H2020-INFRAEDI-2018-2020

Concerted action for the European HPC CoEs

Project Number: 823964

D2.6 – Final Summary of Consolidated CoE Input Provided to the EU HPC Ecosystem

WP2 – The HPC CoE General Assembly

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

Workpackage:	2	The HPC CoE General Assembly
Due Date	31.03.2022	
Author(s):	Lilit Axner	KTH
Authorized by	Edouard Audit	CEA
Reviewer	H. S. Martin	UCL
Reviewer	H-Ch. Hoppe	Scapos AB
Reviewer	B. Koller	HLRS
Final Version Date	31.03.2022	
Dissemination Level	Public	

List of abbreviations

BioExcel	Centre of Excellence for Computational Biomolecular
CASTIEL	Coordination and Support for National Competence Centres project
CEF	Connecting Europe Facility
ChEESE	Centre of Excellence for Exascale in Solid Earth
CoE	Centres of Excellence
CoEC	Center of Excellence in Combustion
CompBioMed	Centre of Excellence in Computational Biomedicine
D	Deliverable
DoW	Description of Work
EC	European Commission
E-CAM	An e-Infrastructure for Software, Training and Consultancy in Simulation and Modelling
EoCoE-II	Energy oriented Centre of Excellence
EOSC	European Open Science Cloud
ESiWACE	For future exascale climate and weather predictions
ETP4HPC	European Technology Platform for High-Performance Computing
EU	European Union
EU13	[Refer to new EU member states since 2004]
EuroCC	European National Competence Centres project
EuroHPC JU	European High Performance Computing Joint Undertaking
EXCELLERAT	The European Centre of Excellence for Engineering Applications
FZJ	Forschungszentrum Jülich (Juelich Research Centre)
H2020	Horizon 2020 –The EC Research and Innovation Programme in Europe
HLRS	Höchstleistungsrechenzentrum Stuttgart (High Performance Computing Centre Stuttgart)
HPC	High Performance Computing
HPC cPPP	HPC contractual Public Private Partnership in HPC
KTH	Kungliga Tekniska högskolan (Royal Institute of Technology, Stockholm)
Max	Material design at eXascale
MFF	Multiannual Financial Frameworks
NCC	National Competence Centre
NOMAD-2	Novel Materials Discovery Centre of Excellence
PerMedCoE	Centre of Excellence for Personalised Medicine in Europe
PRACE	Partnership for Advanced Computing in Europe
POP2	Performance Optimisation and Productivity Centre of Excellence Computing Applications
HiDALGO	HPC and Big Data Technologies for Global Challenges
SRA	Strategic Research Agenda
TREX	Targeting Real chemical accuracy at the EXascale Centre
UCL	University College London
WP	Work Package

Executive Summary

The European High Performance Computing (HPC) Centres of Excellence (CoEs) are one of the three pillars of the European HPC Ecosystem. They ensure that extreme scale applications will result in tangible benefits for addressing scientific, industrial, or societal challenges. The FocusCoE project created a productive platform for the CoEs to coordinate strategic directions, establish collaboration and provided support services for industrial outreach and promotion of CoE results, competences, and services [1].

This deliverable is the final report of FocusCoE Work Package 2, Task 2.3 (referred to as T2.3 below) which did support the CoEs in their interactions with the rest of the European HPC ecosystem. Specific objectives included establishing contacts, consolidating input provided by the CoEs to the the other ecosystem players, and facilitating a steady flow of information in both directions.

T2.3 created an effective forum based on HPC3 platform to coordinate strategic directions involving the CoEs in conjunction with key European HPC stakeholders in the HPC ecosystem, including the EuroHPC JU, PRACE and ETP4HPC, the EuroCC and CASTIEL projects and their national HPC competency centres (NCCs), and other coordination and support actions.

The CoEs are regularly contacted by European HPC ecosystem players (including the ones listed above) to provide input on matters of mutual interest. Task 2.3 coordinated and consolidated these requests, contacting all relevant CoE players, bundling responses and giving substantial, clear feedback to requestors. This materially increases the impact of the CoE community on the HPC ecosystem at large while minimising disturbance of the primary CoE activities. In addition, T2.3 worked with the CoEs to support the CoEs in preparing for comprehensive European stakeholder workshops on HPC training and skills development needs. Specific activities included

1. Answering an ETP4HPC/EXDCI-2 questionnaire for the contractual Public Private Partnership in HPC (HPC cPPP).
2. Providing use case and application input for the ETP4HPC Strategic Research Agendas (SRAs) 4 and 5
3. Contributing to legacy code studies of the EXDCI-2 project
4. Replying to a questionnaire for the Multiannual Financial Frameworks (MFF) and Connecting Europe Facility (CEF).
5. Submitting CoE contact points to the European Open Science Cloud (EOSC)
6. Establish discussions with EuroHPC JU pre-Exascale and Petascale hosting sites, formulating CoEs access requirements
7. Driving the communication to the EuroCC and CASTIEL projects and to their HPC NCCs, covering mutual competencies and needs, a series of thematic workshops, and specific collaborations
8. Shepherding sessions on the CoEs at the EuroHPC Summit Weeks 2021 and 2022.

T2.3 proactively worked on these other exchanges not triggered by specific requests, adapting, and communicating questions to the CoEs in a coordinated manner. The timely input received back was a clear indication of CoE interest in and appreciation of T2.3. These results were communicated back to the requestors and constituted important input to the HPC Ecosystem.

Table of Contents

Table of Figures

Table of Tables

1 Introduction

As was mentioned above the 15 European High Performance Computing (HPC) Centres of Excellence (CoEs) are one of the three pillars of the European HPC Ecosystem. They ensure that extreme scale applications will result in tangible benefits for addressing scientific, industrial, or societal challenges. The FocusCoE project created a productive platform for the CoEs to coordinate strategic directions, establish collaboration and provided support services for industrial outreach and promotion of CoE results, competences, and services [1].

The CoEs embody the leading-edge European excellence in HPC application development across many scientifically and industrially relevant areas, and as such are a main part of the European HPC application pillar.

