EURO²

SUCCESS STORIES 2023



EuroCC Success Stories

1st Edition, May 2023

Publisher:

HLRS on behalf of the EuroCC consortium

(Nobelstraße 19, 70569 Stuttgart)

Published: in 2023

Edition: 1st

Copyright © 2023 Members of EuroCC Consortium

Text Copyright © 2023 Members of EuroCC Consortium

Editor: Members of EuroCC Consortium

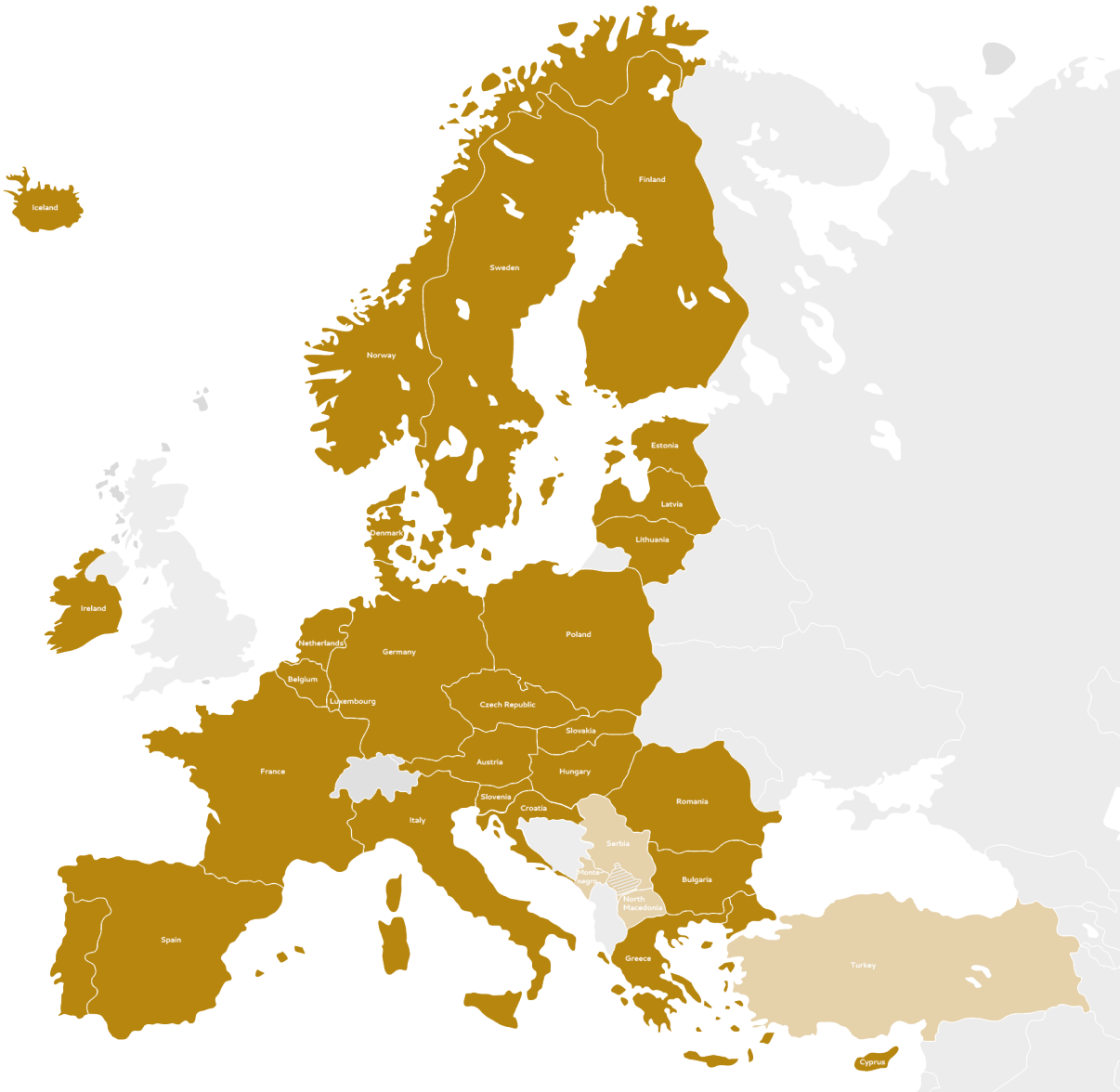
Number of copies: 1000

About the EuroCC2 Project

The mission of EuroCC 2 is to continue the establishment of a network of National Competence Centres (NCC) in the most efficient way, while continuing to address the differences in the maturity of HPC deployment in Europe, for which improvement has already been noted.

Therefore, in addition to high-level management to monitor progress in the NCCs' development, the main task of the overall activity is to support national centres in setting up their individual operational frameworks, while accessing and making the most of the experience and expertise currently available at national and European level.

Learn more about EuroCC2 here!



Index

Section 1
IT and Software – 6

Section 2
Environment, Energy and Agriculture – 28

Section 3
Natural Sciences and Aeronautics – 60

Section 4
Manufacturing and Engineering – 78

Section 5
Finance – 100

Section 6
Pharmacy and Medicine – 110

Section 7
Traffic and Logistics – 130

Section 8
Public and Communication – 144

Section 1

IT and Software

Using VEGA GPU for Natural Language Understanding – 8

Icarus Simulations – 10

Thermo-Calc Software AB to access VEGA – 12

Analysis and Visualization of Social Media Data using HPC – 14

Automated Voucher Generator - 16

Instagram Data Email Value Generator - 18

DigitalSmart and DunavNET - 20

High-Perfomance Data Management on HPC Infrastrucutre - 22

Voxo AB SHAPE Access- 24

Large-Scale Real-Time Image Content Moderation- 26



Photo: Mohammad Rahmani // Unsplash

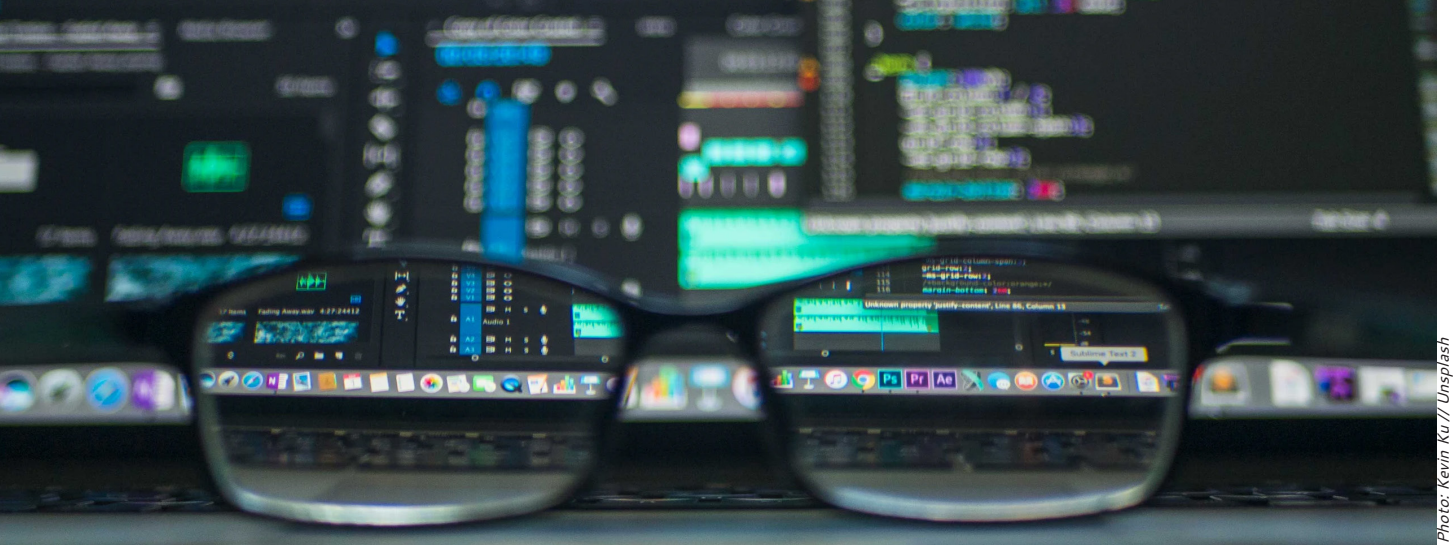


Photo: Kevin Ku // Unsplash

Using VEGA GPU for Natural Language Understanding

ENCCS (EuroCC National Competence Centre Sweden).
ENCCS provides high-performance computing training and support for industry, academia and public administration for free.
www.enccs.se

Organisations Involved

AI Sweden is the Swedish national center for applied artificial intelligence. Their mission is to accelerate the use of AI for the benefit of society and competitiveness for everyone living in Sweden. https://www.ai.se

Technical Challenge

Natural Language Understanding (NLU) is an area in machine learning that transforms human natural language into easily understood machine language and has been recently largely being developed for popular languages around the world. For such models to work, the algorithm must train on a large number of texts and for that it needs large computing power as well as GPUs.



Photo: Compare Fibre // Unsplash



Solution

The NLU team at AI Sweden has gained access to the EuroHPC JU system VEGA in order to experiment with a Swedish GPT (Generative Pre-trained Transformer) model.
The access will be used to test the domain-specific capabilities of a Swedish GPT model using fine-tuning and p-tuning. More concretely, in this research project, AI Sweden wants to fine-tune the model on conversational data to create a chatbot. They will also explore the more efficient technique of p-tuning to improve the conversational capacity of the model.

Business Impact

Comparing these techniques will allow AI Sweden to define the most resource-efficient way to adapt the model to specific use cases of the public sector.

Benefits

- Natural language understanding gives the opportunity to analyse easily massive amounts of text, which can be translated into behavioural patterns, sentiment analysis, interests etc.
- Allow the stakeholders to make detailed data driven decisions faster and more efficiently, saving significant time and resources, and improve the overall level of service provided
- NLU can help organisations to automate document analysis and data extraction, making it easier to gain insights from large volumes of unstructured data.

- Natural Language Understanding (NLU)
- Natural Language Processing (NLP)
- Transformers
- Deep Learning (DL)

- Industry Sector: IT/HPC Systems and Software Providers

- Technology: AI

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

Icarus Simulations

to optimize workflows thanks to VEGA access

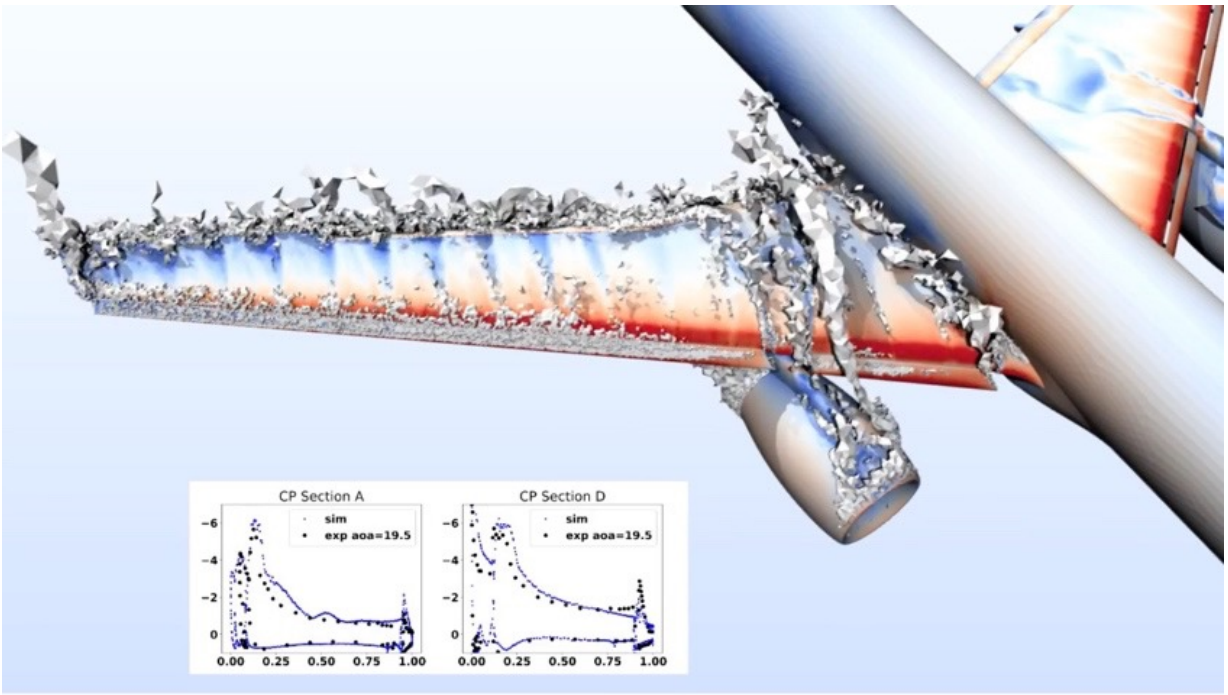
ENCCS (EuroCC National Competence Centre Sweden). ENCCS provides high-performance computing training and support for industry, academia and public administration for free.

Organisations Involved

Icarus Simulations AB provides reliable aerodynamic simulations with Adaptive Euler, directly simulating natural laws.
http://icarusmath.com

The Scientific Challenge

Icarus Simulations AB provides reliable aerodynamics simulations using just a web browser, but aerodynamics need computing power to efficiently simulate CFD (Computational Fluid Dynamics).



Simulation example of aircraft aerodynamics.

The Solution

With the support of ENCCS, the Swedish SME Icarus Digital Math (http://icarusmath.com, Icarus Simulation AB), a spin-off from KTH, received access to the EuroHPC JU system Vega at **Izum**, Slovenia. During this project, Icarus aims to focus on the development of their software for industrial and commercial applications in the aviation, marine and automotive industry.

The Business Impact

Access to the EuroHPC system Vega gives Icarus Simulations AB a unique opportunity to develop their software by utilizing multiple cores and save time in their workflows.

Icarus is involved in the ELISE project for electric aviation, described as a success in Swedish mainstream media (https://www.di.se/hallbart-naringsliv/sverige-ska-bli-ledande-inom-el-flyg-eneroth-tillsatter-utredning/), where their breakthrough simulation technology is used for design of electric aircraft. They are also collaborating with the Formula 1 industry related to this project.

Benefits

- ✔ Software development workflow optimisation



Photo: Nick Morales // unsplash

›CFD
›Aerodynamics
›Dynamics
›Airplane Design
›VEGA

›Industry Sector: Software Providers

›Technology: HPC

Contact:
Apostolos Vasileiadis
Email: apostolos.vasileiadis@it.uu.se

Thermo-Calc Software AB to access VEGA

TENCCS (EuroCC National Competence Centre Sweden).
ENCCS provides high-performance computing training and support for industry, academia and public administration for free.
https://enccs.se

Industrial Organisations Involved

Thermo-Calc Software AB is a world-leading developer and supplier of software and databases for materials design and process optimization. The company is committed to improving and further expanding its databases of thermodynamic and kinetic models as the R&D of materials proceeds by exploring increasingly more complex alloy chemistries and broader ranges of processing conditions than ever before.
https://thermocalc.com

Technical Challenge

There is a need to predict and understand the properties of materials at all stages in the materials' life cycle and at the same time to reduce dependence on expensive, time-consuming experiments.



Photo: DeepMind // Unsplash

The Solution

Thermo-Calc Software AB (TCSAB) has been awarded development access to EuroHPC JU petascale system Vega at Izum, Slovenia with the help of ENCCS.

The proposed project leverages the power of HPC and first-principles-based calculations to accelerate the development of CALPHAD thermodynamic and kinetic databases that bring the feasibility of alloys-by-design to reality. TCSAB researchers will utilize Density Functional Theory (DFT) packages such as VASP and QuantumEspresso (QE) together with TC-Python-based workflows that maximize the potential of a High-Performance Computing (HPC) environment. The allocation awarded is 1.9 million core hours which will help the researchers at TCSAB to shorten the database development cycle significantly.

Business Impact

By shortening the database development cycle, TCSAB can focus on developing other functions in CALPHAD as well as client acquisition and support.

Benefits

- ✔ Shorten database cycle significantly

- Material
- Molecular Simulations
- Quantum Chemistry
- Software Development
- Data

- Industry Sector: Software Providers

- Technology: HPC

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



Photo: Headway // Unsplash

Analysis and Visualization of Social Media Data using HPC

The Slovak National competence centre for High Performance Computing (NCC for HPC) hosted by the Slovak Academy of Sciences acts as the single point of contact and coordination in the field of HPC on a national level. It is the first initiative in Slovakia to support HPC usage beyond academia and is thus very important for the development of the HPC ecosystem.
NCC Slovakia provides non-financial support via consultations, R&D project collaboration, training and education as well as access to HPC resources.

EURO SLOVAKIA

Organisations Involved

MEMO 98 is an internationally recognized, independent, and non-profit specialist media institution with 22 year-long of experience conducting media monitoring and research and assisting civil society groups. Using approved methodologies and tools, media analysis and monitoring is provided with tangible results, in particular during election periods. MEMO 98 participated in about 150 media and election-related projects and trainings in more than 55 countries.

MEMO 98

Technical Challenge

The goal of this proof of concept (PoC) was to set up a high-performance workflow for processing and visualization of large graph-like data sets acquired from several social media platforms. The particular challenge was to utilize HPC resources in an interactive way and, in addition, with graphical user interface. These conditions were required both by the client and the algorithm use for solving the problem itself. The proposed workflow was designed to be reusable in client's future projects and scalable, as the amount of data is expected to grow.



Photo: Campaign Creators // Unsplash

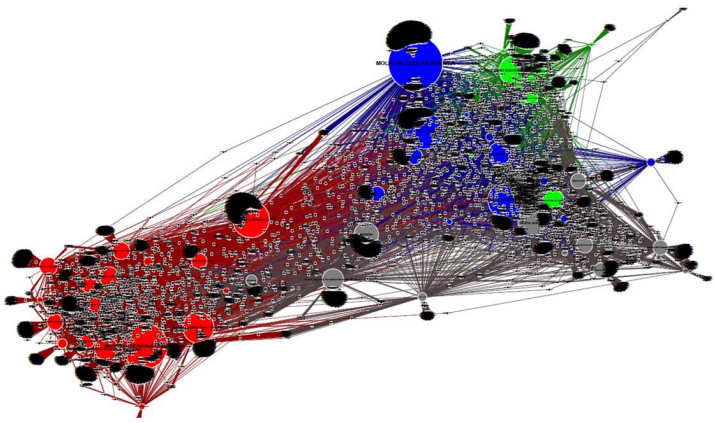


Figure 1: Sample graph of interactions from Moldovia elections in 2021 – data from Facebook.

Solution

MEMO 98 had certain experience with data processing and visualisation, but no experience with HPC environment.
The solution proposed by the NCC relied on three components: first, it is the "Open OnDemand" middleware that provides graphical UI in user's browser to access compute nodes, and integration with job scheduler (Slurm) on the HPC system. Second, it is the 3D rendering of the output from graph processing program via VirtualGL layer, which is done locally on the GPUs of the compute nodes. Last component is the graph processing program itself, Gephi. It is a visualization and exploration package for general graphs and networks, and it's worth noting that for applications without a need of real-time user interactions, more efficient "batch" parallel processing is possible.

Business Impact

Informational environment has evolved with the advent of social networks (SNs). SNs allow for rapid and low-cost spreading of disinformation with huge impact on the society. The MEMO 98 organization analysed the parliamentary election campaign in Moldova on 11th of July 2021, on five SNs: Facebook, Instagram, Odnoklassniki, Telegram and YouTube. Data from SNs (posts and interactions) were collected using CrowdTangle and converted to interaction graphs. Such graph consists of nodes (accounts, and URLs of posts) and edges (URLs shared between accounts), and captures information, e.g. which accounts share URLs of interest the most frequently. The Gephi software used in the client's company, ran on a standard desktop, could not handle the size of such data anymore. The continuous layout graph processing algorithm (ForceAtlas2 – Jacomy et al. PLoS ONE 2014, 29(6): e98679, doi: 10.1371/journal.pone.0098679) takes advantage of user's intervention during the progressing. Facilitating the HPC resources in real-time interactive setup is a great opportunity not only for SN analysis, but also in other areas, such as biology, genomics, etc.

Benefits

- ✓ Overcoming computational time, memory and storage bottlenecks
- ✓ Interactive computation on HPC infrastructure
- ✓ Remote 3D rendering
- ✓ Reusable and scalable workflow

- Social Network
- Interaction Graph
- Graph Computation
- Interactive Computation on HPC
- Remote Rendering

- Industry Sector: Media

- Technology: HPC, Big Data, Machine Learning

Contact:
Michal Pitonak
Email: michal.pitonak@nscc.sk

Automated Voucher Generator

The leading partner of the NCC is the Faculty of Computer Science and Engineering (FCSE).
In addition to FCSE, partners for the establishment of the National Center for Competence for HPC (High Performance Computing) in North Macedonia are: Infinite Solutions and Newman's Business Accelerator.

Organisations Involved

InterTec (with offices in Germany, North Macedonia, and Kosovo) is a visionary and a reliable technology partner collaborating with clients across multiple business domains. By unifying the engineering talent, technical expertise, and essential project management, InterTec mainly deliver IT services and consulting for companies that come from the shopping rewards industry. They build a software platform for generation, life-cycle management and tracking of shopping vouchers and rewards that provides people access to the best savings, cashback, deals, product inspirations and reviews.

The Scientific Challenge

The client deals with the problem of real-time voucher title and content creation, out of a large pool of crawled content data. The existing solution incorporates semi-automatic voucher generation using recent advancements in NLP and availability of pretrained models for language generation.

Since the voucher generation latency was not suitable for on demand generation, the main goal was to speed up the creation process and improve the voucher content quality while freeing up more resources.



Photo: John Schnobrich // Unsplash

The Solution

Based on the problem analysis and the needs of the client, the solution included fine tuning of separate models for automated text generation and paraphrasing for the title and the content of the voucher. Also, the implementation of the execution environment was decoupled from the business process and horizontally scaled. Furthermore, the model's time execution complexity was reduced by using onnx runtime instead of a native execution runtime and by decreasing the float number representation precision.

The Business Impact

NCC North Macedonia successfully managed to improve the quality of the generated content and to decrease the voucher generation latency for up to 30 times. The proposed solution fully automated the voucher creation process and enabled real-time, on-demand vouchers creation on high loads. That potentially will lead to improved voucher quality and improved customer satisfaction.

Benefits

- ✓ The model's time execution complexity was reduced for up to 30 times
- ✓ The model's predictive performance for vouchers content generation was improved for 5%
- ✓ The proposed models and the suggested model serving architecture for voucher generation is appropriate for real-time on-demand generation on high loads



Photo: Brooke Lark // Unsplash

- NLP
- Real-time
- Voucher
- Automation
- Paraphrasing
- Text

- Industry Sector: IT/HPC Systems, Services & Software Providers

- Technology: AI

Contact:
Email: gjorgji.madjarov@finki.ukim.mk

Instagram Data Email Value Generator

The leading partner of the NCC is the Faculty of Computer Science and Engineering (FCSE).
In addition to FCSE, partners for the establishment of the National Center for Competence for HPC (High Performance Computing) in North Macedonia are: Infinite Solutions and Newman's Business Accelerator.

Industrial Organisations Involved

Influencers Club is a company which builds a platform that enables people to search and connect with millions of creators (artists, chefs, coaches, influencers, podcasters, or any other type of creator from Instagram, YouTube and TikTok) at scale. The platform is a solution to one of the biggest problems in the creator economy: "How to get creators to notice your messages?". With a team of 35+ people and 3M+ emails sent per month, the platform gets on average a 20% positive reply rate. Furthermore, the platform contains the largest database of creators and allows searching by 100+ parameters.

Technical Challenge

The challenge of the Influencers Club company is to reach as many as possible creators or customers that match given search criteria. The database contains large amount of data for various creators and customers. The creators and the consumers have more than 100 different attributes available for searching and filtering.

The existing solution of the company includes running full text search queries over the PostgreSQL database based on various criteria for a given input requirements. Running all the queries and filtering in some scenarios can take few hours. The challenge was to increase the efficiency of the data team of the company, with respect to the time needed to search through the database and obtain valuable results, find the creators and the customers that match the search criteria.



Photo: Soundtrap // Unsplash

The Solution

The focus of the NCC North Macedonia team was to investigate several aspects of the given problem: usage of AI techniques for data analytics, improving Postgres performance, improving scraping speed, increasing the scale of the scraping power, implementing real-time collection of data, utilizing massive amounts of processing power to process and analyse real-time data at scale. Based on the company requirements we provided a solution using the ELK (Elasticsearch, Logstash, and Kibana) stack. The data was indexed and imported in Elasticsearch. Elasticsearch is a search and analytics engine. All the pre-processing of the data is done by using Logstash, which is a server-side data processing pipeline that ingests data from multiple sources simultaneously, transforms it, and then sends it to a „stash" like Elasticsearch. Part of the solution was implementation of several visualizations and dashboards in Kibana. Kibana lets users visualize data with charts and graphs in Elasticsearch. The implemented Elastic Stack processes data upon ingest, it uses the inference ingest processor to apply given machine learning model to the incoming data at ingest time without ever leaving Elasticsearch. Additionally, several models were included in the implemented ELK stack to support various natural language processing (NLP) tasks: language identification to determine the language of text, sentiment analysis to identify positive vs. negative sentiment, Named entity recognition (NER) and text classification.

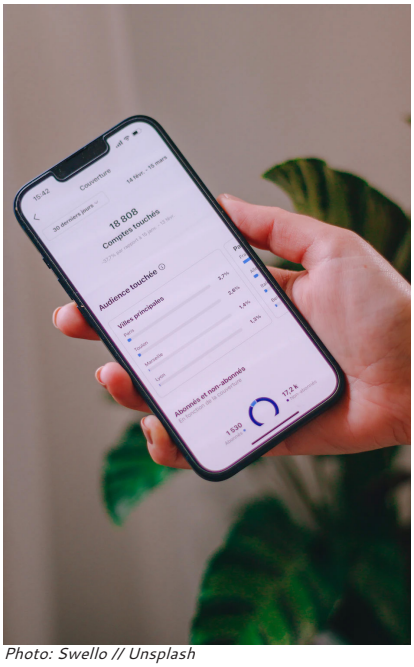


Photo: Swello // Unsplash

Business Impact

NCC North Macedonia successfully managed to increase the efficiency of the data team of the company, with respect to the time needed to search through the database and obtain valuable results, find the creators and the customers that match the search criteria. The proposed solution fully automated the entire business process, from scraping/collecting the data, pre-processing the data, giving additional value to the data by applying machine learning algorithms, storing the data, searching/filtering the data, and visualizing the obtained results.

Benefits

- ✔ The implemented solution can easily be scalable to handle petabytes of data both in a structured and unstructured format
- ✔ High productivity with parallel processing by allocating primary and replica shards across all available Elasticsearch nodes
- ✔ Near real time search, analytics, and visualization of large-scale data (after one second the added document is searchable in this engine)
- ✔ Improved search performance
- ✔ Machine learning support – pre-trained data frame analytics model or a model deployed for natural language processing tasks to infer against the data that is being ingested in the pipeline

- Social Media
- Instagram
- NLP
- Elasticsearch
- Near Real Time Search

- Industry Sector: IT/HPC Systems, Services & Software Providers
- Technology: AI

Contact:
Email: ivica.dimitrovski@finki.ukim.mk



Photo: Homepage NG1oT

DigitalSmart and DunavNET

to provide AI edge solution for poultry industry, done in context of FF4EuroHPC experiments

Among its other activities as NCC, EuroCC Montenegro organized in 2021 promotional events, including a presentation of the FF4EuroHPC Open Call #1 for Experiments. This sparked interest and a local start-up DigitalSmart partnered with IoT provider DunavNET, UDG and local poultry producers to apply for the call. The project was selected for funding and is currently being implemented.

Organisations Involved

DigitalSmart is the coordinator of an experiment called AIMHiGH implemented in the context of the European project FF4EuroHPC. The title of the project is AI/ML Enabled by HPC for Edge Camera Devices for the Next Generation Hen Farms and is funded as an application experiment within Horizon 2020 FF4EuroHPC project.

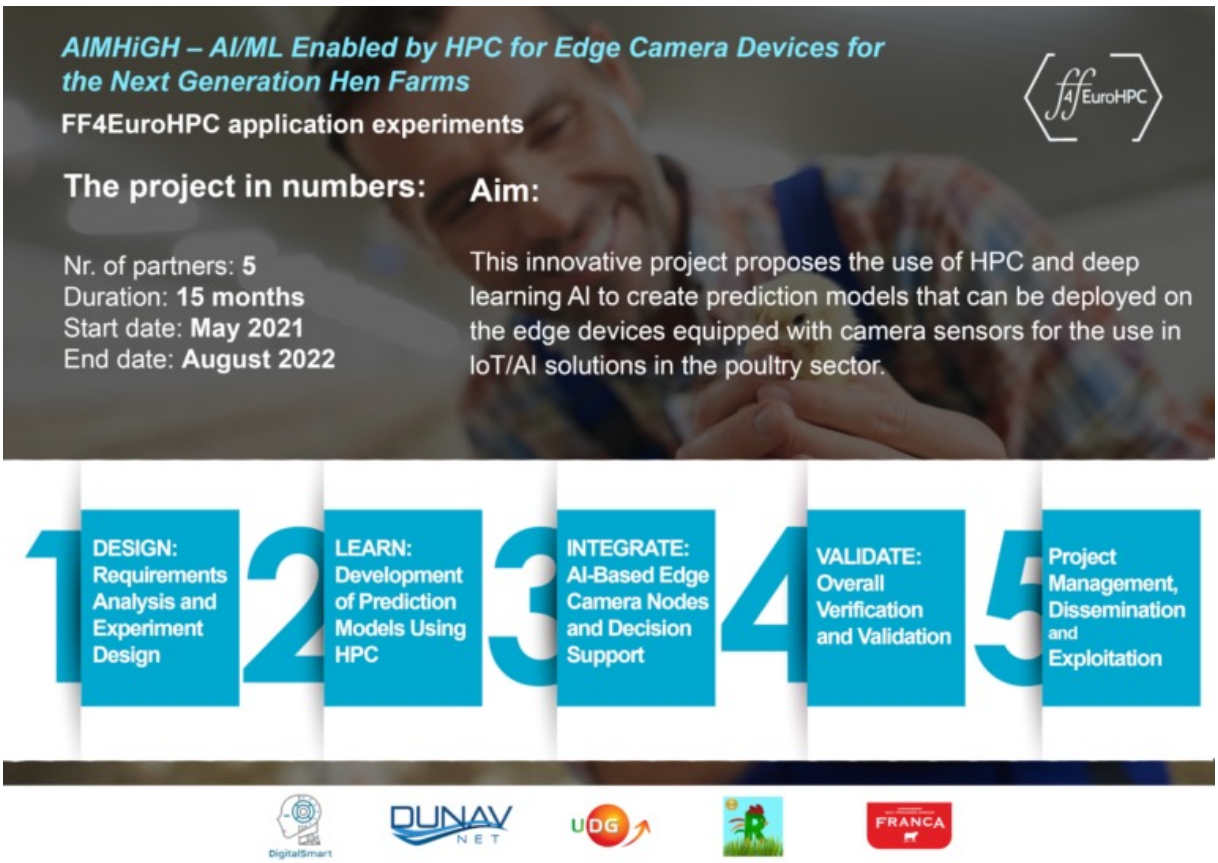
DunavNET provides an expertise in AI/ML, IOT and software development, while the University of Donja Gorica will be providing HPC and domain expertise through NCC Montenegro and FoodHub Centre of Excellence. Montenegrin companies Mesopromet Franca and Radinović Company will be taking part in the evaluation and piloting process. The project is fully aligned with the priorities of S3 Smart Specialization strategy for Montenegro.

Technical Challenge

The AIMHiGH project proposes the use of HPC and Deep Learning AI to create prediction models that can be deployed on devices equipped with camera sensors for the use in IoT/AI solutions in the poultry sector. More precisely, the devices will be equipped with camera sensor nodes in order to create new advanced sensors that can count chickens, help assessing the weight of animals, detect the dead ones, thus improving early disease detection.



Photo: Zoe Schaeffer // Unsplash



Solution

The idea is to utilize HPC to support development of prediction models that can be deployed on platforms such as NVIDIA Jetson Nano equipped with camera and integrated into digital farming platforms such as PoultryNET from the company DunavNET.

Business Impact

Through this project, DigitalSmart obtained access to HPC resources and the solution will be integrated in the already commercialized digital farming platform created and marketed by DunavNET. The solution can also be interesting for other IOT vendors, as well as for the end users.

Benefits

- ✓ More efficient farm management.
- ✓ Early detection of diseases or uneven growth

- DigitalSmart Montenegro
- Innovation
- HPDA
- AI / Edge AI
- IOT

- Industry Sector: SME, Start-up

- Technology: HPC/HPDA, AI

Contact:
Stevan Čakić, MSc,
Email: stevan.cakic@digitalsmart.me

High-Perfomance Data Management on HPC Infrastrucutre

EuroCC Italy is the Italian competence center part of the EuroHPC Competence Centers Network of the EuroCC Project. The long-term objectives of EuroCC Italy are to structure and strengthen the collaboration between research groups and institutes, HPC data centers and companies, to support SMEs in innovation projects and to create awareness about the productive advantages possible through the use of the different solutions in the HPC / HPDA / AI field. EuroCC Italy is partnered by CINECA; ABD, Bi-REX, Leonardo Company and Dompè.

Organisations Involved

Cherrydata is a start-up that offers technology and solutions for high-performance data intelligence, with exceptionally fast response time and extremely low hardware (HW) requirements. This performance edge is granted by AyraDB, a next-generation database with outstanding performance and scalability. This allows Cherrydata's customers to work on their business KPIs with no technical or economic scalability limitations. It also reduces energy consumption, with a positive environmental impact.

The Scientific Challenge

To enable a variety of big data applications, the market needs highly efficient data management technologies: satellite data are notoriously big, raising challenges not only from the perspective of analytic applications, but also to be efficiently stored and retrieved.

AyraDB was designed to address these challenges. The goal of the present research is to perform a large-scale test of AyraDB on Davinci-1 supercomputer, belonging to Leonardo Company. These tests can act as a Proof of Concept (PoC) of the technical features of AyraDB and of its suitability for challenging applications.

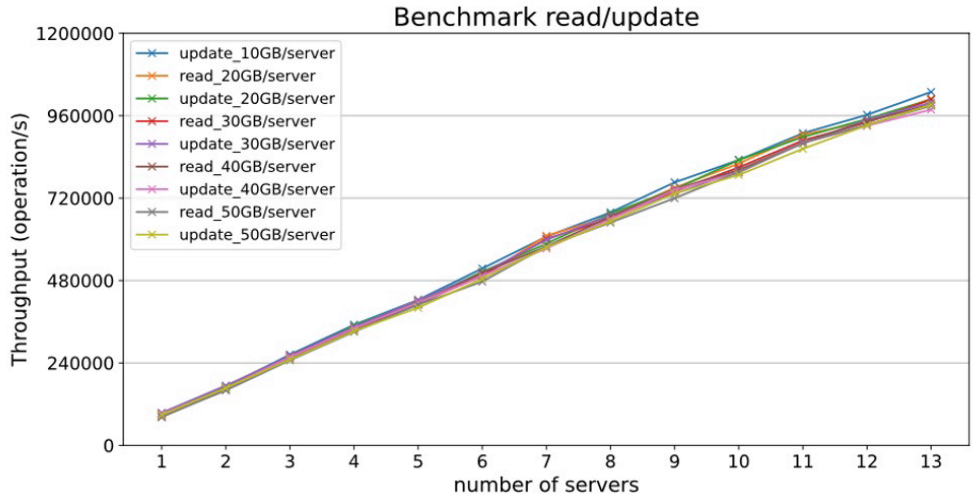


Figure 1. Linear scalability: the sensitivity of performance to table size is not significant.

The Solution

Tests were performed on Davinci-1 with a total of 500 runs on a number of 8-core servers growing from 1 to 20, a data size growing from 10 GB to 500 GB. Results show that AyraDB reaches 106 requests/s with 13 servers, with consistent read/write performance, beating the state of the art by a 5x factor (see chart).
AyraDB has not only best-in-class cost-to-performance ratio. Tests demonstrated that it has linearly scalable performance: if HW capacity doubles, AyraDB's performance doubles. Competing solutions require more than double capacity to double performance. Above a certain threshold, their performance stops increasing even if the infrastructure provides additional capacity, the database cannot exploit it. In fact, AyraDB can use additional capacity with no upper bounds to scalability.

The Business Impact

The PoC has provided third-party performance benchmarks. Obtaining third-party performance benchmarks is a time-consuming and costly activity and the PoC has provided these benchmarks, acting as proof of the performance of AyraDB beyond Cherrydata's own claims.
This is clearly important for Cherrydata and can boost adoption by helping overcome the skepticism that is inevitably associated with a new technology. In the context of an IT deep tech is even more critical. The quasi monopoly of global cloud providers and their aggressive marketing have created the belief that they provide the best possible technology.
With this PoC, it was proved that it is possible to substantially improve the quality of cloud services and that greater performance can be obtained with a technology architecture that makes a better use of RAM and, thus, reduces costs by a factor of 20.



Photo: Campaign Creators // Unsplash

Benefits

- Performance improvement of data management
- Cost reduction of cloud services
- Enabling technical and economic feasibility of several big data applications
- Energy savings

- Keywords Database
- Scalability
- HPC
- Big Data
- Geospatial Data

- Industry Sector: IT/HPC Systems, Services & Software Providers

- Technology: Database, HPC

Contact:
Chiara Francalanci
Email: chiara.francalanci@cherry-data.com

Paolo Giacomazzi
Email: paolo.giacomazzi@cherry-data.com

Paolo Ravanelli
Email: paolo.ravanelli@cherry-data.com

Voxo AB SHAPE Access
to Improve Swedish Text-to-Speech Algorithms

ENCCS (EuroCC National Competence Centre Sweden). ENCCS provides high-performance computing training and support for industry, academia and public administration for free.
www.enccs.se

Industrial Organisations Involved

Voxo AB (https://www.voxo.ai/) is a Stockholm-based startup that specializes in extracting, analysing, and visualising voice data. Their services are used in multiple industries to provide insights and enable data-driven business development.

Technical Challenge

Tools such as Apple's Siri, Amazon's Alexa, and Google Home have brought text-to-speech capabilities to the masses. These conversational assistants respond to natural-language requests and reply in kind. They use machine-learning models trained on large amounts of recorded speech samples matched with the corresponding text. When the assistant wants to say something, the model is able to build new utterances that sound natural.

These big tech companies also provide APIs to access such capabilities, and those support many languages. To use them, the user has to send the text to their server and receive the generated speech back. This is not relevant when the text pertains to someone's personal data. In the European Union (EU), GDPR requires that such data be handled correctly, and in particular not transmitted outside the EU. Using a third-party API of a trans-national company cannot provide the required transparency.



Johan Wadenholt, CEO Voxo AB, at the ENCCS Industry days 2021

The Solution

VOXO is keen to build on their existing voice expertise to enter market sectors that need the capability to synthesize voices speaking Swedish. The speech generation must not compromise the integrity of the data, which might be personal to a user. Thus, existing programmatic APIs are unsuitable, and Voxo is building its own solution using HPC.
As first-time HPC users, Voxo applied to the pan-European program for introductory industrial HPC access, called SHAPE (https://prace-ri.eu/hpc-access/shape-access/). They were delighted to receive help from ENCCS to write their proposal to build a Swedish-language voice-to-text capability. Ultimately they were awarded 25,000 core hours on the JUWELS Booster cluster. This cluster is housed at the Jülich Supercomputing Centre (https://www.fz-juelich.de/ias/jsc/EN/Home/home_node.html) in Germany and includes over 3700 latest-model NVIDIA A100 GPUS. These will be invaluable for training text-to-speech models for Swedish.
The model will generate audio streams quickly, so that users will be comfortable with natural conversation flow, without pauses for generating long replies. It will be implemented using existing Tacotron and WaveGlow technology, such as described in this blog post from NVIDIA https://developer.nvidia.com/blog/how-to-deploy-real-time-text-to-speech-applications-on-gpus-using-tensorrt/.



Photo: Fa Barboza // Unsplash

Business Impact

The speech model will be a key component of a conversational assistant capable of providing information in real time in response to spoken natural-language questions. It will be capable of learning to pronounce jargon relevant to particular domains, such as banking. More than 14 businesses across multiple industrial sectors, such as news organizations, energy and IT companies and insurance organizations are already using the model to automate processes and improve current solutions.

