H2020-INFRAEDI-2018-2020

Concerted action for the European HPC CoEs

Project Number: 823964

D3.5 – Final Report on WP3 activities (including a success story booklet)

WP3 – CoE-Industry Interaction

Focus CoE has received funding from the European Union's Horizon 2020 research and innovation programme under the grant agreement Nº 823964

Workpackage:	WP3	CoE-Industry Interaction
Due Date	2022-03-30	
Author(s):	Marie-Françoise Gerard Lucile Garnier Marta Chinnici Andrea Quintiliani Xavier Salazar Alexander Hektor Teona Macharadze Hugh Martin	TERATEC TERATEC ENEA ENEA BSC HLRS HLRS UCL
Authorized by	Coordination team	SCAPOS
Reviewer	Julien Thélot	CEA
Reviewer	Simon Wong	ICHEC
Reviewer	Massimo Celino	ENEA
Final Version Date	2022-03-25	
Dissemination Level	Public	

List of Acronyms and Abbreviations

AI	Artificial Intelligence
B2B	Business to Business
CoE	Centres of Excellence (CoEs) for High Performance Computing (HPC) applications
CSA	Coordination and Support Action
D	Deliverable
EC	European Commission
EU	European Union
H2020	Horizon 2020 – The EC Research and Innovation Programme in Europe
HPC	High Performance Computing
HPCCOE	HPC CoE Council
KPI	Key-Performance Indicator
NCC	National Competence Centres
PCC	Project Coordination Committee
R&D	Research and Development
SME	Small and medium-sized enterprise
Tech. Session	Technical Session
WP	Work Package

Executive Summary

This deliverable reports on the final status of all WP3 activities and summarizes the main outcomes and conclusions with regards to interaction of the CoEs with industry.

This deliverable reports on the final status of all WP3 activities and summarizes the main outcomes and conclusions with regards to interaction of the CoEs with industry.

During the FocusCoE project and as stated in the FocusCoE Grant Agreement, the WP3 team has worked toward developing the interaction between the CoEs and industry with the aim of increasing the impact of the CoE effort on the competitiveness of European companies. The main objectives of WP3 were:

- outreach and promotion of the CoEs' competences and service offerings;
- general promotion of HPC use by SMEs including interactions with other actions/undertakings;
- monitoring of success stories and interaction with WP5 (dissemination of success stories);
- business development support by the analysis of "customer/user" sectors and market "requirements capture" and feedback to the CoEs;
- exchange of best practices regarding industry-CoE interactions.

To achieve these objectives, Task 3.1 "Action planning and monitoring of success" focused on strategic planning of the interactions with industry, selecting the most relevant actions and also monitoring the successes of the WP3 actions by ensuring the general management of WP3. Task 3.2 "Sectorial communication activities" implemented part of the plan elaborated in Task 3.1 by organising and executing participation in industrial sectorial events and by promoting the CoEs' competences and service offerings in these events. Task 3.3 "Coordination with HPC SMEs initiatives" provided a dedicated effort toward SMEs and the connection with other actions targeting the development of HPC use by SMEs. Task 3.4 "Industrial feedback to CoEs" gathered the needs, challenges and requests coming from industry that could be answered by the CoEs and gave feedback about this information to the CoEs.

The previous deliverables, D3.1 "Report on the sectorial approach priorities" and D3.2 "Interim report on the WP3 activities", reported on all the work and outcomes achieved in each task of WP3 until the Midterm review of the whole FocusCoE project. D3.1 was initially submitted at M6 and D3.2 was initially submitted at M18- (both deliverables were re-submitted after the Midterm review in November 2020). Then the deliverable D3.3 "Report of the actions towards SMEs"), submitted in September 2021, reported on the actions performed by Task 3.3 until September 2021. Finally, the deliverable D3.4 "Report on the sectorial communication activities" (submitted in parallel of this deliverable D3.5) reports about the actions performed by Task 3.2 until the end of the extension of the FocusCoE project in March 2022. Therefore, this final deliverable D3.5 - except for the new actions performed in the meantime - will mainly provide the high-level conclusions and main outcomes of each task. Since some details for each task are already available in the relevant previous or parallel deliverables, direct references will be made to those deliverables where suitable and only brief descriptions or explanations will be repeated when needed to ensure consistency in this final deliverable.

Table of Contents

1 Introduction 8
2 Developing the interaction between the CoEs and the industry 9
2.1 Methodological approach used in WP3 9
3 Main outcomes and results per task achieved by WP3 during the FocusCoE project 9
3.1 Task 3.1: Action planning and monitoring of success 9
3.2 Task 3.2: Sectorial communication activities 11
3.3 Task 3.3: Coordination with HPC SME initiatives 13
3.4 Task 3.4: Industrial feedback to the CoEs 14
4 Alternative solutions and additional plans set up by WP3 following the COVID-19 outbreak 19
5 Final conclusions and final feedbacks from WP3 with regards to the CoEs' interactions with industry 19
6 References 21
7 Annexes 22
7.1 Success stories 22
7.2 Annex: Guidelines for interaction with industry (extract from the generic common part of the "Industry Questionnaire" report sent to the CoEs) 53

Table of Tables

No table.

1 Introduction

The main goal of WP3 was to support, assist and provide guidance to the COEs with regards to the development of their interactions with industry. This was realised by: supporting and guiding the CoEs in the outreach and promotion of their competence and service offering; interacting with other actions/undertakings in favour of general promotion of HPC use by SMEs; contributing to the success stories production and the monitoring of success stories in interaction with WP5 (responsible for the dissemination of success stories); providing analysis of "customer/user" sectors and market "requirements capture" and giving feedback about them to the CoEs; and providing best practices regarding promotion of services to industry through sectorial activities.

The present document describes the global strategy set up by WP3 and outlines the main results, outcomes and conclusions issued from the actions achieved by WP3. This document is organized as follows: first a short reminder about the methodological approach used in WP3 is given in Section 2 (since it has been exhaustively detailed in the previous deliverables D3.11 and D3.22). Then, Section 3 presents the main results in each task of WP3 since the Midterm Review. Section 4 then summarizes the issues, obstacles (mainly due to the COVID-19 crisis encountered during the FocusCoE project), solutions, and additional plans decided by WP3. The final conclusions and final feedbacks from WP3 with regards to the CoEs' interactions with industry are displayed in Section 5.

[1]FocusCoE Deliverable D3.1 "Report on the sectorial approach priorities"
[2] FocusCoE Deliverable D3.2 "Interim report on the WP3 activities"

2 Developing the interaction between the CoEs and the industry

2.1 Methodological approach used in WP3

The methodological approach used in WP3 has already been explained in more details in the previous deliverables D3.1 and D3.2, but below is a short reminder of the global methodological approach used in WP3, aimed at promoting the CoEs' competences and offerings toward industry. The strategy was based on two main approaches:

1. Leveraging existing contacts of the WP3 partners with industrial companies;
2. Creating contacts with new companies not yet strongly connected with the HPC ecosystem.

The WP3 partners had been selected based on their expertise and collaboration with industrial HPC users. Through their past activities, they had developed industrial contacts with both large companies and SMEs. The CoEs have benefited from the WP3 partners' networks in particular via the "Direct Contact Approach" (as part of Task 3.1), allowing the CoEs to benefit – if interested – from the industrial contacts of the WP3 partners and via the coordinated effort towards SMEs (Task 3.3).
The activities in Task 3.2 on "Sectorial events" followed the second approach. Here, the idea was to select sectorial events where WP3 could develop new contacts of interest to the CoEs. These industrial sectorial events initially included trade fairs, brokerage events, or local industrial Open Days. Due to COVID-19, online events were targeted in replacement of the cancelled or postponed events, and additionally thematic webinars were organized. WP3 also provided feedback to the CoEs on their offerings. This has been done via the questionnaire towards industry in Task 3.4.

3 Main outcomes and results per task achieved by WP3 during the FocusCoE project

3.1 Task 3.1: Action planning and monitoring of success

The activities of Task 3.1 had three components:

1) Identify the most relevant actions toward industries that can be beneficial for the promotion of the CoEs;
2) Prepare the plan for these actions;
3) Monitor the execution of the plan and identify successes and best practices.

The first actions of Task 3.1 (performed from M1 to M6 of the FocusCoE project) have been devoted to the setup of a strategic plan to support the CoEs in the promotion of their offerings and interactions with industry. The methodology and the results are detailed in the deliverable D3.1 ('Report on the sectorial approach priorities'). Then the results for M7 to M18 are detailed in D3.2 ('Interim report on the WP3 activities').

Since the deliverable D3.2 (re-submitted at end of 2020) and following the first online events which took place in 2021, Task 3.1 proceeded with the following activities (envisaged and presented as "upcoming" in D3.2):

- in interaction with the CoEs, Task 3.1 selected and updated the next set of sectorial events for the CoEs to attend with support of FocusCoE through to 2022. The selection took into consideration the online format (due to COVID-19) and the specific interests of the CoEs (for having a booth, a talk, or being part of an experts panel for instance). Thematic webinars were also proposed and organized to target a specific topic or industry sector. For these thematic webinars, WP3's existing contacts in industrial companies were used to promote the events and to invite the most relevant industry stakeholders.
- FocusCoE initiated interactions of the relevant CoEs with the NAFEMS community, with the Enlit community, with the SIMAI community in order to get the CoEs in contact with industry-oriented events specific to their own communities and to promote their offerings to the most industry-comprised audiences possible. The CoEs, with the support and guidance of FocusCoE, participated in these events and interacted with these communities through the Task 3.2 sectorial events actions.
- An analysis of the WP3 KPIs, in conjunction with direct discussions with CoE representatives, confirmed that they were appropriate for assessing and monitoring our activities. In addition, direct feedbacks from the CoEs-industry contact points were regularly collected through two channels: firstly, after each sectorial event the WP3 partner in charge of coordinating the participation of the CoEs collected the CoEs' feedbacks; also, WP3 asked the CoEs'-industry contact points during the Joint All CoEs-WP3 telcos which were organized every three months to provide feedbacks on the long-term outcomes after the sectorial events.

Regarding the project management of WP3: the activities of WP3 were managed through online meetings, held once a month. The meetings were regularly attended by representatives of all organisations participating in the WP. Occasionally, minor issues that could not be handled internally within the WP were reported to the Project Coordination Committee (PCC) for resolution. No critical specific problems were to be reported about these meetings.
Additionally, every three months, WP3 organized online meetings with all the CoEs-industry contact points in order to collect their needs, to share the plans for the selected sectorial events, to promote the documents produced for the CoEs to guide them to interact with industry, and to collect the CoEs' feedback with regards to the past events.
Task 3.1 continued until the end of the extension in March 2022 to ensure that all the planned actions and the additional ones decided after the Midterm Review were achieved.

Regarding success stories, there are now 9 success stories available on the HPCCOE website: https://www.hpccoe.eu/success-stories/. The full contents of the stories are displayed in a "raw" version in annex 7.1. It was initially planned as a booklet of success stories to be used as communication material during physical events. However, due to COVID-19 all events went online or in a very limited hybrid mode. Also due to the travel restrictions of the various countries, the WP3 partners or the CoEs' representatives were not able to travel and attend any physical events. Therefore, the webpage format appeared more relevant, and in agreement with the Contract Amendment #2 signed by the Commission on 21.09.2021, the success stories were promoted through this online webpage format. These stories were also regularly promoted through social media and the FocusCoE newsletter by WP5.

3.2 Task 3.2: Sectorial communication activities

Task 3.2 was dedicated to the outreach and promotion of the CoEs' competences and service offerings towards potential industrial users to help the CoEs to connect with industry. The initial strategy and the results of this task for M1 to M18 are detailed in D3.2 ('Interim report on the WP3 activities'). Then the updated strategy and the following results from M19 to M40 are detailed in D3.4 ('Report on the sectorial communication activities') which has been submitted in parallel to this final deliverable D3.5.
The sectorial events were managed by Task 3.2 until the very end of the project at M40 with the last event, AI4EU web cafe, held on March 17th 2022 after several postponements due to internal reasons within the AI4EU project.

To assess the results for the CoEs regarding the efforts provided within Task 3.2 and also more generally by WP3, WP3 interacted with the CoEs' representatives during a last "Joint ALL CoEs-WP3" telco on February 9th 2022 and via follow-ups emails to collect all of the CoEs feedback. The feedback received from the CoEs about the Task 3.2 actions, and the WP3 provided support and guidance for the sectorial events were:

- Regarding the first sectorial event, Biofit2019, which took place in December 2019 before the COVID-19 outbreak, the CoEs were fully satisfied with the coordination carried out by WP3 (see details in D3.4[3]). For CompBioMed, the main benefit was the exposure to an industry audience; for POP there were 16 new leads, contacts with 5 professional bodies and this experience helped them to build their understanding of the sector; for BioExcel, there was 1 SME lead that has had an ongoing interaction with the CoE following the event.
- For the first online events in early 2021, not all CoEs were ready (with regards to their maturity and how advanced they were in the project) to reap the full benefits from the events as all of their communication material were not always finalized and available for a complete interaction during the events. Moreover, the events were not always appropriately prepared by the organizers of the conferences or the trade fairs due to a lack of experience in early 2021 with online events (e.g. with virtual platforms not practical to interact with visitors). This was, for instance, the case for the 'International CAE Conference CAE2020" and partly for the 'European Materials Modelling Council event EMMC2021' (see details in D3.4). However:
    - For CAE2020 in November 2020, for EXCELLERAT and MaX, this event was the opportunity to create all the communication materials (with FocusCoE WP3 and WP5 support) that they used afterwards during other events and actions. Later on in 2021, POP e-mailed the attendees after the event and they were invited to contribute with an article[4] to Enginsoft's "Newsletter Magazine" which provided concrete visibility to POP in the Enginsoft community and was a nice opportunity to promote their offerings.
    - For EMMC2021 in March 2021, POP had 10 leads, 1 of whom became a user; for Nomad-2 several discussions were started thanks to the event and one is still ongoing.
- Regarding the thematic webinars organized by WP3 (see details in D3.4):

[3] FocusCoE Deliverable D3.4 "Report on the sectorial communication activities"
[4] https://www.enginsoft.com/assets/pdf/newsletter/newsletter2021_1.pdf#page=52

  - in February 2021, about 'Opportunities and Challenges for Industrial Applications'[5] the CoEs had not specific industry leads after this first organized webinar. However, for POP, even if it had no direct impact for their business development, this webinar was a chance to publicize the work of POP.
  - In May 2021, during "The HPC CoE services and applications" [6] webinar organized within the HiPEAC sessions , the participating CoEs reported that they benefited from this webinar through their interactions with the other CoEs in the same field and with an appropriate audience. They also relished the opportunity to learn how other CoEs deal with industry and share best practices. CompBioMed and BioExcel also reported that this webinar was a valuable means for the CoEs to raise their dissemination figures.

