H2020-FETHPC-3-2017 - Exascale HPC ecosystem development

EXDCI-2

European eXtreme Data and Computing Initiative - 2

Grant Agreement Number: 800957

D3.2
First Report on joint brainstorming sessions among scientific and industrial users communities

Final

Version: 1.1
Author(s): Mathis Bode, RWTH Aachen University
Date: 24/09/2019

Project and Deliverable Information Sheet

EXDCI Project	**Project Ref. №: FETHPC-800957**	
	Project Title: European eXtreme Data and Computing Initiative - 2	
	Project Web Site: http://www.exdci.eu	
	Deliverable ID: < **D3.**2>	
	Deliverable Nature: <DOC_TYPE: Report>	
	Dissemination Level: PU	**Contractual Date of Delivery:** 24 / June / 2019
		Actual Date of Delivery: 21 / June / 2019 (Revised version resubmitted 30 / 09 /2019)
	EC Project Officer: Evangelos Floros	

* - The dissemination level are indicated as follows: **PU** – Public, **CO** – Confidential, only for members of the consortium (including the Commission Services) **CL** – Classified, as referred to in Commission Decision 2991/844/EC.

Document Control Sheet

Document	**Title: First Report on joint brainstorming sessions among scientific and industrial users communities**	
	ID: D3.2	
	Version: <1.1>	**Status:** ***Final***
	Available at: http://www.exdci.eu	
	Software Tool: Microsoft Word 2013	
	File(s): D3.2-v1.1.docx	
Authorship	**Written by:**	Mathis Bode,
	Contributors:	Stéphane Requena, GENCI/PRACE Jean Claude André, JCA Consultance Rossen Apostolov, KTH Lee Margetts, Univ. of Manchester
	Reviewed by:	Mark Asch, Université de Picardie John Clifford, PRACE aisbl
	Approved by:	MB/TB

Document Status Sheet

Version	Date	Status	Comments
0.1	05/06/2019	Draft	Before internal review
0.2	12/06/2019	Draft	After internal review 1
0.3	18/06/2019	Draft	After internal review 2
1.0	21/06/2019	Final version	Approved by TB/MB

1.1	24/09/2019	Final version	Revised version following Interim Review.

Document Keywords

Keywords:	PRACE, , Research Infrastructure, Roadmaps for new usages, CoEs, FETHPC, BDEC, HPC, HDPA

Table of Contents

List of Figures

References and Applicable Documents

List all external documents referenced in this document

[1] http://www.exdci.eu
[2] http://www.prace-project.eu
[3] http://www.etp4hpc.eu
[4] https://aioti.eu/
[5] https://www.exascale.org/bdec/

[6] http://www.bdva.eu/

List of Acronyms and Abbreviations

Below is an extensive the List of Acronyms used in previous deliverables. Please add additional ones specific to this deliverable and delete unrelated ones.

AI	Artificial Intelligence
AIOTI	Alliance for Internet of Things Innovation
BDEC	Big Data and Extreme-scale Computing
BDVA	Big Data Value Association
CoE	Centres of Excellence for Computing Applications
CSA	Coordination and Support Action
D	Deliverable
DL	Deep Learning
DoE	Department of Energy (US)
EC	European Commission
ECP	Exascale Computing Project (US)
EIP	European Innovation Partnership
ETP4HPC	European Technology Platform for High Performance Computing
FET	Future and Emerging Technologies
H2020	Horizon 2020 – The EC Research and Innovation Programme in Europe
HPC	High Performance Computing
HPDA	High Performance Data Analytics
HW	Hardware
M	Month
ML	Machine Learning
PRACE	Partnership for Advanced Computing in Europe
SME	Small and Medium Enterprise
SRA	Strategic Research Agenda
SSC	Scientific Steering Committee
SW	Software
WP	Work Package

Executive Summary

It is essential for WP3 to stay in contact with the various players/users in/of the European HPC infrastructure. For that, several brainstorming/data gathering activities have been carried out, which are summarized in this deliverable. Contacts have been initiated with a number of scientists in various fields likely to be either impacted by new High Performance Computing (HPC)/High Performance Data Analytics (HPDA) technologies or at the origin of new applications. Among them: representatives of European Centres of excellence (CoEs), scientists involved in relevant Future and Emerging Technologies (FET) projects, international researchers active in a variety of other domains like nuclear fusion, artificial intelligence, social sciences, etc.

WP3 experts are also collaborating with Work Package 2 (WP2) in the process of updating the SRA-4 (fourth version of the Strategic Research Agenda) by providing inputs from applications research communities coming from academia and industry. The SRA-4 will be issued at the end of 2019 and be delivered for the first time to the EuroHPC Joint Undertaking. The SRA-4 will involve ETP4HPC (European Technology Platform for High Performance Computing, [3]) for the hardware and software HPC industries, BDVA (Big Data Value Association, [6]) for the Big Data applications, AIOTI (Alliance for Internet of Things Innovation, [4]) for the IoT/AI applications and CoE / Partnership for Advanced Computing in Europe (PRACE, [2]) Scientific Steering Committee (SSC) for the scientific and industrial HPC applications.

Also, during the last 3 BDEC (Big Data and Extreme-scale Computing, [5]) meetings in Indiana (USA, November 2018), Kobe (Japan, February 2019), and Poznan (Poland, May 2019) experts from WP3 presented or were involved in several application use case presentations including data analytics in climate research, digital transition of Material Nano-Characterization and end-to-end workflow supports in earth sciences. Such international meetings were also the opportunity to exchange with applications experts from other continents including those from the DoE's Exascale Computing Project (ECP) who launched in early 2019 a new center of co-design toward Artificial Intelligence (AI) called ExaLearn aiming to bridge scalable AI and the massive amount of data generated by DoE's numerical simulations and large-scale scientific instruments.