Benefits

- Make customer support much more accessible, by removing reuirements like visiting a bank branch, or having and being able to use a computer or mobile computing device.
- Quick audio generation
- Natural language for better user experience
- Accessibility

- NLP
- AI
- Speech Analysis
- Speech Generation

- Industry Sector: IT/HPC Systems, Services & Software Providers
- Technology: AI

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



Photo: Mike van den Bos // Unsplash

Large-Scale Real-Time Image Content Moderation

TRUBA HPC Center, operating under TUBITAK ULAKBIM, coordinates NCC Turkey. 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.
Our objective is to level up the industry, public sector, and academia in High-Performance Computing (HPC), High-Performance Data Analytics (HPDA), Artificial Intelligence (AI), and prepare for the developing EuroHPC ecosystem. Our competencies include HPC, HPDA, AI, CUDA, Materials Science, and Computational Fluid Dynamics (CFD). The particular third party presenting this success story is TRUBA.

Organisations Involved

Founded in 2010, Machinetutors (https://machinetutors.com/) provides machine learning consultancy and customized AI software development services. Machinetutors empowers businesses all over the world by solving real-world problems.
Machinetutors has two products:
• mtDATA, a data collection and annotation services platform,
• mtAPI, SaaS AI solutions with pre-trained models, customization options, and scalable infrastructure.
The client is a British content moderation SaaS start-up. Company name and further details are confidential.

machinetutors

Technical Challenge

This project addresses the problem of large-scale real-time image-based content moderation. The system is deployed to a production environment where tens of thousands of users browse the internet daily. The system must be both accurate and run in real-time to meet the business requirements. Moreover, the model size must be small so that multiple copies of the model can be run simultaneously on a GPU to reduce server costs. A major challenge has been making several models work efficiently together.

TRUBA Turkish Science e-Infrastructure

ODTÜ METU

Sabancı Üniversitesi

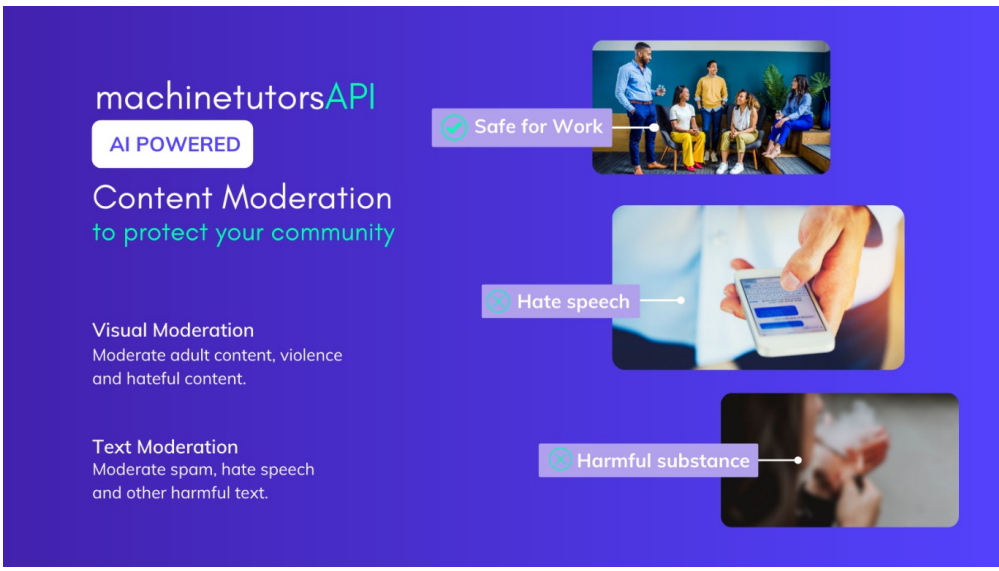
uhem



Photo: Karl Pawlowicz // Unsplash

Solution

In order to solve the problem defined, we developed three main models.
In the first model, we propose a multi-label NSFW classifier that can detect the NSFW levels (light, medium, hard) and predict other labels, such as the real person and clothing characteristics.
The second model is a one-stage body, based age & gender detection model. Current age & gender methods are both face-based i.e., they use face bounding boxes and are two-stage processes, they first run a face detector and then run the model on these boxes. When multiple faces are present in an image, this approach fails to meet the real-time requirement.
The third one is a segmentation model. These three models run in a pipeline via which we can run various scenarios.



Business Impact

Our client, a SaaS online content moderation start-up, is currently the number one content filter in their specific target market around the world thanks to the success of AI-supported high-tech features developed during this project.
User feedback on all our models from our clients' users are positive and they consider it to be the best product on the market. Thousands of users now browse the Internet with their adjusted moderation level. Our client has already reached a breakeven point financially.

HPC's speed and cost benefits enabled the project to be successfully delivered on time. All of the engineers on Machinetutors team are now proficient in using the TRUBA infrastructure due to this collaboration. We were able to work effectively and efficiently with our colleagues from TRUBA and look forward to the next project.

Benefits

- With this collaboration, we were able to run many experiments in parallel and quickly see the effects of the model updates
- With the ability to run large batch size trainings on newer GPUs, our experiments completed much faster
- Being able to access many GPUs at the same time enabled us to tune the hyper-parameters of each model to improve the results
- The speed and cost-efficiency provided by this support have helped us gain a considerable competitive advantage in the global AI ecosystem

- Artificial Intelligence
- Machine Learning
- Deep Learning
- Content Moderation
- Classification
- Segmentation
- Object Detection
- Data Collection
- Annotation

- Industry Sector: SME, Start-up

- Technology: HPC/HPDA, AI

Contact:
Email: contact@machinetutors.com

Section 2

Environment, Energy and Agriculture

WeatherAI – 30

CFD of Selective Catalytic Reaction – 32

Turbulent combustion modeling explosion safety assessment – 34

Wheater Forecast WRF-NMME Model – 36

Supercomputer Applications in Environmental Modelling - 38

An accurate AI-based Cloud Mask Processor for Sentinel-2 - 40

QualeAria-Local - 42

Machine learning can help with weed detection - 44

Railway Freight Car- 46

Predict the Damage- 48

Migrating AI workload to HPC – 50

Parallel Computing on TRUBA cluster – 52

High Efficiency Impeller Design - 54

Weather Research and Forecasting Model (WRF) - 56

Environmental Data- 58



Photo: Artuem.ro // Unsplash



Figure 1. Sample flooding event detection based on the image segmentation model of FloodCAM.

WeatherAI
by EuroCC Italy

WeatherAI – EuroCC Italy is the Italian competence center part of the EuroHPC Competence Centers Network of the EuroCC Project. The long-term objectives of EuroCC Italy are to structure and strengthen the collaboration between research groups and institutes, HPC data centers and companies, to support SMEs in innovation projects and to create awareness about the productive advantages possible through the use of the different solutions in the HPC / HPDA / AI field. EuroCC Italy is partnered by CINECA; ABD, BiREX, Leonardo Company and Dompè.

Industrial Organisations Involved

WaterView blends expertise in meteorology, hydrology, environmental modeling, computer vision and AI, looking for new solutions to solve challenging problems.
WaterView combines its cutting-edge computer vision technologies with simple cameras, to turn each cam into a non-dedicated multivariable sensor.
That is one small step for a camera, one giant leap to reduce the costs of granular, real-time weather and environmental monitoring.

The Technical Challenge

The first generation of models are based on physical models, require both device calibration and installation calibration, trained on large, annotated dataset.
In addition, these models require great hardware and software development.
WeatherAI wants to develop a second generation of AI analytics, more portable and scalable than the first-generation ones, trained on optimized datasets.

The Solution

Four GPUs were used for the model training. Parallel CPUs were used for image augmentation. Hyperparameters and datasets were optimized for 5 models.
To complete the project, a TensorFlow 2.X library was developed, to simplify training, exporting and quantization of AI models for edge applications (https://github.com/waterviewsrl/swiss-army-keras).
Randomized image augmentation pipelines, parallelized on node CPUs, were realized using the Albumentations library. To track dataset evolution using standard versioning tools (GIT), the Data Version Control System DVC was adopted. This helped the integration of Machine Learning / Data Science procedures and tools into the existing DevOps pipelines.

EUROCC ITALY
HPC Competence Center

waterview
your eye on climate change

Business Impact

The access to HPC resources helped in the development of the second generation of weather analytics, based on AI technologies.
The solutions provide telemetries related to the current weather conditions (rain precipitations, visibility distance), as well as alerts related to specific dangerous conditions (flooding or snow settling on road pavement):
1) RainCAM: detects the presence of rain precipitations, according to three severity classes
2) VisCAM: measures the visibility distance, in the range 25 to 1500 meters
3) FloodCAM: detects flooding events
4) SnowCAM: detects snow settling on the ground
5) SmoCAM: detects the presence of outdoor and indoor smoke and fire

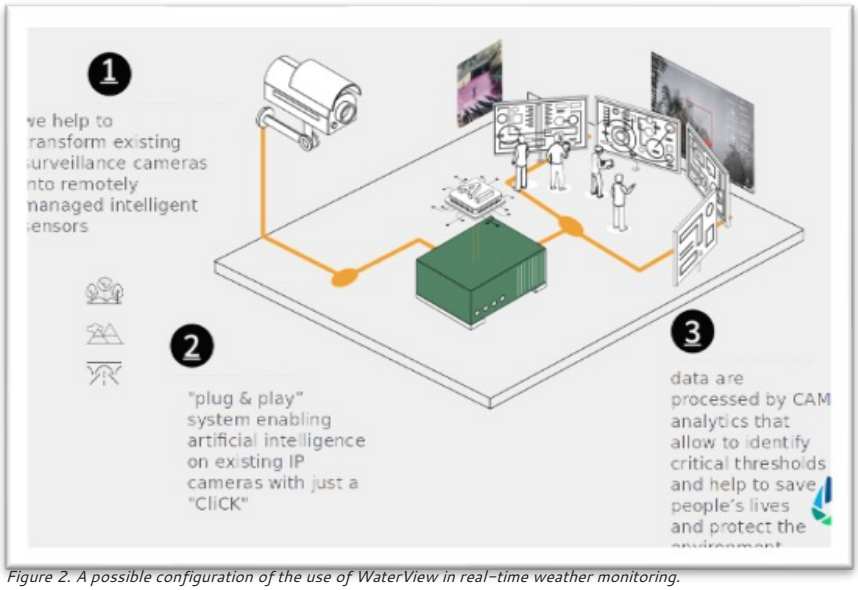


Figure 2. A possible configuration of the use of WaterView in real-time weather monitoring.

The following table reports the evaluation metrics for five models. Mean Intersection Over Union (mIOU) is reported for image segmentation models, while Accuracy and F1 metrics are given for the classification models

Modello	mIOU	Acc	F1
FloodCAM	0.94	-	-
SnowCAM	0.91	-	-
SmoCAM	0.78	-	-
RainCAM	-	0.96	0.95
VisCAM	-	0.91	0.89

Benefits

✓ Test-before-invest.
It is possible to test the feasibility of using AI to develop weather monitoring analytics, before buying the hardware
✓ Faster time to market (3/6 months)
✓ Support and tutoring provided by EuroCC/CINECA staff

- Image Augmentation
- Embarrassingly Parallel
- Dataset Optimization
- GPU
- Environmental Processing
- Real-Time Weather Monitoring

- Industry Sector: Environment/Climate/Weather

- Technology: HPC, AI

The use of cameras by WaterView, weather analytics can run on smart cameras or Edge-HPC / IoT gateways to provide early warnings to local authorities or infrastructure managers (Fig. 1).

Contact:
Matteo Ferrabone
+39 333 6373681
Email: matteo.ferrabone@waterview.it

Computational Fluid Dynamics of Selective Catalytic Reaction

The Czech National Competence Centre (NCC Czechia) for High-Performance Computing (HPC) and Data Analysis (HPDA) is represented by IT4Innovations National Supercomputing Center at VSB – Technical University of Ostrava. Its mission is to analyse, implement, and coordinate all necessary activities and to offer end users its services to meet their needs: from access to supercomputers and technology consulting to providing training for industry, public administration, and academia.

EURO CZECHIA

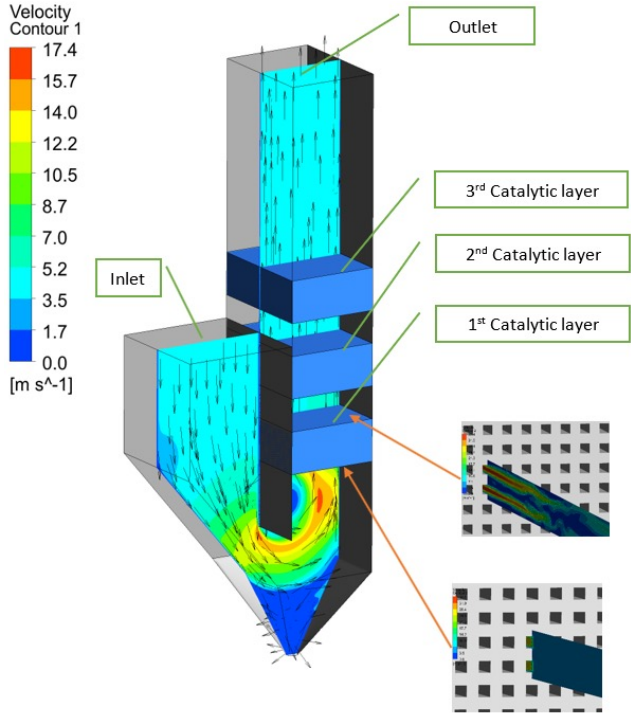
ORGREZ

What is ORGREZ?

The ORGREZ company (founded in 1992 and located in Ostrava in the Czech Republic) provides services and supplies in several specific fields of power engineering, thermal engineering and ecology, generally in the processes of fuel energy conversion and electricity production and distribution. The Ecological Systems Division deals with greenhouse gas issues, both in terms of technologies for controlling and reducing pollutant emissions and systems for monitoring and evaluating greenhouse gas emissions, especially CO2 and NOx.

The Scientific Challenge

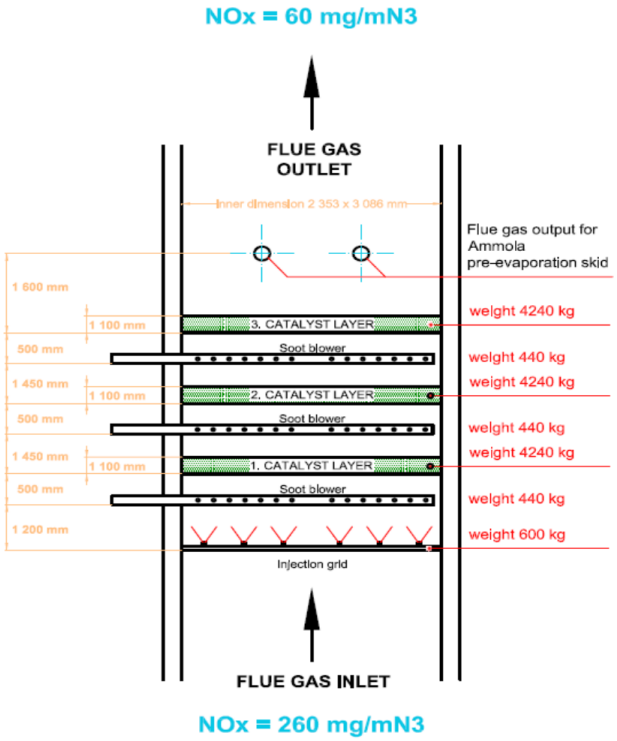
The main objective of the proof-of-concept was to determine whether Computational Fluid Dynamics simulations could be used for the fast and efficient description of the catalysis process of the selective catalytic reduction technology (SCR), and therefore can be used as a tool for the design of this technology, which is one of the existing technologies used to achieve the required reduction of the concentration of nitrogen oxides (NO, NO2, N2O) in the exhaust gases of combustion plants to the required level given by the applicable emission limits of specific combustion plants.



CFD simulation of SCR process

How we made it work

Numerical simulation of the SCR includes a problematic definition of inlet boundary conditions. CFD simulation of combustion process or measured field data can be used as a boundary condition for the SCR simulation. In this case SCR model was tested on the geometry of the coal combustion boiler and project design data specified by the catalyst manufacturer. The boiler is a waste incineration type where municipal waste is used as a fuel. The geometry of the combustion chamber part with catalyst was created according to drawing documentation. Results of CFD simulation were compared with task data. The main parameters were reduction of NO, pressure drop reduction of CO, and conversion of SO2 to SO3. The concentration of all species was calculated as an area-weighted average up and a downstream active catalytic layer or layers.



Basic sketch of the test reactor

The Impact on the Business

Currently, the design of SCR equipment is based on experience and simplified calculations. Using numerical modelling and simulations will make the process more efficient and faster and allow subsequent optimisation of the proposed solution. The use of HPC will enable these simulations to be finished in a reasonable time.
The outcomes of this proof-of-concept served as a base for the preparation and submission of the joint project to develop software for the efficient evaluation of catalytic processes and comprehensive support in designing SCR technologies for industrial applications. This project has been supported by the Ministry of Industry and Trade and will be implemented over the next 18 months. The knowledge gained from this proof-of-concept will be fully used in this follow-up project. In addition to numerical modelling and HPC, the algorithms and methods of artificial intelligence will be used to develop the software tool mentioned above.

Benefits

- ✔ Confirmation of the applicability of CFD for SCR design and optimisation;
- ✔ Time and costs saving due to speed up of SCR design process;
- ✔ Environmental impact due to optimised SCR design leading to increase of NOx emission reduction.

- Selective Catalytic Reduction (SCR)
- Computational Fluid Dynamics (CFD)
- High Performance Computing (HPC)
- Surface Chemical Reactions
- NOx reduction,

- Industry Sector: Energy, Environment

- Technology: HPC, CFD

Contact:
Tomas Karasek ,
Email: tomas.karasek@vsb.cz

Turbulent combustion modeling for explosion safety assessment

The Greek NCC "EuroCC@Greece" is run by a consortium of 5 institutions and coordinated by GRNET. The objective of the NCC is 3-fold: i) advance competitiveness in research ii) improve effectiveness of government services and iii) promote innovation in industry. In this particular success story, the hardware provided by GRNET and expertise of NCSRD in the field of CFD-based simulations were combined to provide services to the client.

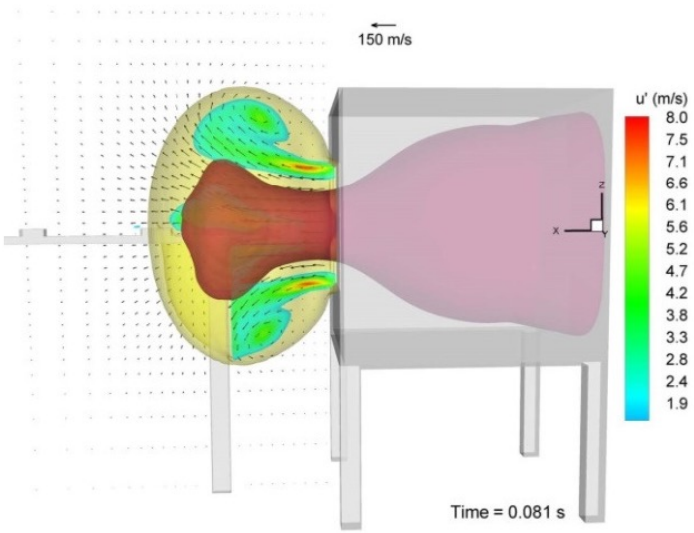
Hellenic Petroleum

Hellenic Petroleum (ELPE) is one of the leading energy groups in South East Europe, with activities spanning across the energy value chain and presence in 6 countries. ELPE has a range of activities that includes: (i) Supply, Refining, and Trading of petroleum products, both in Greece and abroad (ii) Fuels Marketing, both in Greece and abroad (iii) Petrochemicals Production and Trading (iv) Oil & Gas Exploration and Production (v) Power Generation & Natural Gas (vi) Provision of Consulting and Engineering services to hydrocarbon related projects.

Technical Challenge

Safety assessment of potential explosion accidents is conducted in industrial areas where flammable gases and vapors are processed to predict the resulting explosion overpressures. This assessment is usually carried out using simple integral codes, such as Phast. These codes require crucial input parameters which may have great impact on the overpressure predictions especially in case of complex geometry. The purpose of the project is to improve the predictive capabilities of these models in explosion scenarios in complex industrial areas using the more advanced and accurate Computational Fluid Dynamics (CFD) methodology.

A detailed digital representation of the industrial infrastructure (buildings, roads, pipelines, vessels etc) needed to be considered, thus, a very high-resolution grid was necessary, leading to high requirements in memory, number of cores and network bandwidth. At the same time, the problem itself, explosions and fire dispersion, demands complex methodologies and approximations that augment the problem of computer resources.



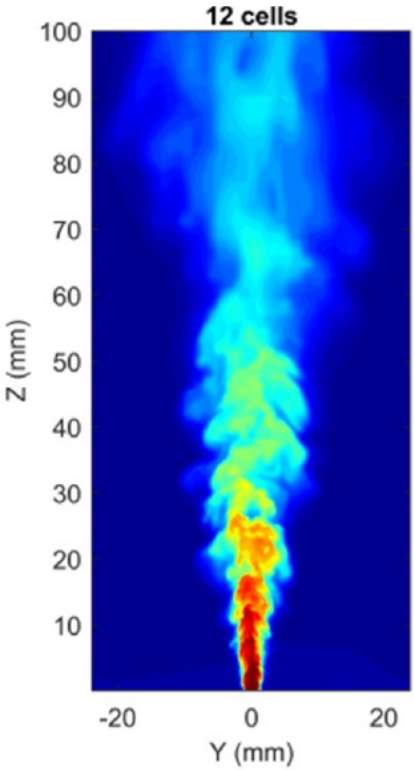
Simulation of an explosion in a vented room: Fuel mass fraction isosurface (yellow) and flame front surface (dark red) along with u' contours and velocity vectors at the onset of external explosion

How it got solved

The HPC problem domain falls within CFD validation methodologies/models against experiments in complex obstructed/confined environment. The problem was solved by using many computational resources in parallel, thus eliminating the need for assumptions and approximations. The ANDREA-HF CFD code that is parallelized with FORTRAN & MPI programming language was utilized. The infrastructure used was ARIS GRNET (20 THIN nodes, 200 cores, 500GB RAM, InfiniBand FDR 56Gb/s network).

Business Impact

The outputs of the project contribute to more accurate and reliable safety assessments. This will benefit the company by enhancing the safety protection, which is of high importance in the industrial area of the company due to the complexity of the facilities, the high number of employees and the hazardous nature of the processes. Moreover, the knowledge of the exact dynamics of the explosion using CFD analysis can indicate specific areas where more violent consequences will occur and as a result can suggest specialized more efficient safety measures.



Large Eddy Simulation of hydrogen accidental release: Instantaneous Hydrogen mole fraction contours for the case of 12 cells discretization of the release diameter (contour levels are from 0 to 1).

Benefits

The main benefits of the project are to further develop and validate existing CFD methodologies/models against experiments in complex obstructed/confined environment and to optimize/develop guidelines for integral code input data in the cases of explosion in complex industrial areas based on the CFD methodology.

In essence, the use of large HPC infrastructure enables the detailed representation of industrial area with complex geometry, the reduction in the modelling assumptions and the more robust solutions by performing many explosion scenarios.

- Computational fluid dynamics
- External explosion; Deflagration
- Physical explosion; Risk assessment
- Vapour cloud explosion
- Dispersion of flammable gases

- Industry Sector: Energy/Petroleum

- Technology: HPC

Contact:
Dr. Ilias Tolias, Post-doc Researcher
Affiliation: National Centre for Scientific Research "Demokritos", Environmental Research Laboratory
Address: Patr. Grigorieo & Neapoleos, Ag. Paraskevi, 15310, Athens, Greece
Email: tolias@ipta.demokritos.gr

Wheater Forecast WRF-NMME Model

Verification of the Temperature, Wind and Precipitation Fields for the High-Resolution WRF-NMME model over the complex terrain of Montenegro

EuroCC Montenegro is implemented at the University of Donja Gorica (UDG). Within the EuroCC project, UDG and other partners established a single National Competence Centre (NCC) in the area of high-performance computing (HPC). The NCC will coordinate activities in all HPC-related fields at the national level and serve as a contact point from customers from industry, science, HPC experts, and the general public. More information on EuroCC project activities in Montenegro can be seen on the https://eurocc.udg.edu.me/.

Institute of Hydrometeorology and Seismology (IHMS)

Hydrometeorological Institute of Montenegro is the national hydrological and meteorological service of Montenegro, controlling precipitation, maritime, meteorological, climatological, synoptical, agrometeorological and hydrological stations. The network of stations include 120 meteorological stations, 40 hydrological stations, 36 water quality stations and 17 air quality stations. There are 49 full-time employees, with over 100 part-time observers being engaged in observing, measuring and collecting of meteorological, hydrological and ecological parameters. More information on http://www.meteo.co.me/ .

Scientific Challenge

This solution presents a verification of the numerical Weather Research and Forecasting Non-Hydrostatic Mesoscale Model (WRF NMM) for weather prediction, using the High-Performance Computing (HPC) cluster over the complex relief of Montenegro.
The Weather Research and Forecasting (WRF) Model is a next-generation mesoscale numerical weather prediction system designed for both atmospheric research and operational forecasting applications. Verification was performed comparing WRF NMM predicted values and measured values for temperature, wind and precipitation for six Montenegrin weather stations in a five-year period using statistical parameters. The difficult task of adjusting the model over the complex Montenegrin terrain is caused by a rapid altitude change in the coastal area, numerous karst fields, basins, river valleys and canyons, large areas of artificial lakes on a relatively small terrain. The subject of the research are meteorological measuring stations in the cities of: Žabljak, Pljevlja (north of Montenegro- mountain area) as well as the Golubovci (near Podgorica), Niksic, (central area of Montenegro) and Bar, Tivat (south of Montenegro - costal area).

The Solution

The data sources used in analysis are eight-hour data reports. By decoding, filtering and sorting them, a database of eight-hour meteorological quantities was created which will be used during verification for each of the stations.
After obtaining prognostic meteorological parameters for 24 hours in advance, a database of hourly modelled and eight-hour temperature values, at a height of 2m and wind speed at a height of 10m was created, while precipitation was summarized in twelve-hour intervals, i.e. from 6 a.m. to 6 p.m. and from 6 p.m. until 6 a.m. the next day. The timeframe on which data was collected for further processing is 5 years, i.e. from 2017 to 2021.

Business Impact

The necessity of setting up high-resolution models is essential to timely forecast dangerous meteorological phenomena. Based on the obtained verification results, it has been proven that WRF NMME model is a good solution for weather prediction in Montenegro.
Basic meteorological parameters of temperature, wind and precipitation and the performance of the model will depend on the time of day, the season, the altitude of the station, as well as the surrounding relief of stations, for which the model results were verified.
Institute for Hydrometeorology will adjust WRF NMME model for national prediction according to results/verification of HPC simulations. Potential business impact could be explored and verified thanks to more precise weather forecasting/intelligence services, also on localized level (subscription-based or dedicated report fee) related to the areas of meteorological phenomena, smart agriculture, tourism activities, green energy, optimized/safe transportation etc.).

Benefits

- ✓ More precise numerical model for weather/climate observations and prediction for Montenegro
- ✓ Multiple application of weather prediction in smart agriculture, tourism, transportation, environment/climate/weather ...

› Weather orecast, Montenegro, WRF-NMME model, model verification

› Industry Sector: Agriculture, Construction, Earth science, Environment/climate/weather, Food and drink, IT/HPC systems, services & software providers, Life sciences, Maritime, Smart City ....)

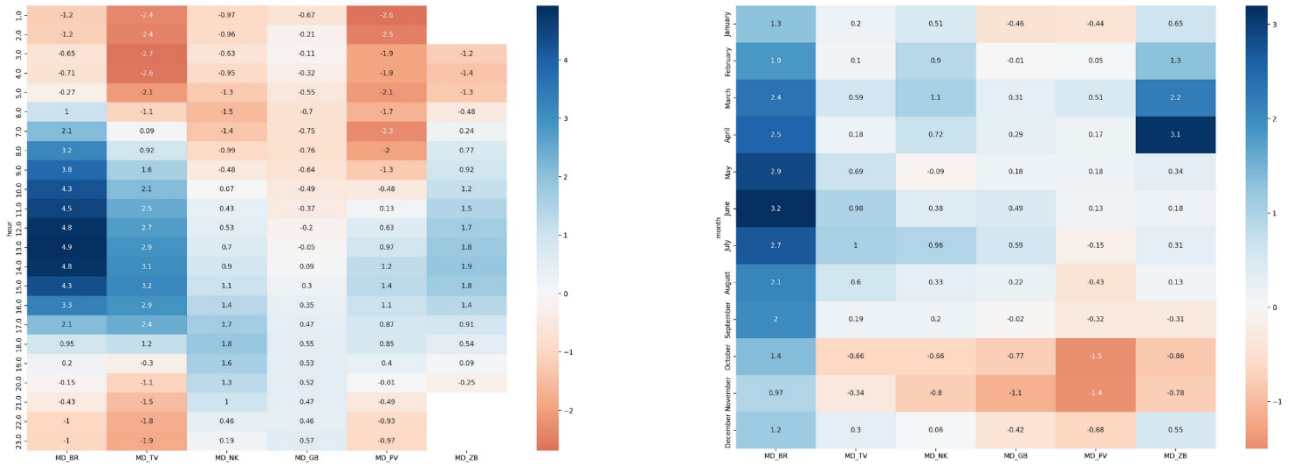
Technology: HPC

›

Contact:
Luka Filipović, NCC Montenegro, UDG,
Email: luka.filipovic@udg.edu.me

Angel Marčev, HMICG,
Email: angel.marcev@meteo.co.me

Aleksandar Zečević, HMICG,
Email: aleksandar.zecevic@meteo.co.me



Mean difference temperature heat map. Shades of blue means underestimation of model while the red shade shows overestimation of modelled temperature.

Supercomputer Applications in Environmental Modelling

NCC-Bulgaria in the area of HPC, HPDA and AI has the goal to enhance and develop the competences of the Bulgarian computational community. NCC-Bulgaria is built by a consortium coordinated by the Institute of Information and Communication Technologies at the Bulgarian Academy of Sciences (IICT-BAS), and 2 members: Sofia University "St. Kliment Ohridski" (SU) and University of National and World Economy (UNWE). The partners are specialized in the area of HPC and ICT in general, to ensure achievement of the project objectives and guarantee the overall success. The partners collaborate with Sofia Tech Park, where the Discoverer EuroHPC supercomputer is operating.





IICT-BAS

The Department of Parallel Algorithms (DPA) is one of the departments of IICT-BAS. The main research activities of DPA are in the areas: Parallel algorithms and distributed computing; Mathematical modelling and scientific computation in Environmental Mathematics, Semi-conductors physics; Monte Carlo algorithms; Development of Monte Carlo approaches for solving Wigner-Boltzmann equation and Barker-Ferry equation.
DPA is an active participant in a number of research and educational projects of the EU programs as well as in NATO Scientific Programs for cooperation between EC and Eastern Europe.

The Challenge

The environmental modelling and air pollution modelling in particular is one of the toughest problems of computational mathematics. All relevant atmospheric processes are to be taken into account, which are mathematically represented by a complex PDE system. In order to simplify it proper splitting procedure is applied. As a result, this system is replaced by several simpler systems (submodels), connected with the main physical and chemical processes. These systems should be calculated in a large spatial domain, as the pollutants can be moved quickly on long distances, driven by the atmosphere dynamics. One major source of difficulty is the high intensity of the atmospheric processes, which require a small time-step to be used in order to get a stable numerical solution. All this makes the treatment of large-scale air pollution models a tuff and heavy computational task.

Scalability of UNI-DEM (480×480×1 grid, 35 species , CHUNKSIZE=32)

NP (MPI)	Time (s)	*Speed-up*	Efficiency
1	195036	--	--
4	48604	*4.0*	100 %
8	25045	*7.7*	96 %
12	17036	*11.3*	94 %
24	9150	*21.1*	88 %
48	4805	*40.1*	84 %
96	2701	*71.4*	74 %
120	2394	*80.6*	67 %
160	2052	*94.0*	59 %

Time (T) in seconds and speed-up (Sp) (with respect to the number of MPI tasks, given in the first column) for running UNI-DEM on the DISCOVERER supercomputer

The Solution

A large-scale environmental model – the Danish Eulerian Model (DEM) was implemented and optimized for performance and efficient work on various long-term scenarios on the new petascale supercomputer DISCOVERER. It is represented by a system of PDE for calculating the concentrations of a number of pollutants in the atmosphere above a large geographical region. The main atmospheric processes (horizontal and vertical wind, diffusion, chemical reactions, emissions and deposition) are adequately represented in the system.
The MPI library is used as a main parallelization tool. MPI parallelization is based on the space domain partitioning. The space domain is divided into sub-domains and each MPI task works on its own sub-domain. There is no data dependency between the MPI tasks on both the chemistry and the vertical exchange stages, which is not the case for the advection-diffusion stage. Spatial grid partitioning between the MPI tasks requires overlapping of the inner boundaries and exchange of certain boundary values on the neighbour subdomains for proper treatment of the boundary conditions. The numerous small calculation tasks are grouped in chunks for efficient cache utilization.

The Scientific impact

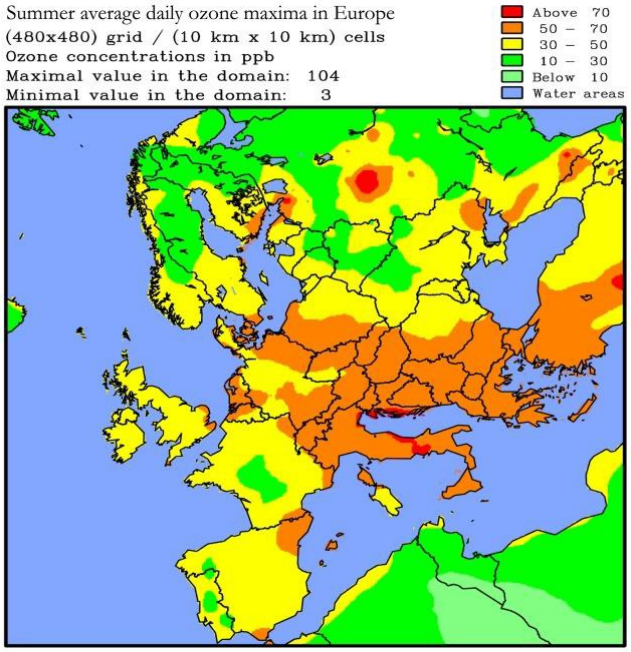
DEM is a powerful and flexible large scale air pollution model, capable of calculating the levels of concentration for a number of dangerous pollutants and other chemically active species interacting with them (precursors), over a long time period. Moreover, various accumulative quantities (AOT40, AOT60, etc.) can be calculated on yearly basis, which have significant impact in the area of agriculture (on the yield of crops in particular), forestry, wildlife and human health. Over the years it has been successfully applied in many scientific and practical problems in various important areas (environmental, medical, social, economic, etc.). Some of the successful applications of DEM include:

- Detailed air pollution evaluation, helping the decision-makers to plan the industry development and take proper protection measures, when necessary;
- The impact of air quality on human health, forestry and wild life;
- The impact of long-term air pollution on agriculture (more specifically, yield & losses of crops estimation) and its economy consequences.

Moreover, DEM has been used for simulations in several critical areas (global climate changes, industrial accidents, natural disasters, etc.).

Benefits

- ✔ Optimization of the unified parallel version of the Danish Eulerian Model (UNI-DEM) on the new petascale EuroHPC supercomputer Discoverer;
- ✔ High scalability, especially on the most time-consuming stages (shown in the table below);
- ✔ Efficient use of nested levels of parallelism (MPI and OpenMP).



Summer average daily ozone maxima in Europe, calculated by the fine grid version of UNI-DEM (with 10x10 km. cells)

- Environmental modelling
- Air pollution
- Supercomputer
- Parallel computing
- Scalability

- Industry Sector: Agriculture, Forestry, Healthcare

- Technology: HPC, MPI, OpenMP

Contact:
Assoc. Prof. Tzvetan Ostromsky,
Email: ceco@parallel.bas.bg

Prof. Ivan Dimov,
Email: ivdimov@bas.bg
Institute of Information and Communication Technologies,
Bulgarian Academy of Sciences

An accurate AI-based Cloud Mask Processor for Sentinel-2

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

EURO ESTONIA

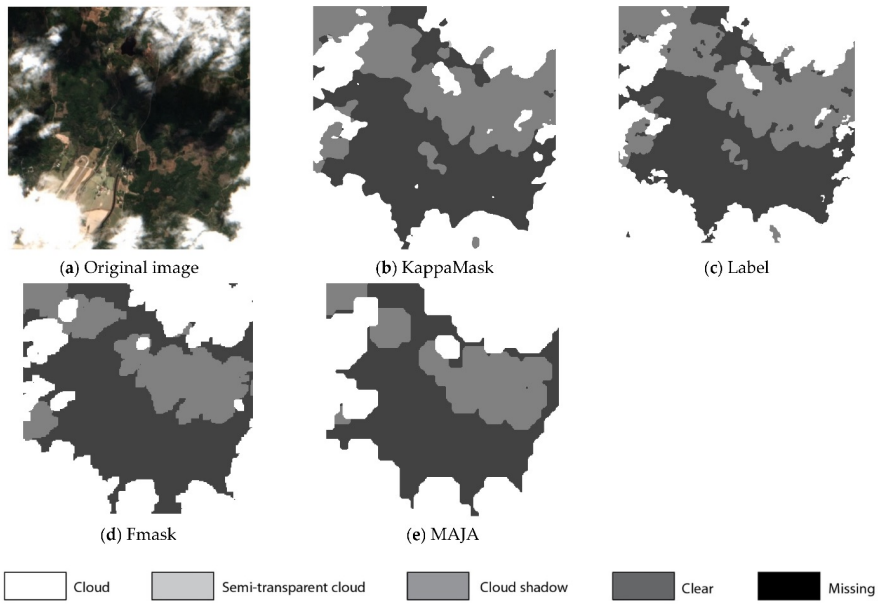
KAPPAZETA

KappaZeta

KappaZeta is a science-driven remote sensing company aiming to make space a valuable asset for everyone. KappaZeta's expertise is in using SAR (radar) satellite data, incorporating it with optical satellite data and providing some of the most accurate AI models on the market. Their key area of focus is agriculture.

Technical Challenge

Sentinel-2 is an Earth observation mission from the Copernicus Programme that systematically acquires optical imagery at high spatial resolution (10 m to 60 m) over land and coastal waters. One of the factors limiting the usage of Sentinel-2 data is clouds. Therefore, the data from Sentinel-2, like all optical satellite imagery, needs to have cloud and cloud-shadow areas identified and filtered out. KappaZeta addresses the problem by introducing KappaMask, an AI-based cloud and cloud shadow masking processor for Sentinel-2. As a cloud detector, KappaMask uses a large convolutional segmentation model. Faster model convergence during training can be achieved by using larger batch sizes of the training data, which means more GPU memory is needed. Additionally, faster CPUs are required for shorter data loading times to increase the training speed even further.



Comparison of L2A prediction output for a 512 × 512 pixels sub-tile in the test dataset. (a) Original Sentinel-2 True-Color Image; (b) KappaMask classification map; (c) Segmentation mask prepared by a human labeller; (d) Fmask classification map; (e) MAJA classification map.

Solution

KappaMask was trained on an open-source dataset and fine-tuned on a Northern European terrestrial dataset which was labelled manually using the active learning methodology. The training was performed on the University of Tartu's HPC Centres' high-performance compute nodes. Powerful GPUs and CPUs were applied to substantially speed up the training of the model. In order to measure the accuracy of Kappamask, a comparison with similar solutions (Sen2Cor, Fmask, MAJA, S2cloudless and DL_L8S2_UV) was carried out on the test dataset, which was not seen by the model at any stage. In this test, the total Dice coefficient of KappaMask model was 80%. With this result, KappaMask outperformed similar AI-approaches currently in use.



Clouds in Bergen, Norway. Photo: Cosmic Timetraveler // Unsplash

Business Impact

KappaMask is an open-source project. All the results, final software and source code will be freely and openly distributed in GitHub. Openness and accessibility of the software should directly translate into higher use by developers.

Benefits

- ✔ Reliable cloud mask processor for Northern Europe region, which is compatible with ESA Sentinel-2 L2 processing chain.
- ✔ Creation of high quality reference dataset for future developments.
- ✔ Innovative application of deep learning techniques in cloud masking.

- Remote sensing
- Machine learning
- Image segmentation
- Cloud mask
- Convolutional neural network

- Industry Sector: Agriculture
- Technology: HPC, HPDA, AI, Deep Learning

Contact:
Ülar Allas,
Email: ylar.allas@ut.ee



Photo: Homepage EuroCC Italy

QualeAria-Local

EuroCC Italy is the Italian competence center part of the EuroHPC Competence Centers Network of the EuroCC Project. The long-term objectives of EuroCC Italy are to structure and strengthen the collaboration between research groups and institutes, HPC data centers and companies, to support SMEs in innovation projects and to create awareness about the productive advantages possible through the use of the different solutions in the HPC / HPDA / AI field. EuroCC Italy is partnered by CINECA; ABD, Bi-REX, Leonardo Company and Dompè.

