EXDCI-2 – Lessons Learned 2018-2020

Florian Berberich
PRACE aisbl

François Bodin
University of Rennes 1

The EXDCI-2 project has received funding from the European Unions Horizon 2020 research and innovation programmed under grant agreement no. 800957.

Introduction

- **Development and advocacy of a competitive European HPC Exascale Strategy**

EXDCI-2 supported the implementation of a common European HPC strategy, opened to synergistic areas including High Performance Data Analytics (HPDA) and Artificial Intelligence (AI). The expertise of the partners and stakeholders permitted to elaborate a transversal prospective vision covering innovation opportunities.

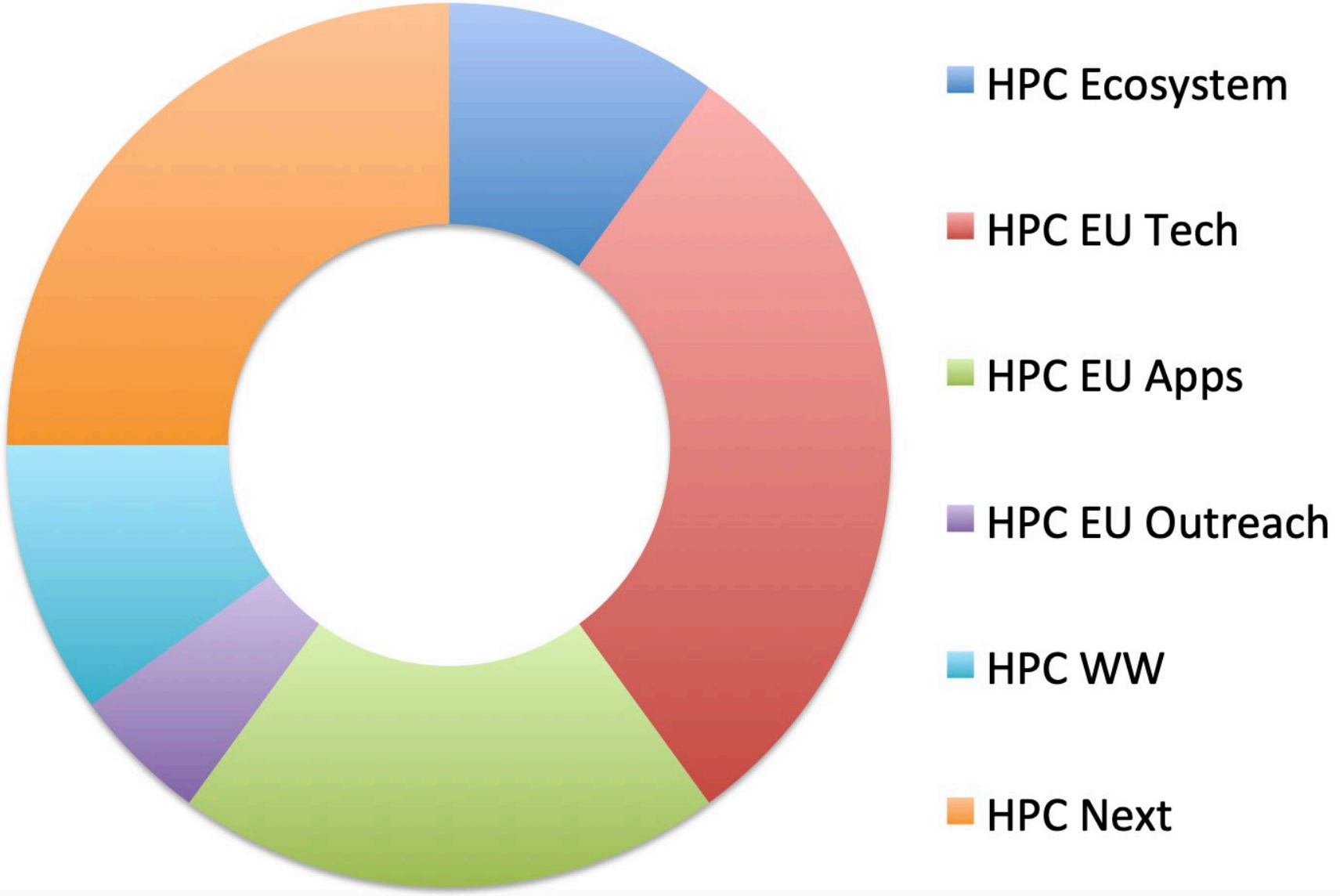
EXDCI-2 Consortium: 4 contractual partners and 20 Linked Third Parties