Information with respect to industry has been collected during a joint workshop involving the PRACE Industrial Advisory Committee and the EuroHPC Infrastructure Advisory Group held on 3rd June 2019. This will be specified during the NAFEMS World Congress 2019, where activities are planned.

As next actions, more regular internal EXDCI communication channels to WP2 and WP4 will be established to synchronize energy. In agreement with this, the International Supercomputing Conference (ISC) 2019 in Frankfurt will be used to discuss current developments with various players/users in/of the European HPC infrastructure. Furthermore, the collaboration with the BDEC Application Group will be intensified.

1 Introduction

Work Package 3 (WP3) "Excellence in HPC applications and usages" of EXDCI-2 [1] focuses on applications and best practice usage in the context of potential requirements towards Exascale platforms. This concerns both classical High Performance Computing (HPC) and High Performance Data Analytics (HPDA) applications. As ascertained during the EXDCI project (i.e. the forerunner to the current project EXDCI-2), Europe is developing a significant fraction of the applications used in the world and the biggest producer of data. Therefore, this effort must be continued as science opportunities are evolving very quickly with the expected availability of Exascale supercomputers (and corresponding HPDA tools). Exemplarily, the following three points can be named:

- Large research infrastructure designs are evolving with new capabilities of HPC and Big Data.
- New tools and approaches are increasingly needed to take into account new technical realities.
- As users get a better understanding of the potential of HPC and Big Data applications, possibly combined, scientific goals are evolving rapidly and new application domains appear.

For scientific applications, WP3 relies on the PRACE (Partnership for Advanced Computing in Europe, [2]) Scientific Case and interacts with the PRACE user communities, and the HPC Centres of Excellence for Computing Applications (CoEs) to focus on the Exascale aspects, as well as on the influence of specific technological or algorithmic innovations. WP3 roadmaps HPC applications and usages and coordinates with HPC user communities and CoEs.

Similar considerations apply to industrial applications, but in this case, it is often required to fit the application in a complex proprietary workflow, for instance a "digital twin" of a product or manufacturing setup. This requires engaging and working directly with the users or their collective organizations including European Innovation Partnerships (EIPs) and national initiatives. The roadmap results are of importance to avoid gaps in the value chain, but also to evaluate scenarios where "disruptive innovation" entails changes in the value chain or permit different entry strategies.

To fulfil this role within the EXDCI-2 project, WP3 is divided into three tasks:

- Task 3.1 – Roadmap of HPC applications and usages
- Task 3.2 – Engagement with HPC users communities and CoEs
- Task 3.3 – Preparation of industrial codes to exascale

Mathis Bode (RWTH Aachen University) is WP3 leader and Stéphane Requena (GENCI/PRACE) is WP3 co-leader.

This deliverable entitled "D3.2 – First Report on joint brainstorming sessions among scientific and industrial users communities" was postponed from M15 to M16 due to the fact that the EuroHPC Summit Week 2019 in Poznan, Poland took place in M15 and was essential for this deliverable. Consequently, the deliverable outlines the work carried out in WP3 until and including M15 of the EXDCI-2 project. Relevant milestones during this period for this deliverable are:

- MS32 – 1st brainstorming session
- MS34 – Workshop during HPC Summit Week 2019 (Workshop held)

A detailed description of MS34 can be found in the deliverable "D3.4 – First report on the organisation of WP3 workshops during HPC Summit Week 2019" also due in M16 (postponed

from M15). This deliverable gives an overview of various brainstorming sessions during various events including workshops during the EuroHPC Summit Week 2019 and the NAFEMS World Congress 2019 (scheduled). It also describes brainstorming with the CoEs through established communication channels and the *ad hoc* PRACE-CoEs-FETHPC-EXDCI workshop in Brühl (Germany, Oct. 30 and 31, 2018).

2 Summary of Deliverable

This deliverable summarizes the brainstorming activities with respect to various topics and as part of various events of WP3. The brainstorming activities are separated into four parts:

- Brainstorming within CoEs and Future and Emerging Technologies (FET) projects – This activity is mainly to prepare the update of the "Roadmap of HPC applications and usages" (cf. Chapter 3).
- Brainstorming within Scientific Communities – This activity was carried out during the 1st European Communities Workshop on Exascale Computing organized by WP3 during the EuroHPC Summit Week 2019 (cf. Chapter 4).
- Brainstorming within EXDCI-2/PRACE – This activity has been continuously performed during various EXDCI-2/PRACE events such as the EuroHPC Summit Week 2019 (cf. Chapter 5).
- Brainstorming within Industrial Communities – This activity is scheduled for the NAFEMS World Congress 2019 and targets the special needs of industry (cf. Chapter 6).

The report ends with some conclusions and a summary of next steps.

3 Brainstorming within CoEs and FET projects

It should first be recalled that the main deliverable for WP3-Task 3.1, i.e. "Roadmap of HPC applications and usages" is due only on March 2020. Work has nevertheless started to identify and contact relevant stakeholders and to collect early inputs for this updated roadmap. There are however two types of reasons which make it difficult to prepare a timely and well-consolidated roadmap at a higher pace.

A first ensemble of reasons results from the fact that other vision documents have been published quite recently, or are underway, and that it may appear a bit early to bring well-documented and significant modifications to these proposed future roadmaps. For example, the PRACE scientific case (*The scientific case for computing in Europe, 2018-2026*) has been made available only in October 2018, after more than one year of consulting and working with the relevant scientific communities. Another example is the ETP4HPC (European Technology Platform for High Performance Computing [3]) strategic research agenda (SRA), which is undergoing regular revision, the last one being dated end of 2017 (SRA3), while the next version is already underway and should be available by the end of 2019 (SRA4). Not to mention the other documents proposed recently at the international level, e.g., the paper "Big data and extreme-scale computing: Pathways to convergence-toward a shaping strategy for a future software and data ecosystem for scientific inquiry" (*M. Asch et al., Int. J. High Performance Computing Applications, 2018, **32**, 435–479*).