Accordingly, the collective expertise of the CoEs is of high potential value to many HPC actors such as key stakeholders in the HPC ecosystem (including EC, EuroHPC JU, PRACE, EuroCC/CASTIEL and ETP4HPC) and the wider HPC research, user, and provider communities, including industry. CoEs are regularly contacted from HPC ecosystem players to give input or feedback for topics of mutual interest, from text for position papers to contributions to research agendas, statistics, and co-design activities.

In this context, FocusCoE T2.3 created an effective forum based on HPC3 platform to coordinate strategic directions involving the CoEs in conjunction with key European HPC stakeholders in the HPC ecosystem, including the EuroHPC JU, PRACE and ETP4HPC, the EuroCC and CASTIEL projects and their national HPC competency centres (NCCs), and the coordination and support actions. coordinates and consolidates requests for CoE inputs or feedback, effectively creating a networking hub between different HPC ecosystem actors and the CoEs. Specifically, T2.3 acts on request of external HPC ecosystem actors.

- Contacts all relevant CoEs and CoE partners
- Solicits, discusses and collates productive answers from these
- Returns a consolidated answer and facilitates any follow-up discussions

This document describes the activities of T2.3 between M19 and M40 in detail, and summarizes the actions taken in the first 18 months[1]. It is organised as follows:

- This first section is the introduction to the document.
- The second section presents the state of the European HPC ecosystem and the role which the CoEs play.
- The third section summarizes the actions taken by T2.3 in the first 18 months, and details T2.3 activities between M19 and M40.
- The final section summarises the T2.3 results of T2.3 and discusses lessons learned.

This deliverable repeats information about the FocusCoE project as a whole where required to explain the context of T2.3 activities.

[1] Deliverable D2.3 contains the full details on these.

2 HPC CoEs and the European HPC Ecosystem

Compared to the situation at M18 (June 2020), the set of HPC CoEs and the European HPC Ecosystem in general have seen significant changes. This section presents the updated set of CoEs supported by FocusCoE (and thus in *extenso* by T2.3), discusses the current structure of the European HPC ecosystem, and covers the role that the CoEs play to further the overarching goals of the EC and EuroHPC JU funding initiatives.

According to their charter, the CoEs bring together the European world-class knowledge and expertise in user-driven HPC application development, performance tools and programming models for HPC, and co-design activities for innovative systems based on leading edge technologies. They work to improve key HPC applications and libraries to meet the critical European scientific, industrial and societal needs, and to create a vibrant application ecosystem for future Exascale systems.

2.1 HPC CoEs Supported by FocusCoE

During the second period of the FocusCoE project five new CoEs were established or initiated a second phase of their activities, and all of these were supported by the FocusCoE project. T2.3 contacted the new CoEs, offered its services in relaying requests for feedback and information to them and consolidate their answers, and all of the new CoEs accepted this offer.

The superset of CoEs (currently active or terminated) [1] is shown in Table 1.

Acronym	Full Title	Runtime	Project web site
BioExcel-2	BioExcel-2 Centre of Excellence for Biomolecular Research	01/2019-06/2022	https://bioexcel.eu/
ChEESE	Centre of Excellence for Exascale in Solid Earth	11/2018-03/2022	https://cheese-coe.eu/
CompBioMed	A Centre of Excellence in Computational Biomedicine	10/2019-09/2023	https://www.compbiomed.eu/
EoCoE-II	Energy oriented Centre of Excellence for computer applications	01/2019-06/2022	https://www.eocoe.eu
EsiWACE2	Excellence in Simulation of Weather and Climate in Europe, Phase 2	01/2019-12/2022	https://www.esiwace.eu/
E-CAM	An e-infrastructure for software, training and consultancy in simulation and modelling	10/2015-03/2021	https://www.e-cam2020.eu/
EXCELLERAT	The European Centre of Excellence for Engineering Applications	12/2018-05/2022	https://www.excellerat.eu

Acronym	Full Title	Runtime	Project web site
HiDALGO	HPC and Big Data Technologies for Global Challenges	12/2018-02/2022	https://hidalgo-project.eu/
MaX	Materials design at the Exascale	12/2018-05/2022	http://www.max-centre.eu/
POP2	Performance Optimisation and Productivity 2	12/2018-05/2022	https://pop-coe.eu/
FocusCoE	Concerted action for the European HPC CoEs	12/2018-03/2022	https://www.hpccoe.eu/
TREX	Targeting Real Chemical accuracy at the Exascale	10/2020-09/2023	https://trex-coe.eu/
NOMAD-2	Novel materials for urgent energy, environmental and societal challenges	10/2020-03/2024	https://www.nomad-coe.eu/
CoEC	Advanced technology for combustion simulation	10/2020-09/2023	https://coec-project.eu/
PerMedCoE	High-performance computing to treat disease at individual level	10/2020-09/2023	https://permedcoe.eu/
RAISE	Research on AI and Simulation-Based Engineering at Exascale	01/2021-12/2023	https://www.coe-raise.eu/

Table 1: CoEs active during the FocusCoE term.

2.2 European HPC Ecosystem Overview

Due to newly launched R&D&I actions from the funding programmes of the European Commission, the EuroHPC Joint Undertaking and the individual EU member states the European HPC ecosystem has become richer. In our context, noteworthy additions in the 2nd half of 2020 include the national HPC competency centres (NCCs) (coordinated by the EuroCC project and supported by the CASTIEL CSA), the FF4EuroHPC project addressing broader HPC take-up by SMEs, and of course the five additional HPC CoEs CoEC, NOMAD-2, PerMedCoE, RAISE and TREX. In addition, the calls for hosting European Peta- and pre-Exascale systems were concluded.

2.2.1 EuroHPC Joint Undertaking

The European High Performance Computing Joint Undertaking (EuroHPC JU) is a legal and funding entity under Council Regulation EU 2018/1488 and located in Luxembourg. The EuroHPC JU allows the European Union and the EuroHPC JU participating countries to coordinate their efforts and pool their resources to make Europe a world leader in supercomputing. Starting from July 13, 2021, this regulation was revoked, and replaced by

Council Regulation (EU) 2021/1173. These new rules are substantially different including the changes related to additional provisions for EuroHPC JU to acquire and own quantum computers and quantum simulators, industrial-grade supercomputers, and mid-range supercomputers. Moreover, the now provisions allow the use of EuroHPC supercomputers for commercial purposes. Finally, upgrading previously acquired supercomputers is now foreseen [4].