- BSC
- CaStoRC
- CEA
- CMCC
- EPCC, Univ. of Edinburgh
- GENCI
- INRIA
- JUELICH
- KTH
- NEOVIA
- RWTH Aachen Univ.
- Univ. College of London
- Univ. of Ljubljana
- Univ. of Manchester

- ESI Group
- FRAUNHOFER
- HPCnow!
- ParTec
- SCAPOS
- TS-JFL

-

Outline of the talk



- HPC Ecosystem
- HPC EU Tech
- HPC EU Apps
- HPC EU Outreach
- HPC WW
- HPC Next

- Analysis
- FETHPC projects IP
- Spin-off experiment

Analysis and assessment of the EU ecosystem

- HPC cPPP progress monitoring the ecosystem

- Process
  - Boundary conditions for impact assessment : ***use the KPI definitions common to all the cPPPs*** (leverage factor, jobs, SMEs, innovations) + HPC cPPP specific KPIs
  - Data have been collected from various sources: online surveys, SME interviews, questionnaire to CoEs, confidential interviews with companies, DG-CNECT, PRACE

- Main findings
  - HPC cPPP programme already produced observable stimulation of the EU HPC ecosystem between 2014 and 2018, with first tangible results
    - Production of promising IP elements in both hardware and software and various open source developments
    - Creation of jobs in the private sector, especially in SMEs, and in the public research sector
    - Stimulation of continued or new private sector R&D investment, matching or complementing EU funding

FETHPC 2014 project results analysis

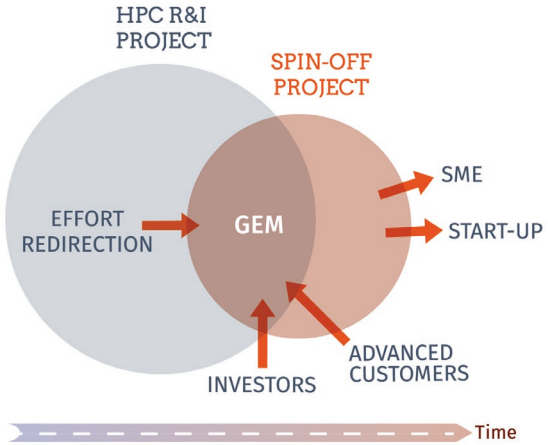
- 171 IP identified
  - Most of them software related
    - 40 related to applications
    - Libraries
    - Runtime
    - FPGA tools and software
  - Miscellaneous hardware results
    - Boards: ARM, FPGA
    - Interconnect, cooling
- Objectives
  - Incentive to reuse the IP
  - Promote synergies

Searchable database https://exdci.eu/activities/fethpc-results

API	**7**
application optimisation	**7**
benchmark suite	**6**
demonstrator	**8**
hardware	**20**
report	**2**
software	**114**
training	**7**

Experimenting the concept of "project spin-off"

- Objective
  - Unlocking innovations in EU-projects by stimulating **project spin-offs**
- Activity: Run a Proof-of-Concept for assessing the spin-off idea
  - Legal preparation (GA amendment for subcontracting)
  - Set-up of a committee of Experts for accompanying and assessing the PoC ("Selection Committee")
  - Run an open call for proposal, and selection of the winning proposal: BigDFT4Chem (https://exdci.eu/newsroom/news/bigdft4chem-bigdft-chemical-industry)
  - Follow-up on of the execution of the PoC
- Main outputs
  - Assessment of the spin-off idea
  - Recommendations on foster innovations and speed-up market access of research project results
  - Contribution to the emergence of a new company to market BigDFT



