical centre in Cyprus, focusing on patient-oriented healthcare, high quality education and research and utilising the latest medical technology.

Technical/Scientific challenge:

The German Medical Institute handles large amounts of medical datasets, whose complexity and volume necessitate substantial computational resources. Data storage and transfer to other systems (e.g. HPC systems) present additional challenges, due to the sensitive nature of the medical data involved, requiring secure, robust, data transfer protocols.

Solution:

NCC Cyprus facilitated access to HPC resources to enable efficient handling of medical images. A robust data transfer protocol was set up to move the large Magnetic Resonance (MRI) datasets to the HPC system securely and efficiently. The use of a Secure Copy Protocol (SCP) and rsync was recommended to ensure smooth data synchronization without risk of loss or corruption. The consultation also covered methodologies for processing MRI data, optimising data preparation for multicore processing and setting up a PyTorch environment for deploying deep learning models for medical image analysis.

Scientific impact:

The German Medical Institute sought to improve its capacity to analyse large volumes of MRI data, which is crucial for both patient care and medical research. The complexity and size of these datasets often exceeded the capacity of standard desktop or workstation computers, limiting the institute's ability to perform deep analyses efficiently.

The consultation through the EuroCC2 project led to substantial improvements, including optimized data transfer, efficient data loading and preprocessing, and enhanced resource utilization. These advancements have not only improved MRI data processing capabilities but have also positioned the German Medical Institute at the forefront of medical imaging research, enabling more precise diagnostics and better patient outcomes.

This success story highlights the transformative role of HPC in medical research, providing the computational power needed to handle the complexity of modern healthcare data and supporting future advancements in the field.

Benefits:

- Significant speed up in processing MRI data
- Consistency and reliability in data handling
- Optimised data transfer, processing and storage
- Enhanced resource utilization

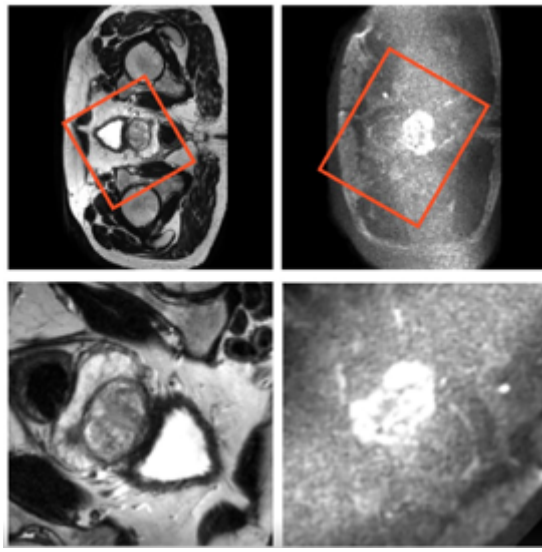


Figure 1: Easier processing of medical images through the use of HPC

Keywords: HPC for medical data, efficient handling of medical images, medical image analysis, secure data transfer

Industry Sector: Health care

Technology: HPC/HPDA

Contact:

eurocc-contact@cyi.ac.cy

Funded by the European Union. This work has received funding from the European High Performance Computing Joint Undertaking (JU) and Germany, Bulgaria, Austria, Croatia, Cyprus (co-funded by the EU within the framework of the Cohesion Policy Programme "THALIA 2021-2027"), Czech Republic, Denmark, Estonia, Finland, Greece, Hungary, Ireland, Italy, Lithuania, Latvia, Poland, Portugal, Romania, Slovenia, Spain, Sweden, France, Netherlands, Belgium, Luxembourg, Slovakia, Norway, Türkiye, Republic of North Macedonia, Iceland, Montenegro, Serbia under grant agreement No 101101903.

Automating literature search using RAG system

NCC Cyprus

Industrial organisations involved:

Medochemie Ltd. is specialized in the development, production, and distribution of generic pharmaceutical products in Cyprus., Medochemie has established itself as a leader in the pharmaceutical field, particularly in Europe, Africa, and Asia.

Technical/Scientific challenge:

Medochemie deals with significant time constraints when developing new products due to the need for a detailed and standardized process. To ensure compliance with health and safety guidelines, they must perform an extensive literature review to address critical questions required for product development. This process is both time-consuming and repetitive, as it follows the same protocol for every product.

Solution:

NCC Cyprus consulted Medochemie on how to automate this process to save time and optimize product development. A Retrieval-Augmented Generation (RAG) system was suggested to allow Medochemie to import all relevant documents into the system. The RAG system could allow processing the documents and automatically answer the protocol's required questions, ensuring that all essential information is correctly extracted and applied to the development process.

Business impact:

Implementing the RAG system could transform Medochemie's product development by automating the extensive literature review needed for compliance and safety. This change could speed up development, enabling faster product launches and boosting their competitive edge. By automating repetitive tasks, the team can focus on more strategic work, improving productivity and reducing costs. The RAG system's accurate data extraction would be used as an additional tool to minimize compliance risks. Plus, it's scalable — allowing Medochemie to expand its product line without needing more resources, making growth more efficient and sustainable.

Benefits:

- Accelerates product development by automating repetitive tasks
- Scales easily to handle increased document volume
- Allows resources to focus on strategic and innovative work

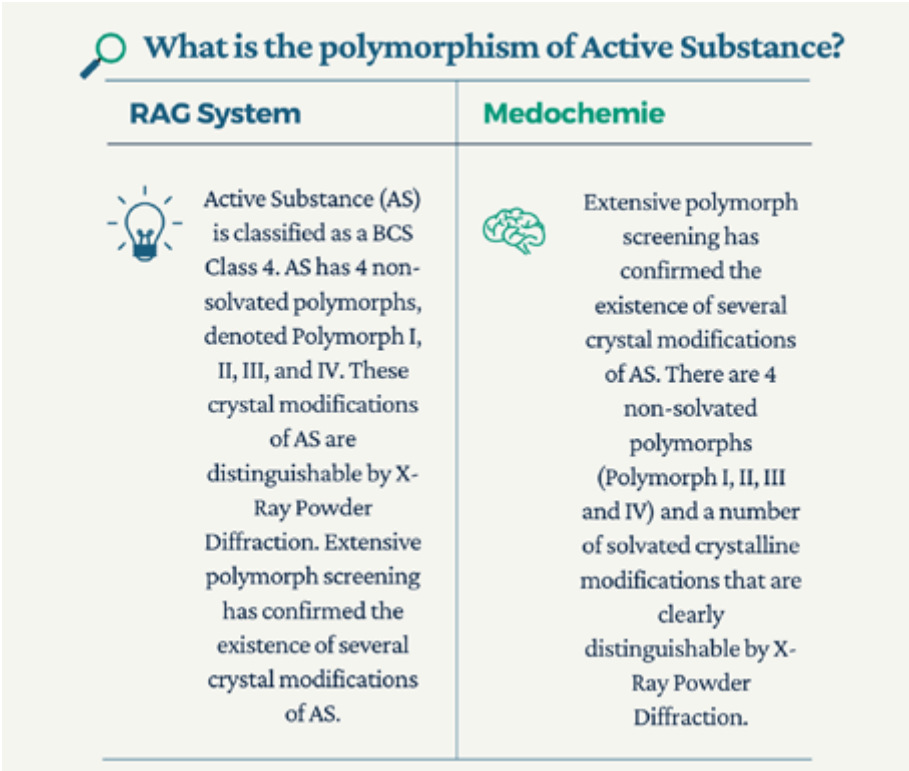


Figure 1: Comparison between Answers Provided by the RAG System and Expected Responses from Medochemie

Keywords: Pharma, Literature automation, RAG-system
Industry Sector: Health care / Pharmaceuticals
Technology: AI, LLM

Contact:
eurocc-contact@cyi.ac.cy

Funded by the European Union. This work has received funding from the European High Performance Computing Joint Undertaking (JU) and Germany, Bulgaria, Austria, Croatia, Cyprus (co-funded by the EU within the framework of the Cohesion Policy Programme "THALIA 2021-2027"), Czech Republic, Denmark, Estonia, Finland, Greece, Hungary, Ireland, Italy, Lithuania, Latvia, Poland, Portugal, Romania, Slovenia, Spain, Sweden, France, Netherlands, Belgium, Luxembourg, Slovakia, Norway, Türkiye, Republic of North Macedonia, Iceland, Montenegro, Serbia under grant agreement No 101101903.

Driving Efficiency in Radiology Workflows with Intelligent Insights

NCC Estonia

Industrial organisations involved:

Better Medicine is an Estonian medical technology company founded in 2020 that focuses on applying artificial intelligence in radiology to detect oncological findings. Unlike traditional single-organ AI solutions, Better Medicine takes a holistic, multi-organ approach, streamlining cancer diagnostics to help detect primary tumours, metastases, and unidentified disease. The company's mission is to transform the landscape of cancer diagnostics and radiology, reducing the burden on professionals and improving the outcomes for the cancer patients.

Technical/Scientific challenge:

Better Medicine faces the challenge of efficiently training advanced AI models for lesion detection, measurement, and segmentation from CT scans. This requires substantial computational power to handle large datasets, optimize model performance, and reduce training time.

Solution:

Using the HPC Centre at the University of Tartu, Better Medicine leverages high-performance GPUs enabled nodes and scalable infrastructure to train deep learning models on large imaging datasets. The computational resources enable faster experimentation, precise parameter optimization, and the development of robust AI models ready for clinical use.

Business impact:

HPC resources significantly enhance the speed and accuracy of Better Medicine's AI models, enabling rapid iteration cycles and model improvements. These models accelerate radiological workflows by automating lesion segmentation and measurement, addressing the global radiologist shortage, and expediting patient care. The adoption of HPC supports scalability, ensuring efficient handling of increasing data volumes as Better Medicine expands its solutions to broader clinical applications.

Benefits:

- Reduced model training time from weeks to days
- Improved AI model performance via rapid iteration cycles
- Enhanced scalability for future clinical datasets
- Enabling accelerated patient care through automated workflows

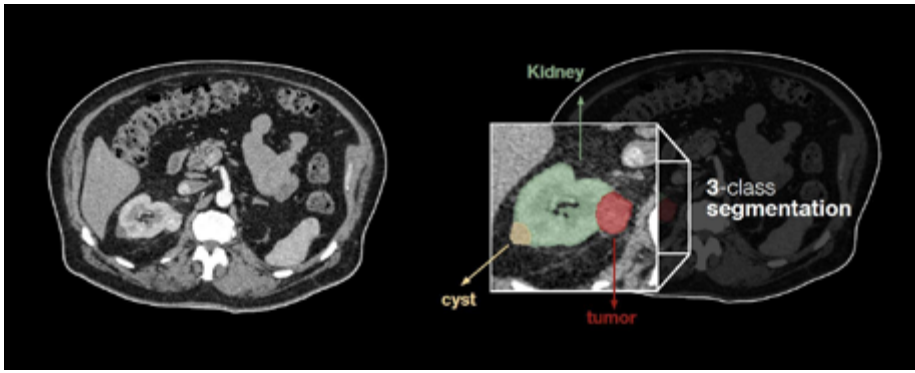


Figure 1: An AI model, developed by Better Medicine that can detect kidneys, kidney lesions, and cysts.

Keywords: AI diagnostics, Medical imaging, CT scans, Cancer diagnostics, Medical technology

Industry Sector: Biotechnology/Bioinformatics, Health care / Pharmaceuticals / Medical devices, Services & software providers

Technology: HPC, HPDA, AI, Radiology

Contact:

Ülar Allas,

ylar.allas@ut.ee

State-of-the-art advancements in quantitative MRI using HPC

NCC Germany

Scientific partners involved:

The Institute of Neuroscience and Medicine 4 (INM-4) at Forschungszentrum Jülich develops innovative methods to advance diagnostics and improve our understanding of the brain with state-of-the-art medical imaging technology, including ultra-high-field (UHF) 7T MRI. Its interdisciplinary approach in close collaboration with Jülich Supercomputing Centre (JSC) leverages cutting-edge computational resources to develop novel imaging methods for visualizing new biomarkers at higher resolutions in shorter acquisition times.

Technical/Scientific challenge:

Quantitative MRI (qMRI) measures underlying MRI parameters, enhancing sensitivity to physiological changes and enabling reliable test-retest comparability, so that observed changes reflect true physiological differences rather than scanner variability. Translating qMRI to UHF, which produces higher-resolution imaging in shorter acquisition times, entails increased field inhomogeneities and specific absorption rate, though. Novel methods developed at INM-4 address these challenges but trigger significantly higher reconstruction complexity and prohibitively long reconstruction times.

Solution:

To address these prohibitively long reconstruction times, INM-4, in collaboration with the Simulation and Data Lab Neuroscience, optimized its reconstruction code for HPC at JSC. By implementing efficient preprocessing, the reconstruction problem was converted into a slice-by-slice process that could be parallelized. Combined with automated slice processing via bash scripts, this reduced compute time from 320 hours to 8 hours per subject using HPC. This optimized workflow overcame previous limitations, enabling the application of a novel qMRI method that achieves faster scans, improved image quality, and precise parametric estimates. Subsequent measurement of a large cohort confirmed HPC-powered qMRI at UHF as a crucial step toward clinical feasibility.