More precisely, the EuroHPC JU aims to:

- develop, deploy, extend, and maintain a world-leading supercomputing and data infrastructure in Europe. The objective is to reach Exascale capabilities by 2022/2024. At the moment it deploys two (pre-) Exascale and five Petascale supercomputers.
- support the development and uptake of innovative and competitive supercomputing technologies and applications based on a supply chain that will reduce Europe's dependency on foreign computing technology.
- widen the use of HPC infrastructures to a large number of public and private users wherever they are located in Europe and support the development of key HPC skills, education and training for European science and industry.

The EuroHPC Joint Undertaking is composed of public and private members [4]:

Public members[2] are

- the European Union (represented by the Commission),
- Member States and Associated Countries that have chosen to become members of the Joint Undertaking: Austria, Belgium, Bulgaria, Croatia, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, the Netherlands, Norway, Poland, Portugal, Romania, Slovakia, Slovenia, Spain, Sweden, and Turkey.

Private members are

- representatives from the three participating private partners, the European Technology Platform for High Performance Computing (ETP4HPC), the Big Data Value Association (BDVA, new name is DAIRO) and the European Quantum Industry Consortium (QuIC).

2.2.2 Pre-Exascale Hosting Activities, HPC National Competency Centres, EuroCC and Castiel

In the 2nd half of 2020, the ecosystem was complemented by the EuroHPC National Competence Centres (supported by the CASTIEL CSA and coordinated by the EuroCC project) and the FF4EuroHPC project addressing the broader HPC take-up by SMEs. There are currently 33 EuroCC NCCs as shown in the figure below and listed in the table.

[2] According to In March 2022

Figure 1: EuroCC 33 NCCs spread across Europe

NCC Country	Website
Competence Centre Germany	https://eurocc-gcs.de/
Competence Centre Bulgaria	http://eurocc-bulgaria.bg/
Competence Centre Austria	https://eurocc-austria.at/
Competence Centre Croatia	https://www.hpc-cc.hr/
Competence Centre Cyprus	https://castorc.cyi.ac.cy/national-hpc-competence-centre
Competence Centre Czech Republic	https://www.it4i.cz/vitejte-v-narodnim-centru-kompetence-pro-hpc
Competence Centre Denmark	https://www.deic.dk/en/supercomputing/EuroCC
Competence Centre Estonia	https://eurocc-estonia.ee/
Competence Centre Finland	https://www.csc.fi/-/eurocc-competence-center-finland
Competence Centre Greece	https://eurocc-greece.gr/
Competence Centre Hungary	https://hpc.kifu.hu/
Competence Centre Ireland	https://www.eurocc-ireland.ie/
Competence Centre Italy	http://www.euroccitaly.it/
Competence Centre Lithuania	https://www.eurocc-lietuva.lt/
Competence Centre Latvia	https://eurocc-estonia.ee/wordpress_root/
Competence Centre Norway	https://www.sigma2.no/norwegian-euro-competence-centre
Competence Centre Poland	https://cc.eurohpc.pl/
Competence Centre Portugal	https://eurocc.fccn.pt/
Competence Centre Romania	https://roncc.ro/
Competence Centre Slovenia	https://www.sling.si/sling/eurocc/
Competence Centre Spain	https://eurocc-spain.res.es/
Competence Centre Sweden	https://enccs.se/
Competence Centre Switzerland	https://www.cscs.ch/about/collaborations/eurocc-ch/
Competence Centre Turkey	http://eurocc.truba.gov.tr/?page_id=2640&lang=en

NCC Country	Website
Competence Centre United Kingdom	https://www.linkedin.com/company/eurocc-uk-the-uk-national-competence-centre-for-hpc-hpda-and-ai/about/
Competence Centre France	https://cc-fr.eu/
Competence Centre Luxembourg	https://www.luxinnovation.lu/national-competence-centre-in-hpc/
Competence Centre Netherlands	https://www.surf.nl/onderzoek-ict/rekendiensten
Competence Centre Belgium	https://www.enccb.be/
Competence Centre Slovakia	https://eurocc.nscc.sk/
Competence Centre North Macedonia	https://www.hpc.mk/index.php/home/
Competence Centre Iceland	https://ihpc.is/community/
Competence Centre Montenegro	https://eurocc.udg.edu.me/

Table 2: The 33 EuroCC HPC NCCs

EuroCC and CASTIEL did both start on 1st of September 2020.

Within the EuroCC project under the European Union's Horizon 2020 (H2020), participating countries are tasked with establishing a single National Competence Centre (NCC) in the area of high-performance computing (HPC) in their respective countries. These NCCs serve as a contact point for customers from industry, science, (future) HPC experts, and the general public alike. Many of them also coordinate activities in all HPC-related fields at the national level. The EuroCC project is funded 50 percent through H2020 (EuroHPC Joint Undertaking [JU]) and 50 percent through national funding programs within the partner countries. The EuroCC activities – with 33 member and associated countries on board – is coordinated by the High-Performance Computing Centre Stuttgart (HLRS). The project aims to elevate the participating countries to a common high level in the fields of HPC, HPDA and artificial intelligence (AI)[7].

The Coordination and Support Action (CSA) CASTIEL leads to cross-European networking activities between National Competence Centres (NCCs) in HPC-related topics addressed through the EuroCC project. CASTIEL emphasizes training, industrial interaction and cooperation, business development, raising awareness of HPC-related technologies and expertise. As a hub for information exchange and training, CASTIEL promotes networking among NCCs and strengthens idea exchange by developing best practices. The identification of synergies, challenges, and possible solutions is implemented through the close cooperation of the NCCs at a European level[7].

2.2.3 ETP4HPC

The ETP4HPC is the European Technology Platform (ETP) in the area of High-Performance Computing (HPC). It is an industry-led think-tank of European HPC technology stakeholders: technology providers, research centres, and end-users. The main objective of ETP4HPC is to define research priorities and action plans in the area of HPC technology. ETP4HPC issues and maintains a Strategic Research Agenda as a mechanism to help the European Commission define the contents of HPC Work Programmes. It also acts as the "one voice" of the European HPC industry in relation to the European Commission, EuroHPC JU and national authorities. ETP4HPC was formed in October 2011[3].