A second ensemble of reasons is linked with the rapid evolution of both hardware and software architectures. The Exascale challenge and the developing needs for addressing within more and more applications significantly increases in HPC and HPDA performances and efficiencies call

for a convergence of HW architectures (CPU, GPU, possibly TPU) and for an hybridation of the SW stacks. These changes are in the process of happening, and many applications, both in the scientific and industrial domains, are still facing difficult questions: At what pace should these evolutions take place? Is there enough manpower to modify the codes? Is there a risk that some important changes do occur rapidly, making previous investments obsolete? Etc.

For the moment, the stakeholders, who have been identified up-to-now, are described here briefly together with the type of vision documents they might have already produced or are presently undertaking. Based on these inputs, a consolidated vision will be produced within the next six months.

3.1 CoEs

The CoEs have been established to explore and prepare for next-generation HPC and HPDA resources, each within its own scientific domain. They are obviously important initiatives to liaise with for preparing the type of roadmap under consideration. This point was first addressed during the first EXDCI-2 technical meeting (Brussels, Sept. 28, 2017), where it was proposed to include one representative of each CoE within an *ad hoc* "roadmap working group". This turned out to be very difficult, as the CoEs were in the process of organizing themselves under a new Coordination and Support Action (CSA) "FocusCoE", and that duplicating tasks for them was not welcome. It was alternatively agreed that EXDCI-2 would first draft its updated roadmap from documents already produced, or soon-to-be-produced, by CoEs, and that the draft would then be reviewed and commented on by the CoEs.

The meeting which took place in Brühl (Germany, Oct. 30-31, 2018), attended by representatives from almost all CoEs, was a first opportunity to meet with them and start identifying the documents to consider. At that time "BioExcel" and "E-CAM" had already available documents, while "EoCoE" agreed to hold an *ad hoc* meeting when necessary. Contacts were pursued during the EuroHPC Summit Week (Poznan, Poland, May 13-17, 2019), where additional information could be gathered from other CoEs: "CompBioMed", "HIDALGO", "Max1-2". Finally, other CoEs, like "ESiWACE" and "EXCELERAT", will be interviewed in the near future.

The collection of relevant information from the various CoEs should then be achieved by mid-summer 2019, in time for starting to draft the updated roadmap.

3.2 New paradigms, coupling with machine-learning

There are other well-established scientific domains, not all of them being covered by specific CoEs, where new paradigms develop rapidly. In some of these fields, like, e.g. nuclear fusion, new applications are presently developed where simulations based on first-principles is completed by Machine Learning (ML) or Deep Learning (DL) based on real-time processing of experimental data. The prediction of disruptive instabilities occurring in controlled fusion plasmas is such an example (see *J. Kates-Harbeck et al., Predicting disruptive instabilities in controlled fusion plasmas through deep learning, Nature, 568, 526-531*).

More generally, disciplines in which data has to be used for initializing or controlling the appearance of phenomena that are otherwise very difficult to predict in real-time from first-principles are obvious candidates for developing hybrid methods. Meteorology is another example: Data assimilation is already well-known as a method to make initial conditions compatible with a numerical weather prediction model. But other issues can possibly be tackled, like early identification of small-scale extreme events, either from experimental or simulated

data, or like ad-hoc parameterization of subgrid-scale phenomena. Applications in internal geophysics and seismic prediction are actively exploring such avenues.

3.3 Social sciences

The work has not started yet, but a number of institutions have been identified, with which contacts will be established shortly to understand their vision, e.g. new problems under reasonable reach with foreseen architectures.

3.4 The FET projects dealing with HPC

Although a bit more complex to collect, as being more dispersed than the above, information from FET projects dealing with new HPC/HPDA methods is also being collected. This search is based on a review of past FETHPC projects already conducted under ETP4HPC [3], and which has kindly been made available to EXDCI-2.

4 Brainstorming within Scientific Communities

During the last three BDEC meetings in Indiana (USA, November 2018), Kobe (Japan, February 2019), and Poznan (Poland, May 2019) experts from WP3 presented or were involved in several use case presentations including data analytics in climate research, digital transition of Material Nano-Characterization or end-to-end workflow supports in earth sciences (in strong collaboration with CNES and attached as Annex 2 of this document). One of the upcoming "demonstrators" of the BDEC2 initiative will focus in 2019-20 on "machine learning everywhere".

Such international meetings were also the opportunity to exchange with applications experts from other continents including the ones from the Exascale Computing Project (ECP) who launched in early 2019 a new center of co-design toward Artificial Intelligence called ExaLearn, aiming to bridge scalable AI and the massive amount of data generated by DoE's numerical simulations and large-scale scientific instruments.

ExaLearn is examining how methods for unsupervised, semi-supervised, and self-supervised learning technologies can be used to build domain science knowledge from large-scale data processed on extreme-scale supercomputers. To ensure broad use, these methods will provide solutions that are both interoperable and reproducible.

Already, ExaLearn is exploring how ML can be used to create realistic surrogate models to replace computationally expensive simulations (e.g. requiring substantial time or memory, such as those from large-scale simulations of the universe). Another target area involves using machine learning to find the optimal mix of protocols and computing parameters that can enhance experimental design.

Furthermore, WP3 organized a workshop during the EuroHPC Summit Week 2019 with the target to discuss important Exascale topics among various scientific communities. The workshop was designed as the start of a series of workshops and therefore denoted as "1st European Communities Workshop on Exascale Computing". It is planned to organize a second version during the EuroHPC Summit Week 2020. In order to support fruitful discussions, HPDA was chosen as the target topic and 12 speakers were invited to present their ideas/experiences/solutions/questions with respect to this topic.