EUROCC ITALY
HPC Competence Center

Involvment of ARIANET

ARIANET is a consultancy company whose purpose is to contribute to the understanding of processes involving pollutants in the atmosphere.
ARIANET's tools are aimed at meteorological reconstruction, treatment of emissions and simulation of processes involving pollutants in the atmosphere: transport and diffusion, chemical transformation and deposition on the ground.
ARIANET supports public and private clients in following fields: impact assessment of industrial plants, implementation of air quality forecasting and assessment systems, source apportionment analysis.

ARIANET
A suez company

Technical Challenge

The pollution inside urban area is determined by following contributions:
· **regional**, due to the sources present in the surrounding areas (substantially uniform over a large portion of the agglomeration);
· **urban**, relating to all sources of pollution within the city whose distribution can be considered uniform (e.g. building heating);
· **street-level**, highly inhomogeneous, exceeding the previous two (due to traffic emissions and to the chemical-physical processes that take place inside the road canyons).
The first two contributions can be simulated by coupling atmospheric chemistry models and Machine Learning models. The third one can be evaluated by means of micro-scale dispersion models.

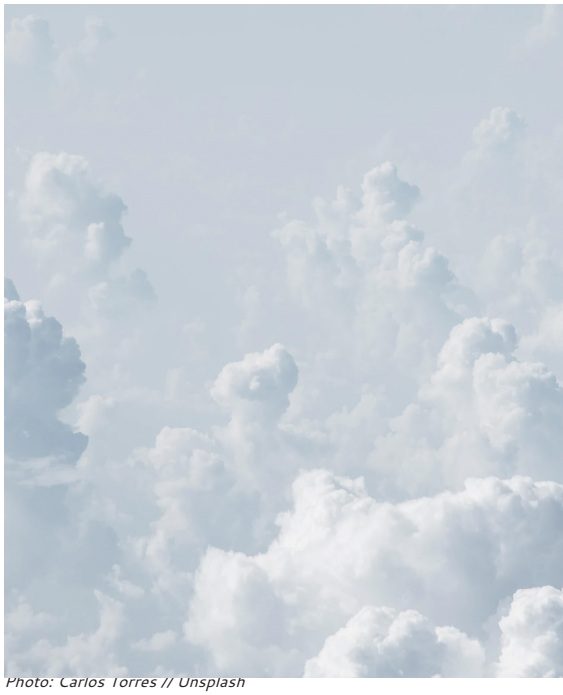


Photo: Carlos Torres // Unsplash

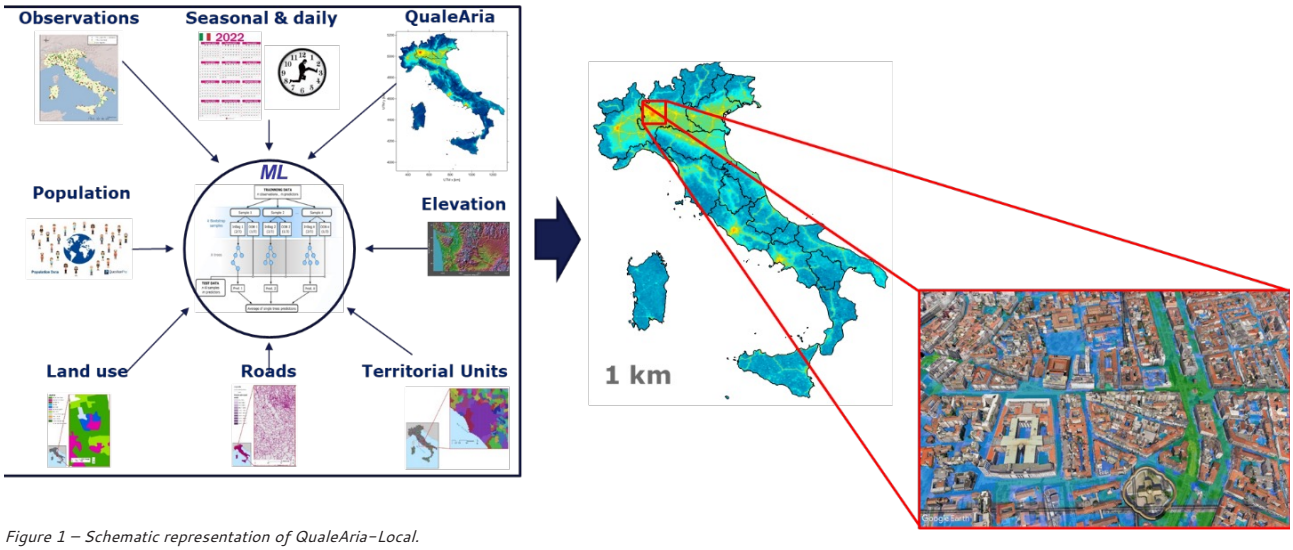


Figure 1 – Schematic representation of QualeAria-Local.

Solution

QualeAria-Local consists of three elements: QualeAria Air Quality Forecast System (AQFS), Random Forest Machine Learning (ML-RF) algorithm and μ-scale Parallel-Micro-Swift-Spray (PMSS) modelling suite. The PM10, PM2.5, NO2 and O3 concentration fields produced by Qualearia for a test period (year 2021), together with a set of spatial-temporal predictors and available air quality observations were processed by the ML-RF algorithm to obtain corresponding higher resolution (1 km) concentration over the national territory. The resulting concentration fields provide the regional and urban background fields to the PMSS suite which is used to produce forecasts of atmospheric particulate matter and nitrogen oxides at very high spatial resolution (4 m) over a selected urban area (Milan conurbation). Figure 1 provides a schematic representation of QualeAria-Local.

Business Impact

The project confirmed the potential of ML models in the field of air pollution, to produce air quality forecasts at high spatial resolutions with a lower computational burden than that required for Chemical Transport Model (CTM) simulations.
The concentration fields produced by the CTM and the seasonal indicators are here the most important spatial-temporal predictors considered by the ML algorithm used (Random Forest).
The performance of the ML model highlighted, in some cases, a reduction of the systematic errors produced by the CTM, without worsening the results obtained.
The project confirmed the need for HPC resources to further increase the spatial resolution of air quality forecasts in urban areas. The constituent elements of QualeAria-Local are computationally intensive and largely already parallelized. For example, a 26-hour simulation, carried out using the PMSS modelling suite, on a domain of dimensions equal to 10x11 km2, which includes the urban area of Milan, at a resolution of 4 m. (2500x2750 grid points) required a calculation time equal to about 90 minutes using 391 cores on the Galileo100 HPC infrastructure.
Estimating the levels of pollution near roads and inside road canyons is important as in these areas some pollutants reach particularly high concentration levels and can be harmful to the exposed population.

Benefits

- ✔ Identification of the computational resources needed to evaluate the air quality within the Milan urban area up to the street level.
- ✔ Procurement of forecasts of air quality over a urban area at a horizontal resolution of 4 m.
- ✔ Procurement of regional air quality assessments at 1 km horizontal resolution for air quality assessment and for epidemiological surveys.

- Air pollution
- Air quality forecasting
- Spatial resolution
- Machine learning
- Random forest
- Time to solution

- Industry Sector: Environment, climate, weather, Health care

- Technology: HPC, ML

Contact:
Camillo Silibello,
Phone +39-02-69331606
Email: c.silibello@aria-net.it

Machine learning can help with weed detection

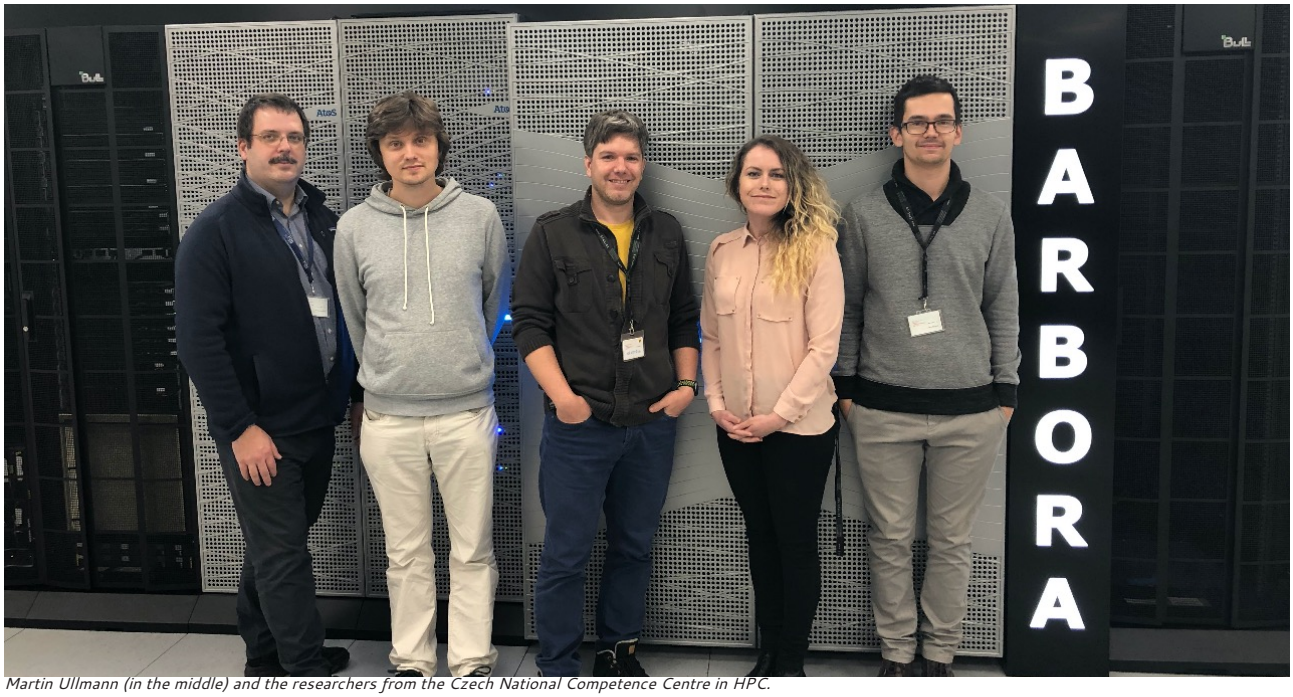
The Czech National Competence Centre (NCC Czechia) for High-Performance Computing (HPC) and Data Analysis (HPDA) is represented by IT4Innovations National Supercomputing Center at VSB – Technical University of Ostrava. Its mission is to analyse, implement, and coordinate all necessary activities and offer end users its services to meet their needs: from access to supercomputers and technology consulting to providing training for industry, public administration, and academia.

Industrial organisation involved

The Ullmanna company was founded in 2019 by IT enthusiast Martin Ullmann and his father-in-law; the farmer Jindřich Ullmann. The company´s located in Opava (Czech Republic). The mission of the company is to assist global farming with the necessary transition towards more sustainable and organic agricultural practices. The Ullmanna company focuses on decreasing or complete abandonment of pesticide utilization in modern farming, which is one of the key elements in addressing the urgent challenge of climate change and environmental degradation.

Technical/Scientific Challenge

The vision of the Ullmanna company is to boost the efficiency and eco-friendliness of farming in conventional, organic, and sustainable farms. Nevertheless, it comes with the complex challenge of growing more organic crops without an enormous increase in the required workforce for weed control. Therefore, the Ullmanna company is developing an agricultural weeding machine that will enable in-line weed control by recognising the target crop using machine learning. This will lead to farming activities without the use of chemical sprays.



Martin Ullmann (in the middle) and the researchers from the Czech National Competence Centre in HPC.

Solution

To automatically identify crops, machine learning was employed, and a neural network was designed and trained on HPC infrastructure. Crop recognition, in this case sugar beet from weeds, allows the weeding machine to remove the weeds while not damaging the grown crop. The goal was to create a neural network model that can be used for inference on the weeding machine while meeting the constraining conditions of deployment (limited hardware, inference speed, recognition accuracy). Two neural network architectures were designed and tested. The first was based on regression; it was easier to train and offered better performance, but only allowed the detection of one crop per photo. Subsequently, an architecture using object detection was proposed since it allows the detection of any number of objects in a photo.



The weeding machine attached to a tractor.

Business impact

The Czech National Competence Centre researchers in HPC have participated in a unique project with the Ullmanna company. Both neural network architectures were optimised to run on the embedded device at 60 frames per second, allowing them to be used in real-time on the weeding machine.

The Barbora supercomputer (using the TensorFlow framework), located at IT4Innovations, was used to train the object detection neural network based on datasets provided by Ullmann. This trained neural network was tested in the field with a weeding machine prototype, and its performance was evaluated.

The cooperation on this project started thanks to the Digital Innovation Hub Ostrava, which connects the activities of IT4Innovations and the Moravian-Silesian Innovation Centre Ostrava, enabling SMEs to examine and solve their needs in the field of digitalisation, and which was financially supported by the Moravian-Silesian Region. The result of this cooperation led to further activities, e.g., obtaining a grant in the open call of the agROBOfood project funded by the Horizon 2020 programme, which will help develop the product's commercial potential.

Benefits

- food production: 3x more profit per hectare
- use of chemicals: 40% pesticide reduction
- environment and health: no contamination of groundwater and no negative effect on workers´ health and safety
- workforce: reduced time and required workforce

- Eco-friendly farming
- Weeding machine
- Machine learning
- Object detection
- Neural network

- Industry Sector: Agriculture, Weeding Machine

- Technology: Machine learning, Object detection, neural network

Contact:
Jan Martinovic ,
Email: jan.martinovic@vsb.cz

Using bulk simulation in the development of a

Railway Freight Car

The Czech National Competence Centre (NCC Czechia) for High-Performance Computing (HPC) and Data Analysis (HPDA) is represented by IT4Innovations National Supercomputing Center at VSB – Technical University of Ostrava. Its mission is to analyse, implement, and coordinate all the activities required to offer end users the services they need: from access to supercomputers and technology consulting to providing training for industry, public administration, and academia.

EURO CZECHIA

ADVANCED ENGINEERING s.r.o.

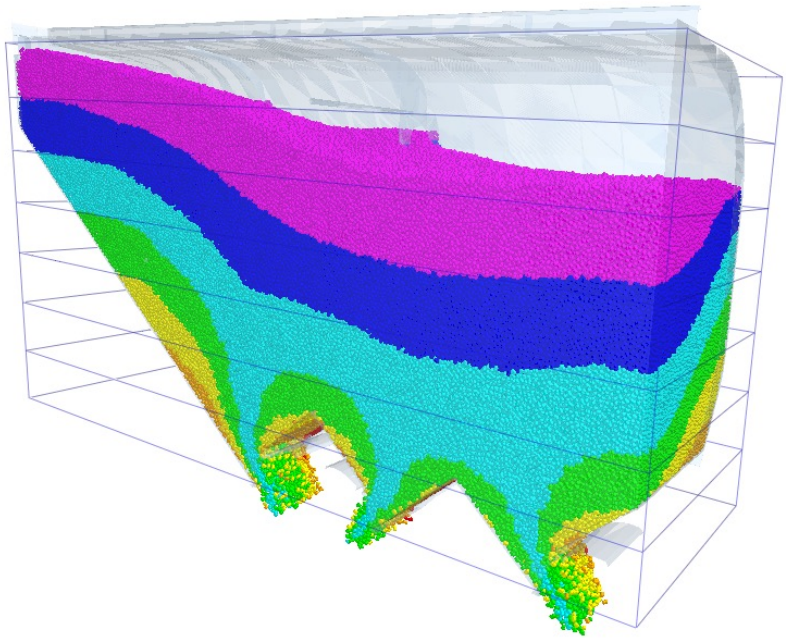
Industrial organisation involved

Advanced Engineering was founded in Prague in 2006 and is a well-established Czech technology and engineering company that focuses on computer simulations, structural analysis and optimisation of structures, and multi-physics modelling and simulation www.advanced-eng.cz/en.

Technical Challenge

NCC facilitated access to the HPC infrastructure operated by the IT4Innovations national supercomputing center and provided support for the access to the supercomputers and during the deployment of the simulation software.

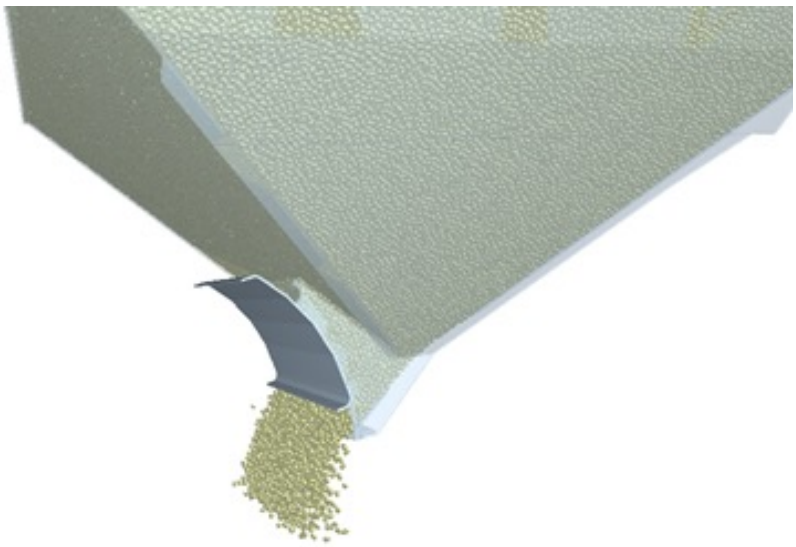
Advanced Engineering carried out a pilot project to simulate the emptying of a newly developed truck; a 103 m³ Tagnpps model designed primarily for grain transport. Because clients expected a guarantee for the time required to empty each wagon during unloading, the development team needed to assure that the designed geometry of the bins and hoppers would ensure complete emptying, without grain sticking to the walls. Knowing the force conditions during emptying, i.e., the action of the grain flow, helped to properly size the closing mechanisms so that the opening and closing forces were suitable for the operator.



Ongoing simulation of emptying a freight wagon box – one-quarter simulation model.

Solution

CAE specialists at Advanced Engineering used the Altair EDEM software tool (www.altair.com/edem), a leading Discrete Element Method (DEM) simulation solution, for computer simulations of bulk material. The particular problem for the freight car was the dimensions given by the ratio of the cargo space and the particle size; in other words, the number of elements representing grains, which for such a problem is in the order of millions.
For each element and each time step, the DEM calculates the interactions of the particles with respect to each other and the environment, and thus calculates the positions of the elements for the next step. In summary, we get a simulation of the grain discharge stream and can observe how each layer of the fill empties, which locations empty first, which empty last, and why. To solve such a problem in a reasonable time, HPC infrastructure is required.



Calculations of the frictional and normal force from the flowing bulk material (grain) on the hopper closing mechanism.

Business Impact

The advantage of computer simulations of bulk material movement compared to physical testing is that the cost of renting the testing space (in our case a grain silo) and the cost of the test fill (grain) are eliminated.
An added benefit is the ability to simulate the interaction of the bulk material with the structure, the ability to compare multiple design options, and to more easily analyse the behaviour of different types of materials (grain) under different external conditions (temperature, moisture).
From the user's point of view, it is important to find out what the real-time requirements for such simulations are, and of course, what costs would have to be considered for further simulations.
From an Advanced Engineering perspective, it was important for the Altair EDEM simulation tool to know where the so-called „performance breaks" were, i.e., how many processors are required to maximise the efficiency of both speed and cost of computation. The second task was to compare the performance and cost-effectiveness of computations on standard (CPU) or graphics processing units (GPU).

Benefits

- ✓ Time and costs savings through numerical modelling and simulations;
- ✓ The ability to simulate the interaction of the bulk material with the structure;
- ✓ The ability to compare multiple design options and to more easily analyse the behaviour of different types of grain materials under different external conditions (temperature, moisture).

› Discrete Element Method (DEM)
› High-Performance Computing (HPC)

› Industry Sector: Railway industry, Agriculture

› Technology: HPC, Discrete Element Method (DEM)

Contact:
Tomas Karasek,
Email: tomas.karasek@vsb.cz



Photo: Joshua Raswon-Harris

Predict the Damage

to infrastructures caused by storms and map more precisely the risks of claims for insurance

The French NCC: CC-FR, dedicated to HPC, HPDA and AI technologies, 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.

CC-FR CENTRE DE COMPÉTENCE HPC.HPDA.IA

Organisations Involved

CC-FR, set up within the framework of EuroHPC, is managed by the European Pole of Competence in high performance digital simulation Teratec, in association with the European Center for **Advanced Research and Training in Scientific Computing** Cerfacs and with the participation of the Big National Equipment Intensive Computing Genci.

Risk Weather Tech, https://www.riskweathertech.com, is a Climate Modelling Specialist. RiskWeatherTech provides risk management and for the study of the vulnerability of territories and companies.

Risk Weather Tech

Technical Challenge

Risk Weather Tech integrated the CC-FR Tailored Program, to benefit from a high-level support in intensive computing. In collaboration with CRIANN (Regional Centre Information Technology and Digital Applications Normandy), Risk Weather Tech was able to access the MYRIA supercomputer, equipped with more than 10,000 computing cores connected in very high throughput, and capable close to 700 trillion transactions per second.

Teratec

CERFACS
CENTRE EUROPÉEN DE RECHERCHE ET DE FORMATION AVANCÉE EN CALCUL SCIENTIFIQUE

GENCI
Le calcul intensif au service de la connaissance

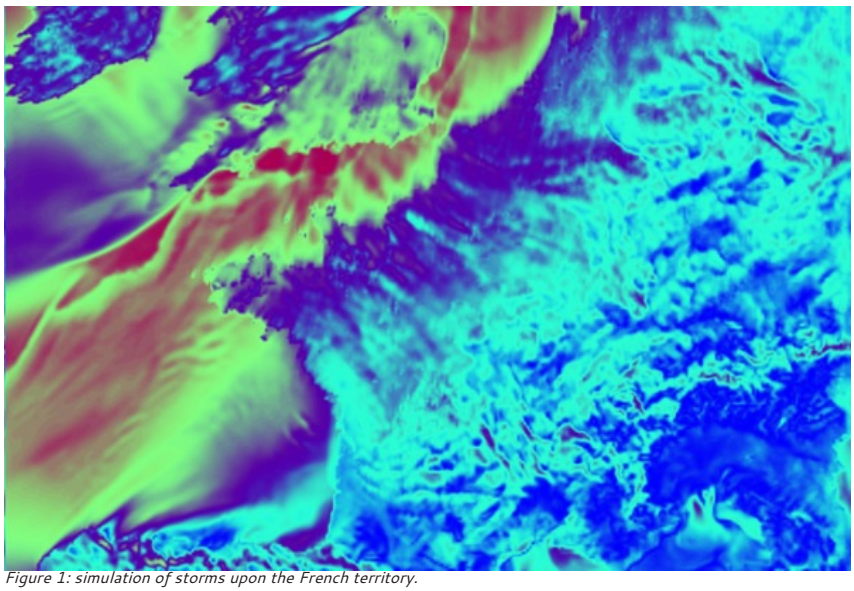


Figure 1: simulation of storms upon the French territory.

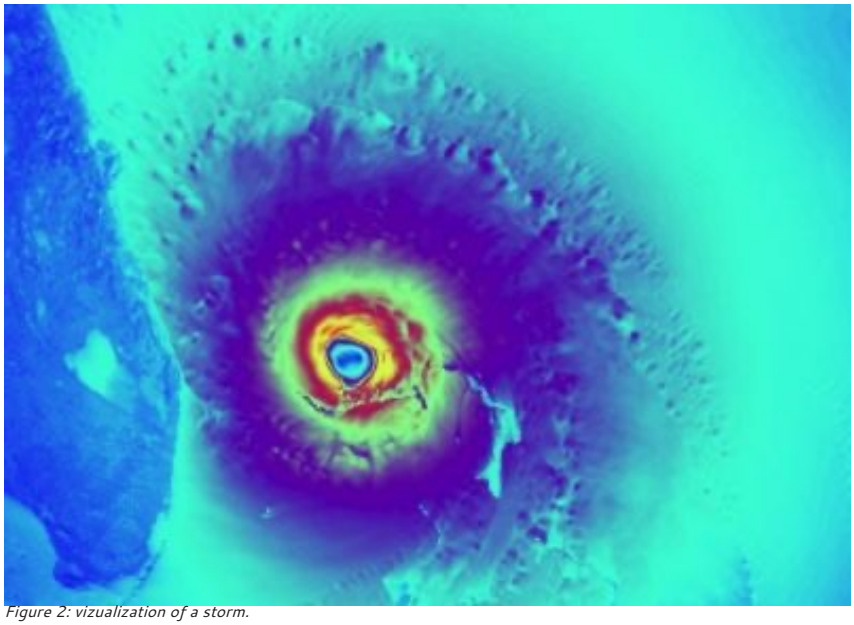


Figure 2: vizualization of a storm.

Solution

Risk Weather Tech has developed, thanks to the Criann's Myria supercomputer and due to the recurrence of intense weather disturbances in France, a vast catalogue of storms physically and statistically plausible. Their main objective in the concerned use case was to anticipate potential damage to infrastructures and to map more precisely the risks of claims for insurance. Using the Myria super calculator of CRIANN, Risk weather Tech has performed 10,000 probable and realistic storms simulation in order to forecast what could happen on the French territory. More than 870,000 CPU hours on the Criann's Myria supercomputer was needed to achieve these simulations.

Business Impact

The catalogue of the 10,000 realistic simulated storms will enable insurers to refine their storm cost estimates as requested by the French Authorities. It could also provide a better understanding of the vulnerability of large infrastructures recently installed on French territory, such as wind turbines or solar panel fields.

Benefits

- This first Catalogue of 10,000 Storms simulation permitted to propose service offerings to predict the damage to infrastructures caused by storms and to map more precisely the risks of claims for insurance.

- HPC
- Simulation
- Algorithm
- Parallelization
- Environment
- Climate and weather

- Industry Sector: Environment/ climate/weather

- Technology: HPC, HPDA, AI

Contact:
Dr. Karim Azoum
Phone +33 7 62 74 03 60
Email: Karim.azoum@teratec.fr
LinkedIn: https://www.linkedin.com/in/karim-azoum-45011710a

Migrating AI workload to HPC

The infrastructure provider Sigma2 and the research companies SINTEF and NORCE have joined forces to establish The Norwegian Euro Competence Centre. The competence centre raises awareness and provides Norwegian SMEs with the expertise necessary to take advantage of innovation possibilities created by HPC (High-Performance Computing) technology. This includes HPDA (High-Performance Data Analytics), ML (Machine learning) and AI (Artificial intelligence), thus increasing the SME's competitiveness.

Industrial organisations Involved

This success story is a result of a proof-of-concept for introducing HPC in an AI workflow. The workflow is detecting field boundaries and seeded areas from sentinel-2 data. DigiFarm is one of Norway's leading agri-tech startups specializing in the development of deep neural networks for automatic detection of field boundaries and sown areas based on high-resolution satellite data. This is primarily used by their customers for verification of subsidies, crop forecasts, variable fertilization/ spraying, carbon mapping, sustainability and traceability verification, grain species detection and regenerative agriculture.

DigiFarm´s core operation consists of the development of technology and model training of deep neural networks that can automatically detect field boundaries, sown area, and other objects within land changes such as trees, bogs, shade (from trees), telephone poles, irrigation divots. The models are trained in several international regions including Europe, South America, APAC (Thailand, Vietnam, and India) as well as the USA and Canada. The SME received EIC Accelerator funding to expand their market outside of Norway, requiring a large amount of sentinel-2 data to be processed within reasonable time.

The Challenge

The workflow was running in cloud with high cost and long runtime, it had a monolithic nature where mixed processes needing CPU, GPU or database access, were executed on the same resources serially, resulting in large cost for high performance resources that was not effectively used. The high level process was typically: download data -> pre-process data on CPU -> analyse data on GPU -> post-process data on CPU -> upload results.



Accurate boundary data can help the estimated 570 milion farmes and growers worldwide boost their yields and lower their input costs.

The Solution

Split the process into multiple subprocesses executed on the appropriate resources to reduce runtime and cost of resource usage, by executing only the jobs requiring GPU on GPU resources and freeing up resources when not needed.

The Business impact

The impact of this investigation is two-fold. The SME could reduce its operational cost, reduce the time to market and expand the market to more areas faster.

Benefits

This collaboration was beneficial for both parties.

Sigma2:

✔ Learning about industrial AI workflows and how they can be optimized.

3B-the fibreglass company:

✔ Higher production capacity through adapting HPC for higher throughput.

✔ Lower operational cost by only using expensive resources when needed by the workflow.



Photo: James Baltz // unsplash

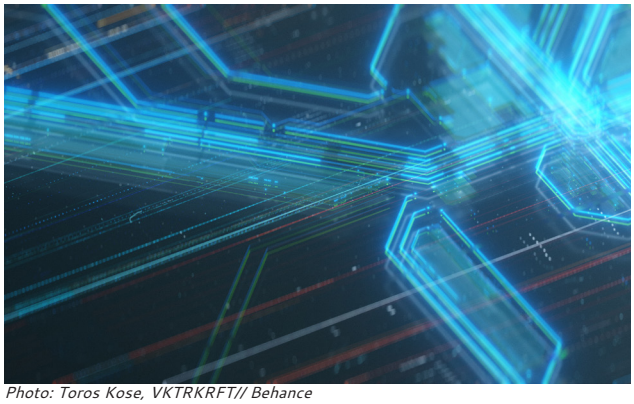


Photo: Toros Kose, VKTRKRFT// Behance

- Industrial AI
- Agri-tech
- Deep neural networks
- Agriculture

- Industry Sector: Agriculture

- Technology: AI

Contact:
Roger Kvam
Email: roger.kvam@sigma2.no

Simulation-Optimisation of a Patented Design with

Parallel Computing on TRUBA cluster

Design-Optimisation of a Disinfection System

TRUBA HPC Center, operating under TUBITAK ULAKBIM, coordinates NCC Turkey. 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. Our objective is to level up the industry, public sector, and academia in High-Performance Computing (HPC), High-Performance Data Analytics (HPDA), Artificial Intelligence (AI), and prepare for the developing EuroHPC ecosystem. The particular third party presenting this success story is TRUBA.

Industrial organisation involved

Design and Simulation Technologies (DSTECH) Inc. is an SME in Turkey supplying and services with customers in the solution of complex engineering problems encountered in different disciplines such as Environment, Energy, and Aerospace sciences. DSTECH develops and employs open-source platforms for the realistic simulation of complex flows on HPC clusters. Novel designs are developed to enhance the efficiency of engineering systems based on high-resolution numerical simulations.

Technical Challenge

Dynamic systems should be optimised using a simulation and optimisation framework to find out design parameters without using a blind-search approach. DSTECH intended to optimise geometrical specifications of a previously patented slot-baffle design for the enhancement of the efficiency of a disinfection contact system. Specifications of the slot-baffle design need to be optimised considering single and dual objectives of reducing short-circuiting effects and increasing mixing. Optimisation of the design requires a series of numerical simulations that may take a long simulation duration.

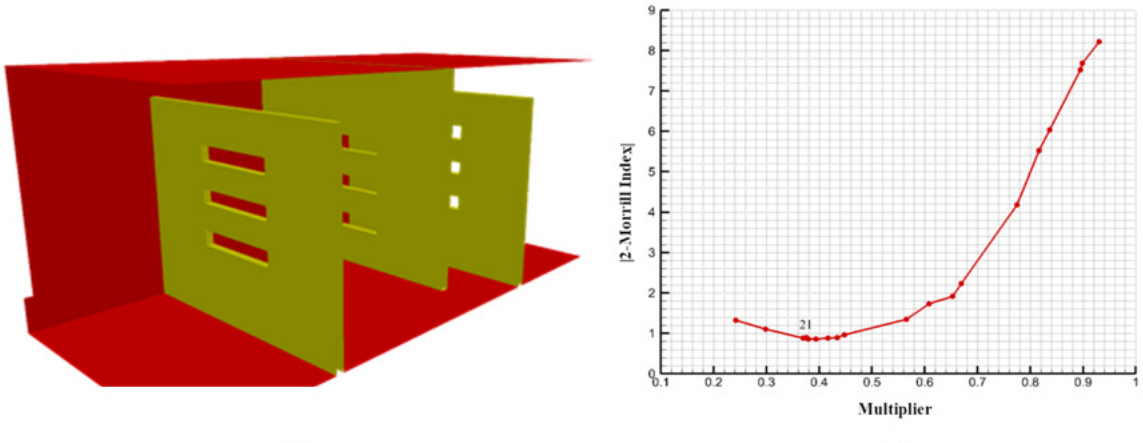


Figure 1. Optimisation of the slot dimensions for constant locations. (a) Three-dimensional view of the optimised baffle and (b) variation of the object function with the iteration

Solution

DSTECH has developed an open-source simulation and optimisation framework augmented with OpenFoam, Dakota, and Python, which can be freely used on HPC clusters without any restrictions such as license, commercialization, and parallelization. The present simulation and optimisation framework is employed to optimise the slot-baffle design in three stages without any convergence issues. Numerical simulations are performed with parallel computing strategies using an intense computational resource allocated on the TRUBA. Numerical simulation results in the present study reveal that the efficiency of the conventional design can be improved by 12.47% when the optimised design is implemented to the present tank.

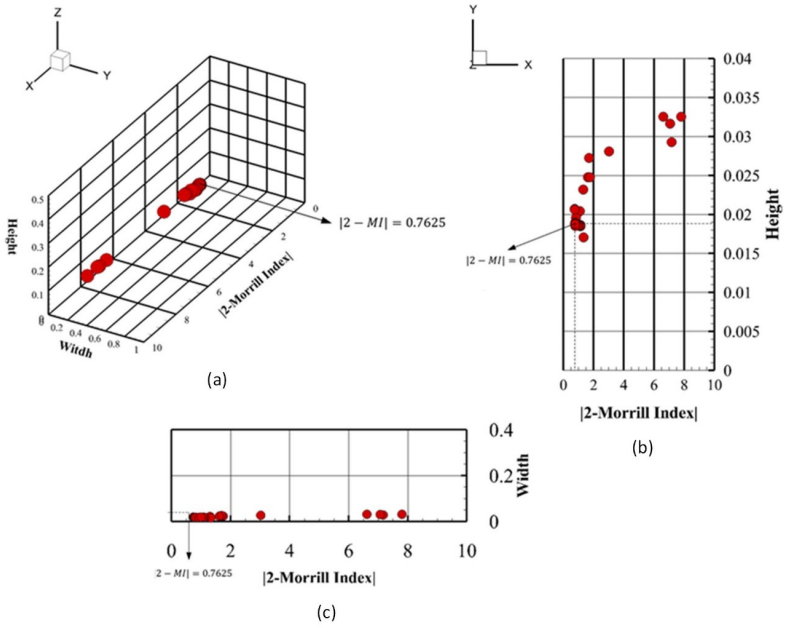


Figure 2. Variations of the parameters during optimisation. (a) Variations of width and height, variation of (b) height and (c) width.

Business Impact

Intense computational power is required for the accurate simulation of turbulent flows in large-scale engineering systems. Use of HPC enables us to perform a series of high-resolution numerical simulations for the optimisation of engineering designs in the project schedule. Moreover, there is no need to install expensive infrastructures and maintenance costs are not required with the applications of parallel computing on the HPC, which are important issues for SMEs.

The case study also started a new partnership between TRUBA and DSTECH. DSTECH has developed an open-source CFD code for the simulation of high-speed turbulent flows over aircrafts with the partnership of TAI (Turkish Aerospace Industries Inc.). In the Horizon 2020 project ECO-Qube, an open-source computational framework has been developed for the simulations of flow and thermal structures in data centres. A series of numerical simulations will be performed for different thermal scenarios to generate extensive data required in the development of an AI-augmented cooling system in ECO-Qube project. DSTECH is developing an AI-augmented disinfection system for the smart management of potable water treatment plants in another research and development project funded by Small and Medium Enterprises Development Organization (KOSGEB) in Turkey. Both TRUBA and DSTECH will evaluate collaboration around these topics in the scope of the EuroCC project.

Benefits

- ✓ An automatic-optimisation framework has been developed
- ✓ A patented design is optimised with parallel computing on the HPC cluster
- ✓ The optimised design has improved the efficiency of the system by 12.47%
- ✓ A new partnership is formed. DSTECH and TRUBA HPC Center will collaborate further around new topics

- CFD
- High-Performance Computing (HPC)
- OpenFoam
- Optimisation
- Dakota

- Industry Sector: Environment
- Technology: CFD Simulation

Contact:
Prof. Dr. Ender Demirel,
Co-founder of DSTECH
Email: edemirel@dstechno.net

High Efficiency Impeller Design
for Wastewater Pumps through High Fidelity CFD simulations

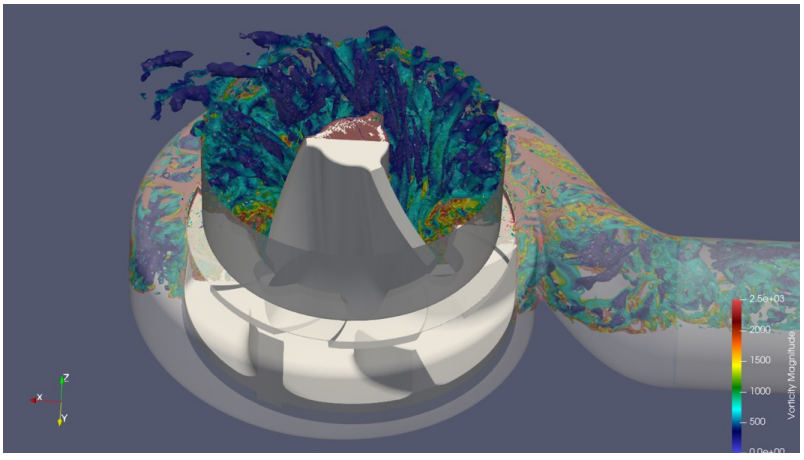
Turkish National e-Science e-Infrastructure (TRUBA), operating under Turkish Academic Network and Information Center (TUBITAK ULAKBIM) is the coordinator of NCC Turkey. Middle East Technical University (METU), Sabancı University (SU), and Istanbul Technical University National Center for High-Performance Computing (UHeM) are part 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. Our competencies include HPC, HPDA, AI, CUDA, Materials Science, Computational Fluid Dynamics (CFD), and several other fields. The particular third-party presenting this success story is UHeM.

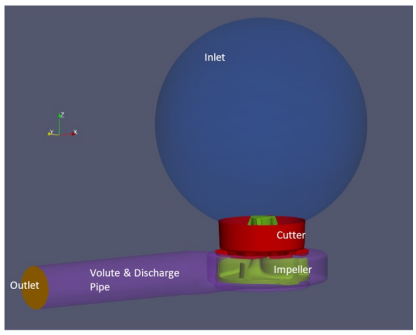
Organisations Involved

EYS, a company based in Turkey, strives to offer practical solutions to organics recycling problems by putting to use their knowledge and experience in manure management, dewatering and composting solutions. The company offers innovative and quality products to shape the future of environmental and agricultural sectors. From scrapers to pumps and mixers, separation to filtration, composting to field applications, it offers the widest range of products and models imaginable in the world of manure management.

Technical Challenge

As the cost of energy increases one has to spend it more efficiently to obtain sustainability. Also, net carbon zero energy production is another aspect of sustainability. On top of these two, recycling water is becoming more and more important and hence the water treatment facilities. EYS is already a manufacturer of rugged and robust submersible manure pumps which are field proven in terms of durability.
The problem to be tackled in this case study is to reduce the hydraulic design process time and increase the accuracy of our predictions with the help of high-fidelity Computational Fluid Dynamics simulations using national resources. These high-fidelity numerical simulations are conducted by using an open-source software, OpenFOAM.





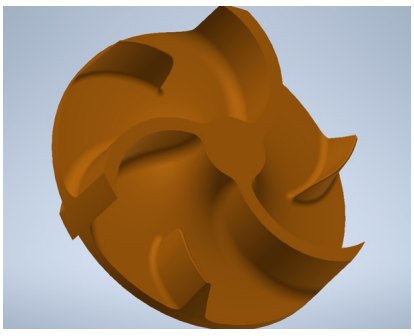


Figure-1: Full computational domain, the impeller and vortical structures inside the suction cone.

Solution

An important aspect of this project is to understand if the dynamic interaction between the rotating impeller blades and the stationary cutter channels can be captured with a reasonable amount of computational effort. Since this is a submersible pump, in order to improve the inflow characteristics a sphere is attached to the suction side. Stator and rotor regions are coupled using the cyclic arbitrary mesh interface method. The full computational domain, the impeller and instantaneous visualization of the vortical structures inside the suction cone can be seen at Figure 1.