The ETP4HPC Strategic Research Agenda (SRA) is a key deliverable that outlines a roadmap for the technological developments within the European HPC ecosystem [3]. As indicated by ETP4HPC, the milestones set in the SRA help the European Commission – under the EuroHPC Joint Undertaking – define the contents of the HPC Technology R&D work programmes as part of the European Research programme.

2.2.4 EuroHPC HPC Summit Weeks

The EuroHPC Summit Weeks are yearly premier European HPC conference events hosted by the EuroHPC JU in a member country. They cover a wide spectrum of keynotes, presentations, panels and workshops on topics relevant to the European HPC ecosystems, and also provide a platform for HPC projects funded by the EC or by EuroHPC to conduct their own sessions and workshops.

3 T2.3 Activities with the European HPC Ecosystem

At the start of the FocusCoE project, Task 2.1 set up the contact list of CoEs that was used and kept up-to-date by T2.3 during the complete reporting period. Later on, with the addition of four new CoEs T2.2 established connections with these and thus enabled a smooth communication flow between the FocusCoE project, including T2.3, and all CoEs. When requests for inputs from CoEs or request to assist with several networking activities between CoEs and other ecosystem representatives were asked for, T2.3 took over the handling of the process. As was described in D2.3, the four main steps to handle these requests were the following:

1. Analyse the request to make sure it fits to the profile and scope of the CoEs.
2. Discuss the level of details of the requested information with the requestor. Accept/reject the request based on its relevance to CoEs.
    a. If needed, adjust and simplify the request to the profiles of CoEs
3. Communicate the request with CoEs and gather their feedback
    a. If needed, communicate the request with each CoE separately based on the type or the delay of the answer from CoE
4. Collect and (if needed) consolidate the answers from CoEs before communicating back to requestors.

These steps ensured quick and qualitative results of networking activities.

3.1 Activities M1 – M18 – Summary

As was described in D2.3, during the first period, T2.3 received five requests for the CoE's input coming from both the EU HPC Ecosystem representatives and industry representatives. The five requests to the CoEs were:

1. Short questionnaire by ETP4HPC/EXDCI-2 concerning the contractual Public Private Partnership in HPC (HPC cPPP)
2. Use case input for the ETP4HPC Strategic Research Agenda (SRA 4)
3. Input for legacy code studies of the EXDCI2 project
4. Questionnaire from WIK-Consult GmbH (Research and Advisory Institute for Communications Services of Germany) for the Multiannual Financial Frameworks (MFF) and Connecting Europe Facility (CEF)
5. Request of the CoEs' contact points by European Open Science Cloud (EOSC) for inclusion in their portal.

T2.3 proactively worked on these requests, analysing them and ensuring their suitability for the CoEs. Furthermore, the requests were adjusted and communicated to the CoEs in a coordinated manner. The timely input received from the CoEs for these requests were clear indications of their interest and appreciation of T2.3 coordinated activities. The results of these requests were communicated back to the requestors and served as part of the important inputs to the HPC Ecosystem from the CoEs.

3.2 Activities M19 – M40 – Details

During the second project period (M19-M40), T2.3 carried out nine activities involving key HPC ecosystem actors and/or initiatives. In the following subsections, we describe these nine requests and present the successful handling of these by T2.3.

3.2.1 EuroHPC JU Pre-Exascale Hosting Entities

Contact establishment

FocusCoE established contacts with all three hosting consortia (Finland, Italy and Spain) and organized online-meetings with both, LUMI and Leonardo. BSC was also contacted, but as their procurement was delayed, no information could be shared at that juncture and thus no further interaction took place. MareNostrum5 procurement is ongoing at the time of writing.

On 11th of November 2021, an HPC3 online meeting with LUMI and Leonardo was organized. The consortia presented details of the architecture, support program, and access mechanisms of their systems. The material was made permanently available to all CoEs, so CoEs that were not able to attend the specific event could study the material afterwards. The CoEs were also provided with subsequent information on future training events specific to LUMI and Leonardo. The material was also discussed in subsequent HPC3 meetings where most of the discussions centred around access mechanisms and the CoEs were introduced to the EuroHPC access policies that were under development at the time.

3.2.2 EuroHPC JU Petascale Hosting Entities

Contact establishment

On 13th of April 2021 the meeting of HPC3 and the representatives of Petascale hosting sites was organised by T2.3. During the meeting the following questions were discussed and clarified:

- EuroHPC JU access policy was discussed, and the points related to the topics like preparation of benchmark applications and support for users from the CoE applications communities were raised. The discussions led an agreement to pursue CoE-Petascale project collaborations around high-level application support: the Petascale projects /hosting entities would provide first-level support to their users but would really be able to benefit from (pragmatically organised, best effort) second-level application support from the CoEs. It was mentioned that the application development and advanced users linked to the CoEs are very keen to support and expand their respective user communities.
- It was discussed that there were indeed application priorities and support expertise that varied between the hosting centres and that support needs from the CoEs (or indeed specific CoEs) would correspondingly differ between the Petascale centres.
- Collaboration between the CoEs and the various players in the ecosystem was stressed to be very important, notably with the NCCs, since there are often close collaborations between the Petascale projects and related NCCs. The potential collaboration with relevant Digital Innovation Hubs (a DIH on HPC-AI was mentioned) was also raised.
- Two-way collaboration (i.e. collaboration with mutual benefit) was stated as an important and feasible objective, and potential interests from the CoEs were the expansion of user communities and possible interaction with national industry sectors.
- It was reported that great progress had been made (beyond expectations at the start of the CoEs) in developing the application codes for the expected EuroHPC machines (e.g. GPGPU porting) and it was explained that there has been (across many CoEs) a great effort to improve software engineering to use software stacks/middleware to

achieve "separation of concern" (e.g. isolating application/modelling as opposed from computer architecture/programming model specifics).

- There was a general agreement that the discussion had been positive and fruitful. The Petascale projects were asked to either contact the individual CoEs directly or ask FocusCoE to assist in making the right contacts, particularly w.r.t the particular applications codes of relevance for their user constituencies.
- The CoE training offer was mentioned as clearly beneficial to the Petascale project users and information about these, the applications and codes, and also other CoE news and activities can be found on the CoE website.