- Regarding the 'ITALIAN SOCIETY OF APPLIED AND INDUSTRIAL MATHEMATICS (SIMAI)' congress in September 2021 (see details in D3.4), EoCoE viewed it as a very interesting opportunity to disseminate to a wide audience and to interact with other CoEs, even if there were no specific outcomes after the event in terms of concrete leads or specific interactions with industry contacts afterwards.
- Regarding the ' EU Sustainable Energy Week (EUSEW)' in October 2021 (see details in D3.4), while the event was complicated to coordinate as this EU-organized event was postponed repeatedly, and the format and proposed modes of participation evolved several times, the overall feedback was very positive and the CoEs confirmed the importance of having participated for increased awareness and to be able to promote their services offerings at the EU level and to interact with policymakers and citizens on major challenges in the energy domain. For HiDALGO, this event created an increase of their social media engagements and they had the opportunity to interact with a good number of participants during their case study presentation. For EoCoE, this was a very relevant opportunity to discuss HPC in a high-level policy setting and to gain high visibility with policymakers and citizens. For MaX, the interest was in being part of the political discussions and to have been part of this high-level event. All three CoEs felt that the session they delivered at this event was a highly valuable 'foot-in-the-door' moment for them, that may lead to an increased presence at subsequent iterations of the event series.
- Regarding the NAFEMS World Congress in October 2021 (see details in D3.4), even if the online platform wasn't fully satisfactory for such a huge event (there were lots of visitors to the online booth but the virtual platform prevented the booth organisers from actively interacting with visitors if the visitors didn't take the initiative or willingly left their business cards), there were positive consequences for the CoEs as EXCELLERAT reported a gain of 5 new members to their services portal, several new subscribers to their newsletter, and 1 company asked for partners through EXCELLERAT for a technical challenge. POP also reported one promising lead afterwards.
- Regarding MEDICA2021 in October 2021 (see details in D3.4), both CoEs CompBioMed and PerMedCoE were satisfied with their participation in the

[5] https://www.hpccoe.eu/2021/03/05/focuscoe-webinar-opportunities-and-challenges-for-industrial-applications/,

[6] https://www.youtube.com/watch?v=CNa2j7HAxmU&feature=youtu.be

experts panel. CompBioMed reported that in this case the main benefit was the exposure to an industry audience to promote the outcomes of the CoE.

- Regarding AI webcafe (see details in D3.4), the CoEs presented several case studies on how AI is leveraging HPC technologies to solve real-life problems and discussed the opportunities and challenges with the audience in a panel discussion. 73 people registered for this event and there were 41 attendees. It was reported as a very good opportunity for CoEs to interact, get together and organize another relevant dissemination activity. Opportunities for interactions, including training activities, how to access technologies were presented. This session raised interest from several complementary initiatives. During the preparation of the session, collaboration opportunities with Stairway and BonsAPPs initiative emerged and were discussed among the participating CoEs. These initiatives were highlighted during the webinar. After the webinar possible synergy with the EIC project Carouseldansing was also spotted. Since the activity took place on the 17th of March 2022, at the time this report was written it was too early to report the further impact of those interactions.

In parallel to the sectorial events and following the review, two best practices guides[7] were produced and shared with the CoEs to share best practices and to be used as reference by the CoEs for their own events, or to attend sectorial events on their own (details about these best practices guide are available in D3.4).

Globally, all the CoEs were very satisfied of the support and guidance provided by WP3 through the sectorial events. They benefited from the assistance and budgetary support for their events, and appreciated the best practices guides.

Task 3.2 continued until the end of the extension in March 2022 to ensure that all the planned sectorial events targeted within WP3 and in interaction with the CoEs were attended or organized.

3.3 Task 3.3: Coordination with HPC SME initiatives

Task 3.3 had the goal of helping CoEs with their outreach activities towards industry, specifically regarding SMEs. Task 3.3 examined channels that were already in place within each CoE, such as associate partnerships, business development and dissemination, and assess their efficacy in SME outreach. It was also part of Task 3.3 to assess why existing programs like Fortissimo have helped in the past.

Therefore, during the first reporting period Task 3.3 conducted a quantitative online survey as well as a heavily qualitative interview series among the CoEs to find answers to these questions and be able to offer better guidance and help (see details in D3.2 and D3.3[8]). The Survey "Action Towards SMEs" therefore had two parts: the first was an online questionnaire containing 23 closed-ended as well as some open-ended questions. The second part was conducted via telephone interviews after the successful completion of the online questionnaire to gather more details.

[7] https://www.hpccoe.eu/download/

[8] FocusCoE Deliverable D3.3 "Report of the actions towards SMEs"

Following the midterm project review and the analysis of the ongoing CoEs' industrial interactions, and co-operations with SMEs (or SME support programmes) in particular, Task 3.3 took the approach to support the CoEs by providing guidance on how to interact with SMEs most effectively. Based on the survey results, online webinars have been organized and conducted to share success stories and lessons learned among the CoEs, focusing on how to most successfully reach out and partner with businesses, especially SMEs. Furthermore, the webinars provided opportunities for relevant European initiatives – FF4EuroHPC, PRACE SHAPE, EuroCC and CASTIEL – to present their activities and experiences to the CoEs as a precursor to direct discussions and the sharing of relevant information. This was finalized by facilitating the sharing of their individual lessons learned as well as providing information on best practices via a "Lessons learned […]" document shared with all CoEs (see details in D3.3).

Task 3.3 also established a collaboration with EuroCC's NCCs through CASTIEL and shared directly with the NCCs the document '"Lessons learned […]"summarizing the gathered information regarding interactions towards SMEs. These lessons learned as well as the useful data from that report will help the NCCs trying to build up relationships to SMEs. This activity was also synergetic to the further CASTIEL activities organising direct communication between the NCCs and the CoEs on several topics (activity managed by FocusCoE-WP2).

Task 3.3 was closed at the end of M36 with the submission of the deliverable D3.3 which presents all the actions performed and the outcomes of Task 3.3.

3.4 Task 3.4: Industrial feedback to the CoEs

The objective of Task 3.4 was to gather the needs and requests coming from industry that could be answered by the CoEs and to feed back this information to the CoEs. With this purpose, a questionnaire targeting industrial users (from both large companies and SMEs) was formulated and launched on March 2nd 2020. The analysed results, including prospects collection and a best practices guide about how to get in contact with industry and how to monitor a CoE's interactions with industry, were shared in individualized reports to each CoE active in Summer 2020 (see all details about the methodology and the outcomes in D3.2 – in the updated part of the deliverable re-submitted in Summer 2021).

The reports sent to all the CoEs had a common generic first part with the results common to all CoEs and an individualized second part with results specific to the CoEs, to its sector with specific advice and best practices recommendation and when there were any, new prospects contacts which were collected through the questionnaire.

The common generic section including best practices recommendation was also shared with the newly awarded Centres of Excellence (PerMedCoe, Raise, Nomad2, Trex) and the National Competence Centers (NCC) of the EuroCC project in order to share the useful common results with the NCCs and to start creating links with this new EuroHPC project.

Feedback received about these reports were varied, and those who responded were in general very thankful and positive: for instance, a representative from Romanian NCC indicated that "The FocusCoE survey came like a golden mine JIT (Just In Time). It certainly provides a lot of help for us, as we started to create our own survey but had big challenges in deciding and formulating the exact questions related to usage (or potential usage) of HPC/HPDA/AI and other technologies. We strongly needed a similar model to inspire from so it is more than helpful."

Some constructive improvements were also suggested on the type of questions, improvements on wording or some missing topics (such as data security or AI/ML related topics) – which

were addressed by organizing the sectorial webinar on AI related topics. We also took the chance to set up several telcos (EuroCC Spain, BioExcel, Cheese, COEC, PerMedCoE) to discuss the questionnaire during the second half of 2021.

The perceived usefulness of the questionnaire varied depending on the maturity of the CoE / CC. It proved to be most useful for those CoEs / CCs that were starting, rather for those who were more mature (in the second project or CC from institutions with long lasting experience dealing with industry). The more mature CoEs / CC have been already undertaken such activities or they have business development specialists. They were also instrumental providing relevant insights to the questionnaire, helping in the promotion, etc. One of the sections of the questionnaire was prepared in coordination with each of the CoEs in order to tailor the questions to their specific needs in their different domains. Hence each of the CoEs had the opportunity to get more in-depth results addressing those specific findings they needed. For the new CoEs and CC only the generic part was shared with because the CoE personalized sections were perceived as too specific (e.g. questions related to specific tools / technologies used by each of the CoEs),

The best ways to collect meaningful data to feed the questionnaire were also discussed, in terms of finding the right target audience answering the questionnaire, in the right variety of responders and expertise. Therefore, the first set of general questions were designed to understand the demographics of the answers requested and hence how to understand how meaningful they were. While preparing the questionnaire it was also of paramount importance to define the expected responders to contact and a strategy was prepared on how to reach them (dissemination of the questionnaire through the WP3 industry contacts, use of the FocusCoE dissemination channels, etc.). In Annex I of the questionnaire, available in Annex 7.2 of this deliverable, some relevant guidelines for interaction with industry and inbound marketing was collected and also shared with all CoEs & CCs.
The best sources of responses were coming from FocusCoE's close network and especially from those partners well-experienced in getting in contact with industry, and hence having many direct, as well as 2nd or 3rd level contacts from those. Specific technical and sectorial networks and industrial associations and initiatives from the EU "New Member States" were also approached but the response rate was significantly lower. As a result, from the feedback gathered and the discussions with CoEs we could also highlight following recommendations for the CoEs:

- First recommendation is to ensure that the exploitation team includes specialized business development expertise in order to be able to understand the importance of preparation of such studies, and how to maximize the return from gathering data from the right people.
- The second recommendation is to take advantage of any public activity where the CoEs are taking part, and ask such questionnaires to be filled during these activities – including sectorial events, participation in tradeshows, presentation, trainings attended by the relevant people.

Task 3.4 also gathered HPC challenges proposed by industrial companies, and promoted them to all the CoEs (and any other research institution) that would like to try to provide a solution. Initial challenges were collected as part of the specific section of the industry questionnaire created in interaction with each of the CoEs in 2021. The CoEs were requested to provide an initial list of challenges they were facing. The list was ranked by the questionnaire responders and completed by the questionnaire responses and then shared back to the CoEs. Following are the challenges that were identified for each of the CoEs:

BioExcel

- Virtual screening for lead discovery
- Antibody engineering and optimisation
- Kinetics of interactions
- Dynamics of molecular complexes
- Molecular determinants of QSAR
- Virtual generative chemistry for lead discovery
- Hit expansion for lead discovery
- Precise binding energy prediction in lead optimisation
- Off-target molecular modelling

CHEESE

- Full-waveform inversion
- High-resolution subsurface imaging
- Near real-time seismic scenarios
- Physics-based probabilistic seismic hazard analysis
- Earth's magnetic field evolution
- Volcanic plumes and pyroclastic density currents
- High-resolution volcanic ash dispersal
- Subsurface thermo-fluid dynamics of magmas
- Probabilistic volcanic hazard analysis
- Faster than real time tsunami computations
- Near real-time tsunami source inversion
- Probabilistic tsunami forecast for early warning and rapid post-event assessment
- Probabilistic tsunami hazard analysis
- Automated array-based statistical detection and restoration of seismic slow-earthquakes

CompBioMed

- High fidelity modelling and simulation of human body (multiscale aspects, validation, verification and uncertainty quantification of models, HPC performance, etc.)
- Access to medical data
- Data security
- Data assimilation and analysis

E-CAM

- Calculation of large systems (thousands of atoms)
- High throughput of new molecules
- Multi-scale coupling
- Bridging timescales
- Simulation of complex chemical reactivity

EoCoE

- Wind simulation over complex terrains
- Simulation of the full rotor model including the actual geometry of the wind turbine blades and tower
- Weather forecast at high resolution and/or global scale

- Continuous probabilistic short-term predictions of optical thickness and wind
- Wind and solar power calculation from meteorological ensembles
- Design of new functional materials
- Modeling and simulating hetero-interfaces in photovoltaic cells
- Modeling electricity production from salinity gradients
- Modeling electricity production from temperature gradient
- Simulating organic and perovskite solar cells
- Construction of hydrologic and/or geothermal models at high resolution over large spatial scales
- Hydropower modeling
- Experimental design of geothermal power
- Modeling geothermal reservoir systems
- Simulation of the turbulent transport and confinement properties of plasma

ESiWACE

- Weather forecast at high resolution
- Weather forecast at global scale
- Climate forecasts at high resolution and for long time scales
- Convection-resolving simulations
- Storm-resolving simulations
- Coupled Atmosphere-Ocean Simulations

EXCELLERAT

- Mesh generation
- Boundary conditions
- Scalability
- Data handling
- Multi-phase computations
- Multi-disciplinary-optimisation
- Multi-physics coupling
- Multi-scale coupling
- Long compute times
- Large data sets
- Difficult validation
- Availability of the right simulation software

HiDALGO

- Crowd dynamics simulation
- Real-time data analytics and visualization
- Workflow design and management
- Coupling simulations from different subdomains
- Integration of HPDA within HPC
- Application lifecycle handling using AI
- Integration of sensor data into simulations' execution

MaX

- Calculation of large systems (thousands of atoms)

- High throughput of new molecules
- Multi-scale coupling
- Bridging timescales
- Simulation of complex chemical reactivity

POP

- Increasing the size of the datasets that can be run
- Increasing the accuracy or resolution of jobs
- Increasing the speed with which results can be obtained
- Adding new functionality without impacting performance

PerMedCoE

- Drug synergies to COVID-19
- Diagnosis Based on omics Information
- Personalised modelling of groups of rare-disease related patients

RAISE

- Methods to generate "realistic" data – when real-time data-collection fails
- Scalability of AI methods & tools
- Parallelizing to decrease computation time while keeping the correctness of the result
- Efficient execution of complex workflows
- Data Protection

Complementary to the challenges collection – and worth mentioning: most of the CoEs have their own ways to address those challenges. An important way to address the challenges this is through the use cases that are planned and executed by the CoEs. Use-cases are concrete cases where the codes and machines of the CoEs are used to get state-of-the-art solutions of typical domain-specific problems. Use-cases show the power of the HPC resources in a real-world context and may inspire users for their specific application cases, and hence become a very powerful tool for CoEs to showcase how those challenges can be addressed. This is why, in collaboration with WP5 of FocusCoE a listing of the relevant use cases was collected and has been highlighted in the Use Cases section of the website in order to promote the CoEs actions and solutions for actively tackling the industry challenges: https://www.hpccoe.eu/use-cases/. Also, the success stories included further details of the challenge that initiated the solution. In many cases the success stores are the successful completion of use-case Success Stories (hpccoe.eu) .