Coordination with European actions on Mathematics and Machine Learning

- Summary
  - Fosters cooperation between HPC community and EU actions/actors from the Mathematics and Machine Learning domains
- Process
  - Contact with EMS (European Mathematical Society) and ECMI (European Consortium for Mathematics in Industry)
  - A survey has been performed to understand the HPC-Mathematics communities articulation
- Main Recommandations
  - **Vital stimulus** for future **mathematics** and **machine learning** should remain **within the framework of the ERC**
  - The ERC's remit should be broadened to allow **smaller grant amounts** and inclusion of mathematics in other Evaluation Panels.
  - To preserve Europe's mathematical heritage and prepare society's future, we should propose attractive research incentives to all mathematicians, especially **younger ones**

HPC Ecosystem

HPC EU Tech

HPC EU Apps

HPC EU Outreach

HPC WW

HPC Next

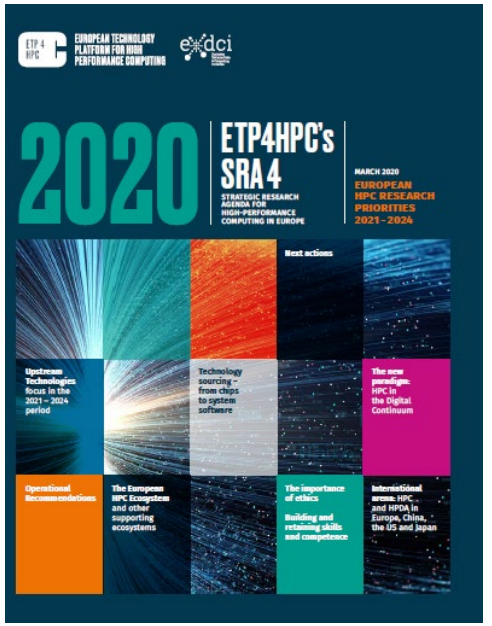
Strategic Research Agenda

Legacy codes analysis

HPC-HPDA technology roadmap / SRA4

- Strategic research Agenda
  - prepare a HPC-HPDA technology roadmap for the Horizon Europe
  - Build a global vision for the research focus areas

- Process
  - A layered approach for identifying SRA-4 research priorities
    - From societal, research and industry challenges to technical research priorities
  - 8 working groups involving ~ 100 experts
  - Connection of these groups through “research clusters”
- Main outputs
  - SRA4 document to drive research of HPC ecosystem
  - Research call inputs for EuroHPC Research and Innovation Advisory Group (RIAG)



Legacy codes and software modernization / transition - 1

- Summary
  - Understanding the technical debt to come with legacy codes
  - Estimating the amount of efforts that will be needed in the future

- Process
  - Specification of **"Code Viscosity", an easy measure of legacy code potential technical debt**
    - Simple enough to be used extensively (require few effort from dev. teams)
    - Less work that the usual metrics used in software engineering
  - Workshop at the EuroHPCSW in order to start gathering feedback
    - Legacy code is a concern to most participants
    - Good feedbacks on viscosity from the audience

Legacy codes and software modernization / transition - 2

- Main findings
  - The European code base represents assets that correspond to a significant amount of money that cannot or should not be wasted
  - The team sizes are too often too small to envision a full rewrite in a timely manner
  - The codes are critical to their owners and, therefore, can't be discarded through obsolescence due to a change of supercomputers
  - Rewriting a code from scratch takes many (> 2) years as soon as the code has a significant size
  - Still some, if not all, codes rely on the availability of FORTRAN (FORTRAN90). This raises the concern with respect to teaching, as FORTRAN is no more studied during the academic syllabus
  - Whereas some teams are eager to start their porting effort very soon, others plan to wait at least 5 years, if not more, before starting the work**. Given the time needed for rewriting, one can expect these codes only to be ready for exascale in 10 to 15 years from now**
- Conclusions of these findings are that some funding needs to be dedicated to code porting by teams in addition to the funds given to CoE
  - Given the size of the codes and their specificities, it is difficult to imagine that the CoEs will do all the jobs