3.2.3 HPC NCCs: Results Sharing and Introductory Workshop

The T2.3 leader is also one of the NCC representatives; it started the communication and networking activities between the two projects EuroCC and FocusCoE by discussing common interests with both during bi-lateral discussions and larger bi-lateral discussions.

Later, during one of the EuroCC workshops on industrial needs, T2.3 shared the information about the industrial questionnaire which FocusCoE conducted at the beginning of 2020. This information was very valuable for NCCs and on their request FocusCoE shared the summary of the results of this questionnaire with NCCs. The results were discussed by NCCs and impacted their future similar queries to industries.

As a follow up action on demand of NCCs, EuroCC organized an informative workshop about several EU projects and initiatives and with the assistance of T2.3. It was agreed that FocusCoE will be invited to be presented to NCCs as well. Prior to these workshops NCCs were asked to provide one question per NCC that could have been answered during the discussion part of the workshop. The FocusCoE leader presented the project to NCCs and detailed the information about the 15 CoEs, their activities, software supported and their relationships with industrial stakeholders. It was mutually agreed that CoEs and NCCs have significant common interests and collaboration options. The questions provided by NCCs were also answered during the workshop. Details of the resulting collaborations are listed in the subsections below.

3.2.4 HPC NCCs: Series of Thematical Workshops

Another initiative taken by FocusCoE and the CASTIEL project together was the organisation of thematic workshops for NCCs to get deeper insight in the activities of each CoE. T2.3 assisted in the organisation of these workshops as well. During the last project year, four workshops were given by CoEs to NCCs:

1. CoEs meet NCCs "2nd session about Quantum mechanical simulations in chemistry and materials science" on 8th of July2021

During this workshop, representatives of the TREX CoE and Max CoE presented the ambition and objectives of the CoEs and, after each presentation, a discussion followed which was triggered by the questions of NCCs. During these discussions NCC Slovakia shared experience of collaborating with TREX. Similarly, TREX representatives expressed their interest in collaborating with NCC Sweden to host a training on the CMake build environment. NCC Cyprus also expressed their interest in collaboration with Max and TREX based on the high interest in Cyprus in materials science.

A question was also raised about the type of training offered by the CoEs – it was explained that these are strongly linked to the applications and supplement the generic trainings offered, for example, by PRACE. It was agreed that it would be helpful to have an overview of general vs application-oriented trainings.

2. CoEs meet NCCs "2nd session about CFD" on 22nd of September2021

First representatives of the CompBioMed, EoCoE, EXCELLERAT, COEC, PerMedCoE, RAISE CoEs presented their activities and ambitions. Afterwards, similar to the previous event, discussion followed based on the questions from NCCs.

It was mentioned that NCC key CFD topics indeed overlap with themes presented by the CoEs in the meeting including (but not limited to): modelling of molten metals and crystals (multi-phase, multi-scale); wind turbines; combustion.

The final discussion concerned distribution of contact details (in particular email) of all participants willing to share them and provided to all participants. There is no plan or expectation to share contact details in a wider context.

3. CoEs meet NCCs "3rd session about Code optimization" on 23rd of November 2021

PerMedCoE, POP and TREX work on detailed code profiling and optimisation tools. The discussion after the presentations concentrated on POP services that are provided free of charge. The current expectation is that the service offering will be within the lifetime of the POP-2 project but it was pointed out that support was continued on a best effort basis in the interim period between POP-1 and -2 funded projects. The possible exchange of Customers in respective regions of NCCs was also discussed. To be considered: GDPR/Confidentiality. NCC interactions with code developers and support requests for code optimisation was another point of interest. As a conclusion, interest stated from CoEs and NCCs in common pool of tools related to performance.

4. CoEs meet NCCs "4th session on Complex Workflows and Data Analysis – relevant tools and developments" on 7th of February 2022

Representatives of the BioEXCEL, CoEC, EoCoE and RAISE CoEs presented news, ideas and collaboration possibilities with NCCs. Tools like Jupyter Lab, building blocks to make the tools easier into pipelines, and data assimilation methods were of particular interest. After their presentations, the discussions continued about synergies between several NCCs and CoE especially related to AI and data analyses on HPC.

3.2.5 HPC NCCs: Specific Collaborations

After the initial meeting between FocusCoE and NCCs initiated by T2.3, there were several ideas for collaborations between CoEs and NCCs and as a first step the training related to the software supported by CoEs was of highest priority for NCCs. We list three collaboration as starting examples:

- NCC Portugal asked BioExcel CoE to organise a Gromacs training event in Portugal. The request was gladly accepted, and the event took place in September 2021 with very positive feedback from participants.
- NCC Sweden asked EXCELLERAT to organise a Nek5000 training event in Sweden. The training was given in August 2021 and was very successful as well.
- NCC Sweden asked HiDalgo to organise a training event related to data management software on HPC. The event took place in April 2021. The participants were mostly from Swedish climate science community, but there were also participants from other NCCs.

In addition to collaborations related to training, there is also significant interest from NCCs in collaborating on software supported by CoEs. In a collaboration between TREX CoE and

NCC Sweden, a library is being developed within the OpenMolcas software which will be used as an input data generator for one of the TREX supported software on to be tested on pre-Exascale machines for acceleration purposes.

Further collaborations between CoEs and NCCs are being discussed and can be anticipated to take place in the future.

3.2.6 EuroHPC Summit Week 2021 and 2022

The EuroHPC Summit Week in 2021 was an online event and FocusCoE project was one of the contributors that organised a workshop titled "Preparing European Applications for the Exascale Era - Successes and Perspectives from the HPC Centres of Excellence" [8].