Scientific impact:

Implementing qMRI at UHF required a novel imaging method to address increased field inhomogeneities and specific absorption rate. This method enabled faster scans and improved image quality but required solving a complex reconstruction problem with high computational demands. Using conventional hardware, reconstruction was prohibitively slow—8 hours per slice for 160 slices per subject—delaying evaluation of its clinical potential.

To overcome this, the team turned to HPC and consulting services at SIDE to adapt their software to harness HPC's power. Parallelizing tasks and automating processes reduced compute time from 320 to just 8 hours per subject, making it feasible to apply the novel qMRI method, which had been constrained by slow reconstruction.

HPC thus enabled the method's practical use, improving image quality and parametric accuracy. This brings qMRI at UHF closer to clinical application, enhancing diagnostics. Further optimization—such as AI-driven image reconstruction—could eventually make it viable in routine clinical settings without direct HPC access, benefiting patients through more precise and timely diagnostics.

Benefits:

- HPC reduced reconstruction time from 320 to 8 hours per subject, making qMRI feasible
- The optimized qMRI method enabled faster scans, improved image quality, and more precise parametric estimates
- Collaboration with JSC enhanced computational efficiency and set the stage for AI-driven qMRI

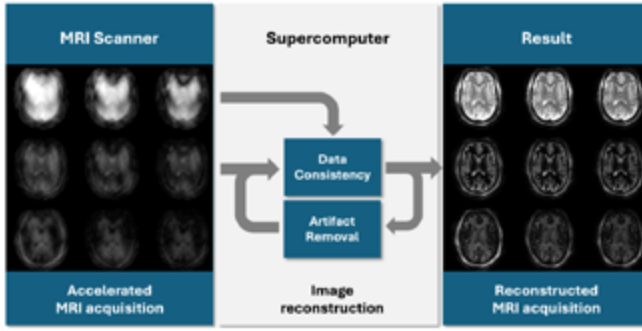


Figure 1: In accelerated acquisitions, undersampling artifacts degrade image quality. The new qMRI method addresses UHF challenges by acquiring numerous MR images with slightly varying scanner settings. To maintain a short acquisition time, each image is highly accelerated, resulting in severe undersampling artifacts. The reconstruction algorithm corrects these artifacts through extensive computation, requiring a supercomputer. The final output is high-quality, artifact-free images, from which quantitative parameters are estimated.

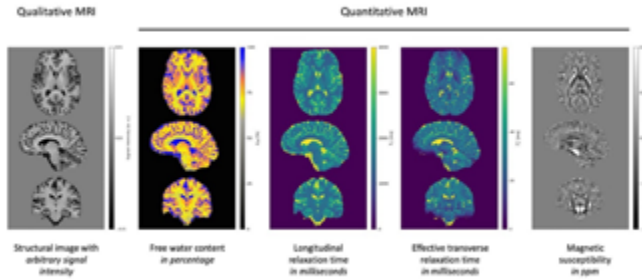


Figure 2: In vivo results showing a qualitative structural image and quantitative water content map Cw, T1 map, T2 map, and magnetic susceptibility map χ acquired with the QRAGE method.*

Keywords: Quantitative MRI (qMRI), Ultrahigh field (UHF) MRI, High-performance computing (HPC), Image reconstruction, AI, Parallelization

Industry Sector: Health care / Pharmaceuticals / Medical devices, IT/HPC systems

Technology: HPC, HPDA, AI

Contact:

Prof. Dr. Markus Zimmermann,

markus.zimmermann@fh-aachen.de

contact@supercomputing-in.de

This project has received funding from the German Federal Ministry of Research, Technology and Space under grant agreement No 16HPC096K. The sole responsibility for the content of this publication lies with the authors.

Harnessing HPC in Cancer Genomics: High Throughput Mutated Protein Structure Prediction using Alphafold2

NCC Hungary

Industrial organisations involved:

DeltaBio 2000 Ltd. is a pioneering genomics and molecular diagnostics company with over 15 years of experience in the field. Utilizing next-generation sequencing (NGS), the company operates five sequencers and maintains a diverse team of molecular biology, sequencing, and bioinformatics experts.

Technical/Scientific challenge:

The integration of next-generation sequencing in cancer diagnostics has revealed thousands of missense variants, many of which are classified as variants of uncertain significance (VUS) due to their unknown effects on protein function and disease risk. Although numerous algorithms have been developed to classify VUS, few rely on protein structure-based features. This project aimed to expand the pool of available mutated protein structures using AlphaFold2, the Nobel Prize-winning algorithm for protein structure prediction. However, the process of generating thousands of native and mutant structures of clinically relevant genes within 6-12 months posed a major computational challenge. Running AlphaFold2 on a standard computer is complex, requiring over 2.6TB of disk space and intensive CPU and GPU computational power for multiple sequence alignment, model generation, and relaxation steps. Access to the Komondor HPC enabled us to timely explore biological questions that are impossible to analyze with conventional computing resources.

Solution:

To predict a final total of 77,713 protein structures (native and mutated), the powerful CPU-only and GPU partitions of DFK's Komondor HPC were utilized. The CPU-only partition, equipped with 184 nodes featuring two AMD EPYC 7763 CPUs each, enabled rapid multiple sequence alignment and structural template searches even for proteins up to 2,500 amino acids long. The GPU partition, consisting of 58 nodes with one AMD EPYC 7763 CPU and four NVIDIA A100 40 GB GPUs per node, enabled the prediction ofAlphaFold2's five model iterations and GPU relaxation for all the proteins in our dataset. Furthermore, the presence of two separate partitions allowed for the utilization of ParaFold (HPC-friendly implementation of AlphaFold2), which further optimized the process by enabling parallel CPU and GPU processing.

Business impact:

Having access to the unique information harboured by the mutated predicted structures enabled the bioinformatics team at Delta Bio 2000 Ltd. to develop novel structure-derived features that are now being used to train machine learning models for interpreting variants of uncertain significance – an ongoing challenge in cancer genomics. Previous efforts focused primarily on the native protein structures, providing a limited perspective. Additionally, this expanded dataset allowed for the establishment of novel collaborations with other leading European laboratories in the field.

In future steps, Komondor's powerful GPU resources will be essential for testing other approaches like unsupervised learning. As unsupervised algorithms for variant classification gain popularity, integrating them with features derived from the generated protein structures is a key step toward DeltaBio's goal of developing novel tools for missense variant classification.

The ability to rapidly train, optimize, and evaluate unsupervised models is crucial for Delta Bio Ltd. to remain competitive. This enables the company to identify potential issues early in the training process and quickly respond by resolving them or exploring alternative approaches.

Benefits:

- Time saving: Rapid prediction of >70,000 native and mutated protein structures
- Pivotal: Enables the exploration of novel biological questions
- High throughput: Rapid and efficient feature extraction from the generated protein structures
- Robustness: HPC's powerful partitions improved the prediction of additional datasets in a short time

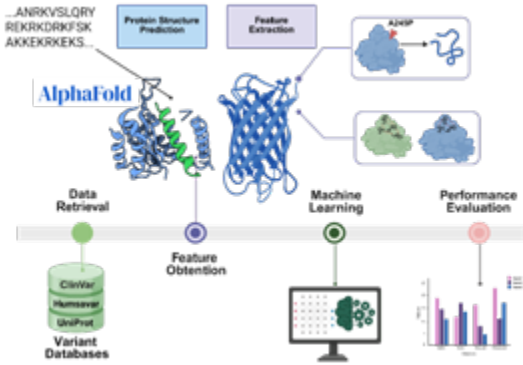


Figure 1: Pipeline of structure prediction of native and mutant proteins using AlphaFold2, structure-feature development and extraction, machine learning training, and evaluation. Created with BioRender. @ 2024 Erda Qorri.

Figure 2: DeltaBio 2000 Ltd. Team. @ 2024 DeltaBio 2000 Ltd.

Keywords: Cancer genomics, AlphaFold2, protein structure prediction, Machine Learning, variants of uncertain significance

Industry Sector: Biotechnology/Bioinformatics

Technology: Machine Learning feature development focused on cancer variant classification, HPC

Contact:

Lajos Pinter, lajos.pinter@deltabio.eu

Erda Qorri, alexandria.kouri@deltabio.eu

Knee cartilage donor-recipient geometric matching with AI

NCC Hungary

Industrial organisations involved:

MedInnoScan Research and Development Ltd. has been creating various applications of machine learning and artificial intelligence since 2017 to offer solutions in medicine, such as diagnosis, therapy suggestion, and automated follow-up monitoring since 2017. The company has cooperated with more than 100 health care institutions on project such as chronic wound assessment and bandage type suggestion, patient self-monitoring of wounds with alerts generated when indicated, and knee cartilage donor-recipient geometric matching, which is the topic of the current story. They have spun out this project into GenuForm AI Ltd., whose logo is visible here.

GenuForm AI

Technical/Scientific challenge:

Matching the geometry is of paramount importance when restoring knee cartilage surfaces with transplantation as it's a rigid tissue. (Immunology is not a factor for cartilages.) Do it well and the patient regains full function for life.

GenuForm's multi-part AI solution optimizes the matching from MR images and recommends which part of the donor cartilage would best serve which recipient, thus optimizing the match and availability of grafts.

Solution:

The first AI part of GenuForm's solution determines the existing geometry of the prospective recipient. The second AI part "digitally restores" the damaged cartilage part, it determines what a healthy tissue should look like. The difference between the two provides the required restoration geometry. Margins are added to this tomatch the available surgical instrument shapes, which yields the final required graft geometry.

The third AI part of the solution determines the existing geometry of the donor when tissue becomes available. The final, fourth part, a combinatorial geometry algorithm then determines if any part of the donor tissue matches the required graft geometry within the provided tolerances. Generally, multiple recipients can be served from one donor cartilage, and the algorithm proposes cutting up the donor tissue into multiple pieces to optimally utilize available tissue while taking provided medical and ethical parameters into account.

Business impact:

The use of HPC allowed the machine learning algorithm and AI to be trained much more quickly and efficiently. Not only did the calculations run faster but with the vastly increased operating memory available the size of the training set that could be processed in each step was significantly enlarged as well. This allowed for much quicker turnaround times on experiments, greatly increasing the hyperparameter exploration spectrum we could traverse. This finally resulted in the AI algorithm's superior performance compared to earlier versions trained in more resource limited environments.

Benefits:

- Perfect geometric match enhances operation success probability
- Improves availability of tissue for transplantation
- Efficient use of a scarce resource: one cartilage can serve 2-3 patients in an optimal case
- Ex-vivo time for cartilage is significantly reduced

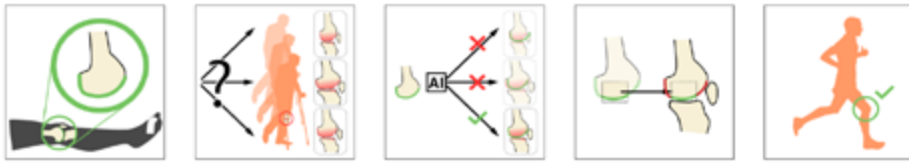


Figure 1: Having harvested available tissue from a suitable cadaver, the AI algorithm determines which patients' geometries the available tissue would match best © 2024 GenuForm AI

Figure 2: A 3d representation of donor and recipient tissues showing the matched areas © 2024 GenuForm AI

Figure 3: The team: Peter Szucs, Dora Torok, Endre Szabo, Prof. Dr. Laszlo Hangody, Tamas Frisch, Peter Szoldan, Andras Solti, Dr. Zsofia Egyed, Christian Szegedy, Adrienn Gogo, Dr. Gyorgy Hangody © 2024 GenuForm AI

Keywords: Knee; knee cartilage; transplantation; MRI; geometry matching; software as medical device
Industry Sector: Health care - Medical device scan
Technology: AI, HPC

Contact:
Peter Szoldan,
peter.szoldan@medinnoscan.com

Developing Optimized Drugs Against COVID-19

NCC Portugal

Scientific partners involved:

BioSIM, a research group, bridges theory and experiment using advanced computational tools for Enzymatic Catalysis and Drug Discovery, while Crowdfight, a non-profit dedicated to fostering scientific collaborations by connecting experts to complement research projects and documenting the credit due to every participant.

Technical/Scientific challenge:

The challenge was to identify molecules capable of inhibiting the interaction between the SARS-CoV-2 Spike protein and human ACE2 receptors to prevent viral infection. With 200,000 molecules to screen, the task required high computational power to identify promising candidates for drug development against COVID-19 and its variants.

Solution:

To tackle this challenge, researchers used high-performance computing (HPC) and specialized Graphical Processing Units (GPUs) to simulate the atomic-scale interaction between the Spike protein and ACE2 receptors. Computational resources enabled the testing of 200,000 molecules, narrowing down the candidates to 20–30 molecules with high inhibitory potential. These molecules were later optimized and experimentally tested, leading to the discovery of two promising inhibitors, which were patented.