Business Impact

· This result obtained via this case study shows that using HPC resources for high fidelity numerical simulations increases the accuracy of predictions for industrial applications and hence make the CFD calculations more reliable.

· An increase of %7 in efficiency is achieved with the new impeller as shown in Figure-2.

· This real-life result encourages us to use more HPC resources in the future and company has already applied for another national grant as a result of this study.

It is shown by the results (e.g., Figure 2) of this preliminary study that the efficiency can be improved faster by using HPC resources. The main reason is that by increasing the accuracy and reliability of the simulations, a significant improvement is achieved in decreasing the number of design iterations, shortening prototyping time and lowering the cost of model production and prototype testing.

EYS is hoping to expand their computational efforts by incorporating an advanced optimization tool in their computational design process and bringing it to a closed loop to achieve most efficient products.

The company increased their competitiveness in the market by directly increasing the efficiency of their product. EYS will be following the same path for different line of products in their wide product range.

Benefits

- EYS has gained extensive experience on conducting high fidelity simulations by using HPC resources.
- EYS manufactured wastewater pump prototypes, measured performance and validated simulations.
- EYS increased hydraulic efficiency of submersible wastewater pump using geometry modifications.

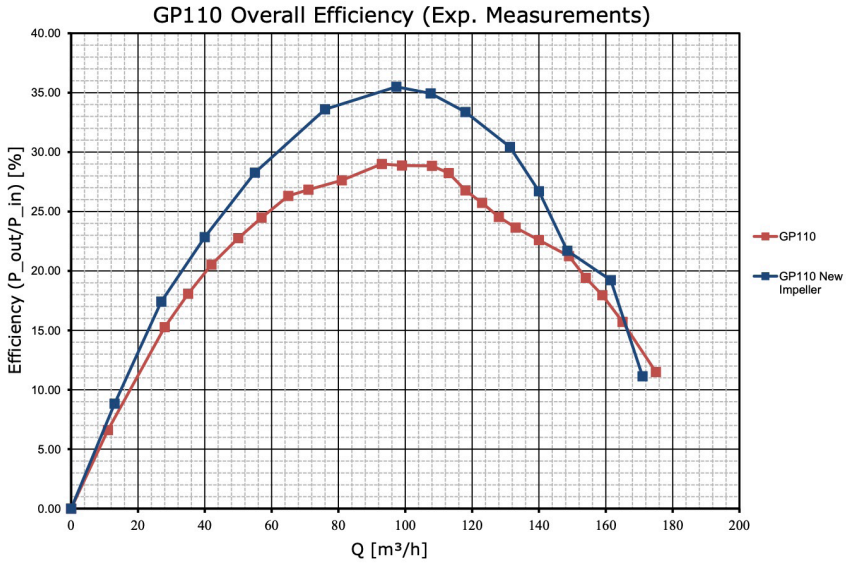


Figure-2 Efficiency comparison between two designs.

- CFD
- WasteWater management
- Energy Efficiency
- Submersible chopper pumps
- Rotodynamic machinery.

- Industry Sector: Agriculture, Energy, Mechanical Engineering, Environment, Manufacturing & Engineering

- Technology: HPC, High Fidelity Simulations, Advanced Turbulence Modeling

Contact:
Ayse G. Gungor,
Email: ayse.gungor@itu.edu.tr

The multiphysics experiments of the **Weather Research and Forecasting Model (WRF)** on precipitation patterns of Turkey

TÜBİTAK-TRUBA, one of the two capacious HPC centers authoritatively in Turkey, coordinates NCC Turkey. The Middle East Technical University (METU), Sabanci University (SU), and Istanbul Technical University National Center for High-Performance Computing (UHeM) are third-party partners and co-operate with TÜBİTAK-TRUBA in the decision-making process. In a project, TÜBİTAK-TRUBA provides technical support to researchers and monitor running jobs. Furthermore, TÜBİTAK-TRUBA regularly organises online meetings with the researchers to discuss drawbacks and progressions.

This project is conducted under the auspices of TÜBİTAK-TRUBA with the inspections mentioned above. On the researcher side, the METU researcher team performs the project and present the success story.

Industrial Organisations Involved

The collaboration between the METU research team and Eriktronik SME was based upon a European Union (EU) project proposal. After two submissions to a European call that is unable to continue due to limited funding, we were directed to the TÜBİTAK-TRUBA to maintain our works. The Eriktronik company and researchers at METU collaborated in this project, where an operational forecast of hydro-meteorological variables are aimed to be acquired at high accuracy. The heavy lifting in terms of Numerical Weather Prediction (NWP) model setup and run has been undertaken by the METU team mainly.

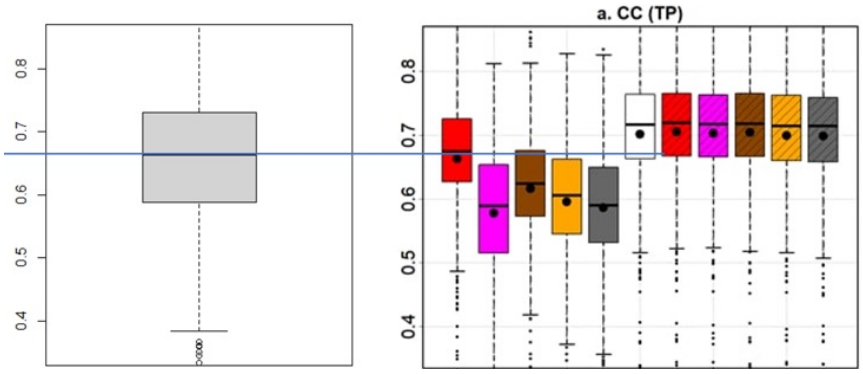


Figure 1. Previously available model performances (solid gray box on the right with a median 0.58) are improved to the level of European state of the art model accuracy levels (left gray box with median 0.66). These are very encouraging results compared with existing state of the art models.

Technical Challenge

The reliable weather research and forecast heavily rely on the choice of the model physics options. There are multiple physics options in the WRF model that should be tested in a combinational manner to improve the forecasts. The weather events are chaotic by nature, and any physic option may be time and domain-specific thus have some uncertainties. Therefore, the physics combinations doing a good job are hired as an ensemble to deal with these uncertainties. A sensitivity test can reveal the ensemble members of the multiphysics combinations. On the other hand, similar sensitivity tests are conducted for climate modelling but with relatively

long runs (~year-long). To our knowledge, there are no studies comprehensively dealing with such sensitivity tests of the multiphysics options over the Turkey domain.
The WRF model is an NWP system designed to solve physics equations in the 3D and 4D architectures. It supports parallel computing in which end users need to utilize HPCs in weather forecasts and climate prediction. To a certain extent, high resolution is needed to achieve improved and creditable results in running WRF. The higher the resolution is tuned, the higher the computational costs are required.



Weather in Giresun, Turkey. Photo: Alp Duran // unsplash

Solution

The project's sensitivity tests were completed over the Turkey domain for 2020, with a 60-combination of model physics in 4-km resolution. The combination number, resolution, and simulation time are rather comprehensive for such sensitivity tests. We have valuable information now about which multiphysics ensemble responds favorably to the Turkey precipitation characteristics. Our findings pave the way for improving weather forecasts and climate predictions over Turkey. This study could not be performed without utilizing TÜBİTAK-TRUBA HPC facilities.

Business Impact

This project is to have several impacts on the private and public sectors. On the private sector side, high accuracy operational forecasts that are tailored to Turkey via specific parametrizations are obtained, where such a forecast was available before. Having such a high accuracy forecast is very valuable for many industries, in particular those are interested in prediction of renewable energy (wind, solar, and hydropower). The private sector companies dealing with renewable energy have an opportunity to access these data.

On the public sector side, institutions that are closely working on the prediction of hydrometeorological variables such as precipitation, temperature, and runoff can utilize the results of this study and improve their short term operational forecasts and climate predictions. In particular, they may improve early flood warning systems based on the new forecasts. The Ministry of Environment, Urbanization and Climate Change can act upon the climate change impacts on urban and rural areas of Turkey when climate predictions are completed based on the best configuration obtained with this project. When the predictions are completed, The General Directorate of Water Management can update their basin management plans. The Ministry of Agricultural and Forestry can monitor drought and flood events more precisely when the seasonal forecasts are performed, depending upon the multiphysics ensemble obtained here.

Benefits

- ✔ The team has gained experience for the first time in the HPC domain.
- ✔ The team has been encouraged to apply to the EuroHPC projects in the seasonal forecast and climate prediction areas through gained experience with this project.
- ✔ The team improved their insight into driving mechanisms of precipitation over Turkey.

- Hydrometeorological forecast
- Precipitation
- Temperature and wind
- Renewable energy
- Numerical weather prediction

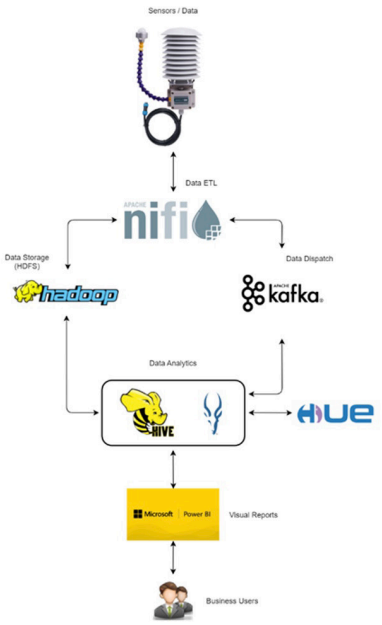
- Industry Sector: Earth Science, Energy, Environment/ Climate/Weather, Public Services/Civil Protection
- Technology: HPC

Contact:
Email: tuyilmaz@metu.edu.tr
Email: iyucel@metu.edu.tr

Environmental Data
from Transferring and Analytics to Decision Making

NCC-Bulgaria was founded by Institute of Information and Communication Technologies at the Bulgarian Academy of Sciences, Sofia University "St. Kliment Ohridski" and the University of National and World Economy.

NCC-Bulgaria is focused on:
- Creating a roadmap for successful work in the field of high-performance computing, big data analysis and artificial intelligence.
- Analyzing the existing competencies and facilitating the use of HPC/HPDA/AI in Bulgaria
- Raising awareness and promoting HPC/HPDA/AI use in companies and the public sector.



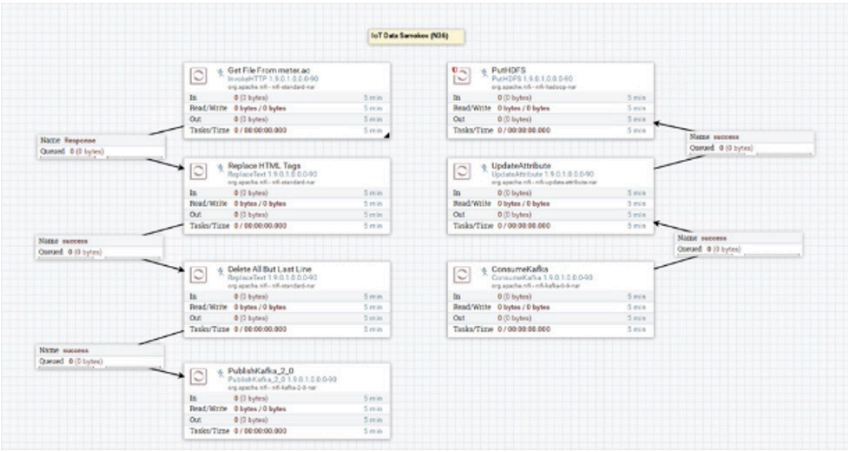
Referential IoT Architecture for transferring and computational analysis of environmental (meteorological) data: *Components and their communication links, within the realized referential IoT architecture.*

Industrial organisation involved

NCC Bulgaria works mainly with agricultural companies, using sensors for soil and meteorological factors monitoring. Also companies working in the tourist sector like tour operators, event organizers are involved.

Technical Challenge

Reliable, fast, and accurate transferring of environmental (meteorological) data, followed by its computational analysis (descriptive and predictive), is a considerable technical challenge.
On one hand, it is difficult to methodologically organise the large volumes of data being transferred, the enormous diversity of environmental sensors used, and the proliferation of the transmission protocols. On the other hand, consistent assurance and maintenance of collected data quality while selecting the right tooling and algorithms for data analytics and presentation, requires specialised knowledge and technological experience.

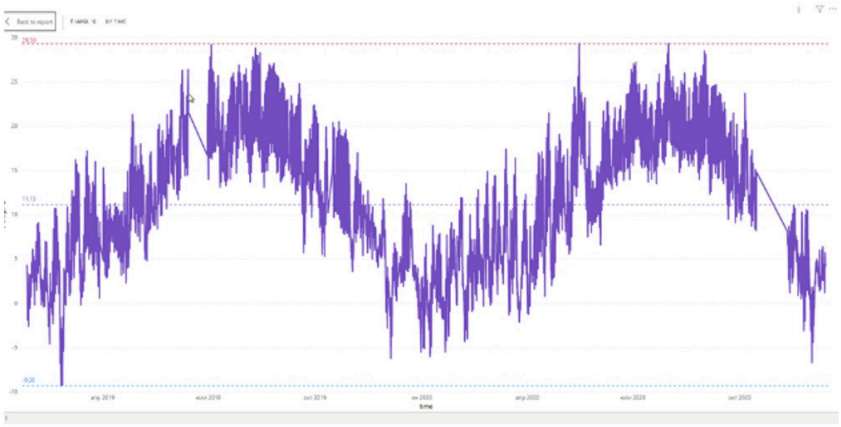


Environmental Data ETL:
NiFi flow created and used for events extraction, transformation and load.

Solution

To solve those challenges, a referential IoT architecture is proposed, designed and implemented with the following components:

- Data extraction and cleaning module NiFi,
- Data distribution module Kafka,
- Data storage module Hadoop HDFS,
- Data analytics modules Hive, Impala and Hue,
- Data presentation module PowerBI.



Environmental Data Report:
PowerBI analytical events report, created by using ODBC connection with the cluster.

Building such an architecture enables flexibility and speed in data loading and clean up from a multitude of meteorological sensors, different networks and protocols, reliable ingestion, dispatch, consumption and storage of the extracted environmental physical values, powerful data analysis engines for both streaming and batch processing, for analysis and design of visual presentation of the results to be easily understood by all different types of users.

Business Impact

At the moment, there is a lack of real-time analytics and predictive instruments for forecasting bad weather conditions, alerting for emergency situations and natural disasters.
Real-time analytics of meteorological data contribute to lowering risks, saving human lives and health, decreasing business losses, preventing unexpected and undesirable consequences.
In addition, collecting, transforming, and analysing environmental big data sets support creating conceptual models and developing practical weather models for local and potentially global effects of the global warming effects. This approach is used to improve the planning of actions for multiple scientific and business-related efforts to control temperatures rise and reduce risks.
The developed solution support:

- Tourist companies, organizing tours and events
- Companies in the hotel industry sector - climate analysis for hotel chain expansion
- Companies in the agricultural sector – for business processes optimization (providing optimal conditions for growing crops), improving product quality, minimizing risks, etc.

Benefits

- ✓ Flexible and reliable, multi-protocol and multi-network data extraction
- ✓ Cost effective Open-Source components
- ✓ Distributed high availability file system and event dispatching system
- ✓ Powerful data analytics including streaming and batch processing
- ✓ Sophisticated visualisation

- IoT (Internet of Things)
- Big Data and Environmental Data
- Meteorology
- Analytics
- Apache Ecosystem
- Reference Architecture

- Industry Sector: Agriculture, Environment/climate/weather, IoT, IT/ HPDA systems, services & software providers, Tourism

- Technology: IoT, Apache Ecosystem and HPC/HPDA Integration

Contact:
Dr. Yavor Christov,
University of National and World Economy,
Email: ychristov@unwe.bg

Prof. Kamelia Stefanova,
University of National and World Economy team,
Email: kstefanova@unwe.bg

Section 3

Natural Sciences and Aeronautics

Using CFD to optimize aircraft design with Airinnova– 62

Optimization of the CHAMPION Software – 64

EnginZyme – 66

Northvolt on EuroHPC system – 68

Estonian COVID-19 Data Portal and the KoroGeno-EST study - 70

Combination of Computation and Experimental Science- 72

Video Super-Resolution - 74

Graphene-Enhanced Polymer Composite Production - 76



Photo: Pawel Czerwinski // Unsplash



Photo: Bao Menglong // Unsplash

Using CFD to optimize aircraft design with Airinnova

ENCCS (EuroCC National Competence Centre Sweden). ENCCS provides high-performance computing training and support for industry, academia and public administration for free.
https://enccs.se

ENCCS
EuroCC National Competence Centre Sweden

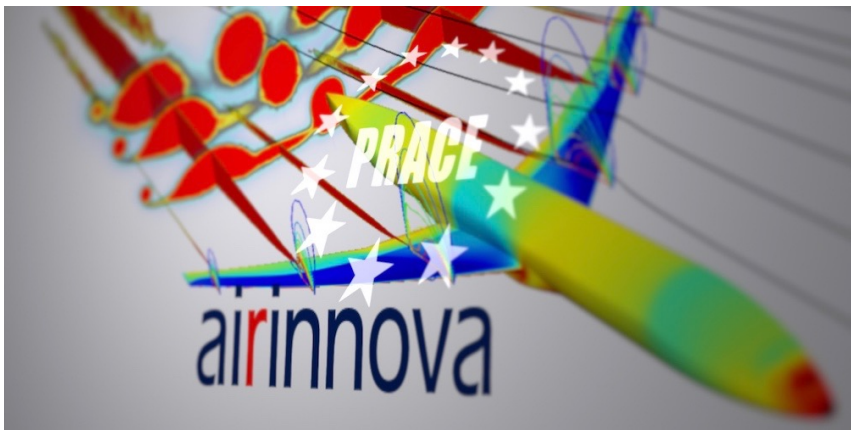
Organisations Involved

Airinnova was founded in 2015 by a team of experts from KTH, the company is focusing on providing advanced computational technology for cutting edge aircraft preliminary design, computational aerodynamics, and multi-disciplinary optimization. It is currently actively involved in European research projects, engineering training and software development for aerospace applications
http://airinnova.se

airinnova
Engineering Solutions and Innovations

Technical Challenge

Aerodynamics needs computing power to efficiently simulate CFD (Computational Fluid Dynamics) for aircraft design. Airinnova will focus on a very high altitude and long endurance solar-electronic aircraft, using the designed models by Conceptual Design Corporate (CRC), simulate the given models, and optimize to find out the best design regarding lift-to-drag ratio.



Solution

During this access, AIRINNOVA will be able to work with:

- **Automation process development**: Test the automation process by starting from a conceptual aircraft design model from a Commercial Software RDS into a working mesh.
- **Benchmark**: performance analysis for the desired model using the proper transition model, compare the predictions with the experiments.
- **Port**: deploy and run on a PRACE Tier1 system and prepare for Tier-0 system.

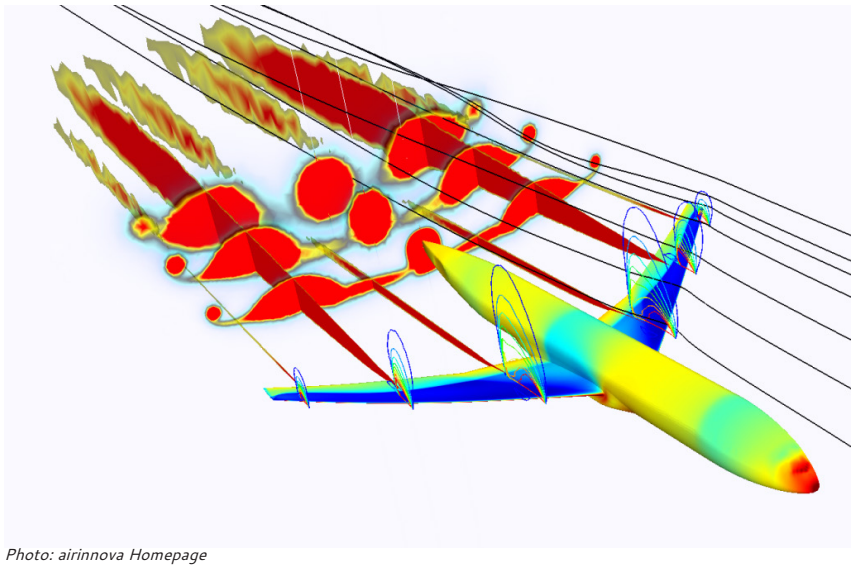


Photo: airinnova Homepage

Business Impact

Computing power is crucial for Airinnova and their activity and their business model since they need to simulate multiple aircraft parts for clients, multiple times in order to come up with the most efficient design.

Benefits

- ✔ Efficient CFD simulations using multiple CPUs
- ✔ Multiple tests runs of aircraft designs
- ✔ Reduction of costs due to gain of time

- Computanional fluid dynamics (CFD)
- Aircraft design
- Aerodynamics
- Perfomance

- Industry Sector: Aeronautics

- Technology: HPC

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

Optimization of the CHAMPION Software

for a FF4EuroHPC experiment

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

Industrial organisations Involved

Compular is a Swedish SME, based in Gothenburg, that develops cutting-edge analysis tools for molecular dynamics simulations.
https://compulartech.com

Compular is part of the FF4EuroHPC experiment consortium composed of Stiftelsen Chalmers Industriteknik, Enerpoly and ENCCS.

The Challenge

Materials science is often using mainly experimental methods in a lab to investigate new material properties. Lab work comes with a lot of disadvantages. It takes a long time to reach optimal compositions, there is a high cost in salaries, equipment and materials, and processes are often uncertain and unreliable. Digitalisation in this field is needed to design and assess new materials efficiently as well as stardardise logistics.

CHAMPION is a molecular dynamics analysis software for finding what moves together in atomistic trajectories and analysing the detected structures statistically in a hierarchical and dynamical manner.

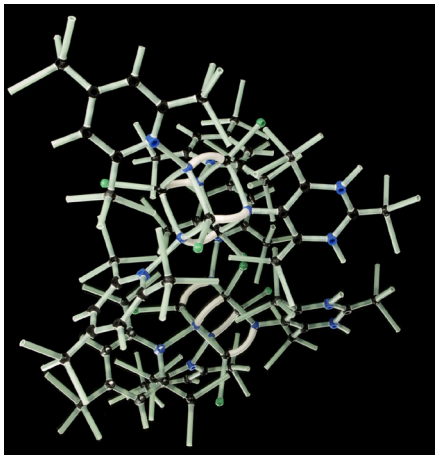


Photo: Photoholgic // Unsplash

The Solution

With the successful application of the experiment with Compular to the FF4EuroHPC program and access to EuroHPC JU supercomputer Karolina, Compular will work on their new and innovative technology that automatically detects, identifies, and analyses the emergent structures and dynamics within materials which give rise to their technologically relevant properties.
The consortium will leverage HPC resources and state-of-the-art simulation and analysis software developed by Compular to create better performing and sustainable battery solutions. This is an area of intense research, with high societal impact and we are excited to be part of this project.



The Business impact

Compular was able to optimize the CHAMPION software, ensuring that it is robust and efficient for usage in a Software-as-a-Service framework. The experiment is moving forward according to the initial plans. For Compular, the FF4EuroHPC and EuroHPC-JU development access programs ensured timely access to world-class computational resources and an accelerated development process for software relevant to, among others, battery design. Such allocations show that cutting-edge HPC infrastructure is critical for R&D-based SMEs and startups. The awarded project will fund a computational experiment in partnership with Stiftelsen Chalmers Industriteknik, Enerpoly, and ENCCS.

Benefits

Currently the whole simulation and analysis state is kept in memory at all times, making memory requirements a severe limitation on the usefulness of the software. With this access, Compular aims to:

- Integrate relevant algorithms with an existing Sqlite3 persistence layer to drastically reduce the memory requirements
- Re-design algorithms to optimise for parallelizability and run-time performance instead of memory usage, and
- Parallelise the algorithms that limit runtime performance the most after the memory footprint has been reduced

- Molecular simulation
- Batteries
- Materials
- Digitalisation
- Innovation

- Industry Sector: Chemicals, Raw materials

- Technology: HPC

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

EnginZyme to increase the stability of enzymes by combining experimental and computational techniques thanks to PRACE DECI17

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

ENCCS
EuroCC National Competence Centre Sweden

Industrial organisation involved

EnginZyme is a start-up company that is a spin-off from Swedish academia, and very much values collaborations between industry and academia that facilitates high-impact scientific discovery.

EnginZyme https://enginzyme.com/

Technical Challenge

EnginZyme uses cell-free synthetic biology in order to harness the power of nature's catalysts, enzymes. This gives a unique insight to address many challenges facing the chemical industry in the 21st century, both by accelerating the industry's transition to sustainability and by innovating solutions for other industries through novel chemical products. EnginZyme has been allocated **1.2 million core-hours** after the **PRACE DECI17** call for a molecular dynamics study in order to evaluate how the wild-type proteins behave in solution, and how different mutations influence the structure and properties of the enzyme



Photo: ANIRUDH // Unsplash

Solution

The overall goal of the project is to increase the stability of enzymes in EnginZyme's cascades for sustainable production of chemical products, by combining experimental and computational techniques. The typical current use of HPC is to run molecular dynamics, primarily utilizing the **GROMACS** and **Amber** software, but also using the **Rosetta** modeling suite.



Business Impact

This allocation would give EnginZyme the power to further evolve their technology and strengthen their mission of offering a more sustainable alternative to the traditional chemical processing industry by designing more stable biocatalysts.

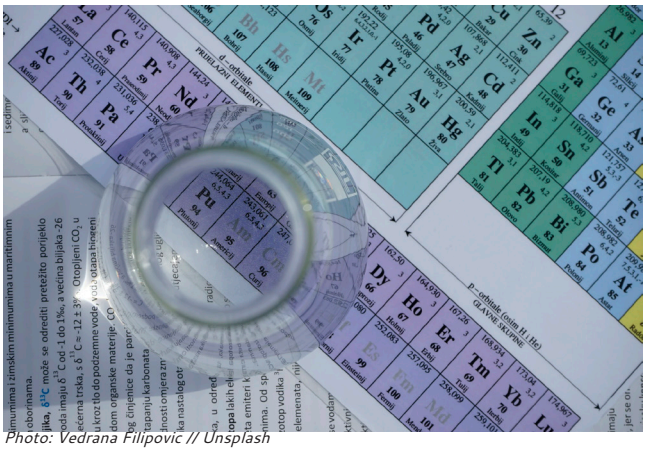


Photo: Vedrana Filipovic // Unsplash

Benefits

- ✓ Create versatile biocatalysts in standard chemical industry fixed-bed reactors that can be used in continuous operation instead of in batches like with bioreactors

- Molecular Dynamics
- Gromacs
- Enzymes
- Amber
- Proteins

- Industry Sector: Biotechnology/Bioinformatics, Chemicals

- Technology: HPC

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



Photo: D koi // Unsplash

Northvolt to be the first industrial actor to access a EuroHPC system

ENCCS (EuroCC National Competence Centre Sweden). ENCCS provides high-performance computing training and support for industry, academia and public administration for free.
https://enccs.se

Organisations Involved

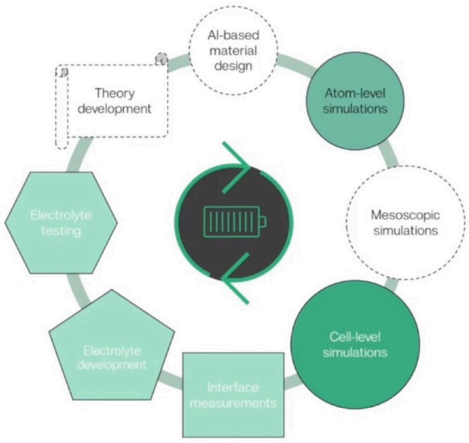
Northvolt AB accelerates the transition to a decarbonized future by supplying sustainable, high-quality battery cells and systems.
https://northvolt.com/

Technical Challenge

Northvolt, a world leader in the field of battery production and development, has been awarded development access to EuroHPC JU petascale supercomputer VEGA hosted at The Institute of Information Science (IZUM) in Maribor, Slovenia.

Northvolt will investigate the use of classical and reactive molecular dynamics and quantum chemical simulations to devise bottom-up design strategies for improved batteries. This is an area of intense research, with high societal impact: electric mobility has the potential to significantly reduce CO2 emissions of the road transportation sector, which currently accounts for about one-fifth of total CO2 emissions in Europe.





Solution

The nearly 2 million core hours awarded were used to perform larger and more realistic simulations of the electrochemistry relevant for battery development using multiple software such as **LAMMPS** for reactive force field simulations, **GROMACS** for dynamical molecular simulations, **psi4** for sapt simulations, and **ADF** for electronic structure calculations.

Business Impact

The atomistic level simulations which were done during this access are an important part of the R&D process in Northvolt (Image 1). They provide important detailed data on the molecular structure and dynamics of battery materials. Together with mesoscopic and cell-level simulations they are a crucial link in designing and producing batteries faster and more efficiently.

Benefits

- With supercomputing access atomic-level simulations are faster, making Northvolt's workflows faster and more efficient
- Northvolt's vision to a more sustainable and decarbonized future adds to the significant work on climate change along with other organisations and companies

- Quantum chemistry
- Simulations
- HPC
- Battery development
- Batteries
- Sustainability

- Industry Sector: Chemicals

- Technology: HPC

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

The Estonian COVID-19 Data Portal and the KoroGeno-EST study

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

Partners Involved

The Estonian Elixir node is organised by the University of Tartu. Other partners are Tallinn University of Technology and National Institute for Physics and Biophysics. KoroGeno-EST, funded by the European Union COVID-19 Pandemic Response, is a collaboration between the University of Tartu, Synlab Eesti OÜ, and the Estonian Health Board.

The Challenge

The aim of the KoroGeno-EST study is to sequence (determine the composition of the genome) and analyze the complete genomes of SARS-CoV-2 that have caused and are causing infections in Estonia and to perform a molecular epidemiological analysis on them. Like living organisms in the wild, viruses can be divided into different groups based on genetic similarity. Grouping with a certain degree of accuracy must be used to obtain a visually perceptible picture of the presence of variants of concern (VOCs). Working with this problem involves a very large-scale data analysis, that can be conducted only by applying High-Performance Computing resources and relevant methods.

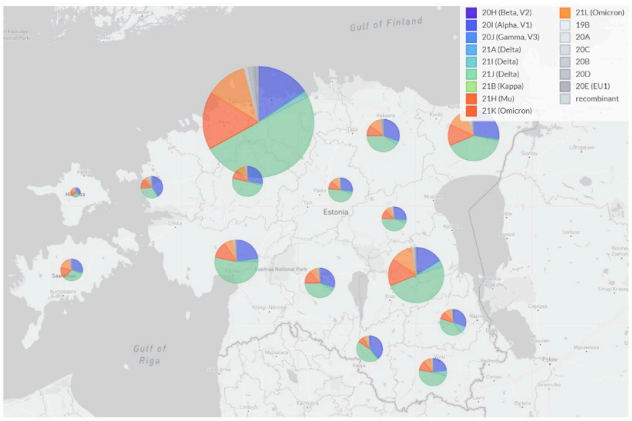


Figure 1. Geographic distribution of SARS-CoV-2 clades in Estonian counties.

Graphics: https://auspice.biit.cs.ut.ee/ncov/est

The Solution

Genome analysis of SARS-CoV-2 variants was performed on the University of Tartu HPC centre's Galaxy instance. The NCC Estonia supported the study by providing necessary infrastructure together with in-depth support, code optimization and technical consultancy for the project. The main results of the KoroGeno-EST study are available on the Estonian COVID-19 Data portal. The portal provides an overview of the proportions of SARS-CoV-2 clades and variants of concern that can be observed by the visitor over time. Visitors can select the time period for which the data is displayed and choose whether to display travel-related or domestic events.

The Business impact

In addition to delivering the results of KoroGeno-EST study, the Estonian COVID-19 Data Portal also provides information, guidelines, tools, and services to support researchers to utilise Estonian and European infrastructures for data sharing. The portal is a national node of the European COVID-19 Data Portal and it is operated by the ELIXIR Estonia. The NCC Estonia is here continuously helping with workflow optimization and data handling, also assisting with process automatization.

Benefits

- ✔ Data collected during KoroGeno-EST study will help to bring together relevant datasets for sharing and analysis in an effort to accelerate coronavirus research.
- ✔ The Estonian COVID-19 data portal enables researchers to upload, access and analyse COVID-19 related reference data and specialist datasets as part of the wider European COVID-19 Data Platform.

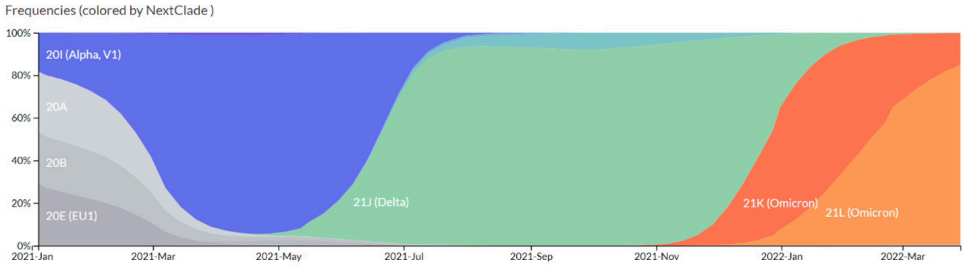


Figure 2. Proportions of SARS-CoV-2 clades by sampling date.

- Geonomics
- COVID-19 Data Portal
- Open data
- Data sharing
- HPC

- Industry Sector: Biotechnology/Bioinformatics

- Technology: HPC, HPDA

Contact:
Ülar Allas
Email: ylar.allas@ut.ee

Combination of Computation and Experimental Science

to reveal snapshots of 2D polymers

The Portuguese National Competence Center (NCC) in EuroCC is coordinated by the Portuguese national funding agency for science, research and technology (FCT) and integrates several entities that through the NCC provide their expertise and capabilities to support the HPC adoption and usage in the country.

EUROCC PORTUGAL

Scientific Partners Involved

This project brought together experts from Belgium, Spain and Portugal. A team from the Belgian university KU Leuven, POLYMAT, integrated in the University of the Basque Country, in Spain, was supported by the computer models of CICECO, a unit of University of Aveiro and also received support from University of Coimbra – both in Portugal – in order to explore the mechanisms that lead to the formation of 2D polymers.

POLYMAT Basque Center for Macromolecular Design and Engineering

universidade de aveiro

1290 UNIVERSIDADE D COIMBRA

Technical Challenge

Polymers have acquired an important role in human life. There are uncountable products that depend on them, such as plastic lenses in glasses or clothing made from synthetic fibres and they can even be used in advanced electronic applications (directly linked to 2D polymers). However, the process of how the polymers are exactly formed contains some gaps and could benefit from clarification. Therefore, a complete understanding of their fabrication is useful to boost the size and structural quality of these materials.



Photo: POLYMAT Homepage

Solution

Through a scanning tunnelling microscopy, an advanced type of non-optical microscopy, researchers were able to follow in real-time the process of birth and growth of 2D polymers on a solid support immersed in a reactive solution. They used computer models and the Navigator Cluster, operating at University of Coimbra, to perform massive quantum calculations – nearly 16k atoms had to be performed – to get the validation and support of the findings.

Business Impact

Using complex computational models developed at the University of Aveiro, they registered the coexistence of several growth mechanisms that, when controlled, allow the formation of polymeric sheets of great quality and dimensions. The findings ended up being reported in the journal *Nature* in 2022: G Zhan et al. *Nature* 2022, DOI: 10.1038/s41586-022-04409-6.

Benefits

✓ This research brought new lights on the formation and growth of 2D polymers and the potential to improve their quality and size and, therefore, the quality of the materials resulted from them.

Fig. 1: Disorder-to-order transition.

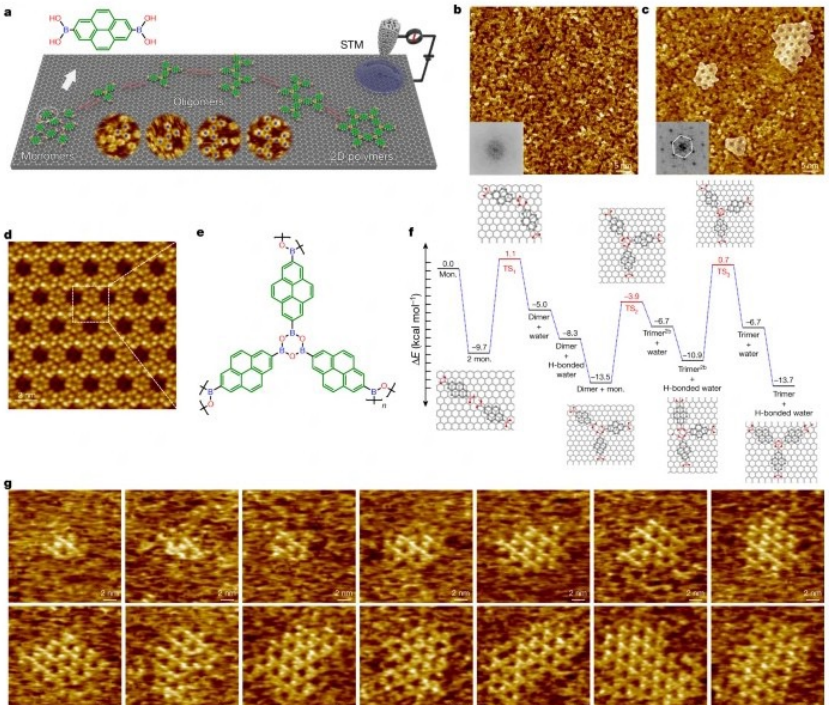


Fig.1: Disorder-to-order transition.

- Materials
- CICECO
- Polymers
- POLYMAT
- Electronics

- Industry Sector: Materials science
- Technology: HPC, HPDA

Contact:
Catarina Ferreira Gonçalves
Email: catarinafg@lip.pt



Photo: Peter Stumpf// Unsplash

Video Super-Resolution

EuroCC Italy is the Italian competence center part of the EuroHPC Competence Centers Network of the EuroCC Project. The long-term objectives of EuroCC Italy are to structure and strengthen the collaboration between research groups and institutes, HPC data centers and companies, to support SMEs in innovation projects and to create awareness about the productive advantages possible through the use of the different solutions in the HPC / HPDA / AI field. EuroCC Italy is partnered by CINECA; ABD, Bi-REX, Leonardo Company and Dompè.

Organisations Involved

KEBULA is a Data Lab born as a spin-off company of the University of Salerno, qualified as innovative start-up. The mission of KEBULA is unlocking the full potential of data to ignite tomorrow's business. KEBULA helps its clients to enable their data journey and enhancing their data engineering, analytics and AI capabilities. KEBULA is a founding member of MPAI (Moving Pictures and Data Coding with AI), which sets new standards for encoding multimedia data using Artificial Intelligence (AI).

Technical Challenge

Video streaming is now ubiquitous and fundamental for numerous business applications, including web streaming, remote sensing and satellite monitoring. Current efforts are being made into designing new video codecs that can reduce the bitrate required for storage and transmission of high-quality videos. The use of AI is a promising direction in this regard. In particular, super-resolution models can enable transmission and storage of low-resolution videos, which can then be upscaled to the desired resolution without a significant loss of quality intrinsic to traditional upscaling methods.



Example of a 4K frame produced by one of the models

Solution

Different models were evaluated using the standard BD-PSNR and BD-RATE. All models were compared with the HEVC codec, tuned for streaming settings. The largest models recorded the best super-resolution quality: IconVSR obtained 1.127 BD-PSNR and -28.511% BD-RATE when computed on the 3 RGB channels. Thus, on average, upscaling 1080p videos to 4K resolution using IconVSR would produce a 1.127 higher PSNR at the same bitrate. EDVR also obtained similarly good results, with 1.101 BD-PSNR and -27.759% BD-RATE. However, these models could only run at 0.79 and 0.73 FPS on Marconi100 respectively, making them unsuitable for real-time applications. On the contrary, smaller models proved promising for real-time applications. For example, the base EVSR model can run at 82.03 FPS on Marconi100.