During that workshop, CoEs presented their activities targeting the Exascale era, discussed the success stories and challenges faced. The workshop was finalized with a panel discussion titled "Applications Support in EuroHPC - Status and Future: Centres of Excellence in Context" with panel members Bastian Koller (EXCELLERAT CoE coordinator), Daniel Opalka (EuroHPC JU officer), Edouard Audit (EoCoE CoE coordinator), Elisa Molinari (Max CoE coordinator, Joost VandeVondele (Forward-looking Software Solutions work-package leader within PRACE), Juan Pelegrin (Head of sector, European Commission), and Sinead Ryan (Vice-Chair of PRACE Scientific Steering Committee). The panel was moderated by Erwin Laure (HPC3 Secretary & FocusCoE WP2). The following questions were discussed:

- Without efficient applications, the impact of EuroHPC investments will be severely limited. Efforts are on the way at multiple levels: within groups & organizations, via national projects, and via EU efforts, most notably the Centres of Excellence but also within PRACE and the recently created national Competence Centres.
    - are those efforts effective and sufficient?
    - how can we reach domains that are currently not covered?
- Application development and optimization is a long-term effort that requires highly skilled people, yet most of the mentioned EU-efforts will end towards the end of this year.
    - how can we avoid a damaging loss of people?
    - how can application development be secured in the long term?
- Related efforts in the US or Japan happen in large, tightly coordinated projects (like the US ECP) while Europe has a more decentralized approach
    - should Europe change its approach?
      what are the pros and cons?
- In the future, disruptive technologies might appear (e.g. Quantum Computing)
    - should application developers already prepare for these?

In a lively discussion, the need for long-term stable financial support for application development was stressed. The CoEs in their current form were deemed to be an excellent setup to cover both, the breadth and depth in scientific application development, yet mechanisms to include disciplines not yet covered as well as ISVs will be needed. Future disruptive technologies will need to be monitored, but they are at a too early stage for serious application development.

At the same event, another workshop "HPC Education & Training – Perspectives from EU13 Member States" was organised by FocusCoE WP4 and is described in D4.2.

During the EuroHPC Summit Week in March 2022, ta workshop took place within the "HPC Ecosystem & EuroHPC Day" (22nd of March) entitled "European HPC CoEs: perspectives for a healthy HPC application eco-system and Exascale" organised by HPC3 with the strong support of FocusCoE WP2 and Task 2.3 in particular [9]. During this workshop, most of the

CoEs will be present and will discuss question related the future activities of CoEs within European HPC ecosystem, their contributions to Exascale software and engagement in support of different types of stakeholders to use such software.

3.2.7 Input to the ETP4HPC Strategic Research Agenda 5

The SRA5, which is still under development at the time of writing this deliverable, identified ten working groups: system architecture, system hardware components, system software and management, programming environment, I/O & storage, mathematics & algorithms, application co-design, centre-to-edge-framework, Quantum Computing for HPC, and non-conventional HPC architectures.

Many CoE members contributed to several of these working groups, and FocusCoE in particular ensured appropriate CoE input to the application co-design chapter. Drafts of the chapter were made available to HPC3 for comments and additions, so CoEs who had no members in the working group were also able to provide their input. In a first step, research trends and current state-of-the-art as well as challenges for 2023-2027 were identified. The current draft of this input can be found in the appendix (note: this might still change). In a second step, overlaps with “research clusters” (whose definition is still ongoing inside ETP4HPC) will be defined. This work, although taking place after the end of FocusCoE, will still be supported and managed through the HPC3 office.

3.2.8 Full-day Webinar about Intel’s oneAPI unified programming model for CoEs

On 2nd of March, 2022, FocusCoE organised a full-day Webinar about Intel’s oneAPI unified programming model, which covers CPUs, GPGPUs and other accelerators. 20+ participants across the active CoEs participated and were briefed on the oneAPI approach, programming, and optimisation and debugging tools [10]. These are available free of charge.

Since Intel and SiPearl announced a technical collaboration to support SiPearl’s CPUs and Intel’s upcoming GPGPUs, targeting future European Exascale systems, this topic has significant relevancy for EPI and the European HPC ecosystem at large.

4 Conclusion

During the second project period, FocusCoE T2.3 handled nine request from different HPC ecosystem stakeholders and contributed to the networking activities between these actors and CoEs. Many of the activities carried out were related to establishment of tight communication between EuroCC, NCCS, and CoEs as well as between EuroHPC JU pre-Exascale and Petascale hosting sites. These activities were rather successful and also contributed to the tight ties between current HPC ecosystem representatives and CoEs.

Based on these activities, there were several important aspects that should be considered in future communications between CoEs and HPC ecosystem. These are listed below:

- learning / understanding the structure, goals, maturity, and needs of the different players in the ecosystem takes time and not all encounters can bear fruit immediately.
- managing large "sets" of players like the CoEs or NCC can be challenging, especially when they should "meet" somehow and produce mutual benefits.
- Intermediates like CSAs can be very helpful, however there also needs to be a one-to-one connection.
- With the series of thematic meetings of CoEs and NCC, we have explored one workable way, but it might be useful to further optimise the approach or experiment with related ideas.

- There is no "one size fits all" and one has to be flexible about approaches, especially when the communities are diverse.

5 References

[1] https://www.hpccoe.eu
[2] https://ec.europa.eu/digital-single-market/en/news/ten-new-centres-excellence-hpc-applications
[3] https://www.etp4hpc.eu/sra.html
[4] https://eurohpc-ju.europa.eu
[5] https://www.lumi-supercomputer.eu/may-we-introduce-lumi/
[6] https://eurohpc-ju.europa.eu/press-release/leonardo-new-eurohpc-world-class-pre-exascale-supercomputer-italy
[7] https://www.eurocc-access.eu/
[8] "FocusCoE organizes two workshops at EHPCSW21", recordings available at https://www.hpccoe.eu/2021/04/30/focuscoe-organizes-two-workshops-at-ehpcsw21/
[9] "FocusCoE at EuroHPC Summit Week 2022", news entry on the FocusCoE/HPC3 session at EHPCSW'22, https://www.hpccoe.eu/2022/03/24/focuscoe-at-eurohpc-summit-week-2022/
[10] "FocusCoE Hosts Intel OneAPI Workshop for the EU HPC CoEs", https://www.hpccoe.eu/2022/03/04/focuscoe-hosts-intel-and-eu-hpc-coes-for-oneapi-workshop/

6 Annex

6.1 Input to ETP4HPC SRA5

1.1.1 ***Application co-design***

1.1.1.1 **Research trends and current state of the art**

Sustaining excellence and European world-leadership in HPC applications is key for European science, industry (incl. SMEs) and the public sector. There is a breadth of applications in the fundamental, applied and social sciences where computing plays a pivotal role[3]:

- A world class European computational infrastructure will expand the frontiers of fundamental sciences like physics and astronomy, supporting and complementing experiments. Researchers will be able to simulate the formation of galaxies, neutron stars and black holes, predict how solar eruptions influence electronics, and model properties of elementary particles. This will explain the source of gamma-ray bursts in the universe, advance our understanding of general relativity, and help us advance understanding of the fundamental structure of matter by means of simulating the theory of strong interactions called quantum chromodynamics. This fundamental research itself leads to advances in the state-of-the-art of scientific computing and helps attract new generations to science, technology, engineering and mathematics.