Scientific impact:

Using HPC allowed the researchers to significantly accelerate the screening process and evaluate a vast number of molecules, which would have been impossible with conventional methods. This approach helped identify two promising candidates for drug development in record time. The computational models are being tested against new SARS-CoV-2 variants, and initial results show the inhibitors are still effective. This research could lead to the creation of optimized drugs with fewer side effects and high efficacy, potentially transforming COVID-19 treatment and providing a defense against future variants. The impact of HPC in this project not only sped up drug discovery but also laid the groundwork for ongoing virus inhibition research.

Benefits:

- Accelerated drug discovery
- Optimized molecule selection
- Enabled 3D modeling of virus-protein interactions
- Facilitated collaboration
- Ensured drug effectiveness against new variants

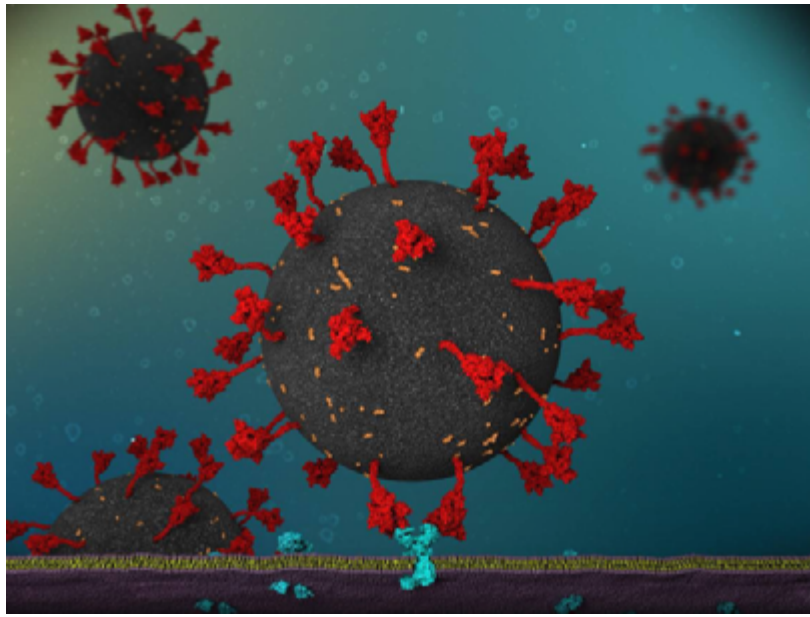


Figure 1: The interaction between SARS-CoV-2 Spike – Human ACE2

Keywords: COVID-19, HPC, drug discovery, molecular inhibitors
Industry Sector: Health Care / Pharmaceuticals /Medical Devices
Technology: HPC, HPDA

Contact:
Lígia Breda Melo / NCC Portugal,
ligiabredamelo@lip.pt / eurocc-portugal@lip.pt

Photo by **Aleksei Smagin** on **Unsplash**

Section 7
Other Fields

Interdisciplinary Studies of Human Past

NCC Estonia

Scientific partners involved:

The Institute of Genomics of the University of Tartu is the leading hub of genomics and related transdisciplinary research in Estonia. The Institute of Genomics was established to promote greater synergy between genomic-based research and other scientific fields. The research group focuses on studying the genetic diversity of modern humans and evolutionary background using genome data from both contemporary and ancient humans in a global and regional context.

Technical/Scientific challenge:

Archaeogenomics aims to trace the migration, adaptation, and interbreeding of ancient human populations. This involves analysing genetic diversity across different geographical regions and time periods. The human genome consists of about 3 billion base pairs, and sequencing technologies generate massive amounts of raw data (often in the range of terabytes or more) that has to be stored, processed, analysed and shared. However, effective analysis and data storage of such huge amounts of data is unthinkable using conventional computers.

Solution:

In archaeogenomics, which focuses on analysing ancient DNA to understand the genetic history of past human populations, high-performance computing (HPC) plays a crucial role in addressing the unique challenges of working with ancient genetic data. Researchers extract and process biomolecules from archaeological material and soil in the ancient DNA cleanroom and analyse data using the bioinformatics toolbox. HPC systems provided by the University of Tartu HPC centre offer computational power needed to efficiently process and analyse these datasets. HPC enables faster and more accurate genome assembly, alignment, and variant calling, which are key tasks in genomic analysis.

Scientific impact:

The Archaeogenomics Research Group is a partner of Estonian Roots, a Centre of Excellence for transdisciplinary studies on ethnogenesis and cultural diversity. The Centre of Excellence brings together researchers from humanities and scientists to investigate the origins of cultural and genetic diversity in Estonia and neighbouring areas. To understand human demographic history and the spread of epidemics, genetic results are analysed in the context of natural and anthropogenic environments to explore the interactions and co-evolution of biological and cultural evolution. All the necessary DNA analyses are carried out using the facilities of the University of Tartu HPC Centre, illustrating the fact that the project would not be feasible without HPC solutions.

Benefits:

- HPC plays a key role in genomic research for efficient handling and processing of large-scale genomic data
- Thanks to powerful HPC solutions, accurate data analysis can be performed in a short time. In addition, population scale analysis can be conducted using HPC.
- Flexibility in analytical modules: it is possible to install different modules that allow performing various analyses that are suitable for modern DNA and for ancient DNA

Figure 1: Rocket, the HPC cluster of the University of Tartu used daily by researchers at the Genomics Institute.

Keywords: Archaeogenomics, Estonian Biobank, Genetic history of Uralic speakers, Interdisciplinary Studies, Genomic Data

Industry Sector: Biotechnology/Bioinformatics, Life sciences

Technology: HPC, HPDA, Next Generation Sequencing

Contact:

Ülar Allas,

ylar.allas@ut.ee

Harnessing AI to reduce emissions: Supersight's vision for the future of real estate

NCC Finland

Industrial organisations involved:

Supersight Oy is a leading developer of privacy-preserving computer vision technology. The Finland-based company is committed to develop cutting-edge technology that provides people flow analytics as a service for real estate occupancy optimisation. The technology brings efficiency to real estate management and helps provide for a more sustainable future.

Technical/Scientific challenge:

"Companies are wasting hundreds of billions on empty office space. If the use of office buildings could be better optimised, numerous client companies could save millions and reduce energy related emissions by eliminating unnecessary space," says Supersight, describing the correlation between real estate business, energy and emissions.

To solve this problem of unused office space, Supersight set out to develop machine vision model. Processing large-scale visual video stream datasets of people entering and leaving the office premises effecting the people flow analysis, combined with the development of a sophisticated AI model requires parallel computing to accelerate the training of the necessary neural networks.

Solution:

Supersight's new improved solution comprises a mobile phone, a supercomputer, and a working AI model. Android mobile phones are used as computing sensors, the most cost-effective and cyber-safe solution for measuring the use of the built environment. Processing of large-scale visual datasets combined with the development of a sophisticated AI model requires parallel computing to accelerate the training of the neural networks for the development of their own AI model. Supersight got access to the LUMI supercomputer and expert support in using the software.

Business impact:

"With the LUMI supercomputer, we significantly accelerated the testing and modification of our model. This was absolutely crucial in our quest for 99% accuracy in modelling real space usage. With LUMI, we were able to run multiple training processes simultaneously, which ensured that our final model was optimised for performance and scaled well to a variety of environments," tells Supersight.

Benefits:

- Breakthrough technology innovation resulting in a real time people flow analysis tool while ensuring privacy
- Accelerated R&D processes in 99% accurate AI model development
- Reduced energy related emissions in real estate business
- Reduced costs for companies renting real estate costs

Keywords: RDI, People flow analytics, Sustainability, Neural networks, Real estate
Industry Sector: AI technology for people flow analytics
Technology: HPC LUMI, AI (model optimisation)

Contact:
Dan Still,
Dan.Still@csc.fi

The acquisition and operation of the EuroHPC supercomputer is funded jointly by the EuroHPC joint Undertaking, through the European Union's Connecting Europe Facility and the Horizon 2020 research and innovation programme, as well as the Participating States FI, BE, CH, CZ, DK, EE, IS, NL, NO, PL, SE.

Leverage from the EU 2014–2020

This project has received support from Business Finland.

Artificial Intelligence-supported Human Pose Estimation

NCC Finland

Industrial organisations involved:

Top Data Science Ltd specializes in artificial intelligence, machine learning, and software engineering services. The majority of its customers are based in the Nordic countries, Germany, and Japan. The biggest solution segment is based on computer vision technologies, which utilize video and other image data to collect information relevant to the customer use case in focus.

Technical/Scientific challenge:

Top Data Science Ltd was chosen as a partner in the AISA (AI-based Situational Awareness) project, funded by Business Finland and led by Nokia. This project aims to enhance AI application and high-speed wireless communications among Finnish industrial companies and research institutes. The company is focusing on urban security and industrial applications by detecting poses, activities, and counting people from video feeds using deep learning.

Solution:

Top Data Science leveraged LUMI for human pose estimation analysis, processing hundreds of video clips in parallel during neural net training to create a dataset of human skeleton trajectories for fast inference.

"Overall, LUMI provided a powerful and effective platform for our projects, meeting requirements, and facilitating workflow," explains Top Data Science. They add: *"We were able to effectively outsource computationally heavy tasks and generate the desired human skeleton trajectory dataset, which is now actively utilized in our work."*

LUMI's capability to run parallel jobs on GPUs and its flexible containerization approach offered the needed scalability. Additionally, its comprehensive documentation with plenty of examples and easy navigation allowed them to find the necessary information without extra support.

Business impact:

Supercomputing enabled the company to train neural nets efficiently for quick inference. *"This [using LUMI] helps us address customer problems and deliver tailored solutions that go beyond what the normal computational setup offers. It's crucial to understand LUMI's strengths and limitations, such as its suitability for inference tasks. Comprehensive documentation and containerization features make it an asset for various projects,"* says Top Data Science. LUMI's advanced capabilities can enhance workflow and drive innovation, provided its resources are leveraged effectively.

Benefits:

- Scalability
- Good functionality and ease of use of AI tools and container environment
- Outsourcing of computationally heavy tasks
- Reduced model training time

Keywords: Machine learning, Automation, Human pose estimation, LUMI, HPC

Industry Sector: AI Techology, services & software providers

Technology: HPC, AI

Contact:

Dan Still,

Dan.Still@csc.fi

The acquisition and operation of the EuroHPC supercomputer is funded jointly by the EuroHPC joint Undertaking, through the European Union's Connecting Europe Facility and the Horizon 2020 research and innovation programme, as well as the Participating States FI, BE, CH, CZ, DK, EE, IS, NL, NO, PL, SE.

This project has received support from Business Finland.

Developing quantum algorithms for the near future with HPC

NCC Finland

Industrial organisations involved:

QMill develops quantum-advantage algorithms which will be executed on near-term quantum computers to solve complex problems that are too large for existing supercomputers.

Technical/Scientific challenge:

The challenge that the company is working on is to find completely new and viable algorithms for future quantum machines, thereby reducing the time taken to beneficial use of quantum computing. Algorithms cannot be easily tested directly on quantum computers. Classical HPC helps the company now to get their quantum algorithms ready in a couple of years, meeting the expected market needs. With their algorithms, QMill algorithm development targets industrial optimization customers from different industries and development aims for optimization, cost savings and efficiency, but also for environmental friendliness and energy savings. Indeed, even if there are still few operational quantum machines globally, but the size of the quantum industry market is estimated to be approximately one billion euros.

Solution:

With LUMI, QMill is testing whether their quantum circuits work correctly and efficiently, and how they scale when the number of qubits are increased. QMill also uses LUMI to benchmark quantum algorithms, i.e. to see how well a problem is solved by classical algorithms on a supercomputer and compare the results with a quantum algorithm.

Business impact:

"We are constantly looking for algorithm ideas, verifying them mathematically and by simulation and then computing with real quantum computers. We also benchmark against classical computing all the time," says QMill describing the development process.

"Our goal is to validate the first useful quantum algorithm in about three years. We will benchmark the algorithm against all the best high-performance computing algorithms and hope that it performs better than any of them. Time will tell if we have a working quantum machine three years from now to demonstrate our results," QMill adds.

Benefits:

- Quantum technology and algorithms benchmark with HPC
- Reduced energy consumption when algorithms run on future quantum systems
- Smooth simulations
- Quantum ecosystem development

Keywords: RDI, Quantum, Quantum algorithms, Emulation, Benchmark
Industry Sector: Quantum, software development
Technology: HPC, Quantum computing

Contact:
Dan Still,
Dan.Still@csc.fi

The acquisition and operation of the EuroHPC supercomputer is funded jointly by the EuroHPC joint Undertaking, through the European Union's Connecting Europe Facility and the Horizon 2020 research and innovation programme, as well as the Participating States FI, BE, CH, CZ, DK, EE, IS, NL, NO, PL, SE.

Leverage from the EU 2014-2020

This project has received support from Business Finland.