During the second half of FocusCoE, task 3.4 completed the work initiated with the industry questionnaire with the organization of thematic webinars on industrial challenges on selected topics. The main idea was to cover challenges from new CoEs that would not have been covered previously and discuss them among different CoEs and related industry stakeholders (ranging from big companies to SMEs). Industrial challenges and opportunities were discussed among several CoEs in the sectors of Bio / Medicine / Pharma (HiPEAC sessions in May 2021) and AI (AI WebCafe March 2022) to foster interactions between companies and solution providers (in this case the CoEs).

Task 3.4 was closed at end of M36 as planned in the DoA.

4 Alternative solutions and additional plans set up by WP3 following the COVID-19 outbreak

During the lifetime of the project, WP3 initiated several alternative solutions and additional plans to continuously support and assist the CoEs in their interaction with industry despite several issues and obstacles encountered between 2020 and 2022. First, there seem to have been a shift regarding the main priorities of the CoEs. In agreement with the directions and the priorities expressed by the European Commission in 2020, the projects were very much concerned with delivering applications for exascale, in particular to the scientific community. Therefore, topics related to "services for industry" and more generally interaction with industry seemed less in the focus of the CoEs' efforts. Second, the COVID-19 pandemic strongly impacted WP3's activities, hence WP3 swiftly adapted the work and interactions with both the CoEs and industry. Indeed, due to the repeated postponements of the physical events during the period 2020-2022, alternative additional actions had to be considered and organized by WP3 in order to continue sectorial communication activities for the CoEs despite the practical impossibility of participating at physical sectorial events. The first tested alternative was participation in online events. Besides the online events, WP3 also organized a series of webinars, in Winter 2021-Spring 2022, as a replacement for some planned sectorial events. This alternative solution was tested to see how this could foster contacts to new potential users and clients for the CoEs. In the organization of these thematic webinars, one main challenge was in how to gather numerically relevant audiences. This was done by widely disseminating the information about them through the FocusCoE website and social media, the relevant CoEs' websites and social media, via FocusCoE and participating CoEs' newsletters. In addition to these classical channels, the WP3 partners also invited their own industry contacts from the relevant industry sectors.

As a first conclusion, these new types of "sectorial events" (i.e. the online events or the thematic webinars), were a little less efficient than the physical typical events to provide industrial prospects to the CoEs. It is indeed always more difficult to interact with the speakers or participants in an online event and the online format is always more formal and rigid than a face-to-face physical meeting. In an online event there is no space for the kinds of informal contacts and conversations that are common in an on-site event, and the virtual "discussion rooms" that have been tested in several events to simulate these types of contacts have proven to be less effective than real ones. However, given the issues of the COVID-19 pandemic active during 2020-2022 and the continuously changing travel regulations (depending on each country), these solutions presented the best alternatives. The KPIs available in D3.4 show that even if travel and real meetings were not possible, these alternative options enabled FocusCoE to put the CoEs in contact with many potential industrial prospects and for each event, resulting with at least a couple of serious potential industrial prospects or collaborations. The key elements for a successful event (e.g., in terms of interactions with visitors, of getting new prospects and of ensuring a good visibility) to be checked upfront for successful events are detailed in D3.4.

5 Final conclusions and final feedbacks from WP3 with regards to the CoEs' interactions with industry

During the FocusCoE project, WP3 focused on developing interactions between the CoEs and industry. Through Task 3.1, this was performed by defining at first a methodological approach of using existing contacts of the WP3 partners with industrial companies and then

by developing a strategy to create new contacts with stakeholders from the industry, either already integrated in the HPC ecosystem or not yet part of this ecosystem. In parallel and in coordination with WP5, 9 success stories of interactions of CoEs with industry were created and published.

Through task 3.2, the implementation of the strategy was realized to promote the CoEs' competences and service offerings to new industrial contacts during the sectorial communication activities. Even if the task was strongly impacted by the COVID-19 crisis for physical events, the WP3 team successfully adapted its strategy with the CoEs' approval to focus on participating to online and hybrid events. In addition, WP3 produced two Best Practices guides shared with all CoEs to be used as reference when interacting with industrial companies, including SMEs during sectorial events, be it physical or online.

For task 3.3, the focus was on helping the CoEs to interact with SMEs specifically; thanks to an online survey and interviews realized during the first months of the FocusCoE project, WP3 supported the CoEs by providing guidance on how to interact with SMEs via online webinars. It was also an opportunity for relevant European initiatives to present their activities and experiences to the CoEs with regards to interactions with SMEs. The exchanged lessons learned and best practices were shared with all CoEs with the production of a "Lessons learned with regards to interactions with SMEs" document to conclude this task.

Within task 3.4, the WP3 team gathered the needs and requests coming from industry to be answered by the CoEs and provided feedback to the CoEs afterwards. With the analysis of the questionnaire targeting industrial users, WP3 could produce for each CoE an individual report personalized to its activities and needs with specific recommendations, contacts for potential prospects and also promotion of the industrial challenges. The common part of this reports was also shared with the NCCs from the EuroCC project to enhance general understanding of the industry challenges and the EU projects' interactions with industry.

Challenges faced by WP3 included the priority shift from the "services for industry" topic to a "developing applications for exascale" topic following the orientation changes initiated by the European Commission. This resulted in the CoEs being focused less on industry interactions and more on exascale readiness. Then there was the impact on WP3 activities and sectorial events by the COVID-19 pandemic. Therefore, WP3 had globally to adapt the initial developed strategies and to organize new types of events or to provide best practices guides to support and guide the CoEs through the new types of challenges encountered when they try to get in contact with industry and when they interact with SMEs particularly.

Regarding the new types of online sectorial events, it was noticed that while they were a little less efficient than physical events, they still provided a way of interacting with industry for the CoEs during a time of travels restrictions, but in the case that key elements are ensured for a successful even (e.g. a reliable efficient digital platform for the event, a communication strategy organized months in advance, etc.) were respected by the organizers of the events and by the CoEs, they still represented opportunities to form new industry prospects. With this global adaption to the new needs and new channels of interaction, FocusCoE managed to create several interactions between the CoEs and industrial companies, and during each event organized or coordinated by FocusCoE WP3, potential industrial prospects were made or collaborations were launched for the CoEs. At the conclusion of the project, the CoEs indicated that they appreciated the support and guidance provided by FocusCoE with regards to the topic of "interactions with industry" and the actions undertaken by the WP3 team were fruitful and useful for the CoEs.

6 References

[1] FocusCoE Deliverable D3.1 "Report on the sectorial approach priorities"
[2] FocusCoE Deliverable D3.2 "Interim report on the WP3 activities"
[3] FocusCoE Deliverable D3.4 "Report on the sectorial communication activities"
[4] Enginsoft The Simulation Based Engineering and Sciences Magazine Newsletter https://www.enginsoft.com/assets/pdf/newsletter/newsletter2021_1.pdf#page=52
[5] FocusCoE Webinar: Opportunities and Challenges for Industrial Applications https://www.hpccoe.eu/2021/03/05/focuscoe-webinar-opportunities-and-challenges-for-industrial-applications/
[6] Recording of the HipEAC session in 2021: https://www.youtube.com/watch?v=CNa2j7HAxmU&feature=youtu.be
[7] Download webpage of the FocusCoE website : https://www.hpccoe.eu/download/
[8] FocusCoE Deliverable D3.3 "Report of the actions towards SMEs"

7 Annexes

7.1 Success stories

Release date	CoE(s)	Title	URL
06.10.2020 (on new website)	CompBioMed	Drug Discovery with Janssen Pharmaceutica NV	https://www.hpccoe.eu/2020/10/06/drug-discovery-janssen-pharmaceutica-nv/
06.10.2020 (on new website)	EoCoE	Renewable Energy - harness the power of vorticity	https://www.hpccoe.eu/2020/10/06/industry-sector-renewable-energy/
06.10.2020 (on new website)	E-CAM	Designing control pulses for superconducting Qubit systems with local control theory	https://www.hpccoe.eu/2020/10/06/designing-control-pulses-for-superconducting-qubit-systems-with-local-control-theory/
12.10.2020	ESiWACE	Improving weather and climate forecasting with a new NEMO configuration	https://www.hpccoe.eu/2020/10/12/coes-success-stories-focus-on-esiwace-improving-weather-and-climate-forecasting-with-a-new-nemo-configuration/
02.12.2020	MaX	AiiDA Platform Accelerates Materials Discovery	https://www.hpccoe.eu/2020/12/02/coe-success-stories-focus-on-max-coe-nccr-marvel-aiida-platform-accelerates-materials-discovery/
19.03.2021	HiDALGO	Assisting decision makers to solve Global Challenges with HPC applications – Migration issues	https://www.hpccoe.eu/2021/03/19/hidalgo-assisting-decision-makers-to-solve-global-challenges-with-hpc-applications-migration-issues/
19.03.2021	HiDALGO	Assisting decision makers to solve Global Challenges with HPC applications – Covid-19 modelling	https://www.hpccoe.eu/2021/03/19/hidalgo-assisting-decision-makers-to-solve-global-challenges-with-hpc-applications-covid-19-modelling/
23.07.2021	EXCELLERAT	Enabling High Performance Computing for Industry through a Data Exchange & Workflow Portal	https://www.hpccoe.eu/2021/07/23/excellerat-enabling-high-performance-computing-for-industry-through-a-data-exchange-workflow-portal/
23.07.2021	EXCELLERAT	Bringing industrial end-users to Exascale computing: An industrial level combustion design tool on 128K cores	https://www.hpccoe.eu/2021/07/23/excellerat-bringing-industrial-end-users-to-exascale-computing-an-industrial-level-combustion-design-tool-on-128k-core/
28.10.2021	POP & PerMedCoE	POP and PerMed Centers of Excellence are Getting Cell-Level Simulations Ready for Exascale	https://www.hpccoe.eu/2021/10/28/pop-and-permed-centers-of-excellence-are-getting-cell-level-simulations-ready-for-exascale/

CoE involved:

CompBioMed is a user-driven Centre of Excellence in Computational Biomedicine. They have users within academia, industry and clinical environments and are working to train more people in the use of their products and methods.

Organisations & Codes Involved:

SURFsara, the domain expert, is the National Supercomputing and e-Science Support Center in the Netherlands. SURFsara provides expertise and services in the areas of High Performance Computing, e-Science & Cloud Services, Data Services, Network support, and Visualisation.

Janssen Pharmaceutica NV, the industrial partner, is an affiliate of the Pharmaceutical branch of the US Johnson & Johnson company. Janssen's interests are in developing and using advanced molecular simulation methods to optimize lead compounds in discovery programs, predicting the activity of compounds with specific targets.

CHALLENGE:

Janssen's primary challenge is in developing and using advanced molecular simulation methods to optimize lead compounds in discovery programs. Such methods, if proven robust and accurate could have a profound impact on the way drug discovery is performed. They would permit reliable computational triaging of very close analogue molecules greatly improving efficiency. Also, this would lead to high-confidence design of synthetically more challenging molecules leading to better drugs in new chemical space.

SOLUTION:

Through the CompBioMed HPC allocations service JAN obtained an allocation of 1.8M core/hours on Cartesius (SURFsara). JAN evaluated the use of non-commercial software for predicting the free energy of a series of compounds and built a protocol which uses open-source software for MD and free energy perturbation simulations. The main simulation code used was GROMACS (version 2016.1) configured to run on GPU, and the results compared to previous simulations performed with Schrodinger's FEP+ commercial software. Slurm support for job array and job dependencies, has been used to orchestrate and streamline the workflow and to optimise the MD simulations execution on multiple compute nodes.

Scientific impact of this result:

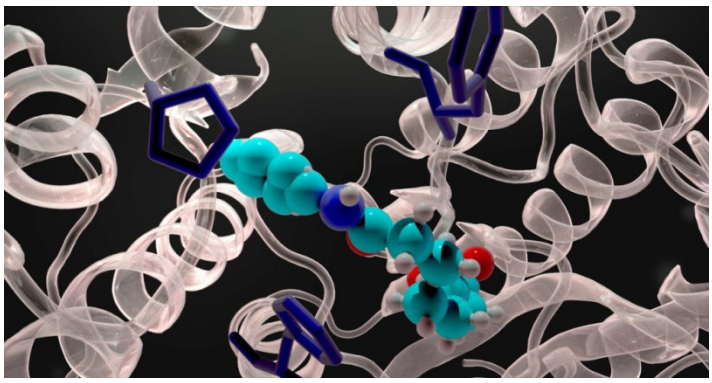
New computational approaches described above can have a major impact on the drug discovery process. Importantly the work studied here will go to the heart of the design-make-test cycle and contribute higher quality methods for compound prioritization.

This is a fundamental issue of drug discovery, because, whilst compounds with acceptable potency can often be found, and found quickly, they do not always come with other desired properties. Hence, during a typical lead optimization (LO) program the challenge often becomes to maintain potency whilst modifying the chemical structure of the lead molecules to overcome these other issues. In this regard computational tools which can accurately predict binding mode combined with accurate binding affinity prediction will be extremely powerful and reduce the number of 'backwards' steps required to subsequently move forwards in an LO program.

The project has introduced Janssen to numerous research groups throughout Europe that are able to assist them in their aims and to combine expertise to produce useful results. The use of HPC is critical for these experiments, without which the compute power does not satisfy the needs of such complex systems.

Benefits for further research:

- Higher quality methods for compound prioritization.
- Increased efficiency during lead optimisation program.
- Reduction in time/cost of synthesising redundant drug molecules
- Higher probability of determining new active drug



Renewable Energy - harness the power of vorticity

CoE involved:

EoCoE is an energy-oriented Centre of Excellence for computing applications that builds on its unique expertise at the crossroads of high-performance computing (HPC) and renewable energy. It brings an impulse to accelerate the digitization of the future energy system. The coding developments are assisted by multi-disciplinary teams with expertise in applied mathematics and high performance computing (HPC).

Organisations & Codes Involved:

The end user VORTEX BLADELESS is a Spanish tech startup that is developing an environmentally friendly aerogenerator which needs no blades. It is a new wind energy technology specially designed for on-site generation in residential areas, being able to work on-grid, off- grid, or along with regular solar panels or other generators.

The domain expert, the BARCELONA SUPERCOMPUTING CENTRE is a Spanish public research center that developed the multi-physics simulation code ALYA and provided the MareNostrum supercomputer.

CHALLENGE:

Vortex Bladeless aims to harness the power of vorticity for a new generation of wind turbines. They develop a single column without bearings or gears. It just oscillates with the wind. Experiments with scaled down prototypes have been encouraging, but the physics behind these devices is highly complex. There is a need to optimize and explore scalability due to the complexity of the flow and the need for time accurate results. The needed Large Eddy Simulation (LES) simulations are computationally demanding.

SOLUTION:

The company has been working with experts at the Barcelona Supercomputing Centre on the MareNostrum supercomputer. The fluid-structure interaction (FSI) between the Vortex Bladeless device and a turbulent flow is simulated with Alya.