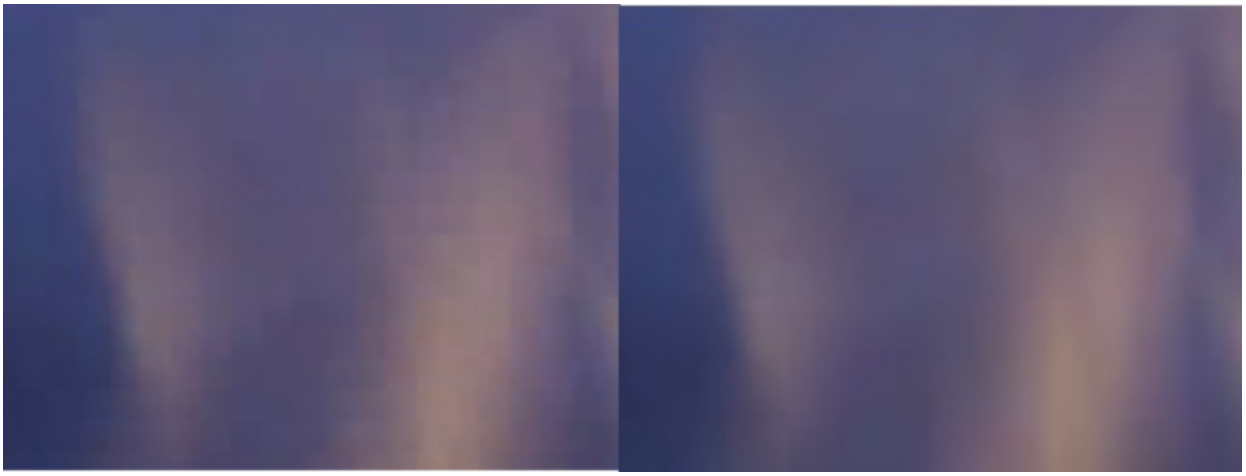
Business Impact

The project has had several positive effects for KEBULA. First, KEBULA gained experience on performing multi-node, multi-GPU training.

Second, KEBULA acquired knowledge and experience on the usage of video super-resolution models. Super-resolution has been gaining traction in recent years since it can be applied to a variety of applications beyond video compression. As more powerful GPUs will be available and existing models improve, more possibilities for real-world applications of super-resolution will open. Since KEBULA aims to cover a wide range of data analysis and transformation pipelines, experience on this kind of models will be more and more relevant to open new business opportunities. Third, since the PoC was born in the framework of our collaboration with MPAI regarding the design of new, better, video codecs enhanced by AI algorithms, the knowledge acquired by testing and evaluating all the described models will concretely help KEBULA to determine possible paths to reach this goal.

Benefits

- ✔ The models developed paves KEBULA's current and future work with MPAI for video compression algorithms.
- ✔ The super-resolution algorithms implemented will be used by KEBULA in the aerospace domain, where applications have stricter hardware and power constraints.



A detail of the previous frame, with different models. Left: HEVC 4K. Right: HEVC HD+ super-resolution 4K

- Video compression
- Video resolution
- Artificial Intelligence

- Industry Sector: Aeronautics, Media

- Technology: HPC, AI

Contact:
Gioele Ciaparrone
+393357191849
Email: gioele.ciaparrone@kebula.it

Improving the Efficiency of the Graphene-Enhanced Polymer Composite Production via Classical Molecular Dynamics

The leader of the NCC is TUBITAK-TRUBA, who is one of the two large HPC centers of Turkey. In addition to TRUBA, two academic partners Middle East Technical University (METU) and Sabanci University (SU) act as third parties. While METU is a public university based in Ankara, SU is a privately-funded university in Istanbul. Our 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. The particular third-party presenting this success story is the Middle East Technical University.

Organisations Involved

This success story is a result of a case study conducted by METU in collaboration with an SME (Nanografi) located in METU's Technopark. METU is a publicly-funded university with approximately 30000 students. It is considered both a teaching and a research university, recognized as such by the government. Founded in 2011, Nanografi (https://nanografi.com/) is a company that specializes in the production of nanomaterials including graphene-polymer composites, CVD-grown graphene, graphene oxide and fullerenes. Nanografi has a strong R&D department and has references from large companies such as Bosch, Intel and Xerox.

The Challenge

One of the most popular products that Nanografi specializes in is graphene, and graphene oxide, enhanced polymer composites. Their protocol is based on a "synthesize-test-improve" with a lot of the human resources spent in the trial-and-error stages to develop the composite with the best mechanical properties.

Since there are a large number of parameters that potentially go into this problem such as the type of the polymer, the sheet number of graphene, and pretreatment conditions, these experiments are both costly and time-consuming. The company would like to reduce the cost and duration of these experiments by means of pre-screening the parameter space via materials simulations. Our academic team from METU (Dr. Erol Yildirim and Dr. Hande Toffoli) in collaboration with our TRUBA specialist (Dr. Sevil Sarıkurt Malcıoğlu) have helped implement an alternative computational route for Nanografi to first model the composites using classical molecular dynamics (MD) to identify strong candidates before the actual synthesis stage.

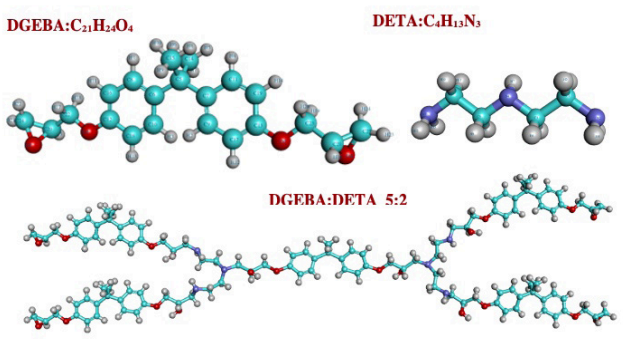


Figure 1: Molecular structure of DGEBA, DETA, and representative crosslinked epoxy chain.

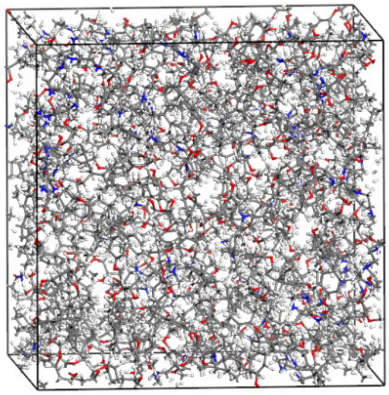


Figure 2: Simulation cell for the epoxy resin model with 90% crosslinking ratio by using 40 DETA and 100 DGEBA monomers.

The Solution

Our collaboration provided the following solutions:

- A company employee was trained and learned how to use HPC resources on the TRUBA computers. She also learned how to use the open-source code LAMMPS (https://www.lammps.org/) for conducting the calculations.
- Our academic expert, Dr. Erol Yıldırım set up a protocol for preparing and running different composites and trained the company employee on how to use this.
- The company had no prior experience in using HPC resources. For the first time in their history, they were introduced to these services.

The Business Impact

The business impact for this collaboration is not expected to be immediate but instead rather long-term, following a gradual transition period. At the moment, the experimental protocol utilized by Nanografi is arduous and time-consuming. There is a vast parameter space related to the synthesis stage to be tried. Our team has proposed an alternative mechanism where the initial elimination of the composites to be synthesized is to be simulated using classical molecular dynamics.

The results of the simulations will then be analyzed and strong composite candidates will be identified. The parameters that can be implemented in the calculations are the type of the polymer matrix, linker type, chain length, properties of the graphene-based enhancement material, and effects of temperature. The synthesis will then process within the much-reduced domain of these candidates. This will, in the long run, benefit the company by means of reducing both human resource needs and synthesis costs.

In addition, with this collaboration, the company was introduced, for the first time, to HPC services. Hopefully, this collaboration will motivate investment in HPC-literate personnel and their training in the long run.

Finally, our academic team, our TRUBA expert, and Nanografi (along with an end-user company) have submitted a proposal to the FF4EuroHPC call. We were able to work efficiently and in harmony. The areas of expertise of the various team members complemented one another rather nicely. This case study can therefore be seen as the beginning of a lifelong collaboration, which will surely benefit all parties involved.

Benefits

This collaboration was mutually very beneficial for both parties.

METU/TRUBA...

- ✔ has now gained access to R&D input from an industrial partner with no prior experience with HPC or infrastructure.
- ✔ could not have formed this sort of a partnership through which they were able to participate in the FF4EuroHPC call.

Nanografi...

- ✔ were introduced to computational materials science and HPC services for the first time,
- ✔ were able to tap into computational work as an alternative or support mechanism to the traditional and rather time-consuming experimental methods,
- ✔ could not have formed this sort of a partnership through which they were able to participate in the FF4EuroHPC call.

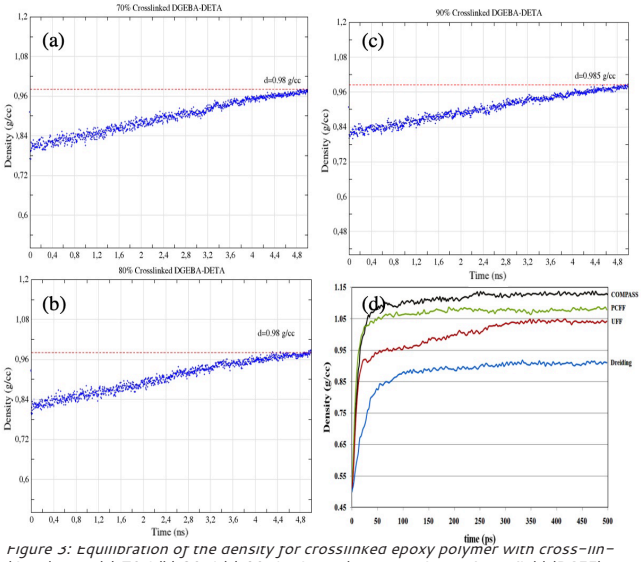


Figure 3: Equilibration of the density for crosslinked epoxy polymer with cross-linking degree (a) 70% (b) 80 % (c) 90 % using polymer consistent force field (PCFF) through 5 ns NPT dynamics. Each system cooled from 600 K to 298.5K (d) Equilibration of the density using different force fields through 500 ps NPT dynamics (Reference: B. Arab, A. Shokuhfar, J. Nano- Electron. Phys. 5 No 1, 01013 (2013))

- Polymer matrix composites
- Graphene materials
- Manufacturing
- Material science
- Raw materials

- Industry Sector: Aeronautics, Chemicals, Manufacturing & Engineering, Material sciences

- Technology: HPC

Contact:
Email: ustunel@metu.edu.tr

Section 4

Manufacturing and Engineering

Nilar AB Accesses VEGA – 80

Multi-Point Aerodynamic Shape Optimization of a Bicycle – 82

Optimizing industrial particle processes through simulation – 84

Multi-scale and multiphysics simulation of a dam – 86

Anomaly Detection of Noisy Time Series - 90

Numerical Analysis and Assessment - 92

Enabling HPC Usage for Expensive ML Tasks - 94

Integrated HPC-HPDA Workflow for Pipe Flow Assessment- 96

Software development and scalability in HPC environment – 98



Photo: Greg Rosenke// Unsplash



Photo: Roberto Sorin // Unsplash

Nilar AB Accesses VEGA

ENCCS (EuroCC National Competence Centre Sweden).
ENCCS provides high-performance computing training and support for industry, academia and public administration for free.
https://enccs.se

Organisations Involved

Nilar AB designs, develops and manufactures Nilar Hydride® batteries in their state-of-the-art factory in Gävle, Sweden. With products that offer unique safety benefits, environmental awareness and long-term power, Nilar enables the optimal battery system, whether it be for your home, business, or infrastructure project.
https://www.nilar.com

Technical Challenge

Nilar AB is a Swedish SME that designs, develops and manufactures batteries for stationary energy storage. To ensure excellent performance, lifespan and safety of the batteries, the manufacturing process involves both optical quality inspection as well as a set of tests that the battery modules must pass before they can be sent to the customers. Manual optical inspection is time-consuming, ineffective and prone to error.



Figure 1: logo of the NILAR-ENCCS collaboration

Solution

Together with researchers at RISE, Nilar aims to improve optical inspection by using AI-based machine vision. This work is part of an ongoing Swedish national funding agency (Vinnova) SiP PiiA project, "AI for increased process efficiency".

Business Impact

Nilar is a modern SME where production efficiency plays a crucial role in their activities. By automating optical inspection of batteries, Nilar is able to save time and resources relating to quality control.



Photo: CHUTTERSNAP // Unsplash

Benefits

With the acquired development access to VEGA in Slovenia with the help of ENCCS, it has been possible...

- ✓ to speed up the training time of the AI models significantly,
- ✓ increase the number of experiments that can be done on the image datasets

- Batteries
- Battery development
- Machine vision
- AI

- Industry Sector: Manufacturing and Engineering

- Technology: AI

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

Multi-Point Aerodynamic Shape Optimization of a Bicycle

The Greek NCC "EuroCC@Greece" is run by a consortium of 5 institutions and coordinated by GRNET. The objective of the NCC is 3-fold: i) advance competitiveness in research ii) improve effectiveness of government services and iii) promote innovation in industry. In this success story, the hardware provided by GRNET and the s/w and expertise of PCOPt/NTUA in the field of CFD-based aerodynamic optimization were combined to provide services to the client.

Industrial organisations Involved

Grammarios Bikes is a Greek SME focusing on building, maintaining and optimizing custom-made bicycles. They recently started evaluating computational tools to analyze and optimize parts of their products. In this context, they wanted to evaluate the suitability of CFD-based aerodynamic optimization in terms of gains in performance of the final product as well as cost and means to conduct the numerical optimization.

The Challenge

Grammarios Bikes asked for a custom-made bicycle with lower drag and side forces than the baseline one. Specifically, the new bicycle should perform better than the current one in a range of bicycle and side wind velocities. The combination of these requirements gave rise to a multi-point, multi-objective aerodynamic optimization problem, relying on Computational Fluid Dynamics (CFD) for analyzing the aerodynamic performance of the bicycle at hand.



Photo: Jacek Dylag // Unsplash

The Solution

An adjoint-assisted, gradient-based optimization algorithm was used in this problem. The adjoin-tOptimisationFoam solver, developed by PCOpt/NTUA and made publicly available through the opensource s/w, OpenFOAM, was utilized. Two objective functions from four operating points were concatenated, using weights provided by the end user based on the frequency of each operating point.

The HPC infrastructure of GRNET was utilized, to highlight the technical capabilities available on a national level and to help introduce the client into HPC. In less than 48100 CPU*hours, drag was reduced from 2% to 3.9%, depending on the operating point, while keeping the side force intact. This result was achieved by enlarging the front fairing and increasing the inner radius and spokes width of the front rim.

The Business Impact

Grammarios Bikes was occasionally using CFD to evaluate potential designs using hardware available to the SME, but adjustments to them were made through a trial-and-error process, which can be quite time consuming and of questionable outcome. Using CFD-based optimization and the hardware provided by the NCC allowed for a significant reduction of the wall-clock time to solution and of person-hours due to the automated nature of the process. Grammarios Bikes is now considering introducing CFD-based optimization into its process of designing new products.

Benefits

- ✔ Increase of performance by 2% to 3.9%, depending on the operating point
- ✔ Reduction of CPU wall-clock to solution, compared to manual trial-and-error, by around 30%
- ✔ Reduction of person-hours, compared to manual trial-and-error, by around 50%

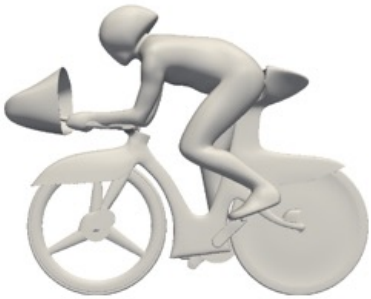


Figure 1: Initial bicycle geometry



Figure 2: Optimized bicycle geometry



Figure 3: Change (in meters) in the various parts of the bicycle

- Aerodynamic Shape Optimization
- Adjoint Methods
- Bycycles
- CFD
- HPC

Industry Sector: Manufacturing and Engineering, Mechanical Engineering

Technology: HPC, Optimization

Contact:
Dr. Evangelos Papoutsis-Kiachagias, Research Engineer
Affiliation: National Technical University of Athens, School of Mechanical Engineering, Parallel CFD & Optimization Unit
Address: Zografou Campus, 9 Iroon Polytechneiou str, 15772, Zografou, Athens, Greece
Email: vpapout@mail.ntua.gr

Optimizing industrial particle processes through simulation

EuroCC Belgium is Belgium's National Competence Centre (NCC) in the area of high-performance computing (HPC) and high-performance data analytics (HPDA) and coordinates activities in all HPC-related fields and serves as a reference contact point on HPC/HPDA/AI at a national level.
The consortium is composed of the following partners: Cenaero, CECI and Innoviris Brussel and the Flemish Supercomputing Center.

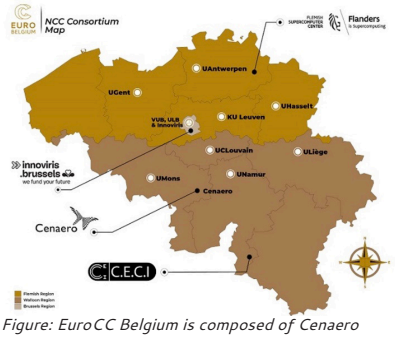


Figure: EuroCC Belgium is composed of Cenaero and Ceci in Wallonia, VSC in Flanders and Innoviris Brussels. The universities are partners in the regional umbrella organizations.

Industrial organisation involved

MPACTS is a spin-off company from KU Leuven that specializes in simulation technology. Founded in 2018 and based in Lievegem, Belgium, the company collaborates closely with the MeBioS research group to develop its software.

The MPACTS software uses the Discrete Element Method (DEM) to simulate the behavior of large numbers of particles in order to improve machine design and performance. By testing designs in a virtual environment, the company is able to save development time and costs associated with building physical prototypes.
To perform simulations for clients, MPACTS relies on the VSC infrastructure to provide the necessary computational resources. This allows MPACTS to understand, predict, and solve industrial problems for its clients with a high degree of accuracy and efficiency.

Technical Challenge

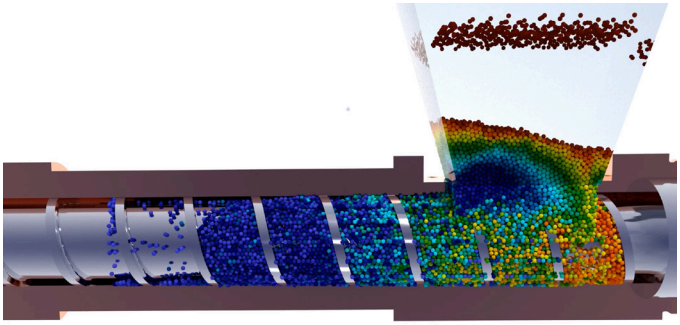
More particles can only be simulated in shorter time spans by making use of the parallel resources available. For 'shared memory' parallelism such as TBB or GP-GPU, the bottleneck for this type of simulation is typically memory bandwidth and memory latency. When switching to an MPI parallel acceleration, across multiple compute nodes, the main challenge becomes an efficient dynamic domain decomposition as the particles move across the simulated domain.

Solution

To make the simulations faster, several strategies can be implemented. First, the software itself can be made more efficient and faster. Second, more computational resources can be used, and acceleration hardware such as GPUs can be used.

Simon Vanmarcke (Mpacts): "As a striking example of the first strategy, the experts at VSC recommended we sort the particles so that particles close in (simulated) space are also close together in computer memory. This increases the likelihood for cache hits and decreases the overall memory bandwidth." The effect on the computational time is illustrated in the YouTube movie for single-core performance. In GPU-accelerated simulations, the effect is much more pronounced as the memory throughput on these devices is even more of a bottleneck.

More particles can only be simulated in shorter time spans by using the parallel resources available. For 'shared memory' parallelism such as TBB or GP-GPU, the bottleneck for this type of simulation is typically memory bandwidth and memory latency. When switching to an MPI parallel acceleration across multiple compute nodes, the main challenge becomes an efficient dynamic domain decomposition as the particles move across the simulated domain. Additionally, the VSC makes it possible for Mpacts to obtain very large amounts of computational power which have GPU acceleration available.



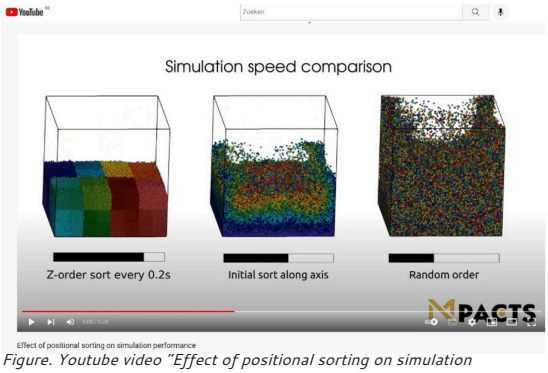


Figure. Youtube video "Effect of positional sorting on simulation performance": https://www.youtube.com/watch?v=QPuAXO-iXOI

Impact/Benefits

Simon Vanmarcke (Mpacts): "Using the supercomputing infrastructure of the VSC matches our usage pattern very well. Typically, we require 'bursts' of computational power when performing simulations for clients. It would be unfeasible and inefficient to purchase the required computational power in-house as it would be costly to do so and idle most of the time. Working with the VSC enables us to deliver results with 'industrial turnaround times' at a cost that scales only with how much actual simulation work is done, not with the peak resources used. A real game-changer."

- DEM
- Particle simulations
- HPC

- Industry Sector: Services and software providers, Manufacturing and Engineering

- Technology: HPC

Contact:
EuroCC Belgium - Jothi Blontrock, Copywriter
Web: www.enccb.be
Email: communication@enccb.be or Jothi.blontrock@uhasselt.be
+32 11 26 86 33
L: E123 - UHasselt Campus Diepenbeek
Twitter: @EuroCC_Belgium
LinkedIn: EuroCC@Belgium



Photo: Gabor Koszegi// Unsplash

Multi-scale and multiphysics simulation of a dam

The French NCC: CC-FR, dedicated to HPC, HPDA and AI technologies, 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.

CC-FR CENTRE DE COMPÉTENCE HPC.HPDA.IA

Organisations Involved

CC-FR, set up within the framework of EuroHPC, is managed by the European Pole of Competence in high performance digital simulation Teratec, in association with the European Center for **Advanced Research and Training in Scientific Computing** Cerfacs and with the participation of the Big National Equipment Intensive Computing Genci.

Teratec

CERFACS
CENTRE EUROPÉEN DE RECHERCHE ET DE FORMATION AVANCÉE EN CALCUL SCIENTIFIQUE

GENCI
Le calcul intensif au service de la connaissance

CEVAA, https://cevaa.com, specialized in acoustics and mechanics, offers its test equipment and its expertise to Industry. Numerical simulation by finite elements is their major tool, with various specialized software and a dedicated team of engineers.

CEVAA
CENTRE D'EXPERTISE ET D'ESSAIS
VIBRATION / ACOUSTIQUE / FIABILITÉ

Technical Challenge

In 2021, CEVAA was asked to study the mechanical strength of a dam. With scales ranging from centimeters to several tens of meters, taking into account loads multiphysics (temperature, pressure, forces), the internal software and computing resources of CEVAA were insufficient to conduct the study.

Therefore, CEVAA requested The French NCC support to use HPC Computing resources available at the CRIANN through GENCI as well as the latest sequential and parallel versions of Code Aster for the realization of a multi-scale simulation and multiphysics simulation of a dam.



Figure 1: dam simulation

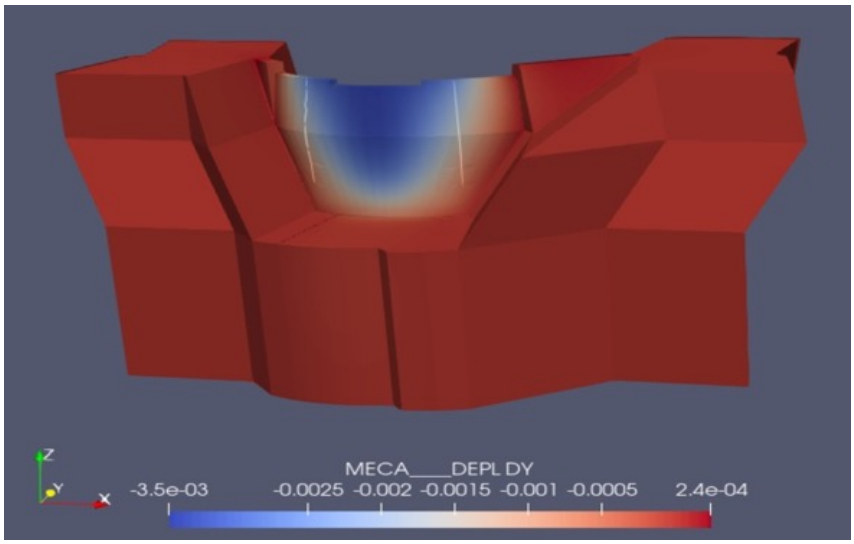


Figure 2: Results of modal deformation on a dam

Solution

The study was carried out on the CRIANN Myria calculator with Code Aster for calculations and SalomeMeca for mesh and post-processing.

The project received advanced support from the CRIANN, with the installation of the latest versions Code Aster sequential and parallel (versions 13 and 14), then help in the creation of the scripts to launch calculations. Several dozen calculations were launched on 8 CPU cores and 120 GB of memory.

Business Impact

The tailored program of CC-FR enable large client projects to be carried out with a significant productivity gain.

Benefits

- The use of the Myria calculator allowed to carry out several simulations simultaneously, and to obtain a delivery time in agreement with the customer needs.
- The remote viewing service allowed to perform pre and post-processing operations very smoothly on the calculator, without having to move the data.

- HPC
- Simulation
- Multi-scale
- Multiphysics
- Mechanical Engineering

- Industry sector: Mechanical Engineering

- Technology: HPC, Code Aster

Download the success stories:
https://cc-fr.eu/wp1/wp-content/mkp-starter/images/accompaniments/Fiche%20simseo%20-%20CEVAA.PDF
More information:
https://cc-fr.eu/accompagnement

Contact:
Dr Karim Azoum
https://www.linkedin.com/in/karim-azoum-45011710a
Email: Karim.azoum@teratec.fr
+33 7 62 74 03 60

Anomaly Detection of Noisy Time Series

The infrastructure provider Sigma2 and the research companies SINTEF and NORCE have joined forces to establish The Norwegian Euro Competence Centre. The competence centre raises awareness and provides Norwegian SMEs with the expertise necessary to take advantage of innovation possibilities created by HPC (High-Performance Computing) technology. This includes HPDA (High-Performance Data Analytics), ML (Machine learning) and AI (Artificial intelligence), thus increasing the SME's competitiveness.

Industrial organisations Involved

This success story is a result of a proof-of-concept evaluation carried out by NORCE in collaboration with 3B-the fibreglass company. 3B (https://www.3b-fibreglass.com) capitalizes on extensive glass fibre expertise developed over 50 years. Its history started with the Belgian plant in Battice, which saw its first production run in 1966 as part of Owens Corning. In 2008, 3B was formed as an independent company integrating three of the sites formerly owned by Owens Corning: the two benchmark manufacturing facilities of Battice (Belgium) and Birkeland (Norway), as well as our Customer Service Centre based in Brussels (Belgium).

The Challenge

3B-the fibreglass company is producing different types of fibreglass. The products are wound onto bobbin shapes at a speed of typically 50km/h. The production is to a high degree automatic. One the remaining manual procedures is the detection of breaking fibres during the winding process.

The aim of the collaboration was two-fold. The first aim was to investigate the possibility to automatically detect breaking fibres using the winding motor's current measured with a frequency of 20Hz. In the effort team members of Dr. Paul Jensen (3B) and Dr. Klaus Johannsen (NORCE) have been collaborating to design time series analysis procedures capable to perform the detection. The second aim was to increase 3B's competence in data analytics, machine learning and modeling. This was achieved by the named collaboration as well through an efficient learning-by-doing approach.

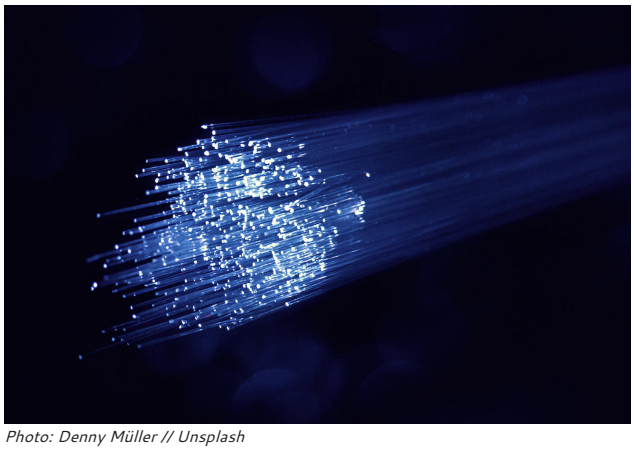


Photo: Denny Müller // Unsplash

The Solution

As a results of our collaboration, we developed an anomaly detection procedure based on the following key elements:

- Semi-supervised learning, i.e., modeling of only the normal behavior of the time series (one class classification) using a k-means bisection based Gaussian mixture model,
- A four-dimensional smoothed feature space derived from the one-dimensional time series of currents using the time and the first three coefficients of a time-local polynomial approximation,
- A smoothing in score-space using exponential smoothing.

The quality of the detection could significantly be increased to reach a precision and recall of 98.5% respectively 77.6%. The relatively low recall is due to end-of-bobbin anomalies, which has to be handled differently.

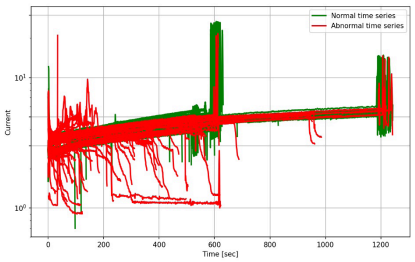


Figure 1: Time Series, normal (green), abnormal (red

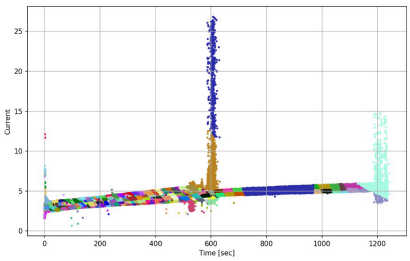


Figure 2: Partitioning of normal time series with minimum size 1000into 362 partitions.

The Business impact

The impact of this investigation is two-fold. The results help the company management to take informed decisions on the development of their break-detection strategy. Further, the company is better prepared to assess the cost-benefit-relationship of further educating their staff in data analytics, machine learning and modeling.

Benefits

This collaboration was beneficial for both parties.

NORCE:

- ✓ has learned about the challenges of an interesting industry partner,
- ✓ understands better the level of machine learning competence in SMEs and
- ✓ understands better to which extent machine learning procedures can provide added value to industry partners.

3B-the fibreglass company:

- ✓ has gained insight into the capabilities of advanced machine learning procedures and
- ✓ understands better the competence level required to apply machine learning in their company.

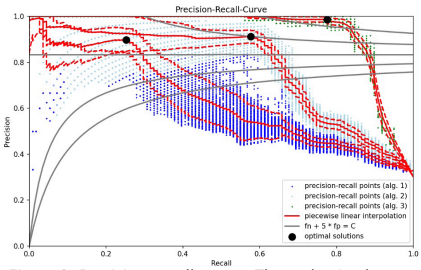


Figure 3: Precision-recall curves. The evaluation has been bootstrapped 100 times. The blue points show the precision-recall points, the red lines give the average precision-recall curve and its uncertainty interval (+- 1 standard deviation.

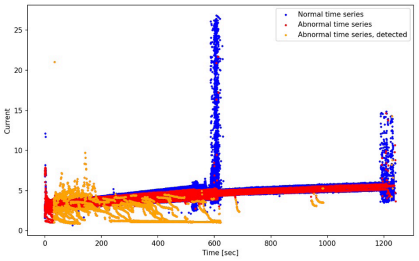


Figure 4: Display of abnormal points as detected by the algorithm. The identification of abnormal points are also possible within the areas covered by blue (normal points). This is due to the modeling of slope and curvature

- Glass Fibre Production
- Time Series Analysis
- Anomaly Detection
- Semi-Supervised Learning

- Industry Sector: Process Inndustry, Material Services

- Technology: Machine Learning

Contact:
K. Johannsen,
NORCE Norwegian Research Centre AS,
Nygårdsgaten 112, 5008 Bergen, Norway,
Email: kljo@norceresearch.no

Code Modernization for the Glass Industry

Turkish National e-Science e-Infrastructure (TRUBA), operating under Turkish Academic Network and Information Center (TUBITAK ULAKBIM) is the coordinator of NCC Turkey. 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. Our 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. The third-party presenting this success story is Sabancı University.

Industrial Organisation Involved

As one of the most powerful industrial conglomerates in Turkey, Şişecam is a global player in all key areas of the glass industry. It is internationally renowned for its ever-growing production power, highly reputable brand image, superior product quality, and value-creating sustainable growth approach. It is currently one of the world's leading glass producers with production operations located in 14 countries on four continents, 24 000 employees, and sales in over 150 countries around the globe.

Technical Challenge

Şişecam fabricates glass products using custom-made furnaces. They use a simulation code written some decades ago that traces the glass particles inside the furnace. In a nutshell, the longer the particles stay inside the furnace, the better the quality of the product. To achieve this, they need to perform simulations with many particles; however, the computational cost increases enormously as the number of particles increases.

To overcome this problem, the company applied for support from the center for consultancy on code modernization and scalable parallelization of the code to decrease the design process of the furnaces.

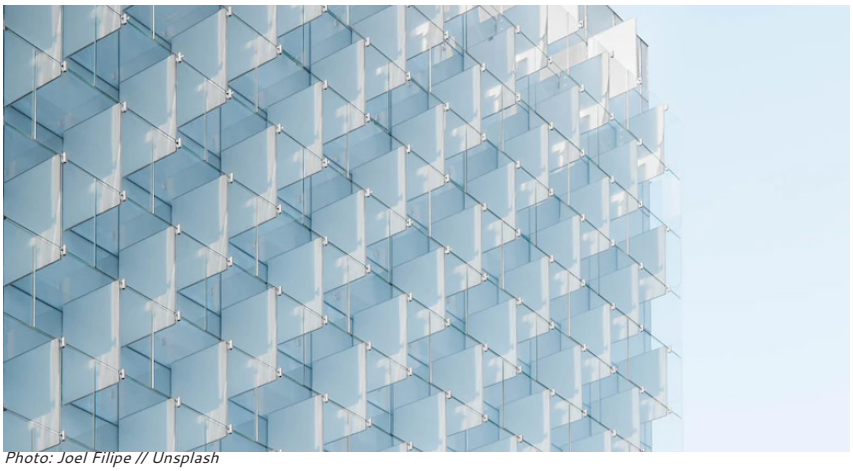


Photo: Joel Filipe // Unsplash

Solution

Before the use-case study, the simulation process was taking days and Şişecam was limited by small models to make the process faster and wanted to increase the size to have more accurate simulations. In this case study, after discussions, we all agreed that converting the whole codebase to C++ was the main path we needed to follow.

The competence center helped Şişecam's engineers to convert the code to a more modern architecture and discuss/recommend parallelization strategies on a multicore system with OpenMP. At last, the Fortran code is translated to modern C++ line by line, except goto statements which are translated as loop constructs for better readability. Then, the dead code is eliminated to reduce code size. Common logical patterns are converted to functions for better readability.

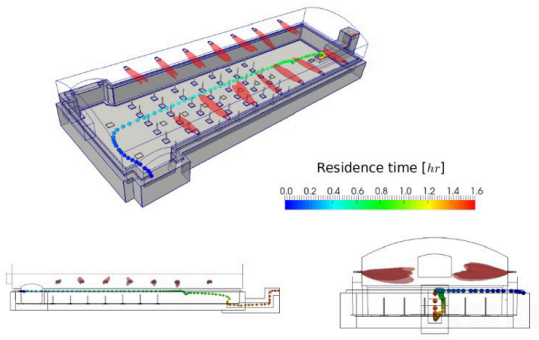


Figure 1. Tracing the glass particles inside the oven

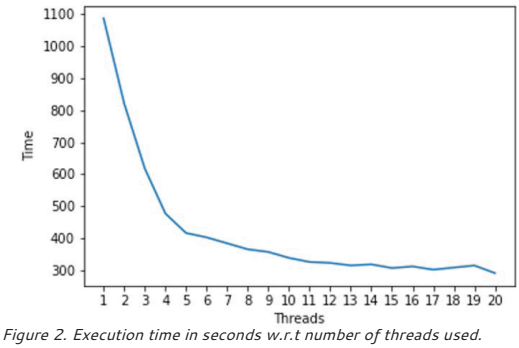


Figure 2. Execution time in seconds w.r.t number of threads used.

Business Impact

In the simulation code, the particles are processed in batches. After the code is parallelized and modernized, the batch size is increased from 10 to 1000. The reason for using a large batch size is the long-tailed running time of some of the particles; using a larger batch size and dynamic scheduling, we were able to increase core utilization to > 90% from < 20% on the development machine (with 16-core/32-thread AMD Ryzen9).

On the longest-running test case provided by Şişecam, the Fortran code was taking 7 hours 14 minutes on the development machine, whereas unoptimized C++ code takes only 27 minutes 21 seconds. On a server with 20-core Xeon 6148 2.40 GHz, the code only takes 286 seconds. Even when a single thread, the runtime is 18 minutes, a huge speed-up compared to 7 hours on the development machine. Using 20 cores, the new implementation was able to achieve 3.86x speedup which can be further improved.

Benefits

- ✔ We were able to translate a complicated numerical code that depends on depreciated technologies to modern C++.
- ✔ While translating, improvements in both maintainability and performance are done.
- ✔ In addition, the implementation is parallelized to use multiple cores with efficiency and the runtime is reduced from hours to minutes.

- Glass Production
- Particle Simulation
- Code Modernization

- Industry Sector: Glass Manufacturing & Engineering

- Technology: HPC

Contact:
Kamer Kaya,
Sabancı University,
Computer Science and Engineering
Email: kaya@sabanciuniv.edu



Photo: Jani Brumat // Unsplash

Numerical Analysis and Assessment

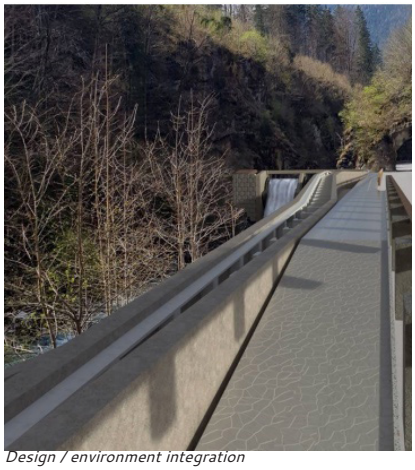
of a conceptual Dam Proposal

Croatian Competence Centre (HR HPC CC) provides end users from scientific and higher education communities, various industries and public administrations, access to innovative solutions adapted to the level of maturity of national and European High-Performance Computing (HPC) infrastructure. HR HPC CC helps strengthen existing and develop new national competencies for High-Performance Computing (HPC), High-Performance Data Analytics (HPDA) and the area of Artificial Intelligence (AI).

HRVATSKI CENTAR KOMPETENCIJA ZA HPC

Organisations Involved

CWKW Felsentunnel GmbH & Co. KG is an SME contractor developing a hydroelectric power plant for the Municipality of Ramsau b. Berchtesgarden. They are a small and agile SME in the engineering field.

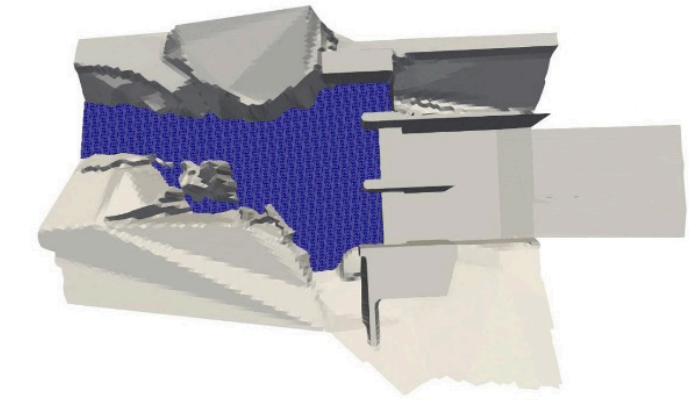


Design / environment integration

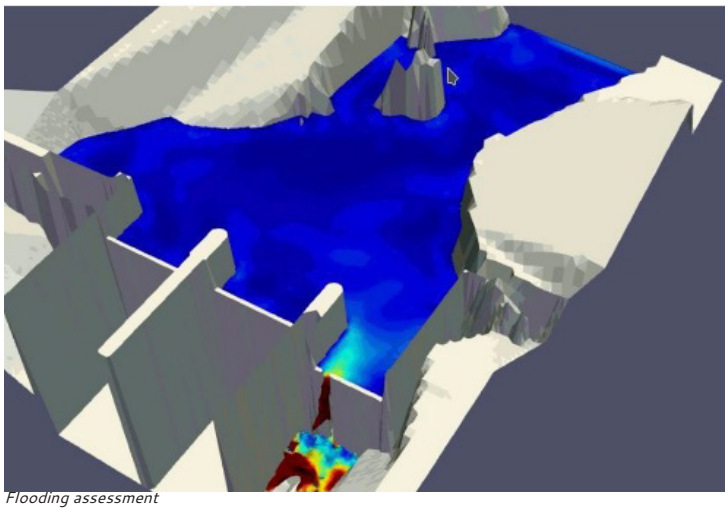
Technical Challenge

SME partner has been developing a small river dam that should meet energy production requirements so as to be a viable option while also integrating into the surrounding environment with minimal disturbance/ changes.