PAID-T: Advanced Trading Simulations powered by HPC

NCC Montenegro

Industrial organisations involved:

PAID MNE specializes in crafting scalable software solutions for investment firms, leveraging advanced algorithms, machine learning, and artificial intelligence to optimize trading strategies and risk management. Their PAID-T trading solution dynamically adapts to market fluctuations, offering optimized trading experience.

INVT provides software development services with an emphasis on AI innovation. Their multidisciplinary team pushes IT solution boundaries to maximize user benefits across various sectors.

Technical/Scientific challenge:

The PAID-T (Price Action Intelligent Detection Trading) project faced a technical challenge in optimizing complex trading algorithms by running millions of simulations to fine-tune parameters for precise asset price prediction. This required processing minute-level OHLC datasets for multiple trading pairs, exceeding standard system capabilities. The solution involved scaling from single-node multiprocess execution to a multimode HPC environment, necessary for performance optimization and efficient resource management.

Solution:

To optimize trading algorithms with complex parameters, the LUMI supercomputer's capabilities were leveraged. Project dependencies were adapted to align with LUMI's default Python module, ensuring seamless compatibility. Multinode execution was enhanced using mpi4py with a task-pull paradigm, implementing a state machine with a master-worker model for efficient task distribution and execution. This enabled real-time result collection and resource optimisation, achieving 1.2 million simulations in 4.7 hours, with single simulations reduced to 54 seconds and parallel runs averaging 58 seconds, showcasing substantial improvements in speed and efficiency.

Business impact:

The integration of the LUMI supercomputer capabilities has transformed PAID's financial market analysis. Previously constrained by standard computing resources and lengthy processing times, the system now enables rapid processing of billions of historical transactions.This enhancement enables the efficient identification of critical market patterns, providing a strong foundation for optimizing trading strategies with precision.

HPC-powered simulations have enabled the refinement of trading strategies across diverse market conditions, consistently achieving and exceeding key performance indicators (KPIs). Target KPIs assumed an average annual return of 30%, with maximum drawdowns remaining below 10% and contained within a single quarter. Additionally, maintaining a Sharpe ratio of at least 1.5 and a Sortino ratio of at least 2 are essential benchmarks, achieved through advanced computational insights and the robustness provided by supercomputer LUMI.

These advancements will support substantial business growth by improving the reliability of investment strategies. Enhanced analytical capabilities, powered by high-performance computing, strengthens market adaptability and competitive edge in the financial sector.

Benefits:

- Time savings - Reduced data analysis time from several days to just hours, enabling faster decision-making and improved responsiveness
- Product optimization – Enhancing the effectiveness, accuracy, and reliability of trading strategies
- Cost Efficiency – Optimized resource utilization, minimizing wastage while maximizing computational performance and output

Figure 1: Orchestration of trading simulations for multinode multicore environment

Figure 2: Obtained results of an optimized trading strategy

Keywords: Financial market analysis, trading algorithms, predictive modelling, data processing optimization, trading simulations

Industry Sector: Finance/Insurance

Technology: HPC, Parallel processing, Grid search, Algorithm fine-tuning, Python

Contact:

nebojsa.janovic@paid-t.com

luka.filipovic@udg.edu.me

Redefining math learning using data and generative AI

NCC Sweden

Industrial organisations involved:

Mappi is a maths learning website for high school students that aims to redefine Maths Learning using Data and Generative AI.

Technical/Scientific challenge:

Mappi follows along students' work, looking for strengths and blind spots, in order to generate tailored studying material and tutoring.

Solution:

Education, and specifically the learning of mathematics, is an individual process and inherently of a non-deterministic nature. This is why the classical approach based on scheduled learning can be a problem. Working with LLMs by fine-tuning a general model and converting it to a math tutor lets bypass this problem. Doing that on a 70B parameter model on your own laptop is a no-go if one wants to avoid frying their not-so-fine-tune-friendly GPU.

Business impact:

In particular, in the development process the biggest difference is how Mappi works towards more qualitative outputs. Before being able to access fine tuning possibilities the path to more qualitative outputs is limited to primarily more advanced prompting. Having access to high-performance computing opens the door to a development process more focused on implementing specific fine-tuned models on specific learning tasks.

Benefits:

- Fine-tune AI models fast and efficiently
- Fast iterations using heavy LLM (70B)

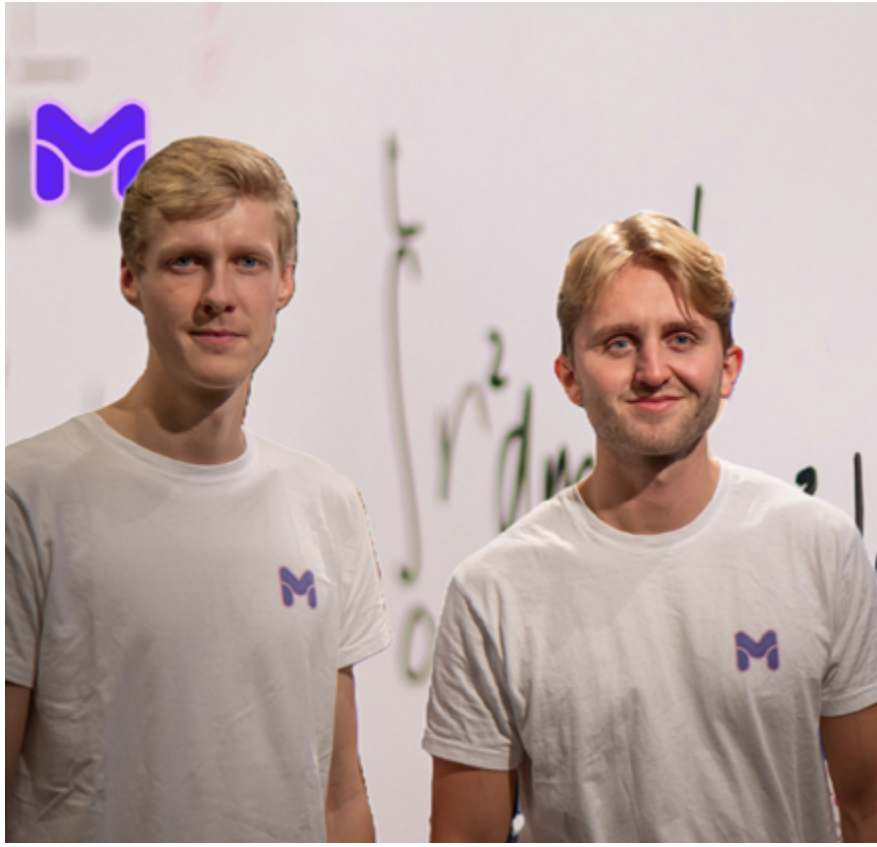


Figure 1: Mappi Co-Founders Karl Flintberg and Andreas Ahlin

Keywords: Mappi, AI math tutor,fine tuning, gen ai
Industry Sector: Education
Technology: AI

Contact:
Erik Holmström,
erik.holmstrom@ri.se

Advancing Sustainable Networking with AI

NCC Sweden

Scientific partners involved:

RISE Research Institutes of Sweden is a state-owned research institute with over 3000 employees and offices located throughout Sweden. The organization is dedicated to fostering innovation and advancing knowledge across various scientific and technological fields, to foster research and innovation to serve Swedish industrial and societal sectors towards strengthening Sweden's global competitiveness, while aligning with sustainability objectives.

Technical/Scientific challenge:

As future communication and computing networks grow in scale, optimizing resource allocation becomes increasingly vital to ensure energy efficiency and sustainability. Modern networks face NP-hard challenges, such as balancing energy, spectrum utilization, and latency while meeting application demands.

A prominent use case is scheduling in IoT networks augmented with backscatter communication. These networks rely on battery-free sensor tags, which require carefully planned schedules for carrier provisioning from neighboring devices. Solving this scheduling problem at scale, while maintaining efficiency and adaptability, is a critical challenge for sustainable digital infrastructure.

Solution:

The project developed an AI-driven scheduler to address resource allocation challenges in IoT networks. Using Deep Learning, Graph Neural Networks, and Transformers, the scheduler generates efficient carrier schedules by learning from optimal solutions in small IoT networks and generalizing to networks up to 100× larger without retraining. With the Meluxina HPC cluster, scaled training pipelines for both simulated and real-world network topologies were developed.

The HPC resources allowed training over 50 ML models and identify the hyperparameters and embedding configurations that allowed the model to robustly generalize to such large networks. Optimizing workflows for GPU acceleration on NVIDIA A100 GPUs, the scheduler achieves polynomial runtime complexity, enabling it to dynamically adapt to changing network conditions. This solution not only scales effectively but also significantly reduces energy and spectrum utilization.

Scientific impact:

Access to the MeluXina supercomputer has been instrumental in accelerating research and innovation in energy-efficient networking. The scheduler achieves up to 2× reduction in energy and spectrum usage, offering a scalable solution for large-scale IoT networks. These improvements support sustainable digitalization while positioning RISE as a leader in AI-powered optimization for network resource allocation.

Moreover, the planned open-source release of the developed code, models, and datasets ensures that the work benefits the wider research community and industrial stakeholders, fostering real-world applications in diverse sectors like smart cities, healthcare, and industrial IoT.

Benefits:

- Time savings: Substantial reduction in training and experimentation cycles using HPC-enabled parallel training, allowed to widen the hyperparameter exploration space
- Cost savings: The developed solution achieves up to 2× better resource utilization in the studied IoT networks
- Product optimization: Scalable scheduling solutions adaptable to dynamic network conditions
- Sustainability: Enhanced energy efficiency for future digital infrastructures in the context of 6G networks

Keywords: Network optimisation, AI, optimisation

Industry Sector: Computer science

Technology: AI, HPC

Contact:

Apostolos Vasileiadis,

apostolos.vasileiadis@ri.se

CHAPTER 3

THE CENTRES OF EXCELLENCE

THE CoEs

The Centres of Excellence (CoEs) are funded through the EuroHPC Joint Undertaking (see p.168 for details).
For an overview of all current CoEs, see https://hpc-portal.eu/coes.

CoE ESiWACE3

ESiWACE3 – the Centre of Excellence in simulation of weather and climate in Europe – focuses on supporting the weather and climate modelling community in achieving a higher readiness level regarding exascale supercomputing and fostering knowledge transfer between different Earth System modelling centres and teams across Europe. To achieve this aim, climate modelling groups have collaborated with High-Performance Computing (HPC) centres and partners from the technology industry to enhance all aspects of the weather and climate modelling workflow.
https://www.esiwace.eu/

CoE EXCELLERAT P2

EXCELLERAT P2 is the European Centre of Excellence for Engineering Applications, driving the application of High-Performance Computing (HPC) in engineering across Europe.
https://www.excellerat.eu

CoE HiDALGO2

HiDALGO2 is the Centre of Excellence in HPC and Big Data technologies for Global Systems funded by the European Commission. It aims to explore synergies between modelling, data acquisition, simulation, data analysis and visualisation along with achieving better scalability on current and future HPC and AI infrastructures. The main application areas cover simulation of urban air pollution and comfort, urban building energy, prediction of renewable energy production, wildfires, and material transport in water.
https://www.hidalgo2.eu/

CoE MaX

MaX is a European Centre of Excellence that enables materials modelling, simulations, discovery, and design at the frontiers of the current High-Performance Computing (HPC), promoting the use of exascale and post-exascale computing capabilities.
MaX's current challenge lies in redesigning the most used open-source community codes in quantum materials simulations and the related data ecosystem to take full advantage of the exascale technology.
https://www.max-centre.eu/

CoE MultiXscale

MultiXscale is a Centre of Excellence funded in exascale-oriented application co-design and delivery for multiscale simulations. One of its targets is to increase productivity for computational scientists through the EESSI software environment, ensuring performance and portability of scientific software across HPC systems.
https://www.multixscale.eu/

Accelerating the Dutch atmospheric large-eddy simulation achieving a 12x speedup for a high-resolution local weather model with modern graphics cards

CoE ESiWACE3

Partners involved:

TU Delft: the team leading the development of DALES, an atmospheric physics simulation code, was awarded support within ESiWACE3 to port the code to GPU, paving the way for ExaScale.
Netherlands eScience Center: A team of research software engineers (RSEs) experienced in HPC and GPU acceleration worked in close collaboration with the TU Delft team to tackle this challenge.

Codes involved:

Open source Dutch Atmospheric Large Eddy Simulation (DALES): https://github.com/dalesteam/dales
Open source RTE-RRTMGP radiation solver: https://github.com/earth-system-radiation/rte-rrtmgp

Technical/Scientific challenge:

High-fidelity weather and climate simulations are computationally expensive due to their intrinsic multiscale and multi-physics nature; to obtain for instance a realistic turbulent boundary layer with emerging clouds over a large spatial domain, atmospheric large-eddy simulations (LES) need to be run on grids containing billions of cells using many processors in parallel. This makes them a perfect use case for exascale computing. Now, sophisticated atmospheric physics models such as the Dutch Atmospheric Large-Eddy Simulation (DALES) have a long development and validation cycle, including fine-tuning physical model parameters and mitigating model uncertainties. In the case of DALES, the large 'legacy' code base in Fortran90 is not well suited for a complete rewrite in a modern, GPU-friendly language. Additionally, DALES is used by researchers and students alike, for whom the Fortran90 syntax is easily accessible compared to more advanced computing-focused languages.