As it is important to get a detailed idea about the complete European Exascale landscape, the EXDCI-2, WP3 workshop was organized in close collaboration with Guy Lonsdale from FocusCoE. Finally, six speakers representing different CoEs and six speakers representing different FETHPC projects were selected. The agenda of the workshop is shown in Figure 1 and Figure 2.

The workshop took place on Thursday, May 16th, 2019 from 2:30pm to 7:00pm, which was after the official end of PRACEdays 2019. About 50 people attended the workshop in total, which was a success. The presentation covered various aspects of HPDA from Big Data processing to reproducibility of simulation results. The discussions and questions showed that many European communities face similar challenges and would be thankful for more joint efforts and combined solutions. For WP3, it was important to get this feedback, which helps to produce a more relevant HPC roadmap.

As mentioned in the beginning, more details of this workshop can be found in the deliverable "D3.4 – First report on the organisation of WP3 workshops during HPC Summit Week 2019".

1st European Communities Workshop on Exascale Computing

Focus on High Performance Data Analytics

May 16, 2019, Poznan, Poland

hosted by EuroHPC Summit Week 2019, May 13 - 17, 2019

Agenda May 16, 2019

14:30	**Welcome and brief introduction**	**Giovanni Aloisio, Mathis Bode**
14:30-16:30	**Session 1 – CoEs & HPDA**	**Chair: Guy Lonsdale**
14:30-14:50	ESiWACE	Sandro Fiore (CMCC)
14:50-15:10	EXCELLERAT	Bastian Koller - Dimitris Liparas (HLRS)
15:10-15:30	HiDALGO	Francisco Javier Nieto de Santos (ATOS)
15:30-15:50	MaX	Sebastiaan Huber (EPFL)
15:50-16:10	CompBioMed	Hector Martinez (Univ. of Oxford)
16:10-16:30	BioExcel	Rossen Apostolov (KTH)
16:30-17:00	*Coffee break*	

Figure 1 Agenda of the 1st European Communities Workshop on Exascale Computing (1/2)

17:00-19:00	**Session 2 – FETHPC & HPDA**	**Chair: Giovanni Aloisio**
17:00-17:20	Maestro: Towards a Memory- and Data-aware Middleware	Dirk Pleiter (Jülich Supercomputing Centre) MAESTRO Project
17:20-17:40	Towards data intensive aware programming models for Exascale systems	Francisco Javier Garcia-Blas (Universidad Carlos III de Madrid) ASPIDE Project
17:40-18:00	The VESTEC project: Fusing HPC and real-time sensor data for urgent decision making	Gordon Gibb (EPCC, Univ. of Edinburgh) VESTEC Project
18:00-18:20	ExaNeSt: Low-Latency Communication and Acceleration in a liquid-cooled energy-efficient Prototype Rack	Manolis G.H. Katevenis (FORTH, Crete, GR) ExaNeSt Project
18:20-18:40	Sage2: Architecting a storage platform for the extreme data era	Sai Narasimhamurthy (Seagate Technology, LLC) Sage2 Project
18:40-18:55	Final Discussion	
18:55-19:00	**Wrap up and closing session**	**Giovanni Aloisio**

Program Committee
Giovanni Aloisio
Guy Lonsdale
Mathis Bode
Stefan Krieg
Jean-Claude André
Sandro Fiore

This event is funded by EXDCI2: the EXDCI2 project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 800957 http://www.exdci.eu

Figure 2 Agenda of the 1st European Communities Workshop on Exascale Computing (2/2)

5 Brainstorming within EXDCI-2/PRACE

WP3 participated in the PRACE/CoEs/FET-HPC/EXDCI Meeting in Brühl, Germany, October 30-31st, 2018. A specific focus was on the coordination between PRACE and CoEs, particularly on services offered by PRACE to the CoEs and how to align services that are offered by both.

In the former category, access to resources offered by PRACE for CoEs was discussed. Historic data showed that the CoEs did not make full use of the resources allocated to them and potential reasons for this were identified and discussed. This ranged from lack of awareness within CoEs to the need of larger allocations for some CoEs. The very nature of how CoEs need to use PRACE resources (primarily porting, scaling studies, and evaluation of new approaches) also contributed to the non-constant use of CoE resources. An awareness campaign was agreed and also the addition of new and different resources to PRACE will likely improve the situation.

In the latter category, the need for alignment of work on applications within PRACE (particularly the high-level support teams and WP8 in PRACE-6IP) and the work ongoing in the CoEs was discussed. In general, there is good alignment as in most cases work is executed in joint teams with CoE and PRACE members, but sometimes there is still the need of better coordination and mutual awareness. Similar topics were also discussed with respect to training. There, particularly the organization of joint PRACE-CoE events is highly valuable for both parties.

WP3 was also represented at the EHPCSW19 conference in Poznan. The HPC Ecosystem workshop gave an opportunity to bring up the needs and clarify the role of the CoEs in the wider HPC landscape. Topics of discussion included tighter coordination of training activities, their planning and efficient dissemination; better coordination of support activities with clear definitions of areas of expertise to be covered by the different providers such as PRACE PATCs and CoEs personnel; coordination around service provisioning and engagement of relevant user communities; incorporating needs of the wider user communities (to a big extent represented by the CoEs) within future policy developments. The meeting was useful for identifying overlaps and streamlining cooperation.

The co-design workshop was another event where WP3 was represented. All major extreme-scale players - Europe, USA, Japan and China - highlighted the need for close interactions with application developers and co-design foundation in the development of future extreme-scale systems. It was shown that many of the codes under development by the European CoEs are part of the application suites that are considered in the process.