NCC's goal in this project was the assessment of the feasibility of such a dam, assessment of potential flooding and overall integration into the environment which entails the development of visual materials. Dam feasibility was determined analytically (by a third party), whereas flooding analyses and visualization were done by the NCC using Computational Fluid Dynamics (CFD).



Dam – numerical model



Flooding assessment

Solution

Numerical simulations were based on the current state-of-the-art approach to flooding assessment. An open-source CFD software stack was employed and a customized solver based on the Volume of Fluid methodology was developed to ascertain the possibility and extent of the flooding region in a realistic environment. Terrain and realistic river inflow data were obtained. Based on findings, improvements in the current dam design were suggested in order to ensure safety.

Business Impact

Due to the complexity of the problem, the only viable approach to solve this task is either by using Computational Fluid Dynamics or a similar numerical approach. However, conventional computers lack the processing power to resolve this problem in a reasonable time frame, which is important due to tight project development schedules. By utilizing HPC resources, simulations were completed in a time frame that is several orders of magnitude smaller. Input by the experts from the Academia, findings and suggested solutions incentivized the SME partner to investigate the use of HPC resources for their proprietary code as well as explore future collaborative projects.

Benefits

- ✓ Assessment time reduced by several orders of magnitude
- ✓ Approach allows iterative design changes to resolve safety concerns

- Computational Fluid Dynamics
- HPC
- Dam
- Flooding
- Safety
- Mechanical Engineering
- Volume of Fluid

- Industry Sector: Mechanical Engineering, Environment

- Technology: HPC

Contact:
Lado Kranjčević
Email: lado.kranjcevic@riteh.hr

Enabling HPC Usage
for Expensive ML Tasks on Manufacturing Environments

Turkish National e-Science e-Infrastructure (TRUBA), operating under Turkish Academic Network and Information Center (TUBITAK ULAKBIM) is the coordinator of NCC Turkey. 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. Our 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.
The particular third-party presenting this success story is Sabancı University.

Industrial organisations Involved

Erste Software is an SME founded in 2017 in Ankara, Turkey. Its focus is smart environments, and its main expertise is on Internet of Things and Big Data solutions. Despite being a young SME, Erste employs an experienced R&D team that has been involved in several EU-funded projects mainly leveraging industrial IoT technologies such as Optimum, Pianism, I2Panema, Machinaide, and Gamma. Overall, the company has three main products (1) IoTWare, an engineering tool suite for enhanced smart environments (2) MobiVisor, a secure mobile platform management tool, and (3) Detangle, an innovative feature-based effort and quality analysis software.

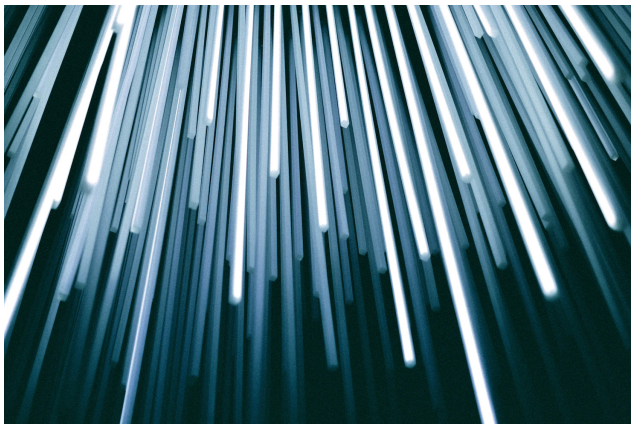


Photo: Christopher Burns // Unsplash

The Challenge

Erste is working on two international manufacturing projects. They require Anomaly Detection and Predictive Maintenance models trained on a huge amount of data. The company is highly fluent in front-end and back-end development; the R&D team was developing prediction and classification algorithms to reduce the maintenance cost and downtime of the press machines used in production.

However, in their use-case, the data can stream from up to 10 production lines, each having 3-10 machines, and each machine is equipped with sensors producing tens of data points in every milli-second. Having expertise on streaming data processing, Erste did not have any HPC skills or know-how on HPC clusters, SLURM, etc. This is why they have applied for support from the center.

The Solution

As ML models get larger, the amount of time and resources needed for training that doesn't reduce the performance gets larger as well. Using an HPC cluster answers Erste's needs in terms of computation resources. In this way, the end-to-end predictive maintenance tool they have can work on-site as is without consuming the critical server resources while offloading heavy ML tasks to leverage an external cluster and being able to perform them faster than regular servers. This being said, there was no such solution enabling HPC usage at the time when we started the use case. With the center's consultancy and guidance, a solution that enables HPC usage is designed.

Given a set of future tasks, the tool can automatically handle data movement, training, downloading models and deploying them back to the on-site server. To implement it, an Erste engineer has learned to use SLURM, which is a cluster management and job scheduling system providing an interface for executing and monitoring tasks, and managing the contention for resources. At the end of the use case, Erste completed a proof-of-concept implementation of this module for their ML-based predictive-maintenance tool.

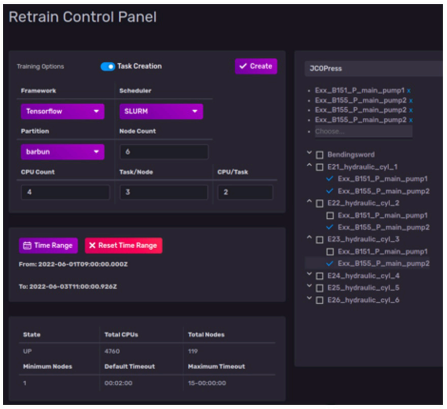


Figure 1: Task creation UI

The Business impact

Considering the size of the manufacturing companies and the data that is produced/gathered from the machines, the tool will improve the applicability of the ML on Erste's solution. For instance, training a medium-sized ML model on the server containing the streaming DB and answering the queries takes around 1 hour (on 8 cores) while no queries are being executed on it. With TRUBA's high-end CPUs, the time reduces to around 30 minutes (on 8 cores) and with a single GPU, the training takes less than 10 minutes. Overall, the tool developed in this use-case allows Erste to keep the scarce on-site computational resources dedicated to their end-to-end Predictive Maintenance tool, basic data processing, e.g., rule-based anomaly detection on streaming data and data queries and leverage external HPC clusters.

Benefits

- ✔ A state-of-the-art tool has been developed to offload streaming-data computations to an HPC cluster has been developed
- ✔ The tool allows on-site computational resources which are usually scarce to be dedicated for the actual task
- ✔ Significant performance improvement has been obtained in ML training

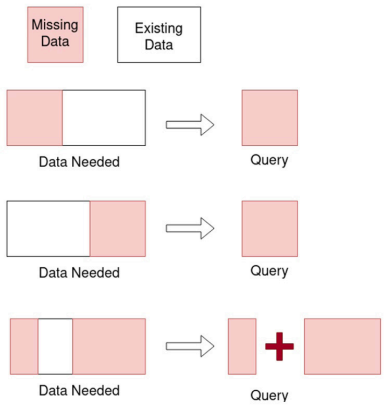


Figure 2: Scheduling of Data Transfers

- Offloading ML
- Manufacturing
- Streaming Data

- Industry Sector: Manufacturing and Engineering

- Technology: HPC, AI

Contact:
Kamer Kaya, Sabancı University,
Computer Science and Engineering,
Email: kaya@sabanciuniv.edu

Integrated HPC-HPDA
Workflow for Pipe Flow Assessment

Croatian Competence Centre (HR HPC CC) provides end users from scientific and higher education communities, various industries and public administrations, access to innovative solutions adapted to the level of maturity of national and European High-Performance Computing (HPC) infrastructure. HR HPC CC helps strengthen existing and develop new national competencies for High-Performance Computing (HPC), High-Performance Data Analytics (HPDA) and the area of Artificial Intelligence (AI).

HRVATSKI CENTAR KOMPETENCIJA ZA HPC

ADRIA ELECTRONIC

Industrial Organisation Involved

Adria-electronic d.o.o. is a company founded in 1974 that specializes in intelligent electrical systems, automation and control. The company's main products are intelligent sensing systems for hotel rooms, building automatization systems and electrical energy sensing equipment and software.

Technical Challenge

SME partner required help with ultrasonic pipe measurements, general sensing and automation. Therefore, a series of experimental test measurements had to be conducted, obtained experimental data had to be analysed, and subsequently used to improve current systems that integrate different sensing technologies.

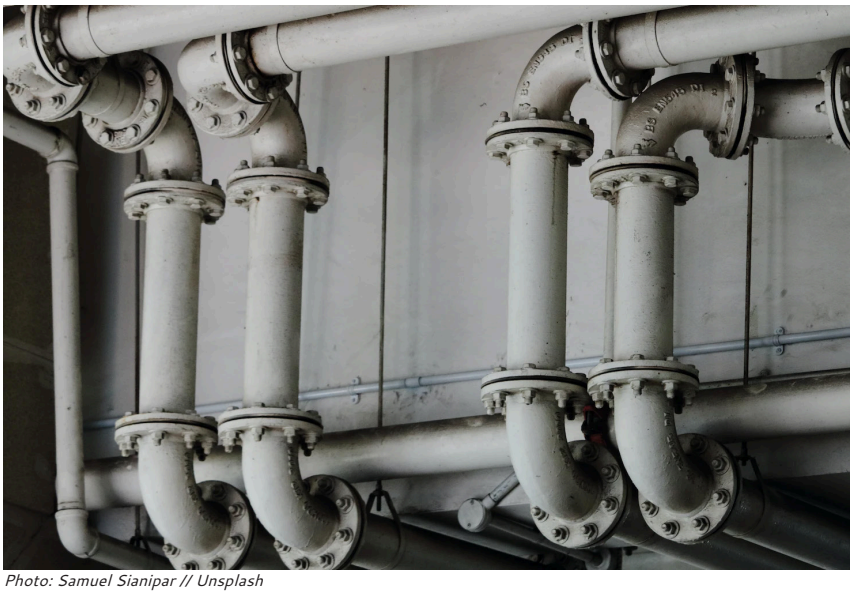
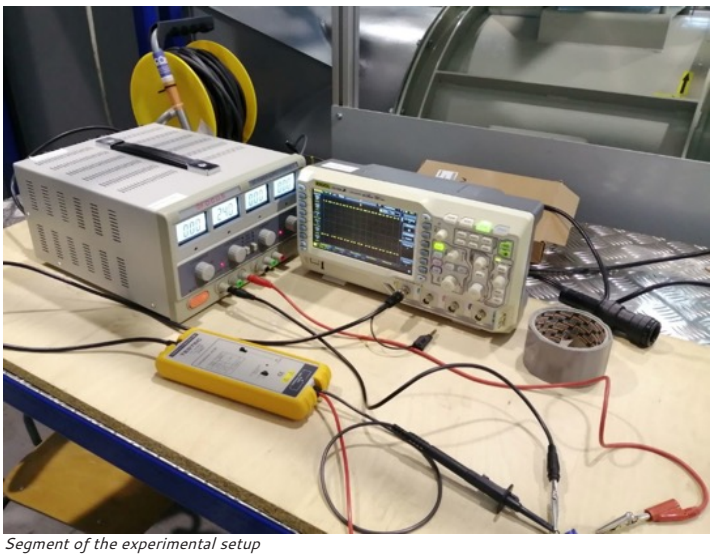


Photo: Samuel Sianipar // Unsplash

Solution

NCC Croatia along with the Department of Fluid Mechanics and Computational Engineering at the Faculty of Engineering, University of Rijeka, assisted with big data analytics of flow meter oscilloscope measurements.

Members of the Faculty conducted experimental measurements in order to obtain large datasets which were subsequently used to statistically analyse an SME-defined system. Due to the complexity of the datasets, computational requirements were significant hence the use of extensive computational resources was required.



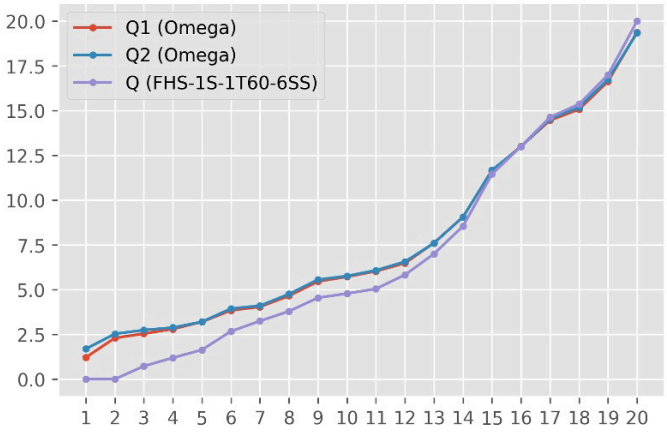
Segment of the experimental setup

Business Impact

This collaboration enabled prototyping and potential developments to be used in a future commercial solution.

Benefits

- Insight/knowledge of the system behaviour
- Gathered information and conclusions will facilitate development of future products and solutions



Results agreement for the conducted calibration and analytics model

- Data Analytics
- Big Data
- Measurements
- Sensors
- Automation

Industry Sector: Mechanical Engineering

- Technology: Big Data

Contact:
Lado Kranjčević
Email: lado.kranjcevic@riteh.hr

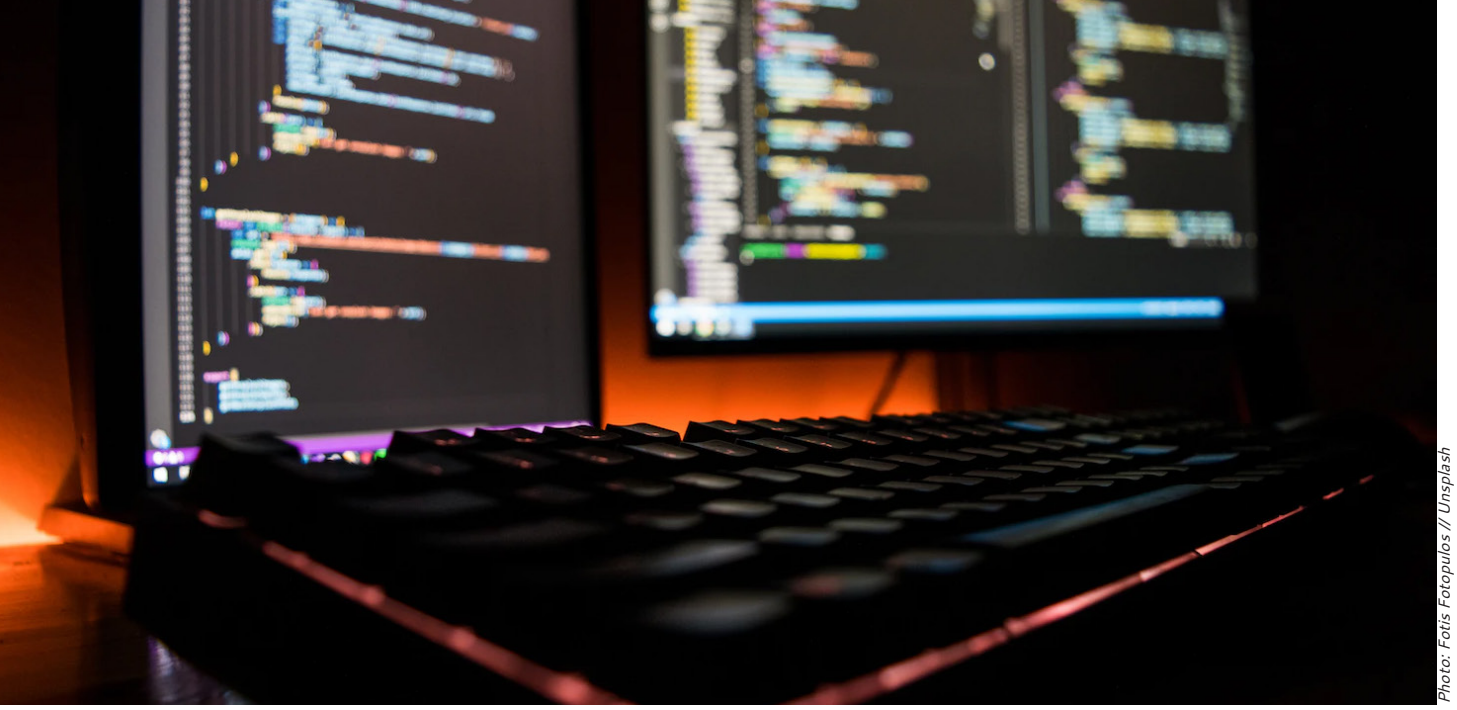


Photo: Fotis Fotopulos // Unsplash

Software development and scalability in HPC environment

Croatian Competence Centre (HR HPC CC) provides end users from scientific and higher education communities, various industries and public administrations, access to innovative solutions adapted to the level of maturity of national and European High-Performance Computing (HPC) infrastructure. HR HPC CC helps strengthen existing and develop new national competencies for High-Performance Computing (HPC), High-Performance Data Analytics (HPDA) and the area of Artificial Intelligence (AI).

HRVATSKI CENTAR KOMPETENCIJA ZA HPC

Organisations Involved

IB-Nujić j.d.o.o. is a specialized office for flow modelling with numerous projects over the past 20 years. The Office has worked on various models for use in water management, including Hydro_AS-2D, Hydro_GS-2D, Hydro_FT-2D, Laser_AS-2D, flow network generator, etc. More than 400 hundred institutions worldwide have successfully implemented the models.

IB-NUJIĆ

Technical Challenge

SME software developer was interested in improving current state of the code that was being internally developed. Three tasks were defined. The code in question is a flow modelling software that utilizes implicit approach. Primary goal was to improve the overall performance i.e. develop a highly-parallelizable code that can be used on HPC systems.

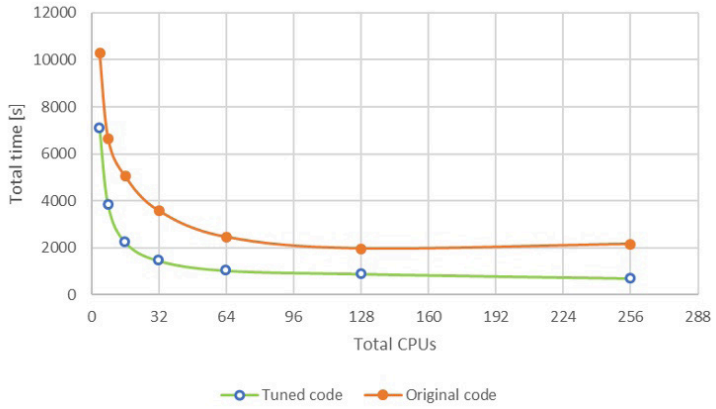
Secondary goal was related to the implementation of a licensing/protection methodology that would allow commercial use of the code on HPC systems with isolated backend. Finally, the assessment of the code and comparative analysis had to be conducted.

Solution

Coding specialists have worked with the code developer to optimize the Fortran code in question using OpenMP and MPI. Achieved parallelization/scalability has been assessed on different platforms and an HPC system. NCC has provided HPC resources for the specific use case. Software licensing has been implemented and evaluated. Accuracy of the results the code provides has been compared with the results calculated by market competitors. Cooperation is still ongoing, with focus on commercialization and HPC-deployed software adoption.

Business Impact

The code in question (early variants) was a commercial success in its own right. However, as the technology moved forward, competitors have been able to provide similarly accurate and swiftly computed results. Implemented parallelism was conducted in order to ensure a competitive edge. Results (performance) incentivized new companies to investigate and adopt newly developed code. Although there is room for improvement, current computational gains in conjunction with overall accuracy have already increased sales and overall interest in the software.



Scaling improvements for a specific test case.

Benefits

- ✓ Increased performance and scalability
- ✓ Competitive edge
- ✓ Use of new methodologies/technologies/hardware

- Computational Fluid Dynamics
- HPC
- Software
- OpenMP
- Code Development

- Industry Sector: Mechanical Engineering, Environmental Engineering

- Technology: HPC

Contact:
Lado Kranjčević
Email: lado.kranjcevic@riteh.hr

Section 5

Finance

Fast Generation of Low-Discrepancy Sequences – 102

Artificial Intelligence and Behavioural Finance – 104

Optimization, parallelization and porting of an AI algorithm – 106

Nested Named Entity Recognition – 108





Photo: Micheile Henderson // Unsplash

Fast Generation of Low-Discrepancy Sequences Using Accelerators

The National Competence Centre of Bulgaria (NCC-Bulgaria) in the area of HPC, HPDA and AI has the goal to enhance and develop the competences of the Bulgarian computational community, making full use of EuroHPC resources and the EuroCC partnership.
NCC-Bulgaria is built by a consortium coordinated by the Institute of Information and Communication Technologies at the Bulgarian Academy of Sciences (IICT-BAS), and two members, Sofia University "St. Kliment Ohridski" (SU), and University of National and World Economy (UNWE). They collaborate with Sofia Tech Park, where the Discoverer EuroHPC supercomputer is operating.

Organisations Involved

The department of Scalable Computing and Applications (SCA) with HPC Centre is one of the departments of the Institute of Information and Communication Technologies at the Bulgarian Academy of Sciences (IICT-BAS). It manages the HPC centre where is located the supercomputer Avitohol.
The department SCA with HPC centre is focused on activities in the development and deployment of Cloud middleware and software components, methods, algorithms, and applications suitable for Cloud and HPC computing systems.

Technical Challenge

Computational accelerators offer significant advantages in terms of computational power when compared with regular CPUs. However, when dealing with the problem of fast generation of low-discrepancy sequences using accelerators, one has to take into account that any complex logic operations and branches in the code require significant rewriting of the code and may become a bottleneck in a quasi-Monte Carlo simulation algorithm. Such algorithms are used in problems with hundreds and frequently thousands of dimensions, arising from pricing problems in Mathematical Finance.

Solution

Modern compilers have powerful built-in vectorisation and optimisation techniques that allow good speed-up to be achieved for codes that have suitable patterns of executing loops and memory accesses. When dealing with our problem, we divided the codes into two parts. For those parts where we expected that the compilers can do better job at vectorisation/optimisation, we added suitable labels at critical points in the code, where some compiler directives or parameters can be varied, and then we considered the performance of the code on certain representative test cases as a loss function that is to be optimised.

The optimisation problem is high-dimensional, with mostly integer variables. For finding the optimum we employed a genetic algorithm and we obtained acceptable performance. Interestingly, in some places it was better to disable the vectorisation capabilities of the compiler for best results.

In order parts of the code, we reorganized the logic operations in a way that decreases branching or at least groups together threads with similar branching patterns.

In some cases branching was replaces with a combination of logic and arithmetic operations that achieves the same overall result. Optimising this second part was more time-consuming but achieved much better improvement in execution speed.

Scientific Impact

Various families of low-discrepancy sequences are studied and used in diverse quasi-Monte Carlo methods. Our codes dealt with the Sobol and Halton sequences, which are one of the most popular families in use. The ability to generate terms of these sequences on systems with accelerators remove a possible bottleneck in many practical applications, e.g., in Mathematical Physics and Finance. In addition to that, the reduced computational cost of generation enabled various efforts for optimising the parameters of these sequences and their distribution properties. Such optimisation procedures employ heuristic optimisation techniques in a setting with very high number of parameters (thousands). The results from the optimisations are stored and can now be used in solving complex practical problems. -

Benefits

- ✔ Fast generation of low-discrepancy sequences enables the use of quasi-Monte Carlo methods on supercomputers equipped with accelerators, with their superior accuracy in complex, high-dimensional problems.
- ✔ Leveraging compiler's ability to optimise and vectorise saves development time and decreases code complexity.
- ✔ Decreasing the computational cost of the generation process enables other algorithms and techniques that otherwise were unfeasible.

Sequence	Time (s)	Speed	Speedup
Halton Vectorized	5.6	455.7	4.8
Halton Auto	315.5	8.1	
Halton Vectorized	22	116.4	14.3

› Supercomputer Applications
› High-dimensional simulations

› Industry Sector: Finance

› Technology: HPC, AI

Contact:
Prof. Emanouil Atanassov,
Email: emanouil@parallel.bas.bg

Prof. Aneta Karaivanova,
Email: anet@parallel.bas.bg

Assoc. Prof. Sofiya Ivanovska

Mariya Durchova,
Email: mabs@parallel.bas.bg

Institute of Information and Communication Technologies,
Bulgarian Academy of Sciences

Decision support tool combining **Artificial Intelligence and Behavioural Finance** adapted to trading and portfolio management

The French NCC: CC-FR, dedicated to HPC, HPDA and AI technologies, 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.

Organisations Involved

CC-FR, set up within the framework of EuroHPC, is managed by the European Pole of Competence in high performance digital simulation Teratec, in association with the European Center for Advanced Research and Training in Scientific Computing Cerfacs and with the participation of the Big National Equipment Intensive Computing Genci.



Photo: shutterstock.com

Horae Technology is an entrepreneurial and independent investment management company, founded in 2021. Their mission is to enhance the value of their clients' savings. Horae Technology addresses both private clients and institutional investors through the mutual fund « Horae Opportunities » and the management of customised mandates. With the support of the Ecole Polytechnique, Horae Technology develops experimental decision-support algorithms inspired by behavioural finance and deep reinforcement learning

The Challenge

In September 2021, Horae Technology contacted the CC-FR Tailored Program, to benefit from a high-level support in intensive computing.

Solution

In collaboration with CRIANN (Regional Centre Information Technology and Digital Applications Normandy), Horae Technology was able to access the MYRIA supercomputer, equipped with more than 10,000 computing cores connected in very high throughput, and capable close to 700 trillion transactions per second.

Thanks to the CRIANN high-level expertise in intensive computing, Horae Technology has been able to significantly increase the power of its artificial intelligence tool and the robustness of its models. Individual and personalized support also enabled the company to quickly become autonomous in the use of the MYRIA supercomputer.

Business Impact

Up to day, Horae Technology continues to improve its tool and they aim for a full integration of artificial intelligence in their investment process over a period of 6 months.

Benefits

This computing power enabled the company to...

- ✔ Significantly improve the calculation time of its intelligence algorithm's artificial training
- ✔ Accelerate the performance of its tool through the parallelization of its code and the ability to analyze thus thousands of financial securities simultaneously



Figure 2: Retroactive evaluation of buy and sell signals

›HPC
›AI
›HPDA
›Algorithm
›Parallelization

›Industry Sector: Finance

›Technology: AI

Contact:
Dr Karim Azoum
LinkedIn: https://www.linkedin.com/in/karim-azoum-45011710a
Email: Karim.azoum@teratec.fr
+33 7 62 74 03 60



Photo: Roberto Sorin // Unsplash

Optimization, parallelization and porting
of an AI algorithm to HPC Architecture

The French NCC: CC-FR, dedicated to HPC, HPDA and AI technologies, 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.

CCFR CENTRE DE COMPÉTENCE HPC.HPDA.IA

Organisations Involved

CC-FR, set up within the framework of EuroHPC, is managed by the European Pole of Competence in high performance digital simulation Teratec, in association with the European Center for Advanced Research and Training in Scientific Computing Cerfacs and with the participation of the Big National Equipment Intensive Computing Genci.

Teratec

CERFACS
CENTRE EUROPÉEN DE RECHERCHE ET DE FORMATION AVANCÉE EN CALCUL SCIENTIFIQUE

GENCI
Le calcul intensif au service de la connaissance

Advestis, https://www.advestis.com, is a European Contract Research Organization (CRO) with practice of data science. Based in Paris, Advestis is a pure player in data science, with 14 data scientists and more than 40 projects since 2011. Advestis provides resources to put into production machine learning pipelines. Their goal is to push the Technology Readiness Level (TRL) of their customers' innovationsin data science projects for: Financial institutions, Predictive maintenance and Medical research.

advestis
one step ahead

Technical Challenge

Advestis integrated the CC-FR Tailored Program, to benefit from a high-level support in intensive computing.

Figure 1: logo of the NILAR-ENCCS collaboration



Photo: CHUTTERSNAP // Unsplash

Solution

Advestis integrated the CC-FR Tailored Program, to benefit from a high-level support in intensive computing.

Business Impact

Thanks to CRIANN high-level expertise in intensive computing, Advestis was able to significantly increase the power of its AdLearn algorithm. Individual and personalized support has also enabled the company to quickly become autonomous in the use of supercomputer.

Benefits

This computing power enabled the company to:

- ✔ Significantly improve the calculation time of its AdLearn algorithm
- ✔ Accelerate the performance of its AdLearn algorithm through the parallelization of its code.

- HPC, AI
- Optimisation, Parralelization
- Predictive Maintenance
- Medical Research

- Industry Sector: Finance, Predictive Maintenance and Medical Research

- Technology: AI and HPC

Contact:
Dr Karim Azoum
Email: Karim.azoum@teratec.fr
+33 7 62 74 03 60

Nested Named Entity Recognition
via Language Model Based Neural Architecture

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

Industrial organisations Involved

Founded in 2015 in Turkey, Artiwise is helping companies improve their customer satisfaction by providing actionable insights. By consolidating all customer communication channels on a single platform, Artiwise enables companies to understand the sentiment of their customers. Artiwise is the 1st company to measure customer satisfaction by analyzing the emotions of customers and offers tools such as root-cause analysis, aspect-based sentiment analysis, instant customer insights, category-based customer satisfaction scores, and periodic comparison reports to help companies identify the factors that contribute to customer dissatisfaction and churn. By providing a holistic view of customer feedback, Artiwise helps companies take a proactive approach improving the customer experience.

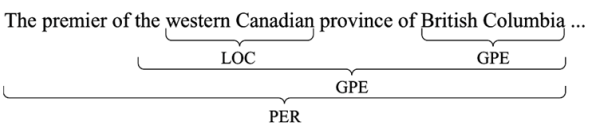


Figure 1: An example for Nested-Ner problem

The Challenge

Named Entity Recognition (NER) is a natural language processing task that extracts predefined categories such as person names, organizations, locations, quantities, monetary values, etc from a text. Most NER models focus on flat entities, which are concepts that can only be associated with one category. However nested NER entities can provide a more complete view of the information. For example, a flat NER model might classify „Chelsea Football Club" as an organization, but in some contexts „Chelsea" may refer to a location. A model using nested entities would be able to capture this distinction. Flat NER models may have limitations in their ability to handle complex situations like this.

The Solution

This research presents a method for identifying named entities in economic news articles written in Turkish. Artiwise used a deep learning model based on transformer-based language models and created a large dataset of annotated articles. The data required cleaning before it could be used, so they implemented a data-cleaning pipeline. Multiple annotators were able to label the data at the same time using the annotation tool, Artiwise Analytics, and a guide was used to ensure consistency in the annotations. The model was trained and evaluated on the TRUBA GPU clusters, achieving an 82% micro-F1 measure. These results suggest that the model is suitable for use in industrial settings and may be able to address various challenges in these domains. Artiwise is grateful for the assistance of Dr. Sefer Baday, their advisor, who provided valuable academic insight. The use of TRUBA's GPU clusters was crucial to the success of the projectI

Business Impact

The current flat-NER models available to the customers can extract meaningful sequences from text, but they do not have the depth of understanding needed for more complex tasks. The company has developed a model that can handle a variety of tasks and have tested it on two different datasets in their project, this model can be used in various industries. It is integrated into the Artiwise Analytics platform, making it easy for users to upload their data, tag the relevant categories, and train the model using user-friendly interfaces. The model is based on a large language model and requires significant hardware resources, such as a GPU cluster, to train. Thanks to the support of EuroCC, the company was able to test their model on a GPU-based cluster and has decided to invest in infrastructure by adding more GPUs to increase their capabilities.

Benefits

- ✔ Operational efficiency thanks to replacing human operations with machine-based systems.
- ✔ Precise analysis than using the flat type of NER models.
- ✔ Understanding a document is key to extending the operation to different departments in the banking domain.

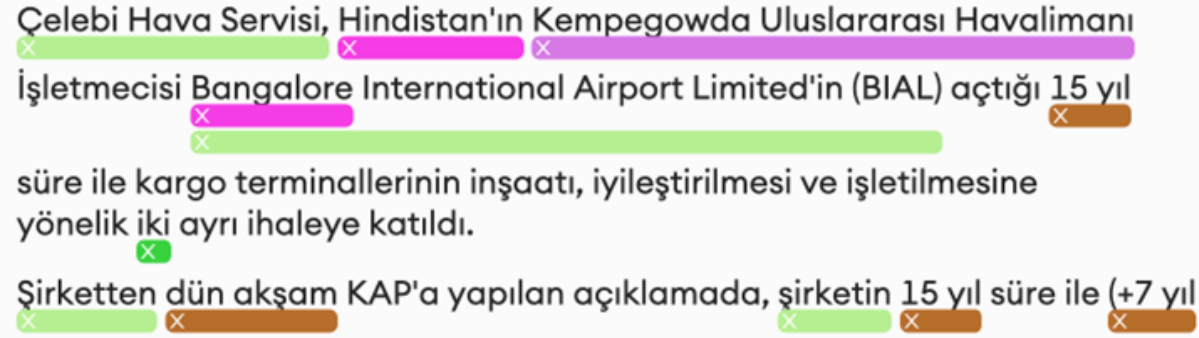
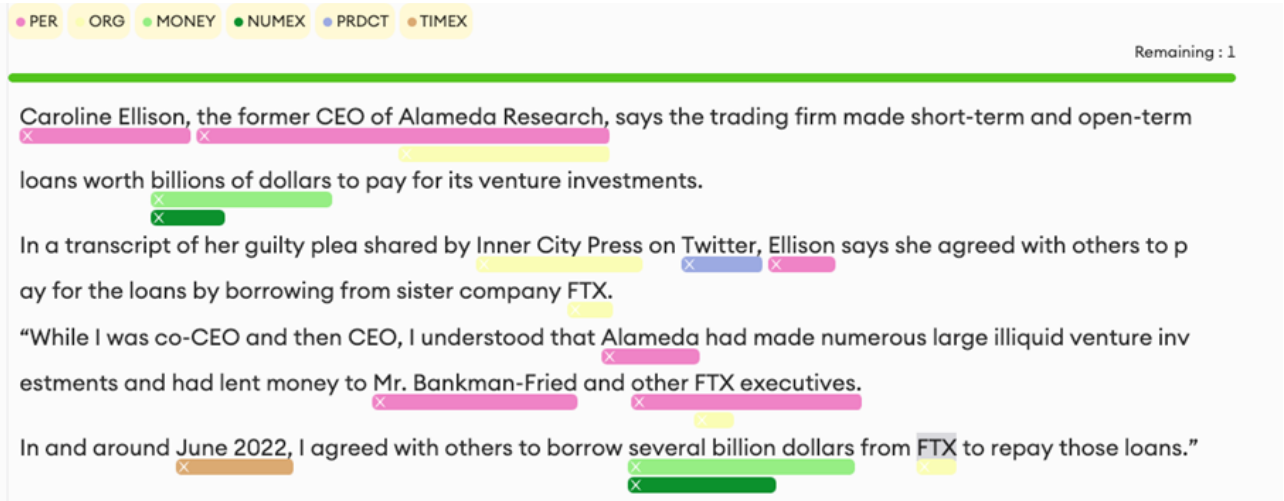


Figure 2: Application UI exapmple. (https://www.bloomberght.com/celebi-hindistan-da-iki-ihaleye-katildi-2303994)



- Nested Ner,
- Named Entity Recognition
- Machine Reading Comprehension
- Information Extraction

- Industry Sector: Finance/Insurance, Automative, Public Services

- Technology: Natural Language Processing, Deep Learning, Machine Learning

Contact:
Fatih Samet Çetin,
Email: fatih.cetin@artiwise.com

Section 6

Pharmacy and Medicine

Developing new transdermal formulations on Vega - 112

Optimized Drugs against COVID-19 – 114

HPC model of human brain visual system with NEST – 116

PaRI -Nordic Pandemic Research Infrastructure – 118

Launch of the Vaccination Centre - 120

Medical Image Processing as a Service – 122

Numerical Simulation of Exoskeleton – 124

AI Medical Device Software – 126

Supercomputer Applications in Biomedical Engineering - 128



Photo: Wengang Zhai // Unsplash



ERCO Pharma to use

The VEGA Supercomputer

for developing new transdermal formulations

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

Organisations Involved

ERCO Pharma is a commercially focused, emerging pharmaceutical services and development company – built on 20 years of research at Karolinska Institutet and strongly committed to increasing the understanding of the human skin barrier and ready to help address transdermal challenges.
https://ercopharma.se

Technical Challenge

ERCO Pharma's proprietary technology and tools, including a high-resolution molecular model of human skin, enable in silico modeling of potential membrane permeability enhancers and drug candidates. Using advanced molecular simulations, they are able to assess the membrane interaction properties and effects of permeation enhancers on a scale inaccessible to conventional in vitro/ex vivo experiments on skin.
Transdermal and topical drug delivery allows for non-invasive, pain-free, continuous drug administration with reduced side effects and increased patient compliance compared to per oral or intravenous drug delivery. Molecular simulations are however computationally expensive and need to run in many-core systems.

Solution

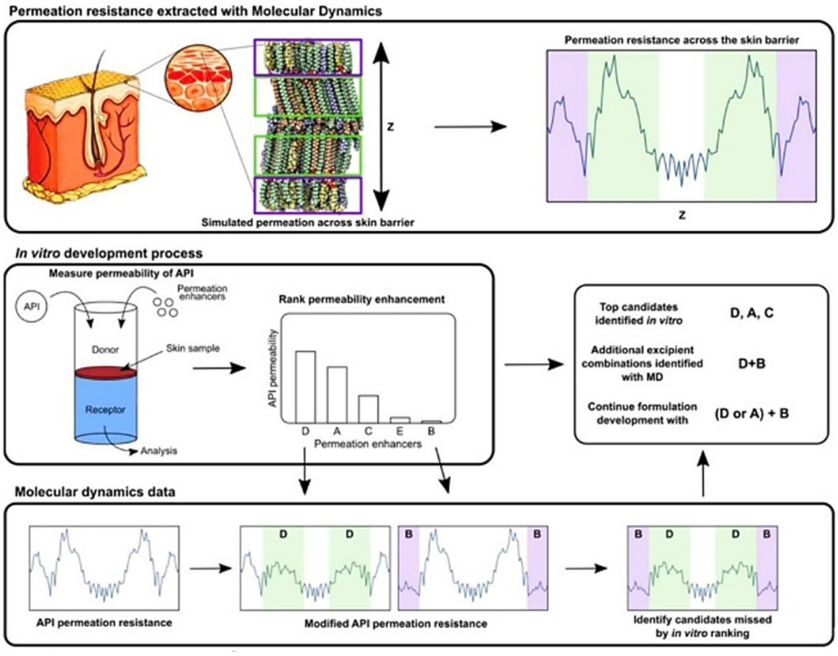
With access to the VEGA supercomputer, ERCO Pharma is now able to use 3 900 000 GPU hours enabling them to explore a new way of developing transdermal, or topical, pharmaceutical formulations. The developed model of the skin barrier structure is combining cryo-electron microscopy (EM) images of skin samples, molecular dynamics (MD) simulations and EM image simulations from the equilibrated MD output systems. (Lundborg et al., 2018)

Business Impact

This atomistic model for skin permeability is patented by ERCO Pharma AB.
Using HPC facilitates selecting combinations of permeation enhancers in a way that is not possible based on experimental data from single permeation enhancers.
ERCO Pharma envisions being able to improve existing transdermal formulations as well as improving the development process of novel delivery vehicles.

Benefits

- The MD skin permeability simulations provide increased knowledge about where, in the skin barrier, a permeant is passing the slowest.
- It also shows where, in the skin barrier, permeability enhancers are most likely to remain and be effective.
- This facilitates selecting combinations of permeation enhancers in a way that is not possible based on experimental data from single permeation enhancers.



A combination of in vitro tests and molecular dynamics simulations can help identify combinations of permeation enhancers not found from in vitro testing alone. (Image from https://ercopharma.se/what-we-offer/topical-formulations/)

- Molecular Simulations
- Medical Science
- Molecular Model
- Skin Model

- Industry Sector: Pharmacy and Medics

- Technology: HPC

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

Researchers Discover Promising Molecules

to develop optimized Drugs against COVID-19

The Portuguese National Competence Center (NCC) in EuroCC is coordinated by the Portuguese national funding agency for science, research and technology (FCT) and integrates several entities that through the NCC provide their expertise and capabilities to support the HPC adoption and usage in the country.

EUROCC PORTUGAL

Industrial Organisations Involved

LIP is both a partner of EuroCC and partner of INCD, the Portuguese Distributed Computing Infrastructure. LIP facilitated the access and supported BioSIM researchers in using the HPC computing resources of INCD.