Solution:

The solution approach employed in DALES relied on using OpenACC, a directive-based programming model, to accelerate the various physics modules and atmospheric dynamics with minimum disruption of the code base. The ESiWACE team also interfaced the code with GPU-accelerated libraries when possible. OpenACC is well suited for this task as it seamlessly integrates in the Fortran90 DALES code. Building upon the initial effort by the TU Delft team, eScience Center RSEs have expanded the work to include all of DALES core modules in two steps: 1) implementing OpenACC directives and replacing the radiation scheme with the GPU-aware RTE-RRTMGP library, 2) iteratively profile and optimize the code obtained after the first step to ensure good performances on the GPU, while maintaining the baseline CPU performances and code readability.

Impact:

The leap in performance (see below) for the Dutch Large-Eddy simulation transforms DALES from primarily a research tool into a viable limited-area weather forecasting and climate prediction tool. This can potentially have huge benefits, as European societies are struggling to mitigate the negative consequences of global warming and transitioning towards sustainable energy sources. The hyper-resolution LES can provide much-needed improvements in predicting summer convective heavy rain or local extreme temperatures in urbanized areas, phenomena that are expected to get more severe over the next decades. At cloud-resolving resolutions, it could elevate the accuracy of yield prediction of photovoltaics and wind turbines, facilitating the balance of power supply and demand in a renewable-based energy system. Finally, the GPU acceleration opens a pathway towards leveraging DALES for local climate projections, providing trustworthy detailed local climate information to societal stakeholders in agriculture, infrastructure and other relevant sectors.

Benefits:

- We have ported the DALES atmospheric physics solver to GPUs using OpenACC. For cases that fit in a single GPU's memory, a speedup of 20x or more has been measured for the Nvidia A100 vs. the 192-core AMD EPYC Genoa CPU. This also translates into substantial energy savings for longer runs.
- We have interfaced DALES with RTE-RRTMGP, a modern GPU-aware community standard radiation solver
- On a node-by-node basis for large test cases, a 12x speedup is obtained (SURF Snellius supercomputer, 4 H100 GPUs vs 192-core AMD EPYC Genoa CPUs)

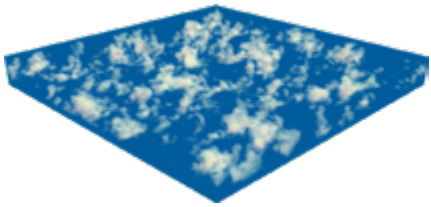


Figure 1: 3d low-level cumulus clouds emerging from a numerical simulation with DALES. By Johanna Grönqvist and Fredrik Jansson, TU Delft.

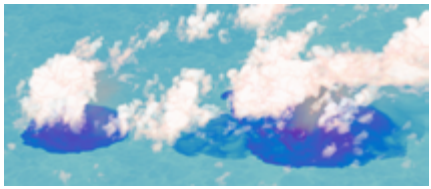


Figure 3: 3D visualization of simulated convective clouds by DALES. Clouds are shown in white and rain is shown in gray. The air temperature near the surface is shown in blue, dark blue areas being cold pools associated with rain. By Johanna Grönqvist and Fredrik Jansson, TU Delft.

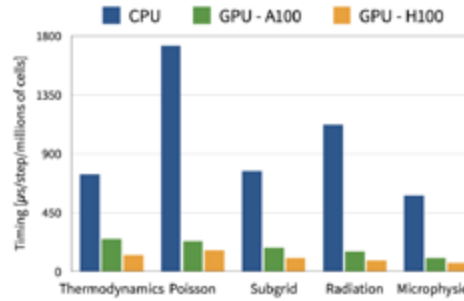


Figure 2: Timing comparison between CPU and GPU execution of DALES main physical processes, using a full Snellius AMD EYPC Genoa node vs. GPU nodes with 4 cards of the indicated type. The time step of the model is dominated by sequential subroutines that compute the effect of these processes.

Source: L. Esclapez et al., „Accelerating the Dutch Atmospheric Large-Eddy Simulation (DALES) Model with OpenACC," 2025 IEEE International Parallel and Distributed Processing Symposium (IPDPS), Milano, Italy, 2025, pp. 678-688, doi: 10.1109/IPDPS64566.2025.00066

Keywords: Weather, high fidelity simulation, large eddy simulation, cloud-resolving

Industry Sector: Environment/climate/weather

Technology: HPC, GPU, OpenACC

Time period: October 2023 - October 2024

Contact:

g.vandenoord@esciencecenter.nl

F.R.Jansson@tudelft.nl

CoE Success Story 2

Speeding up atmospheric chemistry simulation for weather forecast by improving the hardware usage of modern supercomputers

CoE ESiWACE3

Partners involved:

Barcelona Supercomputing Center (BSC): https://bsc.es/

BSC is the national supercomputing centre in Spain. It specialises in high-performance computing (HPC) and manages Marenostrum (MN), one of the most powerful supercomputers in Europe. It focuses on five different research fields: Computer Sciences, Life Sciences, Earth Sciences, Computer Applications in Science and Engineering, and Computational Social Sciences and Humanities.

Codes involved:

Chemistry Across Multiple Phases (CAMP): https://github.com/open-atmos/camp

Multiscale Online Nonhydrostatic Atmosphere Chemistry model (MONARCH): https://gmd.copernicus.org/articles/10/609/2017/

Technical/Scientific challenge:

Earth science models for accurately predicting weather and climate (such as MONARCH) need to employ detailed models (like CAMP) of atmospheric chemistry. As such, chemistry components can account for 80% of the total computation time of Earth system model runs on supercomputers; maximising their computational efficiency is paramount. Today, GPUs account for a significant portion of the compute capacity in current supercomputers. For CAMP, researchers at BSC had already achieved a substantial speed-up over more conventional uses of GPUs. Yet, this solution (like that present in state-of-the-art models in general) also left parts of the machines unused: while the work can be performed on either the GPUs or the CPUs of the computational nodes, the other part is essentially idle. In addition, for CAMP, which was developed on the Marenostrum 4 system, it was unclear how the gains would carry over to the new Marenostrum 5 system and how they would affect the performance of coupled runs as part of a complete atmospheric model. Notably, such models exhibit substantial dynamic variations in the computational load of the various components, e.g., due to day and night differences, making it challenging to fully exploit the heterogeneous hardware consistently over time.

Solution:

After the CAMP code was deployed in Marenostrum 5, the team began development to enable the simultaneous use of both GPU and CPU components. In particular, they implemented a dynamic load balancing algorithm to maximise the utilization of both elements. This is particularly important, as both GPUs and CPUs exhibit significantly different performance characteristics, and to address the aforementioned changing computational load over time. To evaluate the performance of a comprehensive system simulation, the improved CAMP model was integrated with the MONARCH atmospheric model. While the mere porting from MN4 to MN5 resulted in a speed-up of 1.5x for the GPU-only version measured on a single MN5 node (4 GPUs and 80 CPU cores), the combined CPU/GPU version with dynamic load balancing achieved a further speed-up of approximately 1.33x, resulting in a combined speed-up of 2x.

Impact:

First of all, the work directly improves the performance of geoscientific models that rely on atmospheric chemical kinetics of CAMP. Moreover, the methods used to improve CPU and GPU computing power simultaneously carry over to other atmospheric chemistry solutions currently running on one of these hardware (CPU or GPU) while leaving the other part essentially idling. Also, the load balancing method is designed to advance the state-of-the-art on balancing CPU-GPU load, focusing on an easy-to-implement approach for an environment with such highly variable load, such as atmospheric simulations, where many variables affect the workload during the simulation due to events such as the cycle of day and night.

Benefits:

The hardware configuration corresponds to the Marenostrum 5 GPU partition, where each node has 4 GPUs and 80 cores.

- The code was updated from Marenostrum 4 to Marenostrum 5, resulting in a 1.5x speedup on a single node for the GPU-only version
- The CPU-GPU version with the dynamic load balance improves single-node speed-up to 1.3x over the GPU-only version

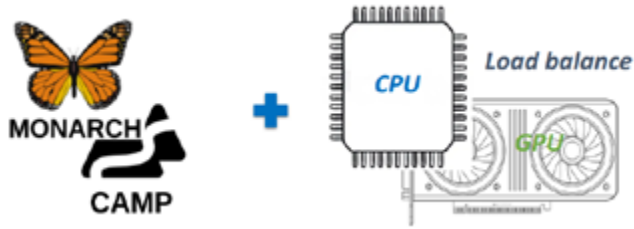


Figure 1: Integrating load balance CPU+GPU in MONARCH-CAMP

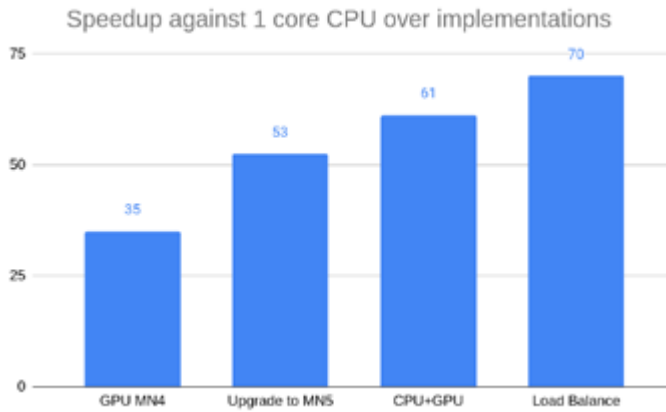


Figure 2: Speedup of the different implementations performed against the CPU-only version. The hardware used is 1 GPU against 1 CPU core. The upgrade to Marenostrum 5 is a hardware upgrade without optimizations. The CPU+GPU case uses both resources simultaneously during the simulations, with the workload distribution between the CPU and GPU being constant. The Load Balance corresponds to the algorithm that balances the workload between CPU and GPU for each time step during the simulation.

Keywords: High-Performance Computing; GPU acceleration; Climate Simulation; Atmospheric chemistry; Algorithm design and analysis; Performance evaluation; Load balance

Industry Sector: Earth sciences, IT/HPC systems

Technology: HPC, CUDA

Time period: March 2024 - September 2024

Contact:

christian.guzman@bsc.es

xavier.yepes@bsc.es

Automated workflow for the Hi-fidelity simulation of propulsion devices: application to after-burners

CoE EXCELLERAT P2

Partners involved:

Cerfacs is a science centre that produces innovative solutions for the simulation of earth physics and engineering. Backed by seven public and private partners, supported by European and national research projects, it is focused on research and training for high performance simulation. Cerfacs developed the simulation workflow.

Safran Aircraft Engines draws on a legacy reaching back over 110 years to design, develop, produce and market, alone or in partnership, engines for civil and military aircraft. It tested and deployed the workflow on one of their actual after-burner studies.

Codes involved:

AVBP: https://services.excellerat.eu/en/services/application-software/avbp/
PyHIP: https://pypi.org/project/pyhip/
Kalpataru: https://gitlab.com/cerfacs
Lemmings: https://pypi.org/project/lemmings/
Tekigo: https://pypi.org/project/tekigo/

Technical/Scientific challenge:

The decarbonation and depollution efforts of Safran Group are pushing the innovation of new burner technologies forward. However, the volume of after-burner flames and the unpredictability of newer fuels such as hydrogen are challenging the meshing practices. Indeed Hydrogen flames can be far thinner than Methane or Kerosene ones. The required high-resolution zones, extremely computer-intensive, must be kept as small as possible. A manual positioning is too difficult in complex configurations. On the other hand, direct remeshing on hydrogen flames is limited to research purposes, because it disregards the affordability of the generated grid. Finally the procedure of running a simulation, decide when to refine, set up the next simulation and start again cannot be done manually in an industrial context. This required automation and portability across machines, in other words: an integrated workflow.

Solution:

The first step was the integration of the "Lemmings" tool developed during EXCELLERAT P1 for packaging run workflows, providing portability, low impact on queues, exhaustive logging, and traceability of job sequences, thus overcoming limitations of mere job schedulers. The second step was the creation of a resolution-independent remeshing criterium using Tekigo (a result from COEC). Finally, the robustness of PyHIP and the performance of Kalpataru, both remeshing tools backed by EXCELLERAT P2, pushed the solution to maturity.

The resulting Automated Static Mesh Refinement workflow (ASMR) developed by Cerfacs has been deployed and successfully tested by Safran engineers. It allows for fast exploration and feedback on new meshes directly based on the physical phenomena observed and computational resources available. The current sequential mesh adaptation (using pyHip) permits to process up to 200 million tetrahedrons. Cerfacs is currently working on parallel mesh adaptation.

Impact:

Safran Group developed the capability to routinely execute high fidelity simulations during the after-burner design process. Without this workflow, the computational cost would have limited these simulations to exceptional runs on a case-by-case basis.