The results from initial simulations of a scaled-down device were very close to the actual wind tunnel tests performed by the Vortex Bladeless team, allowing them to develop the idea of a range of devices at the micro scale and the utility scale. Then, the behaviour of the device at a more realistic scale was studied by means of numerical simulations, helping in the design of real scale experiments and reducing costs.

Business impact:

In order to understand the aerodynamic behaviour of the devise, experimental, or numerical studies can be
performed. For the real scale model, that is around three meters high, experimental studies would have been too expensive.

It was therefore decided to resort to high accuracy numerical simulations of the flow coupled to the oscillating turbine. Since the flow is highly transient, and it is important to capture the dynamically important scales of the flow accurately, Large Eddy Simulation (LES) was selected as the modeling technique.

Modeling of turbulent flow is still nowadays one of the most computational demanding problems. The interaction with the Barcelona Supercomputing Centre was, therefore, crucial to develop a much better understanding of the aerodynamics of the device. BSC provided the computation resources and the inhouse code Alya that can efficiently run on high-end Supercomputers.

Moreover, they contributed with their know-how on LES turbulence modeling and Fluid-Structure Interaction. Technological advantages the HPC provides are important for small to medium-sized enterprises (SME) to remain competitive. For instance, the savings regarding the CPU time are key for SMEs to be able to use highly advanced techniques such as Large Eddy Simulation.

Benefits for further research:

- Reduced costs by preparing the real scale experiments by means of numerical simulations
- The highly optimized code enabled to avoid throwing away costly computational resources
- The cpu time for assembly has been reduced up to 38%.
- A new solver has provided speed ups of up to five times with respect to Alya's own solvers.

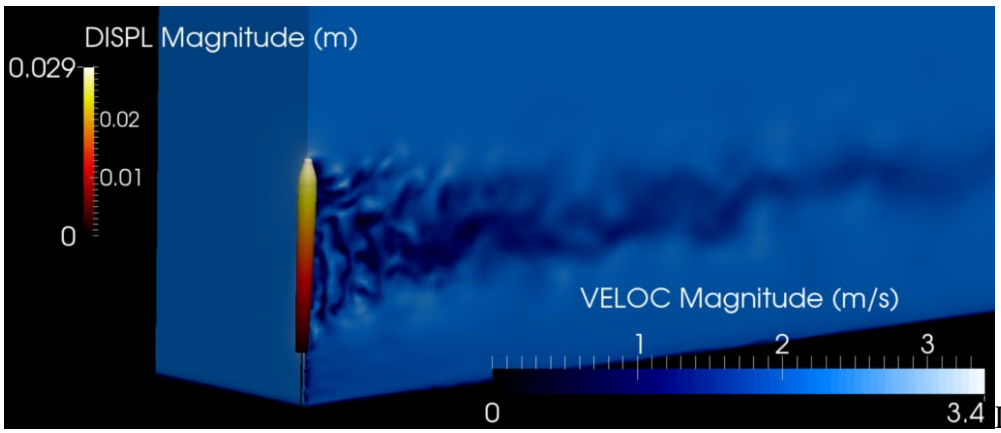


Image: Simulation of the fluid-structure interaction (FSI) between the Vortex Bladeless device and a turbulent flow from Alya

Designing control pulses for superconducting Qubit systems with local control theory

CoE involved:

The European HPC Centre of Excellence (E-CAM) is an e-infrastructure for software development, training, and industrial discussion in simulation and modelling that started in October 2015. E-CAM focuses on four scientific areas of interest to computational scientists: Classical Molecular Dynamics, Electronic Structure, Quantum Dynamics, Meso- and MultiScale Modelling

Organisations & Codes Involved:

CECAM Centre Européen de Calcul Atomique et Moléculaire (Host beneficiary), located at the EPFL in Lausanne, is an organization devoted to the promotion of fundamental research on advanced computational methods and to their application to important problems in frontier areas of science and technology.

IBM Research Laboratory – Zurich (Industrial partner) is the European branch of IBM research, which is the research and development division of the American multinational information technology company IBM.

CHALLENGE:

The aim of this pilot project was to develop a new method and dedicated software for designing control pulses to manipulate qubit systems (see Fig.1A) based on the local control theory (LCT) The system is composed of two fixed frequency superconducting transmon qubits (Q1 and Q2) coupled to a tunable qubit (TQ) whose frequency is controlled by an external magnetic field. Changing the frequency, the TQ behaves as a targeted quantum logic gate, effectively enabling an operation on the qubit states. The system schematizes an approach to construct real quantum universal gates currently investigated by IBM.

SOLUTION:

Local control theory (LCT), the main theoretical tool used, originates from physical chemistry where it is used to steer chemical reactions towards predetermined products, but it had never been used to design a quantum gate. To create the software, researchers added new functionalities to the open source QuTip program package. Two main modules were developed during the project: *LocConQubit*, which implements the LCT and accompanying procedures, and *OpenQubit*, a patch to the first module which introduces Lindblad master equation propagation scheme into the LCT which also enables direct construction of pulses under the presence of decoherence effects. All modules were written in Python and expand the functionalities of the QuTip program package.

Business impact:

In order to understand the aerodynamic behavior of the devise, experimental, or numerical studies can be performed. For the real scale model, that is approximately three meters high, experimental studies would have been too expensive. It was therefore decided to resort to high accuracy numerical simulations of the flow coupled to the oscillating turbine. Since the flow is highly transient, and it is important to capture the dynamically important scales of the flow accurately, Large Eddy Simulation (LES) was selected as the modeling technique.

Modeling of turbulent flow is still nowadays one of the most computational demanding problems. The interaction with the Barcelona Supercomputing Center was, therefore, crucial to develop a much better understanding of the aerodynamics of the device. BSC provided the computation resources and the inhouse code Alya that can efficiently run on high-end Supercomputers. Moreover, they contributed with their know-how on LES turbulence modeling and Fluid-Structure Interaction.

Technological advantages the HPC provides are important for small to medium-sized enterprises (SME) to remain competitive. For instance, the savings regarding the CPU time are key for SMEs to be able to use highly advanced techniques such as Large Eddy Simulation.

Benefits for further research:

- Reduced costs by preparing the real scale experiments by means of numerical simulations
- The highly optimized code enabled to avoid throwing away costly computational resources
- The cpu time for assembly has been reduced up to 38%.
- A new solver has provided speed ups of up to five times with respect to Alya's own solvers.

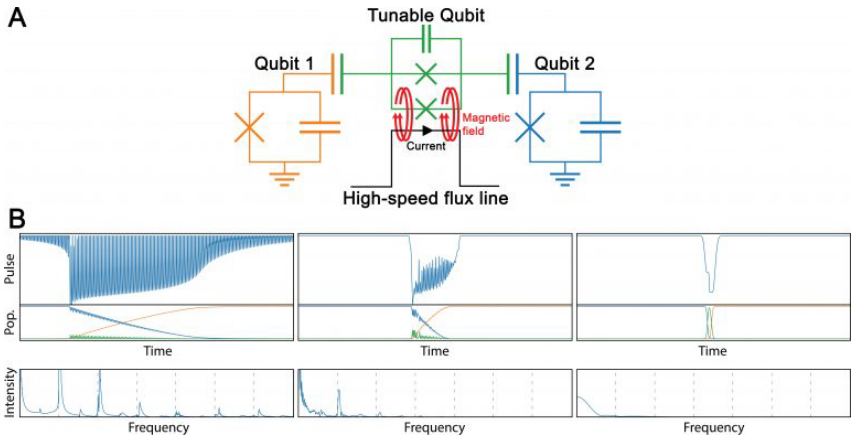


Figure 1: A) Schematic representation of the qubit system; B) Pulses obtained by LCT (left), by a frequency filtering LCT procedure (middle), and by analytic function fitted to LCT pulse parameters (right) with their corresponding population transferring from qubit 2 (blue) to qubit 1 (orange), and frequency spectra.

Improving weather and climate forecasting with a new NEMO configuration

Highlighted Centre of Excellence

ESiWACE, the "Centre of Excellence in Simulation of Weather and Climate in Europe" has been funded by the European Commission to substantially improve efficiency and productivity of numerical weather and climate simulation. Overall, the Centre of Excellence prepares the European weather and climate community to make use of future exascale systems in a co-design effort involving modelling groups, computer scientists and HPC industry. .

Organisations & Codes Involved:

Atos Center for Excellence in Performance Programming (CEPP) provides expertise and services in High Performance Computing, Artificial Intelligence and Quantum Computing.

LOCEAN laboratory, part of the CNRS-IPSL, conducts studies on the physical and biogeochemical processes of the ocean and their role in climate in interaction with marine ecosystems.

CERFACS research center is specialized in modelling and numerical simulation, through its facilities and expertise in high-performance computing.

CHALLENGE:

A key model for weather and climate comprehension is NEMO (Nucleus for European Modelling of the Ocean), a modelling framework for research activities and forecasting services in ocean and climate sciences. NEMO is used by a wide variety of applications at global or regional focus, with different resolutions, different numerical schemes, parameterizations and therefore with different performance constraints. The technical challenge here was to find a way to ease the profiling and benchmarking of NEMO for its versatile uses in order to increase the performance of this framework.

SOLUTION:

In response to this challenge, ESiWACE has developed a new configuration, adapted to HPC benchmarking, that is polymorphic and can very simply reproduce any use of NEMO.
Thanks to a close collaboration between the Atos CEPP, LOCEAN and CERFACS, a dedicated configuration of NEMO to ease its profiling and benchmarking has been set up.
Many tests of this configuration, including large-scale experiments on the Atos Bull supercomputers at Météo France and CEA's Very Large Computing Centre (TGCC) have been performed. This resulted in several optimisations improving the performance of the NEMO code in this configuration by up to 38%.

The open source NEMO 4.0, which was released in 2019, benefited from this work and included the following improvements: an automatic MPI sub-domain decomposition, and a rewriting of the communication routines with an optimisation of treatment of the North pole singularity.

Business impact:

Some of the NEMO uses such as weather or climate forecasts are among the key challenges our society must address. Improvements of NEMO numerical performance allow to refine model results, to reduce forecast uncertainties and to better predict high-impact extreme events, thus saving lives. On Earth the Ocean has a huge impact on the atmosphere. Thus, NEMO is widely used coupled with atmospheric models. For example, it is used to simulate ocean eddies, temperature and salinity that play a key role in cyclone intensity and trajectory forecast. Therefore, business impacts of improving NEMO may be indirect but there are significant as they concern everyone and all kinds of companies and entities.

This work benefits society by improving the efficiency and productivity of numerical weather and climate simulation and by preparing them for future exascale systems. It fosters expertise exchanges which enable researchers and industry to be more productive, leading to scientific excellence in Europe, through direct and tight collaborations.

Benefits for further research:

- Simplified profiling and benchmarking of NEMO at different scales to find the most relevant optimisation paths.
- Up to 38% efficiency and scalability increase of NEMO with the optimized configuration on HPC systems.
- Reduction in time/cost of ocean simulations for both research and production purpose improving weather and climate predictions, allowing to protect property, interests and saving lives.

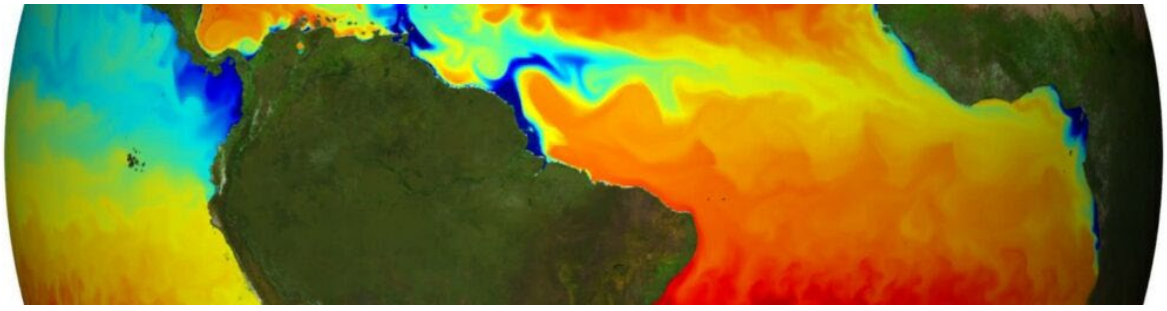


Image: Example of NEMO High Resolution Simulation (ORCA 1/12°) showing here the Sea Surface Salinity in both Atlantic and Pacific Ocean on the equator. Source: https://www.nemo-ocean.eu/

AiiDA Platform Accelerates Materials Discovery

Highlighted Centre of Excellence

MAX (MAterials design at the eXascale) is a European Centre of Excellence which enables materials modelling, simulations, discovery and design at the frontiers of the current and future High Performance Computing (HPC), High Throughput Computing (HTC) and data analytics technologies.
>> Learn more about MAX

Quick Summary

- **Industry Sector Involved:**

  Materials for energy

- **Software and hardware used:**

  AiiDA (Automated Interactive Infrastructure and DAtabase for computational science), developed by MaX CoE and partners, including NCCR MARVEL and Bosch Research, on the Piz Daint supercomputer at Swiss National Supercomputing Centre (CSCS).

- **Challenge:**

  Finding new candidate materials for application as solid-state electrolytes in next generation batteries.

- **Solution:**

  A simple and efficient framework to predict the diffusion of lithium ions (Li ions) in solid-state materials, then using the AiiDA platform to employ it in a large-scale computational screening.

Organisations Involved:

NCCR MARVEL is a center on Computational Design and Discovery of Novel Materials created by the Swiss National Science Foundation in May 2014. MARVEL targets the accelerated design and discovery of novel materials, via a materials' informatics platform of database-driven high-throughput quantum simulations, powered by:

- advanced electronic-structure capabilities, for predictive accuracy
- innovative sampling methods to explore configuration/composition space
- application of big-data concepts to computational materials science.

Bosch Research and Technology Center: Founded in 1999, the North American division of Corporate Research at Bosch has been shaping the technology of Bosch's future for nearly 20 years. The team has worked in close collaboration with its colleagues and counterparts in Germany and around the world. The center is committed to providing technologies and systems for the four business sectors of Bosch — Mobility Solutions, Energy and Building Technology, Industrial Technology and Consumer Goods – by scouting and collaborating with top universities and industry partners in North America.

CHALLENGE:

Solid-state electrolytes have the potential to enhance both safety and performance of Li-ion batteries, allowing for novel cathode and anode chemistry, preventing the growth of Li–metal dendrites — the needle-like formations of lithium that grow inside batteries, causing devices to lose power more quickly, short out, or sometime even catch fire — and pushing the miniaturization of battery cells.

Despite intense research in this field for decades though, no known solid-state ionic conductor satisfies all the requirements needed for battery applications. This makes the search for new materials a worthwhile endeavor. Computational approaches in the search for new materials are less human-intensive and easily parallelizable and precede synthesis and characterization in the laboratory. Computational screening relies on simulations of the electronic structure, to determine the insulating character of a material, and molecular dynamics simulations to predict the Li-ion diffusion coefficients.