INCD Infraestrutura Nacional de Computação Distribuída

The Challenge

The challenge was the identification of drugs to prevent the interaction between the SARS-CoV-2 virus and the human cells with the screening of thousands of molecules using HPC resources, to identify and choose a set of molecules that could bind to the Spike-ACE2 complex of SARS-CoV-2 and evaluate how strong the association was aiming at experimental validation.

LIP

The Solution

To design drugs that can prevent the interaction between the virus protein and the human one, the BioSIM researchers created a model of this process, using computational resources from the Portuguese National Distributed Computing Infrastructure (INCD).

They used GPUs to study how molecular recognition occurs between the Spike protein of the virus and the ACE2 receptors. The model served as the basis for testing a large number of molecules collected from "virtual libraries", to check if some of them could prevent the virus from infecting human cells. From this huge list of molecules, it was important to find the ones that could bind to Spike-ACE2 complex and to see how strong this association was. To do this, additional CPU based processing was performed against the database of molecules. This enabled to assess, in atomic detail, the ability of each molecule to block or limit recognition between the two proteins.

The Scientific impact

Using the HPC resources provided by INCD, 200 thousand molecules were screened. Once the molecules that presented the greatest potential were identified, the information was made available to other research groups that collaborated with the laboratory. Those molecules were tested both in vitro and in cellular models. Two molecules with good capacity to inhibit the virus were identified resulting in giving origin to two patents.

BioSIM researchers are currently testing the two molecules in other variants and the preliminary results show that they have the power to prevent the spike protein from binding to human cells. If this is confirmed, the next step will be to conduct animal trials.

Benefits

According to Sérgio Sousa from BioSIM, *"these computational resources allowed us to make a ranking of molecules, from the least to the most promising,"* he explains.
"If we didn't have access to them, we would have tested a much smaller number of these molecules and it would mean much less chance of success".

- SARS-CoV-2
- COVID-19
- BioSIM,
- INCD

- Industry Sector: Biotechnology, Health Care, Pharmaceuticals

- Technology: HPC, HPDA

Contact:
Catarina Ferreira Gonçalves
Email: catarinafg@lip.pt

HPC implementation of a model of human brain visual system using **NEST Simulator**

The National Competence Centre of Bulgaria (NCC-Bulgaria) in the area of High-Performance Computing (HPC), High-Performance Data Analytics (HPDA) and Artificial Intelligence (AI) has the goal to enhance and develop the competences of the Bulgarian computational community, making full use of EuroHPC resources and the EuroCC partnership.
The NCC-Bulgaria is built by a consortium coordinated by the Institute of Information and Communication Technologies at the Bulgarian Academy of Sciences (IICT-BAS), and two members, Sofia University "St. Kliment Ohridski" (SU), and University of National and World Economy (UNWE). The three partners carry diverse technical and scientific background in the area of HPC and ICT in general, so as to ensure achievement of the project objectives and guarantee the overall success. The partners collaborate with Sofia Tech Park, where the Discoverer EuroHPC supercomputer is operating.

Scientific Partners Involved

Two departments from IICT-BAS were involved in this study: (1) the Scalable Computing and Applications (SCA) with HPC Centre and (2) The Information Technologies for Sensor Data Processing (ITSDP). The SCA manages the HPC centre and the main research is focused on the development and deployment of Cloud middleware and software components suitable for Cloud and HPC computing systems while ITSDP is focused on development and evaluation of algorithms and software libraries in the field of smart multisensing and artificial neural networks.

Technical Challenge

This collaborative research project funded by the Bulgarian Science Fund and performed by teams from the Institute of Information and Communication Technologies and the Institute of Neurobiology at the Bulgarian Academy of Sciences led to the development of a hierarchical model of the human brain visual system.

It was implemented using NEST Simulator – an open source platform running on desktop, HPC, or supercomputer architectures allowing parallel simulation of user-defined structures of spiking neural networks models of any size. The running time of such a model on a desktop PC increased dramatically with its dimension. Besides, the overall model included also a Python module whose running speed appeared to be the bottleneck in simulation investigations.

Although parallel implementation on a desktop PC with an 8-core processor using mpi4py allowed a significant drop in its run time, it still remained the slowest part of the model. Another challenge to be solved was related to merging of NEST module with mpi4py parallel simulation. Hence the need for more powerful computer architecture appeared obvious for further development and refinement of the overall model.

Solution

Since supercomputers offer much more powerful computing hardware than single processor desktop PCs, implementation of the model on supercomputer Avitohol appeared the proper solution to the challenge described above. Three algorithms were developed and tested:

- Algorithm 1: Python module parallelized into multiple processes.
- Algorithm 2: Python module parallelized via spawning within a single process.
- Algorithm 3: The combined run of NEST and Python modules in a parallel simulation using spawning.

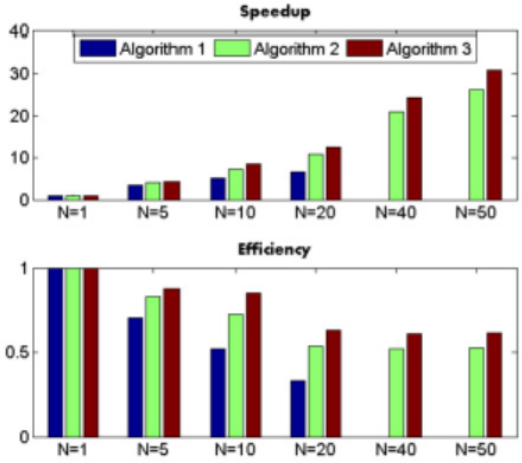


Figure 1: Speedup and efficiency of algorithms

Comparison of simulation times for different number of processes N and nodes

Table 1 Simulation times of all tested algorithms in hours:minutes:seconds format.

configuration	*N*	*nodes : total cores*	*Algorithm 1*	*Algorithm 2*	*Algorithm 3*
desktop	1	1:8	64:19:12		
desktop	5	1:8	3:36:36		
Avitohol	5	1:16	0:52:41	2:56:12	3:16:26
Avitohol	10	1:16	0:35:41	1:40:50	1:41:28
Avitohol	20	1:16	0:38:19	1:43:30	1:41:06
Avitohol	20	2:32	0:28:05	1:07:52	1:08:20
Avitohol	40	3:48		0:34:59	0:34:21
Avitohol	50	4:64			0:27:57

Scientific Impact

Running of the parallel Python module on the HPC facility of IICT – the supercomputer Avitohol – allowed usage of a much higher number of cores and thus a higher number of parallel processes, so a significant drop in its computation time was achieved. The problem arising from merging NEST and Python modules in a common MPI environment was overcome using a dynamic process management version of the Python module (Algorithm 3).

Thus, the overall model run time was significantly decreased allowing for its further application for in-silico investigation of the human visual system and brain structures involved in it by varying their parameters.

Benefits

- ✓ The team developed a pipeline for HPC implementation of spike timing neural network models of various brain structures.
- ✓ HPC implementation of such models will allow for a variety of simulation investigations of brain functioning.

- Supercomputer applications
- High-dimensional simulations
- Brain modelling

- Research Area: Neuroscience

- Technology: HPC, AI

Contact:
Prof. Petia Koprinkova-Hristova
Email: pkoprinkova@bas.bg
Assist. Prof. Simona Nedelcheva,
ITSDP Department, IICT-BAS

Assoc. Prof. Sofiya Ivanovska,
Email: sofia@parallel.bas.bg,

Mariya Durchova,
Email: mabs@parallel.bas.bg,
SCA Department with HPC Centre,
IICT-BAS



HPC cluster "Rocket" at the University of Tartu

PaRI – Nordic Pandemic Research Infrastructure

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

Organisations Involved

NeIC is hosted by NordForsk, which provides for and facilitates Nordic cooperation on research and research infrastructure across the Nordic region. The core partners of the PaRI project included University of Oslo, University of Bergen, Technical University of Denmark, Estonian Compute Infrastructure, German Network for Bioinformatics Infrastructure, and National Bioinformatics Infrastructure Sweden.

The Collaboration

NoPaRI brought together the needs, interests and resources to create e-infrastructure to support research excellence in the Nordic region. The scope of the project in spite of its focus on COVID-19 pandemic related research, is significantly wider including preparedness for all pandemic related research in the future. The project activities included genome-wide association studies, drug testing, viral sequencing and human viral gene expression.

Results of Collaboration

The PaRI team has successfully encouraged researchers and data producers to upload their data to the European Nucleotide Archive (ENA) — the primary European research database for nucleotide sequence information that includes raw sequencing data. This means that more data has been available for researchers studying the COVID-19 virus and its behaviour. Moreover, the Galaxy Nordic COVID-19 portal monitors the public output of viral genome sequencing projects that is submitted to the ENA and makes the data more accessible for further use. Yet another output of the PaRI project is the PaRI dashboard, a visualisation tool designed to help stakeholders such as epidemiologists, state institutions and researchers monitor the pandemic locally, down to regional and municipality levels.

Scientific Impact

The COVID-19 outbreak has induced an unprecedented near real-time exchange of information with an exceptional number of individual researchers performing multiple analyses in parallel, using publicly available, as well as their own, data. It has, however, also become clear that there is an unmet demand for services based on the real-time comparison of different epidemiological analyses of the same data, as well a major need for exchange of more person-sensitive information, such as geo-localisation, patient data, etc., between authorities, researchers and different countries. The ideal solution, described by the domain experts working with the data, will allow integration of patient-specific and population-wide levels of epidemiological data and combine it with sequencing and later serological information.

The success of the PaRI project is based on good collaboration between the participating partners. The NCC prioritised access to UT-Rocket for pandemic related workload. UT-Rocket is a HPC cluster with 6,000+ cores. It is also accessible via Galaxy web interface. In addition, it has extensions to allow processing of sensitive data.

Benefits

- ✔ Access to compute resources for analysing pandemic data.
- ✔ Setup and configuration of compute resources to accept workload from the user front-end.
- ✔ Streamlined, shared and up-to-date pandemic analysis tools and workflows (mainly as containers), with addition of a user-friendly front-end, Galaxy portal, wherever applicable.
- ✔ Help Nordic users to publish pandemic research data in public databases in a FAIR manner, thereby offering better reusability and more citations.
- ✔ Synergies with other initiatives in timely addressing of the pandemics.

- Data processing,
- FAIR data,
- E-Infrastructure,
- Bioinformatics,
- Galaxy

- Industry Sector: Biotechnology, Pharmaceuticals

- Technology: HPC, HPDA

Contact:
Ülar Allas
Email: ylar.allas@ut.ee

Launch of the Vaccination Centre

The Czech National Competence Centre (NCC Czechia)for High-Performance Computing (HPC) and Data Analysis (HPDA) is represented by IT4Innovations National Supercomputing Center at VSB – Technical University of Ostrava. Its mission is to analyse, implement, and coordinate all necessary activities and offer end users its services to meet their needs: from access to supercomputers and technology consulting to providing training for indutry, public administration, and academia.

Organisations Involved

The University Hospital Ostrava is a leading healthcare provider in the Czech Republic. It offers specialised medical care in the Moravian-Silesian Region. In cooperation with the Medical Faculty of Ostrava University, it educates young doctors and other healthcare professionals.

The Regional Authority of the Moravian-Silesian Region is one of the regional bodies. It is a regional administrative government and, at the same time, part of the State Administration of the Czech Republic. It has 18 departments and one independent division. The Regional Authority of the Moravian-Silesian Region is located in Ostrava.

The metropolis of the Moravian-Silesian region is the 2nd largest city in the Czech Republic and has the third-largest population. The statutory city of Ostrava is implementing a range of projects, creating an energised town that offers a high quality of life for all generations and attracts talented people to live and work here.

The Faculty of Medicine of the University of Ostrava is the youngest medical faculty in the Czech Republic. The Faculty ranks among Central Europe's most modern teaching and research centres for medical and non-medical fields. Currently, the faculty has around 1700 students in 27 degree programmes, with over 400 students graduating every year.

The Challenge

During the Covid-19 pandemic the Czech National Competence Centre in HPC was invited to participate in a project to build a large-scale vaccination centre at Černá louka in Ostrava. The temporary centre became operational in March 2021.

During the planning of the vaccination centre, many different aspects had to be considered, such as personnel, space organisation, number of other stations, etc. The main point was to maximise the number of vaccinated people in one day while keeping the minimum critical issues and number of employees. IT4Innovations was asked to find an optimal configuration of individual stages to maximise the efficiency of the vaccination centre.



Photo : A consultation before the vaccination process // Beáta Kapošváry, University Hospital Ostrava

The Solution

To find the optimal solution which can be easily presented to an audience of all technical levels, we created a simulation of the people flow through the vaccination centre. In this simulation, the centre is modelled as a directed graph through which people must pass.

Such representation was chosen because it is easily modified at any simulation step. It can create different scenarios, such as people going for lunch, a blackout, or failure at a given station. The developed simulation made it possible to determine the required number of operators and staff positions for each station. The simulation shows how many people can be in the vaccination centre at a particular time, at different stages such as registration, vaccination itself, or the waiting room after immunisation.

The Business Impact

The Czech National Competence Centre in HPC contributed to building a large-scale vaccination centre in Ostrava. Other partners in this unique and urgent project included the Moravian-Silesian Region, the Statutory City of Ostrava, the Faculty of Medicine of the University of Ostrava, and the University Hospital Ostrava.

The team led by Tomáš Martinovič from IT4Innovations created a simulation of the passage through the vaccination centre, making it possible to detect critical points. In the original design, these were the waiting room after vaccination and the printing of the confirmation at the exit. Even „little things" such as lunch breaks for staff were considered. To minimise the number of staff, rotations of support staff during holidays were prepared, with individuals taking turns at each post.

The supporting simulation the scientists at IT4Innovations developed was a significant element in designing a large-scale vaccination centre. The centre's capacity testing would have taken much longer without the simulation. Based on the simulation, it was found that the centre can vaccinate up to 2,600 people a day.



Photo: A registration process in the vaccination centre // Beáta Kapošváry, University Hospital Ostrava

The Czech National Competence Centre in HPC contributed to building a large-scale vaccination centre in Ostrava. Other partners in this unique and urgent project included the Moravian-Silesian Region, the Statutory City of Ostrava, the Faculty of Medicine of the University of Ostrava, and the University Hospital Ostrava.

The team led by Tomáš Martinovič from IT4Innovations created a simulation of the passage through the vaccination centre, making it possible to detect critical points. In the original design, these were the waiting room after vaccination and the printing of the confirmation at the exit. Even „little things" such as lunch breaks for staff were considered. To minimise the number of staff, rotations of support staff during holidays were prepared, with individuals taking turns at each post.

The supporting simulation the scientists at IT4Innovations developed was a significant element in designing a large-scale vaccination centre. The centre's capacity testing would have taken much longer without the simulation. Based on the simulation, it was found that the centre can vaccinate up to 2,600 people a day.

Benefits

- ✔ Vaccination, according to a simulation prepared by scientists at IT4Innovations, can be administered to up to 2,600 people per day.
- ✔ The simulation helps to organise the centre's operation and minimise the number of staff.
- ✔ The simulation significantly helped with the centre's testing and saved a considerable amount of time.

- Healthcare,
- Vaccination centre
- Simulation

- Industry Sector: Vaccination

- Technology: Simulation

Contact:
Tomas Martinovic
Email: tomas.martinovic@vsb.cz

Medical Image Processing as a Service

The Czech National Competence Centre (NCC Czechia)for High-Performance Computing (HPC) and Data Analysis (HPDA) is represented by IT4Innovations National Supercomputing Center at VSB – Technical University of Ostrava. Its mission is to analyse, implement, and coordinate all necessary activities and offer end users its services to meet their needs: from access to supercomputers and technology consulting to providing training for industry, public administration, and academia.

EURO CZECHIA

Organizations Involved

FAKULTNÍ NEMOCNICE OSTRAVA

University Hospital Ostrava is a state-funded organisation established in 1912 by the Ministry of Health of the Czech Republic. The primary purpose of this organisation is to provide health services. Other activities include the development of science and research. The organisation conducts both basic and applied research.

Technical Challenge

The main goal of cooperation between NCC and University Hospital Ostrava was to deploy and test a tool providing remote automatic tissue segmentation from patient image data obtained from computed tomography (CT) or magnetic resonance imaging (MRI) on HPC infrastructure at IT4Innovations. The methods used to enable tissue segmentation were based on deep learning (DL). The objectives were...

...to provide a service based on state-of-the-art algorithms for automatic segmentation of the desired tissues as an AI-based annotation service and

...to collect the data after the automatic segmentation and validation by medical doctors, and to provide HPC-based training of new models or enhancement of the existing models through fine-tuning.

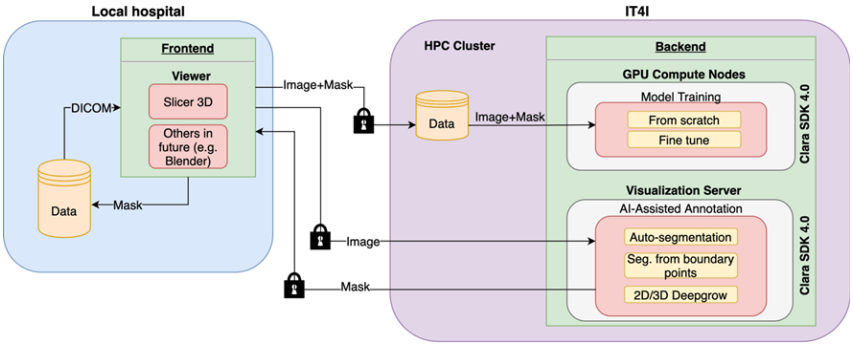


Figure 1: The main concept of the tool vfor medical image processing and analysis

Solution

The evaluated concept consists of two main components: the first component runs at a medical doctor's site in the hospital, and the other one operates at IT4Innovations National Supercomputing Center. The part at the hospital is represented by a frontend that mediates the interaction between the doctor and the data.

An open-source 3D Slicer was chosen to provide this functionality, and it facilitates the interconnection to a backend part. The backend part provides the computational power of an HPC cluster and other required features. It enables the training of deep learning models for automatic segmentation and model inference of incoming image data. The whole concept evaluated by NCC researchers have adopted the NVIDIA's Clara Train SDK to create the backend part.

Business Impact

Analysis of the patient´s image data is used to establish or refine the patient´s diagnosis and is performed by medical doctors. When detailed data analysis such as image segmentation is required, e.g., to prepare a patient for surgery, the process is lengthy and demands the physician´s attention given that it is often entirely or largely manual. Consequently, there is a high demand for supporting tools that enable the automation of this radiological examination, especially when the quality evolves with each additional case analysed.

Patients and clinicians can benefit from the toolkit since automation provides a higher level of health care.

Benefits

- ✔ The time of the segmentation process is reduced by automation.
- ✔ Spared time can be used for the physician´s benefit.
- ✔ The automated segmentation process can be applied to specific tissues of interest to the physician.

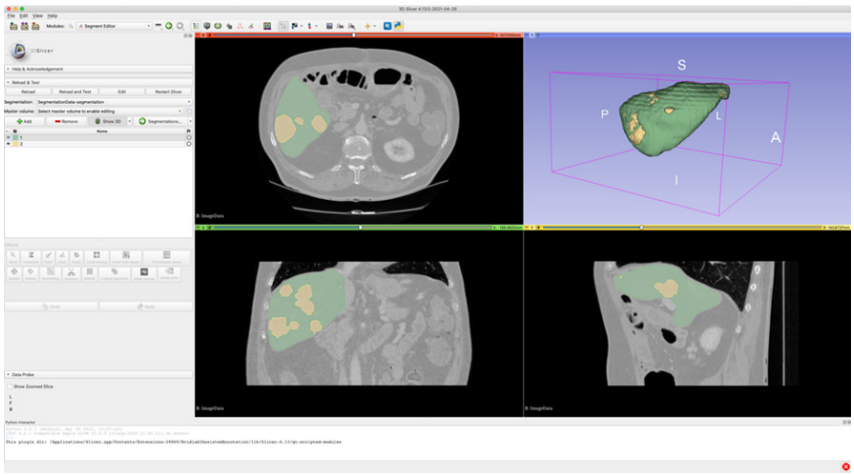


Figure 2: 3D Slicer working environment with possible output obtained by automatic segmentation on HPC cluster

- Automatic Segmentation,
- Deep Learning (DL),
- Computed Tomography (CT),
- MRI
- HPC

- Industry Sector: Health, Surgery, Medical Sector

- Technology: Information Technology, Deep Learning (DL), HPC

Contact:
Tomas Karasek
Email: tomas.karasek@vsb.cz

Numerical Simulation of Exoskeleton

The Czech National Competence Centre (NCC Czechia) for High-Performance Computing (HPC) and Data Analysis (HPDA) is represented by IT4Innovations National Supercomputing Center at VSB – Technical University of Ostrava. Its mission is to analyse, implement, and coordinate all necessary activities and offer end users its services to meet their needs: from access to supercomputers and technology consulting to providing training for industry, public administration, and academia.

EURO CZECHIA

Organisations Involved

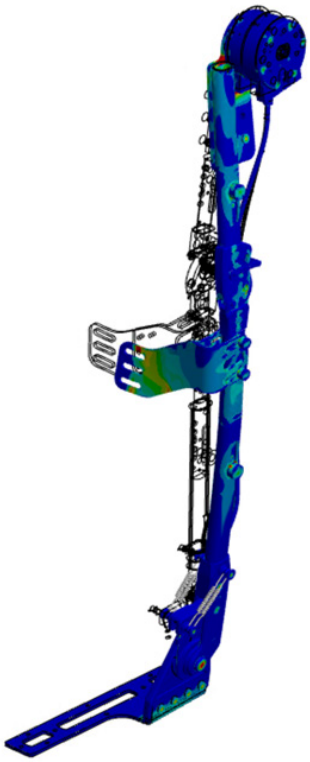
MEBSTER

The MEBSTER is a research & development company and through its cost-effective assistive devices...
...it supports medical care providers to improve and streamline their services for their clients with mobility disorders.
...it aids individual clients with mobility disorders to build healthy habits and motivate them to actively improve their quality of life.

It is an international team of engineers, doctors, and physiotherapists based in the Czech Republic in 2018. They are committed to intensively developing innovative assistive technologies and medical devices for rehabilitation spanning over six years. They design their devices directly with users, and their research is solely patient-centred. They rigorously test all products to ensure they are simple and as comfortable as possible. In addition, they collaborate with a team of world experts to set trends in their industry.

Technical Challenge

The objective of this proof-of-concept was to demonstrate the use of numerical modelling and simulation in the design process of an innovative UNILEXA exoskeleton for gait assistance designed for people with partial or complete loss of lower limb



Numerical model of exoskeleton

Solution

The consultancy provided by the NCC consisted in the revision of the methodology used by the Mebster company for the numerical modelling and simulation of the exoskeleton.

The computational model used was based on the finite element method (FEM) and the exoskeleton assembly was simulated with the inclusion of effects resulting in the strength/stiffness of the system. Since the exoskeleton assembly is mathematically a complex nonlinear problem with a wide variation of boundary conditions, the use of HPC is necessary to solve such a complex numerical model. In this case NCC showcased to the company advantages of the HPC in the term of the possibility of solving much complex models which are not solvable using conventional personal computers used in the company, by running simulations of the exoskeleton on HPC infrastructure.

The ANSYS software was used to create the computational model and numerical simulation. A simplified model of the part of the UNILEXA exoskeleton assembly was created to estimate the required HPC resources. The four nodes of the Barbora system were used to solve a complex numerical model, including nonlinearities such as contact interfaces with friction, large displacements, etc.

Business Impact

The UNILEXA exoskeleton is a risk class I medical device. The placing of similar products on the market is subject to strict regulations under a new EU regulation known as the Medical Device Regulations (MDR). Each product must be validated through simulations and tests to verify and guarantee its safety and usability for patients. Testing medical devices are very costly and time-consuming, hence the use of simulations. Thanks to the detailed numerical simulation, it will be possible to repeatedly ascertain and verify the safety and effectiveness of the existing product and its other versions and variants. In addition, the simulation will help to reveal optimisation opportunities in terms of weight, usability and cost for the production and servicing of the UNILEXA exoskeleton.

The consequence of the use of numerical simulation will be, in particular, an increase in the quality of care provided to the end user with an emphasis on user comfort not only for the client (patient) but also for the medical staff using the technology, as well as an increase in the competitiveness of the product on the market.

Benefits

- ✓ Possibility of verification the safety and effectiveness of the existing product and its further versions and variants
- ✓ Increase in the quality of care provided to the end user with an emphasis on user´s comfort
- ✓ Increase in the competitiveness of the product on the market.



Picture of the exoskeleton prototype

- Finite Element Method (FEM)
- exoskeleton, contact problems,
- bolt pretension

- Industry Sector: Medical Devices

- Technology: HPC, CSD

Contact:
Tomas Karasek
Email: tomas.karasek@vsb.cz

AI Medical Device Software
to quantify brain damage and clinical prognosis

The French NCC: CC-FR, dedicated to HPC, HPDA and AI technologies, 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.

Organisations Involved

CC-FR, set up within the framework of EuroHPC, is managed by the European Pole of Competence in high performance digital simulation Teratec, in association with the European Center for Advanced Research and Training in Scientific Computing Cerfacs and with the participation of the Big National Equipment Intensive Computing Genci.

Braintale, www.braintale.eu, is an innovative company opening a new era in medicine by providing physicians with clinically validated prognostic solutions for the management of brain injured patients through sensitive and reliable measurements of brain white-matter microstructure alterations. Building on more than 15 years in clinical development, BrainTale's products are developed collaboratively to address medical need and fulfil expectations of healthcare professionals and patients alike. Since its inception in 2018, the company has set up a complete quality management system and is now ISO:13485:2016 certified, with a suite of products available on the European market.

The Challenge

Since Fall 2021, Braintale has integrated the CC-FR tailored program as they needed to benefit from high-level support in intensive computing. Braintale needed to analyse under efficient conditions the data of more than 500 patients in order to develop their research tools to optimize brain diseases studies. They also wanted to optimize the retrospective analysis of large research cohorts that require intensive processing of very large volumes of data.

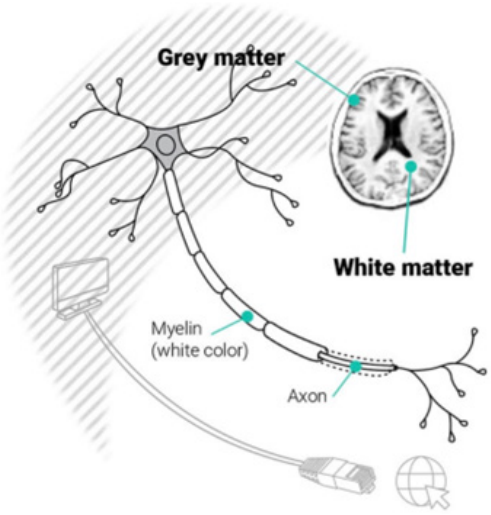


Figure 1: Concept of white matter

The Solution

In collaboration with CRIANN (Regional Numerical Centre and Digital Applications of Normandy), Braintale was able to access to expertise and to the MYRIA supercomputer, equipped with more than 10,000 computer cores very high-speed, and capable of executing nearly 600,000 billion transactions per second.

The Business Impact

TheThe use of the MYRIA supercomputer and the support of CC-FR allowed Braintale to develop and produce its portfolio of AI-based digital medical device and products that aim to improve the care of patients with neurological diseases and in intensive care

Benefits

The power and agility of MYRIA supercomputer combined with the company's servers allow it to envisage a rapid commercial deployment of its solutions while increasing its product portfolio.

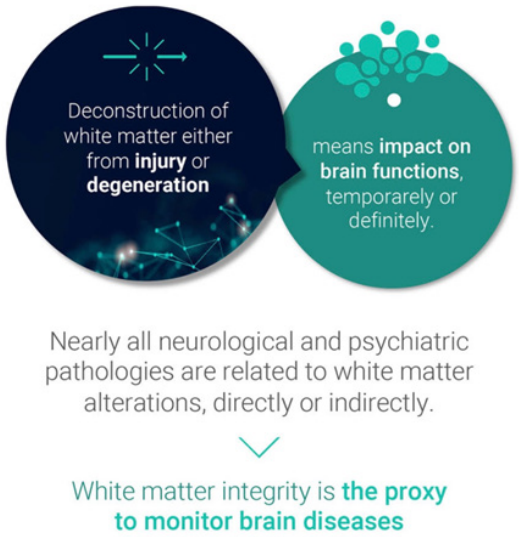


Figure 2: braintale objectives

- HPC, AI
- Algorith
- Medical Devices
- Health care
- HPDA

- Industry Sector: Pharmaceutical

- Technology: HPC, HPDA, AI

Contact:
Dr Karim Azoum
LinkedIn: https://www.linkedin.com/in/karim-azoum-45011710a
Email: Karim.azoum@teratec.fr
+33 7 62 74 03 60

Supercomputer Applications in Biomedical Engineering

The National Competence Centre of Bulgaria (NCC-Bulgaria) in the area of High-Performance Computing (HPC), High-Performance Data Analytics (HPDA) and Artificial Intelligence (AI) has the goal to enhance and develop the competences of the Bulgarian computational community, making full use of EuroHPC resources and the EuroCC partnership.
The NCC-Bulgaria was built by a consortium coordinated by the Institute of Information and Communication Technologies at the Bulgarian Academy of Sciences (IICT-BAS), and two members, Sofia University "St. Kliment Ohridski" (SU), and University of National and World Economy (UNWE). The three partners carry diverse technical and scientific background in the area of HPC and ICT in general, so as to ensure achievement of the project objectives and guarantee the overall success.

Organisations Involved

AMET Ltd. is a company dedicated to development, modern manufacturing and distribution of electronic medical equipment and modules. It is a reliable and desired partner in both Bulgarian and foreign market. The partnership with the Department of Scientific Computations with Laboratory on 3D Digitization and Microstructure Analysis at IICT-BAS is directly based on the development of computer models and their efficient implementation on HPC systems.

The results are new products or/and upgraded ones including smart electrosurgical instruments and physiotherapy equipment.

Technical Challenge

One challenge was to develop a supercomputer model of the processes of Radio-Frequency (RF) Liver Ablation (a treatment based on RF alternating current flow, delivered by a needle-like probe, that destroys the undesired tumor tissue via heating). Up to billions unknowns with respect to space variables and thousands of time steps are needed to resolve the complex geometry.

Another challenge was to num rically analyse the temperature field within the impact region of a portable device for contact-free ticks' removal. Such a device decreases the probability of blood infection. A computer model was developed that took into account the multiple skin layers and the thin film with respect to the applied gel.

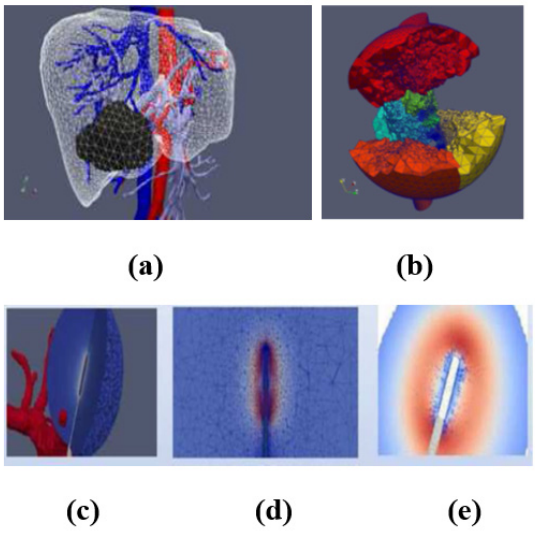


Figure 1: Radio-frequency liver ablation simulations: (a) Triangulated liver with a tumor inside via Netgen; (b) Domain decomposition via ParMETIS; (c) Visualization of the bipolar probe within the tumor; (d) The computational domain; (e) Visualization of the results: isolines on the temperature field

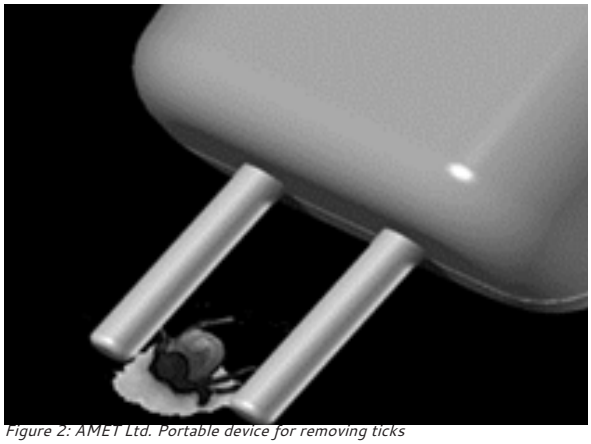


Figure 2: AMET Ltd. Portable device for removing ticks

Solution

The parallel realization of the developed computer models involves integration of open-access parallel software packages, such as: Netgen, ParMETIS, Hypre, BoomerAMG, and ParaView. It is built upon a computational model with distributed memory, that allows for solving large-scale problems with billions unknowns. The realized models are strongly coupled. This determines the key role of the communications realized by MPI.
For the parallel implementation of the model, the computational domain needs to be decomposed into various sub-regions. It is shown that for the considered class of problems 1D and/or 2D domain decomposition (DD) leads to highly restricted parallelism with parallel efficiency dropping even below 1%. Thus, the application of ParMETIS for efficient 3D is vital.

A complete 6-minute process of RF liver ablation with a bipolar probe has been simulated, taking into account the blood circulation in the whole body.

Business Impact

High-frequency electro surgery is one of the most commonly used in practice surgical technique. Today this method is present in every surgery in the world. One of the specific areas of impact of the developed HPC solutions is the radio-frequency tumor ablation. The aim of the joint work with AMET Ltd is to respond to the ever-increasing need for more precise intervention, both for intervention in hardly accessible places and for maximally low-invasive intervention.

The developed software tools for supercomputer simulation of coupled physical processes of radiofrequency electro surgery manipulations are beyond the scope of available commercial packages.

Measurable indicators are applied to assess the reliability of the obtained results, thus providing proven criteria for minimizing subjective inaccuracies. The impact of using large-scale HPC models reached more than two times improved precision of evaluating the volume of effectively ablated tissue.
The business impact for the industrial partner includes...
...Improving the technology characteristics of existing products (e.g., hepatic tumor ablation);
...Development of new products (e.g., contactless removal of blood-sucking parasites).
...an increase in the market share of AMET Ltd in the field of the studied electrosurgical technologies has been achieved.

Benefits

- ✓ Fully realistic computer simulation of strongly coupled processes of RF electrosurgical technologies
- ✓ Time/cost saving of parameter optimization of high tech low-invasive procedures
- ✓ Assessment of complex processes hardly assessable for medical imaging

- Supercomputer Applications
- Smart Electrosurgical Instruments
- Physiotherapy Equipment
- Contact-Free Tick Removal

- Industry Sector: Pharmaceuticals

- Technology: HPC, AI, Big Data

Contact:
Prof. Svetozar Margenov
Prof. Stanislav Harizanov,
Email: sharizanov@parallel.bas.bg
Institute of Information and Communication Technologies,
Bulgarian Academy of Sciences

Section 7

Mobility and Logistics

Supercomputer Applications in Hydrodynamics – 132

Autonomous Navigation at Sea – 134

Public Transport Analysis on HPC Infrastructure – 136

Traffic Events and Alerts Data – 138

RISE Accesses VEGA – 140



Photo: Sawyer Bengtson // Unsplash



Photo: Ibrahim Boran // Unsplash

Supercomputer Applications in Hydrodynamics

The National Competence Centre of Bulgaria (NCC-Bulgaria in the area of HPC, HPDA and AI has the goal to enhance and develop the competences of the Bulgarian computational community, making full use of EuroHPC resources and the EuroCC partnership.

NCC-Bulgaria is built by a consortium coordinated by the Institute of Information and Communication Technologies at the Bulgarian Academy of Sciences (IICT-BAS), and two members, Sofia University "St. Kliment Ohridski" (SU), and University of National and World Economy (UNWE). The partners collaborate with Sofia Tech Park, where the Discoverer EuroHPC supercomputer is operating.

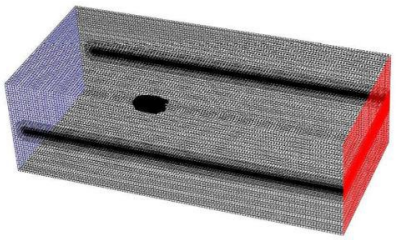
Organisations Involved

The department of Scalable Computing and Applications (SCA) with HPC Centre is part of IICT-BAS. It manages the HPC centre with the supercomputer Avitohol. The department is focused on the development and deployment of Cloud middleware and software components, methods, algorithms, and applications suitable for Cloud and HPC computing systems.

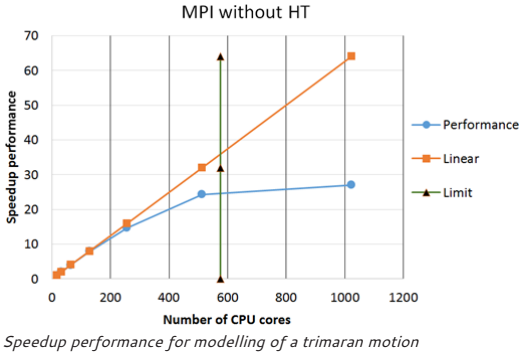
The Bulgarian Ship Hydrodynamics Centre (BSHC) of the Bulgarian Academy of Sciences (IMS-BAS) is a national research and development centre. BSHC's mission is to perform fundamental and applied research and train graduate and post-graduate students in the areas of ship hydrodynamics; environmental protection; sea/river energy utilization; national security and defense.

Technical Challenge

The subject of this investigation is a simulation of the motion of a trimaran/ship in multiphase flow conditions – which is one of the most complicated problems in the ship hydrodynamics. One of the challenges is the choice of the number of cells in the generated fine mesh. This number is important in order to receive an accurate simulation. Another challenge is the pre-processing in which grid cells are split proportionally so that simulations can be performed on different CPU cores in order to improve scalability.



Generation of the mesh, the total number of cells 9 631 050



Speedup performance for modelling of a trimaran motion

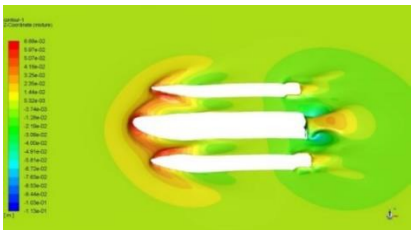
Solution

The mathematical model includes Navier-Stokes equations. We apply the finite element method using OpenFoam CFD software to compute the solution.
A given geometry of the trimaran/ship and the corresponding number of cells in the network is considered. With different geometries, the ship „behaves" differently in calm waters and in the presence of multiphase flow (wind/waves).

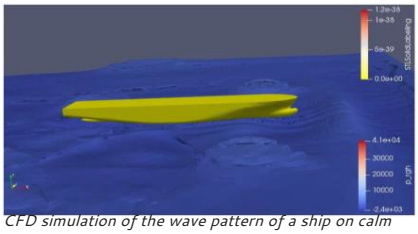
The Rhino 5.0 product is used for the 3D model generation. The total number of cells for generation of the mesh is 9 631 050. To obtain a 3-minute simulation of the ship in calm waters, we need computer time of up to 1-2 weeks using a single server. That is why the team had to distribute the calculations in parallel using MPI on several computational nodes in order to obtain a complete simulation.



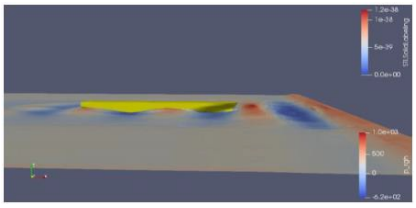
CFD simulation of the wave pattern of a trimaran



CFD simulation of the interactions between the trimaran hulls



CFD simulation of the wave pattern of a ship on calm water



CFD simulation of the wave pattern of a ship on multiphase flow

Business Impact

Results show a good parallel efficiency up to 4 and 8 nodes. Using a supercomputer Avitohol, the computation time was reduced from a week to a day to simulate a 1-minute video of ship/trimaran motion in the presence of waves and wind. The parallel CFD simulation tests for modelling a trimaran were done to demonstrate the parallel performance of the OpenFoam software and to define efficient usage of Avitohol resources based on the introduced scalability criterion.

The business impact: (i) Study of ship hull elements to decrease drag and regulate the ship's turbulent flow. (ii) Development of new hull elements with simple and scalable solutions that reduces fuel consumption and environmental impact.

Benefits

- ✓ Reduction of simulation time by using parallel computations;
- ✓ Improving the stability of a ship/trimaran in calm waters or in multiphase flow (presence of winds and waves);
- ✓ Study of ship hull elements to decrease drag and regulate the ship's turbulent flow.