The same workflow is being evaluated for other Safran design fields: ignition phases, cruise runs, turbomachinery. An interesting feature is emerging: The meshes created by the workflow are born of physical meshing criteria, while meshing practices commonly used by industry outside this workflow come from years of experience. Comparing the two kinds of meshes is shedding a new light on present simulation practices.

The increased automation in the workflow also reduced user support requests through higher homogeneity and better traceability of simulation runs. Thus, this workflow, initially required by industrial users, is now also slowly gaining traction with academic users for simulations with lower technology readiness levels, especially for highly unpredictable hydrogen burners.

Benefits:

- Affordable simulations: The simulation can be adapted to available resources because the number of final degrees of freedom is directly input by the user
- Lower manpower: The automation decreases the need for run supervision, while the traceability and logs accelerate runtime. This saves up to three quarters (75%) of the supervision time previously necessary.
- Resource optimisation: The automation ensures a progressive and reasonable usage of resources. This avoids run outliers in performance or set-ups.

Keywords: Automated workflow, reactive simulation, high-fidelity meshing, after-burners, simulation automation

Industry Sector: Aeronautics, Energy, Manufacturing & engineering, Mechanical engineering

Technology: HPC Simulation Workflow

Time period: Ongoing deployment started in 2024

Contact:

Antoine Dauptain (Cerfacs),

dauptain@cerfacs.fr

Luis Carbajal Carrasco (Safran Aircraft),

luis.carbajal-carrasco@safrangroup.com

High-fidelity simulation using Adaptive Mesh Refinement with Spectral Element Method solver

CoE EXCELLERAT P2

Partners involved:

KTH is a leading Scandinavian research centre and the largest technical university in Sweden. It conducts research and education in engineering and technology, collaborating with prominent partners worldwide and participating in a variety of key projects.
https://www.kth.se/en

Codes involved:

Neko: https://neko.cfd/
Nek5000: https://nek5000.mcs.anl.gov/

Technical/Scientific challenge:

High-fidelity modelling of realistic, turbulent flows everyone encounter every day, like flows around various types of rotors, is challenging due to the wide range of the flow features that must be resolved. This makes running simulations computationally expensive and poses a meshing problem, as the flow dynamics may not be a priori known. A common solution is to use turbulence modelling, where small vortices are no longer resolved in the simulation, but their effect is approximated by a mathematical model. This lowers resolution requirements and thus reduces computational costs, but introduces modelling errors, making simulations of flows with complex dynamics less reliable. It is important to be able to perform high accuracy simulations of these complex flows since they could be used to develop new turbulence models.

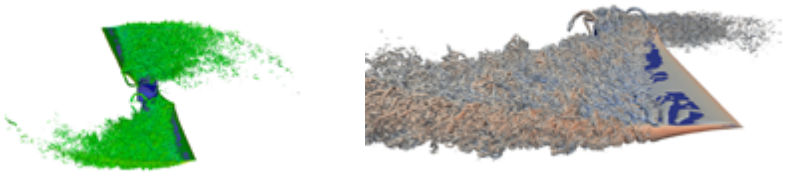


Figure 1: Visualisation of the vortical structures in the flow around the "Iowa" rotor at Reynolds number 15000 for the whole rotor (left) and zoom in to the blade tip (right). Flow complexity and the wide range of the vortex scales are clearly visible.

Solution:

Placing higher resolution only where it is needed is the core principle of Adaptive Mesh Refinement (AMR), which dynamically adjusts mesh resolution during runtime guide computational error control, which is well suited to flows with not a priori known dynamics. KTH is implementing an AMR workflow in the CFD solver Neko based on the Spectral Element Method (SEM). SEM can be considered a special version of widely known Finite Element Method. To implement AMR, they use native to SEM decomposition of the domain into non-overlapping spectral subdomains called elements, and perform recursive splitting/merging of these elements to create new ones. The resulting mesh is complex and requires special treatment of nontrivial combination of element's interfaces, but allows one to complete the simulation at minimal cost increasing at the same time the robustness and reliability of the solution. An important aspect here is its efficient use of the computing resources.

Impact:

As AMR captures flow features during simulation, initial mesh generation is considerably simplified. The significant reduction in computational cost and increased solution reliability will let researchers run bigger and more complex cases of industrial relevance. A good example is high-fidelity simulation of a drone rotor, which would give more insight into its flow dynamics and noise generation.

In addition, a simple method to increase local grid resolution through AMR can be a stepping stone for more advanced algorithmic developments, like an immersed boundary method (IBM), in which object surfaces are not represented directly by the mesh but rather modelled inside the computational domain, providing great flexibility but requiring sufficient resolution near the boundary.

Benefits:

- Control of computational error at optimal computational cost
- Reduced mesh size resulting in reduced computational cost
- Simplified meshing and possibility to model flow cases with unknown dynamics

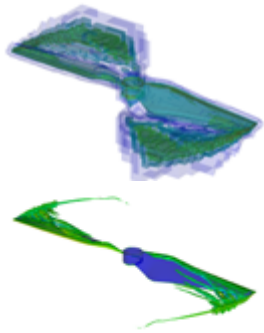


Figure 2: Locally increased resolution resulting from use of AMR framework. Upper plot shows volume occupied by various refinement levels and the lower figure presents the corresponding flow structure.

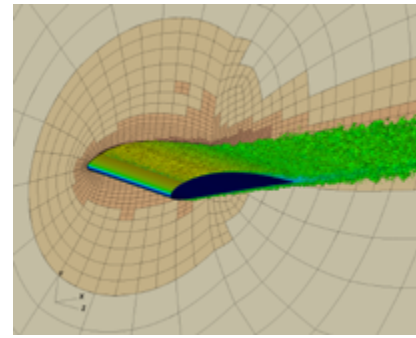


Figure 3: Vortical structures for the flow around a periodic wing. Structure of the AMR mesh is presented by black lines (element interfaces) and a colourmap showing various refinement levels.

Keywords: Hpc, cfd, high fidelity simulations, spectral element method, adaptive mesh refinement, Neko

Industry Sector: Automotive, Aerospace

Technology: HPC

Time period: Ongoing

Contact:

Adam Peplinski (KTH, Royal Institute of Technology), adampep@kth.se

Advanced scalable workflow of heat load assessment inside nuclear fusion reactors

CoE EXCELLERAT P2

Partners involved:

The LeCAD laboratory (University of Ljubljana) manages one of the few supercomputers in Slovenia, which it uses for engineering tasks. This has enabled LeCAD to enter several major international science and infrastructure programs related to supercomputing and the development of new high-performance scientific codes for simulations of complex physical processes, mostly related to nuclear fusion technology.

Codes involved:

L2G, OpenFOAM and Raysect: Engineering Design and Digital Twin of the First Wall of a Tokamak Fusion Reactor

Technical/Scientific challenge:

One of the key challenges in designing the "first wall" of a tokamak fusion reactor (the component directly exposed to the extremely hot plasma inside) is accurately capturing complex physical phenomena, one of them being intense heat flux deposition to the wall. The plasma effects are especially difficult to model due to extreme heat gradients and high energy particle interactions with the wall. These make physical experiments costly and technically demanding. Moreover, the complex geometry of the wall, designed with the goal of reducing the heat flux, adds difficulty to simulations. In preparation of digital twin of the first wall of a tokamak fusion reactor, one of the challenges is to address these physical phenomena (e.g. heat flux deposition, photonic movement, and heat transfer) with sufficient complexity to arrive at a digital solution that is comparable to an actual experiment using an IR camera . This complexity comes at the cost of time and memory heavy computation.

Solution:

The three codes (L2G, OpenFOAM and Raysect) have been connected by the UL team in an initial proof-of-concept workflow to calculate a synthetic camera signal as digital twin Heat fluxes and temperatures were calculated for the full wall of the tokamak fusion reactor, taking into account the magnetic fields of the external coils and the plasma itself, heat flux of particles moving along field lines, and resulting heat transfer, to finally compute the optical effects due to the resulting radiation from the wall.

EXCELLERAT is now transforming the workflow by integrating the HPC framework that enables execution of all three simulation steps in parallel without the need for user intervention. The project is implementing automated data management systems and optimizing code for exascale computing, focusing on the parallel processing for field line tracing and thermal modeling.

Impact:

The integration of field line tracing, thermal and optical simulations into a digital workflow advances the analysis of reactor components and reduces computational overhead through an integrated and parallel execution. This permits to run predictive simulations that would otherwise be too time-consuming. The capability to predict heat flux distribution permits to identify the heat loads on the wall early on, which can extend the lifetime of the wall and reduce maintenance costs. The workflow for synthetic infrared camera imaging is also crucial for real-time monitoring applications of fusion reactors. Thus, this work also brings economic and societal benefits by improving the safety of the nuclear fusion reactors.

Benefits:

- Integration of field line, thermal, and optical simulations enables reactor monitoring
- Reduction in computational overhead through automated workflow
- Capability for accurate temperature detection despite surface reflections
- Identification of hotspots before component damage occurs

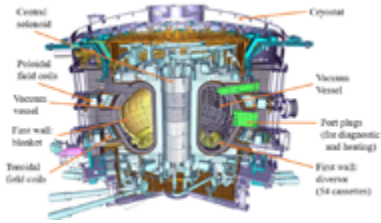


Figure 1: Main components of ITER fusion reactor[1]

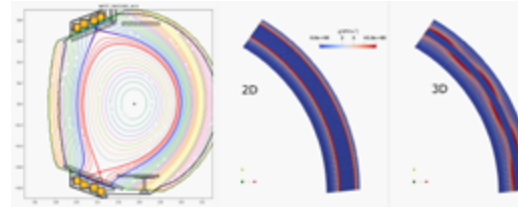


Figure 3: (left) Magnetic flux cross-section of WEST tokamak (right) Distribution of heat flux in ø direction assuming constant magnetic field in ø direction (2D magnetic field) and with varying magnetic field in ø direction (3D magnetic field) where axisymmetry is broken and perodicity occurs, introducing higher level of complexity.

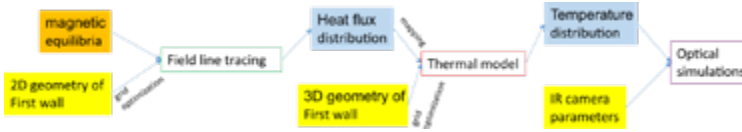


Figure 2: Workflow consisting of three steps: field-line tracing, thermal model, ray-tracing

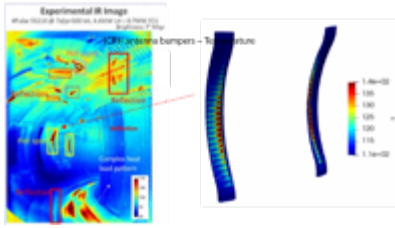


Figure 4: (left) Experimental IR camera image in WEST tokamak. [2] *Temperature distribution on ICRH antenna bumpers by solving first two steps of the workflow: field-line tracing and thermal model. The result matches experimental result to some level of degree.*

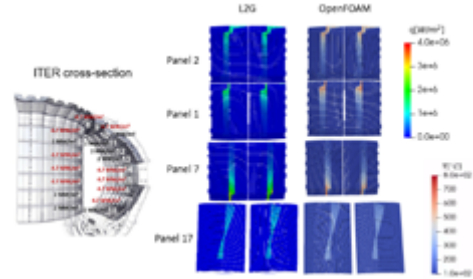


Figure 5: (left) Cross-section of ITER fusion reactor (right) heat flux (L2G code) and temperature (OpenFOAM) distribution on individual panels.

[1] *VG. Dubus, "From Plain Visualisation to Vibration Sensing: Using a Camera to Control the Flexibilities in the ITER Remote Handling Equipment (Doctoral Dissertation)," PhD Thesis, 2014.*

[2] *Aumeunier M.-H. et al 2021 Nucl. Mater. Energy 26 100879*

Keywords: Connecting codes, digital twins
Industry Sector: Energy
Technology: HPC
Time period: Middle of 2025

Contact:
Matic Brank (University of Ljubljana, LECAD),
Matic.Brank@fs.uni-lj.si

High-resolution urban wind comfort computation for the entire city of Stockholm

CoE HiDALGO2

Partners involved:

SLB Analys: Stockholms Luft- och Bulleranalys (SLB-analys) is a unit at the Environment and Health Administration of the City of Stockholm. The unit is responsible for monitoring outdoor air quality in the city. SLB-analys also runs the regional system of air quality monitoring on behalf of the East Sweden Air Quality Management Association (Östra Sveriges Luftvårdsförbund) and assists its municipalities with services including measurements and model calculations.

ENCCS: The EuroCC National Competence Center Sweden (ENCCS) provides high-performance computing training and support for industry, academia and public administration for free.

ENCCS can guide their partners before, during and after they have gotten access to high-performance computers by providing useful knowhow, best practices as well as hands-on support for your code.