Overall, thousands of calculations need to be performed, making automatization and reproducibility a key requirement. In addition, methods need to be computationally inexpensive enough to be run on thousands of materials, yet accurate enough to be predictive.

SOLUTION:

We first reduced the computational burden of modelling the potential energy surface of lithium diffusing in a solid-state ionic conductor to develop a workable framework. We then demonstrated a procedure for running these extensive molecular dynamics simulations in a largescale computational screening. AiiDA made this possible by allowing the automation and explicit storage of the provenance. The novelty of AiiDA in the field of materials informatics is that every calculation is stored as a node in a graph, with input data forming incoming nodes, and output data stored as outcoming nodes, that can again be input to a different calculation.

In addition, AiiDA allows for a high degree of automation and parallelization via its daemon. Every calculation presented in the paper "High-throughput computational screening for solid-state Li-ion conductors" was run with AiiDA.

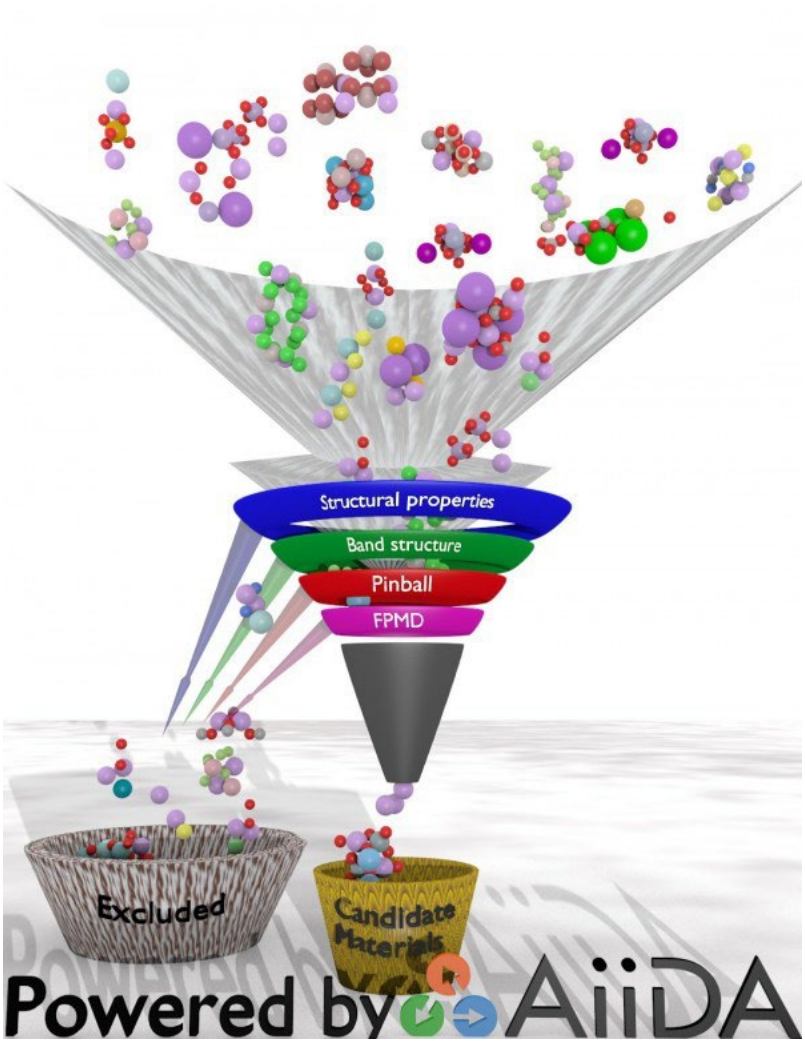
Business impact:

We found five materials with fast ionic diffusion, some in the range of the well-known superionic conductor Li10GeP2S12, such as for example the Li-oxide chloride Li5Cl3O, the doped halides Li2CsI3, LiGaI4, and LiGaBr3, or the Li-tantalate Li7TaO6. In addition, we found 40 materials that show significant diffusion at 1000 K, though they will need to be investigated more thoroughly before their suitability can be determined. All of these potential fast-ionic conductors could be studied further, in more detail, by experiment and simulation, and could result in new fast-ionic conductors or even electrolytes for next generation solid-state Li-ion batteries. Our data could also serve to search for descriptors of fast ionic conduction, which would be of significant interest to the community.

This work benefits society by identifying inorganic solid-state lithium-ionic conductor compounds that could be used as electrolytes to mitigate or overcome the severe safety challenges posed by the use of volatile and flammable liquid or polymer electrolytes in today's Li-ion batteries. Complete replacement of the liquid electrolyte by a solid ceramic would result in an all-solid-state Li-ion battery, highly beneficial due to the higher electrochemical stability of inorganic electrolytes, compared to their organic counterparts.

Benefits for further research:

- We developed efficient ways of simulating the diffusion of lithium in the solid state and gained physical insight into how charge-density rearrangements or lattice vibrations affect it.
- We developed a framework for predicting the diffusion of Li ions in solid-state materials and a process for applying it in largescale computational screening.
- We identified new ceramic compounds for in-depth experimental investigation



Related Images :

The figure shows a schematic representation of the screening funnel. Structures from experimental repositories go sequentially through several computational filters. Each stage of the screening discards unsuitable structures based on properties ever more complex to calculate. The final outcome is of a few tens of viable structures that could be potential candidates for novel solid-state Li-ion conductors.

Assisting decision makers to solve Global Challenges with HPC applications – Migration issues

HIGHLIGHTED CENTRE OF EXCELLENCE

HiDALGO is the Center of Excellence in HPC and Big Data technologies for Global Systems funded by the European Commission. Understanding global challenges to assist decision making by addressing multi-dimensional problems is a real need nowadays. HiDALGO develops novel methods, algorithms and software for HPC and HPDA enabling highly accurate simulations, data analytics and data visualisation to accurately model and simulate these complex processes.

Organisations & Codes Involved:

Save the Children is an NGO promoting policy changes to gain more rights for young people, especially by enforcing the UN Declaration of the Rights of the Child.

Brunel University London is a dynamic institution that plays a significant role in the higher education sector. It carries out applied research on different topics, such as software engineering, intelligent data analysis, human computer interaction, information systems, and systems biology.

CHALLENGE:

At the same time of the current COVID-19 pandemic, other crises have not stopped like forced migration due to conflicts. In fact, the number of forcibly displaced people is still very high, with over 70 million persons being forced to leave their homes. Save the Children provides support in these countries and needs more accurate estimations on people flows and even destinations to send the appropriate amount of help to the right place.

SOLUTION:

To tackle these challenges HiDALGO provides the tool Flee 2.0, a forced migration model which includes refugees and internally displaced people. It places agents that represent displaced persons in areas of conflict and uses movement rules to mimic their behavior as they attempt to find safety.

The code extracts location and route data from OpenStreetMap, and can be validated against UNHCR data for historical conflicts. As output, Flee provides a forecast of the amount (and location) of people that can be displaced given different conflict development scenarios.

SOCIETAL & ECONOMIC IMPACT:

Although it is not a market topic itself, global challenges have been gaining importance in the last years, especially those related to climate change but also others such as peace and conflict. HiDALGO offers a set of tools, services, datasets and resources to define models that may predict situations under certain scenarios that can influence any decision to be taken.

More specifically, Flee 2.0 provides the needed information to decision makers so they can set up the appropriate measure and provide the necessary means at any time.

For instance Flee 2.0 is used to estimate the expected number of refugee arrivals when conflicts occur in North Ethiopia, and to investigate how different conflict developments could affect the number of arrivals. The model development is partially guided by on-site experts and typically provides forecasts of approximately 3 month duration. Although work is ongoing on this project and the models are still basic, the aim is to establish a systematically developed mathematical model to improve the understanding of Save the Children about the migration situation, and support the preparation to mitigate the humanitarian impact of a potential upcoming crisis.

BENEFITS FOR FURTHER RESEARCH:

- Accurate predictions of where people may arrive in Sudan and how quicky, if new violence occurs in North
- Estimate of the effect of different conflict developments on the expected number of arrivals.
- Estimate of the effect of specific policy decisions on the expected number of arrivals.

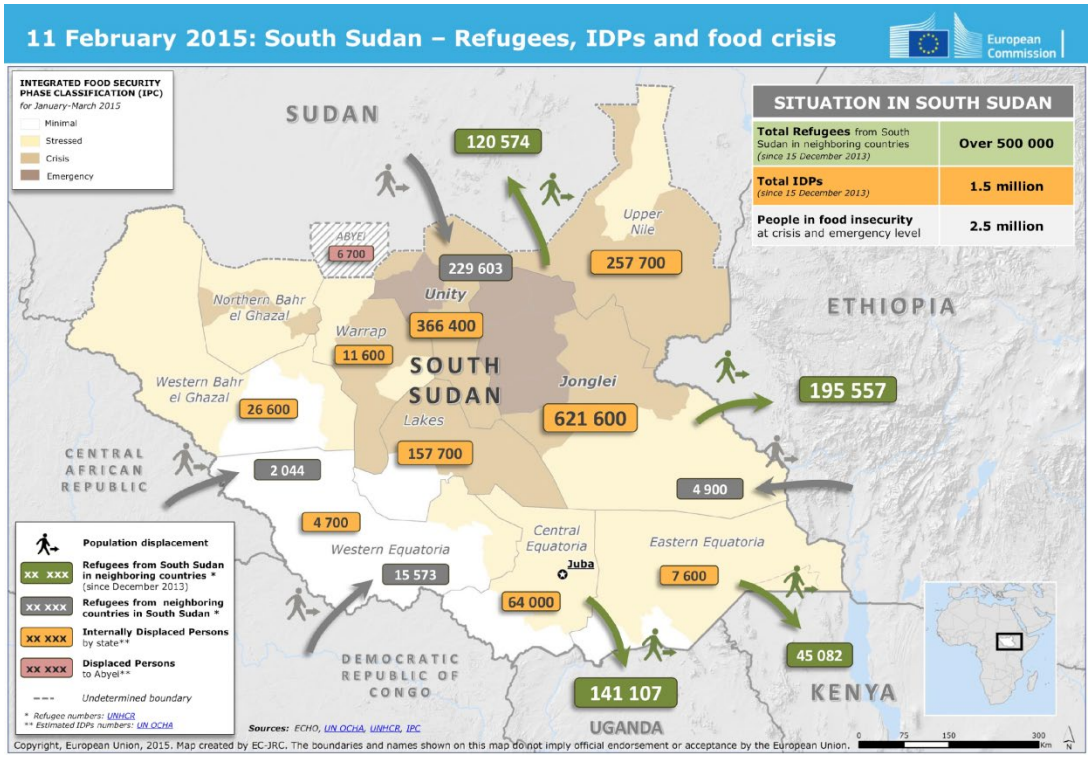


Figure 1: Challenge: overview of a historical emergency situation in Sudan. Reports such as these and those from the UNHCR can be used to validate our modelling approach against.

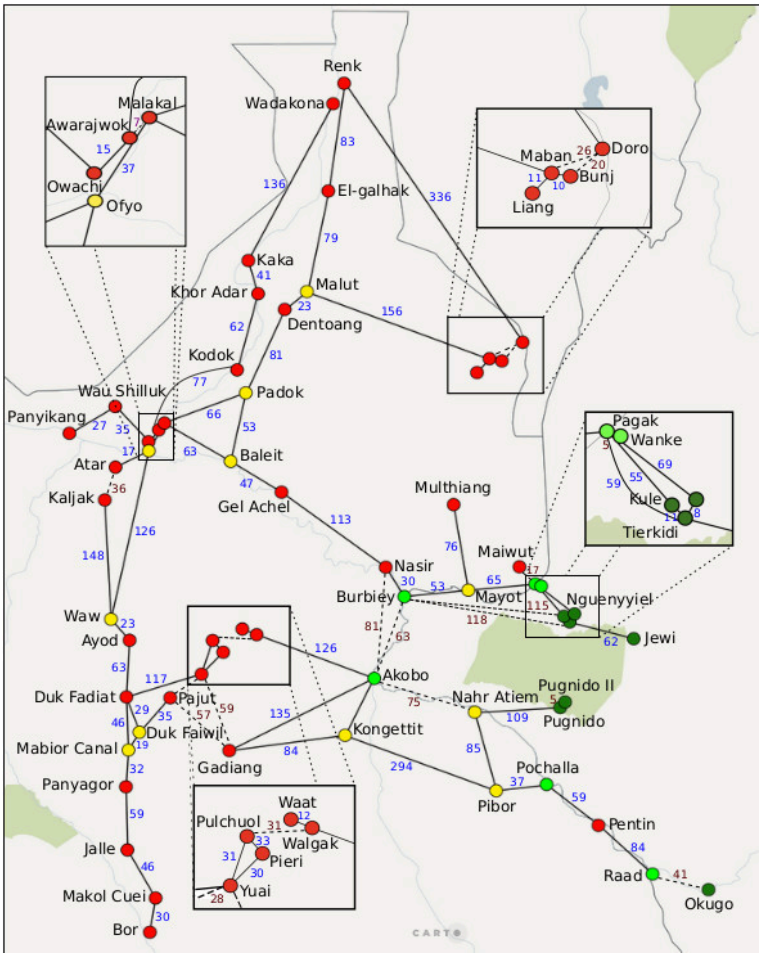


Figure 2: Solution: South Sudan Microscale location graph for the conflict occurred in 2016-2017 representing in red: conflict locations, in yellow: towns, and in green: camps.

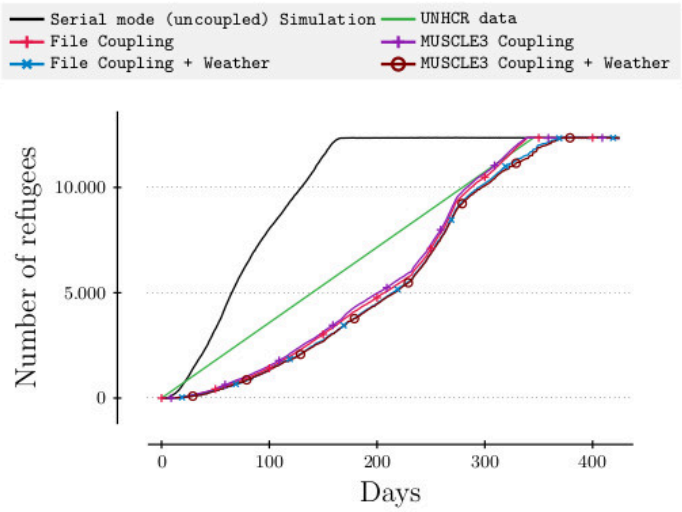


Figure 3: Evolution of refugees' number in camps during the simulation period depending on several coupling approaches (File I/O, MUSCLE3) with or without weather data coupling, comparing to UNHCR data and single-mode simulation.