- Supercomuting Applications
- Hydrodynamics
- Computational Fluid Dynamics (CFD)
- Simulation

- Industry Sector: Shipbuilding, Maritime Industry

- Technology: HPC, HPDI

Contact:
Prof. Todor Gurov,
Email: gurov@bas.bg
Vyara Koleva-Efremova,
Email: viarakoleva@yahoo.com
Institute of Information and Communication Technologies at the Bulgarian Academy of Sciences

Dr. Grigor Nikolov
Email: g.nikolov@bshc.bg
Dr. Dobrin Efremov
Email: d.efremov@bshc.bg
Bulgarian Ship Hydrodynamics Centre (BSHC-BAS), Varna, Bulgarian Academy of Sciences

Autonomous Navigation at Sea

EuroCC National Competence Centre Finland
The Competence Centre Finland will heavily focus on paving the users' way on LUMI, the EuroHPC flagship supercomputer located in Finland.
LUMI is one of the key elements in Finnish and European HPC competence development.

Industrial organisations Involved

Groke Technologies https://www.groke-tech.com/ is a Finnish company that focuses on developing intelligent methods for autonomous navigation to improve safety at sea.

The Challenge

High-performance computing (HPC) plays a vital role in optimizing machine-learning models for computer vision. Detection of all vessels and dangers around navigating personnel is a challenging task on-board any vessel, the company Groke is creating an awareness system that monitors the surroundings along with the watch keepers - giving them peace of mind even in a challenging environment. The technical solution combines radar technology, Automatic Identification System (AIS), visual and thermal cameras, as well as sea charts, Inertial Measurement Units (IMU) and dual band GNSS.

At the moment, AI and machine learning technology are mostly used to process images from a highly specialized camera system that consists of a 225-degree visual camera and a 180-degree thermal camera. By detecting and identifying objects around a vessel from the side of another vessel to a far-away sea buoy, it helps to improve the captains' situational awareness and perception of their vessels' surroundings.



Figure 1: Groke Technologies is developing intelligent methods for autonomous navigation to improve safety at sea. High-performance computing (HPC) plays a vital role in optimizing machine-learning models for computer vision. Copyright: Groke Technologies

Video: https://www.youtube.com/watch?v=zEFYrFl67xo&t=5s

The Solution

With the financial support from Business Finland´s AI Business programme and computing resources provided through NCC Finland, Groke Technologies is training their machine learning models and investigating how to build deep neural networks for object detection with the optimal trade-off between accuracy and performance.

- Building scalable data processing pipelines for training,
- Evaluation and optimization of machine learning algorithms.
- Run large scale training on GPU equipped HPC system.

The Business impact

Due to the strategic investment from Mitsubishi Corporation, Groke technologies' solution will be initially tailored to meet the needs of the Japanese seafaring where it is expected to alleviate the inevitable shortage of sea captains. Japanese vessels are not allowed to sail under Japanese flag unless the captains´ nationality is Japanese, and many of the local captains are already in the later stages of their careers. The data-driven high-technology navigation solution is expected to draw younger generations into the maritime industry.

Benefits

Improved speed of training resulting in lower unit cost and ultimately improved object detection:

- ✔ Improved detection increases safety

Autonomous navigation improves:

- ✔ Time to destination
- ✔ Savings in personnel and energy expenditures

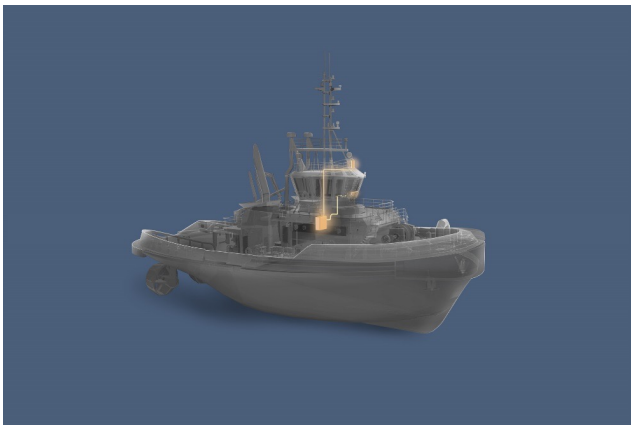


Figure 2: Preferred installation location for sensor unit on tug boats is either mast or on compass deck above bridge. Copyright: Groke Technologies

- HPC
- AI
- Neural Net Training
- Autonomous Navigation
- Fully Automated Vessels

- Industry Sector: Maritime Technology: HPC, AI

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

Public Transport Analysis on HPC Infrastructure

The Turkish NCC is led by TUBITAK-TRUBA, one of the two main HPC centers in the country. Besides TRUBA, three partners, Middle East Technical University (METU), Sabanci University (SU), and National Center for High-Performance Computing (UHeM) of Istanbul Technical University (ITU), act as third- party. METU is a public university based in Ankara and SU is a private one located in Istanbul. ITU UHeM, also based in Istanbul, provides supercomputing and data storage services for academic and industrial users. Our skills span HPC, HPDA, AI, CUDA, Materials Science, Computational Fluid Dynamics, and several other areas. This achievement is presented by the third-party METU.

Industrial organisation Involved

This success story results from the 1st case study jointly led by METU and TRUBA in cooperation with the SME Parabol:
It (paraboly.com) has been conducting R&D activities in Intelligent Transportation Systems since 2011 and developing a public transportation analysis platform (https://tr.cermoni.app) to analyze public transport (PT) demand in a city and to make more efficient PT investment decisions.
By merging the company's main areas of expertise (big data analysis, smart mobility algorithms, mobility management, cloud computing) with R&D activities, they have been able to provide mobility management and analysis software tools in several countries and more than 50 cities in three continents willing to improve various modes of transportation such as PT mobility.

Technical Challenge

Parabol has been developing a PT analysis platform (cermoni.app) to analyze the passengers' boarding data and the vehicles' GPS location data collected from a PT system. Currently, it performs spatio-temporal processing and analysis on the PT data, around 10M rows per month. Moreover, the PT accessibility analysis should be performed by finding activity regions for each commuter. Creating value from such big data to understand the mobility patterns involves 3D (space, time, and commuter) analysis which is time-consuming and is challenging. Hence, the spatiotemporal clustering algorithm for PT user activities (STCAPT) needs parallel execution.

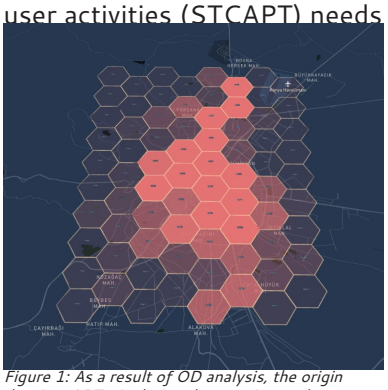


Figure 1: As a result of OD analysis, the origin density of PT trip during the morning peak.

Solution

Considering the fact that STCAPT requires a large amount of memory to run, which an HPC environment can provide, we worked on developing an approach to run STCAPT on HPC. The algorithmic approach followed for parallelizing STCAPT contains first grouping the data by passenger id and then finding regions for each passenger.

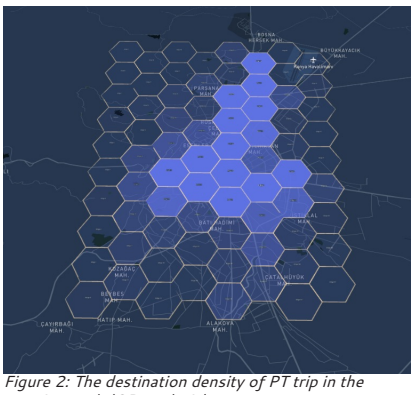


Figure 2: The destination density of PT trip in the morning peak (OD analysis)

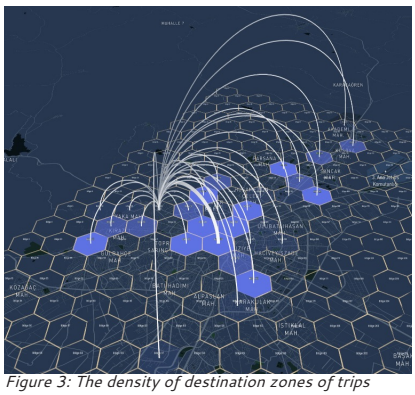


Figure 3: The density of destination zones of trips starting from the selected region.

We first developed an Apache Beam pipeline and deployed it on the Spark runner for its execution on HPC. However, since the execution of the Beam model failed, the pipeline was then ported to Apache Spark and executed successfully. Here, we developed two different Spark pipelines. The first one used Spark transforms and custom map operations on RDDs to generate passenger commute sequences. The next one used SparkSQL to aggregate those sequences to find an Origin-Destination matrix for some given grid. The output of the first pipeline is used as an input for the second one because showing outputs of the first on a map is not practical or useful. All these steps were guided by the METU team and supported by the TRUBA team, who installed and managed the Apache Spark platform on the HPC environment.

While the analysis cannot be completed reasonably on a typical server within the company data center, it took just minutes to run it with a test dataset on TRUBA. Being able to use an HPC infrastructure for this experiment gave us very impressive results. On a Spark cluster created with only three nodes of TRUBA, the test runs resulted in a 90% decrease in runtime.

Business Impact

Advanced algorithms have already been developed to provide more comprehensive analyses, such as Origin-Destination analysis to understand mobility patterns as a unique, high-tech decision support mechanism. The TRUBA HPC environment helps us reach this aim with highly improved performance. It is now possible to run these algorithms and take the results in a reasonable amount of time. Thus, such algorithms with parallel execution on HPC allow working with large-scale networks for NP-hard optimization issues. This creates a really fast, highly accurate, less costly, and more effective decision support mechanism not only for strategic or tactical but also for operational decisions. It gives our algorithms and hence the company a competitive advantage in the market.

Benefits

- ✔ Ability to work in the HPC environment, especially to run Apache Spark pipelines.
- ✔ Understanding urban public transport mobility patterns much better.
- ✔ Saving time in getting results for complex optimization problems and operational decisions with highly-increased performance even for large-scale cities.
- ✔ Increasing the efficiency of public transport systems and accessing better public transport service levels.

- Apache Spark
- Big Data
- HPDA
- Data Analysis
- Smart City

- Industry Sector: Smart City, Intelligent Transportation System, Origin-Destination, Public Transportation

- Technology: Big Data, HPC, HPDA

Contact:
Metin BARIŞ -Parabol,
Email: metin.baris@paraboly.com
Cevat ŞENER -METU,
Email: sener@ceng.metu.edu.tr

Timetable Optimization in Public Transport

The Turkish NCC is led by TUBITAK-TRUBA, one of the two main HPC centers in the country. Besides TRUBA, three partners, Middle East Technical University (METU), Sabanci University (SU), and National Center for High-Performance Computing (UHeM) of Istanbul Technical University (ITU), act as third- party.

METU is a public university based in Ankara and SU is a private one located in Istanbul. ITU UHeM, also based in Istanbul, provides supercomputing and data storage services for academic and industrial users. Our skills span HPC, HPDA, AI, CUDA, Materials Science, Computational Fluid Dynamics, and several other areas. This achievement is presented by the third-party METU.

Industrial organisation Involved

This success story results from the 2nd case study jointly led by METU and TRUBA in cooperation with the SME Parabol:
It (paraboly.com) has been conducting R&D activities in Intelligent Transportation Systems since 2011 and developing a public transportation analysis platform (https://tr.cermoni.app) to analyze public transport (PT) demand in a city and to make more efficient PT investment decisions. By merging the company's main areas of expertise (big data analysis, smart mobility algorithms, mobility management, cloud computing) with R&D activities, they have been able to provide mobility management and analysis software tools in several countries and more than 50 cities in three continents willing to improve various modes of transportation such as PT mobility.

Technical Challenge

Utilizing the TRUBA environment enabled Parabol to have a faster R&D cycle and continue their first case study related to origin-destination analysis. They have used those results to get optimal timetables for each route governed by different terminals.

However, optimizing Public Transport services and exploring passenger mobility requires handling large amounts of data, solving complex combinatorial optimization problems, and dealing with uncertainty. Solving this NP-hard problem requires a significant amount of time, and even impossible to get a solution in a reasonable time as the problem space gets larger.



Solution

The initial heuristic algorithm, dealing with all routes iteratively to get a feasible solution in the company's production environment, was first extended to achieve the global optimum solution. After redefining the algorithm, a Dask cluster was set up on TRUBA to parallelize the optimization algorithm so that each instance solves the timetables of the bus routes originating from a separate regional terminal.

Hence, the Dask cluster size in this data-parallelization approach will be proportional to the number of regional terminals involved, which is generally also proportional to the number of bus routes. Parabol chose Dask to parallelize their algorithm as it requires no code modification in the original algorithm, providing a very time-efficient solution.

When they ran the optimization algorithm on TRUBA with the actual PT data, they achieved an optimal solution in minutes to a couple of hours, depending on the problem size. They also observed that a near-optimal solution could be reached much earlier.

HOCAFAKIH TIP FAKÜLTESİ ANASULTAN	MERAM YAKA ANASULTAN	MERAM YENİYOL DÖRTOKKA
06:00	06:00	06:00
06:05	06:06	06:03
06:10	06:12	06:06
06:15	06:18	06:09
06:20	06:24	06:12
06:25	06:30	06:15
06:30	06:36	06:18
06:35	06:42	06:21
06:40	06:48	06:24
06:45	06:54	06:27
06:50	07:00	06:30
06:55	07:06	06:33
07:00	07:12	06:36
07:05	07:18	06:39
07:10	07:24	06:42
07:15	07:30	06:45
07:20	07:36	06:48
07:25	07:42	06:51
07:30	07:48	06:54
07:35	07:54	06:57

Figure 1: An Optimized Timetable

Business Impact

The Dask-based solution developed in this case study can provide optimized timetables for bus routes to increase fuel savings and passenger satisfaction, as well as improve timetable governance. A sample timetable optimized using the algorithm developed is shown in Figure 1. Using the TRUBA HPC environment allows the company to achieve this goal with significantly improved performance. In addition to this high-tech solution allowing Parabol to work with large-scale PT networks, the parallelization framework they had here can be inspired and easily applied to other complex optimization problems in the PT domain.

All these create a very fast, highly accurate, less expensive, and more efficient decision support mechanism, which gives a competitive advantage in the market.

Benefits

- ✔ Saving time in the R&D process
- ✔ Achieving optimized timetables of bus lines within a reasonable time
- ✔ Having a parallel execution framework for complex optimization algorithms

- Dask Library
- Big Data
- HPDA
- Optimization
- Smart City

- Industry Sector: Smart City, Intelligent Transportation System, Bus Route Timetable Optimization, Public Transportation

- Technology: Big Data, HPC, HPDA

Contact:
Metin BARIŞ -Parabol,
Email: metin.baris@paraboly.com
Cevat ŞENER -METU,
Email: sener@ceng.metu.edu.tr

Traffic Events and Alerts Data

Integration and Analytics to Support Managerial Decisions

NCC-Bulgaria was founded by the Institute of Information and Communication Technologies at the Bulgarian Academy of Sciences, the Sofia University "St. Kliment Ohridski" and the University of National and World Economy.

NCC-Bulgaria is focused on:
- Creating a roadmap for successful work in the field of high-performance computing, big data analysis and artificial intelligence.
- Analyzing the existing competencies and facilitating the use of HPC/HPDA/AI in Bulgaria
- Raising awareness and promoting HPC/HPDA/AI use in companies and the public sector.

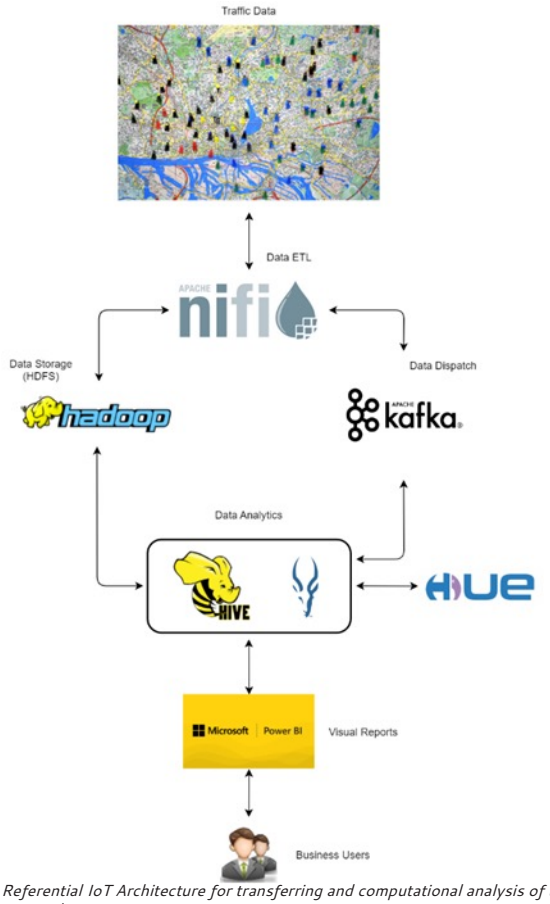
Industrial organisations Involved

- Companies in the transport sector
- Companies in the Logistics and Supply Chain Management sector
- Electrical vehicles users
- Companies developing electrical charging stations
- Companies managing autonomous vehicles.

The Challenge

The considerable technical challenge is to accurately identify, extract, transform and integrate reliable and fast changing data from traffic events and alerts to be followed by sophisticated and exhaustive analysis (descriptive and predictive).

On one hand, the large volumes of data being transferred, the enormous diversity of traffic sensors, events, and alerts to be encompassed, and the proliferation of the transmission protocols is a very difficult technical and scientific task. On the other hand, consistent assurance and maintenance of collected data quality while selecting the right tools and algorithms for data analytics and presentation, is not a trivial task.



Referential IoT Architecture for transferring and computational analysis of traffic events data:
Components and their communication links, within the realized referential IoT architecture

Video: https://www.youtube.com/watch?v=zEFYrFl67xo&t=5s

The Solution

To solve those challenges, a referential IoT architecture is proposed, consisting of the following components:

- Data extraction and cleaning module NiFi,
- Data distribution module Kafka,
- Data storage module Hadoop HDFS,
- Data analytics modules Hive, Impala and Hue,
- Data presentation module PowerBI.

Building such an architecture enables flexibility and speed in data loading and cleaning up from a multitude of traffic information platforms, different payload formats (CSV, JSON, XML, etc.) and transmission protocols (HTTPS, FTPS, SFTP, etc.), reliable ingestion, dispatch, consumption and storage of the extracted traffic events, powerful data analysis engines for both streaming and batch processing, and, finally, convenient visual presentation of the results ready to be used by the business stakeholders.

The Business impact

The big data, generated in the Transport and Logistics & SCM fields, is not adequately collected, organized and used for the optimization of business processes, activities and services.
Expanding the usage and geographical scope of autonomous vehicles such as self-driving cars and self-driving cargo trucks it is very important to apply new technologies to process the generated big data that should be collected, transformed, and analysed in real time. In this regard, becomes critical to design and develop complex IoT architectures and analytical models to monitor and manage all the processes related to the autonomous vehicles.

In addition, such IoT architectures are suitable solutions for collecting data on the practical use of personal and public electrical vehicles. One aspect of these big data is the analysis for planning the optimal distribution of electrical charging stations and predicting the capacity of their use. As a result, based on analytical reports, consumers plan and optimize their routes, navigating electric vehicles to minimize tails and charging time.

- IoT
- Big Data
- Traffic Events
- Traffic Alerts
- Analytics
- Apache Ecosystem
- Reference Architecture

- Industry Sector: Trannsport, Logistics & Supply Chain Management

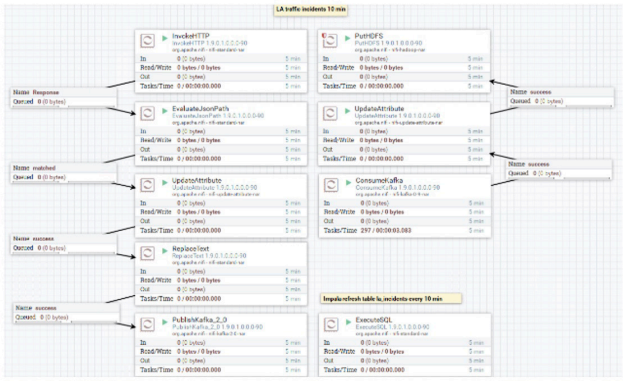
- Technology: IoT, Apache Ecosystem and HPC/HPDA Integration

The companies, operating in Transport and Logistics & SCM sectors, benefit from the real-time analytics and predictive instruments, for:

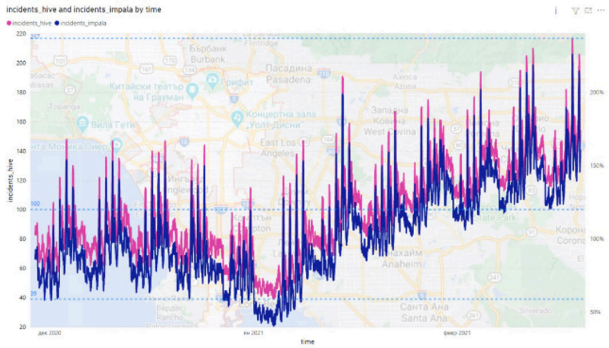
- Optimizing transportation routes
- Decreasing transportation and logistics expenses
- Optimizing staff management
- Warnings and alerts for traffic incidents, natural disasters and extraordinary events
- Real-time management of autonomous vehicles

Benefits

- ✔ Flexible and reliable, multi-protocol and multi-network data extraction
- ✔ Distributed high availability file system and event dispatching system
- ✔ Powerful data analytics including streaming and batch processing
- ✔ Cost effective Open-Source components



Traffic Events ETL: NiFi flow created and used for events extraction, transformation and load



Traffic Events Report:
PowerBI analytical events report, created by using ODBC connection with the cluster

Contact:
Dr. Yavor Christov,
University of National and World Economy,
Email: ychristov@unwe.bg
Prof. Kamelia Stefanova,
University of National and World Economy team
Email: kstefanova@unwe.bg

RISE Accesses VEGA

ENCCS (EuroCC National Competence Centre Sweden).
ENCCS provides high-performance computing training and support for industry, academia and public administration for free.
https://enccs.se

Industrial organisation involved

Trafikverket (Swedish Transport Administration) is responsible for long-term planning of the transport system for all types of traffic, as well as for building, operating and maintaining public roads and railways.
https://www.trafikverket.se

Technical Challenge

While the modern trend towards producing electric vehicles (EVs) is rising, a total change of fleet to EVs will require a substantial amount of time. For these reasons, the traffic flow at highway E4 in Sweden has been the subject of a collaboration between ENCCS, KTH, and Trafikverket.



Photo: Jacek Dylag // Unsplash

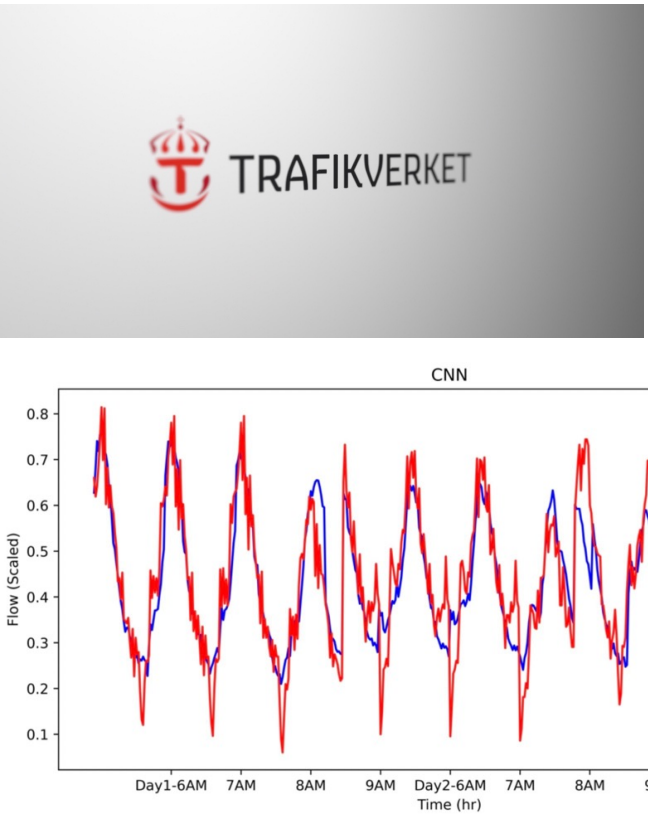
Solution

As a case study, the traffic flow will be modeled on a specific part of the E4S using deep neural networks. The results will give us a hint at the most appropriate models for such modeling. In the second part of the project, the already available model Temporal Graphical Convolutional Networks (T-GCN), will be tested and ported to an HPC system for further analysis of a much larger dataset corresponding to the E4S more extensive structure.
This undertaking is a collaboration between Assoc. Prof. Xiaoliang from KTH, Mr. Christian Edfjall from Trafikverket, and Dr. Hossein Ehteshami from ENCCS.

Business Impact

Trafikverket is the Swedish public authority responsible for the transport system. AI methods in public administration are crucial for the development of social actions that benefit the public. Such methods in the branch of traffic administration will play an important role for future political strategies.

Most inter- and intra-city transportations occur via vehicles. Managing the traffic caused by such vehicles is of great importance for several primary reasons. Proper traffic control saves plenty of time and resources, assures the safety of drivers and passengers, and last but not the least, could help mitigate the issue of air pollution.



Result: The comparison of measured traffic at E4S for two consecutive days between 6 to 10 AM obtained from CNN (Convolutional Neural Network)

Benefits

- Pollution control
- Road design planning
- Traffic safety

- AI
- Traffic
- Neural Networks
- Vehicle

- Industry Sector: Public Administration
- Technology: AI

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

Section 8

Public and Communication

Natural Language Processing models using MeluXina - 146

Low-latency AI-based Gesture Recognition – 148

9 Million Scanned Images for Access to the Public – 150

RISE Accesses VEGA – 152

Vinnova Obtained Access to Vega EuroHPC JU System - 154

Improving Speech-Recognition Technology with Speechly – 156



Photo: krakenimages // Unsplash



Photo: Nick Morrison // Unsplash

The National Library of Sweden to train

Natural Language Processing models using MeluXina

ENCCS (EuroCC National Competence Centre Sweden).
ENCCS provides high-performance computing training and support for industry, academia and public administration for free.
https://enccs.se

Organisations Involved

KBLab at the Swedish National Library (Kungliga Bibliotek, KB). KBLab is a national research infrastructure for digital humanities and social science. The lab provides access to KB's collections in structured and quantitative form. This makes it possible for researchers to both seek new answers and pose new questions in their research.
https://www.kb.se/in-english/research-collaboration/kblab.html

Technical Challenge

Due to the high financial cost of training Natural Language Processing (NLP) models, most of the researchers have only been trained by large companies on English data. In developing a competitively-sized model for Swedish, the Royal Library of Sweden hopes to enable commercial and non-commercial usage of this technology, while allowing researchers to understand what these models learn not only for English.

Solution

With access to MeluXina supercomputer, KBLab is now able to use 10 000 000 GPU hours to train their Natural Language Processing models using their enormous text, audio, and video archive.

It's the second time KB gains access to a EuroHPC JU supercomputer and that makes it the first public administration entity to be accepted through the EuroHPC JU Regular Call which gives the ability to use an even larger amount of CPU and GPU cores.

The transformer neural network and the subsequent transformer language models (LM) such as GPT and BERT have become the new standard of a pretraining-finetuning paradigm. This approach leverages large amounts of unannotated text data in a self-supervised pretraining step, producing generalist models that can then be finetuned on specific tasks.

The largest of these models consists of several hundred billion parameters requiring the model itself to be split over multiple GPUs, further needing up to multiple thousand GPUs to be trained in a reasonable amount of time. Due to the extreme size of the models and their ability to absorb massive datasets, they have shown to be very adept at learning new tasks with only a few training examples or even none, instead only needing a prompt describing the task (Brown et al. 2020).

Business Impact

KBLab holds high-quality massive humanistic data including text, audio, images and video. This makes KBLab a perfect candidate to create NLP models by using hundreds of years' diverse material (approximately 26 PB of data). KBLab's open source VoxRex audio-to-text model has been downloaded more than 200 000 times for many different use cases.

Benefits

- Having a public authority like KBLab developing such open-source models is crucial to societal development. The models can be used by any person without limitations and it is therefore hard to follow in detail the specific benefits that it will bring.
- KB managed to utilise their large material and train their models for free by accessing a EuroHPC JU system. This paves the way for other public sector organisations to do the same and develop their services further.

- Machine Learning
- AI
- Natural Language Processing
- Cultural Heritage

- Industry Sector: Public Administration

- Technology: AI

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



Photo: Usman Yousaf // Unsplash

Low-latency AI-based Gesture Recognition

NCC-Bulgaria was founded by the Institute of Information and Communication Technologies at the Bulgarian Academy of Sciences, the Sofia University "St. Kliment Ohridski" and the University of National and World Economy.

NCC-Bulgaria is focused on:
- Creating a roadmap for successful work in the field of high-performance computing, big data analysis and artificial intelligence.
- Analyzing the existing competencies and facilitating the use of HPC/HPDA/AI in Bulgaria
- Raising awareness and promoting HPC/HPDA/AI use in companies and the public sector.

Organisations Involved

Software Company Ltd is a Bulgarian SME specialised in software development. Since 1996, the company has offered a wide range of high-quality services in the development, delivery, and maintenance of AI and HPDA software in Bulgaria. The company is also ISO 9001:2015 certified and ISO 13485:2016 certified and has CE Mark certificate.

The shared project (SME + NCC BG_SofiaUni team) was completed in relation with the application of the company in the BONSAPPS Open Call for AI talents.

Technical Challenge

During this challenge (BONSAPPS Hackathon), a so-called AI Asset was developed that can recognise a driver's gestures by using a single, cheap and low-resolution thermal camera. This asset was then optimised and deployed to low-power edge platforms such as the NVidia jetson nano.

The final implementation was based on the approach named Embedded vision, which gives visual intelligence to machines and teaches them how to see and think. As a result, a small footprint neural network was created, which consisted of a Convolution Neural Network (CNN) and a multi-stage Temporal Convolutional Network (TCN). Such a network design allows to work with low-resolution images taken by cheap thermal cameras that can be installed in private cars at a low cost. This solution was implemented inside the BONSAPPS AIAsset framework, as part of the BONSAPPS challenge. This framework abstracts away the hardware platform on which the AI app runs on allowing us to focus on the AI app instead of hardware specifics. The hybrid HPC cluster also allowed the team to work with much larger datasets (HPDA).

Solution

Since the aim has been to run the optimised neural network (AI) on low-power devices, while ensuring both high accuracy (HPDA) and low latency (HPC), the final model had to be well-optimised to run on the pre-selected devices.

That is why we automated the training, quantisation, export and deployment to the testing platform parts of the development process by using GitLab CI/CD pipelines running on a HPC cluster with multiple Nvidia V100 compute units. That way we could iterate much faster since the time between each change and deployment had been significantly shorted by said CI/CD pipelines.

Using the HPC cluster also allowed the team to work with much larger datasets that a regular workstation computer would not even be able to load into memory.



Collaboration team from Software Company Ltd



Certificate of attendance at the Hackathon

Business Impact

The Software Company Ltd. was recognised by the BONSAPPS as a reliable partner due to support by our team, resulted in a solution of an important problem using cutting-edge scientific research: implementing HPC technology to improve both the development process and the quality of the final AI application.

Since AI can think faster than a human can due to its deep-learning algorithm maps, it makes perfect business sense to implement AI in various business processes in order to help the company successful in the current competition between the SME.

Benefits

For the Sofia University Team:

- ✓ Acquisition of new technical skills;
- ✓ Improving communication skills;
- ✓ Networking;
- ✓ Inspiration.

For the Software Company:

- ✓ Success and accomplishment feelings.

- Computer Vision
- HMI
- Deep Edge
- Gesture Recognition

- Industry Sector: Automotive

- Technology: HPC, AI

Contact:
Sofia University
Email: anap@phys.uni-sofia.bg

Swedish National Archives will use HPC to process

9 Million Scanned Images for Access to the Public

ENCCS (EuroCC National Competence Centre Sweden).
ENCCS provides high-performance computing training and support for industry, academia and public administration for free.
https://enccs.se

Industrial organisations Involved

Riksarkivet: the Swedish National Archives primarily manage the government records for future generations. As a supervisory authority, they also offer services related to their public task.
https://riksarkivet.se

Riksarkivet

The Challenge

The Swedish National Archives have a large collection of images which contain huge amounts of information, but they are not easily searchable. It would need a large team of employees digitalising the information on those images. Furthermore, manual analysis would be extremely slow.



Photo: Emile Perron // Unsplash

The Solution

The Swedish National Archives will use their access to EuroHPC JU supercomputer VEGA to train and run **HTR** (Handwritten Text Recognition)/ **OCR** (Optical Character Recognition) pipelines consisting of segmentation models, text-recognition models and post-correction models.

The Business impact

This will enable the Swedish National Archives to process approximately 9 million scanned images, index them with metadata, and make them searchable and more accessible. It will improve access to their archives and streamline their internal case handling.



Benefits

Access to Vega, and the ability to train models specialized on archival texts, will:

- ✔ Further the larger goal of adapting AI-technology for use in archives.
- ✔ Opening up new possibilities for academic researchers, genealogists and others to explore and use the archival holdings.
- ✔ This project saves us about 700,000 euros in manual labor costs, and the indexing database is created significantly quicker.

- AI
- Deep Learning
- Image Segmentation
- Image to Text
- Machine Learning
- Public Sector

- Industry Sector: Public Sector

- Technology: AI

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

RISE Accesses VEGA

ENCCS (EuroCC National Competence Centre Sweden).
ENCCS provides high-performance computing training and support for industry, academia and public administration for free.
https://enccs.se

Industrial organisation involved

RISE Research Institutes of Sweden is Sweden's research institute and innovation partner. Through international collaboration with industry, academia and the public sector, they ensure business competitiveness and contribute to a sustainable society.
https://www.ri.se/en

Technical Challenge

Natural Language Understanding (NLU) is an area in machine learning that transforms human natural language into easily understood machine language and has been recently largely developed for large languages around the world. For such models to work, the algorithm must train in large number of texts and for that it needs large computing power, specifically GPUs.



Photo: Gabrielle Henderson // Unsplash

Solution

Evangelia Gogoulou from RISE NLU Group, Digital Systems got access to the EuroHPC JU system Vega at Izum, Slovenia. She will be training from scratch an English BERT model using multiple GPUs and then evaluate its downstream performance on the GLUE benchmark (https://gluebenchmark.com/).

The model performance will be compared with the English BERT trained on one GPU. The second task is to start from the English BERT model and continue pre-training it on Russian, on multiple GPUs. The transferred Russian model will be then evaluated on GLUE and the results will be compared with existing Russian Language models.

Business Impact

RISE, being the public institute for research and innovation in Sweden, has the role of connecting academic research with the industrial sector for the benefit of society and creating progress. This type of modern language AI has the potential to radically improve the handling of language data in both the private and public sectors. In particular, RISE and the NLU Group are committed to support Swedish actors in developing and integrating the latest NLP technologies into their operations.

Benefits

- ✔ Categorise free text effectively
- ✔ Multi-language translation
- ✔ Automate procedures using natural language

- Natural Language Understanding (NLU)
- AI
- Language Model
- Deep Learning

- Industry Sector: Public Administration

- Technology: AI

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



Photo: Kvalifik // Unsplash

The Swedish Innovation Agency Vinnova

Obtained Access to Vega EuroHPC JU System

ENCCS (EuroCC National Competence Centre Sweden).
ENCCS provides high-performance computing training and support for industry, academia and public administration for free.
https://enccs.se

Organisations Involved

Vinnova is Sweden's innovation agency. They help to build Sweden's innovation capacity, contributing to sustainable growth. Vinnova's vision is that Sweden is an innovative force in a sustainable world.
https://www.vinnova.se/en

Technical Challenge

Vinnova is a government agency under the Ministry of Enterprise and Innovation, and the national contact authority for the EU framework programme for research and innovation. Being a governmental agency Vinnova handles a large number of documentations and enhanced analyses of these are of crucial importance.
The AI analysis tool used by Vinnova was originally developed by the Swedish Innovation agency and the Swedish agency for Growth Policy Analysis, Tillväxtverket. Features of this tool include topic modeling, keyword extraction, an AI assisted free text search as well as speech tagging to isolate target audience and purpose.

Solution

The successful application was a joint effort between Vinnova Data Scientist Dr. Leonora Vesterbacka Olsson and ENCCS. With the awarded HPC time on the Vega (https://doc.vega.izum.si/) EuroHPC JU petascale system within EuroHPC JU development call, Vinnova intends to further develop this AI tool for analyses of new types of government agency documents and enhance it with new features.

Business Impact

The text analysis tool in its current state constitutes an aid for analysis of calls for proposals for funding and helps the analyzer to identify trends in topics and purpose found in funding proposals and applications for funding. The various functionalities provide an easy way to perform exploratory analysis and draw well-based conclusions from the results.

Benefits

- ✔ Streamline and automate analysis that has traditionally been done manually.
- ✔ Analysing larger amounts of data and thus contribute to Sweden's strategy regarding innovation.
- ✔ Enhance Vinnova's tools to get important insights for future use regarding successful projects that lead to added value for the Swedish society.





Photo: Jason Goodman // Unsplash

- Data Science
- AI
- Innovation
- Data Analysis
- Strategy

- Industry Sector: Public Administration

- Technology: AI

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

Improving the Speech-Recognition Technology with Speechly

EuroCC National Competence Centre Finland
The Competence Centre Finland will heavily focus on paving the users' way on LUMI, the EuroHPC flagship supercomputer located in Finland. LUMI is one of the key elements in Finnish and European HPC competence development.

Industrial organisations Involved

Speechly is a Finnish company specializing in speech interfaces, https://www.speechly.com a start-up company established in Helsinki, Finland in 2016, has developed technology that allows software developers to easily add an intuitive voice interface to any platform, whether it be Android, iOS, website or video game.

The Challenge

Building speech interfaces is difficult and current solutions on the market are only able to perform fairly simple functions, such as playing Spotify and turning lights on and off. Speechly combines speech recognition with the technology of understanding natural language in a new way to offer a comprehensive solution that enables more complex voice control commands and responds to users' needs.

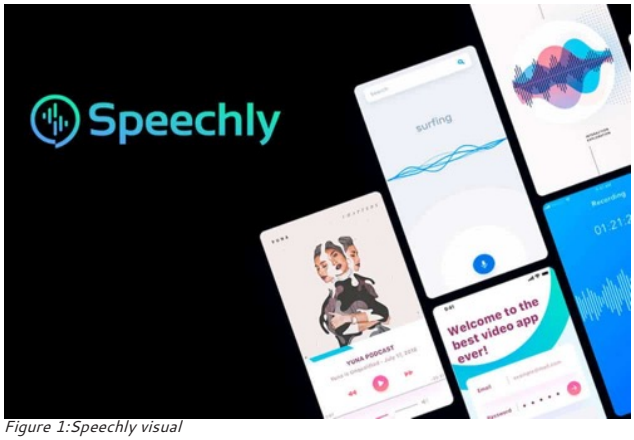


Figure 1:Speechly visual

Video: https://www.youtube.com/watch?v=yzdSVV4xjb0

The Solution

Speechly, has developed a new type of technology for understanding real-time spoken language. With CSC's Puhti supercomputer and Business Finland's AI Business program, Speechly was able to further improve its speech recognition technology.

Speechly's product development work involves testing different machine learning algorithms to find the best speech recognition models. During the nine-month AI Business project, Speechly improved the accuracy of its speech recognition model by almost 60% in terms of word error rate (**Hannes Heikinheimo**, Speechly co-founder). Speed is an asset in product development, as training one model can take anything from days to weeks. Speechly used CSC's Puhti-AI computing environment with support from NCC Finland, making it possible to work faster and utilize larger sets of data.



Photo: Ben Moreland // Unsplash

The Business impact

Speechly is competing with the world's largest tech giants where some of Speechly's people have worked in the past. To strengthen their competitive advantage, the company sought support for product development from Business Finland's AI Business program, allowing them access to CSC's computing resources free of charge with the AI Computing Grant.

Speechly's product development work involves testing different machine learning algorithms to find the best speech recognition models. Modern machine learning models for deep neural networks require high computing capacity.

How quickly the staff can train and test one model is one of the most important factors affecting product development results. Training one model can take anything from days to weeks. Thanks to CSC's computing capacity, Speechly staff was able to work much faster than before and with significantly larger data sets.

Benefits

- Speechly's product development work involves testing different machine learning algorithms to find the best speech recognition models.
- During the nine-month AI Business project, Speechly improved the accuracy of its speech recognition model by almost 60% in terms of word error rate
- Speed is an asset in product development – training one model can take anything from days to weeks.

- HPC
- AI
- Neural Net Training
- Speech
- Natural Language Understanding (NLP)
- Speech-to-Text

- Industry Sector: Speech Recognition

- Technology: HPC, AI

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

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, 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.



Co-funded by
the European Union



EURO^2

Learn more about ***EuroCC2*** *here!*