HPC Ecosystem

HPC EU Tech

HPC EU Apps ← Applications and usages

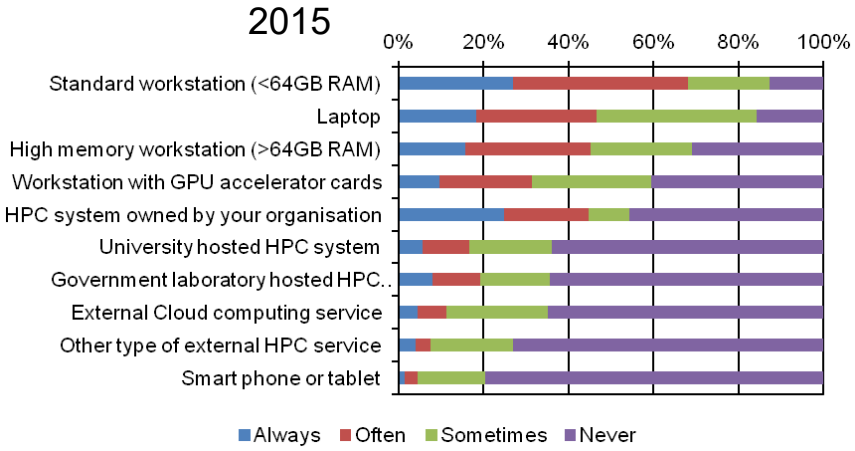
HPC EU Outreach

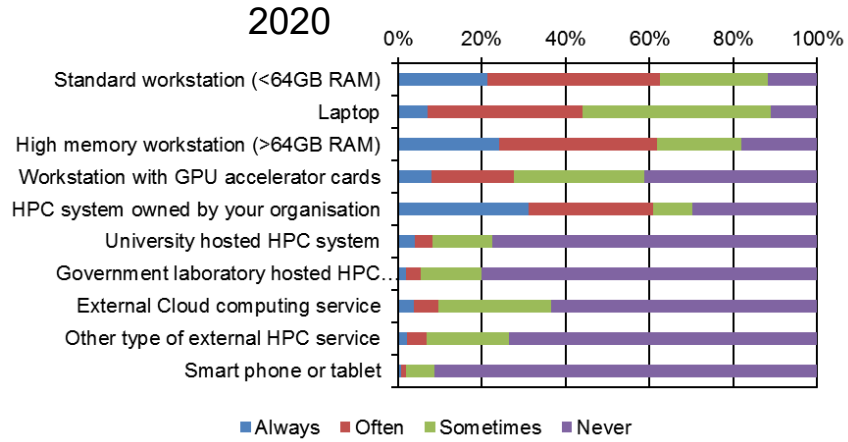
HPC WW

HPC Next

HPC applications and usages / communities analysis

- New methods (AI, edge, …) have huge potential but make cases even more complex
- **A lack of support for hybrid workloads and co-design** processes in Europe was identified
- Amount of computing resources is ok**, but funding for code development and especially code updating difficult**
- **Different communities work on different solutions for partly similar problems**
- **Small differences in industrial usages observed**
- Feedback with respect to training so-so






HPC Ecosystem
HPC EU Tech
HPC EU Apps
HPC EU Outreach
HPC @ schools
HPC WW
HPC Next

HPC Outreach

- Promoting HPC at schools
  - Integrate school teachers point of view
- Status
  - Workshop 8 & 9 July 2019 at Edinburgh
  - Attendees from primary schools and universities
- Main findings
  - It is necessary to emphasise the benefits of mathematics and consequently the computer skills personally and in all careers. **Children must be able to explain to their parents**
  - We must implement the knowledge of programming in all the STEM subjects and evaluate them in the **examinations from the age of 10**
  - Integrating data, AI and IoT related knowledge could be a great support in primary school to develop environment related pupils' projects**. Collaborations are required between universities and primary schools to set up such an initiative**

HPC Ecosystem

HPC EU Tech

HPC EU Apps

HPC EU Outreach

HPC WW ← Big Data and Extreme-scale Computing (BDEC)

HPC Next

International liaison and workshops

The activities of the second period were:

- BDEC workshop in San Diego, USA on 15/10/2019-17/10/2019.
- BDEC BoF during SuperComputing 2019 in Denver, USA, held on 20/11/2019.
- Participation in NSF Smart CyberInfrastructure workshop held on 25/02/2020-27/02/2020 in Washington DC, USA.
- Co-organisation of BDEC virtual workshop (USA), held on 10/06/2020.
- Organisation of BDEC virtual workshop (EU), held on 10/07/2020.
- Participation in BDEC virtual workshop (Japan), held on 01/09/2020
- Co-organisation of the final executive virtual workshop, held on 15/10/2020.