Assisting decision makers to solve Global Challenges with HPC applications – Covid-19 modelling

HIGHLIGHTED CENTRE OF EXCELLENCE

HiDALGO is the Center of Excellence in HPC and Big Data technologies for Global Systems funded by the European Commission. Understanding global challenges to assist decision making by addressing multi-dimensional problems is a real need nowadays. HiDALGO develops novel methods, algorithms and software for HPC and HPDA enabling highly accurate simulations, data analytics and data visualisation to accurately model and simulate these complex processes.

Organisations & Codes Involved:

The National Health Service (NHS) oversees offering public health services in United Kingdom. It is now dealing with the COVID-19 pandemic.

Brunel University London is a dynamic institution that plays a significant role in the higher education sector. It carries out applied research on different topics, such as software engineering, intelligent data analysis, human computer interaction, information systems, and systems biology.

CHALLENGE:

The current pandemic situation has increased the NHS need of supporting tools to detect, predict and even prevent the virus spread behaviour. Knowing in advance this information will support them to take the appropriate decisions while considering health and care capabilities. In addition, the advance warning of new pandemic waves (or when they may subside) can help health authorities to rescale the capacity for non-urgent care, and ensure the timely arrangement of surge intensive-care capacity.

SOLUTION:

To tackle these challenges HiDALGO developed a tool: FACS, the Flu and Coronavirus Simulator, which is an agent-based model that also incorporates SEIRDI (Susceptible-Exposed-Infectious-Recovered-Dead-Immunized) states for all agents.

FACS approximates viral spread on the individual building level, and incorporates geospatial data sources from OpenStreetMap. In this way COVID-19 spread is modelled at local level, providing estimations of the spread of infections and hospital arrivals, given a range of public health interventions. Lastly, FACS supports the modelling of vaccination policies, as well as the introduction of new viral strains and changes in vaccine efficacy.

SOCIETAL & ECONOMIC IMPACT:

Although it is not a market topic itself, global challenges have been gaining importance in the last years, especially those related to climate change but also others such as peace and conflict. HiDALGO offers a set of tools, services, datasets and resources to define models that may predict situations under certain scenarios that can influence any decision to be taken.

More specifically, this tool FACS provides the needed information to decision makers so they can set up the appropriate measure and provide the necessary means at any time.

Indeed the tool helps the NHS to identify peaks of contagion in order to avoid sanitary collapses. Taking the appropriate decisions at the right moment represents a better investment of public resources and, what is more important, saving lives. Moreover, it supports to make better decisions and at appropriate time, to limit the problematic economic consequences of lockdowns and of the other measures taken in pandemic times.

BENEFITS FOR FURTHER RESEARCH:

- Support for the preparatory efforts by the health service for the second and third waves of the pandemic in West London.
- Better understanding of the nature of the current situation and the effect of different measures, such as lockdowns and vaccine efficacy levels.
- Provide models and elements about the foreseen evolution to limit the problematic economic consequences of the lockdowns and the various limitations due to the pandemic.

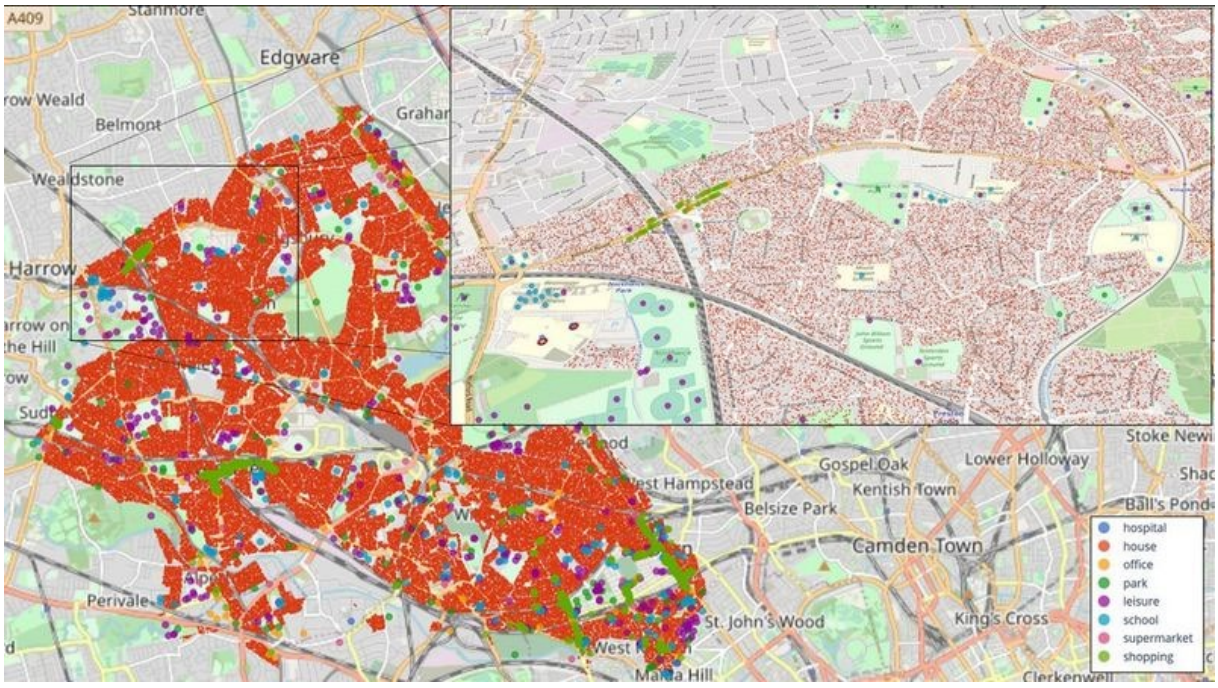


Figure 1: Scatter plot of the spread of COVID-19 cases across the London borough of Brent in UK. In this image we show all the buildings that are incorporated into a FACS model for the Borough of Brent. Each house location contains one or more households, each containing one or more persons.

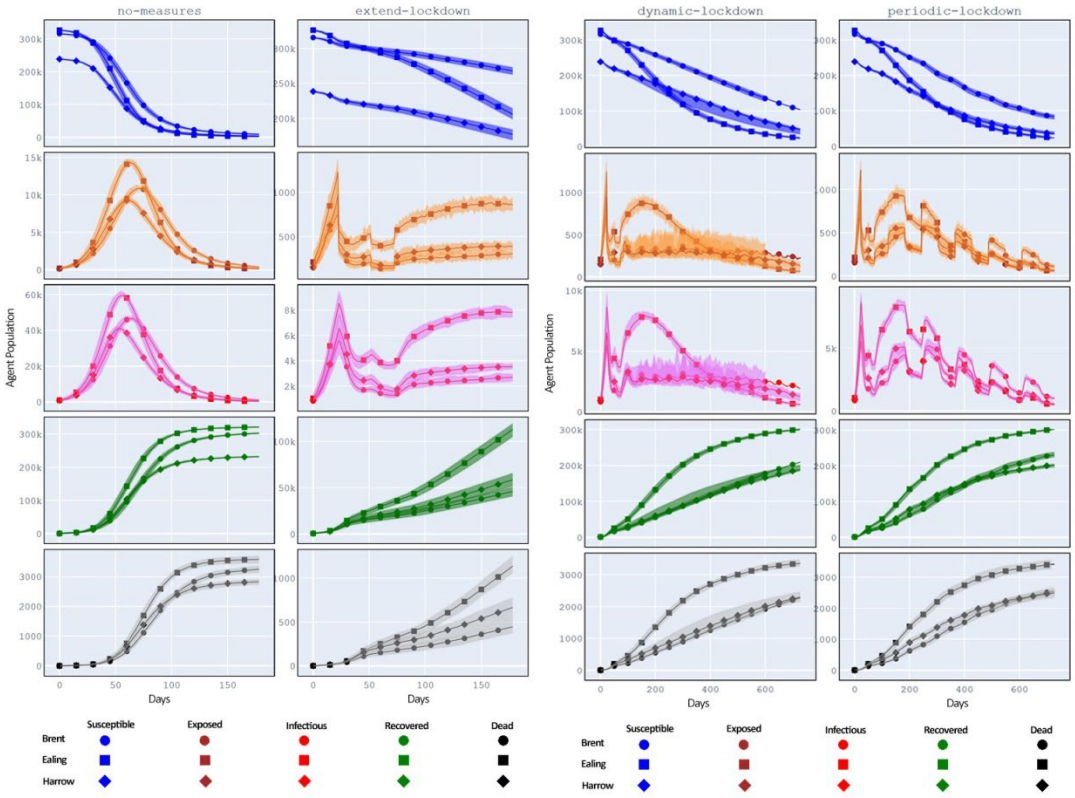
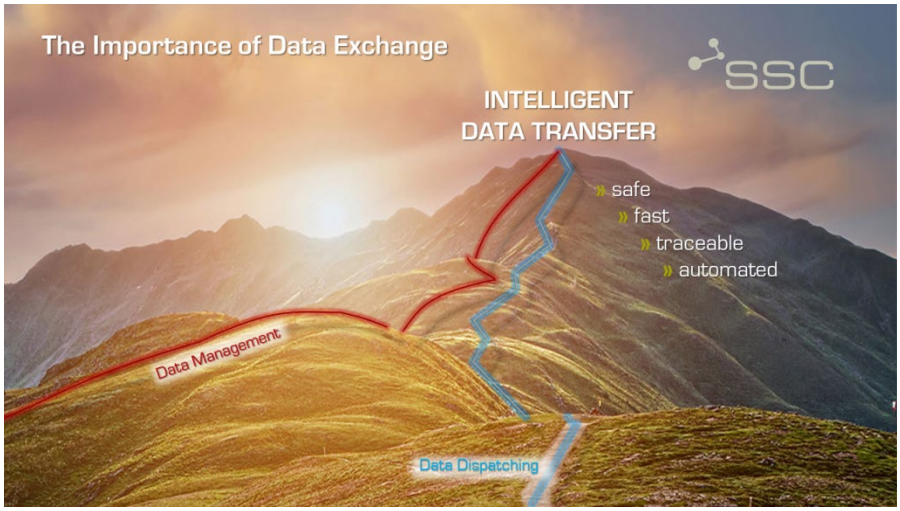


Figure 2: Simulation Plots: Susceptible, Exposed, Infected, Recovered, Dead for the London borough of Brent, Ealing, and Harrow under different lockdown measures. (Blue=Susceptible, Orange=Exposed, Rose=Infected, Green=Recovered, Black=Dead. The square, circle and diamond symbols represent Brent, Ealing and Harrow boroughs.)

Enabling High Performance Computing for Industry through a Data Exchange & Workflow Portal

CoE involved:



Success story # Highlights:

- Keywords:
  - Data Transfer
  - Data Management
  - Data Reduction
  - Automatisation, Simplification
  - Dynamic Load Balancing
  - Dynamic Mode Decomposition
  - Data Analytics
  - combustor design
- Industry sector: Aeronautics
- Key codes used: Alya

Organisations & Codes Involved:

As an IT service provider, SSC-Services GmbH develops individual concepts for cooperation between companies and customer-oriented solutions for all aspects of digital transformation. Since 1998, the company, based in Böblingen (Germany), has been offering solutions for the technical connection and cooperation of large companies and their suppliers or development partners. The corporate roots lie in the management and exchange of data of all types and sizes.

RWTH Aachen University is the largest technical university of Germany. The Institute of Aerodynamics of RWTH Aachen University possesses extensive expertise in the simulation of turbulent flows, aeroacoustics and high-performance computing. For more than 18 years large-eddy simulations with an advanced analysis of the large scale simulation data are successfully performed for various applications.

Barcelona Supercomputing Center (BSC) is the national supercomputing centre in Spain. BSC specialises in High Performance Computing (HPC) and manages MareNostrum IV, one of the most powerful supercomputers in Europe. BSC is at the service of the international scientific community and of industry that requires HPC resources. The Computing Applications for Science and Engineering (CASE) Department from BSC is involved in this application providing the application case and the simulation code for this demonstrator.

The code used for this application is the high performance computational mechanics code Alya from BSC designed to solve complex coupled multi-physics / multi-scale / multi-domain problems from the engineering realm. Alya was specially designed for massively parallel supercomputers, and the parallelization embraces four levels of the computer hierarchy. 1) A substructuring technique with MPI as the message passing library is used for distributed memory supercomputers. 2) At the node level, both loop and task parallelisms are considered using OpenMP as an alternative to MPI. Dynamic load balance techniques have been introduced as well to better exploit computational resources at the node level. 3) At the CPU level, some kernels are also designed to enable vectorization. 4) Finally, accelerators like GPU are also exploited through OpenACC pragmas or with CUDA to further enhance the performance of the code on heterogeneous computers. Alya is one of the only two CFD codes of the Unified European Applications Benchmark Suite (UEBAS) as well as the Accelerator benchmark suite of PRACE.

Technical Challenge:

SSC is developing a secure data exchange and transfer platform as part of the EXCELLERAT project to facilitate the use of high-performance computing (HPC) for industry and to make data transfer more efficient. Today, organisations and smaller industry partners face various problems in dealing with HPC calculations, HPC in general, or even access to HPC resources. In many cases, the calculations are complex and the potential users do not have the necessary expertise to fully exploit HPC technologies without support. The developed data platform will be able to simplify or, at best, eliminate these obstacles.

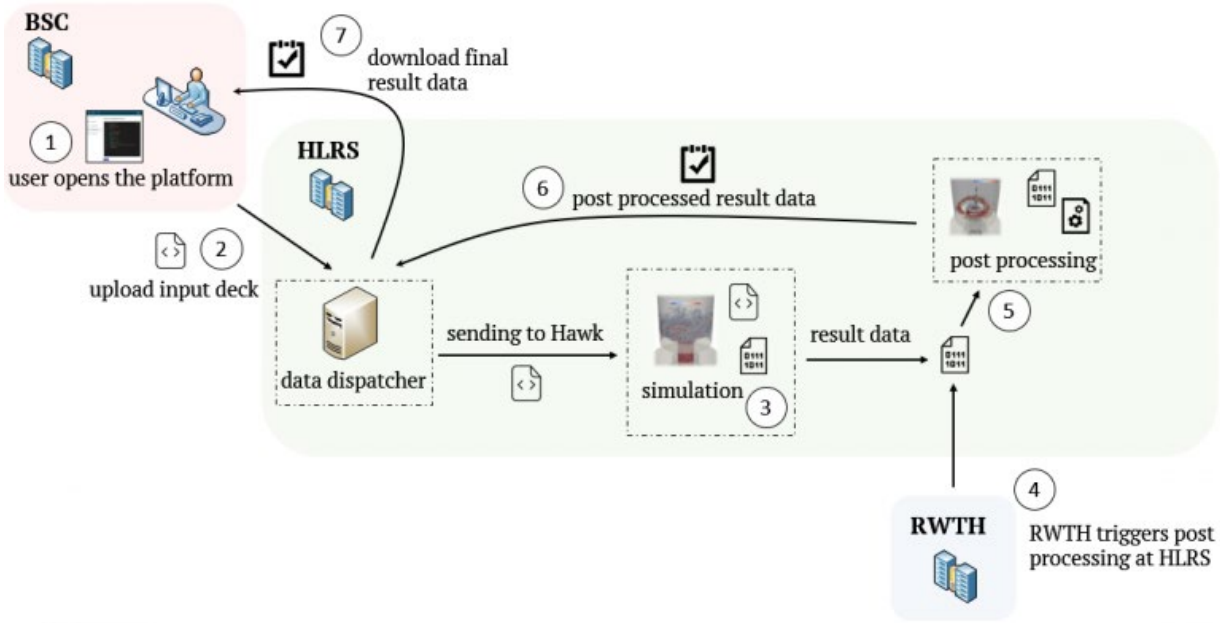


Figure 1: Data Roundtrip with an EXCELLERAT use case

The data roundtrip starts with a user at the Barcelona Supercomputing Center that wants to simulate the Alyafiles. Therefore, the user uploads the corresponding files through the data exchange and workflow platform and selects Hawk at HLRS as a simulation target. After the files have been simulated, RWTH Aachen starts the post-processing process at HLRS. In the end the user downloads the post processed data through the platform.