SZE: The HiDALGO2 partner Széchenyi István University (SZE) is located in Győr, Hungary, in a centre of an industrial region. In HiDALGO2, SZE is developing the Urban Air Project (UAP) and leads the work-package for the applications. The SZE-team is lead by a mathematician and includes a professional scientific programmer and computer scientists, all working for several decades in the field of HPC.

Codes involved:

RedSim (Reduced Simulations - CFD solution), CFDR (CFD rendering – HPC visualisation), RedSim-Preproc-Urban (preprocessing toolbox for urban air simulations), https://redsim.mathso.sze.hu/.

Technical/Scientific challenge:

Wind in cities has a considerable impact on the comfort and safety of their inhabitants. Especially near tall buildings, wind speeds might be much larger than the average speed, which may cause discomfort, or put pedestrians at risk. Thus, there is a need to evaluate construction plans with respect to their effect on the wind comfort, enabling city planners to optimise construction by evaluating several building scenarios under different weather conditions.

SLB-analys, in need to perform related analyses for Stockholm was put into contact with the HiDALGO2 team through ENCCS, as the challenges are manifold:. The requested resolution (1-2 m horizontally of the full area of 100 km2) requires HPC resources to simulate, combining building and terrain data result in a complex geometry handling challenge, and visualization of terabyte-size simulation data remotely fitting the standard work environment of SLB-analys staff need a scalable solution.

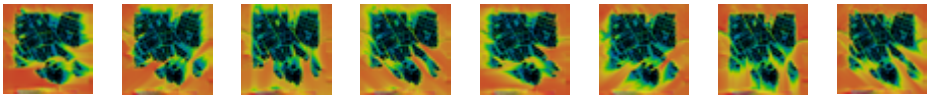


Figure 3: Results of a scenario analysis: visualization of the wind magnitude on a surface of 2 meters height from the ground level for 8 wind directions. © SZE.

Solution:

To address the technical challenges, the Széchenyi István University (SZE) RedSim-team has adapted their HiDALGO2 Urban Air Project (UAP) solution for Stockholm. The basis for the construction of the computational geometry of the CFD model of the city was a data set provided by SLB-analys containing precise building and terrain data (isolines 0-90 m). The preprocessing toolset has been developed significantly to achieve a computational domain for the city with valid surface geometry and reasonably sized polygons for building floors, using Gmsh for the 3D computational mesh. To provide a sufficiently fast simulation, RedSim, the HiDALGO2 highly optimized multi-GPU CFD-solver has been used. For the visualization of the large amounts of data in a user-friendly way, the team used the HiDALGO2 visualization tool called CFDR that renders the raw CFD-related 3D data residing on the HPC platform in real time and provides visualization on a web browser (even on low-end laptops).

Impact:

The presented work provides city planners and policymakers of the Stockholm area with a fast, high-resolution wind data to guide sustainable urban development. To create the computational geometry of an entire city at 2.5 meters resolution takes less than 15 minutes, surface meshing 6-8 hours, 3D meshing 15 minutes. Simulation of one wind scenario takes 15 minutes on 12 nodes of a HPC machine. This innovative GIS and terrain data process together with the scalable HPC-backed CFD solution created a showcase for other large-scale urban areas (with, e.g. 100 km2 area): Optimized HPC and web-based tools reduce the need for physical testing or third-party software, saving time and resources. There is a commercial potential toolchain, comprised by the scalable CFD engine (RedSim) and visualization tool (CFDR), together with the geometry pre-processing tools (RedSim-Preproc-Urban) which can be offered as commercial services or licensing models for environmental consultancies, city authorities, or industry partners.

Benefits:

- The city's computational model was generated based on precise terrain and building GIS data
- High-resolution simulation of an entire urban district with complex terrain becomes possible, due to scalability of RedSim up to 512 CPU nodes and 32 GPUs
- Time-to-solution became hours for city planners for large-scale scenario analysis, which is not available with other solvers

Figure 1: Region of interest of a test computation. Building data provided by SLB-analys (Stockholm City), layout from the Google Earth. © SLB-analys, Google Earth, SZE (for the overall image).

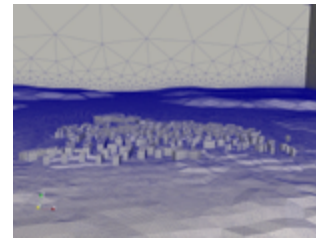


Figure 2: Mesh geometry based on the Stockholm City data of a test computation generated by SZE. © SZE.

Keywords: HPC, CFD, Meshing, CUDA-Acceleration, SIMD-Acceleration, Urban planning, Urban environment, Urban modelling and simulation

Industry Sector: Earth science, Environment/climate/weather, Urban planning

Technology: Python, Shapely, GeoPandas, Gmsh, MPI, manual SIMD, AVX2, AVX512 acceleration on CPU, manual CUDA acceleration on NVIDIA GPUs, WebGL, WASM

Time period: Completed by May 2025

Contact:
Zoltán Horváth (SZE),
horvathz@sze.hu

Predicting potential SARS-CoV-2 mutations of concern via full quantum mechanical modelling

CoE MaX

Partners involved:

CEA

https://www.cea.fr/

Codes involved:

BigDFT: https://bigdft.org

Technical/Scientific challenge:

The receptor ACE2 in human cells can be regarded as the entry door for the SARS-CoV-2 virus, and its spike protein binding domain (RBD) acts like a key by binding to ACE2. Vaccines work by initiating the production of antibodies which recognise and bind to this key, thus disabling it. However, the spike protein evolves fast, making it harder for antibodies to recognise. Now, traditional experimental methods for characterizing protein-receptor interactions are resource-intensive and slow to identify existing mutations of concern. In contrast, being able to predict likely mutations would give us a headway in e.g. developing vaccines. Computational methods become therefore important in view of the understanding of virus binding mechanisms and, more in general, its evolutionary pressure. Recently, the so-called ab initio simulations have become capable of simulating the molecular structure of the whole virus protein and its binding to other molecules like ACE. Yet, these simulations – while very powerful and accurate – are also enormously costly in terms of computational resources.

In order to to predict and characterize binding of SARS-CoV-2 spike variants to the human ACE2 (hACE2) receptor, simulating the electronic structure of systems of approximately 13 000 atoms is necessary – a formidable challenge.

Solution:

The challenge was solved using the BigDFT software, a code that employs a mathematical formalism designed to overcome computational bottlenecks of an ab initio approach on such systems, optimized within the MaX Centre of Excellence for exascale architectures. A team of BigDFT code developers, by leveraging the Quantum Mechanics Complexity Reduction (QM-CR), simulated the electronic structure the Spike RBD vs ACE2 system. This was made possible by combining advanced quantum mechanical algorithms with high-performance computing capabilities, as provided by MaX effort. This approach successfully identified

critical SARS-CoV-2 spike protein mutations, such as A484K, months ahead of epidemiological evidence, demonstrating the power of exascale-ready quantum simulations for real-world biological challenges.

Impact:

Understanding and predicting viral mutations is critical for global health, as demonstrated by the SARS-CoV-2 pandemic. Traditional experimental methods for characterizing protein-receptor interactions are resource-intensive and slow to identify mutations of concern.

The HPC simulations of the MaX team correctly identified the A484K mutation as enhancing spike RBD-ACE2 binding 20 months before its epidemiological detection (variant BA.2.86), proving the predictive power of HPC-driven approaches. In perspective, this capability accelerates the identification of potential viral threats, enabling faster responses for vaccine design and antiviral strategies, thereby improving resilience to future pandemics in humans and animals.

Benefits:

- 20 months advance prediction of the A484K mutation, accelerating the identification of viral threats
- Simulation of ~13,000 atoms, executed in the timeframe of a few hours after the discovery of Omicron mutation, surpassing conventional electronic structure calculations models by over an order of magnitude
- Accurate description of binding mechanisms, paving the way to computational approaches leading to faster vaccine/antibody design
- Usage of HPC resources for a societal problem, enabling large-scale simulations efficiently on exascale-ready platforms

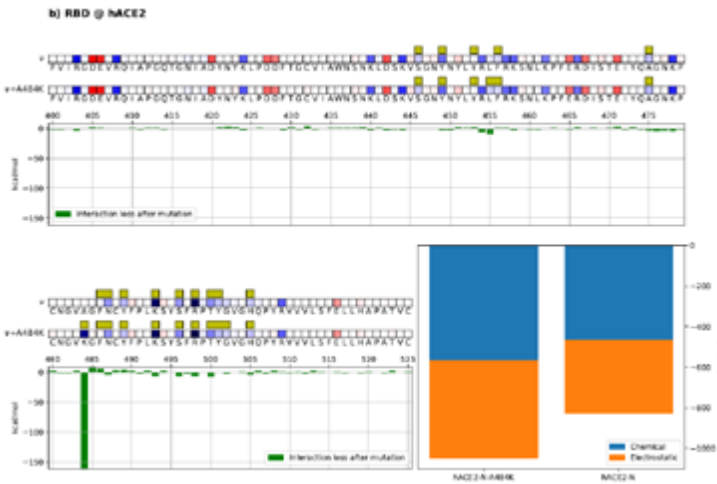


Figure 1: Impact of the A484K mutation on the interaction energy between SARS-CoV-2 RBD and human ACE2. Residue-level interaction loss after mutation is shown for both the RBD (top) and hACE2 (middle); negative values indicate enhanced binding (panel continues in the second row, note the large negative value at position 484). The bar chart at the bottom right decomposes the total interaction energy into chemical and electrostatic contributions, highlighting that the A484K mutation strengthens the binding affinity by improving both components.

Keywords: SARS-CoV-2; ab initio modelling; viral evolution; protein–protein interactions; quantum mechanical modelling; spike protein

Industry Sector: Materials science, Computational Biology

Technology: HPC

Time period: Completed by the end of 2023

Contact:

Luigi Genovese,

luigi.genovese@cea.fr

luigi.genovese@max-centre.eu

Co-design and optimisation of a Quantum ESPRESSO mini-app for future EuroHPC clusters

CoE MaX

Partners involved:

EVIDEN: The team involved in this collaboration is the Eviden's Center for Excellence in Performance Programming (CEPP) made of highly experienced experts in HPC.

SISSA: SISSA is a scientific centre based in Trieste. Its Condensed-Matter Theory group, led by Stefano Baroni, is considered to be the cradle of computer simulation of materials in Italy.

CINECA: CINECA is a non-profit consortium of several Italian institutions and host of supercomputing facilities, equipped with cutting edge technology. It has a strong relationship with its own stakeholders and collaborates with the communities to enable and develop new applications and tools to better address the challenges of the High-end PC systems.

CNR: CNR is the main public research institution in Italy, gathering 110 institutes distributed in the entire country. Three of them participate in MAX: CNR-NANO (Modena) with the "Theory and Simulation" group led by Elisa Molinari, Andrea Ferretti, and Daniele Varsano; CNR-ISM (Rome) with the "Theoretical Spectroscopy" group of Andrea Marini; and CNR-IOM (Trieste) with the "Democritos" group of Paolo Giannozzi.

Codes involved:

Quantum ESPRESSO (Quantum opEn-Source Package for Research in Electronic Structure, Simulation, and Optimisation) is the major open-source (set of) code(s) for quantum materials modelling using the plane-wave pseudopotential method; it has been the development platform for important methodological innovations such as Car-Parrinello molecular dynamics and Density-Functional Perturbation Theory.

https://www.quantum-espresso.org/

Technical/Scientific challenge:

Currently, in Europe new exascale HPC systems are developed, employing novel European Technologies. In particular, the EUPEX prototype features technologies like Rhea 1 processor using heterogeneous memory hierarchy (i.e. HBM and DDR) and based on 80 ARM Neoverse V1 cores including Scalable Vector Extension (SVE).

Co-design means to mutally adapt designs on both sides (hardware and applications) to achieve best efficiencies, in close interaction between software developers and hardware experts. Now, as cutting-edge simulation applications are very complex and difficult to fully understand by anybody but very specialised scientist developers, a much simplified version is needed for hardware-oriented experts to work with it and optimise the software-hardware interface, yet maintains the important performance characteristics of the original software, such that optimisations made here are relevant also for the original application as well as feedback to hardware provider.

Solution:

The Quantum ESPRESSO developers from SISSA extracted a mini-app[1] from the full application, drastically reducing complexity by focusing on the performance-critical parts. This mini-app generates the essential 3D data sets distributed in the Wave function grid of QE together with associated 3D FFT transforms, allowing to explore trade-off between parallelization strategies, band distribution, and the speedup for the execution with many functions modality inside QE with many FFTs in batched mode. The Eviden team focused on the vectorisation and the sensibility to HBM of this mini-app, showing an improvement up to 33% when running on HBM. They proposed two code optimisations further reducing the time to solution by 4% up to 21% depending on the targeted architecture and compiler used. Finally, they showed that SVE vectorisation is properly generated for the mini-app on SVE-enabled processors.