Moreover, WP3 experts are collaborating with WP2 in the process of updating ETP4HPC's SRA (fourth version of the Strategic Research Agenda) by providing inputs from applications research communities coming from academia and industry. The SRA-4 will be issued at the end of 2019 and be delivered for the first time to the EuroHPC Joint Undertaking. The SRA-4 will involve ETP4HPC for the hardware and software HPC industries, BDVA (Big Data Value Association, [6]) for the Big Data applications, AIOTI (Alliance for Internet of Things Innovation, [4]) for the IoT/AI applications and CoE PRACE SSC for the scientific and industrial HPC applications.

WP3 experts provided in February 2019 an initial synthesis of applications requirements and challenges based on the inputs from the 3rd version of the PRACE Scientific Case (produced at the end of 2018) and documents produced for the EuroHPC working groups on user requirements.

This initial contribution (provided as Annex 1 of this document) will be completed by a series of four to six pertinent Exascale use cases.

In that sense, WP3 experts are involved in the SRA-4 editorial team and liaising with CoE and PRACE SSC for providing inputs, based on a joint template elaborated initially by US NIST committee and then completed during the BDEC meeting by US, Asian, and European experts. This template and a set of 77 answers already provided can be found here:

https://docs.google.com/document/d/1YBkdKP4TOXprNy9YOPAsfsFZey5r5jb71P7eQdfHgD8

6 Brainstorming within Industrial Communities

Various forms of information gathering with respect to industrial communities will take place at the NAFEMS World Congress in Quebec Canada. The event runs from 17 to 20th June 2019. NAFEMS is an international trade association that was set up in the 1980s in the UK, later expanding its activities in France, Iberia, Germany, Italy and the Nordic countries, so it has a strong European focus. Its membership comprises more than 1500 firms, ranging from individual consultants to SMEs and large organizations such as Daimler, Airbus, and the European Space Agency. The focus of the association is to promote best practice in the use of engineering simulation tools for design, analysis, and optimization in various sectors including manufacturing, aerospace, automotive, energy, and the built environment. A number of activities will take place at the event that will provide valuable information contributing to the output of EXDCI-2. These include parallel tracks with talks on HPC, simulation data management, optimization, and ML. It is important to note that this is not an academic conference and the talks are delivered by practicing engineers working in industry. A brainstorming workshop led by WP 3 has been arranged as part of the official program in collaboration with the NAFEMS HPC Working Group. The congress also includes exhibition space for software vendors and there are usually 30+ vendors in attendance. We will be talking to these companies during the event, collecting data on current HPC capabilities and their roadmaps for the future. The outcome of this data gathering exercise will be a written report. We consider that it will be much more effective to attend an event of this type, where industry is already gathering, than to try and encourage firms to attend a workshop organized and hosted by an HPC center. Our experience is that such events rarely attract industrial users.

A joint workshop involving the PRACE Industrial Advisory Committee and the EuroHPC Infrastructure Advisory Group was held on 3rd June 2019. The purpose was to brainstorm effective methods for improving the uptake of HPC in industry. The PRACE Industrial Advisory Committee comprises C-level industrial users from Small and Medium Enterprises (SMEs) and large firms. It is an independent body that gives advice, but is not responsible for ensuring the advice is acted upon. The discussion focused on the differences between academia and industry from the perspectives of culture and user needs. The outcomes are being written up for a white paper. Given that the workshop took place very close to the deadline for this report, the white paper is not yet available. Nonetheless, there has been time to prepare a summary of highlights. This is presented in Figure 3.

	Industry	Academia
Requirements	Problem to be solved Support to get started Services to be procured	Compute Architecture (CPU, GPU) Memory (GB per core, bandwidth) Interconnect (Ethernet, Infiniband)
Security	Protection of IP Confidentiality Requires trust (and contract)	Protection of IP Confidentiality Assumed
Training	Business benefits of HPC Writing an effective business case How to procure services Online course spread over weeks	Parallelisation strategies Programming new algorithms Use of optimisation tools Attend intensive course in person
Timescale	Very short – weeks to months Quarterly financial report Expects decision weeks after applying Time to market	Very long – years to decades Review of 3-5 year grant Expects decision months after applying Time to publication

Figure 3 Comparison of academic and industry culture/user needs

7 Conclusions and Next Steps

It is essential for WP3 to stay in contact with the various players/users in/of the European HPC infrastructure. For that, several brainstorming activities have been carried out to understand common issues and challenges including scalable meshing tools, couplers and solvers, disruptive numerical methods, mini apps development, smart in-situ/in-transit analysis/processing coupled with AI, fault tolerance, and programming languages including DSLs. One focus was on HPDA as it plays a more and more important role for many use cases.

An important next step is the Birds of a Feather session at the upcoming NAFEMS World Congress 2019. It will help to communicate gained knowledge to industry and also better understand their particular needs.

It is planned to organize a “2nd European Communities Workshop on Exascale Computing” during the next EuroHPC Summit Week in 2020. The brainstorming activities will continue using the established channels.

8 Annex 1

Initial version of European scientific and industrial applications requirements and challenges provided by WP3 to WP2 task force involved in the 4th version of the SRA

Between 2007 and 2018, three European scientific case studies were completed for PRACE. Together they demonstrate the importance of HPC provision at the highest level to ensure the European leadership in computational science. The third edition of the PRACE Scientific Case (PSC) spanning from 2018 to 2026, was issued mid-October 2018 by the PRACE SSC. Comments and results from the PSC 2018 are highlighted in the next paragraphs.