With the increasing availability of computational resources, high-resolution numerical simulations have become an indispensable tool in fundamental academic research as well as engineering product design. A key aspect of the engineering workflow is the reliable and efficient analysis of the rapidly growing high-fidelity flow field data. RWTH develops a modal decomposition toolkit to extract the energetically and dynamically important features, or modes, from the high-dimensional simulation data generated by the code Alya. These modes enable a physical interpretation of the flow in terms of spatio-temporal coherent structures, which are responsible for the bulk of energy and momentum transfer in the flow. Modal decomposition techniques can be used not only for diagnostic purposes, i.e. to extract dominant coherent structures, but also to create a reduced dynamic model with only a small number of degrees of freedom that approximates and anticipates the flow field. The modal decomposition will be executed on the same architecture as the main simulation. Besides providing better physical insights, this will reduce the amount of data that needs to be transferred back to the user.

Scientific Challenge:

Highly accurate, turbulence scale resolving simulations, i.e. large eddy simulations and direct numerical simulations, have become indispensable for scientific and industrial applications. Due to the multi-scale character of the flow field with locally mixed periodic and stochastic flow features, the identification of coherent flow phenomena leading to an excitation of, e.g., structural modes is not straightforward. A sophisticated approach to detect dynamic phenomena in the simulation data is a reduced-order analysis based on dynamic mode decomposition (DMD) or proper orthogonal decomposition (POD).

In this collaborative framework, DMD is used to study unsteady effects and flow dynamics of a swirl-stabilised combustor from high-fidelity large-eddy simulations. The burner consists of plenum, fuel

injection, mixing tube and combustion chamber. Air is distributed from the plenum into a radial swirler and an axial jet through a hollow cone. Fuel is injected into a plenum inside the burner through two ports that contain 16 injection holes of 1.6 mm diameter located on the annular ring between the cone and the swirler inlets. The fuel injected from the small holes at high velocity is mixed with the swirled air and the axial jet along a mixing tube of 60 mm length with a diameter of D = 34 mm. At the end of the mixing tube, the mixture expands over a step change with a diameter ratio of 3.1 into a cylindrical combustion chamber. The burner operates at Reynolds number Re = 75,000 with pre-heated air at T_air = 453 K and hydrogen coming at T_H2 = 320 K. The numerical simulations have been conducted on a hybrid unstructured mesh including prisms, tetrahedrons and pyramids, and locally refined in the regions of interest with a total of 63 million cells.

SOLUTION:

The developed Data Exchange & Workflow Portal will be able to simplify or even eliminate these obstacles. First activities have already started. The new platform enables users to easily access the two HLRS clusters, Hawk and Vulcan, from any authorised device and to run their simulations remotely. The portal provides relevant HPC processes for the end users, such as uploading input decks, scheduling workflows, or running HPC jobs.

In order to be able to perform data analytics, i.e. modal decomposition, of the large amounts of data that arise from Exascale simulations, a modal decomposition toolkit has been developed. An efficient and scalable parallelisation concept based on MPI and LAPACK/ScaLAPACK has been used to perform modal decompositions in parallel on large data volumes. Since DMD and POD are data-driven decomposition techniques, the time resolved data has to be read for the whole time interval to be analysed. To handle the large amount of input and output, the software tool has been optimised to effectively read and write the time resolved snapshot data parallelly in time and space. Since different solution data formats are utilised by the computational fluid dynamics community, a flexible modular interface has been developed to easily add data formats of other simulation codes.

The flow field of the investigated combustor exhibits a self-excited flow oscillation known as a precessing vortex core (PVC) at a dimensionless Strouhal Number of Sr=0.55, which can lead to inefficient fuel consumption, high level of noise and eventually combustion hardware damage. To analyse the dynamics of the PVC, DMD is used to extract the large-scale coherent motion from the turbulent flow field characterised by a manifold of different spatial and temporal scales shown in Figure 2 (left). The instantaneous flow field of the turbulent flame is visualised by an iso-surface of the Q-criterion coloured by the absolute velocity. The DMD analysis is performed on the three-dimensional velocity and pressure field using 2000 snapshots. The resulting spectrum of the DMD, showing the amplitude of each mode as a function of the dimensionless frequency is given in Figure 2 (top). One dominant mode, marked by a red dot, at Sr=0.55 matching the dimensionless frequency of the PVC is clearly visible. The temporal reconstruction of the extracted DMD mode showing the extracted PVC vortex is illustrated in Figure 2 (right). It shows the iso-surface of the Q-criterion coloured by the radial velocity.

Scientific impact of this result:

The Data Exchange & Workflow Portal is a mature solution for providing seamless and secure access to high-performance computing resources by end users. The innovative thing about the solution is that it combines the know-how about secure data exchange with an HPC platform. This is fundamental because the combination of know-how provision and secure data exchange between HPC and SMEs is unique. Especially the web interface is very easy to use and many tasks are automated, which leads to a simplification of the whole HPC complex and there is an easier entry for HPC engineers.

In addition, the data reduction programming technology ensures a more intelligent, faster transfer of files. There will be a highly increased performance speed when transferring the same data sets over and over. If the file is already known by the system and there is no need to transfer it again. Only the changed parts need to be exchanged.

The developed data analytics, i.e. modal decomposition, toolkit provides an efficient and user-friendly way to analyse simulation data and extract energetically and dynamically important features, or modes, from complex, high-dimensional flows. To address a broad user community having different backgrounds and expertise in HPC applications, a server/client structure has been implemented allowing an efficient workflow. Using this structure, the actual modal decomposition is done on the server running in parallel on the HPC cluster which is connected via TCP with the graphical user interface (client) running on the local machine. To efficiently visualise the extracted modes and reconstructed flow fields without writing large amounts of data to disk, the modal decomposition server can be connected to a ParaView server/client configuration via Catalyst enabling in-situ visualisation.

Finally, this demonstrator shows an integrated HPC-based solution that can be used for burner design and optimisation using high-fidelity simulations and data analytics through an interactive workflow portal with an efficient data exchange and data transfer strategy.

Benefits for further research:

- Higher HPC customer retention due to less complex HPC environment
- Reduction of HPC complexity due to web frontend
- Shorter training phases for inexperienced users and reduced support effort for HPC centres
- Calculations can be started from anywhere with a secure connection
- Time and cost savings due to a high degree of automation that streamlines the process chain
- Efficient and user-friendly post-processing/ data analytics

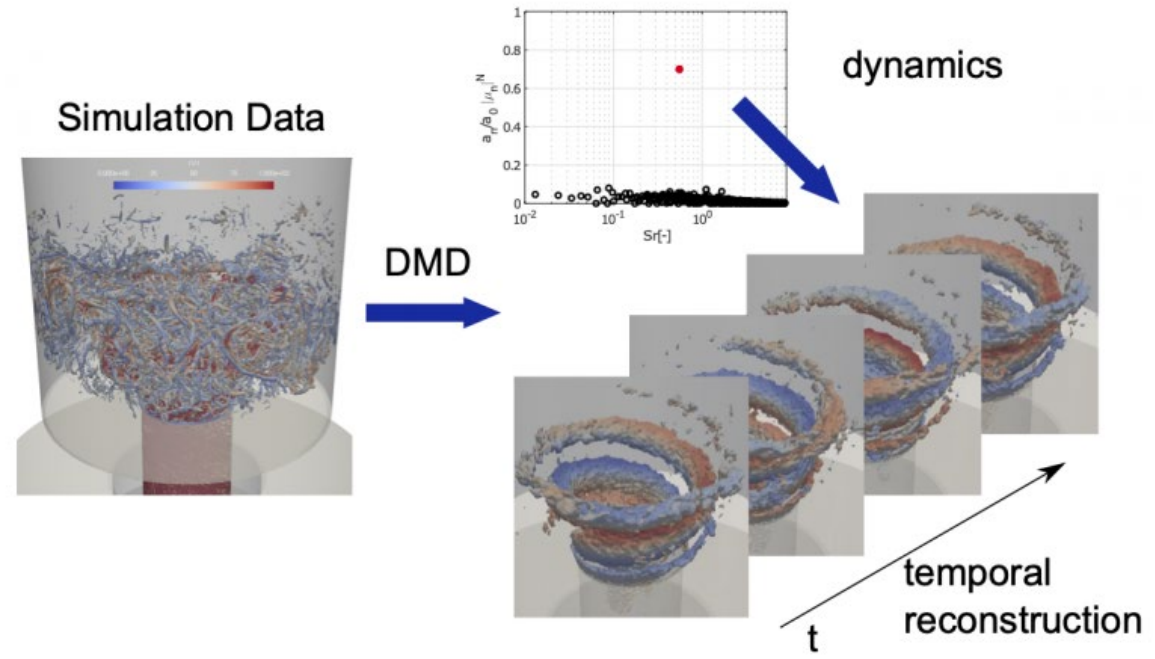
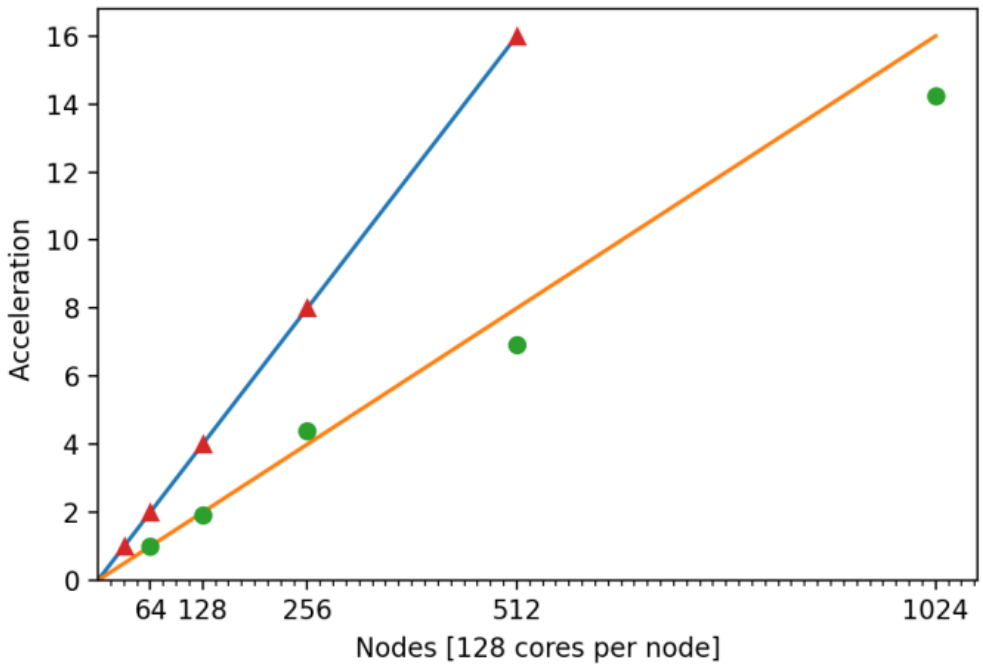


Figure 2: DMD analysis performed on the flow field of a turbulent flame. Instantaneous flow field (left), Spectrum of the DMD (top), Reconstruction of the dominant DMD-mode (right).

Bringing industrial end-users to Exascale computing: An industrial level combustion design tool on 128K cores

CoE involved:



Strong scaling for turbulent channel (tri) and rocket engine simulations (circle). Performance (symbols) versus ideal (line) acceleration.

Success story # Highlights:

- Exascale
- Industry sector: Aerospace
- Key codes used: AVBP

Organisations & Codes Involved:

CERFACS is a theoretical and applied research centre, specialised in modelling and numerical simulation. Through its facilities and expertise in High Performance Computing (HPC), CERFACS deals with major scientific and technical research problems of public and industrial interest.

GENCI (Grand Equipement National du Calcul Intensif) is the French High-Performance Computing (HPC) National Agency in charge of promoting and acquiring HPC, Artificial Intelligence (AI) capabilities, and massive data storage resources.

Cellule de Veille technologique GENCI (CVT) is an interest group focused on technology watch in High Performance Computing (HPC) pulling together French public research, CEA, INRIA among Others. It offers first time access to novel architectures and access to technical support towards preparing the codes for the near future of HPC.

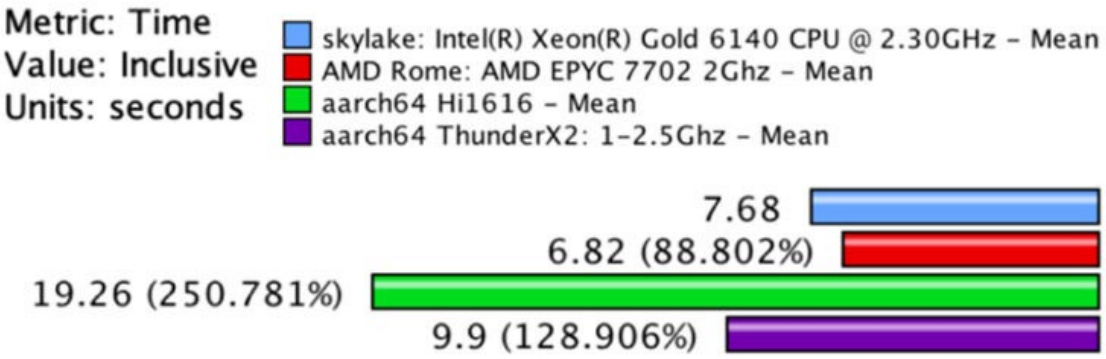
AVBP

AVBP is a parallel Computational Fluid Dynamics (CFD) code that solves the three-dimensional compressible Navier-Stokes on unstructured and hybrid grids. Its main area of application is the modelling of unsteady reacting flows in combustor configurations and turbomachines. The prediction of unsteady turbulent flows is based on the Large Eddy Simulation (LES) approach that has emerged as a prospective technique for problems associated with time dependent phenomena and coherent eddy structures.