The final BDEC activity has been the preparation of a last community report that will be published at the end of February 2021 and simultaneously submitted to the NSF. This report will propose an international strategy for addressing the current challenge of combining HPC, Big Data and Machine Learning.

https://www.exascale.org/bdec/

One report per period

- petascale to exascale
- Big Data convergence
- AI/ML

The Internation
Project roadma

Jack Dongarra, Pete Be
Giovanni Aloisio, Jean-C
Jean-Yves Berthou, Tais
Franck Cappello, Barbar
Thom Dunning, Sandro
Michael Heroux, Adolfy
Sanjay Kale, Richard Ken

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, T. Moore, M. Asch, R. Badia, M. Beck, P. Beckman, T. Bidot,
A. Choudhary, B. de Supinski, E. Deelman, J. Dongarra, A. Dubey, G.
W. Gropp, M. Heroux, Y. Ishikawa, K. Keahey, D. Keyes, W. Kramer, J.
Matsuoka, B. Mohr, D. Reed, S. Requena, J. Saltz, T. Schulthess, R. S
Szalay, W. Tang, G. Varoquaux, J.-P. Vilotte, R. Wisniewski, Z. Xu and

BIG DATA AND EXTREME-SCALE COMPUTING 2: Two Models of Convergence for Cyberinfrastructure for Science

Finding a Path Forward to Software and Data Ecosystems for 21st Century Science

Tech Report No. ICL-UT-xx-xx

Edited by Terry Moore and Mark Asch

January 2021

Prepared by:

EXDCI2 PROJECT OF EU-H2020 PROGRAM
and
UNIVERSITY OF TENNESSEE,
DEPARTMENT OF ELECTRICAL ENGINEERING AND COMPUTER SCIENCE,
INNOVATIVE COMPUTING LABORATORY