Strategic importance of HPC

High Performance Computing (HPC) is a strategic tool for competitive science, fostering innovation and supporting public decision-making based on facts. Having been used for more than 30 years in climate research, numerical weather prediction, particle and astrophysics, earth sciences and chemistry, HPC is now a cornerstone of all scientific fields from biology, life sciences and health, energy, geosciences, high-fidelity combustion, materials science, to social sciences, and humanities. In industry, HPC is widely used in oil & gas exploration, aeronautics, automotive, and finance, and is now becoming crucial for ensuring personalized medicine, developing nano-technologies, or enabling the development and the management of renewable energies. Initially the preserve of large companies, HPC is now used more and more by European SMEs, which represent a very important part of European industrial competitiveness. In order to accelerate this trend, PRACE and some of its partners have launched several initiatives at national or European level (SHAPE programme, Fortissimo projects) offering a unique set of services spanning from HPC dissemination, and open R&D, to confidential R&D, and commercial cloud-based activities.

Finally, HPC is becoming a tool of growing importance for supporting public rational decision making by simulating scenarios and allocating urgent computing resources in a “red button mode” in the case of natural hazards (extreme events such as earthquakes, thunderstorms, flooding or evacuations), industrial risks (rupture of a dam, failure of a power-plant, spills), biological risks (diffusion of pandemics) or (cyber)terrorism attacks.

Convergence of simulation and big data analysis

Beyond its importance for accelerating numerical simulation, HPC is now also used at the convergence of simulation and big data workloads, this need being driven by the increasing amounts of data coming from next generation scientific instruments (satellites, (radio)telescopes, accelerators, microscopes, sequencers), from the Internet of Things (IoT), social media, and from large scale simulations (including massive 3D simulations, multi-scale and multi-physics coupled simulations, ensemble/optimisation/scenario studies, uncertainty quantification).

Many science communities are now combining HPC and High Performance Data Analytics (HPDA) applications and methods in large-scale workflows that orchestrate simulations and incorporate them into the stages of large-scale analysis pipelines of data generated by simulations, experiments or observations. The short time scales required for some of the applications constitute a crucial challenge.

Artificial intelligence

Exploiting and valuing such enormous amounts of structured or unstructured data in a reasonable and competitive time is no longer possible for human beings, leading to the rise of HPDA supported by new data assimilation, interpretation, extraction, and prediction-techniques benefiting from artificial intelligence (AI) and machine learning (ML).

There is an ongoing convergence between HPC and HPDA/AI. Scientific communities generating large amounts of data will require HPDA/AI tools allowing them, when possible, to (in-situ or in-transit) infer experimental/simulation outputs just after being generated. This makes it possible to classify and identify pertinent structures to be stored, saving time and energy, because the raw data does not have to be stored and processed after. In the future, this also could lead to smart AI-driven computational steering techniques or using AI techniques in coupling simulations and learnt models in order to accelerate convergence of scientific models or optimisation/uncertainties studies.

On the other hand, modern AI (coming after a 25-year AI winter with the support of DNN on GPU and FPGA) will benefit from converged HPC/AI architectures for addressing new scale-out challenges including the development of more complex and deeper neural networks (like CapsNet Capsule neural networks), AutoML techniques or explainable AI (XAI) methods.

In summary, many of the industrial, scientific and societal challenges that communities face will be simulated before being actually executed. This scenario requires extensive access to HPC resources, robust and reengineered computational codes and data analytics for which the next generation of scientists, technologists need to be trained in multidisciplinary areas for which new educational plans will be required.

Future directions

Due to the rapid convergence between HPC, HPDA and AI as a result of the explosion of data generated by large scale instruments or numerical simulations, PRACE and the European computational ecosystem must develop new architectures and services addressing mixed HPC/AI workloads. This has been already understood by countries like US (installation in 2018 of the Summit 200 PF converged system at ORNL), China and Japan who will address this convergence in their roadmap.

Such convergence will also be fostered if major EU-wide efforts are made in education and training towards new skills in data science and numerical simulation to train and retain a new generation of researchers and technologists in both science and industry, and if basic training in the technologies is provided and/or their impact is transferred to the general public.

9 Annex 2

Proposal of CNES "DATA TERRA Toward a fully integrated earth sciences data distributed platform" submitted to BDEC2 meeting in Poznan in May 2019

Introduction

Satellites produce a wealth of data and information regarding the Earth sub-systems (land, atmosphere, oceans, solid Earth and biodiversity) and cross-cutting processes (climate change, sustainable development and security).

Space agencies cooperate together in order to optimize the usage of their data; this is done for example in the context of the CEOS (Committee on Earth Observation Satellites). CEOS is for example a forum of technical exchange on data services interoperability (discover, access, subset, visualization & process), Future Data Architecture (datacube, cloud, analysis ready data, exploitation platform…), etc.

Most of space data are open and free (with an exception for very high resolution imagery or for some countries or for cooperation with private entities).

The volume of space data is increasing exponentially with some programs like Copernicus, SWOT, NISAR or the last generation of weather forecast satellites (GOES, HIMAWARI, MTG). It represents dozens of PB and will reach several hundreds of PB in the next 3 three years (For example, for NASA alone, the growth rate of the archive will be around 50 PB from mid-2022.).

Given the explosion of data, it is now advisable not to download large volumes of data at home, but rather to move the processing where large data are hosted.

CNES organisation and services

CNES is the French space agency. It operates several satellites (JASON, CFOSAT, SARAL, Megha-Tropiques, Calipso, SMOS, IASI, PLEIADES, etc.) and develops new ones (MERLIN, MICROCARD, IASI-NG, CO3D, etc.).

Most of the missions are done in cooperation with other space agencies (ESA, EUMETSAT, ISRO, NASA, NOAA, etc.). In terms of processing, CNES is only responsible for the despatialization of data; that is to say up to a product level that does not require an expertise of the satellite or its instruments. In some cases, depending on the cooperation agreements, the treatments are carried out by our partners. Depending on the case, CNES may be responsible for the distribution of products and their long-term archiving or it may delegate these activities to its partners.

CNES is also very committed to promoting the use of its data and spatial data in general.

Some processing can be executed in the satellites themselves to reduce the amount of data to be transferred to Earth (Edge Computing); this is the case, for example, for the IASI instrument on METOP. But this is not a widespread practice now.