CHALLENGE:

Some physical processes like soot formation are so CPU intensive and non deterministic that their predictive modelling is out of reach today, limiting our insights to ad hoc correlations, and preliminary assumptions. Moving these runs to Exascale level will allow simulation longer by orders of magnitudes, achieving the compulsory statistical convergence required for a design tool.

The complexity at the code level to unlock node level and system level performance is challenging and requires code porting, optimisation and algorithm refactoring on various architectures in the way to enable Exascale performance.



Single core performance for a Karman street simulation measures via gprof

SOLUTION:

In order to prepare the AVBP code to architectures that were not available at the start of the EXCELLERAT project, CERFACS teamed up with the CVT from GENCI, Arm, AMD and the EPI project to port, benchmark and (when possible) optimise the AVBP code for Arm and AMD architectures. This collaboration ensures an early access to these new architectures and prime information to prepare our codes for the wide spread availability of systems equipped with these processors. The AVBP developers got access to the IRENE AMD system of PRACE at TGCC with support from AMD and Atos, which allowed to characterise the application on this architecture and create a predictive model on how the different features of the processor (frequency, bandwidth) could affect the performance of the code. They were also able to port the code to flavors of Arm processors singling out compiler dependency and identify performance bottlenecks to be investigated before large systems become available in Europe.

Business impact:

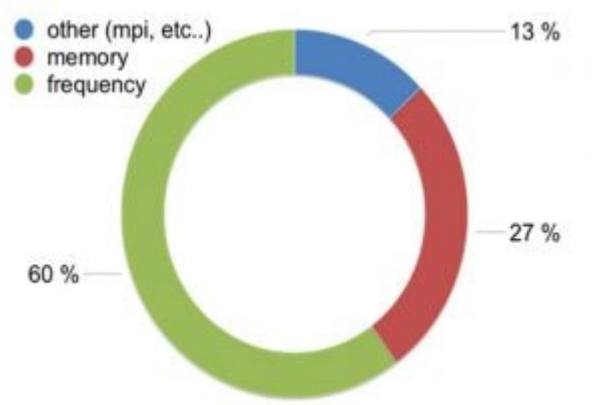
The AVBP code was ported and optimised for the TGCC/GENCI/PRACE system Joliot-CURIE ROME with excellent strong and weak scaling performance up to 128,000 cores and 32,000 cores respectively. These optimisations impacted directly four PRACE projects on this same system on the following call for proposals.

Beside AMD processors, the EPI project and GENCI's CVT as well as EPCC (EXCELLERAT's partner) provided access to Arm based clusters respectively based on Marvell ThunderX2 (Inti Prototype hosted and operated by CEA) and Hi1616 (early silicon from Huawei) architectures. This access provided important feedback on code portability using the Arm and gcc compilers, single processor and strong scaling performance up to 3072 cores.

Results from this collaboration have been included on the Arm community blog [1]. A white paper on this collaboration is underway with GENCI and CORIA CNRS.

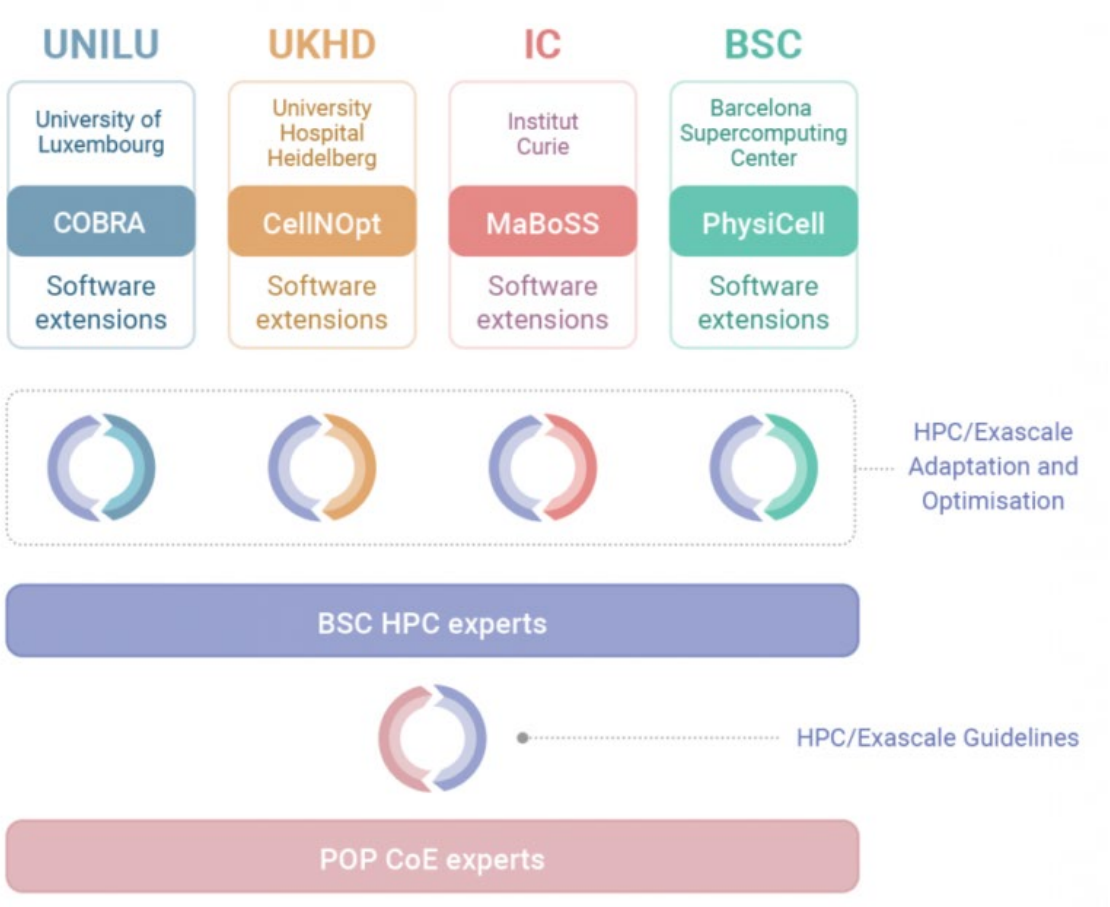
Benefits for further research:

- Code ready for the wide spread access of the Rome Architecture.
- Strong and weak scaling measurements up to 128,000 cores
- Initial optimisations for Arm architectures



Code characterisation on AMD Epyc 2 for AVBP

POP and PerMed Centers of Excellence are Getting Cell-Level Simulations Ready for Exascale



Collaboration of CoEs workflow

Success story # Highlights:

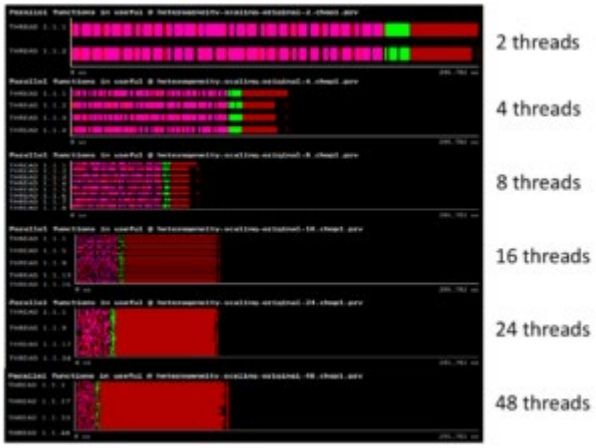
- Industry sector: Computational biology
- Key codes used: PhysiCell
- Keywords:
    - Memory management
    - Cell-cell interaction simulation
    - OpenMP
    - Good practice

Organisations & Codes Involved:

Technical Challenge:

One of PerMedCoE's use cases is the study of tumour evolution based on single-cell omics and imaging using PhysiCell. In order to simulate such large-scale problems that replicate real-world tumours, High Performance Computing (HPC) is essential. However, memory usage presents a challenge in HPC architectures and is one of the obstacles to optimizing simulations for running at Exascale.

With a high number of threads computing in parallel, performance can be degraded because of concurrent memory allocations and deallocations. We observe that as the number of threads goes up, runtime is not reduced as expected. This is not a memory-bound problem, but a memory management one.



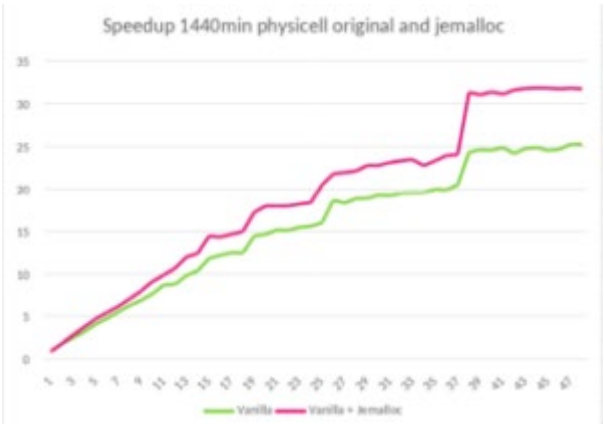
Solution:

Following the POP methodology and using the BSC tools (Extrae, Paraver, and BasicAnalysis), we determine that an overloading of a specific implementation of C++ operators is causing a high number of concurrent memory allocations and deallocations, causing the memory management library to perform synchronization and "steal" cycles of CPU from the running application.

The proposal from POP is to avoid the overloading of operators that need allocation and deallocation of data structures. However, achieving this implies a major code change. To demonstrate the potential of the suggestion without doing the major code change, we suggest using an external library to improve the memory management: Jemalloc.

Jemalloc is "a general purpose malloc implementation that emphasizes fragmentation avoidance and scalable concurrency support" [1]. It can be integrated easily into any code by preloading the library, with LD_PRELOAD in Linux for example. After applying this solution to PhysiCell, we execute the same experiment and obtain a 1.45x speed up when using 48 OpenMP threads.

[1] cited from Jemalloc website. http://jemalloc.net



Business impact:

As the only transversal CoE, one of POP's objectives is to advise and help the other HPC CoEs prepare their codes for the Exascale. This collaboration between POP and PerMed is a good example of the potential of these kinds of partnerships between CoEs.

POP provides the performance analysis expertise and tools, the experience of hundreds of codes analysed, and the best practices gathered from those analyses. PerMedCoE bring state of the art use

cases and real-world problems to be solved. Together, they improve the performance and efficiency of cell-level agent-based simulation software: in this case PhysiCell.

One of the goals of PerMedCoE is the scaling-up of four different tools that address different types of simulations in personalized medicine and that were coded in different languages (C++, R, python, julia). These tools are being re-factored to scale up to Exascale. In this scaling-up, audits such as the ones POP can offer are essential to evaluate past developments and guide future ones.

Currently, only one of PerMed's tools is able to use several nodes to run a single simulation, whereas the rest of the tools can only use all of the processors of a single node. Our scaling-up strategy is a heterogeneous one. We are implementing MPI on some tools, re-factoring others to other languages that can ease the use of HPC clusters such as Julia, or even targeting "many-task computing" paradigms like in the case of model fitting.

The main motivation to have HPC versions of these simulation tools is to be able to simulate bigger, more complex cell-level agent-based models. Current models can obtain up to 10^6 cells, but it has been proven that most of the problems addressed in computational biology (cancerogenesis, cell lines growth, COVID-19 infection) usually target from 10^9 to 10^12 cells. In addition, most of these current simulations consider an over-simplified environment that is nowhere close to a real-life scenario. We aim to have complex multi-scale simulation frameworks that target these bigger, more complex simulations.

Benefits:

- Speedup of 1.45x on runtime with 48 threads in PhysiCell execution
- A good practice for High Performance C++ applications exported
- This behaviour in memory management systems will be detected more easily in future analysis

7.2 Annex: Guidelines for interaction with industry (extract from the generic common part of the "Industry Questionnaire" report sent to the CoEs)

This Appendix -included in the generic common part of the personalized reports sent to the CoEs- is intended to provide guidelines to the CoEs Exploitation teams on the best practices in the interaction with industry. Each CoE is different, with different degrees of TRL in their services and different amount of efforts devoted to attract collaborations with industry. To provide an idea of the content shared with all CoEs, and the NCCs, the guidelines provided in the common section were:

"

The first step is (i) to define a clear list of services offered by the CoE and (ii) to define the target sectors that could be interested in these services, naming specific companies that could be potential clients. For example, a table like this one

Service	**Target sector**	**Target companies**
Service A	Pharmaceuticals	Company 1, Company 2, …
Service B	Biotechnology	Company 3, Company 4, …
Service B	Chemicals	Company 5, Company 6, …
…	…	…

Then there are two marketing strategies for approaching potential customers:

- **Outbound marketing** consists in proactive activities to reach out potential customers, such as direct email, cold calling, trade fairs, and seminar series.
- **Inbound marketing** consists in "being found" by people which are already looking for solutions in your target industry.

Best practices for Outbound marketing:

1. Identify persons to contact in the target companies. Preferably technical persons (R&D engineers, application scientists, technologists, R&D managers, etc.) or Open Innovation managers.
2. Contact them per email or via Linkedin. Start by the people provided in Annex I.
3. Attend sectorial events such as trade fairs and industry workshops and show there the CoE capabilities.
4. Prepare a 10-slides presentation about the CoE and its services. Use this presentation in the first meeting with the contacts that have shown interest in learning more about the CoE.
5. This is B2B (business to business) relationship, in contrast to the B2C (business to consumer) marketing we are used to as consumers. In B2B, **trust** is essential to close an agreement. This is a "people business" and the customer will spend time and money in a collaboration with the CoE only if the customer trusts the people in the CoE. To generate trust requires listening to the problems and challenges of the customer and help them to find a solution. It also requires a professional interaction, e.g. quick answer to the customer requests, formal language, etc.

Best practices for Inbound marketing:

1. In order to do inbound marketing, it is needed to set up the CoE's website like a "hub" for industry, that attracts visitors naturally through search engines, blogging, and social media.
2. Include an independent section for Services in the CoE's website.
3. Add in this section a call-to-action button for requesting a demo or further information.
4. Post in your website blog and social networks at least one article per month related to industrial applications.
5. Identify the sectorial media most relevant in the target sectors (magazines, newsletters, industry association communications, etc.) and publish one article per semester there.

We recommend recording all the interactions with industry in a CRM (Customer Relationship Management) system such as Hubspot or Pipedrive. It is useful to classify the contacts depending on the "sale" stage they are at:

- Prospect: initial contact via email, Linkedin or CoE's form.
- Lead: the contact has shown interest in having a meeting or demo. In the first meeting the goal is to collect as much information as possible about the problem of the customer.
- Opportunity: there are further meetings to present the CoE's capabilities, propose a solution for the customer's problem, address questions and doubts of the customer.
- Customer: agreement closed.