HPC Ecosystem

HPC EU Tech

HPC EU Apps

HPC EU Outreach

HPC WW

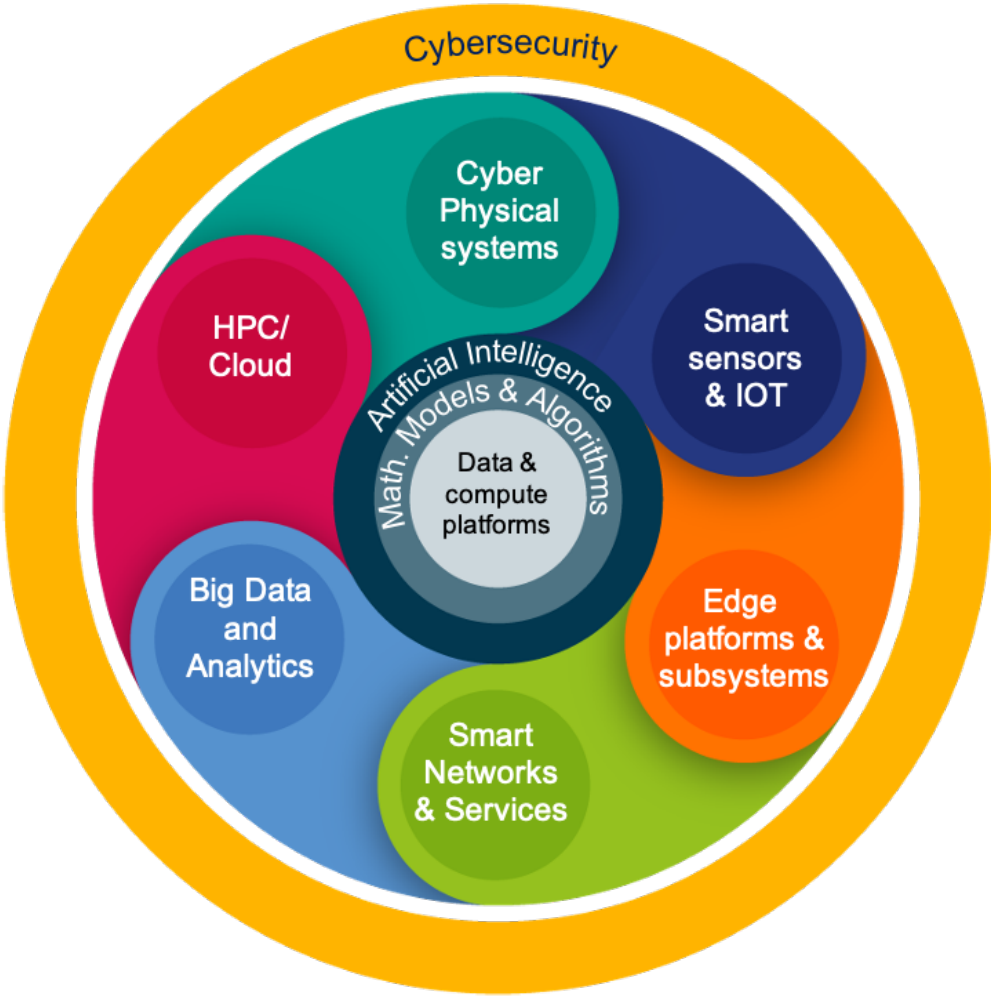
HPC Next ← TransContinuum Initiative

Extending the vision: TransContinuum Initiative (TCI)

The TransContinuum concept

1) Captures the idea that infrastructure elements cannot be used independently

2) Provides a coherent and effective view of the cyberinfrastructure for end-to-end applications



Original courtesy HiPEAC

A continuous dynamic workflow

Between
Smart Sensors and IOT devices

and
HPC / cloud centers

passing through
Edge platforms & subsystems

as well as
Smart Networks and Services

executing
Simulation & Modelling,
Big Data Analytics and ML*

based on
Math. Methods & Algorithms incl. MSODE**

pervasively augmented by
Artificial Intelligence

protected and secured by
Cybersecurity

back to
Cyber-Physical Systems,

all based on
Data and compute platforms (hw and sw)

* ML: Machine Learning
** MSODE: Modelling, Simulation and Optimization in Data-rich Environment

TCI scope and objectives

- Identify priorities and recommendations for European R&I workprogrammes
- Engage with European R&I funding agencies and R&D programmes
- Generate and foster an interdisciplinary network of experts in science and industry
- Contribute to SR(I)As and other partner's documents
- Contribute to the 5 Horizon Europe Missions

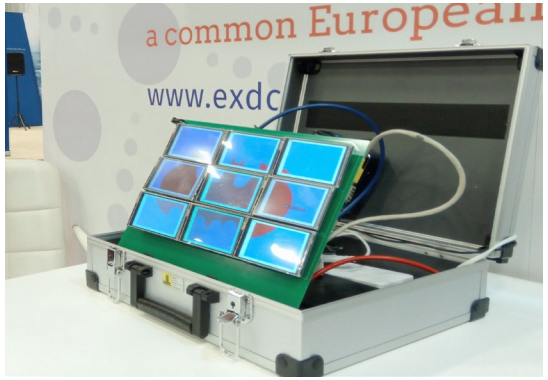
➢ see also "TransContinuum Initiative (TCI): our vision"*

*https://www.etp4hpc.eu/news/247-transcontinuum-initiative-joint-vision-docume.html

Long-term Impact

Conclusion & Long Term Impacts

Domain	*Impact*
Political / EU	• Comprehensive analysis of the HPC/HPDA/AI landscape at large • Analysis addressing technologies (current and future) & applications • Proposing a new holistic initiative (TCI) to embrace future societal and application challenges
Economic	• Assessment of legacy codes assets and stress the need to preserve these assets • Prototyped the creation of spin-off for project - shorten the innovation to market path • Inventoried the IP assets created by HPC projects • Analysis of ecosystem KPI
Social	Facing the lack of human resources: • Analysed the need for connecting HPC to pupils • Propose actions to promote STEM at school
Technology	• Building a new SRA • Providing an extensive survey of HPC projects IP
Science and Innovation	Proposed the TransContinuum Initiative (TCI): • Implements the EXDCI-2 vision • Proposes a new understanding for secured and seamless scientific application • Support infrastructures • Implementing the AI/HPDA/HPC convergence



Archie Frog prototype