- Simulations are critical in Climate, Weather, and Earth Sciences. Exascale resources will enable sub-kilometre scale resolutions, a more realistic representation of all Earth-system components, better mathematical models, and larger ensembles of simulations for uncertainty quantification. This will extend the reliability of forecasts to the extent needed for the mitigation and adaptation to climate change at global, regional and national levels, in particular with respect to extreme events. In addition, coupling of mesoscale and microscale models will enhance the accuracy of pollutant propagation in cities, mid-term power generation in wind farms, etc. In analogy to weather and climate prediction, much enhanced simulation capabilities of solid Earth physics from higher spatial resolution and seismic frequencies down to 10Hz will enable a break-through in the detection and prediction of the precursors of volcanic eruptions and earthquakes, and their impact on infrastructures. A prediction capability at this level of detail is crucial for a wide range of societal impact sectors for food and agriculture, energy, water management, natural hazard response and mitigation, and finance and insurance.

- High-end computing capabilities are becoming increasingly important for life sciences, medicine, and bioinformatics and will have tremendous impact, e.g. for enabling personalised medicine. Researchers are already able to rapidly identify genetic disease variants, and it will become possible to identify diseases that are caused by combinations of variants and design treatments tailored both to the patient and state of the disease. Structural biology will increasingly rely on computational tools, allowing researchers to predict how the flexibility and motion of molecules influences function and disease. Deep learning techniques will provide more specific diagnoses and treatment plans than human doctors, making medical imaging one of the largest future computing users. At anatomical scales, organ level simulations present both a challenge and an opportunity: the Virtual Patient modelling for precision medicine. It is the ultimate example of supercomputer usage for

[3] See for instance: The Scientific Case for Computing in Europe (period 2018-2026), PRACE

multiphysics/multiscale modelling problems. Tumors can be simulated by following the fate of thousands, millions or even billions of cells as their internal circuitry of signaling molecules respond to each other and their environment. This approach can be used to predict the effect of targeted therapies, which is particularly useful to triage potential combination therapies based on the patient's molecular profile. At organ level, modelling is done by tightly coupling different physics (e.g. fluid, tissue, electrophysiology, chemical reactions, heat, transport of large bodies, particles or species) with contributions from different temporal and spatial scales (cells, tissue, organ, system). To make things more complex, these problems present issues such as patient variability and comorbidities in complex geometries, with extremely difficult validation. A strategy to address these two issues is to run problems in a virtual patients' population. All these things together make the use of supercomputers a decisive factor.

- For Energy applications, the oil and gas industries are moving to full waveform inversion combined with neural networks for accurate detection. Exascale resources will make this technique feasible, allowing more accurate predictions of reservoirs and, although still fossil fuels, oil and gas have much reduced CO_2 emissions and air pollution compared to those produced by coal, which is still the dominant source of energy in the world. Simulations are of importance to improve the efficiency of nuclear power, hydropower, wind turbines and, not least, batteries and high-voltage cables to enable transmission and storage. Likewise, simulations are essential for the discovery and optimization of renewable forms of energy, their storage and distribution. Accurate magnetohydrodynamic simulations of plasma are critical for fusion energy, as in the ITER project.

- Computing is already used widely in Engineering & Manufacturing. Engineering applications based on fluid dynamics, combined with orders-of-magnitude-faster resources, will enable direct numerical simulations of the governing equations of fluid mechanics with better accuracy, leading to improved designs and thus significantly better fuel efficiency e.g. for cars and airplanes, while also helping us understand phenomena such as cavitation, flow separation and pollutant formation. New data-driven approaches will enable scientists in academia and industry to integrate all aspects of design in models, use information from internet-of-things sensors, include uncertainty quantification in predictions, and consider the entire life cycle of a product rather than merely its manufacture.

- Chemistry & Materials Science will remain one of the largest users of computing, with industry increasingly relying on simulation to design, for example, catalysts, lubricants, polymers, liquid crystals, and also materials for solar cells and batteries. Electronic structure-based methods and molecular dynamics will handle systems, properties and processes of increased complexity, and drive towards extreme accuracy. These methods are being complemented both with multi-scale models and data-driven approaches using high-throughput and deep learning to predict properties of materials and accelerate discovery. This will enable researchers to address the grand challenge of designing and manufacturing all aspects of a new material from scratch, ushering in a new era of targeted manufacturing.

- Global Challenges, like urban air pollution, financial crises, behaviours in social networks, pandemics (like COVID-19), management of disasters or conflicts, and migration streams require complex simulations, using agent-based systems with a huge number of elements, and specific coupling with other simulations (such as weather and climate predictions and other CFD simulations in 3D models representing digital twins of real locations). These simulations can be used by policy makers in order to understand how to solve existing global issues or how to mitigate their negative effects. Therefore, it is necessary to advance research on agent-based simulations, coupling technologies and executing ensemble methods, as well as leveraging state-of-the-art HPDA and AI techniques. All of this significantly contributes to the retrieved models, their understanding, and the further extraction of knowledge for scientific-based decision making.

- Throughout this spectrum of applications, handling of complex data plays an increasingly important role. Existing research fields are beginning to use deep learning to generate knowledge directly from data instead of first formulating models of the process, so next-generation infrastructure must be able to handle these applications with dramatically increased data storage and I/O bandwidth capabilities. This will be needed for e.g. autonomous driving, Industry 4.0 and the Internet of Things, but will also enable the application of computing across a whole range of non-traditional areas including the humanities, social sciences, epidemiology, finance, promoting healthy living, determining return-on-investments for infrastructure by considering behavioural patterns and, not least, in helping to develop society and secure democracy, of which Cyber Security is an important aspect.