CNES also hosts a mirror of Copernicus data (French Copernicus Collaborative Ground-Segment) which represents about 10 PB (4PB on line and 14 PB capacity on tape).

When a data is processed by CNES, it is on its own computer center. A presentation (made in July 2019) of the CNES computer center can be found here (https://sedoo-cloud.omp.obs-mip.fr/owncloud/index.php/s/NWc9qZi7wSgXhb4/download). It allows numerical simulations (HPC), but the data processing is made in a specific HTC/HDA environment. For data reprocessing (very resource intensive), it is foreseen to use cloud-bursting solutions on commercial means.

The goal of CNES and other space agencies is to promote the use of data by the widest possible user communities. These categories of users can be classified macroscopically into two broad categories: research and the downstream sector.

For the downstream sector, the preferred solution to further exploit the data is commercial cloud computing; This for any type of treatments, including AI. For that, there are a large number of solutions:

- In Europe, 5 CDIAS (Copernicus Data and Information Access Services) have been initiated by the European Commission. They propose processing capabilities (cloud computing), services, and additional data. They rely on commercial clouds; namely CloudFerro, Orange, OVH and T-System.
- Several SME initiatives – for example Sinergise Sentinel Hub or Terradue Ellip.
- Amazon and Google propose a wide range of satellite data
    - All sentinel 1&2 (Copernicus) data are already hosted by Amazon. US agencies (NASA, USGS and NOAA) are moving their data to Amazon to promote its use and facilitate treatment with very large data by users (research & downstream).

Figure 4 DIAS concept

For the research sector in France, satellite data are exploited in five thematic Data & Services hubs:

- AERIS for the atmosphere thematic
- FOR@MATER for the solid Earth thematic
- ODATIS for the ocean thematic
- PNDB for the biodiversity thematic
- THEIA for the land-surface thematic

Figure 5 Earth System Data & Services Hubs in France

Each Data & Services Hub is geographically distributed across multiple data & services centers that all have their own computing capabilities (which mostly are HTC/HDA type clusters). They do not only deal with satellite data, but also a very large amount of in-situ heterogeneous data (ground, sea, airborne, etc.). These data are less bulky than satellite data, but they are much more varied in terms of content, size and formats.

With the rise of IoT and 5G, the volume of in-situ data may explode. In this case it may be necessary to fully review the computing hierarchy of this data and turn to technologies such as Edge Computing, data compression/reduction, use of software defined networks for a smart orchestration of the successive level of resources, support of workflows from end to end (from the edge to the tape of the data center), etc.

In that sense, some Data & Services Centers use data from models that are processed in HPC centers located in other research infrastructures.

Globally, our capacity to adopt an integrated inter- and trans-disciplinary approach is hampered by the fact that Earth Science data is today highly compartmentalised between these different scientific disciplines and communities.

DATA TERRA: toward a fully integrated earth sciences data distributed platform

The Earth is a living system encompassing multi-scale and multi-physics internal dynamical processes and interactions with its external fluid envelopes (e.g., ocean, atmosphere, etc.) and continental interfaces (lands, biosphere and anthroposphere). Understanding, monitoring, and predicting the evolution of the Earth's systems in their environments is a fundamental scientific challenge with important societal and economical applications in terms of natural hazards (e.g. volcanoes, earthquakes, tsunamis, landslides), environmental and climate change, new energy resources, sustainable development.

Some countries have undertaken major efforts to restructure this heterogeneous ecosystem, mutualise resources and expertise, and to provide platform of services enabling from end-to-end (edge to the tape) the efficient data logistics including discovery, access, interoperability and wider reuse of data within and beyond Earth Science communities, to society as a whole. The French "DATA TERRA" Research Infrastructure is an example, and a recognized global leader. System Earth was established in France in 2017 to integrate existing data and service hubs and provide easy access to Earth Science data and associated products across the board for scientists and decision makers. With the core mission to facilitate and foster integrated & interdisciplinary research to understand DATA TERRA processes and Global Changes, DATA

TERRA is committed to implementing the FAIR principles, creating distributed services, tools and workflows for data management, curation and scientific use; and promoting dialogue and international agreement on best practice.