[1] *https://gitlab.com/max-centre/components/mini-apps/-/tree/main/Quantum_Espresso/fftxlib_wfcs*

Impact:

Electronic structure calculations are widely used for predicting and modelling materials' properties in many fields, such as microelectronics, electrochemistry, catalysis, metallurgy, etc. representing a significant part of the European demand for computational resources. As Quantum ESPRESSO is one of the most used codes in this area, its benchmarks are often considered when procuring new HPC machines. Mini-apps making quickly the main computational kernels is thus crucial for an effective codesign of HPC technologies and constitutes an instrumental output of the MaX CoE co-design effort. The results obtained show that the EUPEX platform provides competitive performance for this important application and further confirms the technological choices made in developing the system. In addition, speed-up potential of 21% for the QE application has been demonstrated by improved implementation.

Benefits:

- Quantum ESPRESSO standalone mini-app to foster exchange with partners and thus ease the co-design
- 33% of performance gain thanks to the High Bandwidth Memory compared to DDR5 on Intel Sapphire Rapids HBM 9480 confirming the relevance of the presence of HBM in the EUPEX platform.
- Up to 21% of performance gain thanks to the two optimizations leading to a better vectorization. Source code modification has been shared with the developers to be tested and integrated into the full application.

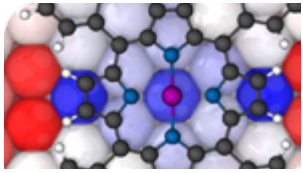


Figure 1: Quantum ESPRESSO is an essential software package for computing atomistic and electronic structures using density functional theory (DFT). It excels in simulating surfaces, junctions, and heterostructures, and its post-processing utilities and open data formats make it ideal for atomic-scale simulations and visualizations.

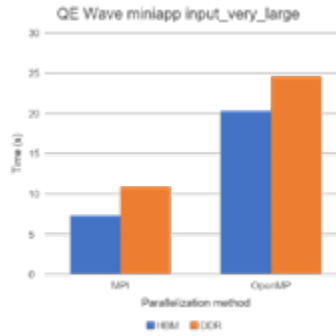


Figure 2: Execution time (in seconds) of Quantum Espresso mini-app on a single SNC4 quadrant (i.e., 14 cores) of an Intel SapphireRapids processor equipped with both HBM and DDR, either by using only the HBM or only the DDR. The parallelization method used is either MPI (14 MPI process, 1 OpenMP thread per process) or OpenMP (1 MPI process, 14 OpenMP threads).
© Eviden.

Keywords: Material Science, co-design, performance optimization

Industry Sector: Materials science, EU hardware technology

Technology: HPC

Time period: Completed in 2024

Contact:

Erwan Raffin, erwan.raffin@eviden.com

Pietro Delugas, pdelugas@sissa.it

Automating the deployment of large software stacks for analysing huge radio astronomy data streams

CoE MultiXscale

Partners involved:

The Square Kilometer Array (SKA) Observatory project is an international collaboration in the field of radio astronomy, building large distributed radio telescope.

Codes involved:

A key technology developed in MultiXscale is the EESSI software environment. EESSI aims to provide a uniform, optimized software stack, available on any system in the world. While created for software deployment in the CoE, it solves a generic technical challenge that many international collaborations face.

Technical/Scientific challenge:

Radio astronomy complements the classical optical astronomy with observations of radio waves from the universe – besides optical signals, these are the only electromagnetic waves from outer space which can penetrate the atmosphere (and which can be measured independently of day time and cloud cover). The radio telescopes targeted by the SKA project will combine the signals from thousands of small antennas, enabling radio astronomers to survey the sky both quickly and with high sensitivity. As the amount of data generated is huge (700 PB per year, about 10 times the estimated size of the entire video library of Netflix), data processing requires likewise amounts of computing resources and is thus distributed over many SKA regional nodes, a global distributed computing infrastructure of the SKA partners called SRCnet.

The software stack used for this analysis contains a considerable number of image and signal processing applications. The challenge for SKA and SRCnet is to maintain up-to-date, tested versions of this software stack optimised to the local hardware throughout their individual regional nodes and their heterogeneous computing facilities. Using standard approaches, this would require an enormous amount of tedious manual (and error prone) work, and likely result in inefficiencies using generic containers. Thus, an automated solution for deploying software across the infrastructure was needed.

Solution:

MultiXscale developed a proof of concept in which the Dutch SKA regional centre added the data analysis software for a full analysis pipeline needed at the SRCNet nodes to the EESSI software stack. The CoE provided the underlying distributed file system infrastructure, build infrastructure, extensive automation of EESSI to enable community contributions, and resolving build issues.

Using this setup, optimised versions of the software were made available automatically through EESSI on 4 SKA regional centres (one of which using a Kubernetes environment), without the need of manual deployment steps. The proof of concept was successful: all SKA regional centres were able to correctly run the analysis pipeline.

Impact:

Containers are the de-facto standard for use in distributed infrastructures like SRCnet. However, since containers are generally focussed solely on software portability and not software performance, applications in those containers are typically not optimized for modern CPU models. Moreover, containers are typically large, thus generating a fair amount of network traffic, startup latency, and requiring large cache size on the client side.

EESSI resolves these issues in two ways. First, multiple builds are made of the same software, each time optimizing for different CPU models. The correct optimization is selected at runtime, which may cause performance increases of up to 30% for certain applications. Second, only the files that are actually opened on an end-user system are transferred from the server that hosts EESSI to the client machine. This drastically reduces network traffic, startup latency and the required cache size on the client side.

All in all, the reduction in required hardware and power through EESSI can reduce the cost and the environmental footprint of scientific analyses.

Benefits:

- Reduced energy consumption for processing data generated by the SKA telescope
- Reduced hardware footprint for processing data generated by the SKA telescope, thus less electronic waste
- Increased productivity for scientists: the EESSI environment (and software) is just 'there' on the system, no need to manually manage / download containers

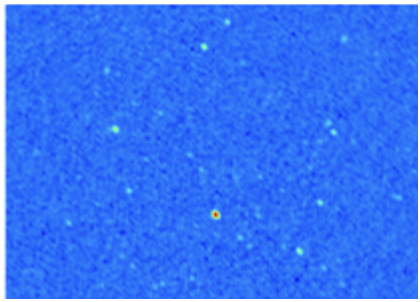


Figure 1: Output of the radio astronomy analysis pipeline, as run on the Canadian SKA regional node, showing clear sources of radio waves. © Tim Kok, SURF

Figure 2: Artistic impression of of the future SKA-Mid telescope, which is composed of many individual telescopes distributed over a large area. Note that radio waves can be received also during day time and through cloudy skies, in contracts to optical waves from outer space. © Square Kilometre Array Observatory, 2021.

Keywords: Astronomy, software distribution, distributed infrastructure, data analysis, signal processing, image processing, radio telescopes

Industry Sector: Space (Astronomy)

Technology: HDPA

Time period: July 2023 – June 2024

Contact:

info@multixscale.eu

CHAPTER 4

FUNDING & INDEX OF ALL CONTRIBUTIONS

Common Acknowledgement

EuroCC 2 and EuroCC4SEE Funding Acknowledgements:
EuroCC 2 and EuroCC4SEE have received funding from the European High-Performance Computing Joint Undertaking (JU) under grant agreement No 101101903 and No 101191697. The JU receives support from the Europan Union's Digital Europe Programme and Germany, Bulgaria, Austria, Croatia, Cyprus, Czech Republic, Denmark, Estonia, Finland, Greece, Hungary, Ireland, Italy, Lithuania, Latvia, Poland, Portugal, Romania, Slovenia, Spain, Sweden, France, Netherlands, Belgium, Luxembourg, Slovakia, Norway, Türkiye, Republic of North Macedonia, Iceland, Montenegro, Serbia, Bosnia and Herzegovina.

CoE Funding Acknowledgements:
ESiWACE3 is funded by the European Union. The project has received funding from the European High Performance Computing Joint Undertaking (JU) and Finland, France, Germany, Netherlands, Spain, Sweden and United Kingdom under grant agreement No 101093054.

EXCELLERAT P2 is funded by the European Union. The project has received funding from the European High Performance Computing Joint Undertaking (JU) and France, Germany, Italy, Slovenia, Spain, and Sweden under grant agreement No 101092621.

HiDALGO2 is funded by the European Union. The project has received funding from the European High Performance Computing Joint Undertaking (JU) and France, Germany, Hungary, Poland and Spain under grant agreement No 101093457.

MaX3 is funded by the European Union. The project has received funding from the European High Performance Computing Joint Undertaking (JU) and Czechia, France, Germany, Italy, Slovenia and Spain under grant agreement No. 101093374.

MultiXscale is funded by the European Union. The project has received funding from the European High Performance Computing Joint Undertaking (JU) and Belgium, Germany, France, Italy, the Netherlands, Norway, Slovenia, Spain under grant agreement No 101093169.


Co-funded by the European Union


EuroHPC Joint Undertaking

This booklet was produced as part of the CASTIEL 2 project. This project has received funding from the European High-Performance Computing Joint Undertaking (JU) under grant agreement No 101102047. The JU receives support from the European Union's Digital Europe Programme and Germany, Italy, Spain, France, Belgium, Austria, Estonia.

Index of All Contributions

Section 2
Agriculture and Food – 44

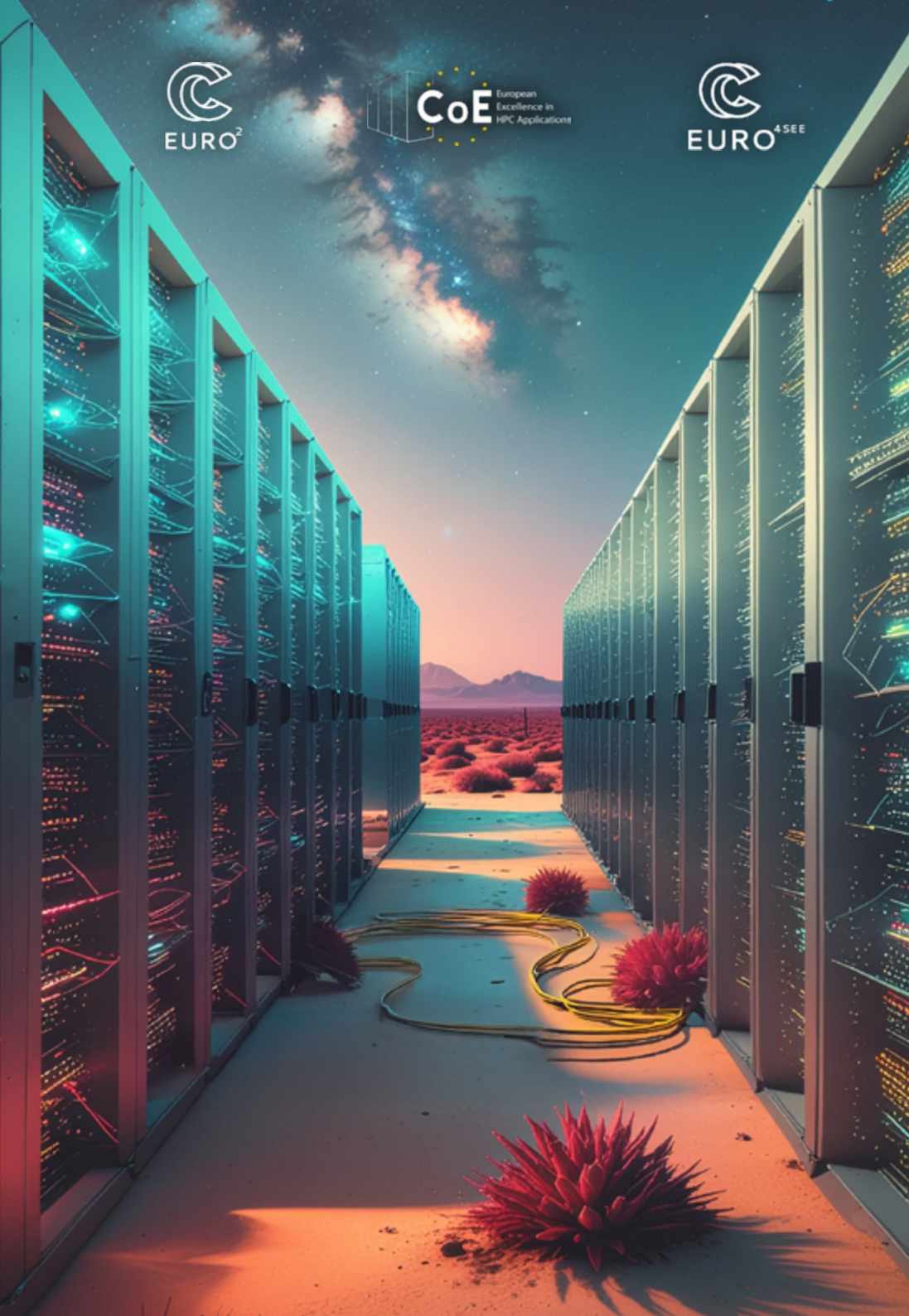
Section 3
Energy – 56

Section 4
Manufacturing and Engineering – 62

For further information, please visit:

www.hpc-portal.eu

HPC IN EUROPE PORTAL


EURO²
CoE
European
Excellence in
HPC Applications
EURO⁴SEE