- As the previous topics show, industry can be found almost everywhere when it comes to large scale simulation science. While in many cases, especially when SMEs are using simulation, data analytics, and AI, industry does not yet require the fastest systems available, it is absolutely crucial that industry has a direct link to those who use these systems and to their technologies in order a) to prepare for their own use as soon as they have a need for this kind of performance and b) to be able to run exceptionally complex tasks, not every day but for particularly large challenges when this compute power is needed to open up new possibilities. If industrial companies are not able to seamlessly connect to the highest-end simulations, data analytics, and AI technologies, they will suffer severe disadvantages compared to those that do.

1.1.1.2 Challenges for 2023-2027

The computing requirements of these scientific and industrial applications drive technological development and applications developers must respond in turn by porting, adapting (including the use of new methods and algorithms), and optimizing application codes for the new technologies. Knowledge of and sufficient access to these new technologies for application developers will assist in ensuring that full scale systems can be used in the most efficient manner. An effective collaboration, including co-design processes, between application and technology developers is crucial for a successful HPC ecosystem to ensure technology is relevant for applications and can be fully exploited when available. It must also be recognised that not all application maintainers possess the knowledge to independently and effectively conduct the necessary code modifications for new technology. Specialist knowledge and support will also be needed to assess the level of revision needed to move applications which are successful at small scales to large-scale HPC infrastructures. Application co-design will involve providing support to application codes and tools as they move through these processes. In addition, the challenge of transferring knowledge between computational and scientific specialists extends beyond porting and optimisation of applications and should include training existing and upcoming researchers in skills critical to fundamental HPC usage. This is even more essential in research fields that have a relatively immature level of uptake and competence in HPC usage. Incorporating all of these issues into a wider approach will foster an ecosystem of skills and expertise in Europe by retaining scientific talent and enhancing the competitiveness of industries on the global market.

Educating new generations of HPC experts and computational scientists is crucial for a healthy HPC environment. In addition, such people need rewarding employment conditions with clear and structured career paths to retain skilled personnel and avoid a brain drain to other parts of the world. Unfortunately, within the traditional academic systems that exist in most European countries this is still difficult for computational experts in certain disciplines and even more so for interdisciplinary experts.

Access to increased computational power and to applications exploiting the features of the system remains crucial to enable more detailed and large-scale (compute intensive) modelling and simulations. At the same time, new approaches like data-driven computing, High Performance Data Analytics, and AI, and their convergence with classical HPC enable new opportunities and require new capabilities. This includes efficient access to large amounts of data with low latencies and high-bandwidth and support for new and large workflows and ensembles encompassing orders of magnitude more active tasks or computational jobs than today. Additionally, some simulations require coupling with others, since their results need to be synchronised, requiring further development in coupling techniques (from data sharing to message passing) and adequate resources access and allocation mechanisms.

Many applications are severely limited by memory bandwidth or communication latency, and as the throughput of floating-point operations has increased faster than the data transport capabilities, even many traditionally floating-point bound applications are now highly memory sensitive. Intra-node and inter-node communication require drastically reduced latencies and high-speed networks, particularly for algorithms based on fast iterations of short tasks so that they can achieve significantly improved performance and strong scaling.

Storage and I/O requirements are expected to grow even faster than compute needs, with much larger data sets being used e.g. for data-driven research and machine learning. This is not limited to the amount of storage, but data-heavy applications will also need exceptionally high-bandwidth parallel file systems, and/or advanced data caching solutions on each node. This increase in storage and I/O resources must be coupled with provisioning of a large-scale end-to-end data e-infrastructure to collect, handle, analyse, visualize, and disseminate petabytes to exabytes of data. In addition, some applications like genome analyses or AI applications require ancillary data like reference datasets or underlying models. Making them available on fast storage with pre-configured access mechanisms can significantly speed-up such applications.

Long-term maintenance and portability of codes (both in terms of performance portability as well as compiling/building the codes on new architectures) are other important factors, requiring standardized, open, and supported programming environments and APIs, including container technologies, supporting a wide range of different hardware technologies. Investing in tools automatizing and tracking deployed software stacks, can help to sort out the challenging combination of codes and architectures. In addition, the software engineering practices like unit testing, continuous integration and deployment need to improve. These efforts typically performed by the development teams, need to be complemented with community benchmarking activities, helping to select the right software for certain environments and ensuring the ~~scientific~~ validity of the results. A particular challenge is the uncertainty about future hardware developments. Porting established codes is a major undertaking and with uncertainties about the long-term future of certain hardware technologies, there is often a reluctance to engage in these expensive endeavours. Performance portability frameworks can shield application developers from hardware changes to some extent, but long term, stable hardware roadmaps are of equal importance.

Many of today's applications are made up of millions of lines of code. Analysing these monolithic codes to identify expensive computational kernels to port to future architectures can be difficult. Creating application dwarfs to simplify the execution and interaction with the algorithmics can be a solution to enhance this task. As an example, in Earth Sciences, dwarfs represent functional units in the forecasting model, such as an advection or a physics parameterization scheme, which also come with specific computational patterns for processor memory access and data communication. To ensure a smooth user experience, dwarfs should be distributed with an accurate user guide and examples of input files or name lists. In addition, modularization of codes and the application of modern software engineering best practises are important.

While in the longer-term computing leadership will require the development of alternatives to the current technologies, including perhaps quantum, dataflow, neuromorphic or RNA computing, there is consensus that even today the fundamental mathematical and computer science algorithms that are needed to meet the requirement of leadership science and industrial competitiveness are not in place. The energy-efficient, application-oriented next-generation computing platforms therefore require an ambitious programme of algorithm development integrated with the co-design/co-development of the overall infrastructure and sustained over longer timescales than the usual 3 to 5 years funding cycles. This is particularly true for quantum computing, where in Europe and globally a vibrant research environment is being built rapidly. World-leading technology and applications innovation in this field requires significant investment in algorithmic and application development, ensuring that the anticipated quantum computers will be useful for the European Research Area. This must include research funding and support for fundamental new algorithms and quantum information theory, and for robustness, reproducibility, data & I/O and the convergence with classical computing.

All these approaches require co-design activities involving architectures, OS, communication libraries, workload management and end-user applications to achieve the intended results.