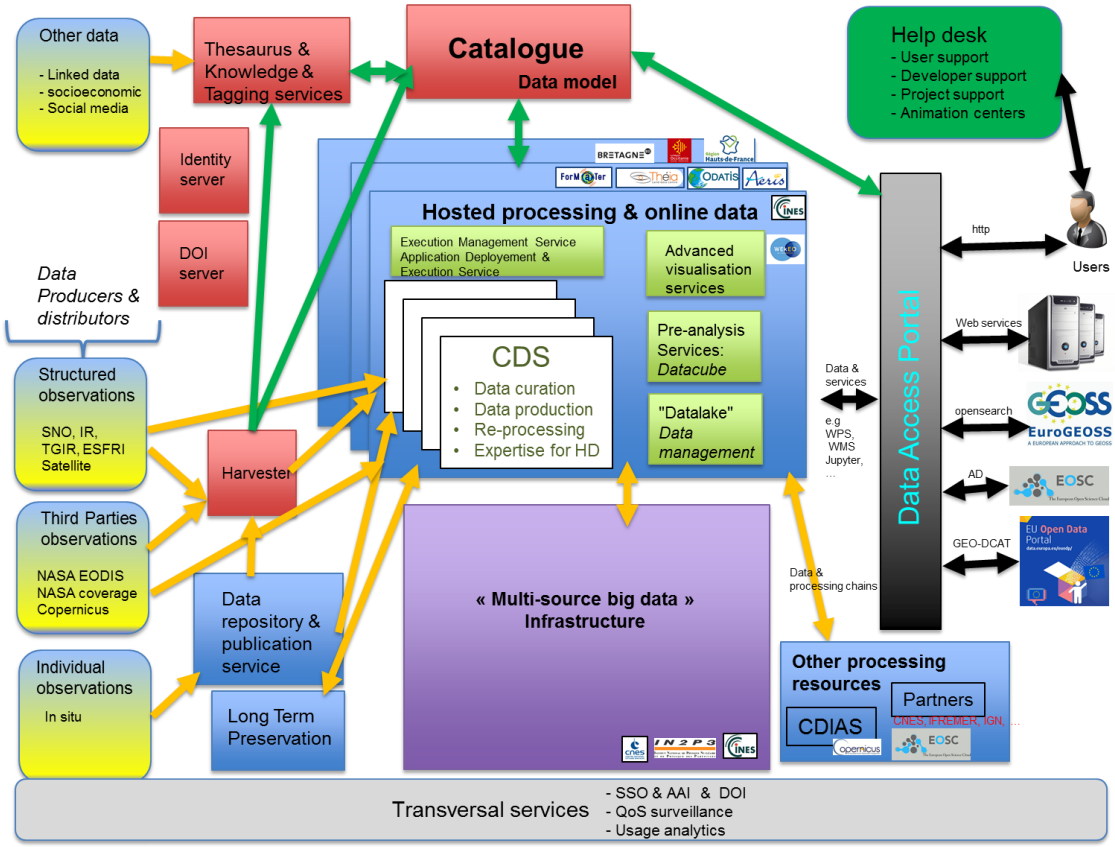

Figure 6 DATA TERRA High level Architecture

DATA TERRA will have to face several challenges:

- the level of FAIRness of the different centers is heterogeneous
- Bulky data are geographically distributed in heterogeneous computing infrastructure with different level of services; there is a need to develop crosscutting applications then to combine different data in different location (in a distributed software platform) with implication on:
    - Workflows
    - Data logistics
    - Network
    - Computing infrastructure
- Stay open to European cooperation (e.g. ESFRI & ENVRI) and international cooperation (e.g. RDA)
- It will not be possible to build a monolithic and centralized system with all data and processing resources
- In a context of convergence of HPC, HPDA and AI, take into account the moving computing infrastructure landscape in France and in Europe
    - Existing processing capabilities in the Thematic Data and Services Hubs
    - French INFRANUM project led by the French Ministry of Research to concentrate the processing infrastructures at the regional and at the national level (GENCI)
    - EOSC which is the natural solution for ESFRI

- DIAS
  PRACE the HPC European Research Infrastructure, EuroHPC and EDI (European Data Infrastructure)

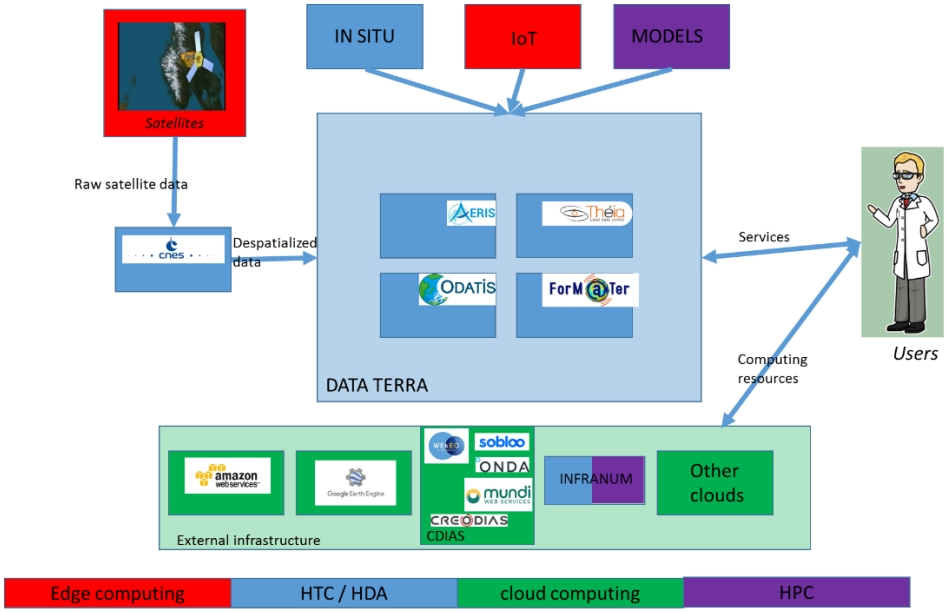

Figure 7 TERRA DATA trans-continuum workflow

The proposed BDEC proposal is a prototype of an architecture that will allow DATA TERRA to fulfill its objectives.

1. What innovative capabilities/functionalities will the proposed candidate platform demonstrate (e.g. transcontinuum workflow, edge computing, data logistics, distributed reduction engine, etc.)?
The prototypes address transcontinuum workflows (cf. figure above), edge computing, data logistics.

2. What applications/communities would/could be addressed?
The communities are the Earth Science communities.

3. What is the "platform vision," i.e. what kind of shared cyberinfrastructure (CI) for science would the further research/design/development of this platform lead to?
Cf. two previous figures.

4. How available/ready/complete is the set of software components to be used to build the demonstrator?
To be developed. The CEF (Connecting European Facilities) OpenData/HPC EC project PHIDIAS (duration 3 years from July 2019 with 16 European partners including CNES, CINES, CSC, CERFACS, IS Terre, IRD, IPSL, SPACIA, …) will allow to develop the first elements of the system.

5. As far as one can tell at this early date, to what extent can this be done with existing and/or otherwise available hardware/software/human resources?
Cf. point 4.

6. What is the potential international footprint of the demonstrator?

System Earth is by nature an international topic. It can be derived at European level in the frame of ESFRI/ENVRI. It can be derived at international level in the frame of GEO or RDA.